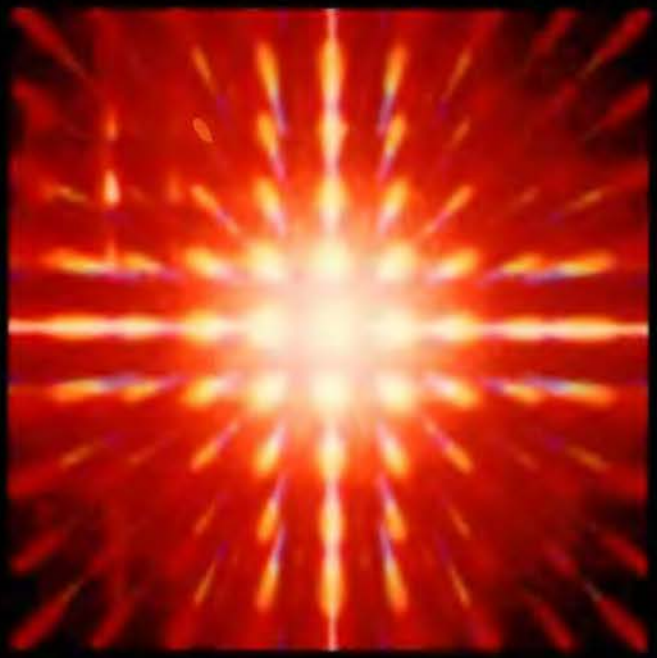
International Edition

OPTICS

FOURTH EDITION

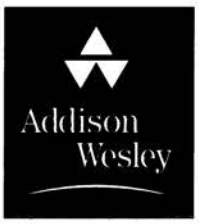


Eugene Hecht

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

4ed OPTICS

Eugene Hecht
Adelphi University



San Francisco Boston New York
Capetown Hong Kong London Madrid Mexico City
Montreal Munich Paris Singapore Sydney Tokyo Toronto



Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

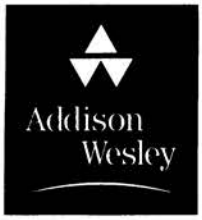
Sponsoring Editor: *Adam Black*
Project Manager: *Nancy Gee*
Manufacturing Supervisor: *Vivian McDougal*
Cover Designer: *Blakeley Kim*
Production Service: **HRS** Interactive
Text Design: **HRS** Interactive
Composition: **HRS** Interactive
Illustration: **HRS** Interactive
Photo Research: *Carolyn Eisen Hecht*

ISBN 0-321-18878-0

If you purchased this book within the United States or Canada you should be aware that it has been wrongfully imported without the approval of the Publisher or the Author.

Copyright ©2002 Pearson Education, Inc., publishing as Addison Wesley, 1301 Sansome St., San Francisco, CA 94111. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, 1900 E. Lake Ave., Glenview, IL 60025. For information regarding permissions, call 847/486/2635.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.



1 2 3 4 5 6 7 8 9 10–MA–04 03 02 01
www.aw.com/physics

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Preface

To *Ca, b. w. l.*

As with previous revisions, this fourth edition has been guided by three distinct imperatives: to fine-tune the pedagogy, modernize the discourse, and update the content.

During the past several years, using the third edition in the classroom, a number of small but significant pedagogical refinements have evolved, and these have been incorporated throughout this revised text. For instance, all symbols representing vectors now have an arrow directly above them. The intent, as ever, is to be responsive to students' needs.

This edition continues the program of gradually modernizing the treatment. In this, there are several goals, vis-à-vis the student: to impart an appreciation of the central role of atomic scattering in almost every aspect of Optics; to provide an understanding, as early as possible, of the insightful perspective offered by Fourier Theory; and to make clear, from the outset, the underlying quantum-mechanical nature of light.

Technological advances have been made in a wide range of areas from lenses and lasers to telescopes and fiberoptics. This fourth edition treats (even if sometimes only in introductory fashion) many of the significant advances that today's students should be aware of.

At the request of users, I have added dozens of new problems throughout the text. Most of these were designed to develop needed analytic skills and are of the "easy-to-intermediate" variety. As in previous editions, the complete solutions to problems without asterisks can be found at the back of the book. It should be noted that the vast majority of the new problems are not provided with such solutions. This was done to increase the selection of potential homework questions.

Much effort was expended in redrawing well over a hundred pieces of existing art to make them clearer and the physics more readable. To further enhance the pedagogy, this edition also contains many new drawings and photographs.

Chapter 2 (*Wave Motion*), which lays the mathematical groundwork for wave theory, has been somewhat revised in order both to make the existing material clearer and to provide a broader foundation for what's to come. For instance, the discussion of the differential wave equation (p. 13) has been fleshed out a bit, with some of the previously missing steps now provided.

Chapter 3 (*Electromagnetic Theory, Photons, and Light*) has been extensively reworked. Nowadays many students studying Optics come to it with little more than the brief exposure to electromagnetic theory afforded by the traditional introductory course in physics. With them in mind, Section 3.1, *Basic Laws of Electromagnetic Theory* (p. 37), has been expanded, making the material far more accessible to these students. In keeping with the commitment to discuss modern applications of Optics, a new section called *Optical Cooling* examines this important technique (p. 65).

The Propagation of Light (Chapter 4) now contains a discussion of the historical origins of the concept of *index of refraction* (p. 88). This helps to make Snell's Law easier to really understand. The chapter is further enhanced with a more thorough treatment of *Fermat and Mirages* (p. 107). A new section, 4.11.2, called *Photons and the Laws of Reflection and Refraction* (p. 141) completes the chapter.

The treatment of *Geometrical Optics* (Chapter 5) was refined here and there (e.g., pp. 161, 164, 166, 178, and 215) to improve its clarity. The field of telecommunications is so important and so rapidly evolving that each edition of this book has had to treat several major technological advances. Accordingly, the discussion of fiberoptics was brought up to date with the consideration of such topics as erbium-doped

Page 4

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

fiber amplifiers (p. 197), wavelength division multiplexing, and optical switching via MOEMS (p. 200). Liquid mirrors are briefly considered on page 223. The chapter ends with a new section on *Gravitational Lensing* (p. 231).

The discussion of wavefront aberrations in Chapter 6 (*More on Geometrical Optics*) has been enlarged (p. 254). There is a new piece concerning the upgraded Arecibo Observatory (p. 258) because it beautifully illustrates an important contemporary approach to dealing with spherical aberration.

Chapter 7 (*The Superposition of Waves*) was reworked to make the material generally more accessible (e.g., p. 303). The phasor representation was used to illuminate the creation of both standing waves (p. 288) and partial standing waves (p. 289). Because of a very significant series of experiments published in the last several years, the discussion of *Group Velocity* (p. 296), Section 7.2.2, was enriched and new subsections on *Superluminal* and *Subluminal Light* were added.

In Chapter 8 (*Polarization*), as elsewhere, the prose was tightened and the analysis clarified, here and there. A few new photographs and several fresh diagrams were included. The section (8.7) on *Retarders* was extended, and the concepts of zero-ordered, multiple-ordered, and compound zero-ordered wave plates were introduced. A section on *Liquid Crystals* (8.12) was added, and the operation of both the liquid crystal variable retarder and liquid crystal display (LCD) were explained.

In addition to a few new photos and the occasional clarifying remark, Chapter 9 (*Interference*) now contains a section, 9.8.4, called *Radar Interferometry*.

Over the last two decades there's been some interesting work done on so-called nondiffracting beams. Accordingly, Chapter 10 (*Diffraction*) contains a new section (10.2.7) entitled *The Zeroth-Order Bessel Beam,* that deals with the phenomenon.

Chapters 11 (*Fourier Optics*), and 12 (*Basics of Coherence Theory*) have undergone a line-by-line fine tuning, but little or no overhaul.

Chapter 13, *Modern Optics: Lasers and Other Topics*, has been revised with the addition of a subsection on *Gaussian Laserbeams* and some updating as required (e.g., it now includes material on the Omega laser).

This fourth edition continues the agenda of unifying the discourse, as much as possible, within the framework of a few grand ideas. Thus the concept of interference, which is one of the premier notions in Optics (and not surprisingly in Quantum Mechanics, as well), is used qualitatively to understand propagation phenomena long before it's studied formally in Chapter 9. Among other benefits, this approach of presenting advanced concepts in simplified form early in the exposition allows the student to develop a cohesive perspective.

Over the years, I have received comments, articles, and photographs, from hundreds of colleagues, and I most sincerely thank them all. I am especially grateful to Professors P. J. Dolan of Northeastern Illinois University, W. A. Mendoza of Jacksonville University, M. W. Coffey of the University of Colorado and H. Fearn of California State University for their contributions and suggestions. Prof. J. R. Peverley of The George Washington University kindly donated several very nice problems on Jones matrices and I thank him for helping to freshen up this edition. Anyone else wishing to contribute their favorite problems, please fell free to do so. Indeed, if you are interested in the discipline and wish to exchange ideas you can contact me by mail at Adelphi University, Physics Department, Garden City, N.Y. 11530 or at genehecht@aol.com.

I'd like to thank the entire team at Addison Wesley, for all their help, without which this edition would never have seen the light of day, as it were. I am especially grateful to Adam Black whose enthusiasm for the project was sustaining, to Joan Marsh whose wise decisions made the whole thing manageable, and to Nancy Gee who handled the day-to-day operation with efficiency and good humor.

The book was produced by HRS Interactive, which did a brilliant job of getting it all together. Lorraine Burke watched over every aspect of the process with incredible patience and skill; Alan Wiener and Jennifer Burke cheerfully brought their production acumen to bear; Ed Burke designed a beautiful book and struggled mightily to maintain the highest standards; Hilda Espreo was the tireless compositor; and as ever, Pat Hannagan, with the able assistance of Chris Burke, produced incomparable art. All have my deepest respect and profound appreciation. Finally, I thank my dear friend, my wife, Carolyn Eisen Hecht for coping with one more edition of one more book.

Freeport, New York E.H.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Content

1 A Brief History 1

1.1 Prolegomenon 1
1.2 In the Beginning 1
1.3 From the Seventeenth Century 2
1.4 The Nineteenth Century 4
1.5 Twentieth-Century Optics 7

2 Wave Motion 10

2.1 One-Dimensional Waves 10
2.2 Harmonic Waves 14
2.3 Phase and Phase Velocity 17
2.4 The Superposition Principle 20
2.5 The Complex Representation 21
2.6 Phasors and the Addition of Waves 23
2.7 Plane Waves 24
2.8 The Three-Dimensional Differential Wave Equation 27
2.9 Spherical Waves 28
2.10 Cylindrical Waves 31
Problems 32

3 Electromagnetic Theory, Photons, and Light 36

3.1 Basic Laws of Electromagnetic Theory 37
3.2 Electromagnetic Waves 44
3.3 Energy and Momentum 47
3.4 Radiation 58
3.5 Light in Bulk Matter 66
3.6 The Electromagnetic-Photon Spectrum 73
3.7 Quantum Field Theory 80
Problems 82

4 The Propagation of Light 86

4.1 Introduction 86
4.2 Rayleigh Scattering 86
4.3 Reflection 95
4.4 Refraction 100
4.5 Fermat's Principle 106
4.6 The Electromagnetic Approach 111
4.7 Total Internal Reflection 122
4.8 Optical Properties of Metals 127
4.9 Familiar Aspects of the Interaction of Light and Matter 131
4.10 The Stokes Treatment of Reflection and Refraction 136
4.11 Photons, Waves, and Probability 137
Problems 141

5 Geometrical Optics 149

5.1 Introductory Remarks 149
5.2 Lenses 150
5.3 Stops 171
5.4 Mirrors 175
5.5 Prisms 186
5.6 Fiberoptics 193
5.7 Optical Systems 201
5.8 Wavefront Shaping 226
5.9 Gravitational Lensing 231
Problems 234

6 More on Geometrical Optics 243

6.1 Thick Lenses and Lens Systems 243
6.2 Analytical Ray Tracing 246
6.3 Aberrations 253
6.4 GRIN Systems 273
6.5 Concluding Remarks 276
Problems 277

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

7 The Superposition of Waves 281
- **7.1** The Addition of Waves of the Same Frequency 282
- **7.2** The Addition of Waves of Different Frequency 294
- **7.3** Anharmonic Periodic Waves 302
- **7.4** Nonperiodic Waves 308

Problems 320

8 Polarization 325
- **8.1** The Nature of Polarized Light 325
- **8.2** Polarizers 331
- **8.3** Dichroism 333
- **8.4** Birefringence 336
- **8.5** Scattering and Polarization 344
- **8.6** Polarization by Reflection 348
- **8.7** Retarders 352
- **8.8** Circular Polarizers 357
- **8.9** Polarization of Polychromatic Light 358
- **8.10** Optical Activity 360
- **8.11** Induced Optical Effects—Optical Modulators 365
- **8.12** Liquid Crystals 370
- **8.13** A Mathematical Description of Polarization 372

Problems 379

9 Interference 385
- **9.1** General Considerations 386
- **9.2** Conditions for Interference 390
- **9.3** Wavefront-splitting Interferometers 393
- **9.4** Amplitude-splitting Interferometers 400
- **9.5** Types and Localization of Interference Fringes 414
- **9.6** Multiple-Beam Interference 416
- **9.7** Applications of Single and Multilayer Films 425
- **9.8** Applications of Interferometry 431

Problems 438

10 Diffraction 443
- **10.1** Preliminary Considerations 443
- **10.2** Fraunhofer Diffraction 452
- **10.3** Fresnel Diffraction 485
- **10.4** Kirchhoff's Scalar Diffraction Theory 510
- **10.5** Boundary Diffraction Waves 512

Problems 514

11 Fourier Optics 519
- **11.1** Introduction 519
- **11.2** Fourier Transforms 519
- **11.3** Optical Applications 529

Problems 556

12 Basics of Coherence Theory 560
- **12.1** Introduction 560
- **12.2** Visibility 562
- **12.3** The Mutual Coherence Function and the Degree of Coherence 566
- **12.4** Coherence and Stellar Interferometry 573

Problems 578

13 Modern Optics: Lasers and Other Topics 581
- **13.1** Lasers and Laserlight 581
- **13.2** Imagery — The Spatial Distribution of Optical Information 606
- **13.3** Holography 623
- **13.4** Nonlinear Optics 639

Problems 644

Appendix 1 649

Appendix 2 652

Table 1 653

Solutions to Selected Problems 658

Bibliography 685

Index 689

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

1 A Brief History

1.1 Prolegomenon

In chapters to come we will evolve a formal treatment of much of the science of Optics, with particular emphasis on aspects of contemporary interest. The subject embraces a vast body of knowledge accumulated over roughly three thousand years of the human scene. Before embarking on a study of the modern view of things optical, let's briefly trace the road that led us there, if for no other reason than to put it all in perspective.

1.2 In the Beginning

The origins of optical technology date back to remote antiquity. Exodus 38:8 (ca. 1200 B.C.E.) recounts how Bezaleel, while preparing the ark and tabernacle, recast "the looking-glasses of the women" into a brass laver (a ceremonial basin). Early mirrors were made of polished copper, bronze, and later on of speculum, a copper alloy rich in tin. Specimens have survived from ancient Egypt—a mirror in perfect condition was unearthed along with some tools from the workers' quarters near the pyramid of Sesostris II (ca. 1900 B.C.E.) in the Nile valley. The Greek philosophers Pythagoras, Democritus, Empedocles, Plato, Aristotle, and others developed several theories of the nature of light. The rectilinear propagation of light (p. 89) was known, as was the *Law of Reflection* (p. 97) enunciated by Euclid (300 B.C.E.) in his book *Catoptrics*. Hero of Alexandria attempted to explain both these phenomena by asserting that light traverses the shortest allowed path between two points. The burning glass (a positive lens used to start fires) was alluded to by Aristophanes in his comic play *The Clouds* (424 B.C.E.). The apparent bending of objects partly immersed in water (p. 102) is mentioned in Plato's *Republic*. Refraction was studied by Cleomedes (50 A.D.) and later by Claudius Ptolemy (130 A.D.) of Alexandria, who tabulated fairly precise measurements of the angles of incidence and refraction for several media (p. 101). It is clear from the accounts of the historian Pliny (23–79 A.D.) that the Romans also possessed burning glasses. Several glass and crystal spheres have been found among Roman ruins, and a planar convex lens was recovered in Pompeii. The Roman philosopher Seneca (3 B.C.E.–65 A.D.) pointed out that a glass globe filled with water could be used for magnifying purposes. And it is certainly possible that some Roman artisans may have used magnifying glasses to facilitate very fine detailed work.

After the fall of the Western Roman Empire (475 A.D.), which roughly marks the start of the Dark Ages, little or no scientific progress was made in Europe for a great while. The dominance of the Greco-Roman-Christian culture in the lands embracing the Mediterranean soon gave way by conquest to the rule of Allah. The center of scholarship shifted to the Arab world, and Optics was studied and extended, especially by Alhazen (ca. 1000 A.D.). He elaborated on the Law of Reflection, putting the angles of incidence and reflection in the same plane normal to the interface (p. 99); he studied spherical and parabolic mirrors and gave a detailed description of the human eye (p. 202).

By the latter part of the thirteenth century, Europe was only beginning to rouse from its intellectual stupor. Alhazen's work was translated into Latin, and it had a great effect on the writings of Robert Grosseteste (1175–1253), Bishop of Lincoln, and on the Polish mathematician Vitello (or Witelo), both of whom were influential in rekindling the study of Optics. Their works were known to the Franciscan Roger Bacon (1215–1294), who is considered by many to be the first scientist in the modern sense. He seems to have initiated the idea of using

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

lenses for correcting vision and even hinted at the possibility of combining lenses to form a telescope. Bacon also had some understanding of the way in which rays traverse a lens. After his death, Optics again languished. Even so, by the mid-1300s, European paintings were depicting monks wearing eyeglasses. And alchemists had come up with a liquid amalgam of tin and mercury that was rubbed onto the back of glass plates to make mirrors. Leonardo da Vinci (1452–1519) described the *camera obscura* (p. 215), later popularized by the work of Giovanni Battista Della Porta (1535–1615), who discussed multiple mirrors and combinations of positive and negative lenses in his *Magia naturalis* (1589).

This, for the most part, modest array of events constitutes what might be called the first period of Optics. It was undoubtedly a beginning—but on the whole a humble one. The whirlwind of accomplishment and excitement was to come later, in the seventeenth century.

1.3 From the Seventeenth Century

It is not clear who actually invented the refracting telescope, but records in the archives at The Hague show that on October 2, 1608, Hans Lippershey (1587–1619), a Dutch spectacle maker, applied for a patent on the device. Galileo Galilei (1564–1642), in Padua, heard about the invention and within several months had built his own instrument (p. 170), grinding the lenses by hand. The compound microscope was invented at just about the same time, possibly by the Dutchman Zacharias Janssen (1588–1632). The microscope's concave eyepiece was replaced with a convex lens by Francisco Fontana (1580–1656) of Naples, and a similar change in the telescope was introduced by Johannes Kepler (1571–1630). In 1611, Kepler published his *Dioptrice*. He had discovered total internal reflection (p. 122) and arrived at the small angle approximation to the Law of Refraction, in which case the incident and transmission angles are proportional. He evolved a treatment of first-order Optics for thin-lens systems and in his book describes the detailed operation of both the Keplerian (positive eyepiece) and Galilean (negative eyepiece) telescopes. Willebrord Snel (1591–1626), professor at Leyden, empirically discovered the long-hidden *Law of Refraction* (p. 100) in 1621—this was one of the great moments in Optics. By learning precisely how rays of light are redirected on traversing a boundary between two media, Snell in one swoop swung open the door to modern applied Optics.



Johannes Kepler (1571–1630). (Burndy Library.)

René Descartes (1596–1650) was the first to publish the now familiar formulation of the Law of Refraction in terms of sines. Descartes deduced the law using a model in which light was viewed as a pressure transmitted by an elastic medium; as he put it in his *La Dioptrique* (1637)

> recall the nature that I have attributed to light, when I said that it is nothing other than a certain motion or an action conceived in a very subtle matter, which fills the pores of all other bodies....

The universe was a plenum. Pierre de Fermat (1601–1665), taking exception to Descartes's assumptions, rederived the Law of Reflection (p. 106) from his own *Principle of Least Time* (1657).

The phenomenon of diffraction, that is, the deviation from rectilinear propagation that occurs when light advances



René Descartes by Frans Hals (1596–1650). (© Musées Nationaux.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Sir Isaac Newton (1642–1727). (Burndy Library.)

beyond an obstruction (p. 443), was first noted by Professor Francesco Maria Grimaldi (1618–1663) at the Jesuit College in Bologna. He had observed bands of light within the shadow of a rod illuminated by a small source. Robert Hooke (1635–1703), curator of experiments for the Royal Society, London, later also observed diffraction effects. He was the first to study the colored interference patterns (p. 400) generated by thin films (*Micrographia*, 1665). He proposed the idea that light was a rapid vibratory motion of the medium propagating at a very great speed. Moreover, "every pulse or vibration of the luminous body will generate a sphere"—this was the beginning of the wave theory. Within a year of Galileo's death, Isaac Newton (1642–1727) was born. The thrust of Newton's scientific effort was to build on direct observation and avoid speculative hypotheses. Thus he remained ambivalent for a long while about the actual nature of light. Was it corpuscular—a stream of particles, as some maintained? Or was light a wave in an all-pervading medium, the aether? At the age of 23, he began his now famous experiments on dispersion.

> I procured me a triangular glass prism to try therewith the celebrated phenomena of colours.

Newton concluded that white light was composed of a mixture of a whole range of independent colors (p. 189). He maintained that the corpuscles of light associated with the various colors excited the aether into characteristic vibrations. Even though his work simultaneously embraced both the wave and emission (corpuscular) theories, he did become more committed to the latter as he grew older. His main reason for rejecting the wave theory as it stood then was the daunting problem of explaining rectilinear propagation in terms of waves that spread out in all directions.

After some all-too-limited experiments, Newton gave up trying to remove chromatic aberration from refracting telescope lenses. Erroneously concluding that it could not be done, he turned to the design of reflectors. Sir Isaac's first reflecting telescope, completed in 1668, was only 6 inches long and 1 inch in diameter, but it magnified some 30 times.

At about the same time that Newton was emphasizing the emission theory in England, Christiaan Huygens (1629–1695), on the continent, was greatly extending the wave theory. Unlike Descartes, Hooke, and Newton, Huygens correctly concluded that light effectively slowed down on entering more dense media. He was able to derive the Laws of Reflection and Refraction and even explained the double refraction of calcite (p. 337), using his wave theory. And it was while working with calcite that he discovered the phenomenon of *polarization* (p. 325).

> As there are two different refractions, I conceived also that there are two different emanations of the waves of light....

Thus light was either a stream of particles or a rapid undulation of aethereal matter. In any case, it was generally agreed that its speed was exceedingly large. Indeed, many believed that light propagated instantaneously, a notion that went back at least as far as Aristotle. The fact that it was finite was



Christiaan Huygens (1629–1695). (Rijksmuseum voor de geschiedenis der natuurwetenschappen, courtesy AIP Emilio Segré Visual Archives.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

determined by the Dane Ole Christensen Römer (1644–1710). Jupiter's nearest moon, Io, has an orbit about that planet that is nearly in the plane of Jupiter's own orbit around the Sun. Römer made a careful study of the eclipses of Io as it moved through the shadow behind Jupiter. In 1676 he predicted that on November 9th Io would emerge from the dark some 10 minutes later than would have been expected on the basis of its yearly averaged motion. Precisely on schedule, Io performed as predicted, a phenomenon Römer correctly explained as arising from the finite speed of light. He was able to determine that light took about 22 minutes to traverse the diameter of the Earth's orbit around the Sun—a distance of about 186 million miles. Huygens and Newton, among others, were quite convinced of the validity of Römer's work. Independently estimating the Earth's orbital diameter, they assigned values to c equivalent to 2.3×10^8 m/s and 2.4×10^8 m/s, respectively.*

The great weight of Newton's opinion hung like a shroud over the wave theory during the eighteenth century, all but stifling its advocates. Despite this, the prominent mathematician Leonhard Euler (1707–1783) was a devotee of the wave theory, even if an unheeded one. Euler proposed that the undesirable color effects seen in a lens were absent in the eye (which is an erroneous assumption) because the different media present negated dispersion. He suggested that achromatic lenses (p. 269) might be constructed in a similar way. Enthused by this work, Samuel Klingenstjerna (1698–1765), a professor at Upsala, reperformed Newton's experiments on achromatism and determined them to be in error. Klingenstjerna was in communication with a London optician, John Dollond (1706–1761), who was observing similar results. Dollond finally, in 1758, combined two elements, one of crown and the other of flint glass, to form a single achromatic lens. Incidentally, Dollond's invention was actually preceded by the unpublished work of the amateur scientist Chester Moor Hall (1703–1771) in Essex.

1.4 The Nineteenth Century

The wave theory of light was reborn at the hands of Dr. Thomas Young (1773–1829), one of the truly great minds of the century. In 1801, 1802, and 1803, he read papers before the Royal Society extolling the wave theory and adding to it a new fundamental concept, the so-called *Principle of Interference* (p. 385):

> When two undulations, from different origins, coincide either perfectly or very nearly in direction, their joint effect is a combination of the motions belonging to each.

He was able to explain the colored fringes of thin films and determined wavelengths of various colors using Newton's data. Even though Young, time and again, maintained that his conceptions had their very origins in the research of Newton, he was severely attacked. In a series of articles, probably written by Lord Brougham, in the *Edinburgh Review*, Young's papers were said to be "destitute of every species of merit."



Augustin Jean Fresnel (1788–1827). (Cultural Service of the French Embassy.)

Augustin Jean Fresnel (1788–1827), born in Broglie, Normandy, began his brilliant revival of the wave theory in France, unaware of the efforts of Young some 13 years earlier. Fresnel synthesized the concepts of Huygens's wave description and the interference principle (p. 444). The mode of propagation of a primary wave was viewed as a succession of spherical secondary wavelets, which overlapped and interfered to reform the advancing primary wave as it would appear an instant later. In Fresnel's words:

> The vibrations of a luminous wave in any one of its points may be considered as the sum of the elementary movements con-

*A. Wróblewski, *Am. J. Phys.* **53,** 620 (1985).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

> veyed to it at the same moment, from the separate action of all the portions of the unobstructed wave considered in any one of its anterior positions.

These waves were presumed to be longitudinal, in analogy with sound waves in air. Fresnel was able to calculate the diffraction patterns arising from various obstacles and apertures (p. 444) and satisfactorily accounted for rectilinear propagation in homogeneous isotropic media, thus dispelling Newton's main objection to the undulatory theory. When finally apprised of Young's priority to the interference principle, a somewhat disappointed Fresnel nonetheless wrote to Young telling him that he was consoled by finding himself in such good company—the two great men became allies.

Huygens was aware of the phenomenon of polarization arising in calcite crystals, as was Newton. Indeed, the latter in his *Opticks* stated,

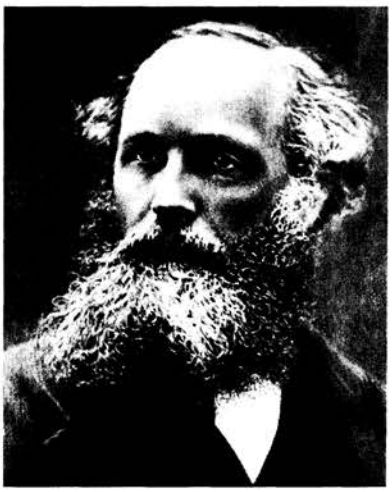
> Every Ray of Light has therefore two opposite Sides....

It was not until 1808 that Étienne Louis Malus (1775–1812) discovered that this two-sidedness of light also arose upon reflection (p. 348); the phenomenon was not inherent to crystalline media. Fresnel and Dominique François Arago (1786–1853) then conducted a series of experiments to determine the effect of polarization on interference, but the results were utterly inexplicable within the framework of their longitudinal wave picture. This was a dark hour indeed. For several years Young, Arago, and Fresnel wrestled with the problem until finally Young suggested that the aethereal vibration might be *transverse* as is a wave on a string. The two-sidedness of light was then simply a manifestation of the two orthogonal vibrations of the aether, transverse to the ray direction. Fresnel went on to evolve a mechanistic description of aether oscillations, which led to his now famous formulas for the amplitudes of reflected and transmitted light (p. 113). By 1825 the emission (or corpuscular) theory had only a few tenacious advocates.

The first terrestrial determination of the speed of light was performed by Armand Hippolyte Louis Fizeau (1819–1896) in 1849. His apparatus, consisting of a rotating toothed wheel and a distant mirror (8633 m), was set up in the suburbs of Paris from Suresnes to Montmartre. A pulse of light leaving an opening in the wheel struck the mirror and returned. By adjusting the known rotational speed of the wheel, the returning pulse could be made either to pass through an opening and be seen or to be obstructed by a tooth. Fizeau arrived at a value of the speed of light equal to 315 300 km/s. His colleague Jean Bernard Léon Foucault (1819–1868) was also involved in research on the speed of light. In 1834 Charles Wheatstone (1802–1875) had designed a rotating-mirror arrangement in order to measure the duration of an electric spark. Using this scheme, Arago had proposed to measure the speed of light in dense media but was never able to carry out the experiment. Foucault took up the work, which was later to provide material for his doctoral thesis. On May 6, 1850, he reported to the Academy of Sciences that the speed of light in water was *less* than that in air. This result was in direct conflict with Newton's formulation of the emission theory and a hard blow to its few remaining devotees.

While all of this was happening in Optics, quite independently, the study of electricity and magnetism was also bearing fruit. In 1845 the master experimentalist Michael Faraday (1791–1867) established an interrelationship between electromagnetism and light when he found that the polarization direction of a beam could be altered by a strong magnetic field applied to the medium. James Clerk Maxwell (1831–1879) brilliantly summarized and extended all the empirical knowledge on the subject in a single set of mathematical equations. Beginning with this remarkably succinct and beautifully symmetrical synthesis, he was able to show, purely theoretically, that the electromagnetic field could propagate as a transverse wave in the luminiferous aether (p. 44).

Solving for the speed of the wave, Maxwell arrived at an expression in terms of electric and magnetic properties of the



James Clerk Maxwell (1831–1879). (AIP Emilio Segrè Visual Archives.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

medium ($c = 1/\sqrt{\epsilon_0\mu_0}$). Upon substituting known empirically determined values for these quantities, he obtained a numerical result equal to the measured speed of light! The conclusion was inescapable—*light was "an electromagnetic disturbance in the form of waves" propagated through the aether*. Maxwell died at the age of 48, eight years too soon to see the experimental confirmation of his insights and far too soon for physics. Heinrich Rudolf Hertz (1857–1894) verified the existence of long electromagnetic waves by generating and detecting them in an extensive series of experiments published in 1888.

The acceptance of the wave theory of light seemed to necessitate an equal acceptance of the existence of an all-pervading substratum, the luminiferous aether. If there were waves, it seemed obvious that there must be a supporting medium. Quite naturally, a great deal of scientific effort went into determining the physical nature of the aether, yet it would have to possess some rather strange properties. It had to be so tenuous as to allow an apparently unimpeded motion of celestial bodies. At the same time, it could support the exceedingly high-frequency ($\sim 10^{15}$ Hz) oscillations of light traveling at 186 000 miles per second. That implied remarkably strong restoring forces within the aethereal substance. The speed at which a wave advances through a medium is dependent on the characteristics of the disturbed substratum and not on any motion of the source. This is in contrast to the behavior of a stream of particles whose speed with respect to the source is the essential parameter.

Certain aspects of the nature of aether intrude when studying the optics of moving objects, and it was this area of research, evolving quietly on its own, that ultimately led to the next great turning point. In 1725 James Bradley (1693–1762), then Savilian Professor of Astronomy at Oxford, attempted to measure the distance to a star by observing its orientation at two different times of the year. The position of the Earth changed as it orbited around the Sun and thereby provided a large baseline for triangulation on the star. To his surprise, Bradley found that the "fixed" stars displayed an apparent systematic movement related to the direction of motion of the Earth in orbit and not dependent, as had been anticipated, on the Earth's position in space. This so-called *stellar aberration* is analogous to the well-known falling-raindrop situation. A raindrop, although traveling vertically with respect to an observer at rest on the Earth, will appear to change its incident angle when the observer is in motion. Thus a corpuscular model of light could explain stellar aberration rather handily. Alternatively, the wave theory also offers a satisfactory explanation provided that *the aether remains totally undisturbed as the Earth plows through it.*

In response to speculation as to whether the Earth's motion through the aether might result in an observable difference between light from terrestrial and extraterrestrial sources, Arago set out to examine the problem experimentally. He found that there were no such observable differences. Light behaved just as if the Earth were at rest with respect to the aether. To explain these results, Fresnel suggested in effect that light was partially dragged along as it traversed a transparent medium in motion. Experiments by Fizeau, in which light beams passed down moving columns of water, and by Sir George Biddell Airy (1801–1892), who used a water-filled telescope in 1871 to examine stellar aberration, both seemed to confirm Fresnel's drag hypothesis. Assuming an aether at *absolute rest*, Hendrik Antoon Lorentz (1853–1928) derived a theory that encompassed Fresnel's ideas.

In 1879 in a letter to D. P. Todd of the U.S. Nautical Almanac Office, Maxwell suggested a scheme for measuring the speed at which the solar system moved with respect to the luminiferous aether. The American physicist Albert Abraham Michelson (1852–1931), then a naval instructor, took up the idea. Michelson, at the tender age of 26, had already established a favorable reputation by performing an extremely precise determination of the speed of light. A few years later, he began an experiment to measure the effect of the Earth's motion through the aether. Since the speed of light in aether is constant and the Earth, in turn, presumably moves in relation to the aether (orbital speed of 67 000 mi/h), the speed of light measured with respect to the Earth should be affected by the planet's motion. In 1881 he published his findings. There was no detectable motion of the Earth with respect to the aether—the aether was stationary. But the decisiveness of this surprising result was blunted somewhat when Lorentz pointed out an oversight in the calculation. Several years later Michelson, then professor of physics at Case School of Applied Science in Cleveland, Ohio, joined with Edward Williams Morley (1838–1923), a well-known professor of chemistry at Western Reserve, to redo the experiment with considerably greater precision. Amazingly enough, their results, published in 1887, once again were negative:

> It appears from all that precedes reasonably certain that if there be any relative motion between the earth and the luminiferous aether, it must be small; quite small enough entirely to refute Fresnel's explanation of aberration.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Thus, whereas an explanation of stellar aberration within the context of the wave theory required the existence of a relative motion between Earth and aether, the Michelson–Morley Experiment refuted that possibility. Moreover, the findings of Fizeau and Airy necessitated the inclusion of a partial drag of light due to motion of the medium.

1.5 Twentieth-Century Optics

Jules Henri Poincaré (1854–1912) was perhaps the first to grasp the significance of the experimental inability to observe any effects of motion relative to the aether. In 1899 he began to make his views known, and in 1900 he said:

> Our aether, does it really exist? I do not believe that more precise observations could ever reveal anything more than *relative* displacements.

In 1905 Albert Einstein (1879–1955) introduced his *Special Theory of Relativity*, in which he too, quite independently, rejected the aether hypothesis.

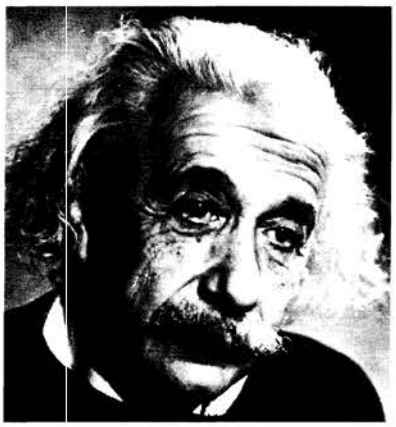
> The introduction of a "luminiferous aether" will prove to be superfluous inasmuch as the view here to be developed will not require an "absolutely stationary space."

He further postulated:

> light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body.

The experiments of Fizeau, Airy, and Michelson–Morley were then explained quite naturally within the framework of Einstein's relativistic kinematics.* Deprived of the aether, physicists simply had to get used to the idea that electromagnetic waves could propagate through free space—there was no alternative. Light was now envisaged as a self-sustaining wave with the conceptual emphasis passing from aether to field. The electromagnetic wave became an entity in itself.

On October 19, 1900, Max Karl Ernst Ludwig Planck (1858–1947) read a paper before the German Physical Society in which he introduced the hesitant beginnings of what was to become yet another great revolution in scientific thought—*Quantum Mechanics*, a theory embracing submicroscopic phenomena (p. 51). In 1905, boldly building on these ideas, Einstein proposed a new form of corpuscular theory in which he asserted that light consisted of globs or "particles" of energy. Each such quantum of radiant energy or *photon*,* as it came to be called, had an energy proportional to its frequency ν, that is, $\mathscr{E} = h\nu$, where h is known as Planck's constant (Fig. 1.1). By the end of the 1920s, through the efforts of Bohr, Born, Heisenberg, Schrödinger, De Broglie, Pauli, Dirac, and others, Quantum Mechanics had become a well-verified theory. It gradually became evident that the concepts of particle and wave, which in the macroscopic world seem so obviously mutually exclusive, must be merged in the submicroscopic domain. The mental image of an atomic particle (e.g., electrons and neutrons) as a minute localized lump of matter would no longer suffice. Indeed, it was found that these "particles" could generate interference and diffraction patterns in precisely the same way as would light (p. 399). Thus photons, protons, electrons, neutrons, and so forth—the whole lot—have both particle and wave manifestations. Still, the matter was by no means settled. "Every physicist thinks that he knows what a photon is," wrote Einstein. "I spent my life to find out what a photon is and I still don't know it."



Albert Einstein (1879–1955).

Relativity liberated light from the aether and showed the kinship between mass and energy (via $\mathscr{E}_0 = mc^2$). What seemed to be two almost antithetical quantities now became interchangeable. Quantum Mechanics went on to establish that

*See, for example, *Special Relativity* by French, Chapter 5.

*The word *photon* was coined by G. N. Lewis, *Nature*, December 18, 1926.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

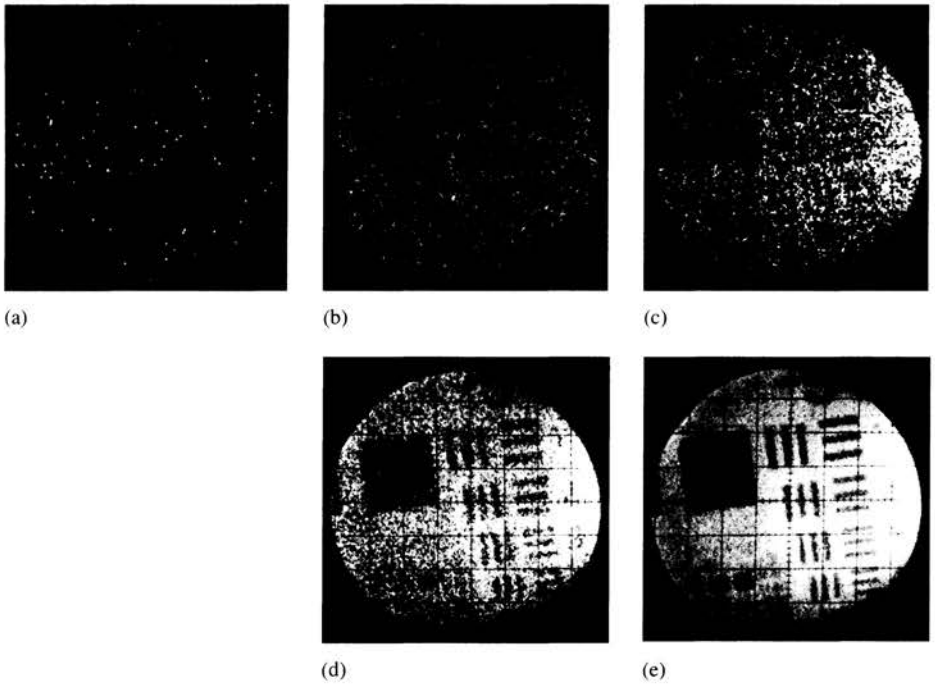


FIGURE 1.1 A rather convincing illustration of the particle nature of light. This sequence of photos was made using a position-sensing photomultiplier tube illuminated by an (8.5×10^3 count-per-second) image of a bar chart. The exposure times were (*a*) 8 ms, (*b*) 125 ms, (*c*) 1 s, (*d*) 10 s, and (*e*) 100 s. Each dot can be interpreted as the arrival of a single photon. (Photos courtesy of ITT Corporation, Electro-Optical Products Division, Tube and Sensor Laboratories, Fort Wayne, Indiana.)

a particle* of momentum p had an associated wavelength λ, such that $p = h/\lambda$. The neutrino, a neutral particle presumably having zero rest mass, was postulated for theoretical reasons in 1930 by Wolfgang Pauli (1900–1958) and verified experimentally in the 1950s. The easy images of submicroscopic specks of matter became untenable, and the wave-particle dichotomy dissolved into a duality.

Quantum Mechanics also treats the manner in which light is absorbed and emitted by atoms (p. 63). Suppose we cause a gas to glow by heating it or passing an electrical discharge through it. The light emitted is characteristic of the very structure of the atoms constituting the gas. Spectroscopy, which is the branch of Optics dealing with spectrum analysis (p. 73), developed from the research of Newton. William Hyde Wollaston (1766–1828) made the earliest observations of the dark lines in the solar spectrum (1802). Because of the slit-shaped aperture generally used in spectroscopes, the output consisted of narrow colored bands of light, the so-called *spectral lines*. Working independently, Joseph Fraunhofer (1787–1826) greatly extended the subject. After accidentally discovering the double line of sodium (p. 270), he went on to study sunlight and made the first wavelength determinations using diffraction gratings (p. 476). Gustav Robert Kirchhoff (1824–1887) and Robert Wilhelm Bunsen (1811–1899), working together at Heidelberg, established that each kind of atom had its own signature in a characteristic array of spectral lines. And in 1913 Niels Henrik David Bohr (1885–1962) set forth a precursory quantum theory of the hydrogen atom, which was able to predict the wavelengths of its emission spectrum. The light emitted by an atom is now understood to arise from its outermost electrons (p. 63). The process is the domain of modern quantum theory, which describes the most minute details with incredible precision and beauty.

The flourishing of applied Optics in the second half of the twentieth century represents a renaissance in itself. In the 1950s several workers began to inculcate Optics with the mathematical techniques and insights of communications theory. Just as the idea of momentum provides another dimension in which to visualize aspects of mechanics, the concept of spatial frequency offers a rich new way of appreciating a broad range of optical phenomena. Bound together by the mathe-

*Perhaps it might help if we just called them all *wavicles*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

matical formalism of Fourier analysis (p. 302), the outgrowths of this contemporary emphasis have been far-reaching. Of particular interest are the theory of image formation and evaluation (p. 529), the *transfer functions* (p. 550), and the idea of *spatial filtering* (p. 318).

The advent of the high-speed digital computer brought with it a vast improvement in the design of complex optical systems. Aspherical lens elements (p. 150) took on renewed practical significance, and the *diffraction-limited* system with an appreciable field of view became a reality. The technique of ion bombardment polishing, in which one atom at a time is chipped away, was introduced to meet the need for extreme precision in the preparation of optical elements. The use of single and multilayer thin-film coatings (reflecting, antireflecting, etc.) became commonplace (p. 425). Fiberoptics evolved into a practical communications tool (p. 197), and thin-film light guides continued to be studied. A great deal of attention was paid to the infrared end of the spectrum (surveillance systems, missile guidance, etc.), and this in turn stimulated the development of infrared materials. Plastics began to be used extensively in Optics (lens elements, replica gratings, fibers, aspherics, etc.). A new class of partially vitrified glass ceramics with exceedingly low thermal expansion was developed. A resurgence in the construction of astronomical observatories (both terrestrial and extraterrestrial) operating across the whole spectrum was well under way by the end of the 1960s and vigorously sustained in the 1980s and 1990s (p. 222).

The first laser was built in 1960, and within a decade laser-beams spanned the range from infrared to ultraviolet. The availability of high-power coherent sources led to the discovery of a number of new optical effects (harmonic generation, frequency mixing, etc.) and thence to a panorama of marvelous new devices. The technology needed to produce a practicable optical communications system developed rapidly. The sophisticated use of crystals in devices such as second-harmonic generators (p. 641), electro-optic and acousto-optic modulators, and the like spurred a great deal of contemporary research in crystal optics. The wavefront reconstruction technique known as *holography* (p. 623), which produces magnificent three-dimensional images, was found to have numerous additional applications (nondestructive testing, data storage, etc.).

The military orientation of much of the developmental work in the 1960s continued in the 1970s, 1980s, and the 1990s with added vigor. That technological interest in Optics ranges across the spectrum from "smart bombs" and spy satellites to "death rays" and infrared gadgets that see in the dark. But economic considerations coupled with the need to improve the quality of life have brought products of the discipline into the consumer marketplace as never before. Today lasers are in use everywhere: reading videodiscs in living rooms, cutting steel in factories, scanning labels in supermarkets, and performing surgery in hospitals. Millions of optical display systems on clocks and calculators and computers are blinking all around the world. The almost exclusive use, for the last one hundred years, of electrical signals to handle and transmit data is now rapidly giving way to more efficient optical techniques. A far-reaching revolution in the methods of processing and communicating information is quietly taking place, a revolution that will continue to change our lives in the years ahead.

Profound insights are slow in coming. What few we have took over three thousand years to glean, even though the pace is ever quickening. It is marvelous indeed to watch the answer subtly change while the question immutably remains—*what is light*?*

*For more reading on the history of Optics, see F. Cajori, *A History of Physics*, and V. Ronchi, *The Nature of Light*. Excerpts from a number of original papers can conveniently be found in W. F. Magie, *A Source Book in Physics*, and in M. H. Shamos, *Great Experiments in Physics*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

2 Wave Motion

The issue of the actual nature of light is central to a complete treatment of Optics, and we will struggle with it throughout this work. The straightforward question "Is light a wave phenomenon or a particle phenomenon?" is far more complicated than it might at first seem. For example, the essential feature of a particle is its localization; it exists in a well-defined, "small" region of space. Practically, we tend to take something familiar like a ball or a pebble and shrink it down in imagination until it becomes vanishingly small, and that's a "particle," or at least the basis for the concept of "particle." But a ball interacts with its environment; it has a gravitational field that interacts with the Earth (and the Moon, and Sun, etc.). This field, which spreads out into space—whatever *it* is—cannot be separated from the ball; it is an inextricable part of the ball just as it is an inextricable part of the definition of "particle." Real particles interact via fields, and, in a sense, the field is the particle and the particle is the field. That little conundrum is the domain of Quantum Field Theory, a discipline we'll talk more about later (p. 139). Suffice it to say now that if light is a stream of submicroscopic particles (photons), they are by no means "ordinary" miniball classical particles.

On the other hand, the essential feature of a wave is its nonlocalization. *A classical traveling wave is a self-sustaining disturbance of a medium, which moves through space transporting energy and momentum.* We tend to think of the ideal wave as a continuous entity that exists over an extended region. But when we look closely at real waves (such as waves on strings), we see composite phenomena comprised of vast numbers of particles moving in concert. The media supporting these waves are atomic (i.e., particulate), and so the waves are not continuous entities in and of themselves. The only possible exception might be the electromagnetic wave. Conceptually, the classical electromagnetic wave (p. 44) is supposed to be a continuous entity, and *it* serves as the model for the very notion of wave as distinct from particle. But in the past century we found that the energy of an electromagnetic wave is *not* distributed continuously. The classical formulation of the electromagnetic theory of light, however wonderful it is on a macroscopic level, is profoundly wanting on a microscopic level. Einstein was the first to suggest that the electromagnetic wave, which we perceive macroscopically, is the statistical manifestation of a fundamentally granular underlying microscopic phenomenon (p. 51). In the subatomic domain, the classical concept of a physical wave is an illusion. Still, in the large-scale regime in which we ordinarily work, electromagnetic waves seem real enough and classical theory applies superbly well.

Because both the classical and quantum-mechanical treatments of light make use of the mathematical description of waves, this chapter lays out the basics of what both formalisms will need. The ideas we develop here will apply to all physical waves from a surface tension ripple in a cup of tea to a pulse of light reaching us from some distant galaxy.

2.1 One-Dimensional Waves

An essential aspect of a traveling wave is that it is a self-sustaining disturbance of the medium through which it propagates. The most familiar waves, and the easiest to visualize (Fig. 2.1), are the mechanical waves, among which are waves on strings, surface waves on liquids, sound waves in the air, and compression waves in both solids and fluids. Sound waves are **longitudinal**—*the medium is displaced in the direction of motion of the wave*. Waves on a string (and electromagnetic waves) are **transverse**—*the medium is displaced in a direction perpendicular to that of the motion of the wave*. In all

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

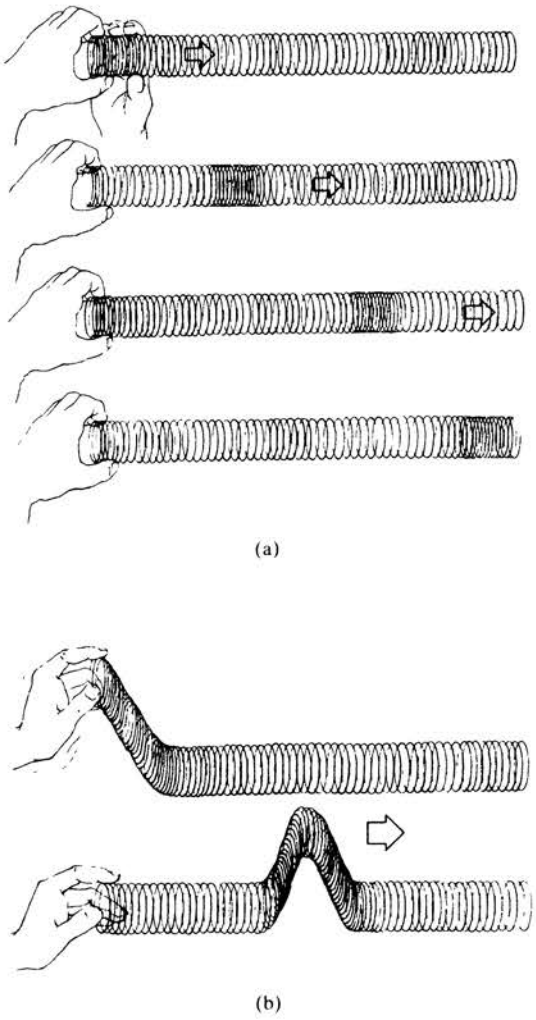


Figure 2.1 (*a*) A longitudinal wave in a spring. (*b*) A transverse wave in a spring.

cases, although the energy-carrying disturbance advances through the medium, the individual participating atoms remain in the vicinity of their equilibrium positions: *the disturbance advances, not the material medium.* That's one of several crucial features of a wave that distinguishes it from a stream of particles. The wind blowing across a field sets up "waves of grain" that sweep by, even though each stalk only sways in place. Leonardo da Vinci seems to have been the first person to recognize that a wave does not transport the medium through which it travels, and it is precisely this property that allows waves to propagate at very great speeds.

Envision some such disturbance ψ moving in the positive x-direction with a constant speed v. The specific nature of the disturbance is at the moment unimportant. It might be the vertical displacement of the string in Fig. 2.2 or the magnitude of an electric or magnetic field associated with an electromagnetic wave (or even the quantum-mechanical probability amplitude of a matter wave).

Since the disturbance is moving, it must be a function of both position and time;

$$\psi(x, t) = f(x, t) \tag{2.1}$$

where $f(x, t)$ corresponds to some specific function or wave shape. This is represented in Fig. 2.3*a* which shows a pulse traveling in the stationary coordinate system S at a speed v. The shape of the disturbance at any instant, say $t = 0$, can be found by holding time constant at that value. In this case,

$$\psi(x, t)|_{t=0} = f(x, 0) = f(x) \tag{2.2}$$

represents the **profile** of the wave at that time. For example, if $f(x) = e^{-ax^2}$, where a is a constant, the profile has the shape of

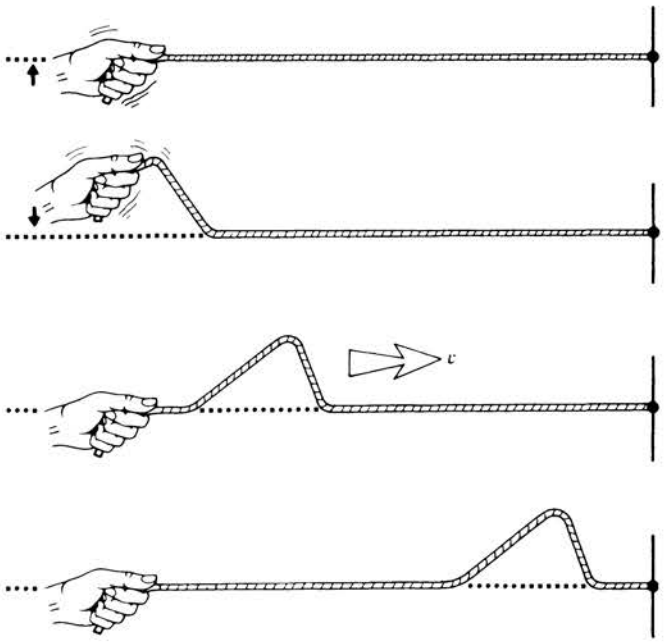


Figure 2.2 A wave on a string.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

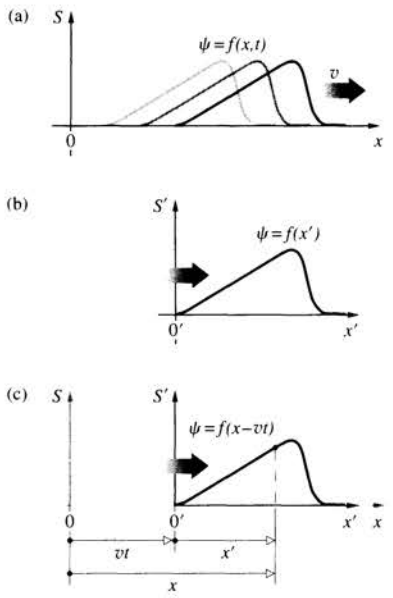


Figure 2.3 Moving reference frame.

a bell; that is, it is a **Gaussian function**. (Squaring the x makes it symmetrical around the $x = 0$ axis.) Setting $t = 0$ is analogous to taking a "photograph" of the pulse as it travels by.

For the moment we limit ourselves to a wave that *does not change its shape* as it progresses through space. After a time t the pulse has moved along the x-axis a distance vt, but in all other respects it remains unaltered. We now introduce a coordinate system S', that travels along with the pulse (Fig. 2.3*b*) at the speed v. In this system ψ is no longer a function of time, and as we move along with S', we see a stationary constant profile described by Eq. (2.2). Here, the coordinate is x' rather than x, so that

$$\psi = f(x') \tag{2.3}$$

The disturbance looks the same at any value of t in S' as it did at $t = 0$ in S when S and S' had a common origin (Fig. 2.3*c*).

We now want to rewrite Eq. (2.3) in terms of x to get the wave as it would be described by someone at rest in S. It follows from Fig. 2.3*c* that

$$x' = x - vt \tag{2.4}$$

and substituting into Eq. (2.3)

$$\psi(x, t) = f(x - vt) \tag{2.5}$$

This then represents the most general form of the one-dimensional **wavefunction**. To be more specific, we have only to choose a shape, Eq. (2.2), and then substitute $(x - vt)$ for x in $f(x)$. ***The resulting expression describes a wave having the desired profile, moving in the positive x-direction with a speed*** v. Thus, $\psi(x, t) = e^{-a(x-vt)^2}$ is a bell-shaped wave.

To see how this all works in a bit more detail, let's unfold the analysis for a specific pulse, for example, $\psi(x) = 3/[10x^2 + 1] = f(x)$. That profile is plotted in Fig. 2.4*a*, and if it was a wave on a rope, ψ would be the vertical displacement and we might even replace it by the symbol y. Whether ψ represents displacement or pressure or electric field, we now have

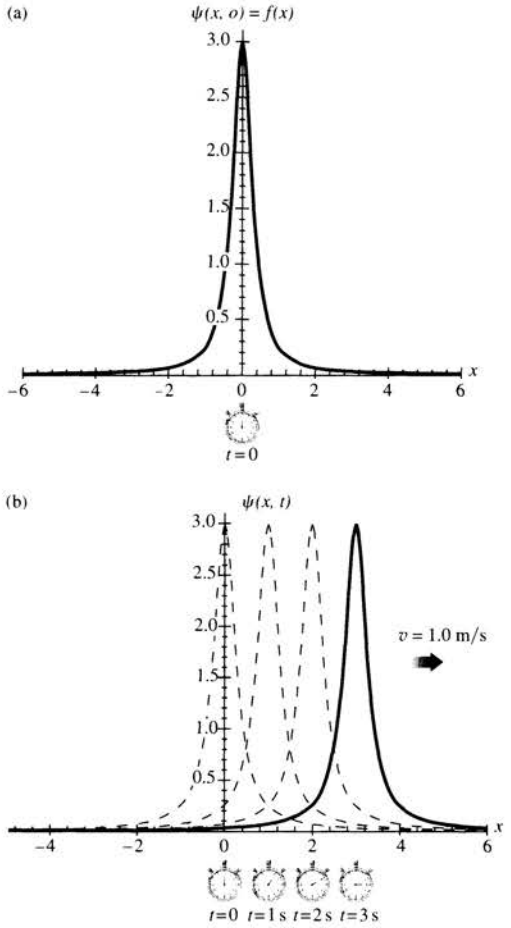


Figure 2.4 (*a*) The profile of a pulse given by the function $f(x) = 3/(10x^2 + 1)$. (*b*) The profile shown in (*a*) is now moving as a wave, $\psi(x, t) = 3/[10(x - vt)^2 + 1]$, to the right. It has a speed of 1 m/s and advances in the positive x-direction.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

the profile of the disturbance. To turn $f(x)$ into $\psi(x, t)$, that is, to turn it into the description of a wave moving in the positive x-direction at a speed v, we replace x wherever it appears in $f(x)$ by $(x - vt)$, thereby yielding $\psi(x, t) = 3/[10(x - vt)^2 + 1]$. If v is arbitrarily set equal to, say, 1.0 m/s and the function is plotted successively at $t = 0$, $t = 1$ s, $t = 2$ s, and $t = 3$ s, we get Fig. 2.4*b*, which shows the pulse sailing off to the right at 1.0 m/s, just the way it's supposed to. Incidentally, had we substituted $(x + vt)$ for x in the profile function, the resulting wave would move off to the left.

If we check the form of Eq. (2.5) by examining ψ after an increase in time of Δt and a corresponding increase of $v\,\Delta t$ in x, we find

$$f[(x + v\,\Delta t) - v(t + \Delta t)] = f(x - vt)$$

and the profile is unaltered.

Similarly, if the wave was traveling in the negative x-direction, that is, to the left, Eq. (2.5) would become

$$\psi = f(x + vt), \quad \text{with} \quad v > 0 \tag{2.6}$$

We may conclude therefore that, regardless of the shape of the disturbance, the variables x and t must appear in the function as a unit, that is, as a single variable in the form $(x \mp vt)$. Equation (2.5) is often expressed equivalently as some function of $(t - x/v)$, since

$$f(x - vt) = F\left(-\frac{x - vt}{v}\right) = F(t - x/v) \tag{2.7}$$

The pulse shown in Fig. 2.2 and the disturbance described by Eq. (2.5) are spoken of as *one-dimensional* because the waves sweep over points lying on a line—it takes only one space variable to specify them. Don't be confused by the fact that in this particular case the rope happens to rise up into a second dimension. In contrast, a two-dimensional wave propagates out across a surface, like the ripples on a pond, and can be described by two space variables.

2.1.1 The Differential Wave Equation

In 1747 Jean Le Rond d'Alembert introduced partial differential equations into the mathematical treatment of physics. That same year, he wrote an article on the motion of vibrating strings in which the so-called *differential wave equation* appears for the first time. This linear, homogeneous, second-order, partial differential equation is usually taken as the defining expression for physical waves in a lossless medium. There are lots of different kinds of waves, and each is described by its own wavefunction $\psi(x)$. Some are written in terms of pressure, or displacement, while others deal with electromagnetic fields, but remarkably all such wavefunctions are solutions of the same differential wave equation. The reason it's a *partial* differential equation is that the wave must be a function of several independent variables, namely, those of space and time. A *linear* differential equation is essentially one consisting of two or more terms, each composed of a constant multiplying a function $\psi(x)$ or its derivatives. The relevant point is that each such term must appear only to the first power; nor can there be any cross products of ψ with its derivatives, or of its derivatives. Recall that the *order* of a differential equation equals the order of the highest derivative in that equation. Furthermore, if a differential equation is of order N, the solution will contain N arbitrary constants.

We now derive the one-dimensional form of the wave equation guided by the foreknowledge (p. 14) that the most basic of waves traveling at a fixed speed requires two constants (amplitude and frequency or wavelength) to specify it, and this suggests second derivatives. Because there are two independent variables (here, x and t) we can take the derivative of $\psi(x, t)$ with respect to either x or t. This is done by just differentiating with respect to one variable and treating the other as if it were constant. The usual rules for differentiation apply, but to make the distinction evident the partial derivative is written as $\partial/\partial x$.

To relate the space and time dependencies of $\psi(x, t)$, take the partial derivative of $\psi(x, t) = f(x')$ with respect to x, holding t constant. Using $x' = x \mp vt$, and inasmuch as

$$\frac{\partial \psi}{\partial x} = \frac{\partial f}{\partial x}$$

$$\frac{\partial \psi}{\partial x} = \frac{\partial f}{\partial x'}\frac{\partial x'}{\partial x} = \frac{\partial f}{\partial x'} \tag{2.8}$$

because $$\frac{\partial x'}{\partial x} = \frac{\partial (x \mp vt)}{\partial x} = 1$$

Holding x constant, the partial derivative with respect to time is

$$\frac{\partial \psi}{\partial t} = \frac{\partial f}{\partial x'}\frac{\partial x'}{\partial t} = \frac{\partial f}{\partial x'}(\mp v) = \mp v\frac{\partial f}{\partial x'} \tag{2.9}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Combining Eqs. (2.8) and (2.9) yields

$$\frac{\partial \psi}{\partial t} = \mp v \frac{\partial \psi}{\partial x}$$

This says that the rate of change of ψ with t and with x are equal, to within a multiplicative constant, as shown in Fig. 2.5. The second partial derivatives of Eqs. (2.8) and (2.9) are

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{\partial^2 f}{\partial x'^2} \tag{2.10}$$

and

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial}{\partial t}\left(\mp v \frac{\partial f}{\partial x'}\right) = \mp v \frac{\partial}{\partial x'}\left(\frac{\partial f}{\partial t}\right)$$

Since

$$\frac{\partial \psi}{\partial t} = \frac{\partial f}{\partial t}$$

$$\frac{\partial^2 \psi}{\partial t^2} = \mp v \frac{\partial}{\partial x'}\left(\frac{\partial \psi}{\partial t}\right)$$

It follows, using Eq. (2.9), that

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \frac{\partial^2 f}{\partial x'^2}$$

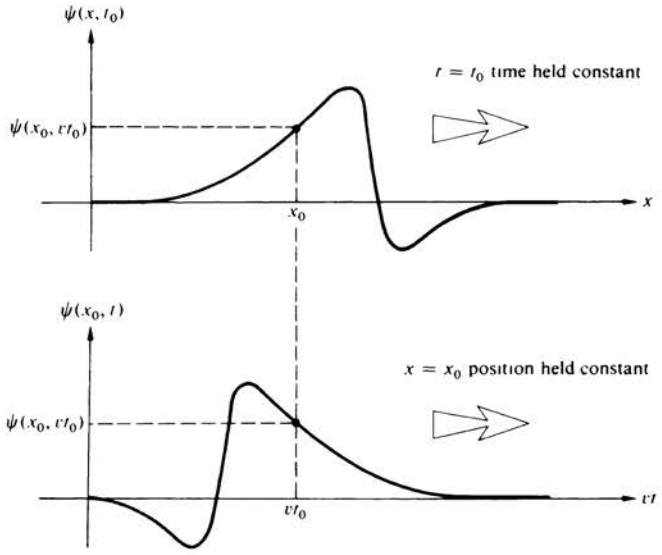


Figure 2.5 Variation of ψ with x and t.

Combining this with Eq. (2.10), we obtain

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} \tag{2.11}$$

which is the desired one-dimensional **differential wave equation**. Note that this is a so-called *homogeneous* differential equation: it doesn't contain a term (such as a "force" or a "source") involving only independent variables. In other words, ψ is in each term of the equation, and that means that if ψ is a solution any multiple of ψ will also be a solution. Equation 2.11 is ***the wave equation for undamped systems*** that do not contain sources in the region under consideration. The effects of damping can be described by adding in a $\partial\psi/\partial t$ term to form a more general wave equation, but we'll come back to that later (p. 71).

As a rule, partial differential equations arise when the system being described is continuous. The fact that time is one of the independent variables reflects the continuity of temporal change in the process under analysis. Field theories, in general, treat continuous distributions of quantities in space and time and so take the form of partial differential equations. Maxwell's formulation of electromagnetism, which is a field theory, yields a variation of Eq. (2.11), and from that the concept of the electromagnetic wave arises in a completely natural way (p. 45).

We began this discussion with the special case of waves that have a constant shape as they propagate, even though, as a rule, waves don't maintain a fixed profile. Still, that simple assumption has led us to the general formulation, the differential wave equation. If a function that represents a wave is a solution of that equation, it will at the same time be a function of $(x \mp vt)$—specifically, one that is twice differentiable (in a nontrivial way) with respect to both x and t.

2.2 Harmonic Waves

Let's now examine the simplest wave form, one for which the profile is a sine or cosine curve. These are variously known as sinusoidal waves, simple harmonic waves, or more succinctly as **harmonic waves**. We shall see in Chapter 7 that any wave shape can be synthesized by a superposition of harmonic waves, and they therefore take on a special significance.

Choose as the profile the simple function

$$\psi(x, t)|_{t=0} = \psi(x) = A \sin kx = f(x) \tag{2.12}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

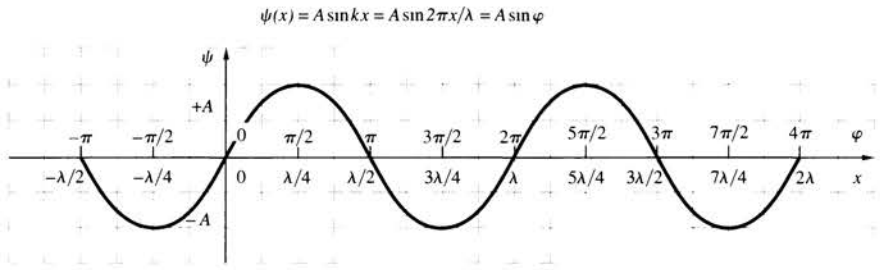


Figure 2.6 A harmonic function, which serves as the profile of a harmonic wave. One wavelength corresponds to a change in phase φ of 2π rad.

where k is a positive constant known as the **propagation number**. It's necessary to introduce the constant k simply because we cannot take the sine of a quantity that has physical units. The sine is the ratio of two lengths and is therefore unitless. Accordingly, kx is properly in radians, which is not a real physical unit. The sine varies from $+1$ to -1 so that the maximum value of $\psi(x)$ is A. This maximum disturbance is known as the **amplitude** of the wave (Fig. 2.6). To transform Eq. (2.12) into a *progressive wave* traveling at speed v in the positive x-direction, we need merely replace x by $(x - vt)$, in which case

$$\psi(x, t) = A \sin k(x - vt) = f(x - vt) \tag{2.13}$$

This is clearly a solution of the differential wave equation (see Problem 2.18). Holding either x or t fixed results in a sinusoidal disturbance; the wave is periodic in both space and time. The **spatial period** is known as the **wavelength** and is denoted by λ. Wavelength is *the number of units of length per wave*. The customary measure of λ is the nanometer, where 1 nm $= 10^{-9}$ m, although the micron (1 μm $= 10^{-6}$ m) is often used and the older angstrom (1 Å $= 10^{-10}$ m) can still be found in the literature. An increase or decrease in x by the amount λ should leave ψ unaltered, that is,

$$\psi(x, t) = \psi(x \pm \lambda, t) \tag{2.14}$$

In the case of a harmonic wave, this is equivalent to altering the argument of the sine function by $\pm 2\pi$. Therefore,

$$\sin k(x - vt) = \sin k[(x \pm \lambda) - vt] = \sin [k(x - vt) \pm 2\pi]$$

and so $$|k\lambda| = 2\pi$$

or, since both k and λ are positive numbers,

$$k = 2\pi/\lambda \tag{2.15}$$

Figure 2.6 shows how to plot the profile given by Eq. (2.12) in terms of λ. Here φ is the argument of the sine function, also called the **phase**. Notice that $\psi(x) = 0$ whenever $\sin \varphi = 0$, which happens when $\varphi = 0, \pi, 2\pi, 3\pi$, and so on. That occurs at $x = 0, \lambda/2, \lambda$, and $3\lambda/2$, respectively.

In an analogous fashion to the above discussion of λ, we now examine the **temporal period**, τ. This is the amount of time it takes for one complete wave to pass a stationary observer. In this case, it is the repetitive behavior of the wave in time that is of interest, so that

$$\psi(x, t) = \psi(x, t \pm \tau) \tag{2.16}$$

and $$\sin k(x - vt) = \sin k[x - v(t \pm \tau)]$$

$$\sin k(x - vt) = \sin [k(x - vt) \pm 2\pi]$$

Therefore,

$$|kv\tau| = 2\pi$$

But these are all positive quantities; hence

$$kv\tau = 2\pi \tag{2.17}$$

or $$\frac{2\pi}{\lambda} v\tau = 2\pi$$

from which it follows that

$$\tau = \lambda/v \tag{2.18}$$

The period is *the number of units of time per wave* (Fig. 2.7), the inverse of which is the **temporal frequency** ν, or *the number of waves per unit of time* (i.e., per second). Thus,

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

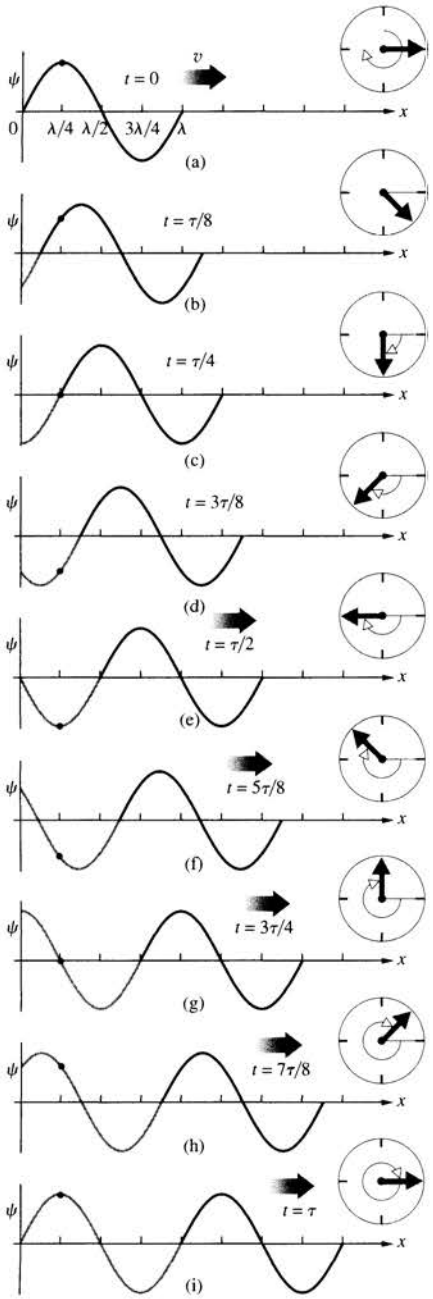


Figure 2.7 A harmonic wave moving along the *x*-axis during a time of one period. Note that if this is a picture of a rope any one point on it only moves vertically. We'll discuss the significance of the rotating arrow in Section 2.6.

$$\nu \equiv 1/\tau$$

in units of cycles per second or Hertz. Equation (2.18) then becomes

$$v = \nu\lambda \tag{2.19}$$

Imagine that you are at rest and a harmonic wave on a string is progressing past you. The number of waves that sweep by per second is ν, and the length of each is λ. In 1.0 s, the overall length of the disturbance that passes you is the product $\nu\lambda$. If, for example, each wave is 2.0 m long and they come at a rate of 5.0 per second, then in 1.0 s, 10 m of wave fly by. This is just what we mean by the speed of the wave (v)—the rate, in m/s, at which it advances. Said slightly differently, because a length of wave λ passes by in a time τ, its speed must equal $\lambda/\tau = \nu\lambda$. Incidentally, Newton derived this relationship in the *Principia* (1687) in a section called "To find the velocity of waves."

Two other quantities are often used in the literature of wave motion. One is the **angular temporal frequency**

$$\omega \equiv 2\pi/\tau \ = 2\pi\nu \tag{2.20}$$

given in units of radians per second. The other, which is important in spectroscopy, is the **wave number** or **spatial frequency**

$$\kappa \equiv 1/\lambda \tag{2.21}$$

measured in inverse meters. In other words, κ is *the number of waves per unit of length* (i.e., per meter). All of these quantities apply equally well to waves that are not harmonic, as long as each such wave is made up of a single regularly repeated **profile-element** (Fig. 2.8).

Using the above definitions a number of equivalent expressions can be written for the traveling harmonic wave:

$$\psi = A \sin k(x \mp vt) \tag{2.13}$$

$$\psi = A \sin 2\pi\left(\frac{x}{\lambda} \mp \frac{t}{\tau}\right) \tag{2.22}$$

$$\psi = A \sin 2\pi(\kappa x \mp \nu t) \tag{2.23}$$

$$\psi = A \sin (kx \mp \omega t) \tag{2.24}$$

$$\psi = A \sin 2\pi\nu\left(\frac{x}{v} \mp t\right) \tag{2.25}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

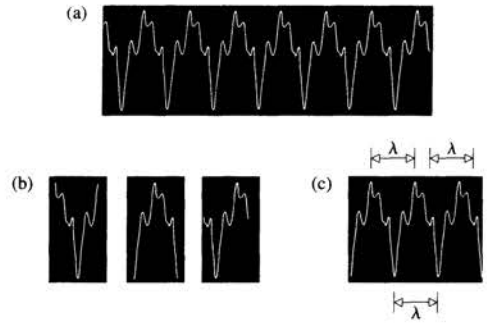


Figure 2.8 (*a*) The waveform produced by a saxophone. Imagine any number of profile-elements (*b*) that, when repeated, create the waveform (*c*). The distance over which the wave repeats itself is called the wavelength, λ.

Of these, Eqs. (2.13) and (2.24) will be encountered most frequently. Note that all these idealized waves are of infinite extent. That is, for any fixed value of t, there is no mathematical limitation on x, which varies from $-\infty$ to $+\infty$. Each such wave has a single constant frequency and is therefore **monochromatic** or, even better, **monoenergetic**. Real waves are never monochromatic. Even a perfect sinusoidal generator cannot have been operating forever. Its output will unavoidably contain a range of frequencies, albeit a small one, just because the wave does not extend back to $t = -\infty$. Thus all waves comprise a band of frequencies, and when that band is narrow the wave is said to be **quasimonochromatic**.

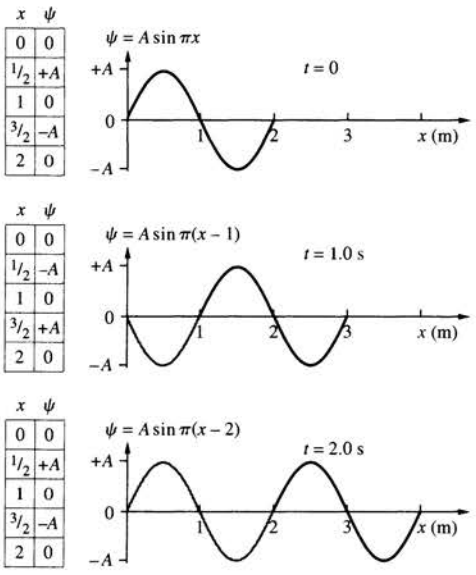


Figure 2.9 A progressive wave of the form $\psi(x, t) = A \sin k(x - vt)$, moving to the right at a speed of 1.0 m/s.

Before we move on, let's put some numbers into Eq. (2.13) and see how to deal with each term. To that end, arbitrarily let $v = 1.0$ m/s and $\lambda = 2.0$ m. Then the wavefunction

$$\psi = A \sin \frac{2\pi}{\lambda}(x - vt)$$

in SI units becomes

$$\psi = A \sin \pi (x - t)$$

Figure 2.9 shows how the wave progresses to the right at 1.0 m/s as the time goes from $t = 0$ [whereupon $\psi = A \sin \pi x$] to $t = 1.0$ s [whereupon $\psi = A \sin \pi(x - 1.0)$] to $t = 2.0$ s [whereupon $\psi = A \sin \pi(x - 2.0)$].

2.3 Phase and Phase Velocity

Examine any one of the harmonic wavefunctions, such as

$$\psi(x, t) = A \sin (kx - \omega t) \tag{2.26}$$

The entire argument of the sine is the phase φ of the wave, where

$$\varphi = (kx - \omega t) \tag{2.27}$$

At $t = x = 0$,

$$\psi(x, t)\big|_{\substack{x=0 \\ t=0}} = \psi(0, 0) = 0$$

which is certainly a special case. More generally, we can write

$$\psi(x, t) = A \sin (kx - \omega t + \varepsilon) \tag{2.28}$$

where ε is the **initial phase**. To get a sense of the physical meaning of ε, imagine that we wish to produce a progressive harmonic wave on a stretched string, as in Fig. 2.10. In order to generate harmonic waves, the hand holding the string would have to move such that its vertical displacement y was proportional to the negative of its acceleration, that is, in simple harmonic motion (see Problem 2.21). But at $t = 0$ and $x = 0$, the hand certainly need not be on the x-axis about to move downward, as in Fig. 2.10. It could, of course, begin its motion on an upward swing, in which case $\varepsilon = \pi$, as in Fig. 2.11. In this

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

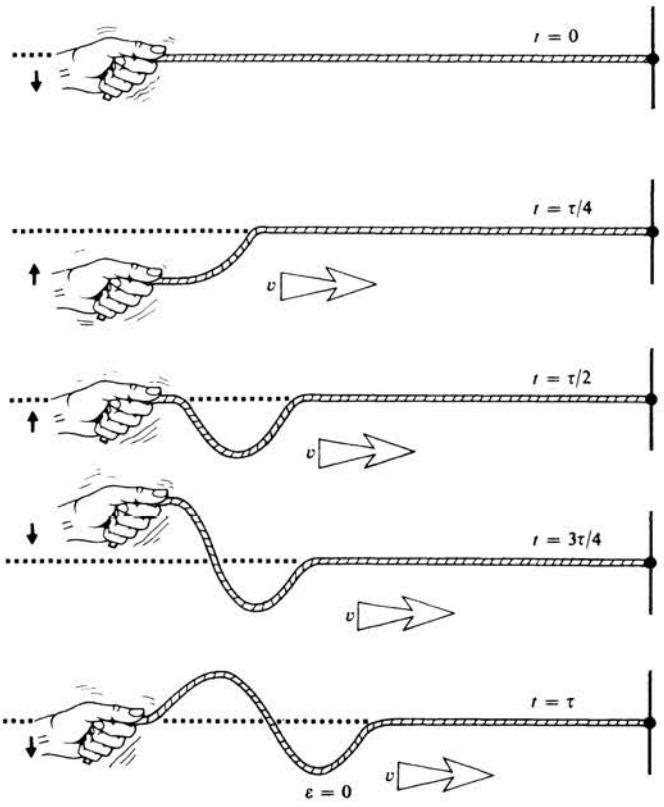


Figure 2.10 With $\varepsilon = 0$ note that at $x = 0$ and $t = \tau/4 = \pi/2\omega$, $y = A \sin(-\pi/2) = -A$.

latter case,

$$\psi(x, t) = y(x, t) = A \sin(kx - \omega t + \pi)$$

which is equivalent to

$$\psi(x, t) = A \sin(\omega t - kx) \tag{2.29}$$

or $$\psi(x, t) = A \cos\left(\omega t - kx - \frac{\pi}{2}\right)$$

The initial phase angle is just the constant contribution to the phase arising at the generator and is independent of how far in space, or how long in time, the wave has traveled.

The phase in Eq. (2.26) is $(kx - \omega t)$, whereas in Eq. (2.29) it's $(\omega t - kx)$. Nonetheless, both of these equations describe waves moving in the positive x-direction that are otherwise identical except for a relative phase difference of π. As is often the case, when the initial phase is of no particular significance in a given situation, either Eq. (2.26) or (2.29) or, if you like, a cosine function can be used to represent the wave. Even so, in some situations one expression for the phase may be mathematically more appealing than another; the literature abounds with both, and so we will use both.

The phase of a disturbance such as $\psi(x, t)$ given by Eq. (2.28) is

$$\varphi(x, t) = (kx - \omega t + \varepsilon)$$

and is obviously a function of x and t. In fact, the partial derivative of φ with respect to t, holding x constant, is the *rate-of-change of phase with time*, or

$$\left|\left(\frac{\partial \varphi}{\partial t}\right)_x\right| = \omega \tag{2.30}$$

The rate-of-change of phase at any fixed location is the angular frequency of the wave, the rate at which a point on the rope

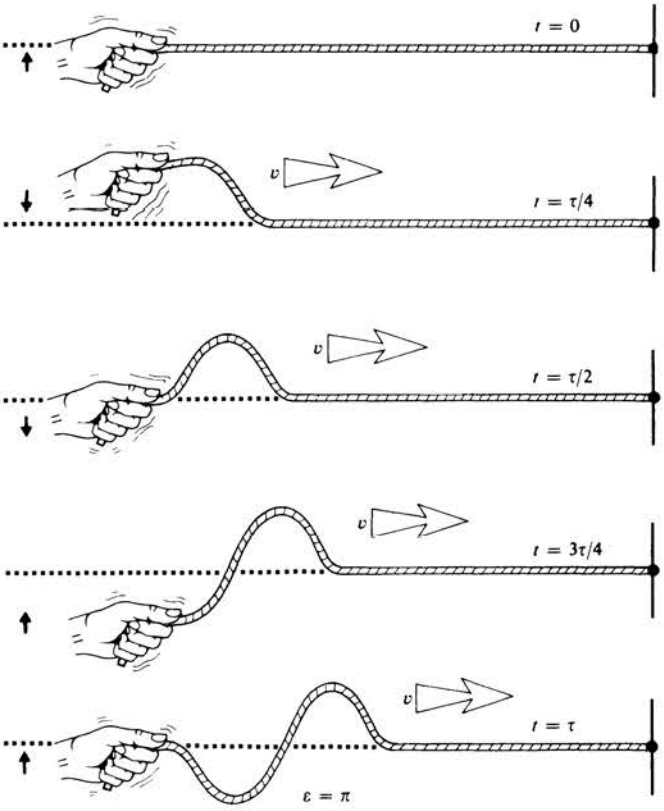


Figure 2.11 With $\varepsilon = \pi$ note that at $x = 0$ and $t = \tau/4$, $y = A \sin(\pi/2) = A$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

in Fig. 2.10 oscillates up and down. That point must go through the same number of cycles per second as the wave. For each cycle, φ changes by 2π.

Similarly, the *rate-of-change of phase with distance*, holding t constant, is

$$\left|\left(\frac{\partial \varphi}{\partial x}\right)_t\right| = k \tag{2.31}$$

These two expressions should bring to mind an equation from the theory of partial derivatives, one used frequently in Thermodynamics, namely,

$$\left(\frac{\partial x}{\partial t}\right)_\varphi = \frac{-(\partial \varphi/\partial t)_x}{(\partial \varphi/\partial x)_t} \tag{2.32}$$

The term on the left represents the *speed of propagation of the condition of constant phase*. Imagine a harmonic wave and choose any point on the profile, for example, a crest of the wave. As the wave moves through space, the displacement y of the crest remains fixed. Since the only variable in the harmonic wavefunction is the phase, it too must be constant for that moving point. That is, the phase is fixed at such a value as to yield the constant y corresponding to the chosen point. The point moves along with the profile at the speed v, and so too does the condition of constant phase.

Taking the appropriate partial derivatives of φ as given, for example, by Eq. (2.29) and substituting them into Eq. (2.32), we get

$$\left(\frac{\partial x}{\partial t}\right)_\varphi = \pm \frac{\omega}{k} = \pm v \tag{2.33}$$

This is the *speed* at which the profile moves and is known commonly as the **phase velocity** of the wave. The phase velocity is accompanied by a positive sign when the wave moves in the direction of increasing x and a negative one in the direction of decreasing x. This is consistent with our development of v as the magnitude of the wave velocity: $v > 0$.

Consider the idea of the propagation of constant phase and how it relates to any one of the harmonic wave equations, say

$$\psi = A \sin k(x \mp vt)$$

with

$$\varphi = k(x - vt) = \text{constant}$$

As t increases, x must increase. Even if $x < 0$ so that $\varphi < 0$, x must increase (i.e., become less negative). Here, then, the condition of constant phase moves in the direction of increasing x. As long as the two terms in the phase subtract from each other, the wave travels in the positive x-direction. On the other hand, for

$$\varphi = k(x + vt) = \text{constant}$$

as t increases x can be positive and decreasing or negative and becoming more negative. In either case, the constant-phase condition moves in the decreasing x-direction.

Any point on a harmonic wave having a fixed magnitude moves such that $\varphi(x, t)$ is constant in time, in other words, $d\varphi(x, t)/dt = 0$, or alternatively, $d\psi(x, t)/dt = 0$. This is true for all waves, periodic or not, and it leads (Problem 2.27) to the expression

$$\pm v = \frac{-(\partial \psi/\partial t)_x}{(\partial \psi/\partial x)_t} \tag{2.34}$$

which can be used to conveniently provide v when we have $\psi(x, t)$. Note that because v is always a positive number, when the ratio on the right turns out negative the motion is in the negative x-direction.

Figure 2.12 depicts a source producing hypothetical two-dimensional waves on the surface of a liquid. The essentially sinusoidal nature of the disturbance, as the medium rises and falls, is evident in the diagram. But there is another useful way to envision what's happening. *The curves connecting all the*

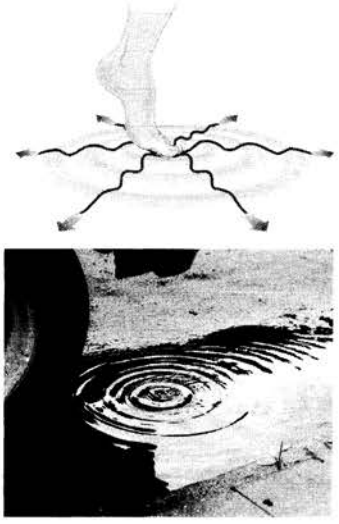


Figure 2.12 Circular waves. (Photo by E.H.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

points with a given phase form a set of concentric circles. Furthermore, given that A is everywhere constant at any one distance from the source, if φ is constant over a circle, ψ too must be constant over that circle. In other words, all the corresponding peaks and troughs fall on circles, and we speak of these as *circular waves*, each of which expands outward at the speed v.

2.4 The Superposition Principle

The form of the differential wave equation [Eq. (2.11)] reveals an intriguing property of waves, one that is quite unlike the behavior of a stream of classical particles. Suppose that the wavefunctions ψ_1 and ψ_2 are each separate solutions of the wave equation; it follows that $(\psi_1 + \psi_2)$ is also a solution. This is known as the **Superposition Principle**, and it can easily be proven since it must be true that

$$\frac{\partial^2 \psi_1}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 \psi_1}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 \psi_2}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 \psi_2}{\partial t^2}$$

Adding these yields

$$\frac{\partial^2 \psi_1}{\partial x^2} + \frac{\partial^2 \psi_2}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 \psi_1}{\partial t^2} + \frac{1}{v^2}\frac{\partial^2 \psi_2}{\partial t^2}$$

and so

$$\frac{\partial^2}{\partial x^2}(\psi_1 + \psi_2) = \frac{1}{v^2}\frac{\partial^2}{\partial t^2}(\psi_1 + \psi_2)$$

which establishes that $(\psi_1 + \psi_2)$ is indeed a solution. What this means is that when two separate waves arrive at the same place in space wherein they overlap, they will simply add to (or subtract from) one another without permanently destroying or disrupting either wave. ***The resulting disturbance at each point in the region of overlap is the algebraic sum of the individual constituent waves at that location*** (Fig. 2.13). Once having passed through the region where the two waves coexist, each will move out and away unaffected by the encounter.

Keep in mind that we are talking about a *linear* superposition of waves, a process that's widely valid and the most commonly encountered. Nonetheless, it is also possible for the wave amplitudes to be large enough to drive the medium in a nonlinear fashion (p. 639). For the time being we'll concentrate on the linear differential wave equation, which results in a linear Superposition Principle.

Much of Optics involves the superposition of waves in one way or another. Even the basic processes of reflection and refraction are manifestations of the scattering of light from

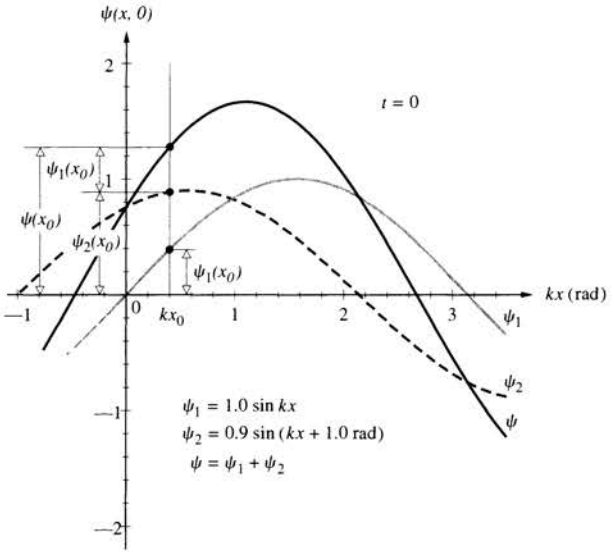


Figure 2.13 The superposition of two equal-wavelength sinusoids ψ_1 and ψ_2, having amplitudes A_1 and A_2, respectively. The resultant, ψ, is a sinusoid with the same wavelength, which at every point equals the algebraic sum of the constituent sinusoids. Thus at $x = x_0$, $\psi(x_0) = \psi_1(x_0) + \psi_2(x_0)$; the magnitudes add. The amplitude of ψ is A and it can be determined in several ways; see Fig. 2.17.

countless atoms (p. 86), a phenomenon that can only be treated satisfactorily in terms of the overlapping of waves. It therefore becomes crucial that we understand the process, at least qualitatively, as soon as possible. Consequently, carefully examine the two coexisting waves in Fig. 2.13. At every point (i.e., every value of kx) we simply add ψ_1 and ψ_2, either of which could be positive or negative. As a quick check, keep in mind that wherever either constituent wave is zero (e.g., $\psi_1 = 0$), the resultant disturbance equals the value of the other nonzero constituent wave ($\psi = \psi_2$), and those two curves cross at that location (e.g., at $kx = 0$ and $+3.14$ rad). On the other hand, $\psi = 0$ wherever the two constituent waves have equal magnitudes and opposite signs (e.g., at $kx = +2.67$ rad). Incidentally, notice how a relative *positive* phase difference of 1.0 rad between the two curves shifts ψ_2 to the *left* with respect to ψ_1 by 1.0 rad.

Developing the illustration a bit further, Fig. 2.14 shows how the resultant arising from the superposition of two nearly equal-amplitude waves depends on the *phase-angle difference* between them. In Fig. 2.14*a* the two constituent waves have the same phase; that is, their phase-angle difference is zero, and they are said to be **in-phase**; they rise and fall in-step, reinforcing each other. The composite wave, which then has a

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

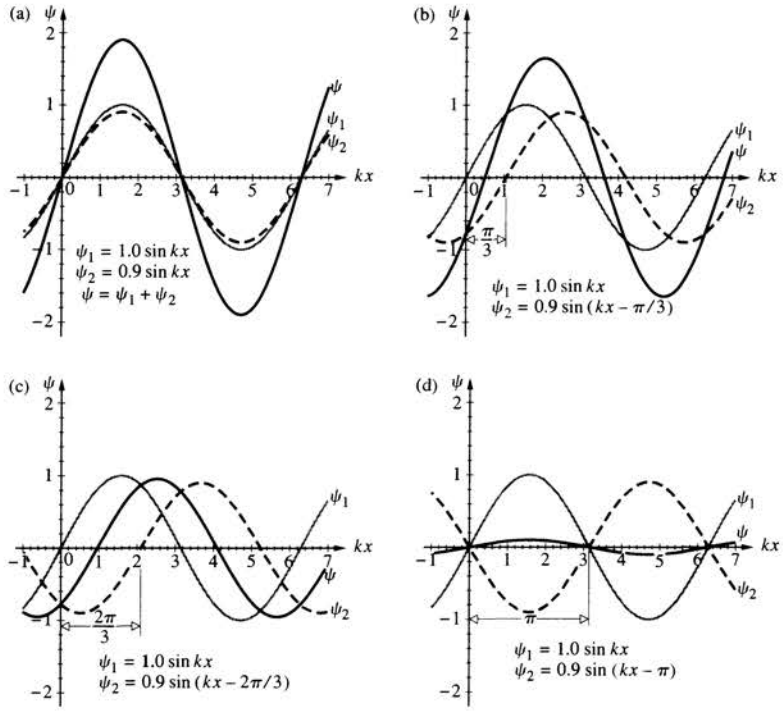


Figure 2.14 The superposition of two sinusoids with amplitudes of $A_1 = 1.0$ and $A_2 = 0.9$. In (*a*) they are in-phase. In (*b*) ψ_1 leads ψ_2 by $\pi/3$. In (*c*) ψ_1 leads ψ_2 by $2\pi/3$. And (*d*) ψ_1 and ψ_2 are out-of-phase by π and almost cancel each other. To see how the amplitudes can be determined, go to Fig. 2.18.

substantial amplitude, is sinusoidal with the same frequency and wavelength as the component waves (p. 285). Following the sequence of the drawings, we see that the resultant amplitude diminishes as the phase-angle difference increases until, in Fig. 2.14*d*, it almost vanishes when that difference equals π. The waves are then said to be 180° **out-of-phase**. The fact that waves which are out-of-phase tend to diminish each other has given the name **interference** to the whole phenomenon.

2.5 The Complex Representation

As we develop the analysis of wave phenomena, it will become evident that the sine and cosine functions that describe harmonic waves can be somewhat awkward for our purposes. The expressions formulated will sometimes be rather involved and the trigonometric manipulations required to cope with them will be even more unattractive. The complex-number representation offers an alternative description that is mathematically simpler to process. In fact, complex exponentials are used extensively in both Classical and Quantum Mechanics, as well as in Optics.

The complex number $\tilde{z}$ has the form

$$\tilde{z} = x + iy \tag{2.35}$$

where $i = \sqrt{-1}$. The real and imaginary parts of $\tilde{z}$ are, respectively, x and y, where both x and y are themselves real numbers. This is illustrated graphically in the Argand diagram in Fig. 2.15*a*. In terms of polar coordinates (r, θ),

$$x = r \cos \theta \qquad y = r \sin \theta$$

and $$\tilde{z} = x + iy = r(\cos \theta + i \sin \theta)$$

The *Euler formula**

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

leads to the expression $e^{-i\theta} = \cos \theta - i \sin \theta$, and adding and subtracting these two equations yields

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

*If you have any doubts about this identity, take the differential of $\tilde{z} = \cos \theta + i \sin \theta$, where $r = 1$. This yields $d\tilde{z} = i\tilde{z}\, d\theta$, and integration gives $\tilde{z} = \exp(i\theta)$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

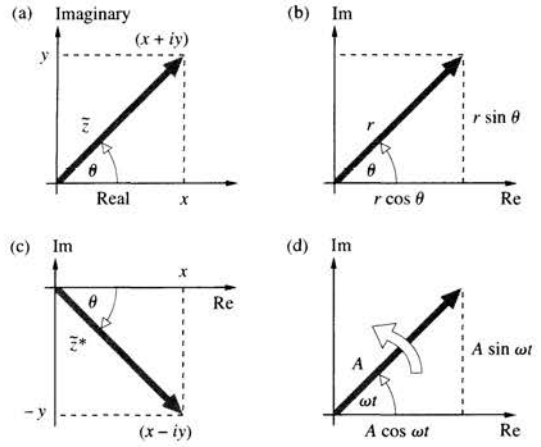


Figure 2.15 An Argand diagram is a representation of a complex number in terms of its real and imaginary components. This can be done using either (*a*) x and y or (*b*) r and θ. Moreover, when θ is a constantly changing function of time (*d*), the arrow rotates at a rate ω.

and $$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Moreover, the Euler formula allows us (Fig. 2.15*b*) to write

$$\tilde{z} = re^{i\theta} = r\cos\theta + ir\sin\theta$$

where r is the *magnitude* of $\tilde{z}$ and θ is the *phase angle* of $\tilde{z}$, in radians. The magnitude is often denoted by $|\tilde{z}|$ and referred to as the *modulus* or *absolute value* of the complex number. The *complex conjugate*, indicated by an asterisk (Fig. 2.15*c*), is found by replacing i wherever it appears, with $-i$, so that

$$\tilde{z}^* = (x + iy)^* = (x - iy)$$

$$\tilde{z}^* = r(\cos\theta - i\sin\theta)$$

and $$\tilde{z}^* = re^{-i\theta}$$

The operations of addition and subtraction are quite straightforward:

$$\tilde{z}_1 \pm \tilde{z}_2 = (x_1 + iy_1) \pm (x_2 + iy_2)$$

and therefore

$$\tilde{z}_1 \pm \tilde{z}_2 = (x_1 \pm x_2) + i(y_1 \pm y_2)$$

Notice that this process is very much like the component addition of vectors.

Multiplication and division are most simply expressed in polar form

$$\tilde{z}_1\tilde{z}_2 = r_1r_2\,e^{i(\theta_1+\theta_2)}$$

and $$\frac{\tilde{z}_1}{\tilde{z}_2} = \frac{r_1}{r_2}\,e^{i(\theta_1-\theta_2)}$$

A number of facts that will be useful in future calculations are well worth mentioning at this point. It follows from the ordinary trigonometric addition formulas (Problem 2.37) that

$$e^{\tilde{z}_1+\tilde{z}_2} = e^{\tilde{z}_1}e^{\tilde{z}_2}$$

and so, if $\tilde{z}_1 = x$ and $\tilde{z}_2 = iy$,

$$e^{\tilde{z}} = e^{x+iy} = e^x e^{iy}$$

The modulus of a complex quantity is given by

$$r = |\tilde{z}| \equiv (\tilde{z}\tilde{z}^*)^{1/2}$$

and $$|e^{\tilde{z}}| = e^x$$

Inasmuch as $\cos 2\pi = 1$ and $\sin 2\pi = 0$,

$$e^{i2\pi} = 1$$

Similarly,

$$e^{i\pi} = e^{-i\pi} = -1 \quad \text{and} \quad e^{\pm i\pi/2} = \pm i$$

The function $e^{\tilde{z}}$ is periodic; that is, it repeats itself every $i2\pi$:

$$e^{\tilde{z}+i2\pi} = e^{\tilde{z}}e^{i2\pi} = e^{\tilde{z}}$$

Any complex number can be represented as the sum of a real part Re $(\tilde{z})$ and an imaginary part Im (z)

$$\tilde{z} = \text{Re}\,(\tilde{z}) + i\,\text{Im}\,(\tilde{z})$$

such that

$$\text{Re}\,(\tilde{z}) = \tfrac{1}{2}(\tilde{z} + \tilde{z}^*) \quad \text{and} \quad \text{Im}\,(\tilde{z}) = \frac{1}{2i}(\tilde{z} - \tilde{z}^*)$$

Both of these expressions follow immediately from the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Argand diagram, Fig. 2.15*a* and *c*. For example, $\tilde{z} + \tilde{z}^* = 2x$ because the imaginary parts cancel, and so Re $(\tilde{z}) = x$.

From the polar form where

$$\text{Re}\,(\tilde{z}) = r\cos\theta \quad \text{and} \quad \text{Im}\,(\tilde{z}) = r\sin\theta$$

it is clear that either part could be chosen to describe a harmonic wave. It is customary, however, to choose the real part, in which case a harmonic wave is written as

$$\psi(x, t) = \text{Re}\,[Ae^{i(\omega t - kx + \varepsilon)}] \tag{2.36}$$

which is, of course, equivalent to

$$\psi(x, t) = A\cos(\omega t - kx + \varepsilon)$$

Henceforth, wherever it's convenient, we shall write the wavefunction as

$$\psi(x, t) = Ae^{i(\omega t - kx + \varepsilon)} = Ae^{i\varphi} \tag{2.37}$$

and utilize this complex form in the required computations. This is done to take advantage of the ease with which complex exponentials can be manipulated. Only after arriving at a final result, and then only if we want to represent the actual wave, must we take the real part. It has, accordingly, become quite common to write $\psi(x, t)$, as in Eq. (2.37), where it is understood that the actual wave is the real part.

2.6 Phasors and the Addition of Waves

The arrow in the Argand diagram (Fig. 2.15*d*) is set rotating at a frequency ω by letting the angle equal ωt. This suggests a scheme for representing (and ultimately adding) waves that we will introduce here qualitatively and develop later (p. 287) quantitatively. Figure 2.16 depicts a harmonic wave of amplitude A traveling to the left. The arrow in the diagram has a length A and revolves at a constant rate such that the changing angle it makes with the reference x-axis is ωt. This rotating arrow and its associated phase angle together constitute a **phasor**, which tells us everything we need to know about the corresponding harmonic wave. It's common to express a phasor in terms of its amplitude, A, and phase, φ, as $A\angle\varphi$.

To see how this works, let's first examine each part of Fig. 2.16 separately. The phasor in Fig. 2.16*a* has a zero phase angle; that is, it lies along the reference axis; the associated sine function can also serve as a reference. In Fig. 2.16*b* the phasor has a phase angle of $+\pi/3$ rad, and the sine curve is shifted to the left by $\pi/3$ rad. That sine curve reaches its first peak at a smaller value of kx than does the reference curve in part (*a*), and therefore it *leads* the reference by $\pi/3$ rad. In parts (*c*), (*d*), and (*e*) of Fig. 2.16, the phase angles are $+\pi/2$ rad, $+2\pi/3$ rad, and $+\pi$ rad, respectively. The entire sequence of curves can be seen as a wave, $\psi = A\sin(kx + \omega t)$, traveling to the left. It is equivalently represented by a phasor rotating counterclockwise such that its phase angle at

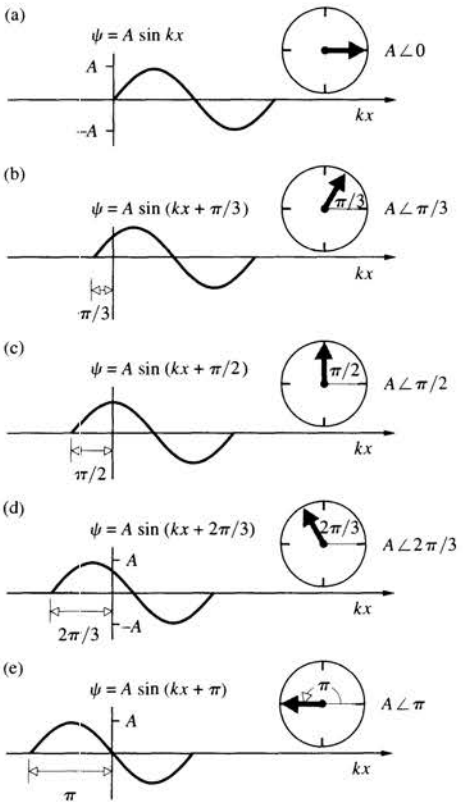


Figure 2.16 A plot of the function $\psi = A\sin(kx + \omega t)$ and the corresponding phasor diagrams. In (*a*), (*b*), (*c*), (*d*), and (*e*), the values of ωt are 0, $\pi/3$, $\pi/2$, $2\pi/3$, and π, respectively.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

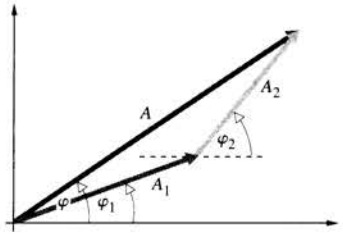


Figure 2.17 The sum of two phasors $A_1\angle\varphi_1$ and $A_2\angle\varphi_2$ equals $A\angle\varphi$. Go back and look at Fig. 2.13 which depicts the overlapping of two sinusoids having amplitudes of $A_1 = 1.0$, and $A_2 = 0.9$ and phases of $\varphi_1 = 0$ and $\varphi_2 = 1.0$ rad.

any moment is ωt. Much the same thing happens in Fig. 2.7, but there the wave advances to the right and the phasor rotates clockwise.

When wavefunctions are combined, we are usually interested in the resulting amplitude and phase. With that in mind, reexamine the way waves add together in Fig. 2.14. Apparently, for disturbances that are in-phase (Fig. 2.14*a*) the amplitude of the resultant wave, A, is the sum of the constituent amplitudes: $A = A_1 + A_2 = 1.0 + 0.9 = 1.9$. This is the same thing we would get if we added two colinear vectors pointing in the same direction. Similarly (Fig. 2.14*d*), when the component waves are 180° out-of-phase $A = A_1 - A_2 = 1.0 - 0.9 = 0.1$ as if two colinear oppositely directed vectors were added. Although phasors are not vectors, they do add in a similar way. Later, we'll prove that two arbitrary phasors, $A_1\angle\varphi_1$ and $A_2\angle\varphi_2$, combine tip-to-tail, as vectors would (Fig. 2.17), to produce a resultant $A\angle\varphi$. Because both phasors rotate together at a rate ω, we can simply freeze them at $t = 0$ and not worry about their time dependence, which makes them a lot easier to draw.

The four phasor diagrams in Fig. 2.18 correspond to the four wave combinations taking place sequentially in Fig. 2.14. When the waves are in-phase (as in Fig. 2.14*a*), we take the phases of both wave-1 and wave-2 to be zero (Fig. 2.18*a*) and position the corresponding phasors tip-to-tail along the zero-φ reference axis. When the waves differ in phase by $\pi/3$ (as in Fig. 2.14*b*), the phasors have a relative phase (Fig. 2.18*b*) of $\pi/3$. The resultant, which has an appropriately reduced amplitude, has a phase φ that is between 0 and $\pi/3$, as can be seen in both Figs. 2.14*b* and 2.18*b*. When the two waves differ in phase by $2\pi/3$ (as in Fig. 2.14*c*), the corresponding phasors almost form an equilateral triangle in Fig. 2.18*c* (but for the fact that $A_1 > A_2$), and so A now lies between A_1 and A_2. Finally, when the phase-angle difference for the two waves (and the two phasors) is π rad (i.e., 180°), they almost cancel and the resulting amplitude is a minimum. Notice (in Fig. 2.18*d*) that the resultant phasor points along the reference axis and so has the same phase (i.e., zero) as $A_1\angle\varphi_1$. Thus it is 180° out-of-phase with $A_2\angle\varphi_2$; the same thing is true of the corresponding waves in Fig. 2.14*d*.

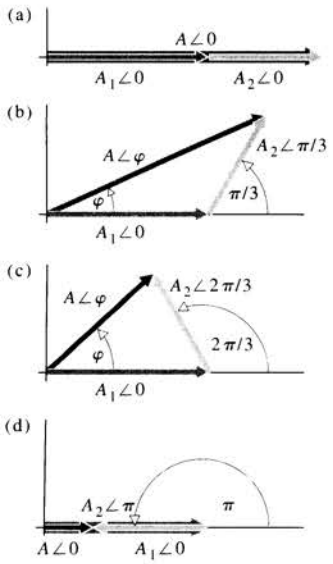


Figure 2.18 The addition of phasors representing two waves having amplitudes of $A_1 = 1.0$, and $A_2 = 0.9$ with four different relative phases, as shown in Fig. 2.14 .

2.7 Plane Waves

The plane wave is perhaps the simplest example of a three-dimensional wave. It exists at a given time, when all the surfaces on which a disturbance has constant phase form a set of planes, each generally perpendicular to the propagation direction. There are quite practical reasons for studying this sort of disturbance, one of which is that by using optical devices, we can readily produce light resembling plane waves.

The mathematical expression for a plane that is perpendicular to a given vector $\vec{\mathbf{k}}$ and that passes through some point (x_0, y_0, z_0) is rather easy to derive (Fig. 2.19). First we write the position vector in Cartesian coordinates in terms of the unit basis vectors (Fig. 2.19*a*),

$$\vec{\mathbf{r}} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$$

It begins at some arbitrary origin O and ends at the point (x, y, z), which can, for the moment, be anywhere in space. Similarly,

$$(\vec{\mathbf{r}} - \vec{\mathbf{r}}_0) = (x - x_0)\hat{\mathbf{i}} + (y - y_0)\hat{\mathbf{j}} + (z - z_0)\hat{\mathbf{k}}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

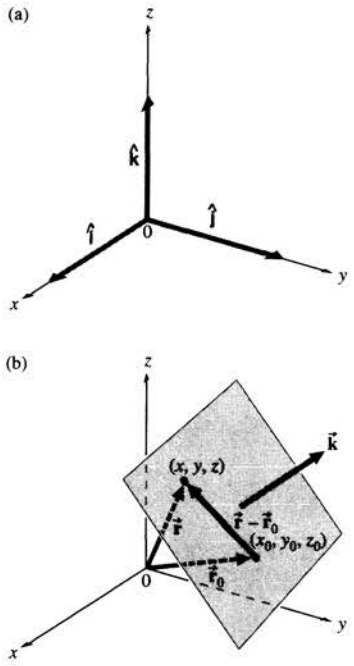


Figure 2.19 (*a*) The Cartesian unit basis vectors. (*b*) A plane wave moving in the $\vec{\mathbf{k}}$-direction.

By setting

$$(\vec{\mathbf{r}} - \vec{\mathbf{r}}_0)\cdot\vec{\mathbf{k}} = 0 \tag{2.38}$$

we force the vector $(\vec{\mathbf{r}} - \vec{\mathbf{r}}_0)$ to sweep out a plane perpendicular to $\vec{\mathbf{k}}$, as its endpoint (x, y, z) takes on all allowed values. With

$$\vec{\mathbf{k}} = k_x\hat{\mathbf{i}} + k_y\hat{\mathbf{j}} + k_z\hat{\mathbf{k}} \tag{2.39}$$

Equation (2.38) can be expressed in the form

$$k_x(x - x_0) + k_y(y - y_0) + k_z(z - z_0) = 0 \tag{2.40}$$

or as

$$k_x x + k_y y + k_z z = a \tag{2.41}$$

where

$$a = k_x x_0 + k_y y_0 + k_z z_0 = \text{constant} \tag{2.42}$$

The most concise form of the equation of a plane perpendicular to $\vec{\mathbf{k}}$ is then just

$$\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} = \text{constant} = a \tag{2.43}$$

The plane is the locus of all points whose position vectors each have the same projection onto the $\vec{\mathbf{k}}$-direction.

We can now construct a set of planes over which $\psi(\vec{r})$ varies in space sinusoidally, namely,

$$\psi(\vec{r}) = A\sin(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}) \tag{2.44}$$

$$\psi(\vec{r}) = A\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}) \tag{2.45}$$

or

$$\psi(\vec{r}) = Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} \tag{2.46}$$

For each of these expressions $\psi(\vec{r})$ is constant over every plane defined by $\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} =$ constant. Since we are dealing with harmonic functions, they should repeat themselves in space after a displacement of λ in the direction of $\vec{\mathbf{k}}$. Figure 2.20 is a rather humble representation of this kind of expression. We have drawn only a few of the infinite number of planes, each having a different $\psi(\vec{r})$. The planes should also have been drawn with an infinite spatial extent, since no limits were put on $\vec{\mathbf{r}}$. The disturbance clearly occupies all of space.

The spatially repetitive nature of these harmonic functions can be expressed by

$$\psi(\vec{r}) = \psi\left(\vec{r} + \frac{\lambda\vec{k}}{k}\right) \tag{2.47}$$

where k is the magnitude of $\vec{\mathbf{k}}$ and $\vec{\mathbf{k}}/k$ is a unit vector parallel to it (Fig. 2.21). In the exponential form, this is equivalent to

$$Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} = Ae^{i\vec{\mathbf{k}}\cdot(\vec{\mathbf{r}}+\lambda\vec{\mathbf{k}}/k)} = Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}e^{i\lambda k}$$

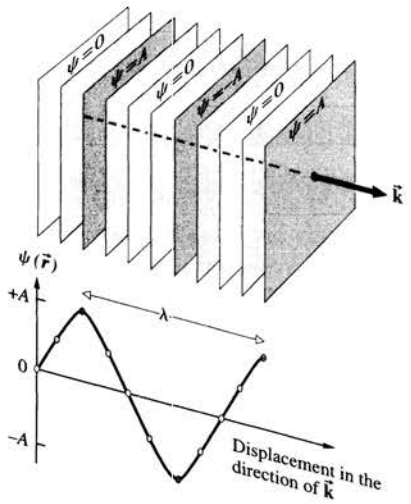


Figure 2.20 Wavefronts for a harmonic plane wave.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

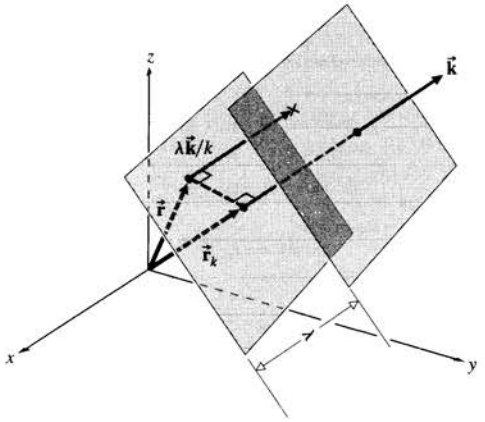


Figure 2.21 Plane waves.

For this to be true, we must have

$$e^{i\lambda k} = 1 = e^{i2\pi}$$

Therefore,

$$\lambda k = 2\pi$$

and

$$k = 2\pi/\lambda$$

The vector $\vec{\mathbf{k}}$, whose magnitude is the *propagation number* k (already introduced), is called the **propagation vector**.

At any fixed point in space where $\vec{\mathbf{r}}$ is constant, the phase is constant as is $\psi(\vec{r})$; in short, the planes are motionless. To get things moving, $\psi(\vec{r})$ must be made to vary in time, something we can accomplish by introducing the time dependence in an analogous fashion to that of the one-dimensional wave. Here then

$$\psi(\vec{r}, t) = Ae^{i(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} \mp \omega t)} \tag{2.48}$$

with A, ω, and k constant. As this disturbance travels along in the $\vec{\mathbf{k}}$-direction, we can assign a phase corresponding to it at each point in space and time. At any given time, ***the surfaces joining all points of equal phase are known as*** **wavefronts**. Note that the wavefunction will have a constant value over the wavefront only if the amplitude A has a fixed value at every point on the wavefront. In general, A is a function of $\vec{\mathbf{r}}$ and may not be constant over all space or even over a wavefront. In the latter case, the wave is said to be *inhomogeneous*. We will not be concerned with this sort of disturbance until later, when we consider laserbeams and total internal reflection.

The phase velocity of a plane wave given by Eq. (2.48) is equivalent to the propagation velocity of the wavefront. In Fig. 2.21, the scalar component of $\vec{\mathbf{r}}$ in the direction of $\vec{\mathbf{k}}$ is r_k. The disturbance on a wavefront is constant, so that after a time dt, if the front moves along $\vec{\mathbf{k}}$ a distance dr_k, we must have

$$\psi(\vec{r}, t) = \psi(r_k + dr_k, t + dt) = \psi(r_k, t) \tag{2.49}$$

In exponential form, this is

$$Ae^{i(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} \mp \omega t)} = Ae^{i(kr_k + k dr_k \mp \omega t \mp \omega\, dt)} = Ae^{i(kr_k \mp \omega t)}$$

and so it must be that $k\,dr_k = \pm\omega\,dt$

The magnitude of the wave velocity, dr_k/dt, is then

$$\frac{dr_k}{dt} = \pm\frac{\omega}{k} = \pm v \tag{2.50}$$

We could have anticipated this result by rotating the coordinate system in Fig. 2.21 so that $\vec{\mathbf{k}}$ was parallel to the x-axis. For that orientation

$$\psi(\vec{r}, t) = Ae^{i(kx \mp \omega t)}$$

since $\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} = kr_k = kx$. The wave has thereby been effectively reduced to the one-dimensional disturbance already discussed.

Now consider the two waves in Fig. 2.22; both have the same wavelength λ such that $k_1 = k_2 = k = 2\pi/\lambda$. Wave-1 propagating along the z-axis can be written as

$$\psi_1 = A_1 \cos\left(\frac{2\pi}{\lambda} z - \omega t\right)$$

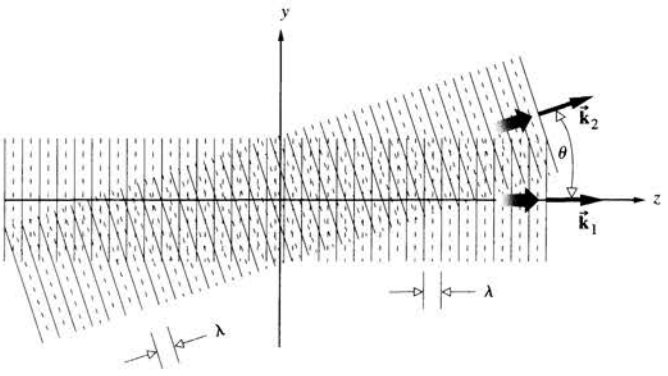


Figure 2.22 Two overlapping waves of the same wavelength traveling in different directions.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

where, because $\vec{\mathbf{k}}_1$ and $\vec{\mathbf{r}}$ are parallel, $\vec{\mathbf{k}}_1 \cdot \vec{\mathbf{r}} = kz = (2\pi/\lambda)z$. Similarly for wave-2, $\vec{\mathbf{k}}_2 \cdot \vec{\mathbf{r}} = k_z z + k_y y = (k \cos \theta)z + (k \sin \theta)y$ and

$$\psi_2 = A_2 \cos\left[\frac{2\pi}{\lambda}(z \cos \theta + y \sin \theta) - \omega t\right]$$

We'll return to these expressions and what happens in the region of overlap when we consider interference in more detail.

The plane harmonic wave is often written in Cartesian coordinates as

$$\psi(x, y, z, t) = Ae^{i(k_x x + k_y y + k_z z \mp \omega t)} \tag{2.51}$$

or

$$\psi(x, y, z, t) = Ae^{i[k(\alpha x + \beta y + \gamma z) \mp \omega t]} \tag{2.52}$$

where α, β, and γ are the direction cosines of $\vec{\mathbf{k}}$ (see Problem 2.39). In terms of its components, the magnitude of the propagation vector is

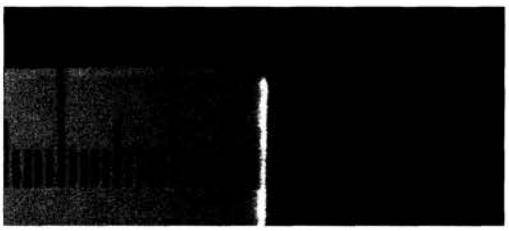
$$|\vec{\mathbf{k}}| = k = (k_x^2 + k_y^2 + k_z^2)^{1/2} \tag{2.53}$$

and of course

$$\alpha^2 + \beta^2 + \gamma^2 = 1 \tag{2.54}$$

We have examined plane waves with a particular emphasis on harmonic functions. The special significance of these waves is twofold: first, physically, sinusoidal waves can be generated relatively simply by using some form of harmonic oscillator; second, ***any three-dimensional wave can be expressed as a combination of plane waves***, each having a distinct amplitude and propagation direction.

We can certainly imagine a series of plane waves like those in Fig. 2.20 where the disturbance varies in some fashion other than harmonically (see photo). It will be seen in the next section that harmonic plane waves are, indeed, a special case of a more general plane-wave solution.



The image of a single collimated laser pulse caught as it swept along the surface of a ruler. This ultrashort burst of light corresponded to a portion of a plane wave. It extended in time for 300×10^{-15} s and was only a fraction of a millimeter long. (Photo courtesy J. Valdmanis and N.H. Abramson.)

2.8 The Three-Dimensional Differential Wave Equation

Of all the three-dimensional waves, only the plane wave (harmonic or not) can move through space with an unchanging profile. Clearly, the idea of a wave as a disturbance whose profile is *unaltered* is somewhat lacking. Alternatively, we can define a wave as any solution of the differential wave equation. What we need now is a three-dimensional wave equation. This should be rather easy to obtain, since we can guess at its form by generalizing from the one-dimensional expression, Eq. (2.11). In Cartesian coordinates, the position variables x, y, and z must certainly appear symmetrically* in the three-dimensional equation, a fact to be kept in mind. The wavefunction $\psi(x, y, z, t)$ given by Eq. (2.52) is a particular solution of the differential equation we are looking for. In analogy with the derivation of Eq. (2.11), we compute the following partial derivatives from Eq. (2.52):

$$\frac{\partial^2 \psi}{\partial x^2} = -\alpha^2 k^2 \psi \tag{2.55}$$

$$\frac{\partial^2 \psi}{\partial y^2} = -\beta^2 k^2 \psi \tag{2.56}$$

$$\frac{\partial^2 \psi}{\partial z^2} = -\gamma^2 k^2 \psi \tag{2.57}$$

and

$$\frac{\partial^2 \psi}{\partial t^2} = -\omega^2 \psi \tag{2.58}$$

Adding the three spatial derivatives and utilizing the fact that $\alpha^2 + \beta^2 + \gamma^2 = 1$, we obtain

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = -k^2 \psi \tag{2.59}$$

*There is no distinguishing characteristic for any one of the axes in Cartesian coordinates. We should therefore be able to change the names of, say, x to z, y to x, and z to y (keeping the system right-handed) without altering the differential wave equation.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Combining this with the time derivative, Eq. (2.58), and remembering that $v = \omega / k$, we arrive at

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \tag{2.60}$$

the *three-dimensional differential wave equation*. Note that x, y, and z do appear symmetrically, and the form is precisely what one might expect from the generalization of Eq. (2.11).

Equation (2.60) is usually written in a more concise form by introducing the *Laplacian* operator

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.61}$$

whereupon it becomes simply

$$\nabla^2\psi = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \tag{2.62}$$

Now that we have this most important equation, let's briefly return to the plane wave and see how it fits into the scheme of things. A function of the form

$$\psi(x, y, z, t) = Ae^{ik(\alpha x + \beta y + \gamma z \mp vt)} \tag{2.63}$$

is equivalent to Eq. (2.52) and, as such, is a solution of Eq. (2.62). It can also be shown (Problem 2.40) that

$$\psi(x, y, z, t) = f(\alpha x + \beta y + \gamma z - vt) \tag{2.64}$$

and

$$\psi(x, y, z, t) = g(\alpha x + \beta y + \gamma z + vt) \tag{2.65}$$

are both plane-wave solutions of the differential wave equation. The functions f and g, which are twice differentiable, are otherwise arbitrary and certainly need not be harmonic. A linear combination of these solutions is also a solution, and we can write this in a slightly different manner as

$$\psi(\vec{r}, t) = C_1 f(\vec{r}\cdot\vec{k}/k - vt) + C_2 g(\vec{r}\cdot\vec{k}/k + vt) \tag{2.66}$$

where C_1 and C_2 are constants.

Cartesian coordinates are particularly suitable for describing plane waves. However, as various physical situations arise, we can often take better advantage of existing symmetries by making use of some other coordinate representations.

2.9 Spherical Waves

Toss a stone into a tank of water. The surface ripples that emanate from the point of impact spread out in two-dimensional circular waves. Extending this imagery to three dimensions, envision a small pulsating sphere surrounded by a fluid. As the source expands and contracts, it generates pressure variations that propagate outward as spherical waves.

Consider now an idealized point source of light. The radiation emanating from it streams out radially, uniformly in all directions. The source is said to be *isotropic*, and the resulting wavefronts are again concentric spheres that increase in diameter as they expand out into the surrounding space. The obvious symmetry of the wavefronts suggests that it might be more convenient to describe them in terms of spherical polar coordinates (Fig. 2.23). In this representation the Laplacian operator is

$$\nabla^2 \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{2.67}$$

where r, θ, ϕ are defined by

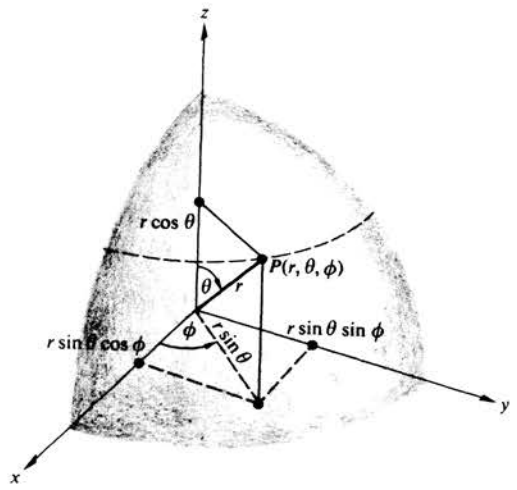


FIGURE 2.23 The geometry of spherical coordinates.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

$$x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r \cos\theta$$

Remember that we are looking for a description of spherical waves, waves that are spherically symmetrical (i.e., ones that do not depend on θ and ϕ) so that

$$\psi(\vec{r}) = \psi(r, \theta, \phi) = \psi(r)$$

The Laplacian of $\psi(r)$ is then simply

$$\nabla^2\psi(r) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) \tag{2.68}$$

We can obtain this result without being familiar with Eq. (2.67). Start with the Cartesian form of the Laplacian, Eq. (2.61), operate on the spherically symmetrical wavefunction $\psi(r)$, and convert each term to polar coordinates. Examining only the x-dependence, we have

$$\frac{\partial\psi}{\partial x} = \frac{\partial\psi}{\partial r}\frac{\partial r}{\partial x}$$

and

$$\frac{\partial^2\psi}{\partial x^2} = \frac{\partial^2\psi}{\partial r^2}\left(\frac{\partial r}{\partial x}\right)^2 + \frac{\partial\psi}{\partial r}\frac{\partial^2 r}{\partial x^2}$$

since

$$\psi(\vec{r}) = \psi(r)$$

Using

$$x^2 + y^2 + z^2 = r^2$$

we have

$$\frac{\partial r}{\partial x} = \frac{x}{r}$$

$$\frac{\partial^2 r}{\partial x^2} = \frac{1}{r}\frac{\partial}{\partial x}x + x\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = \frac{1}{r}\left(1 - \frac{x^2}{r^2}\right)$$

and so

$$\frac{\partial^2\psi}{\partial x^2} = \frac{x^2}{r^2}\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\left(1 - \frac{x^2}{r^2}\right)\frac{\partial\psi}{\partial r}$$

Now having $\partial^2\psi/\partial x^2$, we form $\partial^2\psi/\partial y^2$ and $\partial^2\psi/\partial z^2$, and on adding get

$$\nabla^2\psi(r) = \frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r}$$

which is equivalent to Eq. (2.68). This result can be expressed in a slightly different form:

$$\nabla^2\psi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) \tag{2.69}$$

The differential wave equation can then be written as

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \tag{2.70}$$

Multiplying both sides by r, yields

$$\frac{\partial^2}{\partial r^2}(r\psi) = \frac{1}{v^2}\frac{\partial^2}{\partial t^2}(r\psi) \tag{2.71}$$

Notice that this expression is now just the one-dimensional differential wave equation, Eq. (2.11), where the space variable is r and the wavefunction is the product $(r\psi)$. The solution of Eq. (2.71) is then simply

$$r\psi(r, t) = f(r - vt)$$

or

$$\psi(r, t) = \frac{f(r - vt)}{r} \tag{2.72}$$

This represents a spherical wave progressing radially outward from the origin, at a constant speed v, and having an arbitrary functional form f. Another solution is given by

$$\psi(r, t) = \frac{g(r + vt)}{r}$$

and in this case the wave is converging toward the origin.* The fact that this expression blows up at $r = 0$ is of little practical concern.

A special case of the general solution

$$\psi(r, t) = C_1\frac{f(r - vt)}{r} + C_2\frac{g(r + vt)}{r} \tag{2.73}$$

is the *harmonic spherical wave*

$$\psi(r, t) = \left(\frac{\mathcal{A}}{r}\right)\cos k\,(r \mp vt) \tag{2.74}$$

*Other more complicated solutions exist when the wave is not spherically symmetrical. See C. A. Coulson, *Waves*, Chapter 1.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

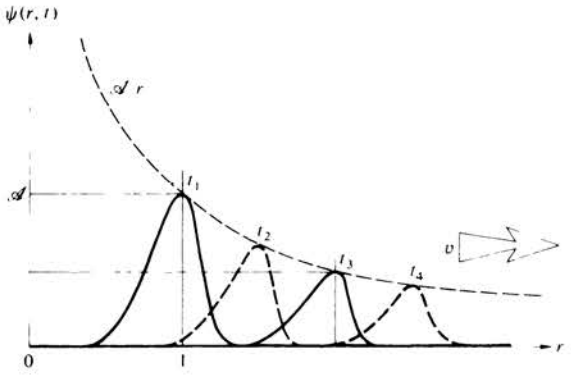


Figure 2.24 A "quadruple exposure" of a spherical pulse.

or
$$\psi(r, t) = \left(\frac{\mathscr{A}}{r}\right) e^{i k(r \mp vt)} \tag{2.75}$$

wherein the constant $\mathscr{A}$ is called the *source strength*. At any fixed value of time, this represents a cluster of concentric spheres filling all space. Each wavefront, or surface of constant phase, is given by

$$kr = \text{constant}$$

Notice that the amplitude of any spherical wave is a function of r, where the term r^{-1} serves as an attenuation factor. Unlike the plane wave, a spherical wave decreases in amplitude, thereby changing its profile, as it expands and moves out from the origin.* Figure 2.24 illustrates this graphically by showing a "multiple exposure" of a spherical pulse at four different times. The pulse has the same extent in space at any point along any radius r; that is, the width of the pulse along the r-axis is a constant. Figure 2.25 is an attempt to relate the diagrammatic representation of $\psi(r, t)$ in the previous figure to its actual form as a spherical wave. It depicts half the spherical pulse at two different times, as the wave expands outward. Remember that these results would obtain regardless of the direction of r, because of the spherical symmetry. We could also have drawn a harmonic wave, rather than a pulse, in Figs. 2.24 and 2.25. In this case, the sinusoidal disturbance would have been bounded by the curves

$$\psi = \mathscr{A}/r \quad \text{and} \quad \psi = -\mathscr{A}/r$$

The outgoing spherical wave emanating from a point source and the incoming wave converging to a point are idealizations. In actuality, light only approximates spherical waves, as it also only approximates plane waves.

As a spherical wavefront propagates out, its radius increas-

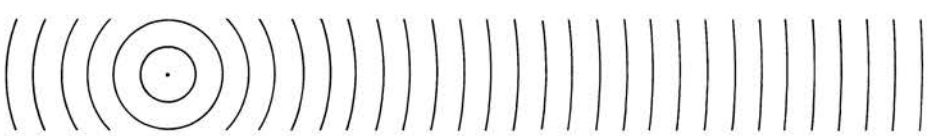


FIGURE 2.26 The flattening of spherical waves with distance.

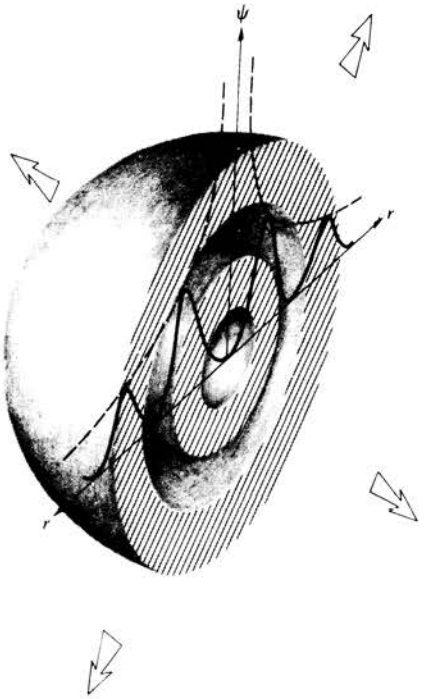


Figure 2.25 Spherical wavefronts.

*The attenuation factor is a direct consequence of energy conservation. Chapter 3 contains a discussion of how these ideas apply specifically to electromagnetic radiation.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

es. Far enough away from the source, a small area of the wavefront will closely resemble a portion of a plane wave (Fig. 2.26).

2.10 Cylindrical Waves

We will now briefly examine another idealized waveform, the infinite circular cylinder. Unfortunately, a precise mathematical treatment is far too involved to do here. We shall, instead, just outline the procedure. The Laplacian of ψ in cylindrical coordinates (Fig. 2.27) is

$$\nabla^2\psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{\partial^2\psi}{\partial z^2} \tag{2.76}$$

where

$$x = r\cos\theta, \quad y = r\sin\theta, \quad \text{and} \quad z = z$$

The simple case of cylindrical symmetry requires that

$$\psi(\vec{r}) = \psi(r, \theta, z) = \psi(r)$$

The θ-independence means that a plane perpendicular to the z-axis will intersect the wavefront in a circle, which may vary in r, at different values of z. In addition, the z-independence further restricts the wavefront to a right circular cylinder centered on the z-axis and having infinite length. The differential wave equation becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \tag{2.77}$$

After a bit of manipulation, in which the time dependence is separated out, Eq. (2.77) becomes something called **Bessel's** equation. The solutions of Bessel's equation for large values of r gradually approach simple trigonometric forms. When r is sufficiently large,

$$\psi(r, t) \approx \frac{\mathscr{A}}{\sqrt{r}}e^{ik(r\mp vt)}$$

$$\psi(r, t) \approx \frac{\mathscr{A}}{\sqrt{r}}\cos k(r \mp vt) \tag{2.78}$$

This represents a set of coaxial circular cylinders filling all space and traveling toward or away from an infinite line source. No solutions in terms of arbitrary functions can now be found as there were for both spherical [Eq. (2.73)] and plane [Eq. (2.66)] waves.

A plane wave impinging on the back of a flat opaque screen containing a long, thin slit will result in the emission, from that slit, of a disturbance resembling a cylindrical wave (see Fig. 2.28). Extensive use has been made of this technique to generate cylindrical lightwaves (p. 394).

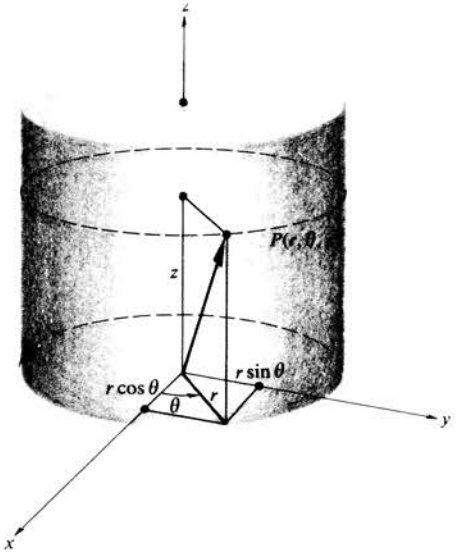


Figure 2.27 The geometry of cylindrical coordinates.

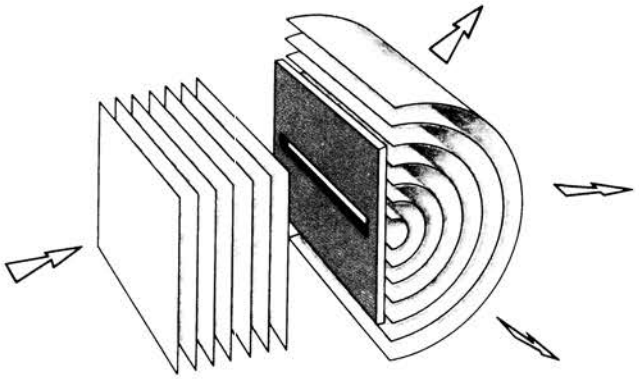


Figure 2.28 Cylindrical waves emerging from a long, narrow slit.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

PROBLEMS

Complete solutions to all problems—except those with an asterisk—can be found in the back of the book.

2.1 How many "yellow" lightwaves ($\lambda = 580$ nm) will fit into a distance in space equal to the thickness of a piece of paper (0.003 in.)? How far will the same number of microwaves ($\nu = 10^{10}$ Hz, i.e., 10 GHz, and $v = 3 \times 10^8$ m/s) extend?

2.2* The speed of light in vacuum is approximately 3×10^8 m/s. Find the wavelength of red light having a frequency of 5×10^{14} Hz. Compare this with the wavelength of a 60-Hz electromagnetic wave.

2.3* It is possible to generate ultrasonic waves in crystals with wavelengths similar to light (5×10^{-5} cm) but with lower frequencies (6×10^8 Hz). Compute the corresponding speed of such a wave.

2.4* A youngster in a boat on a lake watches waves that seem to be an endless succession of identical crests passing with a half-second interval between each. If every disturbance takes 1.5 s to sweep straight along the length of her 4.5-m-long boat, what are the frequency, period, and wavelength of the waves?

2.5* A vibrating hammer strikes the end of a long metal rod in such a way that a periodic compression wave with a wavelength of 4.3 m travels down the rod's length at a speed of 3.5 km/s. What was the frequency of the vibration?

2.6 A violin is submerged in a swimming pool at the wedding of two scuba divers. Given that the speed of compression waves in pure water is 1498 m/s, what is the wavelength of an A-note of 440 Hz played on the instrument?

2.7* A wavepulse travels 10 m along the length of a string in 2.0 s. A harmonic disturbance of wavelength 0.50 m is then generated on the string. What is its frequency?

2.8* Show that for a periodic wave $\omega = (2\pi/\lambda)v$.

2.9* Make up a table with columns headed by values of θ running from $-\pi/2$ to 2π in intervals of $\pi/4$. In each column place the corresponding value of $\sin\theta$, beneath those the values of $\cos\theta$, beneath those the values of $\sin(\theta - \pi/4)$, and similarly with the functions $\sin(\theta - \pi/2)$, $\sin(\theta - 3\pi/4)$, and $\sin(\theta + \pi/2)$. Plot each of these functions, noting the effect of the phase shift. Does $\sin\theta$ lead or lag $\sin(\theta - \pi/2)$. In other words, does one of the functions reach a particular magnitude at a smaller value of θ than the other and therefore lead the other (as $\cos\theta$ leads $\sin\theta$)?

2.10* Make up a table with columns headed by values of kx running from $x = -\lambda/2$ to $x = +\lambda$ in intervals of x of $\lambda/4$—of course, $k = 2\pi/\lambda$. In each column place the corresponding values of $\cos(kx - \pi/4)$ and beneath that the values of $\cos(kx + 3\pi/4)$. Next plot the functions $15\cos(kx - \pi/4)$ and $25\cos(kx + 3\pi/4)$.

2.11* Make up a table with columns headed by values of ωt running from $t = -\tau/2$ to $t = +\tau$ in intervals of t of $\tau/4$—of course, $\omega = 2\pi/\tau$. In each column place the corresponding values of $\sin(\omega t + \pi/4)$ and $\sin(\pi/4 - \omega t)$ and then plot these two functions.

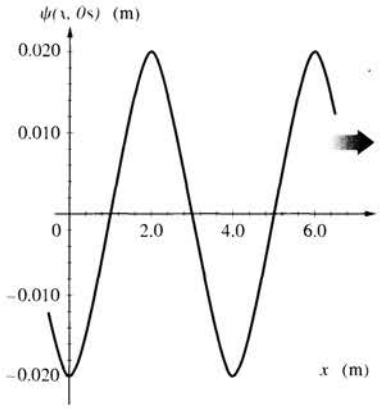
2.12* The profile of a transverse harmonic wave, traveling at 1.2 m/s on a string, is given by

$$y = (0.02 \text{ m}) \sin(157 \text{ m}^{-1})\, x$$

Determine its amplitude, wavelength, frequency, and period.

2.13* Figure P.2.13 represents the profile ($t = 0$) of a transverse wave on a string traveling in the positive x-direction at a speed of 20.0 m/s. (a) Determine its wavelength. (b) What is the frequency of the wave? (c) Write down the wavefunction for the disturbance. (d) Notice that as the wave passes any fixed point on the x-axis the string at that location oscillates in time. Draw a graph of the ψ versus t showing how a point on the rope at $x = 0$ oscillates.

Figure P.2.13

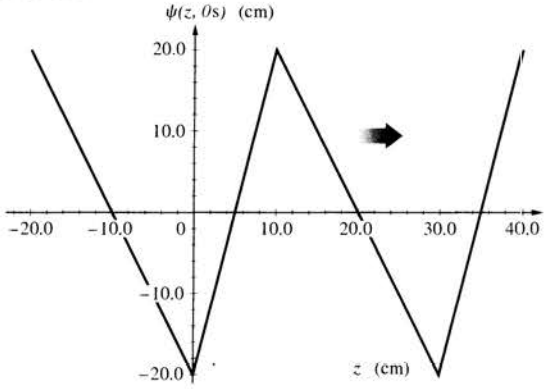


2.14* Figure P.2.14 represents the profile ($t = 0$) of a transverse wave on a string traveling in the positive z-direction at a speed of 100 cm/s. (a) Determine its wavelength. (b) Notice that as the wave passes any fixed point on the z-axis the string at that location oscil-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

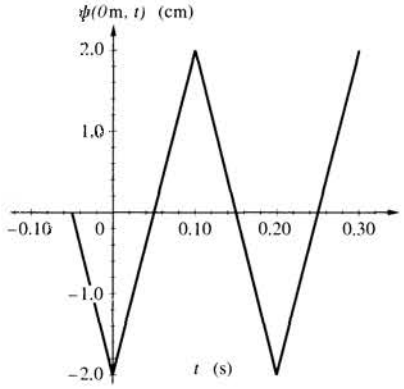
lates in time. Draw a graph of the ψ versus t showing how a point on the rope at $x = 0$ oscillates. (c) What is the frequency of the wave?

Figure P.2.14



2.15* A transverse wave on a string travels in the negative y-direction at a speed of 40.0 cm/s. Figure P.2.15 is a graph of ψ versus t showing how a point on the rope at $y = 0$ oscillates. (a) Determine the wave's period. (b) What is the frequency of the wave? (b) What is the wavelength of the wave? (d) Sketch the profile of the wave (ψ versus y).

Figure P.2.15



2.16 Given the wavefunctions

$$\psi_1 = 4 \sin 2\pi(0.2x - 3t)$$

and

$$\psi_2 = \frac{\sin (7x + 3.5t)}{2.5}$$

determine in each case the values of (a) frequency, (b) wavelength, (c) period, (d) amplitude, (e) phase velocity, and (f) direction of motion. Time is in seconds and x is in meters.

2.17* The wavefunction of a transverse wave on a string is

$$\psi(x, t) = (30.0 \text{ cm}) \cos [(6.28 \text{ rad/m})x - (20.0 \text{ rad/s})t]$$

Compute the (a) frequency, (b) wavelength, (c) period, (d) amplitude, (e) phase velocity, and (f) direction of motion.

2.18* Show that

$$\psi(x, t) = A \sin k(x - vt) \qquad [2.13]$$

is a solution of the differential wave equation.

2.19* Show that

$$\psi(x, t) = A \cos (kx - \omega t)$$

is a solution of the differential wave equation.

2.20* Prove that

$$\psi(x, t) = A \cos (kx - \omega t - \pi/2)$$

is equivalent to

$$\psi(x, t) = A \sin (kx - \omega t)$$

2.21 Show that if the displacement of the string in Fig. 2.7 is given by

$$y(x, t) = A \sin [kx - \omega t + \varepsilon]$$

then the hand generating the wave must be moving vertically in simple harmonic motion.

2.22 Write the expression for the wavefunction of a harmonic wave of amplitude 10^3 V/m, period 2.2×10^{-15} s, and speed 3×10^8 m/s. The wave is propagating in the negative x-direction and has a value of 10^3 V/m at $t = 0$ and $x = 0$.

2.23 Consider the pulse described in terms of its displacement at $t = 0$ by

$$y(x, t)|_{t=0} = \frac{C}{2 + x^2}$$

where C is a constant. Draw the wave profile. Write an expression for the wave, having a speed v in the negative x-direction, as a function of time t. If $v = 1$ m/s, sketch the profile at $t = 2$ s.

2.24* Please determine the magnitude of the wavefunction $\psi(z, t) = A \cos [k(z + vt) + \pi]$ at the point $z = 0$, when $t = \tau/2$ and when $t = 3\tau/4$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

2.25 Does the following function, in which A is a constant,

$$\psi(y, t) = (y - vt)A$$

represent a wave? Explain your reasoning.

2.26* Use Eq. (2.33) to calculate the speed of the wave whose representation in SI units is

$$\psi(y, t) = A \cos \pi(3 \times 10^6 y + 9 \times 10^{14} t)$$

2.27 Beginning with the following theorem: If $z = f(x, y)$ and $x = g(t)$, $y = h(t)$, then

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$

Derive Eq. (2.34).

2.28 Using the results of the previous problem, show that for a harmonic wave with a phase $\varphi(x, t) = k(x - vt)$ we can determine the speed by setting $d\varphi/dt = 0$. Apply the technique to Problem 2.26 to find the speed of that wave.

2.29* A Gaussian wave has the form $\psi(x, t) = Ae^{-a(bx + ct)^2}$. Use the fact that $\psi(x, t) = f(x \mp vt)$ to determine its speed and then verify your answer using Eq. (2.34).

2.30 Create an expression for the *profile* of a harmonic wave traveling in the z-direction whose magnitude at $z = -\lambda/12$ is 0.866, at $z = +\lambda/6$ is $1/2$, and at $z = \lambda/4$ is 0.

2.31 Which of the following expressions correspond to traveling waves? For each of those, what is the speed of the wave? The quantities a, b, and c are positive constants.

$$\text{(a) } \psi(z, t) = (az - bt)^2$$

$$\text{(b) } \psi(x, t) = (ax + bt + c)^2$$

$$\text{(c) } \psi(x, t) = 1/(ax^2 + b)$$

2.32* Determine which of the following describe traveling waves:

$$\text{(a) } \psi(y, t) = e^{-(ax^2 + bt^2 - 2abtx)}$$

$$\text{(b) } \psi(z, t) = A \sin(az^2 - bt^2)$$

$$\text{(c) } \psi(x, t) = A \sin 2\pi\left(\frac{x}{a} + \frac{t}{b}\right)^2$$

$$\text{(d) } \psi(x, t) = A \cos^2 2\pi(t - x)$$

Where appropriate, draw the profile and find the speed and direction of motion.

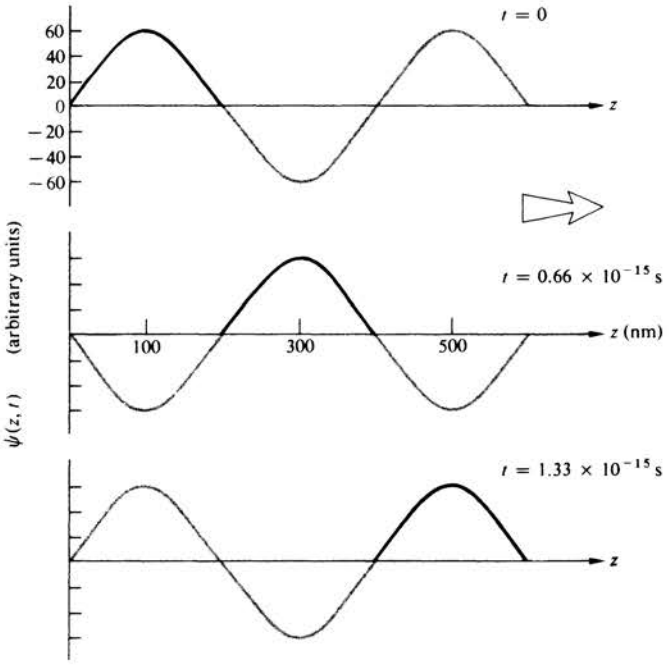
2.33 Given the traveling wave $\psi(x, t) = 5.0 \exp(-ax^2 - bt^2 - 2\sqrt{ab}\,xt)$, determine its direction of propagation. Calculate a few values of ψ and make a sketch of the wave at $t = 0$, taking $a = 25\ \text{m}^{-2}$ and $b = 9.0\ \text{s}^{-2}$. What is the speed of the wave?

2.34* Imagine a sound wave with a frequency of 1.10 kHz propagating with a speed of 330 m/s. Determine the phase difference in radians between any two points on the wave separated by 10.0 cm.

2.35 Consider a lightwave having a phase velocity of 3×10^8 m/s and a frequency of 6×10^{14} Hz. What is the shortest distance along the wave between any two points that have a phase difference of 30°? What phase shift occurs at a given point in 10^{-6} s, and how many waves have passed by in that time?

2.36 Write an expression for the wave shown in Fig. P.2.36. Find its wavelength, velocity, frequency, and period.

Figure P.2.36 A harmonic wave.



2.37* Working with exponentials directly, show that the magnitude of $\psi = Ae^{i\omega t}$ is A. Then rederive the same result using Euler's formula. Prove that $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$.

2.38* Show that the imaginary part of a complex number $\tilde{z}$ is given by $(\tilde{z} - \tilde{z}^*)/2i$.

2.39 Beginning with Eq. (2.51), verify that

$$\psi(x, y, z, t) = Ae^{i[k(\alpha x + \beta y + \gamma z) \mp \omega t]}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

and that $\alpha^2 + \beta^2 + \gamma^2 = 1$

Draw a sketch showing all the pertinent quantities.

2.40* Show that Eqs. (2.64) and (2.65), which are plane waves of arbitrary form, satisfy the three-dimensional differential wave equation.

2.41 De Broglie's hypothesis states that every particle has associated with it a wavelength given by Planck's constant ($h = 6.6 \times 10^{-34}$ J · s) divided by the particle's momentum. Compare the wavelength of a 6.0-kg stone moving at a speed of 1.0 m s with that of light.

2.42 Write an expression in Cartesian coordinates for a harmonic plane wave of amplitude A and frequency ω propagating in the direction of the vector $\vec{\mathbf{k}}$, which in turn lies on a line drawn from the origin to the point (4, 2, 1). *Hint:* First determine $\vec{\mathbf{k}}$ and then dot it with $\vec{\mathbf{r}}$.

2.43* Write an expression in Cartesian coordinates for a harmonic plane wave of amplitude A and frequency ω propagating in the positive x-direction.

2.44 Show that $\psi(\vec{k} \cdot \vec{r}, t)$ may represent a plane wave where $\vec{\mathbf{k}}$ is normal to the wavefront. *Hint:* Let $\vec{\mathbf{r}}_1$ and $\vec{\mathbf{r}}_2$ be position vectors drawn to any two points on the plane and show that $\psi(\vec{r}_1, t) = \psi(\vec{r}_2, t)$.

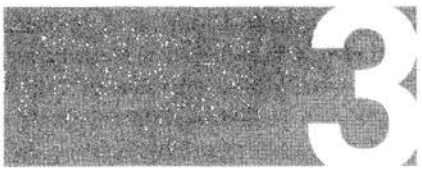
2.45* Make up a table with columns headed by values of θ running from $-\pi$ 2 to 2π in intervals of π 4. In each column place the corresponding value of sin θ, and beneath those the values of 2 sin θ. Next add these, column by column, to yield the corresponding values of the function sin θ + 2 sin θ. Plot each of these three functions, noting their relative amplitudes and phases.

2.46* Make up a table with columns headed by values of θ running from $-\pi$ 2 to 2π in intervals of π 4. In each column place the corresponding value of sin θ, and beneath those the values of sin (θ − π 2). Next add these, column by column, to yield the corresponding values of the function sin θ + sin (θ − π 2). Plot each of these three functions, noting their relative amplitudes and phases.

2.47* With the last two problems in mind, draw a plot of the three functions (a) sin θ, (b) sin (θ − 3π 4), and (c) sin θ + sin (θ − 3π 4). Compare the amplitude of the combined function (c) in this case with that of the previous problem.

2.48* Make up a table with columns headed by values of kx running from $x = -\lambda$ 2 to $x = +\lambda$ in intervals of x of λ 4. In each column place the corresponding values of cos kx and beneath that the values of cos (kx + π). Next plot the three functions cos kx, cos (kx + π), and cos kx + cos (kx + π).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Electromagnetic Theory, Photons, and Light

The work of J. Clerk Maxwell and subsequent developments since the late 1800s have made it evident that light is most certainly electromagnetic in nature. Classical electrodynamics, as we shall see, unalterably leads to the picture of a continuous transfer of energy by way of electromagnetic waves. In contrast, the more modern view of Quantum Electrodynamics (p. 80) describes electromagnetic interactions and the transport of energy in terms of massless elementary "particles" known as *photons*. The quantum nature of radiant energy is not always apparent, nor is it always of practical concern in Optics. There is a range of situations in which the detecting equipment is such that it is impossible, and desirably so, to distinguish individual quanta.

If the wavelength of light is small in comparison to the size of the apparatus (lenses, mirrors, etc.), one may use, as a first approximation, the techniques of *Geometrical Optics*. A somewhat more precise treatment, which is applicable as well when the dimensions of the apparatus are small, is that of *Physical Optics*. In Physical Optics the dominant property of light is its wave nature. It is even possible to develop most of the treatment without ever specifying the kind of wave one is dealing with. Certainly, as far as the classical study of Physical Optics is concerned, it will suffice admirably to treat light as an electromagnetic wave.

We can think of light as the most tenuous form of matter. Indeed, one of the basic tenets of Quantum Mechanics is that both light and material particles display similar wave-particle properties. As Erwin C. Schrödinger (1887–1961), one of the founders of quantum theory, put it:

> In the new setting of ideas the distinction [between particles and waves] has vanished, because it was discovered that all particles have also wave properties, and *vice versa*. Neither of the two concepts must be discarded, they must be amalgamated. Which aspect obtrudes itself depends not on the physical object, but on the experimental device set up to examine it.*

The quantum-mechanical treatment associates a wave equation with a particle, be it a photon, electron, proton, or whatever. In the case of material particles, the wave aspects are introduced by way of the field equation known as Schrödinger's Equation. For light we have a representation of the wave nature in the form of the classical electromagnetic field equations of Maxwell. With these as a starting point one can construct a quantum-mechanical theory of photons and their interaction with charges. The dual nature of light is evidenced by the fact that it propagates through space in a wavelike fashion and yet displays particlelike behavior during the processes of emission and absorption. Electromagnetic radiant energy is created and destroyed in quanta or photons and not continuously as a classical wave. Nonetheless, its motion through a lens, a hole, or a set of slits is governed by wave characteristics. If we're unfamiliar with this kind of behavior in the macroscopic world, it's because the wavelength of a material object varies inversely with its momentum (p. 56), and even a grain of sand (which is barely moving) has a wavelength so small as to be indiscernible in any conceivable experiment.

The photon has zero mass, and therefore exceedingly large numbers of low-energy photons can be envisioned as present in a beam of light. Within that model, dense streams of photons act on the average to produce well-defined classical fields (p. 52). We can draw a rough analogy with the flow of commuters through a train station during rush hour. Each commuter presumably behaves individually as a quantum of humanity, but all have the same intent and follow fairly simi-

*Erwin C. Schrodinger, *Science Theory and Man.*

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

lar trajectories. To a distant, myopic observer there is a seemingly smooth and continuous flow. The behavior of the stream *en masse* is predictable from day to day, and so the precise motion of each commuter is unimportant, at least to the observer. The energy transported by a large number of photons is, *on the average*, equivalent to the energy transferred by a classical electromagnetic wave. For these reasons the classical field representation of electromagnetic phenomena has been, and will continue to be, so useful. Nonetheless, it should be understood that the apparent continuous nature of electromagnetic waves is a fiction of the macroscopic world, just as the apparent continuous nature of ordinary matter is a fiction—it just isn't that simple.

Quite pragmatically, then, we can consider light to be a classical electromagnetic wave, keeping in mind that there are situations for which this description is woefully inadequate.

3.1 Basic Laws of Electromagnetic Theory

Our intent in this section is to review and develop some of the ideas needed to appreciate the concept of electromagnetic waves.

We know from experiments that charges, even though separated in vacuum, experience a mutual interaction. Recall the familiar electrostatics demonstration in which a pith ball somehow senses the presence of a charged rod without actually touching it. As a possible explanation we might speculate that each charge emits (and absorbs) a stream of undetected particles (*virtual photons*). The exchange of these particles among the charges may be regarded as the mode of interaction. Alternatively, we can take the classical approach and imagine instead that every charge is surrounded by something called an electric field. We then need only suppose that each charge interacts directly with the electric field in which it is immersed. Thus if a point charge q_0 experiences a force $\vec{\mathbf{F}}_E$, the *electric field* $\vec{\mathbf{E}}$ at the position of the charge is defined by $\vec{\mathbf{F}}_E = q_0\vec{\mathbf{E}}$. In addition, we observe that a moving charge may experience another force $\vec{\mathbf{F}}_M$, which is proportional to its velocity $\vec{\mathbf{v}}$. We are thus led to define yet another field, namely, the *magnetic induction* or just the *magnetic field* $\vec{\mathbf{B}}$, such that $\vec{\mathbf{F}}_M = q_0\vec{\mathbf{v}} \times \vec{\mathbf{B}}$. If forces $\vec{\mathbf{F}}_E$ and $\vec{\mathbf{F}}_M$ occur concurrently, the charge is moving through a region pervaded by both electric and magnetic fields, whereupon $\vec{\mathbf{F}} = q_0\vec{\mathbf{E}} + q_0\vec{\mathbf{v}} \times \vec{\mathbf{B}}$.

As we'll see, electric fields are generated by both electric charges and *time-varying magnetic fields*. Similarly, magnetic fields are generated by electric currents and by *time-varying electric fields*. This interdependence of $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ is a key point in the description of light.

3.1.1 Faraday's Induction Law

"Convert magnetism into electricity" was the brief remark Michael Faraday jotted in his notebook in 1822, a challenge he set himself with an easy confidence that made it seem so attainable. After several years doing other research, Faraday returned to the problem of electromagnetic induction in 1831. His first apparatus made use of two coils mounted on a wooden spool (Fig. 3.1*a*). One, called the primary, was attached to a battery and a switch; the other, the secondary, was attached to a galvanometer. He found that the galvanometer deflected in one direction just for a moment whenever the switch was closed, returning to zero almost immediately, despite the con-

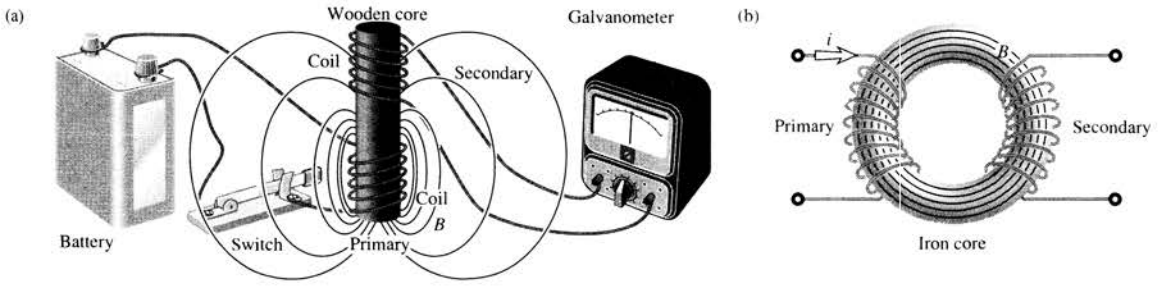


Figure 3.1 (*a*) The start of a current in one coil produces a time varying magnetic field that induces a current in the other coil. (*b*) An iron core couples the primary coil to the secondary.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

stant current still in the primary. Whenever the switch was opened, interrupting the primary current, the galvanometer in the secondary circuit momentarily swung in the opposite direction and then promptly returned to zero.

Using a ferromagnetic core to concentrate the "magnetic force," Faraday wound two coils around opposing sections of a soft iron ring (Fig. 3.1*b*). Now the effect was unmistakable—***a changing magnetic field generated a current***. Indeed, as he would continue to discover, *change* was the essential aspect of electromagnetic induction.

By thrusting a magnet into a coil, Faraday showed that there is a voltage—otherwise known as the *induced electromotive force* or **emf**—across the coil's terminals. (Electromotive force is a dreadful, outmoded term—it's not a force, but a voltage—so we'll avoid it and just use emf.) Furthermore, the amplitude of the emf depends on how rapidly the magnet is moved. *The induced emf depends on the rate-of-change of B* through the coil and not on *B* itself. A weak magnet moved rapidly can induce a greater emf than a strong magnet moved slowly.

When the same changing *B*-field passes through two different wire loops, as in Fig. 3.2, the induced emf is larger across the terminals of the larger loop. In other words, here where the *B*-field is changing, ***the induced emf is proportional to the area A of the loop penetrated perpendicularly by the field***. If the loop is successively tilted over, as is shown in Fig. 3.3, the area presented perpendicularly to the field ($A_\perp$) varies as $A \cos \theta$, and, when $\theta = 90°$, the induced emf is zero because no amount of *B*-field then penetrates the loop: when $\Delta B/\Delta t \neq 0$, emf $\propto A_\perp$. The converse also holds: ***when the field is constant, the induced emf is proportional to the rate-of-change of the perpendicular area penetrated***. If a coil is twisted or rotated or even squashed while in a constant *B*-field so that the perpendicular area initially penetrated is altered, there will be an induced emf $\propto \Delta A_\perp / \Delta t$ and it will be proportional to *B*. In summary, when $A_\perp$ = constant, emf $\propto A_\perp \Delta B/\Delta t$ and, when B = constant, emf $\propto B\Delta A_\perp/\Delta t$.

All of this suggests that the emf depends on the rate-of-change of both $A_\perp$ and *B*, that is, on the rate of change of their product. This should bring to mind the notion of the flux of the field—the product of field and area where the penetration is perpendicular. Accordingly, the **flux of the magnetic field** through the wire loop is

$$\Phi_M = B_\perp A = BA_\perp = BA \cos \theta$$

More generally, if *B* varies in space as it's likely to, the flux of the magnetic field through any open area *A* bounded by the conducting loop (Fig. 3.4) is given by

$$\Phi_M = \iint_A \vec{\mathbf{B}} \cdot d\vec{\mathbf{S}} \tag{3.1}$$

The induced emf, developed around the loop, is then

$$\text{emf} = -\frac{d\Phi_M}{dt} \tag{3.2}$$

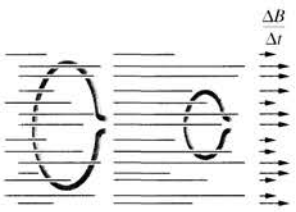
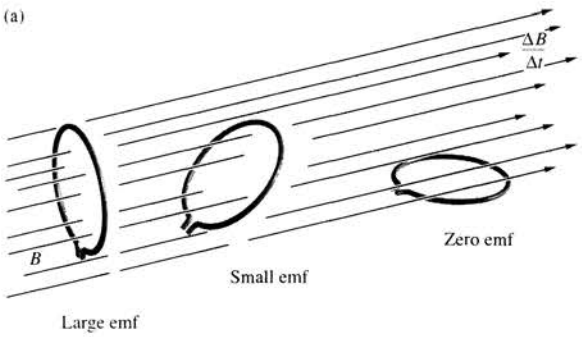


Figure 3.2 A larger time-varying magnetic flux passes through the larger loop and induces a greater emf across its terminals.



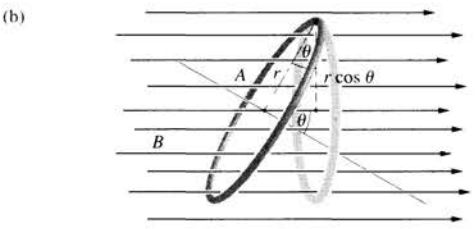


Figure 3.3 (*a*) The induced emf is proportional to the perpendicular area intercepted by the magnetic field. (*b*) That perpendicular area varies as $\cos \theta$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

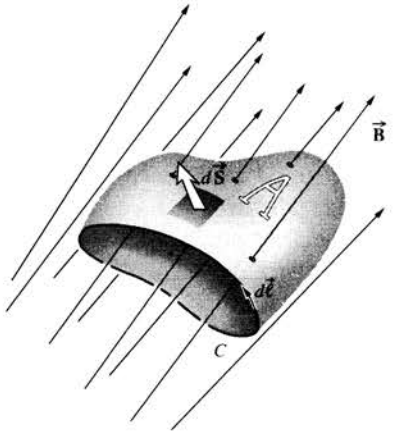


Figure 3.4 $\vec{\mathbf{B}}$-field through an open area *A*.

We should not, however, get too involved with the image of wires and current and emf. Our present concern is with the electric and magnetic fields themselves.

In very general terms, an emf is a potential difference, and that's a potential-energy difference per unit charge. A potential-energy difference per unit charge corresponds to work done per unit charge, which is force per unit charge times distance, and that's electric field times distance. The emf exists only as a result of the presence of an electric field:

$$\text{emf} = \oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} \tag{3.3}$$

taken around the closed curve *C*, corresponding to the loop. Equating Eqs. (3.2) and (3.3), and making use of Eq. (3.1), we get

$$\oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} = -\frac{d}{dt} \iint_A \vec{\mathbf{B}} \cdot d\vec{\mathbf{S}} \tag{3.4}$$

We began this discussion by examining a conducting loop, and have arrived at Eq. (3.4); this expression, except for the path *C*, contains no reference to the physical loop. In fact, the path can be chosen arbitrarily and need not be within, or anywhere near, a conductor. The electric field in Eq. (3.4) arises not from the presence of electric charges but rather from the time-varying magnetic field. With no charges to act as sources or sinks, the field lines close on themselves, forming loops (Fig. 3.5).

We are interested in electromagnetic waves traveling in space where there are no wire loops, and the magnetic flux changes because $\vec{\mathbf{B}}$ changes. The *Induction Law* (Eq. 3.4) can then be rewritten as

$$\oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} = -\iint_A \frac{\partial \vec{\mathbf{B}}}{\partial t} \cdot d\vec{\mathbf{S}} \tag{3.5}$$

A partial derivative with respect to t is taken because $\vec{\mathbf{B}}$ is usually also a function of the space variables. This expression in itself is rather fascinating, since it indicates that ***a time-varying magnetic field will have an electric field associated with it.***

3.1.2 Gauss's Law—Electric

Another fundamental law of electromagnetism is named after the German mathematician Karl Friedrich Gauss (1777–1855). Gauss's Law is about the relationship between the flux of the electric field and the sources of that flux, charge. The ideas derive from fluid dynamics, where both the concepts of field and flux were introduced. The flow of a fluid, as represented by its velocity field, is depicted via streamlines, much as the electric field is pictured via field lines. Figure 3.6 shows a portion of a moving fluid within which there is a region isolated by an imaginary closed surface. The discharge rate, or *volume flux* (Av), is the volume of fluid flowing past a point in the tube per unit time. The volume flux through both end surfaces

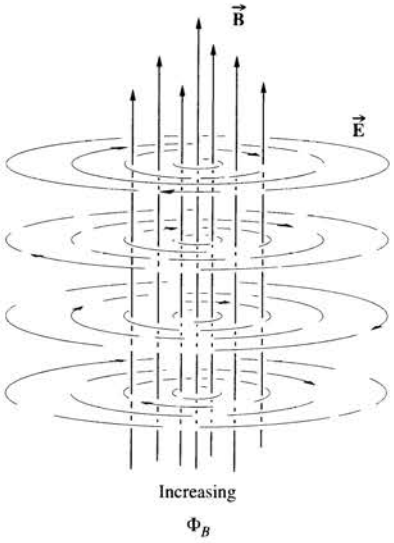


Figure 3.5 A time-varying $\vec{\mathbf{B}}$-field. Surrounding each point where Φ_M is changing, the $\vec{\mathbf{E}}$-field forms closed loops.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Johann Karl
Friedrich Gauss

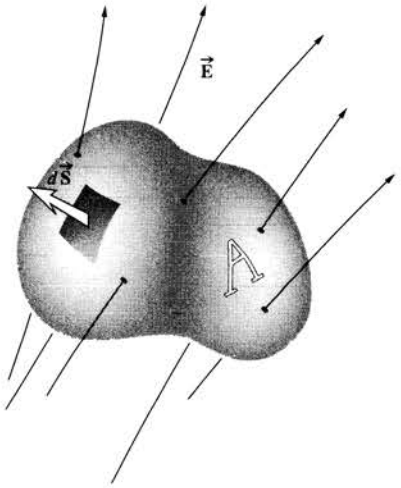


Figure 3.7 $\vec{E}$-field through a closed area A

is equal in magnitude—what flows in per second flows out per second. The net fluid flux (into and out of the closed area) summed over all the surfaces equals zero. If, however, a small pipe is inserted into the region either sucking out fluid (a sink) or delivering fluid (a source), the net flux will then be non-zero.

To apply these ideas to the electric field, consider an imaginary closed area A placed in some arbitrary electric field, as depicted in Fig. 3.7. The flux of electric field through A is taken to be

$$\Phi_E = \oiint_A \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} \tag{3.6}$$

The circled double integral serves as a reminder that the surface is closed. The vector $d\mathbf{S}$ is in the direction of an *outward normal*. ***When there are no sources or sinks of the electric field within the region encompassed by the closed surface, the net flux through the surface equals zero***—that much is a general rule for all such fields.

In order to find out what would happen in the presence of internal sources and sinks, consider a spherical surface of radius r centered on and surrounding a positive point-charge ($q_\bullet$) in vacuum. The E-field is everywhere outwardly radial, and at any distance r it is entirely perpendicular to the surface: $E = E_\perp$ and so

$$\Phi_E = \oiint_A E_\perp \, dS = \oiint_A E \, dS$$

Moreover, since E is constant over the surface of the sphere, it can be taken out of the integral:

$$\Phi_E = E \oiint_A dS = E 4\pi r^2$$

But we know from Coulomb's Law that the point-charge has an electric field given by

$$E = \frac{1}{4\pi\epsilon_0} \frac{q_\bullet}{r^2}$$

and so

$$\Phi_E = \frac{q_\bullet}{\epsilon_0}$$

This is the electric flux associated with a single point-charge $q_\bullet$ within the closed surface. Since all charge distributions are made up of point-charges, it's reasonable that the net flux due to a number of charges contained within any closed area is

$$\Phi_E = \frac{1}{\epsilon_0} \sum q_\bullet$$

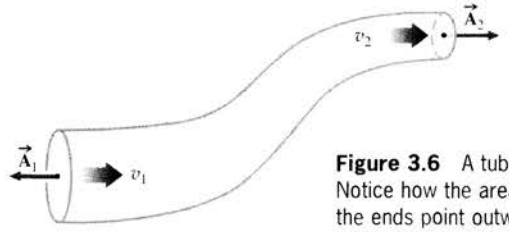


Figure 3.6 A tube of fluid flow. Notice how the area vectors on the ends point outward.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Combining the two equations for Φ_E, we get **Gauss's Law**,

$$\oint_A \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = \frac{1}{\epsilon_0} \sum q.$$

In order to apply the calculus, it's useful to approximate the charge distribution as being continuous. Then if the volume enclosed by A is V and the charge distribution has a density density ρ, Gauss's Law becomes

$$\oint_A \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = \frac{1}{\epsilon_0} \iiint_V \rho dV \tag{3.7}$$

Electric Permittivity

For the special case of vacuum, the electric permittivity *of free space* is given by $\epsilon_0 = 8.854\,2 \times 10^{-12}\ \mathrm{C}^2\ \ \mathrm{N \cdot m}^2$. If the charge is embedded in some material medium its permittivity (ϵ) will appear in Eq. (3.7) instead of ϵ_0. One function of the permittivity in Eq. (3.7) is, of course, to balance out the units, but the concept is basic to the description of the parallel plate capacitor (see Section 3.1.4). There it's the medium-dependent proportionality constant between the device's capacitance and its geometric characteristics. Indeed ϵ is often measured by a procedure in which the material under study is placed within a capacitor. Conceptually, the permittivity embodies the electrical behavior of the medium: in a sense, it is a measure of the degree to which the material is permeated by the electric field in which it is immersed.

In the early days of the development of the subject, people in various areas worked in different systems of units, a state of affairs leading to some obvious difficulties. This necessitated the tabulation of numerical values for ϵ in each of the different systems, which was, at best, a waste of time. The same problem regarding densities was neatly avoided by using specific gravity (i.e., density ratios). Thus it was advantageous to tabulate values not of ϵ but of a new related quantity that is independent of the system of units being used. Accordingly, we define K_E as ϵ/ϵ_0. This is the **dielectric constant** (or *relative permittivity*), and it is appropriately unitless. The permittivity of a material can then be expressed in terms of ϵ_0 as

$$\epsilon = K_E \epsilon_0 \tag{3.8}$$

Our interest in K_E anticipates the fact that the permittivity is related to the speed of light in dielectric materials, such as glass, air, and quartz.

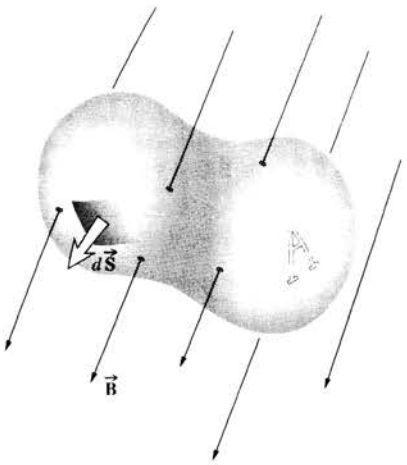


Figure 3.8 $\vec{\mathbf{B}}$-field through a closed area A.

3.1.3 Gauss's Law—Magnetic

There is no known magnetic counterpart to the electric charge; that is, no isolated magnetic poles have ever been found, despite extensive searching, even in lunar soil samples. Unlike the electric field, the magnetic field $\vec{\mathbf{B}}$ does not diverge from or converge toward some kind of magnetic charge (a monopole source or sink). Magnetic fields can be described in terms of current distributions. Indeed, we might envision an elementary magnet as a small current loop in which the lines of $\vec{\mathbf{B}}$ are continuous and closed. Any closed surface in a region of magnetic field would accordingly have an equal number of lines of $\vec{\mathbf{B}}$ entering and emerging from it (Fig. 3.8). This situation arises from the absence of any monopoles within the enclosed volume. The flux of magnetic field Φ_M through such a surface is zero, and we have the magnetic equivalent of Gauss's Law:

$$\Phi_M = \oint_A \vec{\mathbf{B}} \cdot d\vec{\mathbf{S}} = 0 \tag{3.9}$$

3.1.4 Ampère's Circuital Law

Another equation that will be of great interest is associated with André Marie Ampère (1775–1836). Known as the *Circuital Law*, its physical origins are a little obscure—it will take a bit of doing to justify it, but it's worth it. Accordingly, imagine a straight current-carrying wire in vacuum and the circular B-field surrounding it (Fig. 3.9). We know from experiments that the magnetic field of a straight wire carrying a current i is $B = \mu_0 i/2\pi r$. Now, suppose we put ourselves back in time to the nineteenth century when it was common to think of mag-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

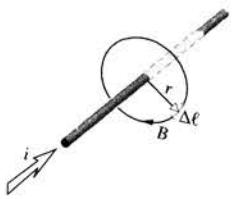


Figure 3.9 The $\vec{\mathbf{B}}$-field surrounding a current-carrying wire.

netic charge (q_m). Let's define this monopole charge so that it experiences a force when placed in a magnetic field B equal to $q_m B$ in the direction of B, just as an electric charge q_e experiences a force $q_e E$. Suppose we carry this north-seeking monopole around a closed circular path perpendicular to and centered on a current-carrying wire and determine the work done in the process. Since the direction of the force changes, because $\vec{\mathbf{B}}$ changes direction, we will have to divide the circular path into tiny segments ($\Delta\ell$) and sum up the work done over each. Work is the component of the force parallel to the displacement times the displacement: $\Delta W = q_m B_{\parallel}\,\Delta\ell$, and the total work done by the field is $\sum q_m B_{\parallel}\,\Delta\ell$. In this case, $\vec{\mathbf{B}}$ is everywhere tangent to the path, so that $B_{\parallel} = B = \mu_0 i/2\pi r$, which is constant around the circle. With both q_m and B constant, the summation becomes.

$$q_m \sum B_{\parallel}\,\Delta\ell = q_m B \sum \Delta\ell = q_m B 2\pi r$$

where $\sum \Delta\ell = 2\pi r$ is the circumference of the circular path.

If we substitute for B the equivalent current expression, which varies inversely with r, the radius cancels—the work is independent of the circular path taken. Since no work is done in traveling perpendicular to $\vec{\mathbf{B}}$, the work must be the same if we move q_m (out away from the wire or in toward it) along a radius, carrying it from one circular segment to another as we go around. Indeed, W is independent of path altogether—the work will be the same for *any closed path* encompassing the current. Putting in the current expression for B and canceling the "charge" q_m, we get the rather remarkable expression

$$\sum B_{\parallel}\,\Delta\ell = \mu_0 i$$

which is to be summed over any closed path surrounding i. The magnetic charge has disappeared, which is nice, since we no longer expect to be able to perform this little thought experiment with a monopole. Still, the physics was consistent, and the equation should hold, monopoles or no. Moreover, if there are several current-carrying wires encompassed by the closed path, their fields will superimpose and add, yielding a net field. The equation is true for the separate fields and must be true as well for the net field. Hence

$$\sum B_{\parallel}\,\Delta\ell = \mu_0 \sum i$$

As $\Delta\ell \to 0$, the sum becomes an integral *around a closed path*:

$$\oint_C \vec{\mathbf{B}} \cdot d\vec{\ell} = \mu_0 \sum i$$

Today this equation is known as **Ampère's Law**, though at one time it was commonly referred to as the "work rule." It relates a line integral of $\vec{\mathbf{B}}$ tangent to a closed curve C, with the total current i passing within the confines of C.

When the current has a nonuniform cross section, Ampère's Law is written in terms of the *current density* or current per unit area J, integrated over the area:

$$\oint_C \vec{\mathbf{B}} \cdot d\vec{\ell} = \mu_0 \iint_A \vec{\mathbf{J}} \cdot d\vec{\mathbf{S}} \qquad (3.10)$$

The open surface A is bounded by C (Fig. 3.10). The quantity μ_0 is called the **permeability of free space** and it's defined as $4\pi \times 10^{-7}\,\mathrm{N \cdot s^2\ C^2}$. When the current is imbedded in a material medium its permeability (μ) will appear in Eq. (3.10). As in Eq. (3.8),

$$\mu = K_M \mu_0 \qquad (3.11)$$

with K_M being the dimensionless *relative permeability*.

Equation (3.10), though often adequate, is not the whole

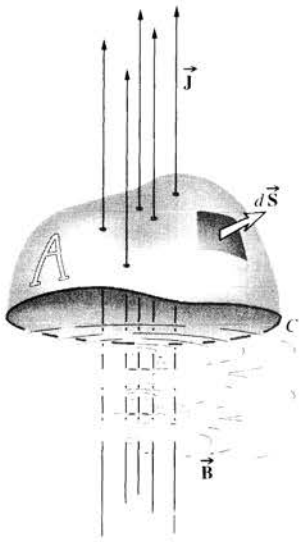


Figure 3.10 Current density through an open area A.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

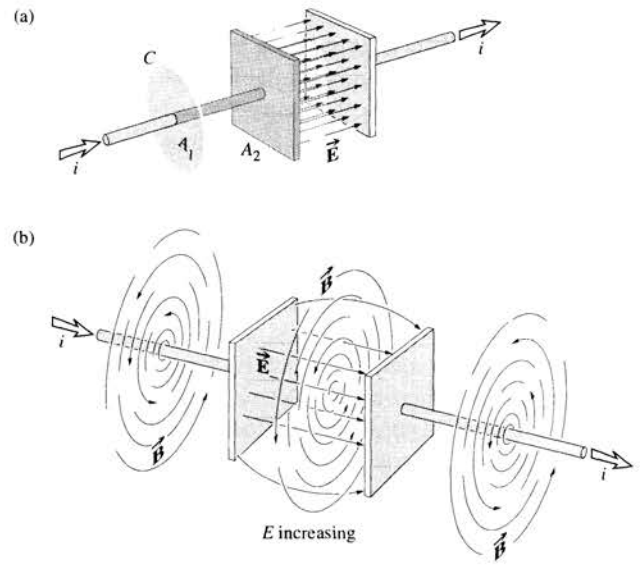


Figure 3.11 (*a*) Ampère's Law is indifferent to which area A_1 or A_2 is bounded by the path C. Yet a current passes through A_1 and not through A_2, and that means something is very wrong. (*b*) $\vec{\mathbf{B}}$-field concomitant with a time-varying $\vec{\mathbf{E}}$-field in the gap of a capacitor.

truth. Ampère's Law is not particular about the area used, provided it's bounded by the curve C, which makes for an obvious problem when charging a capacitor, as shown in Fig. 3.11*a*. If flat area A_1 is used, a net current of i flows through it and there is a $\vec{\mathbf{B}}$-field along curve C. The right side of Eq. (3.10) is nonzero, so the left side is nonzero. But if area A_2 is used instead to encompass C, no net current passes through it and the field must now be zero, even though nothing physical has actually changed. Something is obviously wrong!

Moving charges are not the only source of a magnetic field. While charging or discharging a capacitor, one can measure a $\vec{\mathbf{B}}$-field in the region between its plates (Fig. 3.11*b*), which is indistinguishable from the field surrounding the leads, even though no electric current actually traverses the capacitor. Notice, however, that if A is the area of each plate and Q the charge on it,

$$E = \frac{Q}{\epsilon A}$$

As the charge varies, the electric field changes, and taking the derivative of both sides yields

$$\epsilon \frac{\partial E}{\partial t} = \frac{i}{A}$$

which is effectively a $\vec{\mathbf{E}}$ current density. James Clerk Maxwell hypothesized the existence of just such a mechanism, which he called the *displacement current density*,* defined by

$$\vec{\mathbf{J}}_D \equiv \epsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \tag{3.12}$$

The restatement of Ampère's Law as

$$\oint_C \vec{\mathbf{B}} \cdot d\vec{\ell} = \mu \iint_A \left(\vec{\mathbf{J}} + \epsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \right) \cdot d\vec{\mathbf{S}} \tag{3.13}$$

was one of Maxwell's greatest contributions. It points out that even when $\vec{\mathbf{J}} = 0$, ***a time-varying $\vec{\mathbf{E}}$-field will be accompanied by a $\vec{\mathbf{B}}$-field*** (Fig. 3.7).

3.1.5 Maxwell's Equations

The set of integral expressions given by Eqs. (3.5), (3.7), (3.9), and (3.13) have come to be known as Maxwell's Equations. Remember that these are generalizations of experimental results. The simplest statement of Maxwell's Equations applies to the behavior of the electric and magnetic fields in free space, where $\epsilon = \epsilon_0$, $\mu = \mu_0$, and both ρ and $\vec{\mathbf{J}}$ are zero. In that instance,

$$\oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} = -\iint_A \frac{\partial \vec{\mathbf{B}}}{\partial t} \cdot d\vec{\mathbf{S}} \tag{3.14}$$

$$\oint_C \vec{\mathbf{B}} \cdot d\vec{\ell} = \mu_0 \epsilon_0 \iint_A \frac{\partial \vec{\mathbf{E}}}{\partial t} \cdot d\vec{\mathbf{S}} \tag{3.15}$$

$$\oiint_A \vec{\mathbf{B}} \cdot d\vec{\mathbf{S}} = 0 \tag{3.16}$$

$$\oiint_A \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = 0 \tag{3.17}$$

Observe that except for a multiplicative scalar, the electric and magnetic fields appear in the equations with a remarkable symmetry. However $\vec{\mathbf{E}}$ affects $\vec{\mathbf{B}}$, $\vec{\mathbf{B}}$ will in turn affect $\vec{\mathbf{E}}$. The mathematical symmetry implies a good deal of physical symmetry.

Maxwell's Equations can be written in a differential form, which will be somewhat more useful for our purposes. The

*Maxwell's own words and ideas concerning this mechanism are examined in an article by A. M. Bork, *Am. J. Phys.* **31**, 854 (1963). Incidentally, Clerk is pronounced *clark*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

appropriate calculation is carried out in Appendix 1, and the consequent equations for *free space*, in Cartesian coordinates, are as follows:

$$\frac{\partial \vec{\mathbf{E}}_z}{\partial y} - \frac{\partial \vec{\mathbf{E}}_y}{\partial z} = -\frac{\partial \vec{\mathbf{B}}_x}{\partial t} \quad \text{(i)}$$

$$\frac{\partial \vec{\mathbf{E}}_x}{\partial z} - \frac{\partial \vec{\mathbf{E}}_z}{\partial x} = -\frac{\partial \vec{\mathbf{B}}_y}{\partial t} \quad \text{(ii)} \qquad (3.18)$$

$$\frac{\partial \vec{\mathbf{E}}_y}{\partial x} - \frac{\partial \vec{\mathbf{E}}_x}{\partial y} = -\frac{\partial \vec{\mathbf{B}}_z}{\partial t} \quad \text{(iii)}$$

$$\frac{\partial \vec{\mathbf{B}}_z}{\partial y} - \frac{\partial \vec{\mathbf{B}}_y}{\partial z} = \mu_0 \epsilon_0 \frac{\partial \vec{\mathbf{E}}_x}{\partial t} \quad \text{(i)}$$

$$\frac{\partial \vec{\mathbf{B}}_x}{\partial z} - \frac{\partial \vec{\mathbf{B}}_z}{\partial x} = \mu_0 \epsilon_0 \frac{\partial \vec{\mathbf{E}}_y}{\partial t} \quad \text{(ii)} \qquad (3.19)$$

$$\frac{\partial \vec{\mathbf{B}}_y}{\partial x} - \frac{\partial \vec{\mathbf{B}}_x}{\partial y} = \mu_0 \epsilon_0 \frac{\partial \vec{\mathbf{E}}_z}{\partial t} \quad \text{(iii)}$$

$$\frac{\partial \vec{\mathbf{B}}_x}{\partial x} + \frac{\partial \vec{\mathbf{B}}_y}{\partial y} + \frac{\partial \vec{\mathbf{B}}_z}{\partial z} = 0 \qquad (3.20)$$

$$\frac{\partial \vec{\mathbf{E}}_x}{\partial x} + \frac{\partial \vec{\mathbf{E}}_y}{\partial y} + \frac{\partial \vec{\mathbf{E}}_z}{\partial z} = 0 \qquad (3.21)$$

The transition has thus been made from the formulation of Maxwell's Equations in terms of integrals over *finite regions* to a restatement in terms of derivatives at *points* in space.

We now have all that is needed to comprehend the magnificent process whereby electric and magnetic fields, inseparably coupled and mutually sustaining, propagate out into space as a single entity, free of charges and currents, sans matter, sans aether.

3.2 Electromagnetic Waves

We have relegated to Appendix 1 a complete and mathematically elegant derivation of the electromagnetic wave equation. Here the focus is on the equally important task of developing a more intuitive appreciation of the physical processes involved. Three observations, from which we might build a qualitative picture, are readily available to us: the general perpendicularity of the fields, the symmetry of Maxwell's Equations, and the interdependence of $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ in those equations.

In studying electricity and magnetism, one soon becomes aware that a number of relationships are described by vector cross-products or, if you like, right-hand rules. In other words, an occurrence of one sort produces a related, perpendicularly directed response. Of immediate interest is the fact that a time-varying $\vec{\mathbf{E}}$-field generates a $\vec{\mathbf{B}}$-field, which is everywhere perpendicular to the direction in which $\vec{\mathbf{E}}$ changes (Fig. 3.12). In the same way, a time-varying $\vec{\mathbf{B}}$-field generates an $\vec{\mathbf{E}}$-field, which is everywhere perpendicular to the direction in which $\vec{\mathbf{B}}$ changes (Fig. 3.5). Consequently, we might anticipate the general transverse nature of the $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields in an electromagnetic disturbance.

Consider a charge that is somehow caused to *accelerate* from rest. When the charge is motionless, it has associated with it a constant radial $\vec{\mathbf{E}}$-field extending in all directions presumably to infinity. At the instant the charge begins to move, the $\vec{\mathbf{E}}$-field is altered in the vicinity of the charge, and this alteration propagates out into space at some finite speed. The time-varying electric field induces a magnetic field by means of Eq. (3.15) or (3.19). If the charge's velocity is constant, the rate-of-change of the $\vec{\mathbf{E}}$-field is steady, and the resulting $\vec{\mathbf{B}}$-field is constant. But here the charge is accelerating. $\partial \vec{\mathbf{E}}/\partial t$

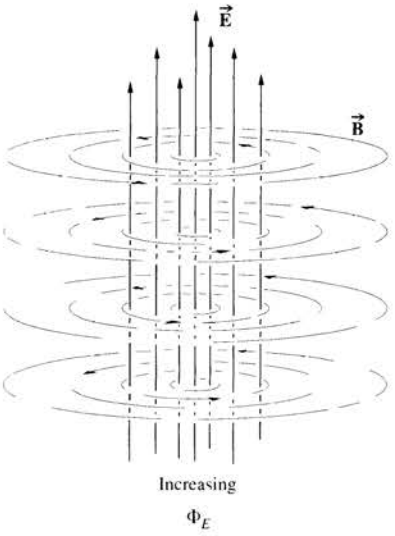


Figure 3.12 A time-varying $\vec{\mathbf{E}}$-field. Surrounding each point where Φ_E is changing, the $\vec{\mathbf{B}}$-field forms closed loops.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

is itself not constant, so the induced $\vec{\mathbf{B}}$-field is time-dependent. The time-varying $\vec{\mathbf{B}}$-field generates an $\vec{\mathbf{E}}$-field, (3.14) or (3.18), and the process continues, with $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ coupled in the form of a pulse. As one field changes, it generates a new field that extends a bit farther, and the pulse moves out from one point to the next through space.

We can draw an overly mechanistic but rather picturesque analogy, if we imagine the electric field lines as a dense radial distribution of strings (p. 59). When somehow plucked, each string is distorted, forming a kink that travels outward from the source. All these kinks combine at any instant to yield a three-dimensional expanding pulse in the continuum of the electric field.

The $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields can more appropriately be considered as two aspects of a single physical phenomenon, the *electromagnetic field*, whose source is a moving charge. The disturbance, once it has been generated in the electromagnetic field, is an untethered wave that moves beyond its source and independently of it. Bound together as a single entity, the time-varying electric and magnetic fields regenerate each other in an endless cycle. The electromagnetic waves reaching us from the relatively nearby Andromeda galaxy (which can be seen with the naked eye) have been on the wing for 2 200 000 years.

We have not yet considered the direction of wave propagation with respect to the constituent fields. Notice, however, that the high degree of symmetry in Maxwell's Equations for free space suggests that the disturbance will propagate in a direction that is symmetrical to both $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$. That implies that an electromagnetic wave cannot be purely longitudinal (i.e., as long as $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are not parallel). Let's now replace conjecture with a bit of calculation.

Appendix 1 shows that Maxwell's Equations for free space can be manipulated into the form of two extremely concise vector expressions:

$$\nabla^2\vec{\mathbf{E}} = \epsilon_0\mu_0\frac{\partial^2\vec{\mathbf{E}}}{\partial t^2} \qquad \text{[A1.26]}$$

and

$$\nabla^2\vec{\mathbf{B}} = \epsilon_0\mu_0\frac{\partial^2\vec{\mathbf{B}}}{\partial t^2} \qquad \text{[A1.27]}$$

The Laplacian,* ∇^2, operates on each component of $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$, so that the two vector equations actually represent a total of six scalar equations. In Cartesian coordinates,

$$\begin{aligned}
\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 E_x}{\partial t^2} \\
\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 E_y}{\partial t^2} \\
\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \frac{\partial^2 E_z}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 E_z}{\partial t^2}
\end{aligned} \qquad (3.22)$$

$$\begin{aligned}
\frac{\partial^2 B_x}{\partial x^2} + \frac{\partial^2 B_x}{\partial y^2} + \frac{\partial^2 B_x}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 B_x}{\partial t^2} \\
\frac{\partial^2 B_y}{\partial x^2} + \frac{\partial^2 B_y}{\partial y^2} + \frac{\partial^2 B_y}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 B_y}{\partial t^2} \\
\frac{\partial^2 B_z}{\partial x^2} + \frac{\partial^2 B_z}{\partial y^2} + \frac{\partial^2 B_z}{\partial z^2} &= \epsilon_0\mu_0\frac{\partial^2 B_z}{\partial t^2}
\end{aligned} \qquad (3.23)$$

Expressions of this sort, which relate the space and time variations of some physical quantity, had been studied long before Maxwell's work and were known to describe wave phenomena (p. 28). Each and every component of the electromagnetic field (E_x, E_y, E_z, B_x, B_y, B_z) obeys the scalar differential wave equation

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \qquad \text{[2.60]}$$

provided that

$$v = 1/\sqrt{\epsilon_0\mu_0} \qquad (3.24)$$

To evaluate v, Maxwell made use of the results of electrical experiments performed in 1856 in Leipzig by Wilhelm Weber (1804–1891) and Rudolph Kohlrausch (1809–1858). Equivalently, nowadays μ_0 is assigned a value of $4\pi \times 10^{-7}$ m·kg/C^2 in SI units, and until recently one might determine ϵ_0 directly from simple capacitor measurements. In any event, in modern units

$$\epsilon_0\mu_0 \approx (8.85 \times 10^{-12}\,\text{s}^2\cdot\text{C}^2/\text{m}^3\cdot\text{kg})(4\pi \times 10^{-7}\,\text{m}\cdot\text{kg}/\text{C}^2)$$

or

$$\epsilon_0\mu_0 \approx 11.12 \times 10^{-18}\,\text{s}^2/\text{m}^2$$

And now the moment of truth—in free space, the predicted

*In Cartesian coordinates,

$$\nabla^2\vec{\mathbf{E}} = \hat{\mathbf{i}}\nabla^2 E_x + \hat{\mathbf{j}}\nabla^2 E_y + \hat{\mathbf{k}}\nabla^2 E_z$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

speed of all electromagnetic waves would then be

$$v = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \approx 3 \times 10^8 \text{ m/s}$$

This theoretical value was in remarkable agreement with the previously measured speed of light (315 300 km/s) determined by Fizeau. The results of Fizeau's experiments, performed in 1849 with a rotating toothed wheel, were available to Maxwell and led him to comment:

> This velocity [i.e., his theoretical prediction] is so nearly that of light, that it seems we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws.

This brilliant analysis was one of the great intellectual triumphs of all time. It has become customary to designate the speed of light in vacuum by the symbol c, which comes from the Latin word *celer*, meaning fast. In 1983 the 17th Conférence Générale des Poids et Mesures in Paris adopted a new definition of the meter and thereby fixed the speed of light in vacuum as exactly

$$c = 2.997\,924\,58 \times 10^8 \text{ m/s}$$

3.2.1 Transverse Waves

The experimentally verified transverse character of light must now be explained within the context of electromagnetic theory. To that end, consider the fairly simple case of a plane wave propagating in vacuum in the positive x-direction. The electric field intensity is a solution of Eq. (A1.26), where $\vec{\mathbf{E}}$ is constant over each of an infinite set of planes perpendicular to the x-axis. It is therefore a function only of x and t; that is, $\vec{\mathbf{E}} = \vec{\mathbf{E}}(x, t)$. We now refer back to Maxwell's Equations, and in particular to Eq. (3.21), which is generally read as *the divergence of* $\vec{\mathbf{E}}$ *equals zero*. Since $\vec{\mathbf{E}}$ is not a function of either y or z, the equation can be reduced to

$$\frac{\partial E_x}{\partial x} = 0 \tag{3.25}$$

If E_x is not zero—that is, if there is some component of the field in the direction of propagation—this expression tells us that it does not vary with x. At any given time, E_x is constant for all values of x, but of course, this possibility cannot therefore correspond to a traveling wave advancing in the positive x-direction. Alternatively, it follows from Eq. (3.25) that for a wave, $E_x = 0$; the electromagnetic wave has no electric field component in the direction of propagation. The $\vec{\mathbf{E}}$-field associated with the plane wave is then exclusively *transverse*.

The fact that the $\vec{\mathbf{E}}$-field is transverse means that in order to completely specify the wave we will have to specify the moment-by-moment direction of $\vec{\mathbf{E}}$. Such a description corresponds to the **polarization** of the light, and it will be treated in Chapter 8. Without any loss of generality, we deal here with *plane* or *linearly polarized waves*, for which the direction of the vibrating $\vec{\mathbf{E}}$-vector is fixed. Thus we orient our coordinate axes so that the electric field is parallel to the y-axis, whereupon

$$\vec{\mathbf{E}} = \hat{\mathbf{j}} E_y(x, t) \tag{3.26}$$

Returning to Eq. (3.18), it follows that

$$\frac{\partial E_y}{\partial x} = -\frac{\partial B_z}{\partial t} \tag{3.27}$$

Therefore B_x and B_y are constant and of no interest at present. The time-dependent $\vec{\mathbf{B}}$-field can only have a component in the z-direction. Clearly then, ***in free space, the plane electromagnetic wave is transverse*** (Fig. 3.13). Except in the case of normal incidence, such waves propagating in real material media are generally not transverse—a complication arising from the fact that the medium may be dissipative or contain free charge.

We have not specified the form of the disturbance other than to say that it is a plane wave. Our conclusions are therefore quite general, applying equally well to both pulses and

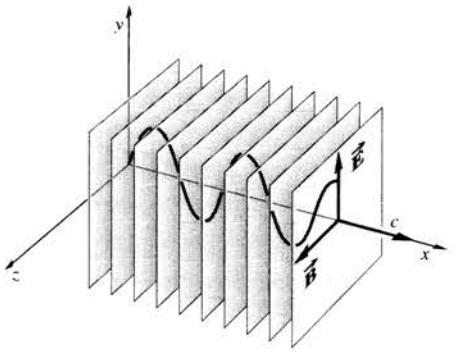


Figure 3.13 The field configuration in a plane harmonic electromagnetic wave.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

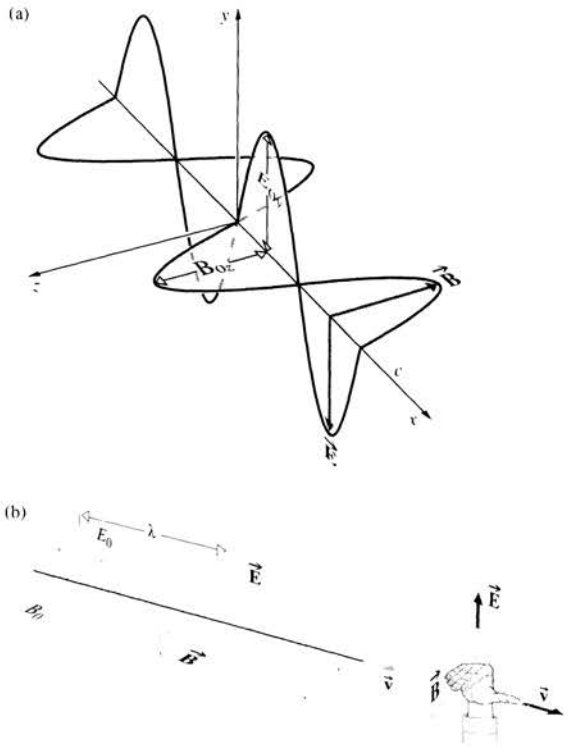


Figure 3.14 (*a*) Orthogonal harmonic $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields for a plane polarized wave. (*b*) The wave propagates in the direction of $\vec{\mathbf{E}} \times \vec{\mathbf{B}}$.

continuous waves. We have already pointed out that harmonic functions are of particular interest because any waveform can be expressed in terms of sinusoidal waves using Fourier techniques (p. 302). We therefore limit the discussion to harmonic waves and write $E_y(x, t)$ as

$$E_y(x, t) = E_{0y} \cos [\omega(t - x/c) + \varepsilon] \tag{3.28}$$

the speed of propagation being c. The associated magnetic flux density can be found by directly integrating Eq. (3.27), that is,

$$B_z = -\int \frac{\partial E_y}{\partial x} dt$$

Using Eq. (3.28), we obtain

$$B_z = -\frac{E_{0y}\omega}{c} \int \sin [\omega(t - x/c) + \varepsilon] dt$$

or

$$B_z(x, t) = \frac{1}{c} E_{0y} \cos [\omega(t - x/c) + \varepsilon] \tag{3.29}$$

The constant of integration, which represents a time-independent field, has been disregarded. Comparison of this result with Eq. (3.28) makes it evident that

$$E_y = cB_z \tag{3.30}$$

Since E_y and B_z differ only by a scalar, and so have the same time-dependence, $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are *in-phase* at all points in space. Moreover, $\vec{\mathbf{E}} = \hat{\mathbf{j}} E_y(x, t)$ and $\vec{\mathbf{B}} = \hat{\mathbf{k}} B_z(x, t)$ are *mutually perpendicular*, and their cross-product, $\vec{\mathbf{E}} \times \vec{\mathbf{B}}$, points in the propagation direction, $\hat{\mathbf{i}}$ (Fig. 3.14).

Plane waves, though important, are not the only solutions to Maxwell's Equations. As we saw in the previous chapter, the differential wave equation allows many solutions, among which are cylindrical and spherical waves (Fig. 3.15).

3.3 Energy and Momentum

One of the most significant properties of the electromagnetic wave is that it transports energy and momentum. The light from even the nearest star beyond the Sun travels 25 million million miles to reach the Earth, yet it still carries enough energy to do work on the electrons within your eye.

3.3.1 The Poynting Vector

Any electromagnetic wave exists within some region of space, and it is therefore natural to consider the *radiant energy per unit volume*, or **energy density**, u. We suppose that the electric field itself can somehow store energy. This is a major logical step since it imparts to the field the attribute of physical reality—if the field has energy it is a thing-in-itself. Moreover,

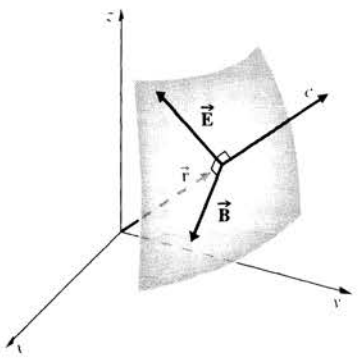


Figure 3.15 Portion of a spherical wavefront far from the source.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

inasmuch as the classical field is continuous, its energy is continuous. Let's assume as much and see where it leads. We can compute (Problem 3.6) the energy density of the $\vec{\mathbf{E}}$-field (e.g., between the plates of a charged capacitor) to be

$$u_E = \frac{\epsilon_0}{2} E^2 \tag{3.31}$$

Similarly, the energy density of the B-field alone (as it might be computed within a current-carrying toroid) is

$$u_B = \frac{1}{2\mu_0} B^2 \tag{3.32}$$

The relationship $E = cB$ was derived specifically for plane waves; nonetheless, it's quite general in its applicability. Using $c = 1/\sqrt{\epsilon_0 \mu_0}$, it follows that

$$u_E = u_B \tag{3.33}$$

The energy streaming through space in the form of an electromagnetic wave is shared equally between the constituent electric and magnetic fields. Inasmuch as

$$u = u_E + u_B \tag{3.34}$$

$$u = \epsilon_0 E^2 \tag{3.35}$$

or equivalently,

$$u = \frac{1}{\mu_0} B^2 \tag{3.36}$$

To represent the flow of electromagnetic energy associated with a traveling wave, let S symbolize the transport of energy per unit time (the power) across a unit area. In the SI system it has units of $\mathrm{W/m^2}$. Figure 3.16 depicts an electromagnetic wave traveling with a speed c through an area A. During a very small interval of time Δt, only the energy contained in the cylindrical volume, $u(c\,\Delta t\,A)$, will cross A. Thus

$$S = \frac{uc\,\Delta t\,A}{\Delta t\,A} = uc \tag{3.37}$$

or, using Eq. (3.35),

$$S = \frac{1}{\mu_0} EB \tag{3.38}$$

We now make the reasonable assumption (for isotropic media) that the energy flows in the direction of the propagation of the wave. The corresponding *vector* $\vec{\mathbf{S}}$ is then

$$\vec{\mathbf{S}} = \frac{1}{\mu_0} \vec{\mathbf{E}} \times \vec{\mathbf{B}} \tag{3.39}$$

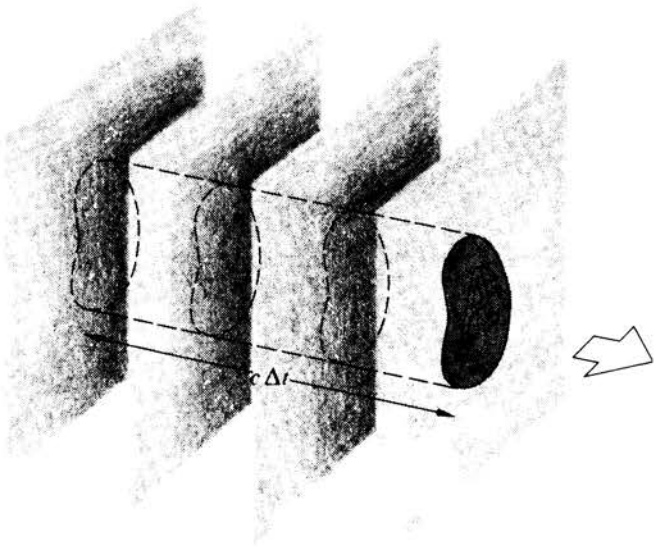


Figure 3.16 The flow of electromagnetic energy.

or

$$\vec{\mathbf{S}} = c^2 \epsilon_0 \, \vec{\mathbf{E}} \times \vec{\mathbf{B}} \tag{3.40}$$

The magnitude of $\vec{\mathbf{S}}$ is the power per unit area crossing a surface whose normal is parallel to $\vec{\mathbf{S}}$. Named after John Henry Poynting (1852–1914), it has come to be known as the **Poynting vector**. Let's now apply these considerations to the case of a harmonic, linearly polarized (the directions of the $\mathbf{E}$- and $\vec{\mathbf{B}}$-fields are fixed) plane wave traveling through free space in the direction of $\vec{\mathbf{k}}$:

$$\vec{\mathbf{E}} = \vec{\mathbf{E}}_0 \cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t) \tag{3.41}$$

$$\vec{\mathbf{B}} = \vec{\mathbf{B}}_0 \cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t) \tag{3.42}$$

Using Eq. (3.40), we find

$$\vec{\mathbf{S}} = c^2 \epsilon_0 \vec{\mathbf{E}}_0 \times \vec{\mathbf{B}}_0 \cos^2(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t) \tag{3.43}$$

This is the instantaneous flow of energy per unit area per unit time.

Averaging Harmonic Functions

It should be evident that $\vec{\mathbf{E}} \times \vec{\mathbf{B}}$ cycles from maxima to minima. At optical frequencies ($\approx 10^{15}$ Hz), $\vec{\mathbf{S}}$ is an extremely

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

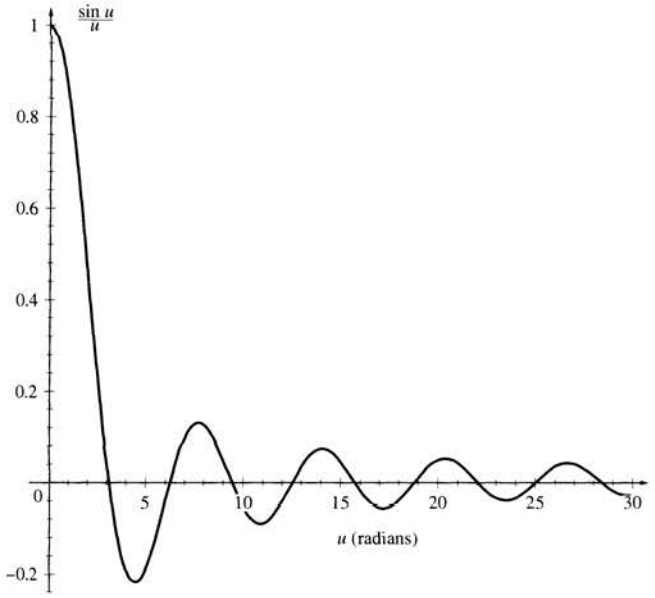


Figure 3.17 sinc u. Notice how the sinc function has a value of zero at $u = \pi$, 2π, 3π, and so forth.

rapidly varying function of time (indeed, twice as rapid as the fields, since cosine-squared has double the frequency of cosine). Therefore, its instantaneous value would be an impractical quantity to measure directly. This suggests that we employ an averaging procedure. That is, we absorb the radiant energy during some finite interval of time using, for example, a photocell, a film plate, or the retina of a human eye.

The specific form of Eq. (3.43), and the central role played by harmonic functions, suggest that we take a moment to study the average values of such functions. The time-averaged value of some function $f(t)$ over an interval T is written as $\langle f(t) \rangle_T$ and given by

$$\langle f(t) \rangle_T = \frac{1}{T} \int_{t-T/2}^{t+T/2} f(t)dt$$

The resulting value of $\langle f(t) \rangle_T$ very much depends on T. To find the average of a harmonic function, evaluate

$$\langle e^{i\omega t} \rangle_T = \frac{1}{T} \int_{t-T/2}^{t+T/2} e^{i\omega t} dt = \frac{1}{i\omega T} e^{i\omega t} \Big|_{t-T/2}^{t+T/2}$$

$$\langle e^{i\omega t} \rangle_T = \frac{1}{i\omega T} (e^{i\omega(t+T/2)} - e^{i\omega(t-T/2)})$$

and

$$\langle e^{i\omega t} \rangle_T = \frac{1}{i\omega T} e^{i\omega t} (e^{i\omega T/2} - e^{-i\omega T/2})$$

The parenthetical term should remind us (p. 22) of $\sin \omega T/2$. Hence

$$\langle e^{i\omega t} \rangle_T = \left(\frac{\sin \omega T/2}{\omega T/2} \right) e^{i\omega t}$$

The ratio in brackets is so common and important in Optics that it's given its own name; $(\sin u)/u$ is called (sinc u). Taking the real and imaginary parts of the above expression yields

$$\langle \cos \omega t \rangle_T = (\text{sinc } u) \cos \omega t$$

and

$$\langle \sin \omega t \rangle_T = (sinc\ u) \sin \omega t$$

The average of the cosine is itself a cosine, oscillating with the same frequency but having a sinc-function amplitude that drops off from its initial value of 1.0 very rapidly (Fig. 3.17 and Table 1 in the Appendix). Since sinc $u = 0$ at $u = \omega T/2 = \pi$, which happens when $T = \tau$, it follows that cos ωt averaged over an interval T equal to one period equals zero. Similarly, cos ωt averages to zero over any whole number of periods, as does sin ωt. That's reasonable in that each of these functions encompasses as much positive area above the axis as negative area below the axis, and that's what the defining integral corresponds to. After an interval of several periods, the sinc term will be so small that the fluctuations around zero will be negli-

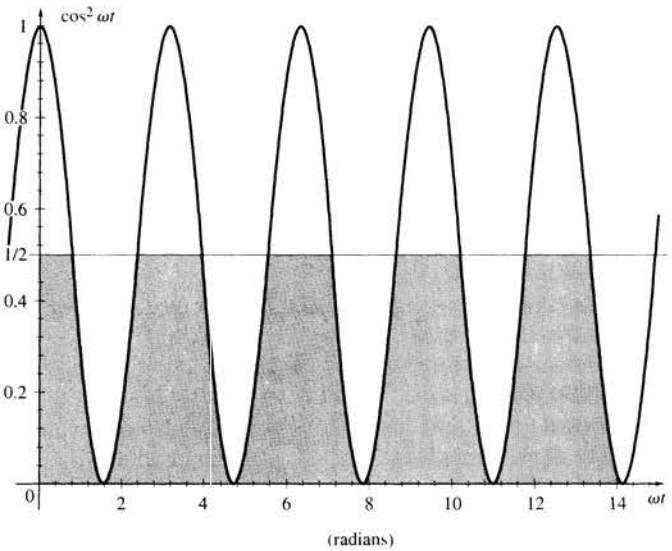


Figure 3.18 Using the peaks above the $\frac{1}{2}$ line to fill the troughs beneath it suggests that the average is $\frac{1}{2}$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

gible: $\langle \cos \omega t \rangle_T$ and $\langle \sin \omega t \rangle_T$ are then essentially zero.

It's left for Problem 3.8 to show that $\langle \cos^2 \omega t \rangle_T = \frac{1}{2}[1 + \text{sinc } \omega T \cos 2\omega t]$, which oscillates about a value of $1/2$ at a frequency of 2ω and rapidly approaches $1/2$ as T increases beyond a few dozen periods. In the case of light $\tau \approx 10^{-15}$ s and so averaging over even a microsecond corresponds to $T \approx 10^9\tau$, far more than enough to drive the sinc function to some totally negligible value, whereupon $\langle \cos^2 \omega t \rangle_T = 1/2$. Figure 3.18 suggests the same result; we chop off the humps above the $1/2$ line and use them to fill in the missing areas beneath the line. After enough cycles, the area under the $f(t)$ curve divided by T, which is $\langle f(t) \rangle_T$, approaches $1/2$.

3.3.2 Irradiance

When we talk about the "amount" of light illuminating a surface, we are referring to something called the **irradiance**,* denoted by I—the ***average energy per unit area per unit time***. Any kind of light-level detector has an entrance window that admits radiant energy through some fixed area A. The dependence on the size of that particular window is removed by dividing the total energy received by A. Furthermore, since the power arriving cannot be measured instantaneously, the detector must integrate the energy flux over some finite time, T. If the quantity to be measured is the *net* energy per unit area received, it depends on T and is therefore of limited utility. Someone else making a similar measurement under the same conditions can get a different result using a different T. If, however, the T is now divided out, a highly practical quantity results, one that corresponds to the average energy per unit area per unit time, namely, I.

The time-averaged value ($T >> \tau$) of the magnitude of the Poynting vector, symbolized by $\langle S \rangle_T$, is a measure of I. In the specific case of harmonic fields and Eq. (3.43),

$$\langle S \rangle_T = c^2 \epsilon_0 |\, \vec{\mathbf{E}}_0 \times \vec{\mathbf{B}}_0 \,| \langle \cos^2 (\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t) \rangle$$

Because $\langle \cos^2(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t) \rangle_T = \frac{1}{2}$ for $T >> \tau$ (see Problem 3.7)

$$\langle S \rangle_T = \frac{c^2 \epsilon_0}{2} |\, \vec{\mathbf{E}}_0 \times \vec{\mathbf{B}}_0 \,|$$

or

$$I \equiv \langle S \rangle_T = \frac{c \epsilon_0}{2} E_0^2 \tag{3.44}$$

*In the past physicists generally used the word *intensity* to mean the flow of *energy per unit area per unit time*. By international, if not universal, agreement, that term is slowly being replaced in Optics by the word *irradiance*.

The irradiance is proportional to the square of the amplitude of the electric field. Two alternative ways of saying the same thing are simply

$$I = \frac{c}{\mu_0} \langle B^2 \rangle_T \tag{3.45}$$

and

$$I = \epsilon_0 c \langle E^2 \rangle_T \tag{3.46}$$

Within a linear, homogeneous, isotropic dielectric, the expression for the irradiance becomes

$$I = \epsilon v \langle E^2 \rangle_T \tag{3.47}$$

Since, as we have learned, $\vec{\mathbf{E}}$ is considerably more effective at exerting forces and doing work on charges than is $\vec{\mathbf{B}}$, we shall refer to $\vec{\mathbf{E}}$ as the **optical field** and use Eq. (3.46) almost exclusively.

The time rate of flow of radiant energy is the **optical power** P or **radiant flux**, generally expressed in watts. If we divide the radiant flux incident on or exiting from a surface by the area of the surface, we have the **radiant flux density** (W m^2). In the former case, we speak of the *irradiance*, in the latter the *exitance*, and in either instance the **flux density**. The irradiance is a measure of the *concentration* of power.

The Inverse Square Law

We saw earlier that the spherical-wave solution of the differential wave equation has an amplitude that varies inversely with r. Let's now examine this same feature within the context of energy conservation. Consider an isotropic point source in free space, emitting energy equally in all directions (i.e., emitting spherical waves). Surround the source with two concentric imaginary spherical surfaces of radii r_1 and r_2, as shown in Fig. 3.19. Let $E_0(r_1)$ and $E_0(r_2)$ represent the amplitudes of the waves over the first and second surfaces, respectively. If energy is to be conserved, the total amount of energy flowing through each surface per second must be equal, since there are no other sources or sinks present. Multiplying I by the surface area and taking the square root, we get

$$r_1 E_0(r_1) = r_2 E_0(r_2)$$

Inasmuch as r_1 and r_2 are arbitrary, it follows that

$$r E_0(r) = \text{constant},$$

and the amplitude must drop off inversely with r. The irradiance from a point source is proportional to $1/r^2$. This is the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

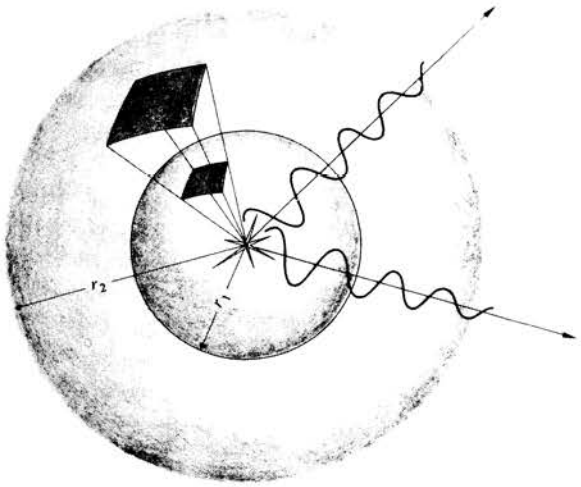


Figure 3.19 The geometry of the Inverse Square Law.

well-known **Inverse Square Law**, which is easily verified with a point source and a photographic exposure meter.

3.3.3 Photons

Light is absorbed and emitted in tiny discrete bursts, in particles of electromagnetic "stuff," known as photons. That much has been confirmed and is well established, but the question of whether or not light is "really" a stream of photons is far from settled, and that issue will be revisited several times to come.*

Ordinarily, a light beam delivers so many minute energy quanta that its inherent granular nature is totally hidden and a continuous phenomenon is observed macroscopically. That sort of thing is commonplace in Nature; the forces exerted by the individual gas molecules in a wind blend into what seems a continuous pressure, but it obviously isn't. Indeed, that analogy between a gas and a flow of photons is one we will come back to presently.

*Thus whenever, for brevity's sake, an expression such as "a beam of photons" is used, the reader is advised to keep in mind that although the light-*is*-corpuscular model has wide acceptance (especially in high-energy physics) and is a crucial part of the contemporary discourse, like almost everything else in physics, it is not yet established beyond all doubt. See, for example, the summary article by R. Kidd, J. Ardini, and A. Anton, "Evolution of the modern photon," *Am. J. Phys.* **57** (1), 27 (1989).

As the great French physicist Louis de Broglie put it, "Light is, in short, the most refined form of matter," and all matter, including light, is quantized. At base it comes in minute elementary units—quarks, leptons, and photons. That overarching unity is among the most appealing reasons to embrace the photon as particle (whatever a "particle" is).

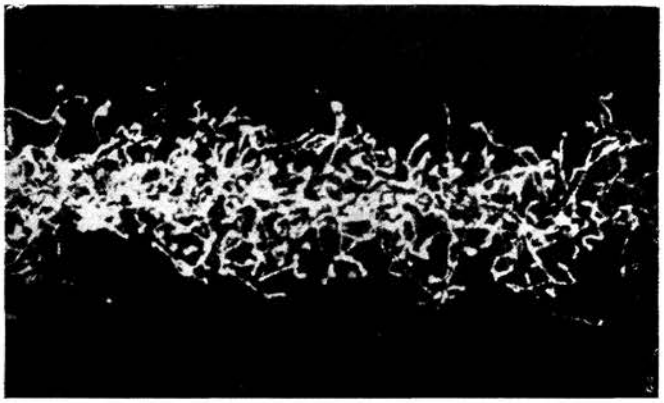
The Failure of Classical Theory

In 1900 Max Planck produced a rather tentative, and somewhat erroneous, analysis of a process known as *blackbody radiation* (p. 581). Nonetheless, the expression he came up with beautifully fit all the existing experimental data, a feat no other formulation had even come close to. Basically, he considered electromagnetic (EM) waves in equilibrium within an isothermal chamber (or cavity). All the EM-radiation within the cavity is emitted and absorbed by the walls of the enclosure—none enters from outside. This ensures that its spectral composition will match that emanating from an ideal black surface. The goal was to predict the spectrum of the radiation that would emerge from a small opening in the cavity. Totally stymied by the problem, as a last resort, Planck turned to the classical statistical analysis of Maxwell and Boltzmann, which was developed as the basis of the kinetic theory of gases. Philosophically, this is a completely deterministic treatment that assumes one can follow, at least in principle, every atom as it moves around in the system. Consequently, each atom is taken to be recognizable, independent, and enumerable. For purely computational reasons, Planck hypothesized that each one of the oscillators lining the walls of the chamber could absorb and emit only discrete amounts of energy proportional to its oscillatory frequency, ν. These energy jolts were equal to whole number multiples of $h\nu$, where h, now called **Planck's Constant**, was found to be 6.626×10^{-34} J·s. Being a rather traditional man, Planck otherwise held fast to the classical wave picture of light, insisting that only the oscillators were quantized.

Prophetically, J. J. Thomson (1903)—the discoverer of the electron—extended the idea, suggesting that electromagnetic waves might actually be radically different from other waves; perhaps local concentrations of radiant energy truly existed. Thomson had observed that when a beam of high-frequency EM-radiation (X-rays) was shone onto a gas, only certain of the atoms, here and there, were ionized. It was as if the beam had "hot spots" rather than having its energy distributed continuously over the wavefront (see photo at the top of page 52).

The concept of the photon in its modern incarnation came

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



A beam of X-rays enters a cloud chamber on the left. The tracks are made by electrons emitted via either the Photoelectric Effect (these tend to leave long tracks at large angles to the beam) or the Compton Effect (short tracks more in the forward direction). Although classically the X-ray beam has its energy uniformly distributed along transverse wavefronts, the scattering seems discrete and random. (From the Smithsonian Report, 1915.)

into being in 1905 by way of Einstein's brilliant theoretical work on the Photoelectric Effect. When a metal is bathed in EM-radiation, it emits electrons. The details of that process had been studied experimentally for decades, but it defied analysis via classical Electromagnetic Theory. Einstein's startling treatment established that ***the electromagnetic field itself is quantized***. Each constituent photon has an energy given by the product of Planck's Constant and the frequency of the radiation field:

$$\mathcal{E} = h\nu \tag{3.48}$$

Photons are stable, chargeless, massless elementary particles that exist only at the speed c.

In 1924 Satyendra N. Bose formulated a new and rigorous proof of Planck's blackbody equation using statistical methods applied to light quanta. The cavity was envisioned to be filled with a "gas" of photons, ***which were taken to be totally indistinguishable, one from the other***. That was a crucial feature of this quantum-mechanical treatment. It meant that the microparticles were completely interchangeable, and this had a profound effect on the statistical formulation. In a mathematical sense, each particle of this quantum "gas" is related to every other particle, and no one of them can be taken as statistically independent of the system as a whole. That's very different from the independent way classical microparticles behave in an ordinary gas. The quantum-mechanical probability function that describes the statistical behavior of thermal light is now known as the Bose-Einstein distribution. The photon, whatever it is, became an indispensable tool of theoretical physics.

Unlike ordinary objects, photons cannot be seen directly; what is known of them comes from observing the results of their being either created or annihilated. Light is never *seen* just sailing along through space. A photon is observed by detecting the effect it has on its surroundings, and it only has an observable effect when it either comes into, or goes out of, existence. Photons begin and end on charged particles; most often they are emitted from and absorbed by electrons. And these are usually the electrons circulating in the clouds around atoms. A number of experiments have directly confirmed the quantal nature of the emission process. For example, imagine a very dim source surrounded, at equal distances, by identical photodetectors each capable of measuring a minute amount of light. If the emission, *no matter how faint*, is a continuous wave, as is maintained classically, all the detectors should register each emitted pulse in coincidence. That does not happen; instead, counts are registered by detectors independently, one at a time, in clear agreement with the idea that atoms emit localized light quanta in random directions.

Furthermore, it has been confirmed that when an atom emits light (i.e., a photon), it recoils in the opposite direction, just as a pistol recoils when it fires a bullet. In Fig. 3.20 atoms pumped up with excess energy (i.e., excited, p. 63) are formed into a narrow beam. They soon spontaneously radiate photons in random directions and are themselves kicked backward, often laterally away from the beam. The resulting spread of

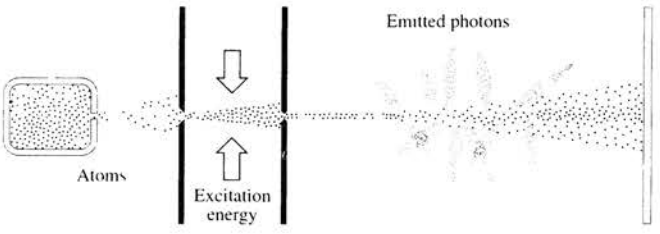


Figure 3.20 When so-called excited atoms forming a narrow beam radiate photons, they recoil laterally and the beam spreads out. Alternatively if the beam is formed of atoms that are not excited (i.e., they are in their ground states), it remains narrow all the way to the screen.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

the beam is a quantum-mechanical effect inconsistent with the classical picture of the emission of a symmetrical wave.

A Barrage of Photons

When we analyze phenomena involving the activity of immense numbers of participants, the use of statistical techniques is often the only practical way to proceed. In addition to the classical Maxwell-Boltzmann statistics (for distinguishable particles), there are two kinds of quantum-mechanical statistics (for indistinguishable particles): Bose-Einstein and Fermi-Dirac. The first treats particles that *are not* subject to the Pauli Exclusion Principle (i.e., particles that have zero or integer spins). Fermi-Dirac statistics treats particles that *are* subject to the Pauli Exclusion Principle (i.e., those that have spins that are odd integer multiples of $\frac{1}{2}$). Photons are called **bosons**, they are spin-1 particles, and the manner in which they group together obeys Bose-Einstein statistics. Similarly, electrons are **fermions**; they are spin-$\frac{1}{2}$ particles that obey Fermi-Dirac statistics.

Microparticles have defining physical characteristics such as mass, charge, and spin—characteristics that do not change. When these are given, we have completely specified the kind of particle being considered. Alternatively, there are alterable properties of any given microparticle that describe its momentary condition such as energy, momentum, and spin orientation. When all of these alterable quantities are given, we have specified the particular **state** the particle happens to be in at the moment.

Fermions are committed loners; *only one fermion can occupy any given state*. By comparison, bosons are gregarious; *any number of them can occupy the same state, and moreover, they actually tend to cluster close together*. ***When a very large number of photons occupy the same state, the inherent granularity of the light beam essentially vanishes and the electromagnetic field appears as the continuous medium of an electromagnetic wave***. Thus we can associate a monochromatic (monoenergetic) plane wave with a stream of photons having a high population density, all progressing in the same state (with the same energy, same frequency, same momentum, same direction). ***Different monochromatic plane waves represent different photon states***.

Unlike the photon, because electrons are fermions, large numbers of them cannot cluster tightly in the same state, and a monoenergetic beam of electrons does not manifest itself on a macroscopic scale as a continuous classical wave. In that regard, EM-radiation is quite distinctive.

For a uniform monochromatic light beam of frequency ν, the quantity $I/h\nu$ is the average number of photons impinging on a unit area (normal to the beam) per unit time, namely, the photon flux density. More realistically, if the beam is quasi-monochromatic (p. 17) with an average frequency ν_0, its **mean photon flux density** is $I/h\nu_0$. Given that an incident quasi-monochromatic beam has a cross-sectional area A, its **mean photon flux** is

$$\Phi = AI/h\nu_0 = P/h\nu_0 \tag{3.49}$$

where P is the **optical power** of the beam in watts. The mean photon flux is the average number of photons arriving per unit of time (Table 3.1). For example, a small 1.0 mW He-Ne laserbeam with a mean wavelength of 632.8 nm delivers a mean photon flux of $P/h\nu_0 = (1.0 \times 10^{-3}\ \text{W})/[(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})(2.998 \times 10^{8}\ \text{m/s})/(632.8 \times 10^{-9}\ \text{m})] = 3.2 \times 10^{15}$ photons per second.

Imagine a uniform beam of light having a constant irradiance (and therefore a constant mean photon flux) incident on a screen. The energy of the beam is deposited on the screen in a random flurry of minute bursts. Individually, the incoming photons register at locations on the plane that are totally unpredictable, and arrive at moments in time that are equally unpredictable. It looks as if the beam is composed of a random stream of photons, but that conclusion, however tempting, goes beyond the observation. What *can* be said is that the light delivers its energy in a staccato of impacts that are random in space and time across the beam.

Suppose that we project a light pattern onto the screen; it might be a set of interference fringes or the image of a woman's face. The barrage of photons forming the image is a

TABLE 3.1 The Mean Photon Flux Density for a Sampling of Common Sources

Light Source	Mean Photon Flux Density Φ/A in units of (photons/s·m^2)
Laserbeam (10 mW, He-Ne, focused to 20 μm)	10^{26}
Laserbeam (1 mW, He-Ne)	10^{21}
Bright sunlight	10^{18}
Indoor light level	10^{16}
Twilight	10^{14}
Moonlight	10^{12}
Starlight	10^{10}

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

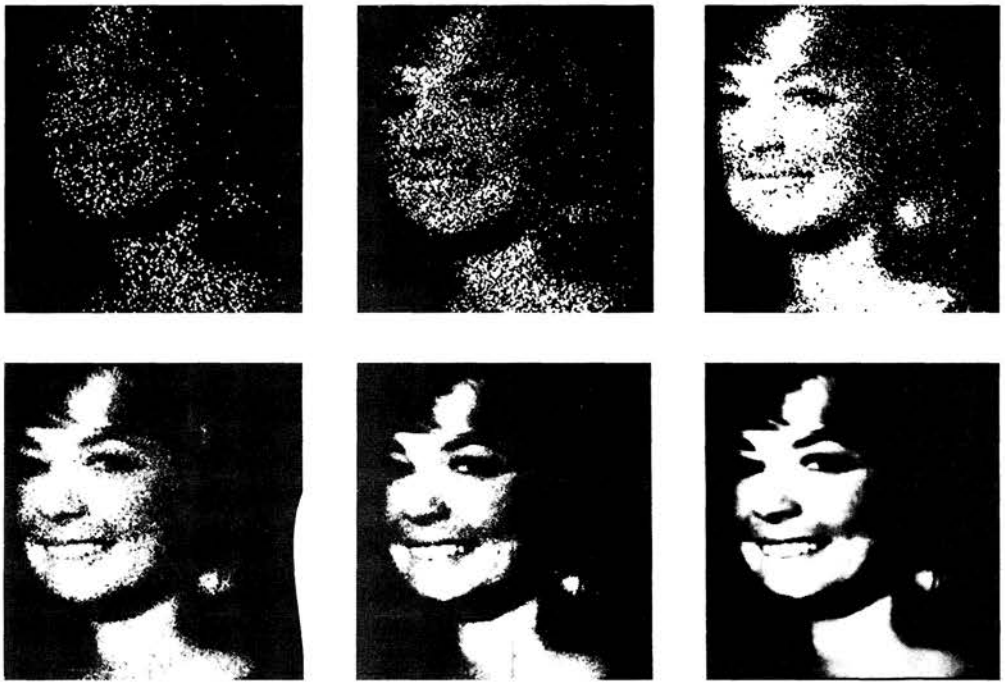


Figure 3.21 These photographs (which were electronically enhanced) are a compelling illustration of the granularity displayed by light in its interaction with matter. Under exceedingly faint illumination, the pattern (each spot corresponding to one photon) seems almost random, but as the light level increases the quantal character of the process gradually becomes obscured. (See *Advances in Biological and Medical Physics* V, 1957, 211–242.) (Photos courtesy Radio Corporation of America.)

statistical tumult; we cannot predict when a photon will arrive at any given location. But we can determine the likelihood of one or more photons hitting any particular point during a substantial time interval. ***At any location on the screen, the measured (or classically computed) value of the irradiance is proportional to the probability of detecting a photon at that location*** (p. 137).

Figure 1.8, which is a pictorial record of the arrival of individual photons, was produced using a special kind of photomultiplier tube. To underscore the inherent photonic nature of radiant energy, let's now use an entirely different and more straightforward photographic approach to record the incidence of light. A photographic emulsion contains a distribution of microscopic ($\approx 10^{-6}$ m) silver halide crystals, each comprising approximately 10^{10} Ag atoms. A single photon can interact with such a crystal, disrupting a silver-halogen bond and freeing up an Ag atom. One or more silver atoms then serve as a development center on the exposed crystal. The film is developed using a chemical reducing agent. It dissolves each exposed crystal, depositing at that site all of its Ag atoms as a single clump of the metal.

Figure 3.21 shows a series of photographs taken with increasing amounts of illumination. Using extremely dim light, a few thousand photons, the first picture is composed of roughly as many silver clumps, making a pattern that only begins to suggest an overall image. As the number of participating photons goes up (roughly by a factor of 10 for each successive picture), the image becomes increasingly smooth and recognizable. When there are tens of millions of photons forming the image, the statistical nature of the process is lost and the picture assumes a familiar continuous appearance.

Photon Counting

What, if anything, can be said about the statistical nature of the barrage of photons delivered as a beam of light? To answer that question, researchers within the last 35 years or so have conducted experiments in which they literally counted individual photons. What they found was that the pattern of arrival of photons was characteristic of the type of source.* We cannot go into the theoretical details here, but it is informative at least to look at the results for the two extreme cases of what is often called *coherent* and *chaotic* light.

*See P. Koczyk, P. Wiewior, and C. Radzewicz, "Photon counting statistics—Undergraduate experiment," *Am. J. Phys.* **64** (3), 240 (1996) and A. C. Funk and M. Beck, "Sub-Poissonian photocurrent statistics: Theory and undergraduate experiment," *Am. J. Phys.* **65** (6), 492 (1997).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

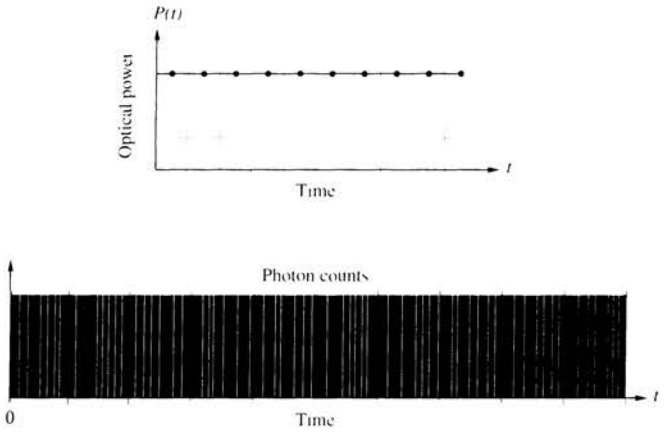


Figure 3.22 A constant optical power and the corresponding random set of photon counts. The arrival of each photon is an independent event.

Consider an ideal continuous laserbeam of *constant irradiance*; remember that irradiance is a time-averaged quantity via Eq. (3.46). The beam has a constant optical power P—which is also a time-averaged quantity—and, from Eq. (3.49), a corresponding mean photon flux Φ. Figure 3.22 depicts the random arrival of photons on a time scale that is short compared to the interval over which the irradiance is averaged. Thus it is possible for the macroscopic quantity P to be measured to be constant, even though there is an underlying discontinuous transfer of energy.

Now pass the beam through a shutter that stays open for a short sampling time T (which might be in the range from about 10 μs to perhaps 10 ms), and count the number of photons arriving at a photodetector during that interval. After a brief pause repeat the procedure, and do it again and again, tens of thousands of times. The results are presented in a histogram (Fig. 3.23), where the number of trials in which N photons were counted is plotted against N. Few trials register either very few photons or very many photons. On average, the number of photons per trial is $N_{av} = \Phi T = PT\ h\nu_0$. The shape of the data plot, which can be derived using probability theory, closely approximates the well-known *Poisson distribution*. It represents a graph of the probability that the detector (during a trial interval lasting a time T) will record zero photons, one photon, two photons, and so forth.

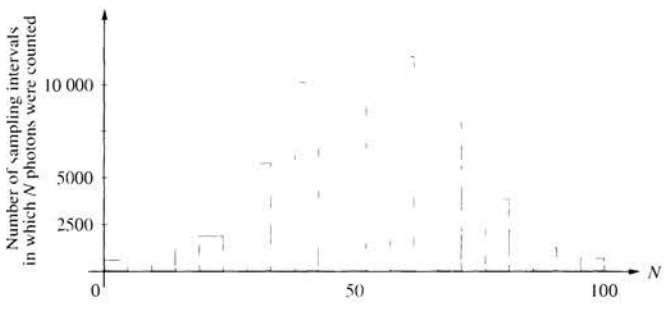


Figure 3.23 A typical histogram showing the probability or photon-count distribution for a beam of constant irradiance.

The Poisson distribution is the same symmetrical curve one gets when counting either the number of particles randomly emitted by a long-lived radioactive sample, or the number of raindrops randomly descending on an area in a steady shower. It's also the curve of the probability of getting a head, plotted against the number of heads occurring (N), for a coin tossed more than about 20 times. Thus with $N_{max} = 20$ the highest probability occurs near the average value N_{av}, namely, at $\frac{1}{2}N_{max}$ or 10 and the lowest at $N = 0$ and $N = 20$. The most probable value will be 10 heads out of 20 tosses, and the likelihood of getting either no heads or all heads is vanishingly small. It would seem that however an ideal laser produces light, it generates a stream of photons whose individual arrival is random and statistically independent. For reasons that will be explored later, an ideal monoenergetic beam—a monochromatic plane wave—is the epitome of what is known as **coherent light**.

Not surprisingly, ***the statistical distribution of the number of photons arriving at a detector depends on the nature of the source***; it is fundamentally different for an ideal *coherent source* at one extreme, as compared to an equally idealized completely incoherent or *chaotic source* at the other extreme. A stabilized laser resembles a coherent source, and an ordinary thermal source such as a lightbulb or a star or a gas discharge lamp more closely resembles a chaotic source. In the case of ordinary light, there are inherent fluctuations in the irradiance and therefore in the optical power (p. 48). These fluctuations are correlated, and the associated number of emitted photons, though random in time, is correspondingly also correlated (Fig. 3.24). The greater the optical power, the greater the number density of photons. Because the arrival of photons at the detector is not a succession of independent events, Bose-Einstein statistics apply (Fig. 3.25). Here the most likely number of counts per interval is zero, whereas, ideally, for laser light the most likely number of photons to be measured during a sampling interval equals the average number recorded. Thus even if a beam of laser light and a beam of ordinary light have the same average irradiance and the same frequency spectrum, they are still inherently distinguishable—a result that extends beyond classical theory.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

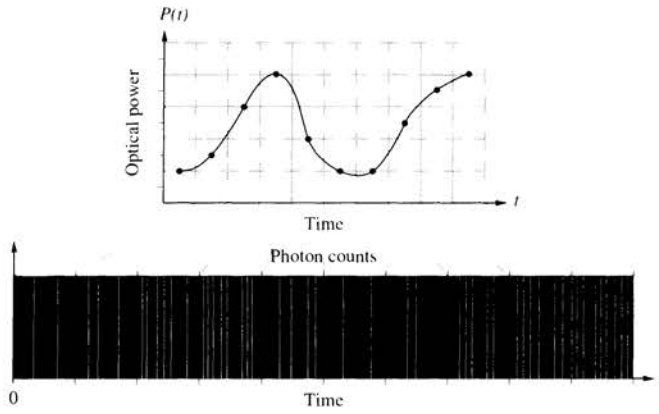


Figure 3.24 A time-varying optical power and the corresponding random set of photon counts. Now there are fluctuations that are correlated, and the photon arrivals are no longer independent.

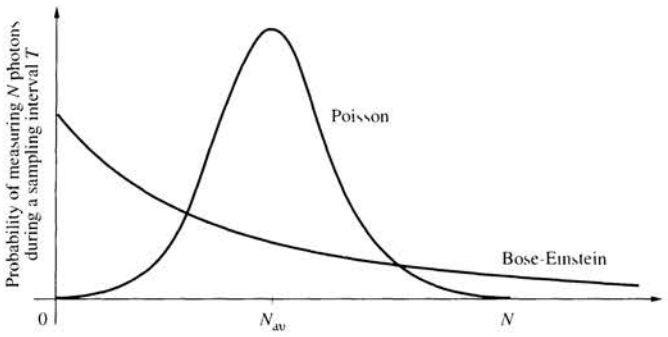


Figure 3.25 Poisson and Bose-Einstein photon counting distributions.

3.3.4 Radiation Pressure and Momentum

As long ago as 1619, Johannes Kepler proposed that it was the pressure of sunlight that blew back a comet's tail so that it always pointed away from the Sun. That argument particularly appealed to the later proponents of the corpuscular theory of light. After all, they envisioned a beam of light as a stream of particles, and such a stream would obviously exert a force as it bombarded matter. For a while it seemed as though this effect might at last establish the superiority of the corpuscular over the wave theory, but all the experimental efforts to that end failed to detect the force of radiation, and interest slowly waned.

Ironically, it was Maxwell in 1873 who revived the subject by establishing theoretically that waves do indeed exert pressure. "In a medium in which waves are propagated," wrote Maxwell, "there is a pressure in the direction normal to the waves, and numerically equal to the energy in a unit of volume."

When an electromagnetic wave impinges on some material surface, it interacts with the charges that constitute bulk matter. Regardless of whether the wave is partially absorbed or reflected, it exerts a force on those charges and hence on the surface itself. For example, in the case of a good conductor, the wave's electric field generates currents, and its magnetic field generates forces on those currents.

It's possible to compute the resulting force via Electromagnetic Theory, whereupon Newton's Second Law (which maintains that force equals the time rate-of-change of momentum) suggests that the *wave itself carries momentum*. Indeed, whenever we have a flow of energy, it's reasonable to expect that there will be an associated momentum—the two are the related time and space aspects of motion.

As Maxwell showed, the **radiation pressure**, $\mathcal{P}$, equals the energy density of the electromagnetic wave. From Eqs. (3.31) and (3.32), for a vacuum, we know that

$$u_E = \frac{\epsilon_0}{2} E^2 \quad \text{and} \quad u_B = \frac{1}{2\mu_0} B^2$$

Since $\mathcal{P} = u = u_E + u_B$,

$$\mathcal{P} = \frac{\epsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2$$

Alternatively, using Eq. (3.37), we can express the pressure in terms of the magnitude of the Poynting vector, namely,

$$\mathcal{P}(t) = \frac{S(t)}{c} \tag{3.50}$$

Notice that this equation has the units of power divided by area, divided by speed—or equivalently, force times speed divided by area and speed, or just force over area. **This is the instantaneous pressure that would be exerted on a perfectly absorbing surface by a normally incident beam.**

Inasmuch as the $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields are rapidly varying, $S(t)$ is rapidly varying, so it is eminently practical to deal with the average radiation pressure, namely,

$$\langle \mathcal{P}(t) \rangle_{\mathrm{T}} = \frac{\langle S(t) \rangle_{\mathrm{T}}}{c} = \frac{I}{c} \tag{3.51}$$

expressed in newtons per square meter. This same pressure is exerted on a source that itself is radiating energy.

Referring back to Fig. 3.16, if p is momentum, the force

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

exerted by the beam on an absorbing surface is

$$A\mathcal{P} = \frac{\Delta p}{\Delta t} \tag{3.52}$$

If p_V is the *momentum per unit volume of the radiation*, then an amount of momentum $\Delta p = p_V(c\,\Delta t\,A)$ is transported to A during each time interval Δt, and

$$A\mathcal{P} = \frac{p_V(c\,\Delta t\,A)}{\Delta t} = A\,\frac{S}{c}$$

Hence the volume density of electromagnetic momentum is

$$p_V = \frac{S}{c^2} \tag{3.53}$$

When the surface under illumination is **perfectly reflecting**, the beam that entered with a velocity $+c$ will emerge with a velocity $-c$. This corresponds to twice the change in momentum that occurs on absorption, and hence

$$\langle \mathcal{P}(t) \rangle_T = 2\,\frac{\langle S(t) \rangle_T}{c}$$

Notice, from Eqs. (3.50) and (3.52), that if some amount of energy $\mathcal{E}$ is transported per square meter per second, then there will be a corresponding momentum $\mathcal{E}/c$ transported per square meter per second.

In the photon picture, each quantum, has an energy $\mathcal{E} = h\nu$. We can then expect a photon to carry a momentum

$$p = \frac{\mathcal{E}}{c} = \frac{h}{\lambda} \tag{3.54}$$

Its vector momentum would be

$$\vec{\mathbf{p}} = \hbar\vec{\mathbf{k}}$$

where $\vec{\mathbf{k}}$ is the propagation vector and $\hbar \equiv h/2\pi$. This all fits in rather nicely with Special Relativity, which relates the mass m, energy, and momentum of a particle by

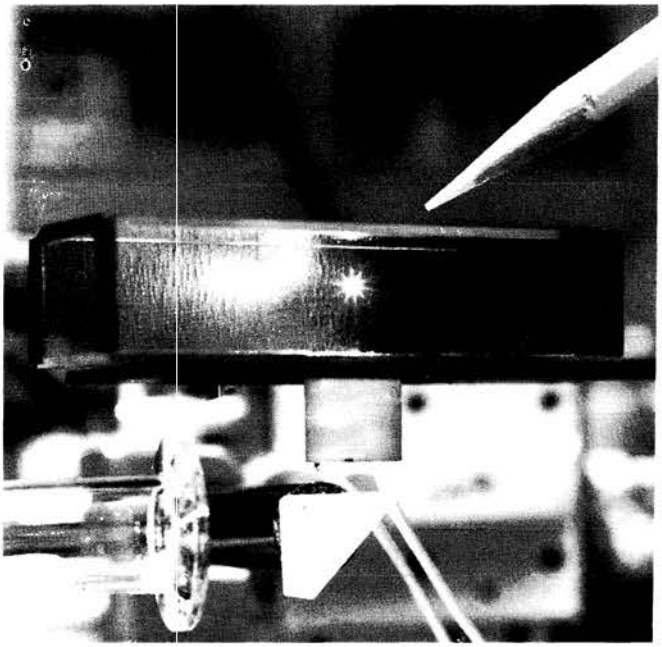
$$\mathcal{E} = [(cp)^2 + (mc^2)^2]^{1/2}$$

For a photon $m = 0$ and $\mathcal{E} = cp$.

These quantum-mechanical ideas have been confirmed experimentally utilizing the Compton Effect, which detects the energy and momentum transferred to an electron upon interaction with an individual X-ray photon (see photo on page 52).

The average flux density of electromagnetic energy from the Sun impinging normally on a surface just outside the Earth's atmosphere is about 1400 W/m^2. Assuming complete absorption, the resulting pressure would be 4.7×10^{-6} N/m^2, or 1.8×10^{-9} ounce/cm^2, as compared with, say, atmospheric pressure of about 10^5 N/m^2. The pressure of solar radiation at the Earth is tiny, but it is still responsible for a substantial planetwide force of roughly 10 tons. Even at the very surface of the Sun, radiation pressure is relatively small (see Problem 3.31). As one might expect, it becomes appreciable within the blazing body of a large bright star, where it plays a significant part in supporting the star against gravity. Despite the Sun's modest flux density, it nonetheless can produce appreciable effects over long-acting times. For example, had the pressure of sunlight exerted on the Viking spacecraft during its journey been neglected, it would have missed Mars by about 15 000 km. Calculations show that it is even feasible to use the pressure of sunlight to propel a space vehicle among the inner planets.* Ships with immense reflecting sails driven by solar radiation pressure may some day ply the dark sea of local space.

*The charged-particle flux called the "solar wind" is 1000 to 100 000 times less effective in providing a propulsive force than is sunlight.



The tiny starlike speck is a minute (one-thousandth of an inch diameter) transparent glass sphere suspended in midair on an upward 250-mW laserbeam. (Photo courtesy Bell Laboratories.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

The pressure exerted by light was actually measured as long ago as 1901 by the Russian experimenter Pyotr Nikolaievich Lebedev (1866–1912) and independently by the Americans Ernest Fox Nichols (1869–1924) and Gordon Ferrie Hull (1870–1956). Their accomplishments were formidable, considering the light sources available at the time. Nowadays, with the advent of the laser, light can be focused down to a spot size approaching the theoretical limit of about one wavelength in radius. The resulting irradiance, and therefore the pressure, is appreciable, even with a laser rated at just a few watts. It has thus become practical to consider radiation pressure for all sorts of applications, such as separating isotopes, accelerating particles, cooling and trapping atoms (p. 65), and even optically levitating small objects.

Light can also transport angular momentum, but that raises a number of issues that will be treated later (p. 330).

3.4 Radiation

Electromagnetic radiation comes in a broad range of wavelengths and frequencies, although in vacuum it all travels at the same speed. Despite the fact that we distinguish different regions of the spectrum with names like radiowaves, microwaves, infrared, and so forth, there is only one entity, one essence of electromagnetic wave. Maxwell's Equations are independent of wavelength and so suggest no fundamental differences in kind. Accordingly, it is reasonable to look for a common source-mechanism for all EM-radiation. What we find is that the various types of radiant energy seem to have a common origin in that they are all associated with *nonuniformly moving charges*. We are, of course, dealing with waves in the electromagnetic field, and charge *is* that which gives rise to field, so this is not altogether surprising.

A stationary charge has a constant $\vec{\mathbf{E}}$-field, no $\vec{\mathbf{B}}$-field, and hence produces no radiation—where would the energy come from if it did? A uniformly moving charge has both an $\vec{\mathbf{E}}$- and a $\vec{\mathbf{B}}$-field, but it does not radiate. If you traveled along with the charge, the current would thereupon vanish, hence $\vec{\mathbf{B}}$ would vanish, and we would be back at the previous case, uniform motion being relative. That's reasonable, since it would make no sense at all if the charge stopped radiating just because you started walking along next to it. That leaves *nonuniformly moving charges*, which assuredly do radiate. In the photon picture this is underscored by the conviction that the fundamental interactions between substantial matter and radiant energy are between photons and charges.

We know in general that free charges (those not bound within an atom) emit electromagnetic radiation when accelerated. That much is true for charges changing speed along a straight line within a linear accelerator, sailing around in circles inside a cyclotron, or simply oscillating back and forth in a radio antenna—***if a charge moves nonuniformly, it radiates***. A free charged particle can spontaneously absorb or emit a photon, and an increasing number of important devices, ranging from the free-electron laser (1977) to the synchrotron radiation generator, utilize this mechanism on a practical level.

3.4.1 Linearly Accelerating Charges

Consider a charge moving at a constant speed. It essentially has attached to it an unchanging radial electric field and a surrounding circular magnetic field. Although at any point in space the $\vec{\mathbf{E}}$-field changes from moment to moment, at any instant its value can be determined by supposing that the field lines move along, fixed to the charge. Thus the field does not disengage from the charge, and there is no radiation.

The electric field of a charge at rest can be represented, as in Fig. 3.26, by a uniform, radial distribution of straight field lines. For a charge moving at a constant velocity $\vec{\mathbf{v}}$, the field lines are still radial and straight, but they are no longer uniformly distributed. The nonuniformity becomes evident at high speeds and is usually negligible when $v << c$.

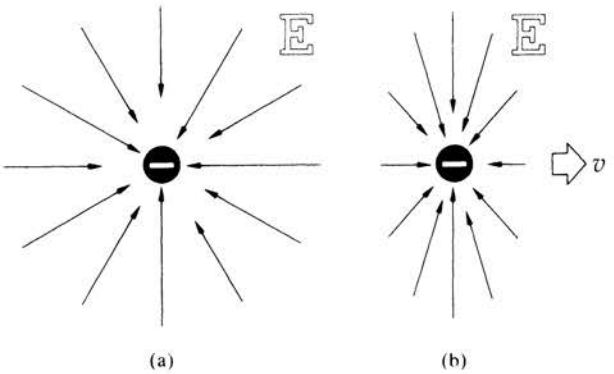


Figure 3.26 (*a*) Electric field of a stationary electron. (*b*) Electric field of a moving electron.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

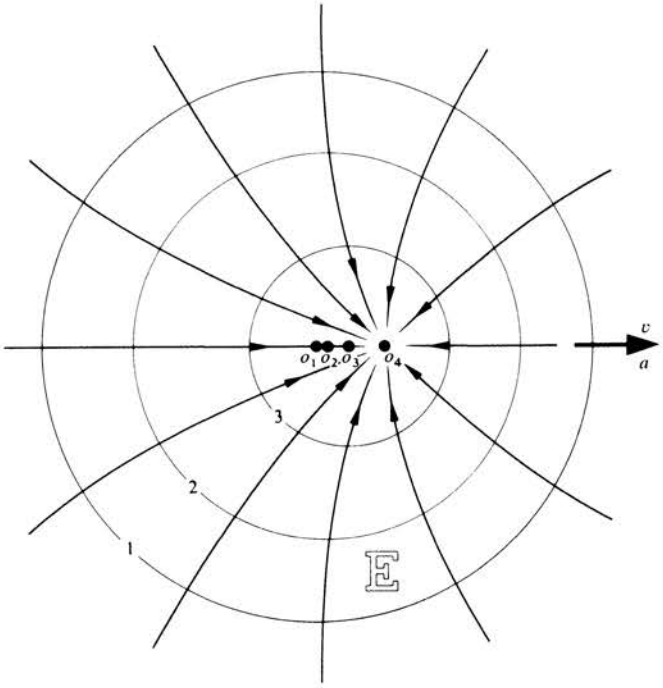


Figure 3.27 Electric field of a uniformly accelerating electron.

In contrast, Fig. 3.27 shows the field lines associated with an electron accelerating uniformly to the right. The points O_1, O_2, O_3, and O_4 are the positions of the electron after equal time intervals. The field lines are now curved, and this is a significant difference. As a further contrast, Fig. 3.28 depicts the field of an electron at some arbitrary time t_2. Before $t = 0$ the particle was always at rest at the point O. The charge was then uniformly accelerated until time t_1, reaching a speed v, which was maintained constant thereafter. We can anticipate that the surrounding field lines will somehow carry the information that the electron has accelerated. We have ample reason to assume that this "information" will propagate at the speed c. If, for example, $t_2 = 10^{-8}$ s, no point beyond 3 m from O would be aware of the fact that the charge had even moved. All the lines in that region would be uniform, straight, and centered on O, as if the charge were still there. At time t_2 the electron is at point O_2 moving with a constant speed v. In the vicinity of O_2 the field lines must then resemble those in Fig. 3.26*b*. Gauss's Law requires that the lines outside the sphere of radius ct_2 connect to those within the sphere of radius $c(t_2 - t_1)$, since there are no charges between them. It is now apparent that during the interval when the particle accelerated, the field lines became distorted and a kink appeared. The exact shape of the lines within the region of the kink is of little interest here. What is significant is that there now exists a *transverse component* of the electric field $\vec{\mathbf{E}}_T$, which propagates outward as a pulse. At some point in space the transverse electric field will be a function of time, and it will therefore be accompanied by a magnetic field.

The radial component of the electric field drops off as $1/r^2$, while the transverse component goes as $1/r$. At large distances from the charge, the only significant field will be the $\vec{\mathbf{E}}_T$-component of the pulse, which is known as the *radiation field.** For a positive charge moving slowly ($v << c$), the electric and magnetic radiation fields can be shown to be proportional to $\vec{\mathbf{r}} \times (\vec{\mathbf{r}} \times \vec{\mathbf{a}})$ and $(\vec{\mathbf{a}} \times \vec{\mathbf{r}})$, respectively, where $\vec{\mathbf{a}}$ is the acceleration. For a negative charge the reverse occurs, as shown in

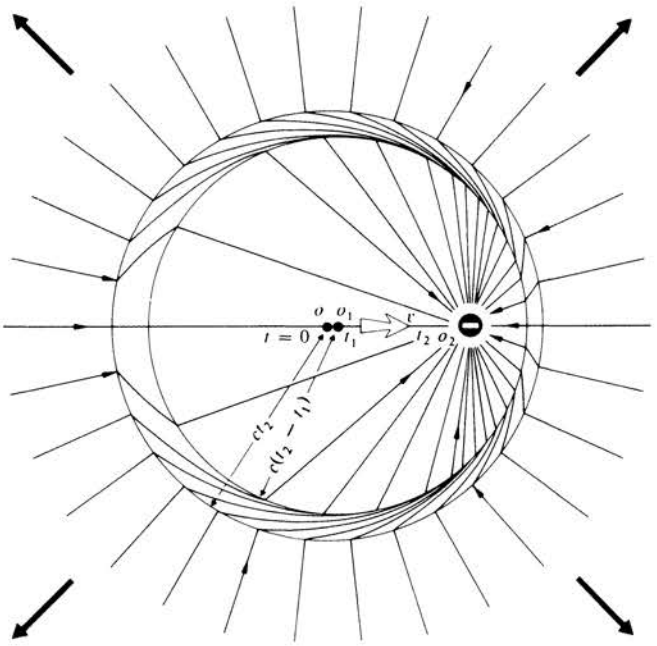


Figure 3.28 A kink in the $\vec{\mathbf{E}}$-field lines.

*The details of this calculation using J. J. Thomson's method of analyzing the kink can be found in J. R. Tessman and J. T. Finnell, Jr., "Electric Field of an Accelerating Charge," *Am. J. Phys.* **35**, 523 (1967). As a general reference for radiation, see, for example, Marion and Heald, *Classical Electromagnetic Radiation*, Chapter 7.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

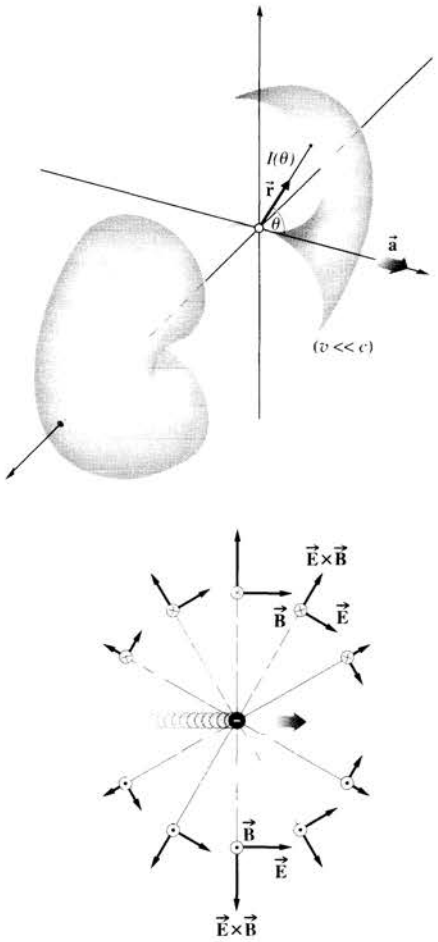


Figure 3.29 The toroidal radiation pattern of a linearly accelerating charge (split to show cross section).

Fig. 3.29. Observe that the irradiance is a function of θ and that $I(0) = I(180°) = 0$ while $I(90°) = I(270°)$ is a maximum. ***Energy is most strongly radiated perpendicular to the acceleration causing it.***

The energy that is radiated out into the surrounding space is supplied to the charge by some external agent. That agent is responsible for the accelerating force, which in turn does work on the charge.

3.4.2 Synchrotron Radiation

A free charged particle traveling on any sort of curved path is accelerating and will radiate. This provides a powerful mechanism for producing radiant energy, both naturally and in the laboratory. The synchrotron radiation generator, a research tool developed in the 1970s, does just that. Clumps of charged particles, usually electrons or positrons, interacting with an applied magnetic field are made to revolve around a large, essentially circular track at a precisely controlled speed. The frequency of the orbit determines the fundamental frequency of the emission (which also contains higher harmonics), and that's continuously variable, more or less, as desired. Incidentally, it's necessary to use clumps of charge; *a uniform loop of current does not radiate.*

A charged particle slowly revolving in a circular orbit radiates a doughnut-shaped pattern similar to the one depicted in Fig. 3.29. Again the distribution of radiation is symmetrical around $\vec{\mathbf{a}}$, which is now the centripetal acceleration acting inward along the radius drawn from the center of the circular orbit to the charge. And once more, ***energy is most strongly radiated perpendicular to the acceleration causing it.*** The higher the speed, the more an observer at rest in the laboratory will "see" the backward lobe of the radiation pattern shrink while the forward lobe elongates in the direction of motion. At speeds approaching c, the particle beam (usually with a diameter comparable to that of a straight pin) radiates essentially along a narrow cone pointing tangent to the orbit in the instantaneous direction of $\vec{\mathbf{v}}$ (Fig. 3.30). Moreover, for $v \approx c$ the radiation will be strongly polarized in the plane of the motion.

This "searchlight," often less than a few millimeters in diameter, sweeps around as the particle clumps circle the machine, much like the headlight on a train rounding a turn. With each revolution the beam momentarily ($< \frac{1}{2}$ ns) flashes through one of many windows in the device. As we will learn (p. 314), when a signal has a short duration it must comprise a broad range of frequencies. The result is a tremendously intense source of rapidly pulsating radiation, tunable over a wide range of frequencies, from infrared to light to X-rays.

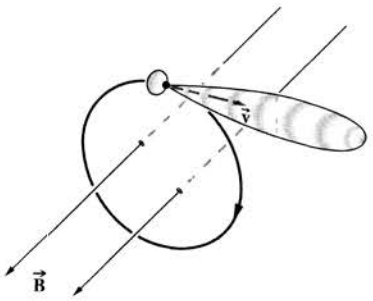
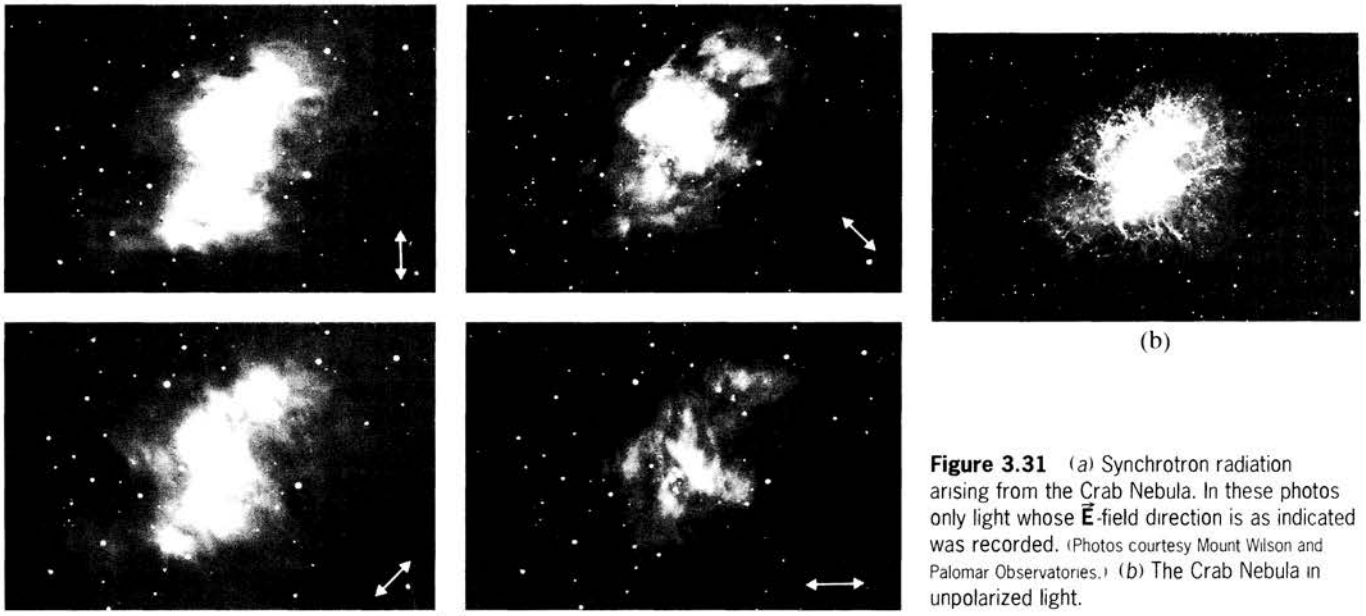


Figure 3.30 Radiation pattern for an orbiting charge.

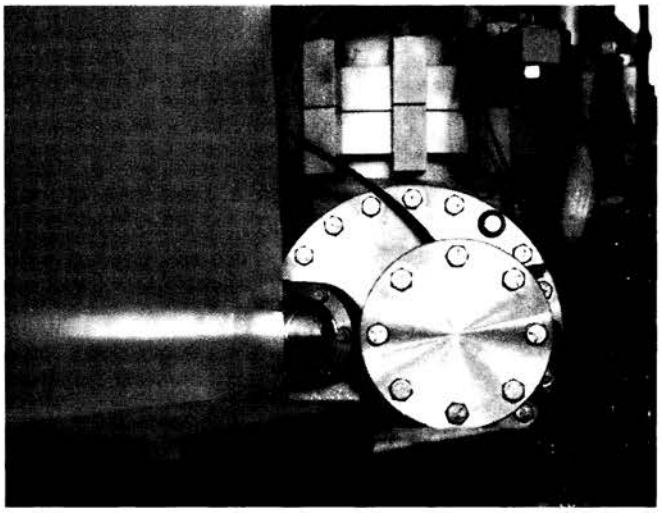
Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



(a) (b)

Figure 3.31 (*a*) Synchrotron radiation arising from the Crab Nebula. In these photos only light whose $\vec{E}$-field direction is as indicated was recorded. (Photos courtesy Mount Wilson and Palomar Observatories.) (*b*) The Crab Nebula in unpolarized light.

When magnets are used to make the circulating electrons wiggle in and out of their circular orbits, bursts of high-frequency X-rays of unparalleled intensity can be created. These beams are hundreds of thousands of times more powerful than a dental X-ray (which is roughly a fraction of a watt) and can easily burn a finger-sized hole through a 3-mm-thick lead plate.



The first beam of "light" from the National Synchrotron Light Source (1982) emanating from its ultraviolet electron storage ring.

Although this technique was first used to produce light in an electron synchrotron as long ago as 1947, it took several decades to recognize that what was an energy-robbing nuisance to the accelerator people might be a major research tool in itself (see photo).

In the astronomical realm, we can expect that some regions exist that are pervaded by magnetic fields. Charged particles trapped in these fields will move in circular or helical orbits, and if their speeds are high enough, they will emit synchrotron radiation. Figure 3.31 shows five photographs of the extragalactic Crab Nebula.* Radiation emanating from the nebula extends over the range from radio frequencies to the

*The Crab Nebula is believed to be expanding debris left over after the cataclysmic death of a star. From its rate of expansion, astronomers calculated that the explosion took place in A.D. 1050. This was subsequently corroborated when a study of old Chinese records (the chronicles of the Beijing Observatory) revealed the appearance of an extremely bright star, in the same region of the sky, in A.D. 1054.

> In the first year of the period Chihha, the fifth moon, the day Chichou [i.e., July 4, 1054], a great star appeared.... After more than a year, it gradually became invisible.

There is little doubt that the Crab Nebula is the remnant of that supernova.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

extreme ultraviolet. Assuming the source to be trapped circulating charges, we can anticipate strong polarization effects. These are evident in the first four photographs, which were taken through a polarizing filter. The direction of the electric field vector is indicated in each picture. Since in synchrotron radiation, the emitted $\vec{E}$-field is polarized in the orbital plane, we can conclude that each photograph corresponds to a particular uniform magnetic field orientation normal to the orbits and to $\vec{E}$.

It is believed that a majority of the low-frequency radiowaves reaching the Earth from outer space have their origin in synchrotron radiation. In 1960 radio astronomers used these long-wavelength emissions to identify a class of objects known as quasars. In 1955 bursts of polarized radiowaves were discovered emanating from Jupiter. Their origin is now attributed to spiraling electrons trapped in radiation belts surrounding the planet.

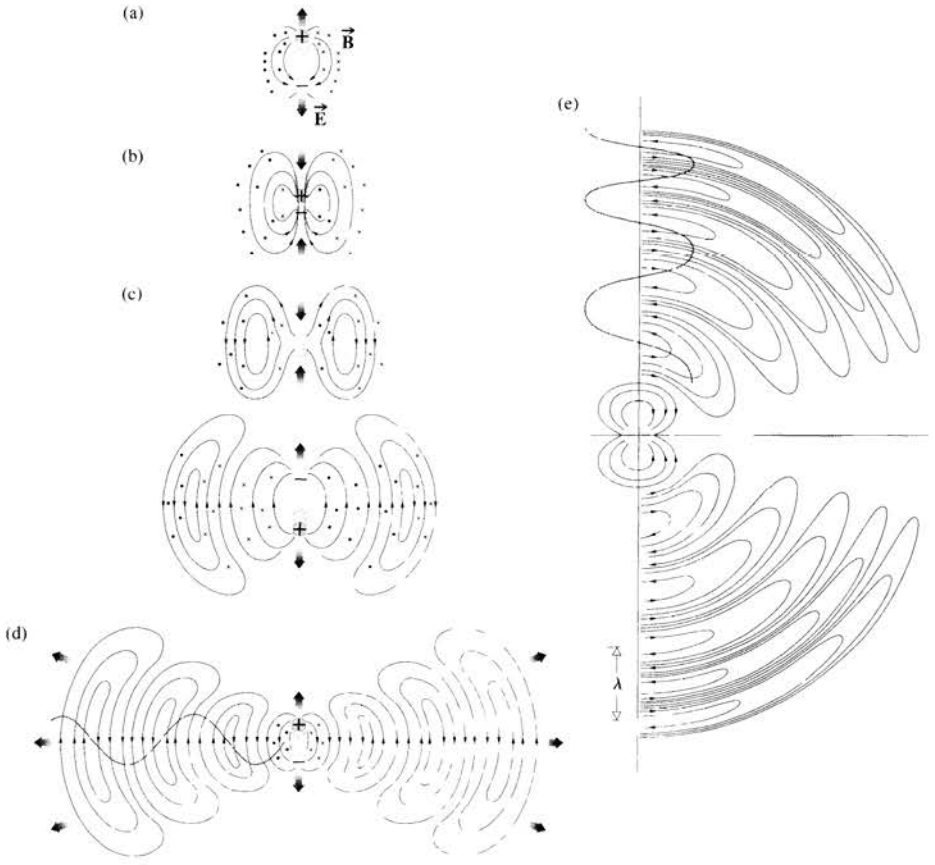


Figure 3.32 The $\vec{E}$-field of an oscillating electric dipole.

3.4.3 Electric Dipole Radiation

Perhaps the simplest electromagnetic wave-producing mechanism to visualize is the oscillating dipole—two charges, one plus and one minus, vibrating to and fro along a straight line. And yet this arrangement is surely the most important of all.

Both light and ultraviolet radiation arise primarily from the rearrangement of the outermost, or weakly bound, electrons in atoms and molecules. It follows from the quantum-mechanical analysis that the electric dipole moment of the atom is the major source of this radiation. The rate of energy emission from a material system, although a quantum-mechanical process, can be envisioned in terms of the classical oscillating electric dipole. This mechanism is therefore central to understanding the way atoms, molecules, and even nuclei emit and absorb electromagnetic waves. Figure 3.32 schematically depicts the electric field distribution in the region of an electric dipole. In this configuration, a negative charge oscillates linearly in simple harmonic motion about an equal stationary positive charge. If the angular frequency of the oscillation is ω, the time-dependent dipole moment $\mathscr{p}(t)$ has the scalar form

$$\mathscr{p} = \mathscr{p}_0 \cos \omega t \qquad (3.55)$$

Note that $\mathscr{p}(t)$ could represent the collective moment of the oscillating charge distribution on the atomic scale or even an oscillating current in a linear television antenna.

At $t = 0, \mathscr{p} = \mathscr{p}_0 = qd$, where d is the initial maximum separation between the centers of the two charges (Fig. 3.32*a*). The dipole moment is actually a vector in the direction from $-q$ to $+q$. The figure shows a sequence of field line patterns as the displacement, and therefore the dipole moment decreases, then goes to zero, and finally reverses direction. When the charges effectively overlap, $\mathscr{p} = 0$, and the field lines must close on themselves.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Very near the atom, the $\vec{\mathbf{E}}$-field has the form of a static electric dipole. A bit farther out, in the region where the closed loops form, there is no specific wavelength. The detailed treatment shows that the electric field is composed of five different terms, and things are fairly complicated. Far from the dipole, in what is called the *wave* or *radiation zone*, the field configuration is much simpler. In this zone, a fixed wavelength has been established; $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are transverse, mutually perpendicular, and in phase. Specifically,

$$E = \frac{\mu_0 k^2 \sin\theta}{4\pi\epsilon_0} \frac{\cos(kr - \omega t)}{r} \tag{3.56}$$

and $B = E/c$, where the fields are oriented as in Fig. 3.33. The Poynting vector $\vec{\mathbf{S}} = \vec{\mathbf{E}} \times \vec{\mathbf{B}}/\mu_0$ always points radially outward in the wave zone. There, the $\vec{\mathbf{B}}$-field lines are circles concentric with, and in a plane perpendicular to, the dipole axis. This is understandable, since $\vec{\mathbf{B}}$ can be considered to arise from the time-varying oscillator current.

The irradiance (radiated radially outward from the source) follows from Eq. (3.44) and is given by

$$I(\theta) = \frac{\mu_0^2 \omega^4}{32\pi^2 c^3 \epsilon_0} \frac{\sin^2\theta}{r^2} \tag{3.57}$$

again an inverse-square-law dependence on distance. The angular flux density distribution is toroidal, as in Fig. 3.29. The axis along which the acceleration takes place is the symmetry axis of the radiation pattern. Notice the dependence of the irradiance on ω^4—*the higher the frequency, the stronger the radiation*. That feature will be important when we consider scattering.

It's not difficult to attach an AC generator between two conducting rods and send currents of free electrons oscillating up and down that "transmitting antenna." Figure 3.34*a* shows

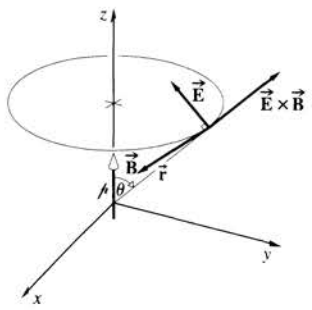


Figure 3.33 Field orientations for an oscillating electric dipole.

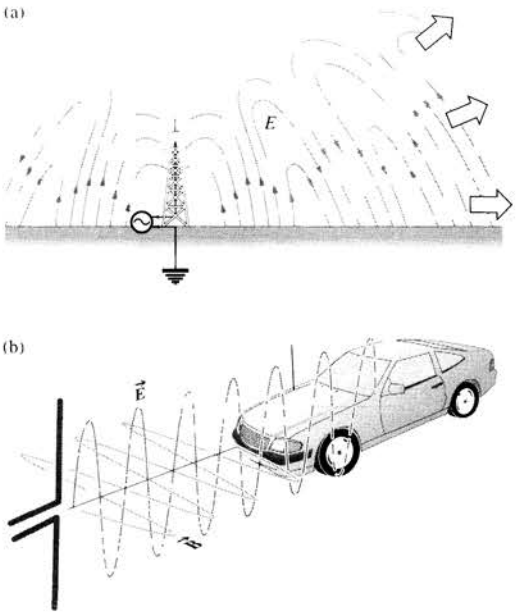


Figure 3.34 (*a*) Electromagnetic waves from a transmitting tower. (*b*) Automobiles often have radio antennas that stick straight up about a meter. The vertically oscillating electric field of a passing radiowave induces a voltage along the length of the antenna, and that becomes the input signal to the receiver.

the arrangement carried to its logical conclusion—a fairly standard AM radio tower. An antenna of this sort will function most efficiently if its length corresponds to the wavelength being transmitted or, more conveniently, to $\frac{1}{2}\lambda$. The radiated wave is then formed at the dipole in synchronization with the oscillating current producing it. AM radiowaves are unfortunately several hundred meters long. Consequently, the antenna shown in the figure has half the $\frac{1}{2}\lambda$-dipole essentially buried in the Earth. That at least saves some height, allowing the device to be built only $\frac{1}{4}\lambda$ tall. Moreover, this use of the Earth also generates a so-called *ground wave* that hugs the planet's surface, where most people with radios are likely to be located. A commercial station usually has a range somewhere between 25 and 100 miles.

3.4.4 The Emission of Light from Atoms

Surely the most significant mechanism responsible for the natural emission and absorption of radiant energy—especially of light—is the *bound charge*, electrons confined within atoms.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

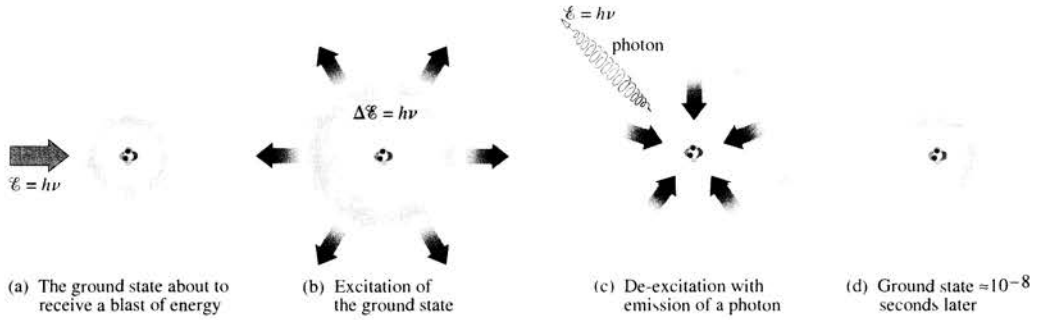


Figure 3.35 The excitation of an atom. (*a*) Energy in the amount $h\nu$ is delivered to the atom. (*b*) Since this matches the energy needed to reach an excited state, the atom absorbs the energy and attains a higher energy level. (*c*) With the emission of a photon, it drops back (*d*) and returns to the ground state in about 10^{-8} s.

These minute negative particles, which surround the massive positive nucleus of each atom, constitute a kind of distant, tenuous charged cloud. Much of the chemical and optical behavior of ordinary matter is determined by its outer or valence electrons. The remainder of the cloud is ordinarily formed into "closed," essentially unresponsive, shells around and tightly bound to the nucleus. These closed or filled shells are made up of specific numbers of electron pairs. Even though it is not completely clear what occurs internally when an atom radiates, we do know with some certainty that light is emitted during readjustments in the outer charge distribution of the electron cloud. This mechanism is ultimately the predominant source of light in the world.

Usually, an atom exists with its clutch of electrons arranged in some stable configuration that corresponds to their lowest energy distribution or *level*. Every electron is in the lowest possible energy state available to it, and the atom as a whole is in its so-called **ground-state** configuration. There it will likely remain indefinitely, if left undisturbed. Any mechanism that pumps energy into the atom will alter the ground state. For instance, a collision with another atom, an electron, or a photon can affect the atom's energy state profoundly. An atom can exist with its electron cloud in only certain specific configurations corresponding to only certain values of energy. In addition to the ground state, there are higher energy levels, the **excited states**, each associated with a specific cloud configuration and a specific well-defined energy. When one or more electrons occupies a level higher than its ground-state level, the atom is said to be **excited**—a condition that is inherently unstable and temporary.

At low temperatures, atoms tend to be in their ground state; at progressively higher temperatures, more and more of them will become excited through atomic collisions. This sort of mechanism is indicative of a class of relatively gentle excitations—glow discharge, flame, spark, and so forth—which energize only the outermost unpaired valence electrons. We will initially concentrate on these outer electron transitions, which give rise to the emission of light, and the nearby infrared and ultraviolet.

When enough energy is imparted to an atom (typically to the valence electron), whatever the cause, the atom can react by suddenly ascending from a lower to a higher energy level (Fig. 3.35). The electron will make a very rapid transition, a **quantum jump**, from its ground-state orbital configuration to one of the well-delineated excited states, one of the quantized rungs on its energy ladder. As a rule, ***the amount of energy taken up in the process equals the energy difference between the initial and final states, and since that is specific and well defined, the amount of energy that can be absorbed by an atom is quantized*** (i.e., limited to specific amounts). This state of atomic excitation is a short-lived resonance phenomenon. Usually, after about 10^{-8} s or 10^{-9} s, the excited atom spontaneously relaxes back to a lower state, most often the ground state, losing the excitation energy along the way. This energy readjustment can occur by way of the emission of light or (especially in dense materials) by conversion to thermal energy through interatomic collisions within the medium. (As we'll soon see, this latter mechanism results in the absorption of light at the resonant frequency and the transmission or reflection of the remaining frequencies—it's responsible for most of the coloration in the world around us.)

If the atomic transition is accompanied by the emission of light (as it is in a rarefied gas), ***the energy of the photon exactly matches the quantized energy decrease of the atom***. That corresponds to a specific frequency, by way of $\Delta\mathscr{E} = h\nu$, a frequency associated with both the photon and the atomic transi-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

tion between the two particular states. This is said to be a **resonance frequency**, one of several (each with its own likelihood of occurring) at which the atom very efficiently absorbs and emits energy. The atom radiates a quantum of energy that presumably is created spontaneously, on the spot, by the shifting electron.

Even though what occurs during the atom-transition interval of 10^{-8} s is far from clear, it can be helpful to imagine the orbital electron somehow making its downward energy transition via a gradually damped oscillatory motion at the specific resonance frequency. The radiated light can then be envisioned in a semiclassical way as emitted in a short oscillatory directional pulse, or **wavetrain**, lasting less than roughly 10^{-8} s—a picture that is in agreement with certain experimental observations (see Section 7.4.2, and Fig. 7.36). It's useful to think of this electromagnetic pulse as associated in some inextricable fashion with the photon. In a way, the pulse is a semiclassical representation of the manifest wave nature of the photon. But the two are *not* equivalent in all respects: the electromagnetic wavetrain is a classical creation that describes the propagation and spatial distribution of light extremely well; yet its energy is not quantized, and that is an essential characteristic of the photon. So when we consider photon wavetrains, keep in mind that there is more to the idea than just a classical oscillatory pulse of electromagnetic wave. Of course, the reason even to introduce the notion of the emission of wavetrains is to have a basis for talking about the frequency of the light. This is perhaps the central problem in any naïve photon model: What agency manifests the frequency?

The emission spectra of single atoms or low-pressure gases, whose atoms do not interact appreciably, consist of sharp "lines," that is, fairly well-defined frequencies characteristic of the atoms. There is always some frequency broadening of that radiation due to atomic motion, collisions, and so forth; hence it's never precisely monochromatic. Generally, however, the atomic transition from one level to another is characterized by the emission of a well-defined, narrow range of frequencies. On the other hand, the spectra of solids and liquids, in which the atoms are interacting with one another, are broadened into wide frequency bands. When two atoms are brought close together, the result is a slight shift in their respective energy levels because they act on each other. The many interacting atoms in a solid create a tremendous number of such shifted levels, in effect spreading out each of their original levels, blurring them into essentially continuous bands. Materials of this nature emit and absorb over broad ranges of frequencies.

Optical Cooling

The linear momentum carried by photons can be transferred to moving atoms or ions, thereby drastically changing their motion. After about ten thousand absorption and subsequent emission cycles, an atom, which was originally moving at perhaps 700 m/s, can be slowed to near zero speed. Since, in general, temperature is proportional to the average kinetic energy (KE) of the particles constituting a system, this process is called **optical** or **laser cooling**. With it, KE temperatures in the microkelvin range are attainable. Laser cooling has become the basis for a variety of applications including the atomic clock, the atom interferometer, and the focusing of atomic beams. For us it brings together the ideas of Sections 3.3.4 and 3.4.4 in a compelling, practical way.

Figure 3.36 depicts a beam of atoms, each of mass m traveling with a velocity $\vec{\mathbf{v}}$, colliding with a counterdirected beam of laser photons having a propagation vector $\vec{\mathbf{k}}_L$. The laser frequency ν_L is selected to be just beneath the resonant frequency (ν_0) of the atoms. Because of its motion, any particular atom "sees" an oncoming photon with a frequency that is Doppler-shifted* upward by an amount $|\vec{\mathbf{k}}_L \cdot \vec{\mathbf{v}}|/2\pi = \nu_L v/c$. When the laser frequency is tuned so that $\nu_0 = \nu_L(1 + v/c)$, collisions with the photons will resonate the atoms. In the process, each photon transfers its momentum of $\hbar\vec{\mathbf{k}}_L$ to the absorbing atom (see p. 57) whose speed is thereupon reduced by an amount Δv where $m\Delta v = \hbar k_L$.

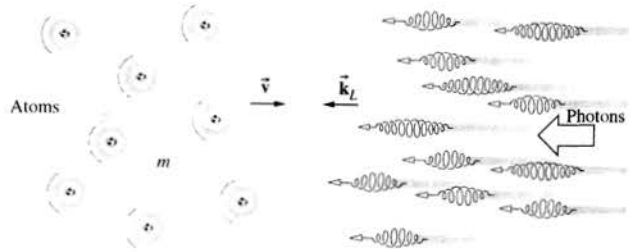


Figure 3.36 A stream of atoms colliding with a laserbeam in a process called laser cooling.

*Imagine an observer moving at v_o, toward a source that is sending out waves having a speed v at a frequency ν_s. As a result of the Doppler Effect, he will experience a frequency $\nu_o = \nu_s(v + v_o)/v$. For more of the details, see almost any introductory physics text, for example, E. Hecht, *Physics: Calculus*, Sect. 11.11.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

The cloud of atoms is not very dense, and each excited atom can drop back to its ground state with the spontaneous emission of a photon of energy $h\nu_0$. This emission is randomly directed, and so although the atom recoils, the average amount of momentum regained by it over thousands of cycles tends to zero. The change in momentum of the atom per photon absorption-emission cycle is therefore effectively $\hbar\vec{\mathbf{k}}_L$, and it slows down. In each cycle (as seen by someone at rest in the lab), the atom absorbs a photon of energy $h\nu_L$, emits a photon of energy $h\nu_0$, and in the process loses an amount of KE corresponding to $h\nu_L v/c$, which is proportional to the Doppler Shift.

By contrast, an atom moving in the opposite direction, away from the light source, sees photons to have a frequency $\nu_L(1 - v/c)$, far enough away from ν_0 that there can be little or no absorption, and therefore no momentum gain.

Notice that the radiation pressure force is frequency dependent and the atoms experience a speed-dependent force via the Doppler Effect. That means that ν_0 and ν_L have to be kept in the proper relationship as v decreases. There are a number of very clever ways this is accomplished.

3.5 Light in Bulk Matter

The response of dielectric or nonconducting materials to electromagnetic fields is of special concern in Optics. We will, of course, be dealing with transparent dielectrics in the form of lenses, prisms, plates, films, and so forth, not to mention the surrounding sea of air.

The net effect of introducing a homogeneous, isotropic dielectric into a region of free space is to change ϵ_0 to ϵ and μ_0 to μ in Maxwell's Equations. The phase speed in the medium now becomes

$$v = 1/\sqrt{\epsilon\mu} \tag{3.58}$$

The ratio of the speed of an electromagnetic wave in vacuum to that in matter is known as the **absolute index of refraction** n:

$$n \equiv \frac{c}{v} = \sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}} \tag{3.59}$$

In terms of the relative permittivity and relative permeability of the medium, n becomes

$$n = \sqrt{K_E K_M} \tag{3.60}$$

There are magnetic substances that are transparent in the infrared and microwave regions of the spectrum. But we are primarily interested in materials that are transparent in the visible, and these are all essentially "nonmagnetic." Indeed, K_M generally doesn't deviate from 1 by any more than a few parts in 10^4 (e.g., for diamond $K_M = 1 - 2.2 \times 10^{-5}$). Setting $K_M = 1$ in the formula for n results in an expression known as *Maxwell's Relation*, namely,

$$n \approx \sqrt{K_E} \tag{3.61}$$

wherein K_E is presumed to be the *static dielectric constant*. As indicated in Table 3.2, this relationship seems to work well only for some simple gases. The difficulty arises because K_E and therefore n are actually *frequency-dependent*. The dependence of n on the wavelength (or color) of light is a well-known effect called **dispersion**. It arises on a microscopic level, and so Maxwell's Equations are quite oblivious to it. Sir Isaac Newton used prisms to disperse white light into its constituent colors over three hundred years ago, and the phenomenon was well known, if not well understood, even then.

TABLE 3.2 Maxwell's Relation

Gases at 0°C and 1 atm		
Substance	$\sqrt{K_E}$	n
Air	1.000294	1.000293
Helium	1.000034	1.000036
Hydrogen	1.000131	1.000132
Carbon dioxide	1.00049	1.00045
Liquids at 20°C		
Substance	$\sqrt{K_E}$	n
Benzene	1.51	1.501
Water	8.96	1.333
Ethyl alcohol (ethanol)	5.08	1.361
Carbon tetrachloride	4.63	1.461
Carbon disulfide	5.04	1.628
Solids at room temperature		
Substance	$\sqrt{K_E}$	n
Diamond	4.06	2.419
Amber	1.6	1.55
Fused silica	1.94	1.458
Sodium chloride	2.37	1.50

Values of K_E correspond to the lowest possible frequencies, in some cases as low as 60 Hz, whereas n is measured at about 0.5×10^{15} Hz. Sodium D light was used ($\lambda = 589.29$ nm).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Scattering and Absorption

What is the physical basis for the frequency dependence of n? The answer to that question can be found by examining the interaction of an incident electromagnetic wave with the array of atoms constituting a dielectric material. An atom can react to incoming light in two different ways, depending on the incident frequency or equivalently on the incoming photon energy ($\mathscr{E} = h\nu$). Generally, the atom will "scatter" the light, redirecting it without otherwise altering it. On the other hand, if the photon's energy matches that of one of the excited states, the atom will absorb the light, making a quantum jump to that higher energy level. In the dense atomic landscape of ordinary gases (at pressures of about 10^2 Pa and up), solids, and liquids, it's very likely that this excitation energy will rapidly be transferred, via collisions, to random atomic motion, thermal energy, before a photon can be emitted. This commonplace process (the taking up of a photon and its conversion into thermal energy) was at one time widely known as "absorption," but nowadays that word is more often used to refer just to the "taking up" aspect, regardless of what then happens to the energy. Consequently, it's now better referred to as **dissipative absorption**. All material media partake in dissipative absorption to some extent, at one frequency or another.

In contrast to this excitation process, *ground-state* or **nonresonant scattering** occurs with incoming radiant energy of other frequencies—that is, lower than the resonance frequencies. Imagine an atom in its lowest state and suppose that it interacts with a photon whose energy is too small to cause a transition to any of the higher, excited states. Despite that, the electromagnetic field of the light can be supposed to drive the electron cloud into oscillation. There is no resulting atomic transition; the atom remains in its ground state while the cloud vibrates ever so slightly at the frequency of the incident light. Once the electron cloud starts to vibrate with respect to the positive nucleus, the system constitutes an oscillating dipole and so presumably will *immediately* begin to radiate at that same frequency. The resulting scattered light consists of a photon that sails off in some direction carrying the same amount of energy as did the incident photon—***the scattering is elastic***. In effect, the atom resembles a little dipole oscillator, a model employed by Hendrik Antoon Lorentz (1878) in order to extend Maxwell's Theory, in a classical way, to the atomic domain. If the incident light is unpolarized, the atomic oscillators scatter in random directions.

When an atom is irradiated with light, the process of excitation and spontaneous emission is rapidly repeated. In fact, with an emission lifetime of $\approx 10^{-8}$ s, an atom could emit upward of 10^8 photons per second in a situation in which there was enough energy to keep reexciting it. Atoms have a very strong tendency to interact with resonant light (they have a large *absorption cross section*). This means that the saturation condition, in which the atoms of a low-pressure gas are constantly emitting and being reexcited, occurs at a modest value of irradiance ($\approx 10^2$ W / m^2). So it's not very difficult to get atoms firing out photons at a rate of 100 million per second.

Generally, we can imagine that in a medium illuminated by an ordinary beam of light, each atom behaves as though it was a "source" of a tremendous number of photons (scattered either elastically or resonantly) that fly off in all directions. A stream of energy like this resembles a classical spherical wave. ***Thus we imagine an atom (even though it is simplistic to do so) as a point source of spherical electromagnetic waves***—provided we keep in mind Einstein's admonition that "outgoing radiation in the form of spherical waves does not exist."

When a material with no resonances in the visible is bathed in light, nonresonant scattering occurs, and it gives each participating atom the appearance of being a tiny source of spherical wavelets. As a rule, the closer the frequency of the incident beam is to an atomic resonance, the more strongly will the interaction occur and, in dense materials, the more energy will be dissipatively absorbed. It is precisely this mechanism of *selective absorption* (see Section 4.9) that creates much of the visual appearance of things. It is primarily responsible for the color of your hair, skin, and clothing, the color of leaves and apples and paint.

3.5.1 Dispersion

Dispersion corresponds to the phenomenon whereby the index of refraction of a medium is frequency dependent. All material media are dispersive; only vacuum is nondispersive.

Maxwell's Theory treats substantial matter as continuous, representing its electric and magnetic responses to applied $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields in terms of constants, ϵ and μ. Consequently, K_E and K_M are also constant, and n is therefore unrealistically independent of frequency. To deal theoretically with dispersion, it's necessary to incorporate the atomic nature of matter and to exploit some frequency-dependent aspect of that nature. Following H. A. Lorentz, the contributions of large numbers of atoms can be averaged to represent the behavior of an isotropic dielectric medium.

When a dielectric is subjected to an applied electric field,

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

the internal charge distribution is distorted. This corresponds to the generation of electric dipole moments, which in turn contribute to the total internal field. More simply stated, the external field separates positive and negative charges in the medium (each pair of which is a dipole), and these charges then contribute an additional field component. The resultant dipole moment per unit volume is called the **electric polarization**, $\vec{\mathbf{P}}$. For most materials $\vec{\mathbf{P}}$ and $\vec{\mathbf{E}}$ are proportional and can satisfactorily be related by

$$(\epsilon - \epsilon_0)\vec{\mathbf{E}} = \vec{\mathbf{P}} \tag{3.62}$$

The redistribution of charge and the consequent polarization can occur via the following mechanisms. There are molecules that have a permanent dipole moment as a result of unequal sharing of valence electrons. These are known as *polar molecules*; the nonlinear water molecule is a fairly typical example (Fig. 3.37). Each hydrogen–oxygen bond is polar covalent, with the H-end positive with respect to the O-end. Thermal agitation keeps the molecular dipoles randomly oriented. With the introduction of an electric field, the dipoles align themselves, and the dielectric takes on an **orientational polarization**. In the case of *nonpolar molecules* and *atoms*, the applied field distorts the electron cloud, shifting it relative to the nucleus, thereby producing a dipole moment. In addition to this **electronic polarization**, there is another process that's applicable specifically to molecules, for example, the ionic crystal NaCl. In the presence of an electric field, the positive and negative ions undergo a shift with respect to each other. Dipole moments are therefore induced, resulting in what is called **ionic** or **atomic polarization**.

If the dielectric is subjected to an incident harmonic electromagnetic wave, its internal charge structure will experience time-varying forces and/or torques. These will be proportional to the electric field component of the wave.* For fluids that are polar dielectrics, the molecules actually undergo rapid rotations, aligning themselves with the $\vec{\mathbf{E}}(t)$-field. But these molecules are relatively large and have appreciable moments of inertia. At high driving frequencies ω, polar molecules will be unable to follow the field alternations. Their contributions to $\vec{\mathbf{P}}$ will decrease, and K_E will drop markedly. The relative permittivity of water is fairly constant at approximately 80, up to about 10^{10} Hz, after which it falls off quite rapidly.

*Forces arising from the magnetic component of the field have the form $\vec{\mathbf{F}}_M = q_e\vec{\mathbf{v}} \times \vec{\mathbf{B}}$ in comparison to $\vec{\mathbf{F}}_E = q_e\vec{\mathbf{E}}$ for the electric component; but $v << c$, so it follows from Eq. (3.30) that $\vec{\mathbf{F}}_M$ is generally negligible.

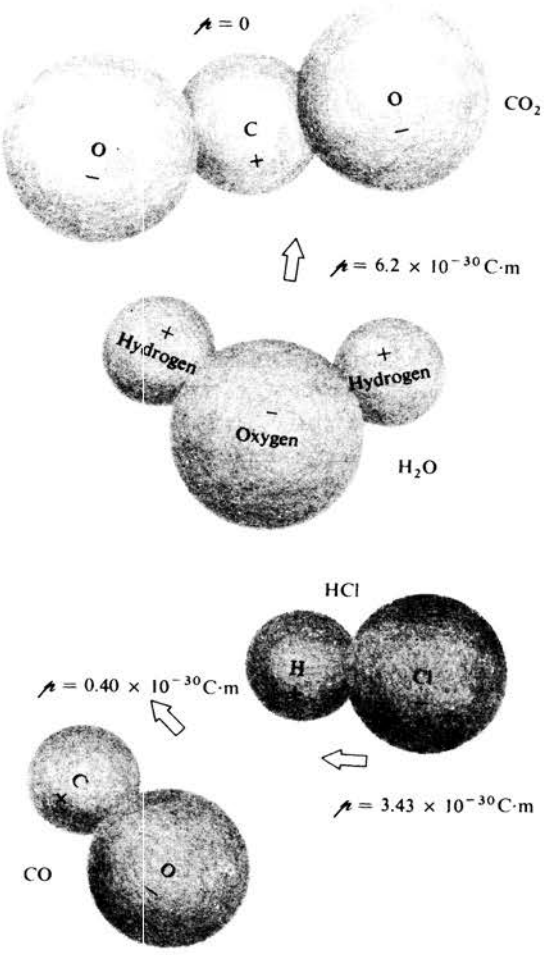


Figure 3.37 Assorted molecules and their dipole moments.

In contrast, electrons have little inertia and can continue to follow the field contributing to $K_E(\omega)$ even at optical frequencies (of about 5×10^{14} Hz). Thus the dependence of n on ω is governed by the interplay of the various electric polarization mechanisms contributing at the particular frequency. With this in mind, it is possible to derive an analytical expression for $n(\omega)$ in terms of what's happening within the medium on an atomic level.

The electron cloud of the atom is bound to the positive nucleus by an attractive electric force that sustains it in some sort of equilibrium configuration. Without knowing much more about the details of all the internal atomic interactions, we can anticipate that, like other stable mechanical systems

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

which are not totally disrupted by small perturbations, a net force, F, must exist that returns the system to equilibrium. Moreover, we can reasonably expect that for very small displacements, x, from equilibrium (where $F = 0$), the force will be linear in x. In other words, a plot of $F(x)$ versus x will cross the x-axis at the equilibrium point ($x = 0$) and will be a straight line very close on either side. Thus for small displacements it can be supposed that the restoring force has the form $F = -k_E x$, where k_E is a kind of elastic constant much like a spring constant. Once somehow momentarily disturbed, an electron bound in this way will oscillate about its equilibrium position with a **natural** or **resonant frequency** given by $\omega_0 = \sqrt{k_E/m_e}$, where m_e is its mass. This is the oscillatory frequency of the *undriven* system.

A material medium is envisioned as an assemblage, in vacuum, of a very great many polarizable atoms, each of which is small (by comparison to the wavelength of light) and close to its neighbors. When a lightwave impinges on such a medium, each atom can be thought of as a classical **forced oscillator** being driven by the time-varying electric field $E(t)$ of the wave, which is assumed here to be applied in the x-direction. Figure 3.38*b* is a mechanical representation of just such an oscillator in an *isotropic medium* where the negatively charged shell is fastened to a stationary positive nucleus by identical springs. Even under the illumination of bright sunlight, the amplitude of the oscillations will be no greater than about 10^{-17} m. The force (F_E) exerted on an electron of charge q_e by the $E(t)$ field of a harmonic wave of frequency ω is of the form

$$F_E = q_e E(t) = q_e E_0 \cos \omega t \tag{3.63}$$

Notice that if the driving force is in one direction the restoring force is in the opposite direction, which is why it has a minus sign. Newton's Second Law provides the equation of motion; that is, the sum of the forces equals the mass times the acceleration:

$$q_e E_0 \cos \omega t - m_e \omega_0^2 x = m_e \frac{d^2x}{dt^2} \tag{3.64}$$

The first term on the left is the driving force, and the second is the opposing restoring force. To satisfy this expression, x will have to be a function whose second derivative isn't very much different from x itself. Furthermore, we can anticipate that the electron will oscillate at the same frequency as $E(t)$, so we "guess" at the solution

$$x(t) = x_0 \cos \omega t$$

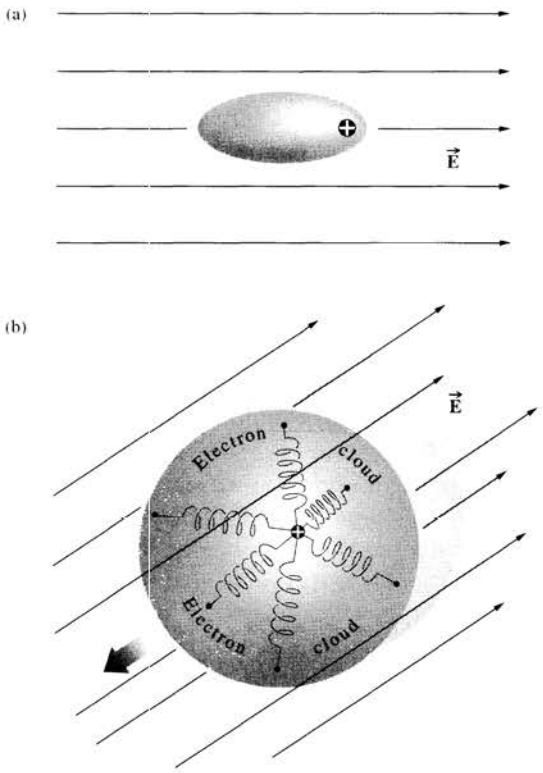


Figure 3.38 (*a*) Distortion of the electron cloud in response to an applied $\vec{E}$-field. (*b*) The mechanical oscillator model for an isotropic medium—all the springs are the same, and the oscillator can vibrate equally in all directions.

and substitute it in the equation to evaluate the amplitude x_0. In this way we find that

$$x(t) = \frac{q_e/m_e}{(\omega_0^2 - \omega^2)} E_0 \cos \omega t \tag{3.65}$$

or

$$x(t) = \frac{q_e/m_e}{(\omega_0^2 - \omega^2)} E(t) \tag{3.66}$$

This is the relative displacement between the negative cloud and the positive nucleus. It's traditional to leave q_e positive and speak about the displacement of the oscillator. Without a driving force (no incident wave), the oscillator will vibrate at its *resonance frequency* ω_0. In the presence of a field whose frequency is less than ω_0, $E(t)$ and $x(t)$ have the same sign, which means that the oscillator can follow the applied force (i.e., is in-phase with it). However, when $\omega > \omega_0$, the displacement $x(t)$ is in a direction opposite to that of the instanta-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

neous force $q_e E(t)$ and therefore 180° out-of-phase with it. Remember that we are talking about oscillating dipoles where for $\omega_0 > \omega$, the relative motion of the *positive* charge is a vibration in the direction of the field. Above resonance the positive charge is 180° out-of-phase with the field, and the dipole is said to lag by π rad.

The dipole moment is equal to the charge q_e times its displacement, and if there are N contributing electrons per unit volume, the electric polarization, or density of dipole moments, is

$$P = q_e x N \tag{3.67}$$

Hence

$$P = \frac{q_e^2 NE/m_e}{(\omega_0^2 - \omega^2)} \tag{3.68}$$

and from Eq. (3.62)

$$\epsilon = \epsilon_0 + \frac{P(t)}{E(t)} = \epsilon_0 + \frac{q_e^2 N/m_e}{(\omega_0^2 - \omega^2)} \tag{3.69}$$

Using the fact that $n^2 = K_E = \epsilon/\epsilon_0$, we can arrive at an expression for n as a function of ω, which is known as a **dispersion equation**:

$$n^2(\omega) = 1 + \frac{Nq_e^2}{\epsilon_0 m_e}\left(\frac{1}{\omega_0^2 - \omega^2}\right) \tag{3.70}$$

At frequencies increasingly above resonance, $(\omega_0^2 - \omega^2) < 0$, and the oscillator undergoes displacements that are approximately 180° out-of-phase with the driving force. The resulting electric polarization will therefore be similarly out-of-phase with the applied electric field. Hence the dielectric constant and therefore the index of refraction will both be less than 1. At frequencies increasingly below resonance, $(\omega_0^2 - \omega^2) > 0$, the electric polarization will be nearly in-phase with the applied electric field. The dielectric constant and the corresponding index of refraction will then both be greater than 1. This kind of behavior, which actually represents only part of what happens, is nonetheless generally observed in all sorts of materials.

We can test the utility of the analysis using a dispersive prism (p. 187) made of the sample material under study, but first we rewrite Eq. (3.70), as is done in Problem 3.53:

$$(n^2 - 1)^{-1} = -C\lambda^{-2} + C\lambda_0^{-2}$$

where, since $\omega = 2\pi c/\lambda$, the multiplicative constant is given by $C = 4\pi^2 c^2 \epsilon_0 m_e/Nq_e^2$. Figure 3.39 is a plot of $(n^2 - 1)^{-1}$ versus λ^{-2} using data from a student experiment. A crown-glass prism was illuminated with the various wavelengths from a He discharge tube, and the index of refraction was measured for each one (Table 3.3). The resulting curve is indeed a straight line; its slope (using $y = mx + b$) equals $-C$, and its y-intercept corresponds to $C\lambda_0^{-2}$. From this it follows that the resonant frequency is 2.95×10^{15} Hz, which is properly in the ultraviolet.

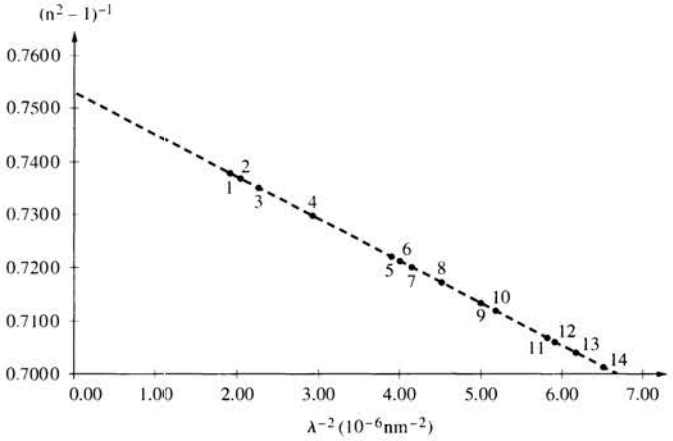


Figure 3.39 Graph of $(n^2 - 1)^{-1}$ versus λ^{-2} for the data shown in Table 3.3. See N. Gauthier, *Phys. Teach.*, **25**, 502 (1987).

TABLE 3.3 Dispersion of Crown Glass. The wavelengths are those of a He discharge tube. The corresponding indices were measured.

	Wavelength λ (nm)	Index of Refraction n
1.	728.135	1.534 6
2.	706.519, 706.570	1.535 2
3.	667.815	1.536 29
4.	587.562, 587.587	1.539 54
5.	504.774	1.544 17
6.	501.567	1.544 73
7.	492.193	1.545 28
8.	471.314	1.546 24
9.	447.148	1.549 43
10.	438.793	1.550 26
11.	414.376	1.553 74
12.	412.086	1.554 02
13.	402.619	1.555 30
14.	388.865	1.557 67

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

As a rule, any given substance will actually undergo several transitions from $n > 1$ to $n < 1$ as the illuminating frequency is made to increase. The implication is that instead of a single frequency ω_0 at which the system resonates, there apparently are several such frequencies. It would seem reasonable to generalize matters by supposing that there are N molecules per unit volume, each with f_j oscillators having natural frequencies ω_{0j}, where $j = 1, 2, 3, \ldots$. In that case,

$$n^2(\omega) = 1 + \frac{Nq_e^2}{\epsilon_0 m_e} \sum_j \left(\frac{f_j}{\omega_{0j}^2 - \omega^2} \right) \tag{3.71}$$

This is essentially the same result as that arising from the quantum-mechanical treatment, with the exception that some of the terms must be reinterpreted. Accordingly, the quantities ω_{0j} would then be the characteristic frequencies at which an atom may absorb or emit radiant energy. The f_j terms, which satisfy the requirement that $\sum_j f_j = 1$, are weighting factors known as *oscillator strengths*. They reflect the emphasis that should be placed on each one of the modes. Since they measure the likelihood that a given atomic transition will occur, the f_j terms are also known as *transition probabilities*.

A similar reinterpretation of the f_j terms is even required classically, since agreement with the experimental data demands that they be less than unity. This is obviously contrary to the definition of the f_j that led to Eq. (3.71). One then supposes that a molecule has many oscillatory modes but that each of these has a distinct natural frequency and strength.

Notice that when ω equals any of the characteristic frequencies, n is discontinuous, contrary to actual observation. This is simply the result of having neglected the damping term, which should have appeared in the denominator of the sum. Incidentally, the damping, in part, is attributable to energy lost when the forced oscillators reradiate. In solids, liquids, and gases at high pressure ($\approx 10^3$ atm), the interatomic distances are roughly 10 times less than those of a gas at standard temperature and pressure. Atoms and molecules in this relatively close proximity experience strong interactions and a resulting "frictional" force. The effect is a damping of the oscillators and a dissipation of their energy within the substance in the form of "heat" (random molecular motion).

Had we included a damping force proportional to the speed (of the form $m_e \gamma \, dx/dt$) in the equation of motion, the dispersion equation (3.71) would have been

$$n^2(\omega) = 1 + \frac{Nq_e^2}{\epsilon_0 m_e} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 + i\gamma_j \omega} \tag{3.72}$$

Although this expression is fine for rarefied media such as gases, there is another complication that must be overcome if the equation is to be applied to dense substances. Each atom interacts with the local electric field in which it is immersed. Yet unlike the isolated atoms considered above, those in a dense material will also experience the induced field set up by their brethren. Consequently, an atom "sees" in addition to the applied field $E(t)$ another field,* namely, $P(t)/3\epsilon_0$. Without going into the details here, it can be shown that

$$\frac{n^2 - 1}{n^2 + 2} = \frac{Nq_e^2}{3\epsilon_0 m_e} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 + i\gamma_j \omega} \tag{3.73}$$

Thus far we have been considering electron-oscillators almost exclusively, but the same results would have been applicable to ions bound to fixed atomic sites as well. In that instance m_e would be replaced by the considerably larger ion mass. Thus, although electronic polarization is important over the entire optical spectrum, the contributions from ionic polarization significantly affect n only in regions of resonance ($\omega_{0j} = \omega$).

The implications of a complex index of refraction will be considered later, in Section 4.8. At the moment we limit the discussion, for the most part, to situations in which absorption is negligible (i.e., $\omega_{0j}^2 - \omega^2 >> \gamma_j \omega$) and n is real, so that

$$\frac{n^2 - 1}{n^2 + 2} = \frac{Nq_e^2}{3\epsilon_0 m_e} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2} \tag{3.74}$$

Colorless, transparent materials have their characteristic frequencies outside the visible region of the spectrum (which is why they are, in fact, colorless and transparent). In particular, glasses have effective natural frequencies above the visible in the ultraviolet, where they become opaque. In cases for which $\omega_{0j}^2 >> \omega^2$, by comparison, ω^2 may be neglected in Eq. (3.74), yielding an essentially constant index of refraction over that frequency region. For example, the important characteristic frequencies for glasses occur at wavelengths of about 100 nm. The middle of the visible range is roughly five times that value, and there, $\omega_{0j}^2 >> \omega^2$. Notice that as ω increases toward ω_{0j}, $(\omega_{0j}^2 - \omega^2)$ decreases and *n gradually increases with frequency*, as is clearly evident in Fig. 3.40. This is called **normal dispersion**. In the ultraviolet region, as ω approaches a natural frequency, the oscillators will begin to resonate. Their amplitudes will increase markedly, and this will be accompanied by damping and a strong absorption of energy from the

*This result, which applies to isotropic media, is derived in almost any text on Electromagnetic Theory.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

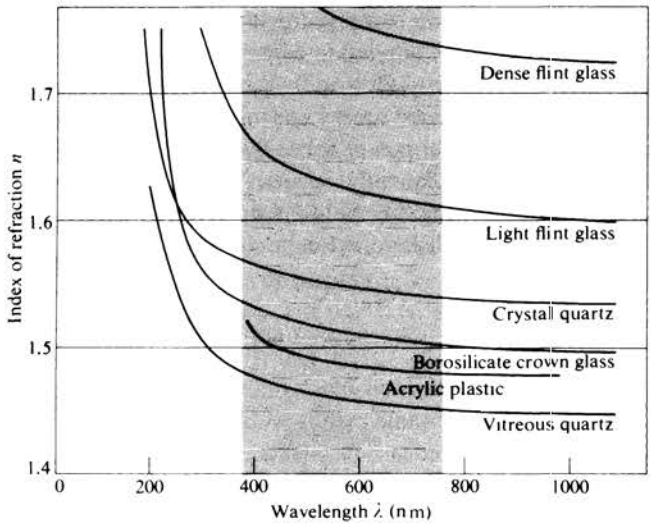


Figure 3.40 The wavelength dependence of the index of refraction for various materials.

incident wave. When $\omega_{0j} = \omega$ in Eq. (3.73), the damping term obviously becomes dominant. The regions immediately surrounding the various ω_{0j} in Fig. 3.41 are called **absorption bands**. There $dn/d\omega$ is negative, and the process is spoken of as **anomalous** (i.e., abnormal) **dispersion**. When white light passes through a glass prism, the blue constituent has a higher index than the red and is therefore deviated through a larger angle (see Section 5.5.1). In contrast, when we use a liquid-cell prism containing a dye solution with an absorption band in the visible, the spectrum is altered markedly (see Problem 3.50). All substances possess absorption bands somewhere within the electromagnetic frequency spectrum, so that the term *anomalous dispersion*, being a carryover from the late 1800s, is certainly a misnomer.

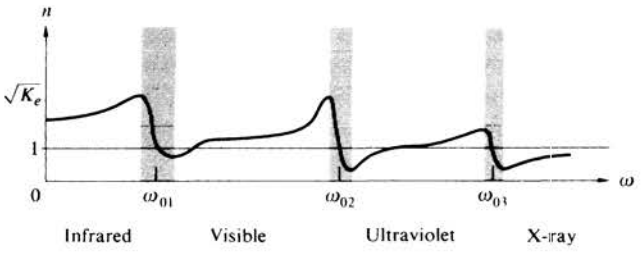


Figure 3.41 Refractive index versus frequency.

As we have seen, atoms within a molecule can also vibrate about their equilibrium positions. But the nuclei are massive, and so the natural oscillatory frequencies are low, in the infrared. Molecules such as H_2O and CO_2 have resonances in both the infrared and ultraviolet. When water is trapped within a piece of glass during its manufacture, these molecular oscillators are available, and an infrared absorption band exists. The presence of oxides also results in infrared absorption. Figure 3.42 shows the $n(\omega)$ curves (ranging from the ultraviolet to the infrared) for a number of important optical crystals. Note how they rise in the ultraviolet and fall in the infrared. At the even lower frequencies of radiowaves, glass is again transparent. In comparison, a piece of stained glass evidently has a resonance in the visible where it absorbs out a particular range of frequencies, transmitting the complementary color.

As a final point, notice that if the driving frequency is greater than any of the ω_{0j} terms, then $n^2 < 1$ and $n < 1$. Such a situation can occur, for example, if we beam X-rays onto a glass plate. This is an intriguing result, since it leads to $v > c$, in seeming contradiction to Special Relativity. We will consider this behavior again later on, when we discuss the group velocity (Section 7.6).

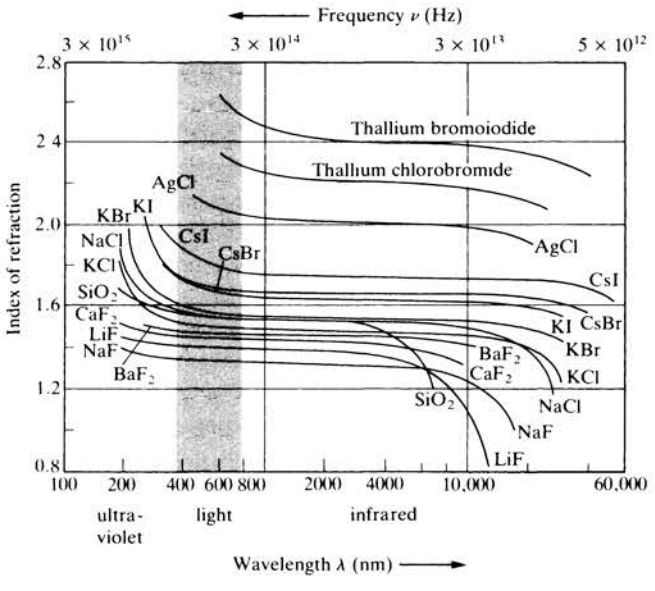


Figure 3.42 Index of refraction versus wavelength and frequency for several important optical crystals. (Adapted from data published by The Harshaw Chemical Co.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

In partial summary then, over the visible region of the spectrum, electronic polarization is the operative mechanism determining $n(\omega)$. Classically, one imagines electron-oscillators vibrating at the frequency of the incident wave. When the wave's frequency is appreciably different from a characteristic or natural frequency, the oscillations are small, and there is little dissipative absorption. At resonance, however, the oscillator amplitudes are increased, and the field does an increased amount of work on the charges. Electromagnetic energy removed from the wave and converted into mechanical energy is dissipated thermally within the substance, and one speaks of an absorption peak or band. The material, although essentially transparent at other frequencies, is fairly opaque to incident radiation at its characteristic frequencies (see photo).

3.6 The Electromagnetic-Photon Spectrum

In 1867, when Maxwell published the first extensive account of his Electromagnetic Theory, the frequency band was only known to extend from the infrared, across the visible, to the ultraviolet. Although this region is of major concern in Optics, it is a small segment of the vast electromagnetic spectrum (see Fig. 3.43). This section enumerates the main categories (there is actually some overlapping) into which the spectrum is usually divided.



A group of semiconductor lenses made from ZnSe, CdTe, GaAs, and Ge. These materials are particularly useful in the infrared (2 μm to 30 μm), where they are highly transparent despite the fact that they are quite opaque in the visible region of the spectrum. (Photo courtesy Two-Six Incorporated.)

3.6.1 Radiofrequency Waves

In 1887, eight years after Maxwell's death, Heinrich Hertz, then professor of physics at the Technische Hochschule in Karlsruhe, Germany, succeeded in generating and detecting electromagnetic waves.* His transmitter was essentially an oscillatory discharge across a spark gap (a form of oscillating electric dipole). For a receiving antenna, he used an open loop of wire with a brass knob on one end and a fine copper point on the other. A small spark visible between the two ends marked the detection of an incident electromagnetic wave. Hertz focused the radiation, determined its polarization, reflected and refracted it, caused it to interfere setting up standing waves, and then even measured its wavelength (on the order of a meter). As he put it:

> I have succeeded in producing distinct rays of electric force, and in carrying out with them the elementary experiments which are commonly performed with light and radiant heat. ... We may perhaps further designate them as rays of light of very great wavelength. The experiments described appear to me, at any rate, eminently adapted to remove any doubt as to the identity of light, radiant heat, and electromagnetic wave motion.

The waves used by Hertz are now classified in the *radiofrequency* range, which extends from a few hertz to about 10^9 Hz (λ, from many kilometers to 0.3 m or so). These are generally emitted by an assortment of electric circuits. For example, the 60-Hz alternating current circulating in power lines radiates with a wavelength of 5×10^6 m, or about 3×10^3 miles. There is no upper limit to the theoretical wavelength; one could leisurely swing the proverbial charged pith ball and, in so doing, produce a rather long, if not very strong, wave. Indeed, waves more than 18 million miles long have been detected streaming down toward Earth from outer space. The

*David Hughes may well have been the first person who actually performed this feat, but his experiments in 1879 went unpublished and unnoticed for many years.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

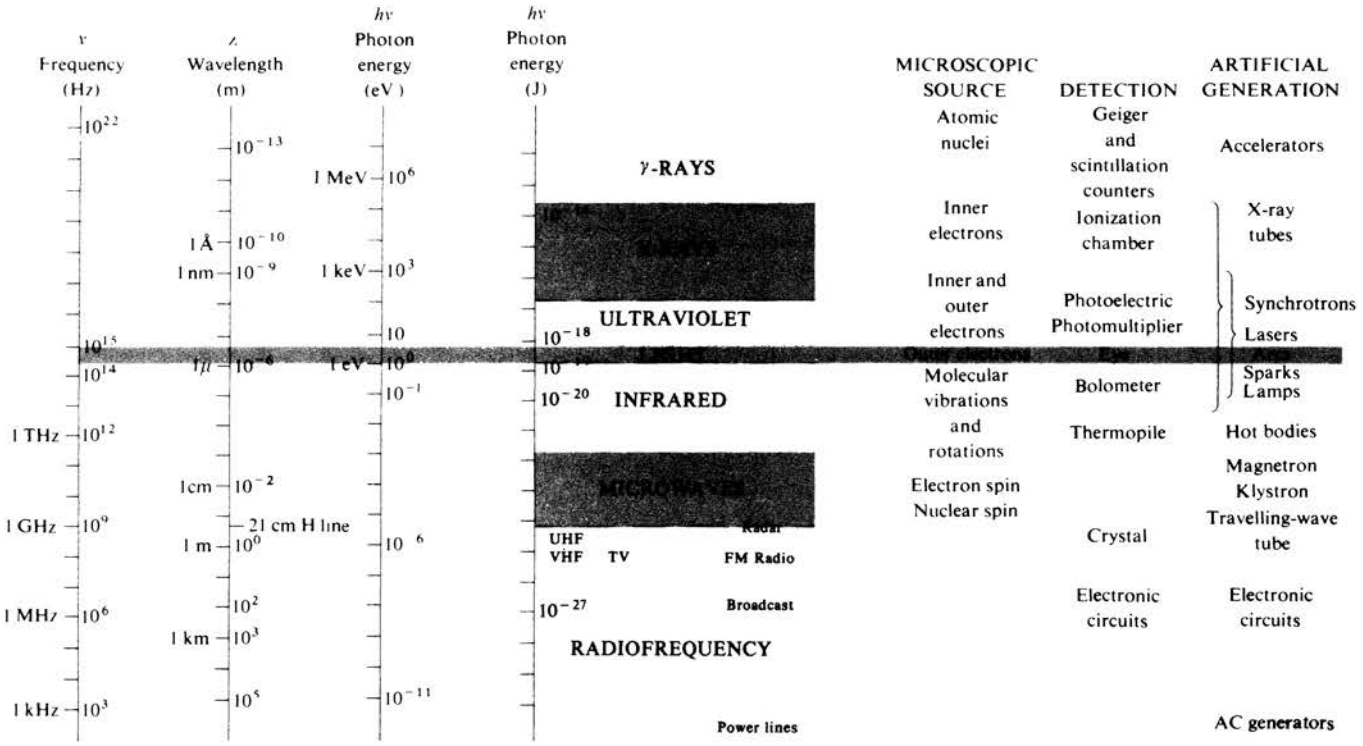


Figure 3.43 The electromagnetic-photon spectrum.

higher frequency end of the band is used for television and radio broadcasting.

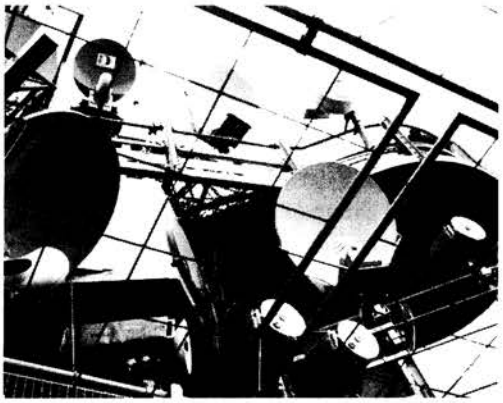
At 1 MHz (10^6 Hz), a radiofrequency photon has an energy of 6.62×10^{-28} J or 4×10^{-9} eV, a very small quantity by any measure. The granular nature of the radiation is generally obscured, and only a smooth transfer of radiofrequency energy is apparent.

3.6.2 Microwaves

The microwave region extends from about 10^9 Hz up to about 3×10^{11} Hz. The corresponding wavelengths go from roughly 30 cm to 1.0 mm. Radiation capable of penetrating the Earth's atmosphere ranges from less than 1 cm to about 30 m. Microwaves are therefore of interest in space-vehicle communications, as well as radio astronomy. In particular, neutral hydrogen atoms, distributed over vast regions of space, emit 21-cm (1420-MHz) microwaves. A good deal of information about the structure of our own and other galaxies has been gleaned from this particular emission.

Molecules can absorb and emit energy by altering the state of motion of their constituent atoms—they can be made to vibrate and rotate. Again, the energy associated with either motion is quantized, and molecules possess rotational and vibrational energy levels in addition to those due to their electrons. Only polar molecules will experience forces via the $\vec{E}$-field of an incident electromagnetic wave that will cause them to rotate into alignment, and only they can absorb a photon and make a rotational transition to an excited state. Since massive molecules are not able to swing around easily, we can anticipate that they will have low-frequency rotational resonances (far IR, 0.1 mm, to microwave, 1 cm). For instance, water molecules are polar (see Fig. 3.37), and if exposed to an electromagnetic wave, they will swing around, trying to stay lined up with the alternating $\vec{E}$-field. This will occur with particular vigor at any one of its rotational resonances. Consequently, water molecules efficiently and dissipatively absorb microwave radiation at or near such a frequency. The microwave oven (12.2 cm, 2.45 GHz) is an obvious application. On the other hand, nonpolar molecules, such as carbon dioxide, hydrogen, nitrogen, oxygen, and methane, cannot make rotational transitions by way of the absorption of photons.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Microwave antennae on the top of the Eiffel Tower in Paris. (Photo by E.H.)

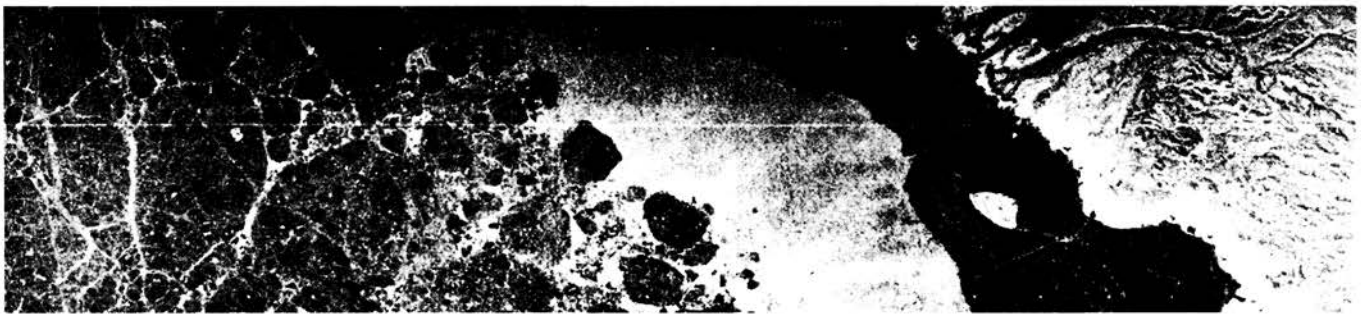


A picture of a candy bar made using T-rays. The nuts, which were hidden beneath the chocolate, are visible as a result of refraction. (Photo courtesy V. Rudd, Picometrix, Inc.)

Nowadays microwaves are used for everything from transmitting telephone conversations and interstation television to cooking hamburgers, from guiding planes and catching speeders (by radar) to studying the origins of the Universe, opening garage doors, and viewing the surface of the planet (see photo). They are also quite useful for studying Physical Optics with experimental arrangements that are scaled up to convenient dimensions.

Photons in the low-frequency end of the microwave spectrum have little energy, and one might expect their sources to be electric circuits exclusively. Emissions of this sort can, however, arise from atomic transitions, if the energy levels involved are quite near each other. The apparent ground state of the cesium atom is a good example. It is actually a pair of closely spaced energy levels, and transitions between them involve an energy of only 4.14×10^{-5} eV. The resulting microwave emission has a frequency of $9.192\ 631\ 77 \times 10^{9}$ Hz. This is the basis for the well-known cesium clock, the standard of frequency and time.

The range of radiation that straddles both microwaves and infrared (roughly 50 GHz to 10 THz) is often called terahertz radiation or T-rays. They're not absorbed by most dry, nonpolar materials such as plastic, paper, or fat. Water will absorb T-rays, and they're reflected by metals because of the free electrons. As a result, they can be used to image internal structure that would otherwise be hidden from view (see photo).



A photograph of an 18 by 75 mile area northeast of Alaska. It was taken by the *Seasat* satellite 800 kilometers (500 miles) above the Earth. The overall appearance is somewhat strange because this is actually a radar or microwave picture. The wrinkled gray region on the right is Canada. The small, bright shell shape is Banks Island, embedded in a black band of shore-fast, first-year sea ice. Adjacent to that is open water, which appears smooth and gray. The dark gray blotchy area at the far left is the main polar ice pack. There are no clouds because the radar "sees" right through them.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

3.6.3 Infrared

The infrared region, which extends roughly from 3×10^{11} Hz to about 4×10^{14} Hz, was first detected by the renowned astronomer Sir William Herschel (1738–1822) in 1800. As the name implies, this band of EM-radiation lies just beneath red light. The infrared, or IR, is often subdivided into four regions: the *near IR*, that is, near the visible (780–3000 nm); the *intermediate IR* (3000–6000 nm); the *far IR* (6000–15000 nm); and the *extreme IR* (15000 nm–1.0 mm). This is again a rather loose division, and there is no universality in the nomenclature. Radiant energy at the long-wavelength extreme can be generated by either microwave oscillators or incandescent sources (i.e., molecular oscillators). Indeed, any material will radiate and absorb IR via thermal agitation of its constituent molecules.

The molecules of any object at a temperature above absolute zero (−273°C) will radiate IR, even if only weakly (see Section 13.1.1). On the other hand, infrared is copiously emitted in a continuous spectrum from hot bodies, such as electric heaters, glowing coals, and ordinary house radiators. Roughly half the electromagnetic energy from the Sun is IR, and the common lightbulb actually radiates far more IR than light. Like all warm-blooded creatures, we too are infrared emitters. The human body radiates IR quite weakly, starting at around 3000 nm, peaking in the vicinity of 10 000 nm, and trailing off from there into the extreme IR and, negligibly, beyond. This emission is exploited by see-in-the-dark sniperscopes, as well as by some rather nasty "heat"-sensitive snakes (Crotalidae, pit vipers, and Boidae, constrictors) that tend to be active at night.

Besides rotating, a molecule can vibrate in several different ways, with its atoms moving in various directions with respect to one another. The molecule need not be polar, and even a linear system such as CO_2 has three basic vibrational modes and a number of energy levels, each of which can be excited by photons. The associated vibrational emission and absorption spectra are, as a rule, in the IR (1000 nm to 0.1 mm). Many molecules have both vibrational and rotational resonances in the IR and are good absorbers, which is one reason IR is often misleadingly called "heat waves"—just put your face in the sunshine and feel the resulting buildup of thermal energy.



Thermograph of the author. Note the cool beard and how far the hairline has receded since the first edition of this book. (Photo by E.H.)

Infrared radiant energy is generally measured with a device that responds to the heat generated on absorption of IR by a blackened surface. There are, for example, thermocouple, pneumatic (e.g., Golay cells), pyroelectric, and bolometer detectors. These in turn depend on temperature-dependent variations in induced voltage, gas volume, permanent electric polarization and resistance, respectively. The detector can be coupled by way of a scanning system to a cathode ray tube to produce an instantaneous televisionlike IR picture (see photo) known as a thermograph (which is quite useful for diagnosing all sorts of problems, from faulty transformers to faulty people). Photographic films sensitive to near IR (< 1300 nm) are also available. There are IR spy satellites that look out for rocket launchings, IR resource satellites that look out for crop diseases, and IR astronomical satellites that look out into space. There are "heat-seeking" missiles guided by IR, and IR lasers and telescopes peering into the heavens.

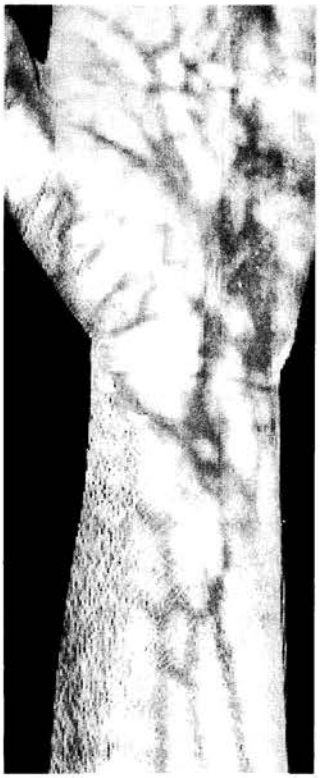
Small differences in the temperatures of objects and their surroundings result in characteristic IR emission that can be used in many ways, from detecting brain tumors and breast cancers to spotting a lurking burglar. The CO_2 laser, because it is a convenient source of continuous power at appreciable levels of 100 W and more, is widely used in industry, especially in precision cutting and heat treating. Its extreme-IR emissions (18.3 μm–23.0 μm) are readily absorbed by human tissue, making the laserbeam an effective bloodless scalpel that cauterizes as it cuts.

3.6.4 Light

Light corresponds to the electromagnetic radiation in the narrow band of frequencies from about 3.84×10^{14} Hz to roughly 7.69×10^{14} Hz (see Table 3.4). It is generally produced by a rearrangement of the outer electrons in atoms and molecules. (Don't forget synchrotron radiation, which is a different mechanism.)*

*There is no need here to define light in terms of human physiology. On the contrary, there is plenty of evidence to indicate that this would not be a very good idea. For example, see T. J. Wang, "Visual Response of the Human Eye to X Radiation." *Am. J. Phys.* **35**, 779 (1967).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



An arm viewed in a broad band of radiant energy extending from 468.5 nm (which is light) to 827.3 nm (which is near-infrared). The technique has many biomedical applications, among which is the early detection of skin cancer.

TABLE 3.4 Approximate Frequency and Vacuum Wavelength Ranges for the Various Colors

Color	λ_0 (nm)	ν (THz)*
Red	780–622	384–482
Orange	622–597	482–503
Yellow	597–577	503–520
Green	577–492	520–610
Blue	492–455	610–659
Violet	455–390	659–769

*1 terahertz (THz) = 10^{12} Hz, 1 nanometer (nm) = 10^{-9} m.

In an incandescent material, a hot, glowing metal filament, or the solar fireball, electrons are randomly accelerated and undergo frequent collisions. The resulting broad emission spectrum is called **thermal radiation**, and it is a major source of light. In contrast, if we fill a tube with some gas and pass an electric discharge through it, the atoms therein will become excited and radiate. The emitted light is characteristic of the particular energy levels of those atoms, and it is made up of a series of well-defined frequency bands or lines. Such a device is known as a gas discharge tube. When the gas is the krypton 86 isotope, the lines are particularly narrow (zero nuclear spin, therefore no hyperfine structure). The orange-red line of Kr 86, whose vacuum wavelength is 605.780 210 5 nm, has a width (at half height) of only 0.000 47 nm, or about 400 MHz. Accordingly, until 1983 it was the international standard of length (with 1 650 763.73 wavelengths equaling a meter).

Newton was the first to recognize that **white light** is actually a mixture of all the colors of the visible spectrum, that the prism does not create color by altering white light to different degrees, as had been thought for centuries, but simply fans out the light, separating it into its constituent colors. Not surprisingly, the very concept of *whiteness* seems dependent on our perception of the Earth's daylight spectrum—a broad frequency distribution that falls off more rapidly in the violet than in the red (Fig. 3.44). The human eye-brain detector perceives as white a wide mix of frequencies, usually with about the same amount of energy in each portion. That is what we mean when we speak about "white light"—much of the color of the spectrum, with no region predominating. Nonetheless, many different distributions will appear more or less white. We recognize a piece of paper to be white whether it's seen indoors under incandescent light or outside under skylight, even though those whites are quite different. In fact, there are many pairs of colored light beams (e.g., 656-nm red and 492-nm cyan) that will produce the sensation of whiteness, and the eye cannot always distinguish one white from another; it can-

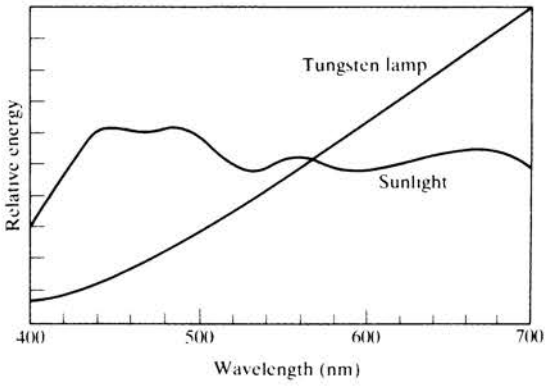


Figure 3.44 A graph of sunlight compared with the light from a tungsten lamp.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

not frequency analyze light into its harmonic components the way the ear can analyze sound (see Section 7.3).

Colors are the subjective human physiological and psychological responses, primarily, to the various frequency regions extending from about 384 THz for red, through orange, yellow, green, and blue, to violet at about 769 THz (Table 3.4). Color is not a property of the light itself but a manifestation of the electrochemical sensing system—eye, nerves, brain. To be more precise, we should not say "yellow light" but rather "light that is seen as yellow." Remarkably, a variety of different frequency mixtures can evoke the same color response from the eye-brain sensor. A beam of red light (peaking at, say, 430 THz) overlapping a beam of green light (peaking at, say, 540 THz) will result, believe it or not, in the perception of *yellow* light, even though no frequencies are actually present in the so-called yellow band. Apparently, the eye-brain averages the input and "sees" yellow (Section 4.9). That's why a color television screen can manage with only three phosphors: red, green, and blue.

In a flood of bright sunlight where the photon flux density might be 10^{21} photons/m^2·s, we can generally expect the quantum nature of the energy transport to be thoroughly masked. However, in very weak beams, since photons in the visible range ($h\nu \approx$ 1.6 eV to 3.2 eV) are energetic enough to produce effects on a distinctly individual basis, the granularity will become evident. Research on human vision indicates that as few as 10 light photons, and possibly even 1, may be detectable by the eye.

3.6.5 Ultraviolet

Adjacent to and just beyond light in the spectrum is the ultraviolet region (approximately 8×10^{14} Hz to about 3.4×10^{16} Hz), discovered by Johann Wilhelm Ritter (1776–1810). Photon energies therein range from roughly 3.2 eV to 100 eV. Ultraviolet, or UV, rays from the Sun will thus have more than enough energy to ionize atoms in the upper atmosphere and in so doing create the ionosphere. These photon energies are also of the order of the magnitude of many chemical reactions, and ultraviolet rays become important in triggering those reactions. Fortunately, ozone (O_3) in the atmosphere absorbs what would otherwise be a lethal stream of solar UV. At wavelengths less than around 290 nm, UV is germicidal (i.e., it kills microorganisms). The particlelike aspects of radiant energy become increasingly evident as the frequency rises.

Humans cannot see UV very well because the cornea absorbs it, particularly at the shorter wavelengths, while the eye lens absorbs most strongly beyond 300 nm. A person who has had a lens removed because of cataracts can see UV ($\lambda >$ 300 nm). In addition to insects, such as honeybees, a fair number of other creatures can visually respond to UV. Pigeons, for example, are capable of recognizing patterns illuminated by UV and probably employ that ability to navigate by the Sun even on overcast days.

An atom emits a UV photon when an electron makes a long jump down from a highly excited state. For example, the outermost electron of a sodium atom can be raised to higher and higher energy levels until it is ultimately torn loose altogether at 5.1 eV, and the atom is ionized. If the ion subsequently recombines with a free electron, the latter will rapidly descend to the ground state, most likely in a series of jumps, each resulting in the emission of a photon. It is possible, however, for the electron to make one long plunge to the ground state, radiating a single 5.1-eV UV photon. Even more energetic UV can be generated when the inner, tightly bound electrons of an atom are excited.

The unpaired valence electrons of isolated atoms can be an important source of colored light. But when these same atoms combine to form molecules or solids, the valence electrons are ordinarily paired in the process of creating the chemical bonds that hold the thing together. Consequently, the electrons are often more tightly bound, and their molecular-excited states are higher up in the UV. Molecules in the atmosphere, such as



An ultraviolet photograph of Venus taken by *Mariner 10*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

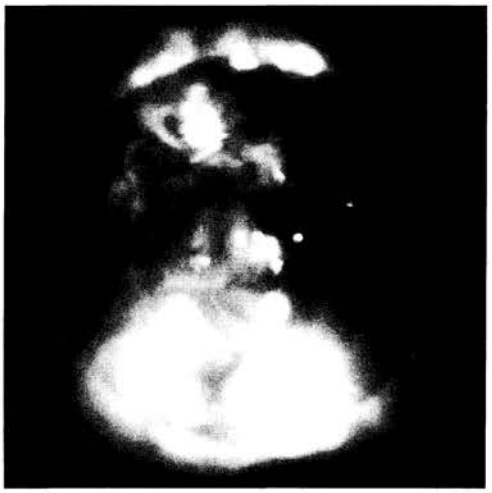
N_2, O_2, CO_2, and H_2O, have just this sort of electronic resonance in the UV.

Nowadays there are ultraviolet photographic films and microscopes, UV orbiting celestial telescopes, synchrotron sources, and ultraviolet lasers.

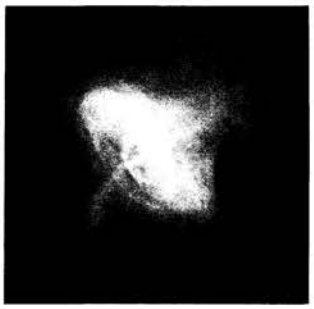
3.6.6 X-rays

X-rays were rather fortuitously discovered in 1895 by Wilhelm Conrad Röntgen (1845–1923). Extending in frequency from roughly 2.4×10^{16} Hz to 5×10^{19} Hz, they have extremely short wavelengths; most are smaller than an atom. Their photon energies (100 eV to 0.2 MeV) are large enough so that X-ray quanta can interact with matter one at a time in a clearly granular fashion, almost like bullets of energy. One of the most practical mechanisms for producing X-rays is the rapid deceleration of high-speed charged particles. The resulting broad-frequency *bremsstrahlung* (German for "braking radiation") arises when a beam of energetic electrons is fired at a material target, such as a copper plate. Collisions with the Cu nuclei produce deflections of the beam electrons, which in turn radiate X-ray photons.

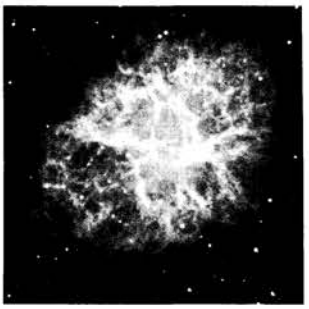
In addition, the atoms of the target may become ionized during the bombardment. Should that occur through removal of an inner electron strongly bound to the nucleus, the atom will emit X-rays as the electron cloud returns to the ground state. The resulting quantized emissions are specific to the target atom, revealing its energy-level structure, and accordingly are called *characteristic* radiation.



An early X-ray photograph of the Sun taken March 1970. The limb of the Moon is visible in the southeast corner. (Photo Courtesy Dr. G. Vaiana and NASA.)



This amazingly detailed X-ray image of the Crab Nebula (see p. 61) was recently taken by the orbiting Chandra X-Ray Observatory. The picture reveals the locations of the most energetic particles in the pulsar.



An optical image of the Crab Nebula. The light forming this picture comes from particles of intermediate energy. The filaments are due to hot gases at temperatures of tens of thousands of degrees.

Traditional medical film-radiography generally produces little more than simple shadow castings, rather than photographic images in the usual sense; it has not been possible to fabricate useful X-ray lenses. But modern focusing methods using mirrors (see Section 5.4) have begun an era of X-ray imagery, creating detailed pictures of all sorts of things, from imploding fusion pellets to celestial sources, such as the Sun (see photo), distant pulsars, and black holes—objects at temperatures of millions of degrees that emit predominantly in the X-ray region. Orbiting X-ray telescopes have given us an exciting new eye on the Universe (see photo). There are X-ray microscopes, picosecond X-ray streak cameras, X-ray diffraction gratings, and interferometers, and work continues on X-ray holography. In 1984 a group at the Lawrence Livermore National Laboratory succeeded in producing laser radiation at a wavelength of 20.6 nm. Although this is more accurately in the extreme ultraviolet (XUV), it's close enough to the X-ray region to qualify as the first soft X-ray laser.

3.6.7 Gamma Rays

Gamma rays are the highest energy (10^4 eV to about 10^{19} eV), shortest wavelength electromagnetic radiations. They are emitted by particles undergoing transitions within the atomic nucleus. A single gamma-ray photon carries so much energy

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

that it can be detected with little difficulty. At the same time, its wavelength is so small that it is now extremely difficult to observe any wavelike properties.

We have gone full cycle from the radiofrequency wavelike response to gamma-ray particlelike behavior. Somewhere, not far from the (logarithmic) center of the spectrum, there is light. As with all electromagnetic radiation, its energy is quantized, but here in particular what we "see" will depend on how we "look."

3.7 Quantum Field Theory

A charged particle exerts forces on other charged particles. It creates a web of electromagnetic interaction around itself that extends out into space. That imagery leads to the concept of the electric field, which is a representation of the way the electromagnetic interaction reveals itself on a macroscopic level. The static electric field is, in effect, a spatial conception summarizing the interaction among charges. Through Faraday's vision, the idea of the field was extended, and it became appropriate to imagine that one charge sets up an $\vec{\mathbf{E}}$-field in space and another charge, immersed in that field, interacts directly with it, and vice versa. What began as a mapping of the force distribution (whatever its cause) became a thing, a field, capable itself of exerting force. Still the picture seems straightforward, even if many questions come to mind. Does the static $\vec{\mathbf{E}}$-field have a physical reality in-and-of itself? If it does, does it fill space with energy and how exactly does that happen? Is anything actually flowing? How does the field produce a force on a charge? Does it take time to exert its influence?

Once the electromagnetic field became a reality, physicists could imagine disturbances of that tenuous medium which so conveniently spans the void of space; ***light was an electromagnetic wave in the electromagnetic field***. Although it's easy enough to envision a wave sweeping through an existing field (p. 59), it's not so obvious how a localized pulse launched into space, like the one shown in Fig. 3.45, might be conceptualized. There is no static field filling space, extending out in front of the pulse; if the pulse advances through the medium of the electromagnetic field, it must first create that medium itself as it progresses. That's not impossible to imagine on some level, but it's hardly what one would call a classical wave. For any traditional wave, a medium in equilibrium is the fundamental starting point; it exists at any location before and after the wave passes. So this idea of an electromagnetic wave, which is so beautiful mathematically, is not quite so transparent conceptually.

As early as 1905, Einstein already considered the classical equations of Electromagnetic Theory to be descriptions of the average values of the quantities being considered. "To me it seems absurd," he wrote to Planck, "to have energy continuously distributed in space without assuming an aether. . . . While Faraday's representation was useful in the development of electrodynamics, it does not follow in my opinion that this view must be maintained in all its detail." Classical theory wonderfully accounted for everything being measured, but it was oblivious to the exceedingly fine granular structure of the phenomenon. Using thermodynamic arguments, Einstein proposed that electric and magnetic fields were quantized, that they are particulate rather than continuous. After all, classical theory evolved decades before the electron was even discovered. If charge (the fundamental source of electromagnetism) is quantized, shouldn't the theory reflect that in some basic way?

Today, we are guided by Quantum Mechanics, a highly mathematical theory that provides tremendous computational and predictive power but is nonetheless disconcertingly abstract. In particular, the subdiscipline that treats microparticles and their interactions, Quantum Field Theory (QFT), in its various forms, is the most fundamental and arguably the most successful of all physical theories. Light quanta come out of the theory in a completely natural way by quantizing the electromagnetic field. The apparent implication of this is

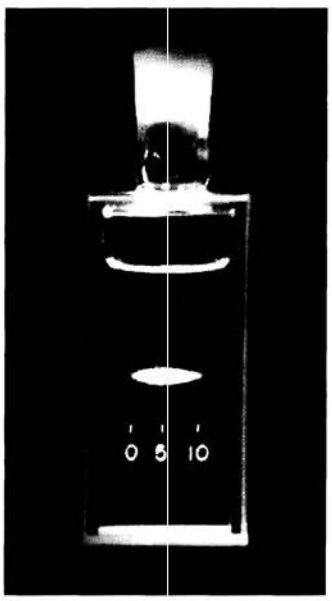


Figure 3.45 An ultrashort pulse of green light from a neodymium-doped glass laser. The pulse passed through a water cell whose wall is marked in millimeters. During the 10-picosecond exposure, the pulse moved about 2.2 mm. (Photo courtesy Bell Laboratories.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

that all microparticles originate in the same way from their own individual fields: **the field's the thing**, as it were. Thus the electron is the quantum of the electron field, the proton is the quantum of the proton field, and so forth. Filling in the details has been the business of field theorists for much of this century.

There are two distinct philosophical currents in contemporary QFT: the field-centered and the particle-centered. In the field-centered view, ***fields are the fundamental entities, particles are just the quanta of the fields***. In the particle-centered view, ***particles are the fundamental entities, fields are just the macroscopic coherent states of particles***. The field tradition goes back to L. de Broglie (1923), E. Schrödinger, P. Jordan, and W. Pauli, whose research laid the foundations of the quantum-mechanical variant sometimes called Wave Mechanics. The particle tradition began with the early work of W. Heisenberg (1925), although its spiritual mentor was P.A.M. Dirac who set the particle agenda with his theory of the electron-positron pair. The particular offshoot of QFT that strives to provide a relativistic quantum-mechanical treatment of the electromagnetic interaction is called Quantum Electrodynamics (QED), and it too has its particle-centered and field-centered proponents. Some of the basic ideas of QED have been made accessible on this level by R. P. Feynman, and insofar as they illuminate Optics we'll explore them later in this text (p. 137).

Contemporary physics by way of QFT holds that all fields are quantized; that each of the fundamental Four Forces (Gravitational, Electromagnetic, Strong, and Weak) is mediated by a special kind of field particle. These *messenger bosons* are continuously absorbed and emitted by the interacting material particles (electrons, protons, etc.). This ongoing exchange *is* the interaction. The mediating particle of the electric field is the **virtual photon**. This massless messenger travels at the speed of light and transports momentum and energy. When two electrons repel one another, or an electron and proton attract, it is by emitting and absorbing virtual photons and thereby transferring momentum from one to the other, that transfer being a measure of the action of force. The messenger particles of the electromagnetic force are called *virtual* photons because they are bound to the interaction. Virtual photons can never escape to be detected directly by some instrument, however unsettling that is philosophically and however hard that makes it to establish their existence. Indeed, virtual photons (as distinct from real photons) exist only as the means of interaction. They are creatures of theory whose metaphysical status is yet to be determined.*

On a macroscopic level, messenger particles can manifest themselves as a continuous field provided they can group in very large numbers. Fundamental particles have an intrinsic angular momentum, or *spin*, that determines their grouping characteristics. Quantum Theory tells us that the desired field behavior can occur only if forces are mediated by messenger particles having angular momenta equal to integer multiples of $h/2\pi$ (i.e., 0, $1h/2\pi$, $2h/2\pi$, $3h/2\pi$, ...). The angular momentum of the virtual photon is $1(h/2\pi)$; it's a spin-1 particle. The exceedingly important class of interactions that have spin-1 messengers are known as **gauge forces**, and the electromagnetic force is the model for all the gauge forces. Today, the magic of action-at-a-distance is understood via the no less mysterious exchange of virtual particles, but at least now a highly predictive mathematical theory is in place that describes the phenomenon.

*For a discussion of the issues being struggled with, see H. R. Brown and R. Harré, *Philosophical Foundations of Quantum Field Theory*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

PROBLEMS

Complete solutions to all problems—except those with an asterisk—can be found in the back of the book.

3.1 Consider the plane electromagnetic wave (in SI units) given by the expressions $E_x = 0$, $E_y = 2\cos[2\pi \times 10^{14}(t - x/c) + \pi/2]$, and $E_z = 0$.

(a) What are the frequency, wavelength, direction of motion, amplitude, initial phase angle, and polarization of the wave?

(b) Write an expression for the magnetic flux density.

3.2 Write an expression for the $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields that constitute a plane harmonic wave traveling in the $+z$-direction. The wave is linearly polarized with its plane of vibration at 45° to the yz-plane.

3.3* Considering Eq. (3.30), show that the expression

$$\vec{\mathbf{k}} \times \vec{\mathbf{E}} = \omega \vec{\mathbf{B}}$$

is correct as it applies to a plane wave for which the direction of the electric field is constant.

3.4* Imagine an electromagnetic wave with its $\vec{\mathbf{E}}$-field in the y-direction. Show that Eq. (3.27)

$$\frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t}$$

applied to the harmonic wave $\vec{\mathbf{B}}$

$$\vec{\mathbf{E}} = \vec{\mathbf{E}}_0 \cos(kx - \omega t) \qquad \vec{\mathbf{B}} = \vec{\mathbf{B}}_0 \cos(kx - \omega t)$$

yields the fact that

$$E_0 = cB_0$$

in agreement with Eq. (3.30).

3.5* An electromagnetic wave is specified (in SI units) by the following function:

$$\vec{\mathbf{E}} = (-6\hat{\mathbf{i}} + 3\sqrt{5}\hat{\mathbf{j}})(10^4 \text{ V/m})e^{i[(4.5 \times 10^7)\pi \times 10^7 - 9.42 \times 10^{15}t]}$$

Find (a) the direction along which the electric field oscillates, (b) the scalar value of amplitude of the electric field, (c) the direction of propagation of the wave, (d) the propagation number and wavelength, (e) the frequency and angular frequency, and (f) the speed.

3.6 The electric field of an electromagnetic wave traveling in the positive x-direction is given by

$$\vec{\mathbf{E}} = E_0\hat{\mathbf{j}} \sin\frac{\pi z}{z_0} \cos(kx - \omega t)$$

(a) Describe the field verbally. (b) Determine an expression for k. (c) Find the phase speed of the wave.

3.7* A 550-nm harmonic EM-wave whose electric field is in the z-direction is traveling in the y-direction in vacuum. (a) What is the frequency of the wave? (b) Determine both ω and k for this wave. (c) If the electric field amplitude is 600 V/m, what is the amplitude of the magnetic field? (d) Write an expression for both $E(t)$ and $B(t)$ given that each is zero at $x = 0$ and $t = 0$. Put in all the appropriate units.

3.8* Calculate the energy input necessary to charge a parallel plate capacitor by carrying charge from one plate to the other. Assume the energy is stored in the field between the plates and compute the energy per unit volume, u_E, of that region, that is, Eq. (3.31). *Hint:* since the electric field increases throughout the process, either integrate or use its average value $E/2$.

3.9* Starting with Eq. (3.32), prove that the energy densities of the electric and magnet fields are equal ($u_E = u_B$) for an electromagnetic wave.

3.10 The time average of some function $f(t)$ taken over an interval T is given by

$$\langle f(t)\rangle_T = \frac{1}{T}\int_t^{t+T} f(t')dt'$$

where t' is just a dummy variable. If $\tau = 2\pi/\omega$ is the period of a harmonic function, show that $\vec{\mathbf{r}}$

$$\langle \sin^2(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t)\rangle = \tfrac{1}{2}$$

$$\langle \cos^2(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t)\rangle = \tfrac{1}{2}$$

and

$$\langle \sin(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t)\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t)\rangle = 0$$

when $T = \tau$ and when $T >> \tau$.

3.11* Show that a more general formulation of the previous problem yields

$$\langle \cos^2\omega t\rangle_T = \tfrac{1}{2}[1 + \text{sinc}\,\omega T \cos 2\omega t]$$

for any interval T.

3.12* With the previous problem in mind, prove that

$$\langle \sin^2\omega t\rangle_T = \tfrac{1}{2}[1 - \text{sinc}\,\omega T \cos 2\omega t]$$

for any interval T.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

3.13* Prove that the irradiance of a harmonic EM-wave is given by

$$I = \frac{1}{2c\mu_0} E_0^2$$

and then determine the average rate at which energy is transported per unit area by a plane wave having an amplitude of 15.0 V/m.

3.14* A light bulb puts out 20 W of radiant energy (most of it IR). Assume it to be a point source and calculate the irradiance 1.00 m away.

3.15* Consider a linearly polarized plane electromagnetic wave traveling in the $+x$-direction in free space having as its plane of vibration the xy-plane. Given that its frequency is 10 MHz and its amplitude is $E_0 = 0.08$ V m,

(a) Find the period and wavelength of the wave.

(b) Write an expression for $E(t)$ and $B(t)$.

(c) Find the flux density, $\langle S \rangle$, of the wave.

3.16* On average, the net electromagnetic power radiated by the Sun, its so-called *luminosity* (L), is 3.9×10^{26} W. Determine the mean amplitude of the electric field due to all the radiant energy arriving at the top of Earth's atmosphere (1.5×10^{11} m from the Sun).

3.17 A linearly polarized harmonic plane wave with a scalar amplitude of 10 V m is propagating along a line in the xy-plane at 45° to the x-axis with the xy-plane as its plane of vibration. Please write a vector expression describing the wave assuming both k_x and k_y are positive. Calculate the flux density taking the wave to be in vacuum.

3.18 Pulses of UV lasting 2.00 ns each are emitted from a laser that has a beam of diameter 2.5 mm. Given that each burst carries an energy of 6.0 J, (a) determine the length in space of each wavetrain, and (b) find the average energy per unit volume for such a pulse.

3.19* A laser provides pulses of EM-radiation in vacuum lasting 10^{-12} s. If the radiant flux density is 10^{20} W m^2, determine the amplitude of the electric field of the beam.

3.20 A 1.0-mW laser has a beam diameter of 2 mm. Assuming the divergence of the beam to be negligible, compute its energy density in the vicinity of the laser.

3.21* A cloud of locusts having a density of 100 insects per cubic meter is flying north at a rate of 6 m min. What is the flux density of locusts? That is, how many cross an area of 1 m^2 perpendicular to their flight path per second?

3.22 Imagine that you are standing in the path of an antenna that is radiating plane waves of frequency 100 MHz and flux density 19.88×10^{-2} W m^2. Compute the photon flux density, that is, the number of photons per unit time per unit area. How many photons, on the average, will be found in a cubic meter of this region?

3.23* How many photons per second are emitted from a 100-W yellow lightbulb if we assume negligible thermal losses and a quasimonochromatic wavelength of 550 nm? In actuality only about 2.5% of the total dissipated power emerges as visible radiation in an ordinary 100-W lamp.

3.24 A 3.0-V flashlight bulb draws 0.25 A, converting about 1.0% of the dissipated power into light ($\lambda \approx 550$ nm). If the beam has a cross-sectional area of 10 cm^2 and is approximately cylindrical,

(a) How many photons are emitted per second?

(b) How many photons occupy each meter of the beam?

(c) What is the flux density of the beam as it leaves the flashlight?

3.25* An isotropic quasimonochromatic point source radiates at a rate of 100 W. What is the flux density at a distance of 1 m? What are the amplitudes of the $\vec{\mathbf{E}}$- and $\vec{\mathbf{B}}$-fields at that point?

3.26 Using energy arguments, show that the amplitude of a cylindrical wave must vary inversely with $\sqrt{r}$. Draw a diagram indicating what's happening.

3.27* What is the momentum of a 10^{19}-Hz X-ray photon?

3.28 Consider an electromagnetic wave impinging on an electron. It is easy to show kinematically that the average value of the time rate-of-change of the electron's momentum $\vec{\mathbf{p}}$ is proportional to the average value of the time rate-of-change of the work, W, done on it by the wave. In particular,

$$\left\langle \frac{d\vec{\mathbf{p}}}{dt} \right\rangle = \frac{1}{c} \left\langle \frac{dW}{dt} \right\rangle \hat{\mathbf{i}}$$

Accordingly, if this momentum change is imparted to some completely absorbing material, show that the pressure is given by Eq. (3.51).

3.29* Derive an expression for the radiation pressure when the normally incident beam of light is totally reflected. Generalize this result to the case of oblique incidence at an angle θ with the normal.

3.30 A completely absorbing screen receives 300 W of light for 100 s. Compute the total linear momentum transferred to the screen.

3.31 The average magnitude of the Poynting vector for sunlight arriving at the top of Earth's atmosphere (1.5×10^{11} m from the Sun) is about 1.4 kW m^2.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

(a) Compute the average radiation pressure exerted on a metal reflector facing the Sun.

(b) Approximate the average radiation pressure at the surface of the Sun whose diameter is 1.4×10^9 m.

3.32* A surface is placed perpendicular to a beam of light of constant irradiance (I). Suppose that the fraction of the irradiance absorbed by the surface is α. Show that the pressure on the surface is given by

$$\mathcal{P} = (2 - \alpha)I/c$$

3.33* A light beam with an irradiance of 2.00×10^6 W/m^2 impinges normally on a surface that reflects 70.0% and absorbs 30.0%. Compute the resulting radiation pressure on the surface.

3.34 What force on the average will be exerted on the (40 m × 50 m) flat, highly reflecting side of a space station wall if it's facing the Sun while orbiting Earth?

3.35 A parabolic radar antenna with a 2-m diameter transmits 200-kW pulses of energy. If its repetition rate is 500 pulses per second, each lasting 2 μs, determine the average reaction force on the antenna.

3.36 Consider the plight of an astronaut floating in free space with only a 10-W lantern (inexhaustibly supplied with power). How long will it take to reach a speed of 10 m/s using the radiation as propulsion? The astronaut's total mass is 100 kg.

3.37 Consider the uniformly moving charge depicted in Fig. 3.26*b*. Draw a sphere surrounding it and show via the Poynting vector that the charge does not radiate.

3.38* A plane, harmonic, linearly polarized light wave has an electric field intensity given by

$$E_z = E_0 \cos\pi\, 10^{15}\left(t - \frac{x}{0.65c}\right)$$

while traveling in a piece of glass. Find

(a) The frequency of the light.

(b) Its wavelength.

(c) The index of refraction of the glass.

3.39* What is the speed of light in diamond if the index of refraction is 2.42?

3.40* Given that the wavelength of a lightwave in vacuum is 540 nm, what will it be in water, where $n = 1.33$?

3.41* Determine the index of refraction of a medium if it is to reduce the speed of light by 10% as compared to its speed in vacuum?

3.42 If the speed of light (the phase speed) in Fabulite ($SrTiO_3$) is 1.245×10^8 m s, what is its index of refraction?

3.43* What is the distance that yellow light travels in water (where $n = 1.33$) in 1.00 s?

3.44* A 500-nm lightwave in vacuum enters a glass plate of index 1.60 and propagates perpendicularly across it. How many waves span the glass if it's 1.00 cm thick?

3.45* Yellow light from a sodium lamp ($\lambda_0 = 589$ nm) traverses a tank of glycerin (of index 1.47), which is 20.0 m long, in a time t_1. If it takes a time t_2 for the light to pass through the same tank when filled with carbon disulfide (of index 1.63), determine the value of $t_2 - t_1$.

3.46* A lightwave travels from point A to point B in vacuum. Suppose we introduce into its path a flat glass plate ($n_g = 1.50$) of thickness $L = 1.00$ mm. If the vacuum wavelength is 500 nm, how many waves span the space from A to B with and without the glass in place? What phase shift is introduced with the insertion of the plate?

3.47 The low-frequency relative permittivity of water varies from 88.00 at 0°C to 55.33 at 100°C. Explain this behavior. Over the same range in temperature, the index of refraction ($\lambda = 589.3$ nm) goes from roughly 1.33 to 1.32. Why is the change in n so much smaller than the corresponding change in K_E?

3.48 Show that for substances of low density, such as gases, which have a single resonant frequency ω_0, the index of refraction is given by

$$n \approx 1 + \frac{Nq_e^2}{2\epsilon_0 m_e(\omega_0^2 - \omega^2)}$$

3.49* In the next chapter, Eq. (4.47), we'll see that a substance reflects radiant energy appreciably when its index differs most from the medium in which it is embedded.

(a) The dielectric constant of ice measured at microwave frequencies is roughly 1, whereas that for water is about 80 times greater—why?

(b) How is it that a radar beam easily passes through ice but is considerably reflected when encountering a dense rain?

3.50 Fuchsin is a strong (aniline) dye, which in solution with alcohol has a deep red color. It appears red because it absorbs the green component of the spectrum. (As you might expect, the surfaces of crystals of fuchsin reflect green light rather strongly.) Imagine that you have a thin-walled hollow prism filled with this solution. What will the spectrum look like for incident white light? By the way, anomalous dispersion was first observed in about 1840 by Fox Talbot, and the effect was christened in 1862 by Le Roux. His work was

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

promptly forgotten, only to be rediscovered eight years later by C. Christiansen.

3.51* Take Eq. (3.71) and check out the units to make sure that they agree on both sides.

3.52 The resonant frequency of lead glass is in the UV fairly near the visible, whereas that for fused silica is far into the UV. Use the dispersion equation to make a rough sketch of n versus ω for the visible region of the spectrum.

3.53* Show that Eq. (3.70) can be rewritten as

$$(n^2 - 1)^{-1} = -C\lambda^{-2} + C\lambda_0^{-2}$$

where $C = 4\pi^2 c^2 \epsilon_0 m_e / N q_e^2$.

3.54 Augustin Louis Cauchy (1789–1857) determined an empirical equation for $n(\lambda)$ for substances that are transparent in the visible. His expression corresponded to the power series relation

$$n = C_1 + C_2/\lambda^2 + C_3/\lambda^4 + \cdots$$

where the Cs are all constants. In light of Fig. 3.41, what is the physical significance of C_1?

3.55 Referring to the previous problem, realize that there is a region between each pair of absorption bands for which the Cauchy Equation (with a new set of constants) works fairly well. Examine Fig. 3.41: what can you say about the various values of C_1 as ω decreases across the spectrum? Dropping all but the first two terms, use Fig. 3.40 to determine approximate values for C_1 and C_2 for borosilicate crown glass in the visible.

3.56* Crystal quartz has refractive indexes of 1.557 and 1.547 at wavelengths of 410.0 nm and 550.0 nm, respectively. Using only the first two terms in Cauchy's Equation, calculate C_1 and C_2 and determine the index of refraction of quartz at 610.0 nm.

3.57* In 1871 Sellmeier derived the equation

$$n^2 = 1 + \sum_j \frac{A_j \lambda^2}{\lambda^2 - \lambda_{0j}^2}$$

where the A_j terms are constants and each λ_{0j} is the vacuum wavelength associated with a natural frequence ν_{0j}, such that $\lambda_{0j}\nu_{0j} = c$. This formulation is a considerable practical improvement over the Cauchy Equation. Show that where $\lambda >> \lambda_{0j}$, Cauchy's Equation is an approximation of Sellmeier's. *Hint:* Write the above expression with only the first term in the sum; expand it by the binomial theorem; take the square root of n^2 and expand again.

3.58* If an ultraviolet photon is to dissociate the oxygen and carbon atoms in the carbon monoxide molecule, it must provide 11 eV of energy. What is the minimum frequency of the appropriate radiation?

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

4 The Propagation of Light

4.1 Introduction

Our present concern is with the basic phenomena of *transmission* (p. 92), *reflection* (p. 95), and *refraction* (p. 100). These will be described classically in two ways: first, via the general notions of waves and rays (p. 104) and then from the more specific perspective of Electromagnetic Theory (p. 111). After that, we'll turn to a highly simplified treatment of Quantum Electrodynamics (QED) for a modern interpretation of what's happening (p. 139).

Most students have already studied these fundamental propagation phenomena in some introductory way and found ideas like the Laws of Reflection and Refraction to be straightforward and simple. But that's only because such treatments are from a macroscopic perspective that tends to be misleadingly superficial. For instance, reflection, which looks as obvious as light "bouncing off a surface," is a wonderfully subtle affair usually involving the coordinated behavior of countless atoms. The more deeply we explore these processes, the more challenging they become. Beyond that, many fascinating questions need to be addressed: How does light move through a material medium? What happens to it as it does? Why does light appear to travel at a speed other than c when photons can only exist at c?

Each encounter of light with bulk matter can be viewed as a cooperative event arising when a stream of photons sails through, and interacts with, an array of atoms suspended (via electromagnetic fields) in the void. The details of that journey determine why the sky is blue and blood is red, why your cornea is transparent and your hand opaque, why snow is white and rain is not. At its core, this chapter is about **scattering**, in particular, the absorption and prompt re-emission of EM-radiation by electrons associated with atoms and molecules. ***The processes of transmission, reflection, and refraction are macroscopic manifestations of scattering occurring on a submicroscopic level.***

To begin the analysis, let's first consider the propagation of radiant energy through various homogeneous media.

4.2 Rayleigh Scattering

Imagine a narrow beam of sunlight having a broad range of frequencies advancing through empty space. As it progresses, the beam spreads out very slightly, but apart from that, all the energy continues forward at c. There is no scattering, and the beam cannot be seen from the side. Nor does the light tire or diminish in any way. When a star in a nearby galaxy 1.7×10^5 light-years away was seen to explode in 1987, the flash of light that reached Earth had been sailing through space for 170 000 years before it got here. ***Photons are timeless***.

Now, suppose we mix a wisp of air into the void—some molecules of nitrogen, oxygen, and so forth. These molecules have no resonances in the visible, no one of them can be raised into an excited state by absorbing a quantum of light, and the gas is therefore transparent. Instead, each molecule behaves as a little oscillator whose electron cloud can be driven into a ground-state vibration by an incoming photon. Immediately upon being set vibrating, the molecule initiates the re-emission of light. A photon is absorbed, and without delay another photon of the same frequency (and wavelength) is emitted; the light is *elastically scattered*. The molecules are randomly oriented, and photons scatter out every which way (Fig. 4.1). Even when the light is fairly dim, the number of photons is immense, and it looks as if the molecules are scattering little classical spherical wavelets (Fig. 4.2)—energy streams out in every direction. Still, the scattering process is quite weak and

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

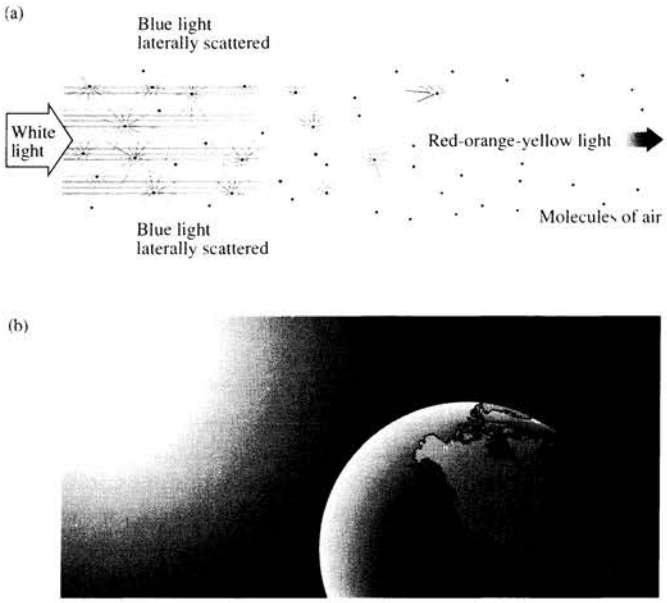


Figure 4.1 (*a*) Sunlight traversing a region of widely spaced air molecules. The light laterally scattered is mostly blue, and that's why the sky is blue. The unscattered light, which is rich in red, is viewed only when the Sun is low in the sky at sunrise and sunset. (*b*) Solar rays reach about 18° beyond the daytime terminator because of atmospheric scattering. Over this twilight band the skylight fades to the complete darkness of night.

the gas tenuous, so the beam is very little attenuated unless it passes through a tremendous volume of air.

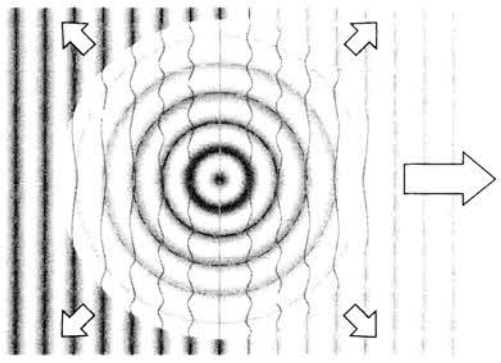


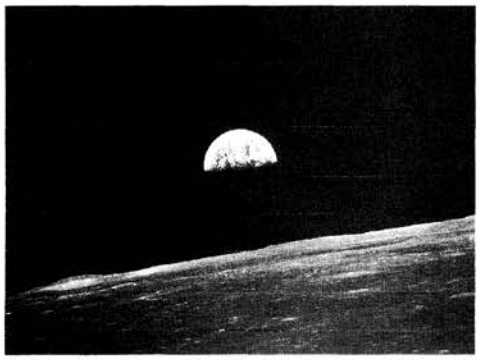
Figure 4.2 A plane wave, incident from the left, sweeps across an atom and spherical wavelets are scattered. The process is continuous, and hundreds of millions of photons per second stream out of the scattering atom in all directions.

The amplitudes of these ground-state vibrations, and therefore the amplitudes of the scattered light, increase with frequency because all the molecules have electronic resonances in the UV. The closer the driving frequency is to a resonance, the more vigorously the oscillator responds. So, violet light is strongly scattered laterally out of the beam, as is blue to a slightly lesser degree, as is green to a considerably lesser degree, as is yellow to a still lesser degree, and so on. The beam that traverses the gas will thus be richer in the red end of the spectrum, while the light scattered out (sunlight not having very much violet in it in the first place) will be richer in blue. That, in part, is why the sky is blue.

Long before Quantum Mechanics, Lord Rayleigh (1871) analyzed scattered sunlight in terms of molecular oscillators. Using a simple argument based on dimensional analysis (see Problem 4.1), he correctly concluded that the intensity of the scattered light was proportional to $1/\lambda^4$ and therefore increases with ν^4. Before this work, it was widely believed that the sky was blue because of scattering from minute dust particles. Since that time, scattering involving particles smaller than a wavelength (i.e., less than about $\lambda/15$) has been referred to as Rayleigh Scattering. Atoms and ordinary molecules fit the bill since they are a few tenths of a nanometer in diameter, whereas light has a wavelength of around 500 nm. A human's blue eyes, a bluejay's feathers, the blue-tailed skinks's blue tail, and the baboon's blue buttocks are all colored via Rayleigh Scattering. Indeed, in the animal kingdom scattering is the cause of almost all the blue, much of the green, and even some of the purple coloration. Scattering from the tiny alveolar cells in the barbs of the jay's feathers make it blue, whereas a parrot's green is a blend of yellow arising from preferential absorption (p. 134) and blue via scattering.

As we will see in a moment, a dense uniform substance will not appreciably scatter laterally, and that applies to much of the lower atmosphere. After all, if blue light were strongly scattered out at sea level, a far-off mountain would appear reddish and that's not the case even over distances of tens of kilometers. Even in the middle regions of the atmosphere, the density is great enough to suppress Rayleigh Scattering; something else must be contributing to the blue of the sky. What happens in the mid-atmosphere is that thermal motion of the air results in rapidly changing *density fluctuations* on a local scale. These momentary, fairly random fluctuations cause more molecules to be in one place than another and to radiate more in one direction than another. M. Smoluchowski (1908) and A. Einstein (1910) independently provided the basic ideas for the theory of scattering from these fluctuations, which

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Without an atmosphere to scatter sunlight, the Moon's sky is an eerie black.

gives similar results to those of Rayleigh. Scattering from inhomogeneities in density is of interest whenever light travels great distances in a medium, such as the glass fiber of a communications link (p. 197).

Sunlight streaming into the atmosphere from one direction is scattered in all directions. Without an atmosphere, the daytime sky would be as black as the void of space, as black as the Moon sky. When the Sun is low over the horizon, its light passes through a great thickness of air (far more so than it does at noon). With the blue-end appreciably attenuated, the reds and yellows propagate along the line-of-sight from the Sun to produce Earth's familiar fiery sunsets.

4.2.1 Scattering and Interference

In dense media, a tremendous number of close-together atoms or molecules contribute an equally tremendous number of scattered electromagnetic wavelets. These wavelets overlap and interfere in a way that does not occur in a tenuous medium. As a rule, ***the denser the substance through which light advances, the less the lateral scattering***, and to understand why that's so, we must examine the interference taking place.

Interference has already been discussed (p. 20) and will be treated in further detail in Chapters 7 and 9; here, the basics suffice. Recall that interference is *the superposition of two or more waves producing a resultant disturbance that is the sum of the overlapping wave contributions*. Figure 2.14 shows two harmonic waves of the same frequency traveling in the same direction. When they are precisely in-phase (Fig. 2.14*a*), the resultant at every point is the sum of the two wave-height values. This extreme case is called ***total constructive interference***. When the phase difference reaches 180°, the waves tend to cancel, and we have the other extreme, called ***total destructive interference*** (Fig. 2.14*d*).

The theory of Rayleigh Scattering has independent molecules randomly arrayed in space so that the phases of the secondary wavelets scattered off to the side have no particular relationship to one another and there is no sustained pattern of interference. That situation occurs when the separation between the molecular scatterers is roughly a wavelength or more, as it is in a tenuous gas. In Fig. 4.3*a* a parallel beam of light is inci-

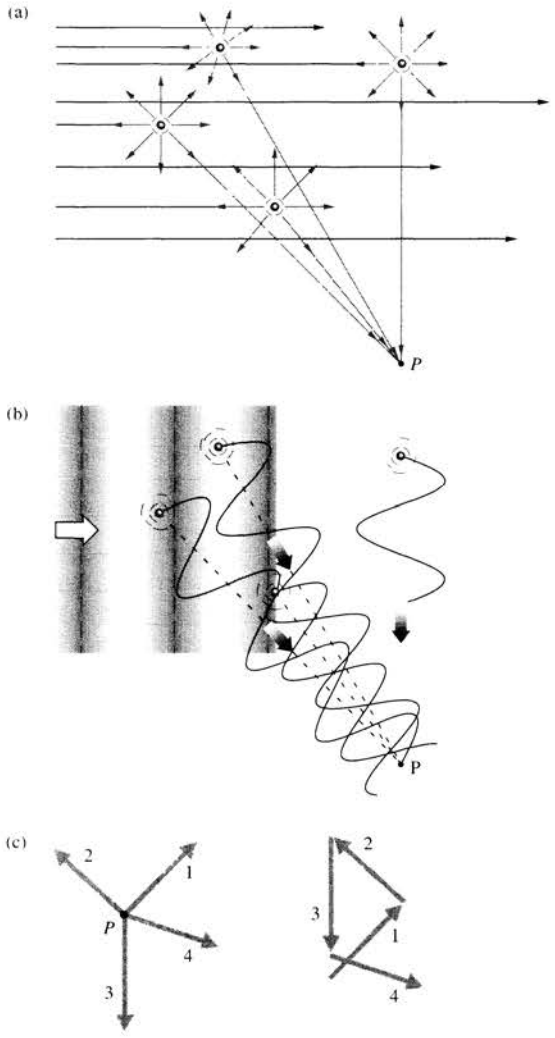


Figure 4.3 (*a*) The scattering of light from a widely spaced distribution of molecules. (*b*) The wavelets arriving at a lateral point *P* have a jumble of different phases and tend not to interfere in a sustained constructive fashion. (*c*) That can probably be appreciated most easily using phasors. As they arrive at *P* the phasors have large phase angle differences with respect to each other. When added tip-to-tail they therefore tend to spiral around keeping the resultant phasor quite small. Remember that we are really dealing with millions of tiny phasors rather than four substantial ones.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

dent from the left. This so-called *primary light field* (in this instance composed of plane waves) illuminates a group of widely spaced molecules. A continuing progression of primary wavefronts sweep over and successively energize and re-energize each molecule, which, in turn, scatters light in all directions, and in particular out to some lateral point P. Because the lengths of their individual paths to P differ greatly in comparison to λ, some of the wavelets arriving at P are ahead of others while some are behind, and that by substantial fractions of a wavelength (Fig. 4.3*b*). In other words, the phases of the wavelets at P differ greatly. (Remember that the molecules are also moving around, and that changes the phases as well.) At any moment some wavelets interfere constructively, some destructively, and the shifting random hodgepodge of overlapping wavelets effectively averages away the interference. ***Random, widely spaced scatterers driven by an incident primary wave emit wavelets that are essentially independent of one another in all directions except forward. Laterally scattered light, unimpeded by interference, streams out of the beam.*** This is approximately the situation existing about 100 miles up in the Earth's tenuous high-altitude atmosphere, where a good deal of blue-light scattering takes place.

That the scattered irradiance should depend on $1/\lambda^4$ is easily seen by returning to the concept of dipole radiation (Section 3.4.3). Each molecule is taken as an electron oscillator driven into vibration by the incident field. Being far apart, they are assumed to be independent of one another and each radiates in accord with Eq. (3.56). The scattered electric fields are essentially independent, and there is no interference laterally. Accordingly, the net irradiance at P is the algebraic sum of the scattered irradiances from each molecule (p. 285). For an individual scatterer the irradiance is given by Eq. (3.57), and it varies with ω^4.

The advent of the laser has made it relatively easy to observe Rayleigh Scattering directly in low-pressure gases, and the results confirm the theory.

Forward Propagation

To see why the forward direction is special, why the wave advances in any medium, consider Fig. 4.4. Notice that for a forward point P all the different paths taken by the light are about the same length; scattering alters the various path lengths by very little. (The scattered wavelets arrive at P more or less in-phase and essentially interfere constructively.) A more detailed description is provided by Fig. 4.5. It depicts a sequence in time showing two molecules A and B, interacting with an incoming primary plane wave—a solid arc represents

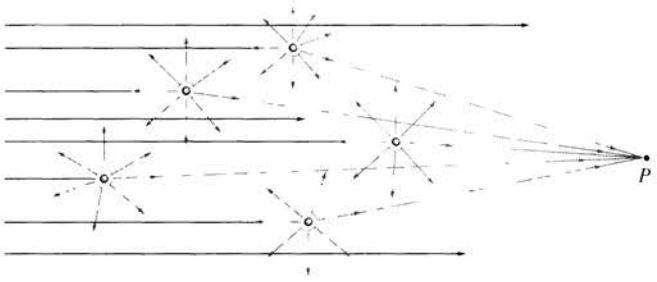


Figure 4.4 Scattering in the forward direction doesn't change the light paths very much, and the waves all arrive at P pretty much in-phase.

a secondary wavelet peak (a positive maximum); a dashed arc corresponds to a trough (a negative maximum). In (*a*), the primary wavefront impinges on molecule A, which begins to scatter a spherical wavelet. For the moment, suppose the

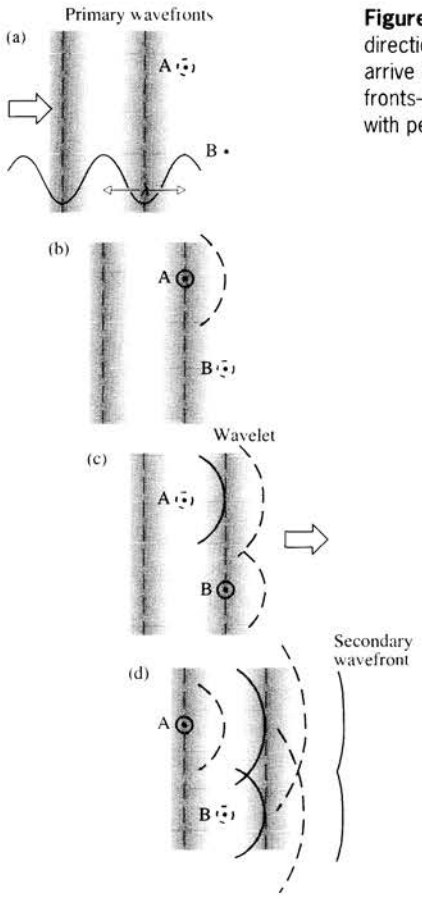


Figure 4.5 In the forward direction the scattered wavelets arrive in-phase on planar wavefronts—trough with trough, peak with peak.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

wavelet is 180° out-of-phase with the incident wave. (A driven oscillator is usually out-of-phase with the driver: p. 93.) Thus A begins to radiate a trough (a negative E-field) in response to being driven by a peak (a positive E-field). Part (b) shows the spherical wavelet and the plane wave overlapping, marching out-of-step but marching together. The incident wavefront impinges on B, and it, in turn, begins to reradiate a wavelet, which must also be out-of-phase by 180°. In (c) and (d), we see the point of all of this, namely, that both wavelets are moving forward—they are in-phase with each other. That condition would be true for all such wavelets regardless of both how many molecules there were and how they were distributed. Because of the asymmetry introduced by the beam itself, ***all the scattered wavelets add constructively with each other in the forward direction***.

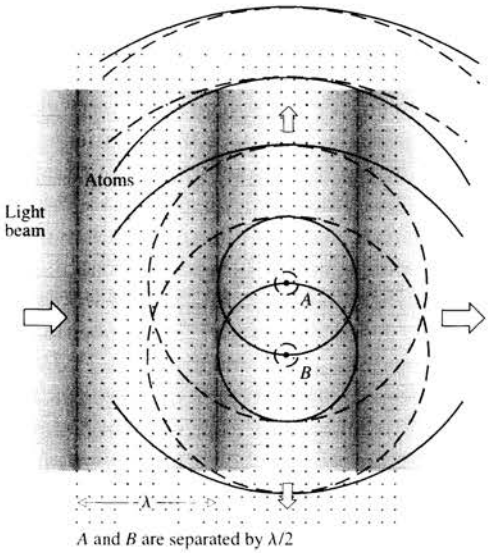


Figure 4.6 A plane wave impinging from the left. The medium is composed of many closely spaced atoms. Among countless others, a wavefront stimulates two atoms, A and B, that are very nearly one-half wavelength apart. The wavelets they emit interfere destructively. Trough overlaps crest, and they completely cancel each other in the direction perpendicular to the beam. That process happens over and over again, and little or no light is scattered laterally.

4.2.2 The Transmission of Light Through Dense Media

Now, suppose the amount of air in the region under consideration is increased. In fact, imagine that each little cube of air, one wavelength on a side, contains a great many molecules, whereupon it is said to have an appreciable *optical density*. (This usage probably derives from the fact that early experiments on gases indicated that an increase in density is accompanied by a proportionate increase in the index of refraction.) At the wavelengths of light, the Earth's atmosphere at STP has about 3 million molecules in such a λ^3-cube. The scattered wavelets ($\lambda \approx 500$ nm) radiated by sources so close together ($\approx$3 nm) cannot properly be assumed to arrive at some point P with random phases—interference will be important. This is equally true in liquids and solids where the atoms are 10 times closer and arrayed in a far more orderly fashion. In such cases, the light beam effectively encounters a uniform medium with no discontinuities to destroy the symmetry. Again, the scattered wavelets interfere constructively in the forward direction (that much is independent of the arrangement of the molecules), but now destructive interference predominates in all other directions. ***Little or no light ends up scattered laterally or backwards in a dense homogeneous medium.***

To illustrate the phenomenon, Fig. 4.6 shows a beam moving through an ordered array of close-together scatterers. All along wavefronts throughout the beam, sheets of molecules are energized in-phase, radiate, and are re-energized, over and over again as the light sweeps past. Thus some molecule A radiates spherically out of the beam, but because of the ordered close arrangement, there will be a molecule B, a distance $\approx\lambda/2$ away, such that both wavelets cancel in the transverse direction. Here, where λ is thousands of times larger than the scatterers and their spacing, there will likely always be pairs of molecules that tend to negate each other's wavelets in any given lateral direction. Even if the medium is not perfectly ordered, the net electric field at a point in any transverse direction will be the sum of a great many *tiny* scattered fields, each somewhat out-of-phase with the next, so that the sum (which will be different from point to point) will always be

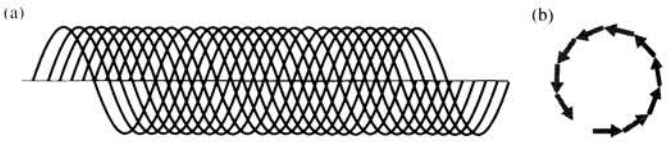


Figure 4.7 (a) When a great many tiny slightly shifted waves arrive at a point in space, there is generally as much positive E-field as negative, and the resultant disturbance is nearly zero. (b) The tiny phasors representing those waves form a very small circular figure, and the resultant (which oscillates in a way that depends on the number of waves) is never large.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

small (Fig. 4.7). This makes sense from the perspective of conservation of energy—we can't have constructive interference in every direction. ***Interference produces a redistribution of energy, out of the regions where it's destructive into the regions where it's constructive.***

The more dense, uniform, and ordered the medium is (the more nearly homogeneous), the more complete will be the lateral destructive interference and the smaller the amount of nonforward scattering. Thus most of the energy will go into the forward direction, and the beam will advance essentially undiminished (Fig. 4.8).

Scattering on a per-molecule basis is extremely weak. In order to have half its energy scattered, a beam of green light will have to traverse ≈150 km of atmosphere. Since about 1000 times more molecules are in a given volume of liquid than in the same volume of vapor (at atmospheric pressure), we can expect to see an increase in scattering. Still, the liquid is a far more ordered state with much less pronounced density fluctuations, and that should suppress the nonforward scattering appreciably. Accordingly, an increased scattering per unit volume is observed in liquids, but it's more like 5 to 50 times as much rather than 1000 times. Molecule for molecule, liquids scatter substantially less than gases. Put a few drops of milk in a tank of water and illuminate it with a bright flash-

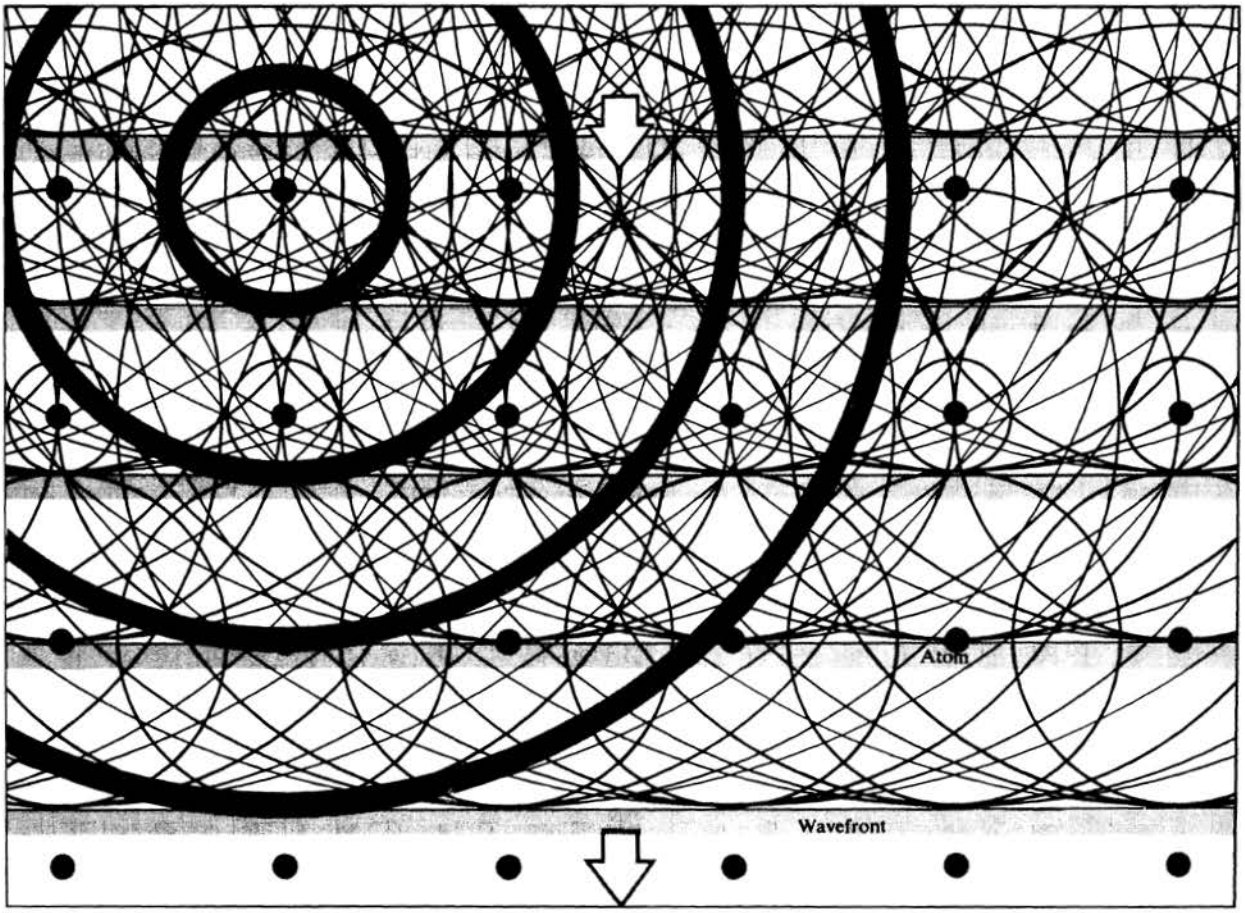


Figure 4.8 A downward plane wave incident on an ordered array of atoms. Wavelets scatter in all directions and overlap to form an ongoing secondary plane wave traveling downward.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

light beam. A faint but unmistakable blue haze will scatter out laterally, and the direct beam will emerge decidedly reddened.

Transparent amorphous solids, such as glass and plastic, will also scatter light laterally, but very weakly. Good crystals, like quartz and mica, with their almost perfectly ordered structures, scatter even more faintly. Of course, imperfections of all sorts (dust and bubbles in liquids, flaws and impurities in solids) will serve as scatterers, and when these are small, as in the gem moonstone, the emerging light will be bluish.

In 1869 John Tyndall experimentally studied the scattering produced by small particles. He found that as the size of the particles increased (from a fraction of a wavelength), the amount of scattering of the longer wavelengths increased proportionately. Ordinary clouds in the sky testify to the fact that relatively large droplets of water scatter white light with no appreciable coloration. The same is true of the microscopic globules of fat and protein in milk.

When the number of molecules in a particle is small, they are all close to one another and act in unison; their wavelets interfere constructively, and the scattering is strong. As the size of the particle approaches a wavelength, the atoms at its extremities no longer radiate wavelets that are necessarily in-phase and the scattering begins to diminish. This happens first at the short wavelengths (blue), and so as the particle size increases, it scatters proportionately more of the red end of the spectrum (and it does so increasingly in the forward direction). The theoretical analysis of scattering from spherical particles of any size was first published by Gustav Mie in 1908. Mie Scattering depends only weakly on wavelength and becomes independent of it (white light in, white light out) when the particle size exceeds λ. Reasonably enough, Rayleigh Scattering is the small-size limiting case of Mie Scattering.

On an overcast day, the sky looks hazy gray because of large water droplets. In the same way, some inexpensive plastic food containers and white garbage-bag plastic look pale blue-white in scattered light and are distinctly orange in transmitted light. The garbage bags, in order to be made opaque, contain (2 to 2.5%) clear TiO_2 spheres ($n = 2.76$) about 200 nm in diameter, and these Mie scatter bluish-white.*

*It has only recently been observed (and that was by chance) that inhomogeneous opaque materials, such as milk and white paint, can reduce the effective speed of light to as little as one-tenth the value anticipated for the medium. See S. John, "Localization of light," *Phys. Today* **44**, 32 (1991).

4.2.3 Transmission and the Index of Refraction

The transmission of light through a homogeneous medium is an ongoing repetitive process of scattering and rescattering. Each such event introduces a phase shift into the light field, which ultimately shows up as a shift in the apparent phase velocity of the transmitted beam from its nominal value of c. That corresponds to an index of refraction for the medium ($n = c/v$) which is other than one, even though **photons exist only at a speed** c.

To see how this comes about, return to Fig. 4.5. Recall that the scattered wavelets all combine in-phase in the forward direction to form what might best be called the *secondary wave*. For empirical reasons alone we can anticipate that the secondary wave will combine with what is left of the primary wave to yield the only observed disturbance within the medium, namely, the **transmitted wave**. *Both the primary and secondary electromagnetic waves propagate through the interatomic void with the speed c.* Yet the medium can certainly possess an index of refraction other than 1. The refracted wave may appear to have a phase velocity less than, equal to, or even greater than c. The key to this apparent contradiction resides in the phase relationship between the secondary and primary waves.

The classical model predicts that the electron-oscillators will be able to vibrate almost completely in-phase with the driving force (i.e., the primary disturbance) only at relatively low frequencies. As the frequency of the electromagnetic field increases, the oscillators will fall behind, lagging in phase by a proportionately larger amount. A detailed analysis reveals that at resonance the phase lag will reach 90°, increasing thereafter to almost 180°, or half a wavelength, at frequencies well above the particular characteristic value. Problem 4.4 explores this phase lag for a damped driven oscillator, and Fig. 4.9 summarizes the results.

In addition to these lags there is another effect that must be considered. When the scattered wavelets recombine, the resultant secondary wave* itself lags the oscillators by 90°.

The combined effect of both these mechanisms is that at frequencies below resonance, the secondary wave lags the pri-

*This point will be made more plausible when we consider the predictions of the Huygens–Fresnel theory in the diffraction chapter. Most texts on E & M treat the problem of radiation from a sheet of oscillating charges, in which case the 90° phase lag is a natural result (see Problem 4.5).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

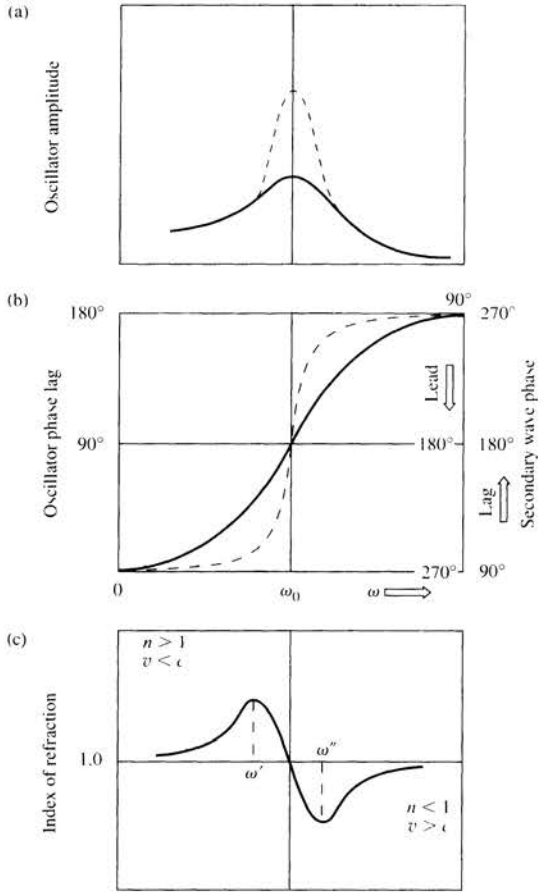


Figure 4.9 A schematic representation of (*a*) amplitude and (*b*) phase lag versus driving frequency for a damped oscillator. The dashed curves correspond to decreased damping. The corresponding index of refraction is shown in (c).

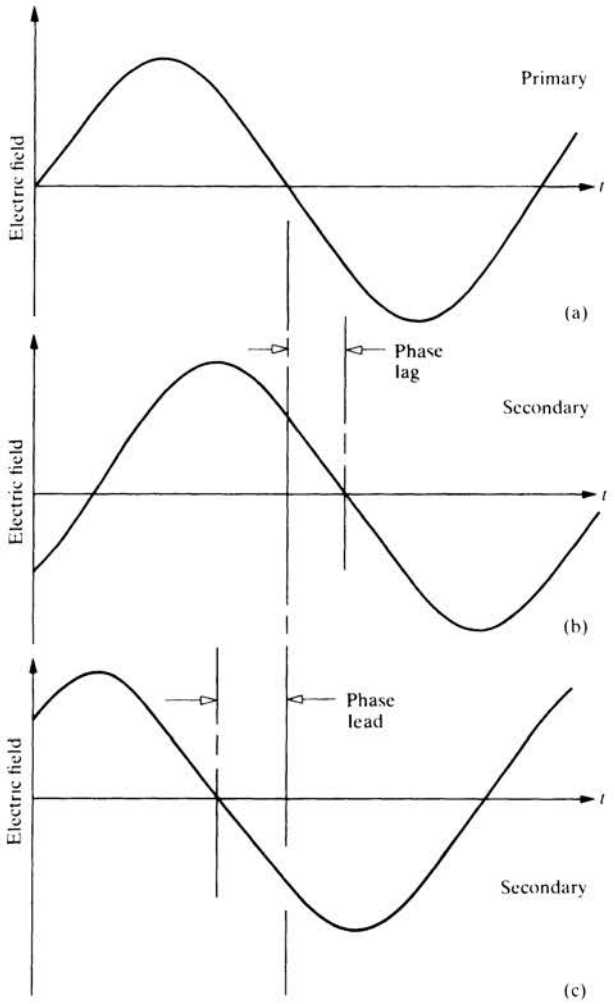


Figure 4.10 A primary wave (*a*) and two possible secondary waves. In (*b*) the secondary lags the primary—it takes longer to reach any given value. In (c) the secondary wave reaches any given value before (at an earlier time than) the primary; that is, it leads.

mary (Fig. 4.10) by some amount between approximately 90° and 180°, and at frequencies above resonance, the lag ranges from about 180° to 270°. But a phase lag of $\delta \geq 180°$ is equivalent to a phase lead of $360° - \delta$, [e.g., $\cos(\theta - 270°) = \cos(\theta + 90°)$]. This much can be seen on the right side of Fig. 4.9*b*.

Within the transparent medium, the primary and secondary waves overlap and, depending on their amplitudes and relative phase, generate the net transmitted disturbance. Except for the fact that it is weakened by scattering, the primary wave travels into the material just as if it were traversing free space. By comparison to this free-space wave, which initiated the process, the resultant transmitted wave is phase shifted, and this phase difference is crucial.

When the secondary wave lags (or leads) the primary, the resultant transmitted wave must also lag (or lead) it by some amount (Fig. 4.11). This qualitative relationship will serve our purposes for the moment, although it should be noted that the phase of the resultant also depends on the amplitudes of the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

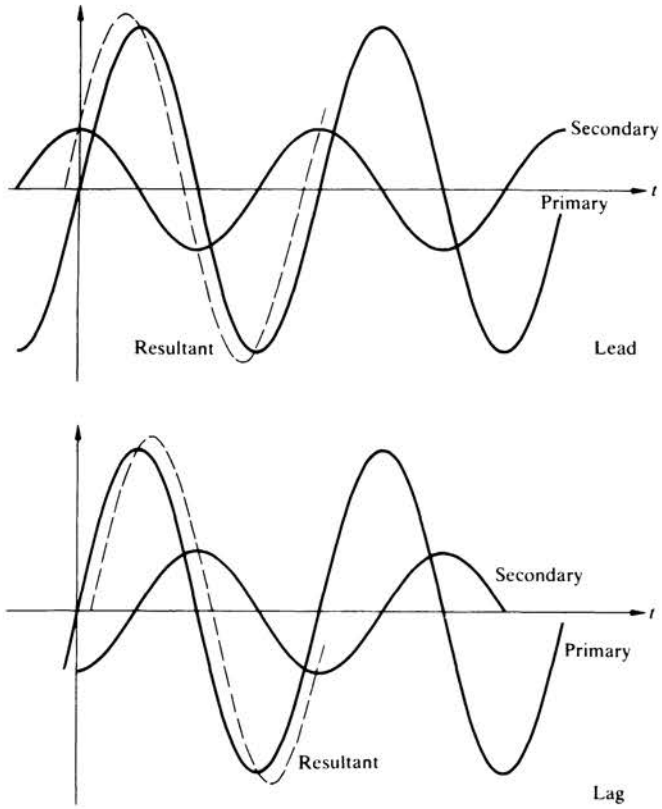


Figure 4.11 If the secondary leads the primary, the resultant will also lead it.

interacting waves [see Eq. (7.10)]. At frequencies below ω_0 the transmitted wave lags the free-space wave, whereas at frequencies above ω_0 it leads the free-space wave. For the special case in which $\omega = \omega_0$ the secondary and primary waves are out-of-phase by 180°. The former works against the latter, so that the refracted wave is appreciably reduced in amplitude although unaffected in phase.

As the transmitted wave advances through the medium, scattering occurs over and over again. Light traversing the substance is progressively retarded (or advanced) in phase. Evidently, since the speed of the wave is the rate of advance of the condition of constant phase, a change in the phase should correspond to a change in the speed.

We now wish to show that a phase shift is indeed tantamount to a difference in phase velocity. In free space, the disturbance at some point P may be written as

$$E_P(t) = E_0 \cos \omega t \tag{4.1}$$

If P is surrounded by a dielectric, there will be a cumulative phase shift ε_P, which was built up as the wave moved through the medium to P. At ordinary levels of irradiance the medium will behave linearly, and the frequency in the dielectric will be the same as that in vacuum, even though the wavelength and speed may differ. Once again, but this time in the medium, the disturbance at P is

$$E_P(t) = E_0 \cos (\omega t - \varepsilon_P) \tag{4.2}$$

where the subtraction of ε_P corresponds to a phase lag. An observer at P will have to wait a longer time for a given crest to arrive when she is in the medium than she would have had to wait in vacuum. That is, if you imagine two parallel waves of the same frequency, one in vacuum and one in the material, the vacuum wave will pass P a time ε_P/ω before the other wave. Clearly then, *a phase lag of ε_P corresponds to a reduction in speed*, $v < c$ and $n > 1$. Similarly, *a phase lead yields an increase in speed*, $v > c$ and $n < 1$. Again, the scattering process is a continuous one, and the cumulative phase shift builds as the light penetrates the medium. That is to say, ε is a function of the length of dielectric traversed, as it must be if v is to be constant (see Problem 4.5). In the vast majority of situations encountered in Optics $v < c$ and $n > 1$; see Table 4.1. The important exception is the case of X-ray propagation where $\omega > \omega_0$, $v > c$, and $n < 1$.

The overall form of $n(\omega)$, as depicted in Fig. 4.9*c*, can now be understood as well. At frequencies far below ω_0 the amplitudes of the oscillators and therefore of the secondary waves are very small, and the phase angles are approximately 90°. Consequently, the refracted wave lags only slightly, and n is only slightly greater than 1. As ω increases, the secondary waves have greater amplitudes and lag by greater amounts. The result is a gradually decreasing wave speed and an increasing value of $n > 1$. Although the amplitudes of the secondary waves continue to increase, their relative phases approach 180° as ω approaches ω_0. Consequently, their ability to cause a further increase in the resultant phase lag diminishes. A turning point ($\omega = \omega'$) is reached where the refracted wave begins to experience a decreasing phase lag and an increasing speed ($dn/d\omega < 0$). That continues until $\omega = \omega_0$, whereupon the transmitted wave is appreciably reduced in amplitude but unaltered in phase and speed. At that point, $n = 1$, $v = c$, and we are more or less at the center of the absorption band.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

TABLE 4.1 Approximate Indices of Refraction of Various Substances*

Air	1.000 29
Ice	1.31
Water	1.333
Ethyl alcohol (C_2H_5OH)	1.36
Fused quartz (SiO_2)	1.4584
Carbon tetrachloride (CCl_4)	1.46
Turpentine	1.472
Benzene (C_6H_6)	1.501
Plexiglass	1.51
Crown glass	1.52
Sodium chloride (NaCl)	1.544
Light flint glass	1.58
Polystyrene	1.59
Carbon disulfide (CS_2)	1.628
Dense flint glass	1.66
Lanthanum flint glass	1.80
Zircon ($ZrO_2 \cdot SiO_2$)	1.923
Fabulite ($SrTiO_3$)	2.409
Diamond (C)	2.417
Rutile (TiO_2)	2.907
Gallium phosphide	3.50

*Values vary with physical conditions—purity, pressure, etc. These correspond to a wavelength of 589 nm

At frequencies just beyond ω_0 the relatively large-amplitude secondary waves lead; the transmitted wave is advanced in phase, and its speed exceeds c ($n < 1$). As ω increases, the whole scenario is played out again in reverse (with some asymmetry due to frequency-dependent asymmetry in oscillator amplitudes and scattering). At even higher frequencies the secondary waves, which now have very small amplitudes, lead by nearly 90°. The resulting transmitted wave is advanced very slightly in phase, and n gradually approaches 1.

The precise shape of a particular $n(\omega)$ curve depends on the specific oscillator damping, as well as on the amount of absorption, which in turn depends on the number of oscillators participating.

A rigorous solution to the propagation problem is known as the *Ewald–Oseen Extinction Theorem*. Although the mathematical formalism, involving integro-differential equations, is far too complicated to treat here, the results are certainly of interest. It is found that the electron-oscillators generate an electromagnetic wave having essentially two terms. One of these precisely cancels the primary wave within the medium. The other, and only remaining disturbance, moves through the dielectric at a speed $v = c/n$ as the transmitted wave.* ***Henceforth we shall simply assume that a lightwave propagating through any substantive medium travels at a speed*** $v \neq c$.

Apparently, any quantum-mechanical model we construct will somehow have to associate a wavelength with the photon. That's easily done mathematically via the expression $p = h/\lambda$, even if it's not clear at this point what is doing the waving. Still the wave nature of light seems inescapable; it will have to be infused into the theory one way or another. And once we have the idea of a photon-wavelength, it's natural to bring in the concept of relative phase. Thus ***the index of refraction arises when the absorption and emission process advances or retards the phases of the scattered photons, even as they travel at speed*** c.

4.3 Reflection

When a beam of light impinges on the surface of a transparent material, such as a sheet of glass, the wave "sees" a vast array of closely spaced atoms that will somehow scatter it. Remember that the wave may be ≈500 nm long, whereas the atoms and their separations (≈0.2 nm) are thousands of times smaller. In the case of transmission through a dense medium, the scattered wavelets cancel each other in all but the forward direction, and just the ongoing beam is sustained. But that can only happen if there are no discontinuities. This is not the case at an interface between two different transparent media (such as air and glass), which is a jolting discontinuity. When a beam of light strikes such an interface, some light is always scattered backward, and we call this phenomenon **reflection**.

If the transition between two media is gradual—that is, if the dielectric constant (or the index of refraction) changes from that of one medium to that of the other over a distance of a wavelength or more—there will be very little reflection; the interface effectively vanishes. On the other hand, a transition from one medium to the other over a distance of 1/4 wavelength or less behaves very much like a totally discontinuous change.

*For a discussion of the Ewald–Oseen theorem, see *Principles of Optics* by Born and Wolf, Section 2.4.2; this is heavy reading. Also look at Reali, "Reflection from dielectric materials." *Am. J. Phys.* **50**, 1133 (1982).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Internal and External Reflection

Imagine that light is traveling across a large homogeneous block of glass (Fig. 4.12). Now, suppose that the block is sheared in half perpendicular to the beam. The two segments are then separated, exposing the smooth flat surfaces depicted in Fig. 4.12*b*. Just before the cut was made, there was no light-wave traveling to the left inside the glass—we know the beam only advances. Now there must be a wave (beam-I) moving to the left, reflected from the surface of the right-hand block. The implication is that a region of scatterers on and beneath the exposed surface of the right-hand block is now "unpaired," and the backward radiation they emit can no longer be canceled. The region of oscillators that was adjacent to these, prior to the cut, is now on the section of the glass that is to the left. When the two sections were together, these scatterers presumably also emitted wavelets in the backward direction that were 180° out-of-phase with, and canceled, beam-I. Now they produce reflected beam-II. Each molecule scatters light in the backward direction, and, in principle, ***each and every molecule contributes to the reflected wave***. Nonetheless, in practice, it is a thin layer ($\approx\lambda/2$ deep) of unpaired atomic oscillators near the surface that is effectively responsible for the reflection. For an air–glass interface, about 4% of the energy of an incident beam falling perpendicularly *in* air *on* glass will be reflected straight back out by this layer of unpaired scatterers (p. 90). And that's true whether the glass is 1.0 mm thick or 1.00 m thick.

Beam-I reflects off the right-hand block, and because light was initially traveling from a less to a more optically dense medium, this is called **external reflection**. In other words, the index of the incident medium (n_i) is less than the index of the transmitting medium (n_t). Since the same thing happens to the unpaired layer on the section that was moved to the left, it, too, reflects backwards. With the beam incident perpendicularly *in* glass *on* air, 4% must again be reflected, this time as beam-II.

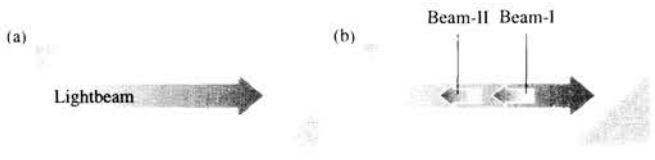


Figure 4.12 (*a*) A lightbeam propagating through a dense homogeneous medium such as glass. (*b*) When the block of glass is cut and parted, the light is reflected backward at the two new interfaces. Beam-I is externally reflected, and beam-II is internally reflected. Ideally, when the two pieces are pressed back together, the two reflected beams cancel one another.

This process is referred to as **internal reflection** because $n_i > n_t$. If the two glass regions are made to approach one another increasingly closely (so that we can imagine the gap to be a thin film of, say, air—p. 402), the reflected light will diminish until it ultimately vanishes as the two faces merge and disappear and the block becomes continuous again. Remember this ***180° relative phase shift between internally and externally reflected light*** (see Section 4.10 for a more rigorous treatment)—we will come back to it later on.

Experience with the common mirror makes it obvious that white light is reflected as white—it certainly isn't blue. To see why, first remember that the layer of scatterers responsible for the reflection is effectively about $\lambda/2$ thick (as per Fig. 4.6). Thus the larger the wavelength, the deeper the region contributing (typically upward of a thousand atom layers), and the more scatterers there are acting together. This tends to balance out the fact that each scatterer is less efficient as λ

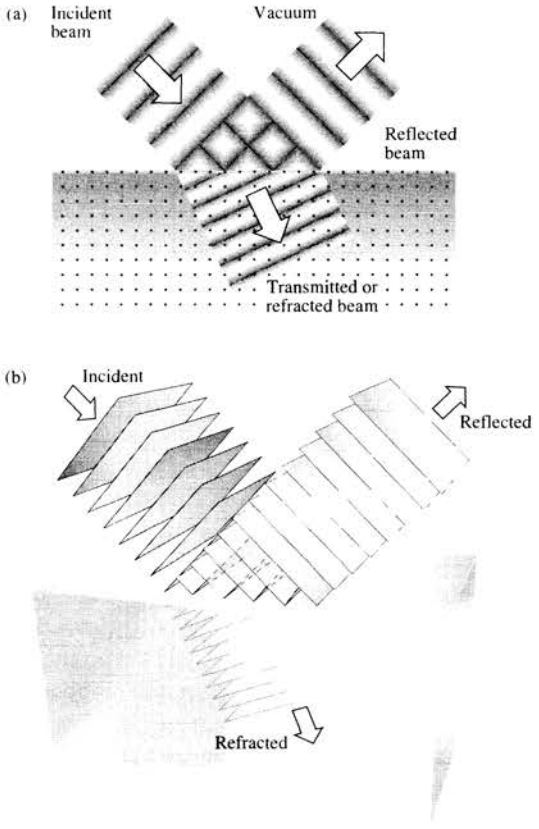


Figure 4.13 A beam of plane waves incident on a distribution of molecules constituting a piece of clear glass or plastic. Part of the incident light is reflected and part refracted.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

increases (remember $1/\lambda^4$). The combined result is that ***the surface of a transparent medium reflects all wavelengths about equally and doesn't appear colored in any way***. That, as we will see, is why this page looks white under white-light illumination.

4.3.1 The Law of Reflection

Figure 4.13 shows a beam composed of plane wavefronts impinging at some angle on the smooth, flat surface of an optically dense medium (let it be glass). Assume that the surrounding environment is vacuum. Follow one wavefront as it sweeps in and across the molecules on the surface. For the sake of simplicity, in Figs. 4.14 and 4.15 we have omitted everything but one molecular layer at the interface. As the wavefront descends, it energizes and re-energizes one scatterer after another, each of which radiates a stream of photons that can be regarded as a hemispherical wavelet in the incident medium. Because the wavelength is so much greater than the separation between the molecules, the wavelets emitted back into the incident medium advance together and add constructively in only one direction, and there is one well-defined *reflected* beam. That would not be true if the incident radiation was short-wavelength X-rays, in which circumstance there would be several reflected beams. And it would not be true if the scatterers were far apart compared to λ, as they are for a diffraction grating (p. 476), in which case there would also be several reflected beams. The direction of the reflected beam is

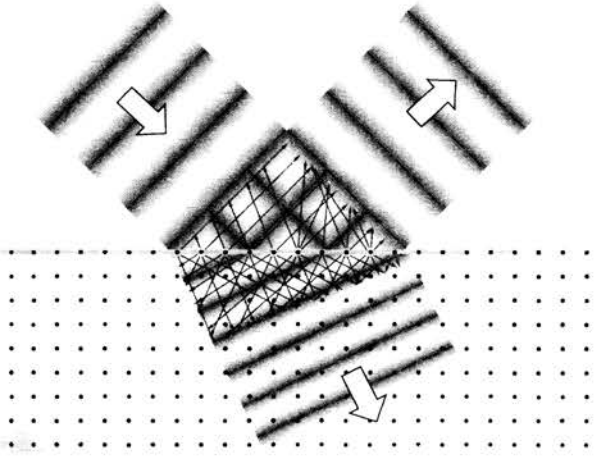


Figure 4.14 A plane wave sweeps in stimulating atoms across the interface. These radiate and reradiate, thereby giving rise to both the reflected and transmitted waves.

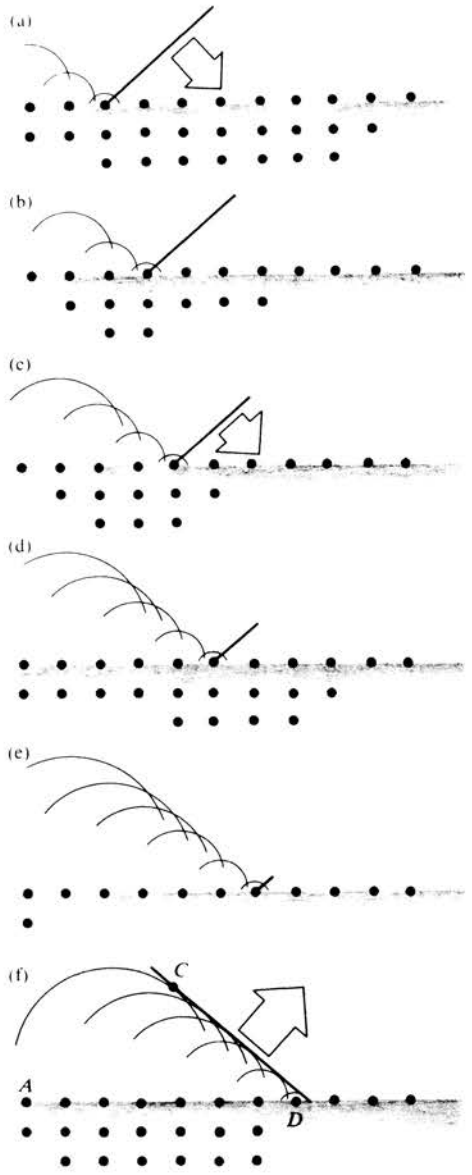


Figure 4.15 The reflection of a wave as the result of scattering.

determined by the constant phase difference between the atomic scatterers. That, in turn, is determined by the angle made by the incident wave and the surface, the so-called **angle-of-incidence**.

In Fig. 4.16, the line $\overline{AB}$ lies along an incoming wavefront, while $\overline{CD}$ lies on an outgoing wavefront—in effect, $\overline{AB}$ transforms on reflection into $\overline{CD}$. With Fig. 4.15 in mind, we see

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027