Learn to:

- Grasp physics terminology
- Get a handle on quantum and nuclear physics
- Understand waves, forces, and fields
- Make sense of electric potential and energy

Steven Holzner, PhD
Author of Physics For Dummies
About the Author

Steven Holzner taught Physics at Cornell University for more than a decade, teaching thousands of students. He’s the award-winning author of many books, including Physics For Dummies, Quantum Physics For Dummies, and Differential Equations For Dummies, plus For Dummies workbooks for all three titles. He did his undergraduate work at MIT and got his PhD from Cornell, and he has been on the faculty of both MIT and Cornell.

Dedication

To Nancy, of course.

Author’s Acknowledgments

The book you hold in your hands is the product of many people’s work. I’d particularly like to thank Acquisitions Editor Tracy Boggier, Senior Project Editor Alissa Schwipps, Senior Copy Editor Danielle Voirol, Technical Editors Laurie Fuhr and Ron Reifenberger, and the many talented folks in Composition Services.
Publisher’s Acknowledgments
We’re proud of this book; please send us your comments at http://dummies.custhelp.com. For other comments, please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993, or fax 317-572-4002.

Some of the people who helped bring this book to market include the following:

**Acquisitions, Editorial, and Media Development**

Senior Project Editor: Alissa Schwipps  
Acquisitions Editor: Tracy Boggier  
Senior Copy Editor: Danielle Voirol  
Contributor: Neil Clark  
Assistant Editor: Erin Calligan Mooney  
Senior Editorial Assistant: David Lutton  
Technical Editors: Laurie Fuhr, Ron Reifenberger  
Senior Editorial Manager: Jennifer Ehrlich  
Editorial Assistants: Rachelle Amick, Jennette ElNaggar  
Cover Photos: © Thorsten | Dreamstime.com  
Cartoons: Rich Tennant  
(www.the5thwave.com)

**Composition Services**

Project Coordinator: Katherine Crocker  
Layout and Graphics: Carrie A. Cesavice, Amy Hassos, Mark Pinto, Erin Zeltner  
Proofreader: Tricia Liebig  
Indexer: Sharon Shock

**Publishing and Editorial for Consumer Dummies**

Diane Graves Steele, Vice President and Publisher, Consumer Dummies  
Kristin Ferguson-Wagstaffe, Product Development Director, Consumer Dummies  
Ensley Eikenburg, Associate Publisher, Travel  
Kelly Regan, Editorial Director, Travel

**Publishing for Technology Dummies**

Andy Cummings, Vice President and Publisher, Dummies Technology/General User

**Composition Services**

Debbie Stailey, Director of Composition Services
Contents at a Glance

Introduction ............................................................................. 1

Part I: Understanding Physics Fundamentals ......................... 7
Chapter 1: Understanding Your World: Physics II, the Sequel .......... 9
Chapter 2: Gearing Up for Physics II ........................................ 19

Part II: Doing Some Field Work:
Electricity and Magnetism .................................................. 35
Chapter 3: Getting All Charged Up with Electricity .................... 37
Chapter 4: The Attraction of Magnetism .................................... 61
Chapter 5: Alternating Current and Voltage ............................ 87

Part III: Catching On to Waves:
The Sound and Light Kinds .................................................. 113
Chapter 6: Exploring Waves .................................................. 115
Chapter 7: Now Hear This: The Word on Sound ........................ 127
Chapter 8: Seeing the Light: When Electricity and Magnetism Combine ............. 155
Chapter 9: Bending and Focusing Light: Refraction and Lenses ............ 175
Chapter 10: Bouncing Light Waves: Reflection and Mirrors ............... 205
Chapter 11: Shedding Light on Light Wave Interference and Diffraction ........ 221

Part IV: Modern Physics .................................................. 247
Chapter 12: Heeding What Einstein Said: Special Relativity ............. 249
Chapter 13: Understanding Energy and Matter as Both Particles and Waves .......... 273
Chapter 14: Getting the Little Picture: The Structure of Atoms ............ 295
Chapter 15: Nuclear Physics and Radioactivity ............................ 319

Part V: The Part of Tens .................................................. 339
Chapter 16: Ten Physics Experiments That Changed the World .......... 341
Chapter 17: Ten Online Problem-Solving Tools .......................... 347

Index ............................................................................. 353
Table of Contents

Introduction ........................................................................................................ 1

About This Book ................................................................................................ 1
Conventions Used in This Book ........................................................................ 1
What You’re Not to Read .................................................................................. 2
Foolish Assumptions .......................................................................................... 3

How This Book Is Organized .......................................................................... 3

Part I: Understanding Physics Fundamentals ............................................ 3
Part II: Doing Some Field Work: Electricity and Magnetism ....................... 4
Part III: Catching On to Waves: The Sound and Light Kinds ....................... 4
Part IV: Modern Physics ............................................................................. 4
Part V: The Part of Tens ............................................................................. 4

Icons Used in This Book ............................................................................... 5
Where to Go from Here ................................................................................... 5

Part I: Understanding Physics Fundamentals ......................................... 7

Chapter 1: Understanding Your World: Physics II, the Sequel ................. 9

Getting Acquainted with Electricity and Magnetism ...................................... 10

Looking at static charges and electric field .................................................. 10
Moving on to magnetism ............................................................................. 11
AC circuits: Regenerating current with electric and magnetic fields ............ 11

Riding the Waves ........................................................................................... 12

Getting along with sound waves ................................................................. 12
Figuring out what light is ............................................................................. 12
Reflection and refraction: Bouncing and bending light ............................. 13
Searching for images: Lenses and mirrors ................................................... 14
Calling interference: When light collides with light .................................... 15

Branching Out with Modern Physics .......................................................... 15

Shedding light on blackbodies: Warm bodies make their own light .......... 15
Speeding up with relativity: Yes, $E = mc^2$ .................................................. 16
Assuming a dual identity: Matter travels in waves, too ............................ 16
Meltdown! Knowing the $\alpha\beta\gamma$’s of radioactivity ...................................... 17
Chapter 2: Gearing Up for Physics II ................................................................. 19
  Math and Measurements: Reviewing Those Basic Skills ......................... 19
  Using the MKS and CGS systems of measurement .................................. 20
  Making common conversions .................................................................... 20
  Keeping it short with scientific notation ................................................. 24
  Brushing up on basic algebra .................................................................. 24
  Using some trig ...................................................................................... 25
  Using significant digits .......................................................................... 26

  Refreshing Your Physics Memory ............................................................. 27
  Pointing the way with vectors ................................................................. 28
  Moving along with velocity and acceleration ......................................... 29
  Strong-arm tactics: Applying some force ............................................... 30
  Getting around to circular motion ........................................................ 30
  Getting electrical with circular motion .................................................... 32

Part II: Doing Some Field Work:
Electricity and Magnetism ................................................................. 35

Chapter 3: Getting All Charged Up with Electricity ............................... 37
  Understanding Electric Charges .............................................................. 37
  Can’t lose it: Charge is conserved ......................................................... 38
  Measuring electric charges ................................................................. 38
  Opposites attract: Repelling and attracting forces .................................. 39

  Getting All Charged Up .......................................................................... 40
  Static electricity: Building up excess charge .......................................... 40
  Checking out charging methods ........................................................... 41
  Considering the medium: Conductors and insulators .......................... 43

  Coulomb’s Law: Calculating the Force between Charges ..................... 44
  Introducing Electric Fields ...................................................................... 45
  Sheets of charge: Presenting basic fields .............................................. 45
  Looking at electric fields from charged objects ................................... 47
  Uniform electric fields: Taking it easy
    with parallel plate capacitors ............................................................... 48
  Shielding: The electric field inside conductors .................................... 50

  Voltage: Realizing Potential ................................................................. 52
  Getting the lowdown on electric potential ............................................ 52
  Finding the work to move charges ........................................................ 53
  Finding the electric potential from charges ......................................... 54
  Illustrating equipotential surfaces for point charges and plates .......... 56

  Storing Charge: Capacitors and Dielectrics .......................................... 57
  Figuring out how much capacitors hold .............................................. 57
  Getting extra storage with dielectrics ............................................... 58
  Calculating the energy of capacitors with dielectrics ......................... 59
**Table of Contents**

**Chapter 4: The Attraction of Magnetism** ................................................................. 61

- All About Magnetism: Linking Magnetism and Electricity .................................. 62
  - Electron loops: Understanding permanent magnets and magnetic materials .......... 62
  - North to south: Going polar ........................................................................... 63
  - Defining magnetic field .................................................................................. 65

- Moving Along: Magnetic Forces on Charges ....................................................... 66
  - Finding the magnitude of magnetic force ....................................................... 66
  - Finding direction with the right-hand rule ..................................................... 67
  - A lazy direction: Seeing how magnetic fields avoid work ............................... 68
  - Going orbital: Following charged particles in magnetic fields ....................... 69

- Down to the Wire: Magnetic Forces on Electrical Currents ................................... 74
  - From speed to current: Getting current in the magnetic-force formula ............... 74
  - Torque: Giving current a twist in electric motors .......................................... 76

**Chapter 5: Alternating Current and Voltage** ......................................................... 87

- AC Circuits and Resistors: Opposing the Flow .................................................. 87
  - Finding Ohm’s law for alternating voltage ..................................................... 88
  - Averaging out: Using root-mean-square current and voltage ....................... 89
  - Staying in phase: Connecting resistors to alternating voltage sources .......... 90

- AC Circuits and Capacitors: Storing Charge in Electric Field ............................. 91
  - Introducing capacitive reactance .................................................................. 92
  - Getting out of phase: Current leads the voltage .......................................... 94
  - Preserving power ......................................................................................... 95

- AC Circuits and Inductors: Storing Energy in Magnetic Field .............................. 95
  - Faraday’s law: Understanding how inductors work ..................................... 96
  - Introducing inductive reactance .................................................................. 97
  - Getting behind: Current lags voltage .......................................................... 102

- The Current-Voltage Race: Putting It Together in Series RLC Circuits ............... 103
  - Impedance: The combined effects of resistors, inductors, and capacitors ....... 104
  - Determining the amount of leading or lagging ............................................ 106

- Peak Experiences: Finding Maximum Current in a Series RLC Circuit ............. 109
  - Canceling out reactance .............................................................................. 109
  - Finding resonance frequency ....................................................................... 109

- Semiconductors and Diodes: Limiting Current Direction ...................................... 110
  - The straight dope: Making semiconductors ................................................. 111
  - One-way current: Creating diodes ............................................................... 112
Part III: Catching On to Waves: The Sound and Light Kinds

Chapter 6: Exploring Waves

- Energy Travels: Doing the Wave
- Up and down: Transverse waves
- Back and forth: Longitudinal waves
- Wave Properties: Understanding What Makes Waves Tick
  - Examining the parts of a wave
  - Relating the parts of a wave mathematically
  - Watching for the sine: Graphs of waves
- When Waves Collide: Wave Behavior

Chapter 7: Now Hear This: The Word on Sound

- Vibrating Just to Be Heard: Sound Waves as Vibrations
- Cranking Up the Volume: Pressure, Power, and Intensity
  - Under pressure: Measuring the amplitude of sound waves
  - Introducing sound intensity
- Calculating the Speed of Sound
  - Fast: The speed of sound in gases
  - Faster: The speed of sound in liquids
  - Fastest: The speed of sound in solids
- Analyzing Sound Wave Behavior
  - Echoing back: Reflecting sound waves
  - Sharing spaces: Sound wave interference
  - Bending rules: Sound wave diffraction
  - Coming and going with the Doppler effect
  - Breaking the sound barrier: Shock waves

Chapter 8: Seeing the Light: When Electricity and Magnetism Combine

- Let There Be Light! Generating and Receiving
  - Electromagnetic Waves
  - Creating an alternating electric field
  - Getting an alternating magnetic field to match
  - Receiving radio waves
- Looking at Rainbows: Understanding the Electromagnetic Spectrum
  - Perusing the electromagnetic spectrum
  - Relating the frequency and wavelength of light
- See Ya Later, Alligator: Finding the Top Speed of Light
  - Checking out the first speed-of-light experiment that actually worked
  - Calculating the speed of light theoretically
- You’ve Got the Power: Determining the Energy Density of Light
  - Finding instantaneous energy
  - Averaging light’s energy density
# Table of Contents

## Chapter 9: Bending and Focusing Light: Refraction and Lenses  
### 175

- Wave Hello to Rays: Drawing Light Waves More Simply  
- Slowing Light Down: The Index of Refraction  
  - Figuring out the slowdown  
  - Calculating the bending: Snell’s law  
  - Rainbows: Separating wavelengths  
- Bending Light to Get Internal Reflection  
  - Right back at you: Total internal reflection  
  - Polarized light: Getting a partial reflection  
- Getting Visual: Creating Images with Lenses  
  - Defining objects and images  
  - Now it’s coming into focus: Concave and convex lenses  
- Drawing ray diagrams  
- Getting Numeric: Finding Distances and Magnification  
  - Going the distance with the thin-lens equation  
  - Sizing up the magnification equation  
- Combining Lenses for More Magnification Power  
  - Understanding how microscopes and telescopes work  
  - Getting a new angle on magnification

## Chapter 10: Bouncing Light Waves: Reflection and Mirrors  
### 205

- The Plane Truth: Reflecting on Mirror Basics  
- Getting the angles on plane mirrors  
- Forming images in plane mirrors  
- Finding the mirror size  
- Working with Spherical Mirrors  
  - Getting the inside scoop on concave mirrors  
  - Smaller and smaller: Seeing convex mirrors at work  
- The Numbers Roundup: Using Equations for Spherical Mirrors  
  - Getting numerical with the mirror equation  
  - Discovering whether it’s bigger or smaller: Magnification

## Chapter 11: Shedding Light on Light Wave Interference and Diffraction  
### 221

- When Waves Collide: Introducing Light Interference  
  - Meeting at the bars: In phase with constructive interference  
  - Going dark: Out of phase with destructive interference  
- Interference in Action: Getting Two Coherent Light Sources  
  - Splitting light with double slits  
  - Gasoline-puddle rainbows: Splitting light with thin-film interference  
- Single-Slit Diffraction: Getting Interference from Wavelets  
  - Huygens’s principle: Looking at how diffraction works with a single slit  
  - Getting the bars in the diffraction pattern  
  - Doing diffraction calculations
Multiple Slits: Taking It to the Limit with Diffraction Gratings ...........241
  Separating colors with diffraction gratings ........................................241
  Trying some diffraction-grating calculations ....................................242
  Seeing Clearly: Resolving Power and Diffraction from a Hole ...............243

Part IV: Modern Physics ...............................................................247

Chapter 12: Heeding What Einstein Said: Special Relativity ............249
  Blasting Off with Relativity Basics ..................................................250
    Start from where you’re standing: Understanding reference frames ..................................................250
    Looking at special relativity’s postulates ......................................252
  Seeing Special Relativity at Work ..................................................253
    Slowing time: Chilling out with time dilation ..................................254
    Packing it in: Length contraction ..................................................259
    Pow! Gaining momentum near the speed of light ..........................262
  Here It Is! Equating Mass and Energy with E = mc² ..........................264
    An object’s rest energy: The energy you could get from the mass ..................................................265
    An object’s kinetic energy: The energy of motion ..................................267
    Skipping PE .............................................................................270
  New Math: Adding Velocities Near Light Speed ..............................270

Chapter 13: Understanding Energy and Matter as Both Particles and Waves ..................................................273
  Blackbody Radiation: Discovering the Particle Nature of Light ........274
    Understanding the trouble with blackbody radiation ..........................274
    Being discrete with Planck’s constant .............................................275
  Light Energy Packets: Advancing with the Photoelectric Effect .........276
    Understanding the mystery of the photoelectric effect ......................276
    Einstein to the rescue: Introducing photons ...................................277
    Explaining why electrons’ kinetic energy is independent of intensity ..................................................279
    Explaining why electrons are emitted instantly ..................................280
    Doing calculations with the photoelectric effect ..................................281
  Collisions: Proving the Particle Nature of Light with the Compton Effect ..................................................282
  The de Broglie Wavelength: Observing the Wave Nature of Matter ......285
    Interfering electrons: Confirming de Broglie’s hypothesis ..................286
    Calculating wavelengths of matter ..................................................286
  Not Too Sure about That: The Heisenberg Uncertainty Principle .........288
    Understanding uncertainty in electron diffraction .............................288
    Deriving the uncertainty relation ....................................................289
    Calculations: Seeing the uncertainty principle in action ......................292
# Table of Contents

## Chapter 14: Getting the Little Picture: The Structure of Atoms  . . . . . .295

Figuring Out the Atom: The Planetary Model ........................................ 296
- Rutherford scattering: Finding the nucleus from ricocheting alpha particles ................................................................. 296
- Collapsing atoms: Challenging Rutherford’s planetary model ... 297
- Answering the challenges: Being discrete with line spectra ...... 298

Fixing the Planetary Model of the Hydrogen Atom:
- The Bohr Model ....................................................................................... 301
  - Finding the allowed energies of electrons in the Bohr atom ...... 302
  - Getting the allowed radii of electron orbits in the Bohr atom.... 303
  - Finding the Rydberg constant using the line spectrum of hydrogen ..................................................... 306
  - Putting it all together with energy level diagrams................... 307
  - De Broglie weighs in on Bohr: Giving a reason for quantization ................................................................. 308

Electron Configuration: Relating Quantum Physics and the Atom .... 309
- Understanding four quantum numbers .......................................... 310
- Number crunching: Figuring out the number of quantum states ............................................................................................ 312
- Multi-electron atoms: Placing electrons with the Pauli exclusion principle ............................................................ 314
- Using shorthand notation for electron configuration .................... 316

## Chapter 15: Nuclear Physics and Radioactivity  . . . . . . . . . . . . . . .319

Grooving on Nuclear Structure .............................................................. 319
- Now for a little chemistry: Sorting out atomic mass and number .............................................................................................. 320
- Neutron numbers: Introducing isotopes ........................................ 321
- Boy, that’s small: Finding the radius and volume of the nucleus .............................................................................................. 323
- Calculating the density of the nucleus ........................................ 323

The Strong Nuclear Force: Keeping Nuclei Pretty Stable ............... 324
- Finding the repelling force between protons ................................. 325
- Holding it together with the strong force ....................................... 325
- Hold on tight: Finding the binding energy of the nucleus .......... 327

Understanding Types of Radioactivity, from α to γ .......................... 328
- Releasing helium: Radioactive alpha decay ................................... 330
- Gaining protons: Radioactive beta decay ........................................ 331
- Emitting photons: Radioactive gamma decay .............................. 332

Grab Your Geiger Counter: Half-Life and Radioactive Decay ....... 333
- Halftime: Introducing half-life ........................................................... 334
- Decay rates: Introducing activity ..................................................... 336
Part V: The Part of Tens ...................................................... 339

Chapter 16: Ten Physics Experiments That Changed the World . . . . 341
  Michelson’s Measurement of the Speed of Light.......................... 342
  Young’s Double-Slit Experiment: Light Is a Wave...................... 342
  Jumping Electrons: The Photoelectric Effect ................................ 343
  Davisson and Germer’s Discovery of Matter Waves...................... 343
  Röntgen’s X-rays ....................................................................... 344
  Curie’s Discovery of Radioactivity ............................................. 344
  Rutherford’s Discovery of the Atom’s Nucleus ............................... 345
  Putting a Spin on It: The Stern-Gerlach Experiment ...................... 345
  The Atomic Age: The First Atomic Pile ...................................... 346
  Verification of Special Relativity ............................................... 346

Chapter 17: Ten Online Problem-Solving Tools .............................. 347
  Vector Addition Calculator ...................................................... 347
  Centripetal Acceleration (Circular Motion) Calculator .................. 347
  Energy Stored in a Capacitor Calculator ...................................... 348
  Electrical Resonance Frequency Calculator ............................... 348
  Capacitive Reactance Calculator ............................................... 349
  Inductive Reactance Calculator .................................................. 349
  Frequency and Wavelength Calculator ....................................... 349
  Length Contraction Calculator .................................................. 350
  Relativity Calculator .................................................................... 350
  Half-Life Calculator .................................................................... 351

Index ......................................................................................... 353
Introduction

For many people, physics holds a lot of terror. And Physics II courses do introduce a lot of mind-blowing concepts, such as the ideas that mass and energy are aspects of the same thing, that light is just a mix of electric and magnetic fields, and that every electron zipping around an atom creates a miniature magnet. In Physics II, charges jump, light bends, and time stretches — and not just because your instructor lost the class halfway through the lecture. Throw some math into the mix, and physics seems to get the upper hand all too often. And that’s a shame, because physics isn’t your enemy — it’s your ally.

The ideas may have come from Albert Einstein and other people who managed to get laws and constants and units of measurement named after them, but you don’t have to be a genius to understand Physics II. After all, it’s only partially rocket science — and those are ultra-cool, nearing-the-speed-of-light rockets.

Many breakthroughs in the field came from students, researchers, and others who were simply curious about their world, who did experiments that often didn’t turn out as expected. In this book, I introduce you to some of their discoveries, break down the math that describes their results, and give you some insight into how things work — as physicists understand it.

About This Book

*Physics II For Dummies* is for the inquiring mind. It’s meant to explain hundreds of phenomena that you can observe all around you. For example, how does polarized light really work? Was Einstein really right about time dilation at high speeds? Why do the electromagnets in electric motors generate magnetism? And if someone hands you a gram of radioactive material with a half-life of 22,000 years, should you panic?

To study physics is to study the world. *Your* world. That’s the kind of perspective I take in this book. Here, I try to relate physics to your life, not the other way around. So in the upcoming chapters, you see how telescopes and microscopes work, and you find out what makes a properly cut diamond so
brilliant. You discover how radio antennas pick up signals and how magnets make motors run. You see just how fast light and sound can travel, and you get an idea of what it really means for something to go radioactive.

When you understand the concepts, you see that the math in physics isn’t just a parade of dreadful word problems; it’s a way to tie real-world measurements to all that theory. Rest assured that I’ve kept the math in this book relatively simple — the equations don’t require any knowledge beyond algebra and trigonometry.

*Physics II For Dummies* picks up where a Physics I course leaves off — after covering laws of motion, forces, energy, and thermodynamics. Physics I and Physics II classes have some overlap, so you do find info on electricity and magnetism in both this book and in *Physics For Dummies*. But in *Physics II For Dummies*, I cover these topics in more depth.

A great thing about this book is that you decide where to start and what to read. It’s a reference you can jump into and out of at will. Just head to the table of contents or the index to find the information you want.

**Conventions Used in This Book**

Some books have a dozen stupefying conventions that you need to know before you can start reading. Not this book. All you need to know is the following:

- New terms are given in italics, like *this*, and are followed by a definition.
- Variables, like *m* for *mass*, are in italics. If you see a letter or abbreviation in a calculation and it isn’t italicized, you’re looking at a unit of measurement; for instance, 2.0 m is 2.0 meters.
- Vectors — those items that have both a magnitude and a direction — are given in bold, like this: *B*.

And those are all the conventions you need to know!

**What You’re Not to Read**

Besides the main text of the book, I’ve included some extra little elements that you may find enlightening or interesting: sidebars and paragraphs marked with Technical Stuff icons. The sidebars appear in shaded gray
boxes, and they give you some nice little examples or tell stories that add a little color or show you how the main story of physics branches out. The Technical Stuff paragraphs give you a little more technical information on the matter at hand. You don’t need this to solve problems; you may just be curious.

If you’re in a rush, you can skip these elements without hurting my feelings. Without them, you still get the main story.

**Foolish Assumptions**

In this book, I assume the following:

✓ You’re a student who’s already familiar with a Physics I text like *Physics For Dummies*. You don’t have to be an expert. As long as you have a reasonable knowledge of that material, you’ll be fine here. You should understand ideas such as mass, velocity, force, and so on, even if you don’t remember all the formulas.

✓ You’re familiar with the metric system, or SI (the International System of Units). You can convert between units of measurement, and you understand how to use metric prefixes. I include a review of working with measurements in Chapter 2.

✓ You know basic algebra and trigonometry. I tell you what you need in Chapter 2, so no need to worry. This book doesn’t require any calculus, and you can do all the calculations on a standard scientific calculator.

**How This Book Is Organized**

Like physics itself, this book is organized into different parts. Here are the parts and what they’re all about.

**Part I: Understanding Physics Fundamentals**

Part I starts with an overview of Physics II, introducing the goals of physics and the main topics covered in a standard Physics II course. This part also brings you up to speed on the basics of Physics I — just what you need for this book. You can’t build without a foundation, and you get the foundation you need here.
Part II: Doing Some Field Work: Electricity and Magnetism

Electricity and magnetism are a big part of Physics II. Over the years, physicists have done a great job of explaining these topics. In this part, you see both electricity and magnetism, including info on individual charges, AC (alternating current) circuits, permanent magnets, and magnetic fields — and perhaps most importantly, you see how electricity and magnetism connect to create electromagnetic waves (as in light).

Part III: Catching On to Waves: The Sound and Light Kinds

This part covers waves in general, as well as light and sound waves. Of the two, light is the biggest topic — you see how light waves interact and interfere with each other, as well as how they manage when going through single and double slits, bouncing off objects, passing through glass and water, and doing all kinds of other things. The study of optics includes real-world objects such as lenses, mirrors, cameras, polarized sunglasses, and more.

Part IV: Modern Physics

This part brings you into the modern day with the theory of special relativity, the particle-wave duality of matter, and radioactivity. Relativity is a famous one, of course, and you see a lot of Einstein in this part. You also see many other physicists who chipped in on the discussion of matter’s travels as waves. You read all about radioactivity and atomic structure, too.

Part V: The Part of Tens

The chapters in this part cover ten topics in rapid succession. You take a look at ten physics experiments that changed the world, leading to discoveries in everything from special relativity to radioactivity. You also look at ten online calculators that can assist you in solving physics problems.
**Icons Used in This Book**

You find icons in this book, and here’s what they mean:

**Remember**

This icon marks something to remember, such as a law of physics or a particularly important equation.

**Tip**

Tips offer ways to think of physics concepts that can help you better understand a topic. They may also give you tips and tricks for solving problems.

**Technical Stuff**

This icon means that what follows is technical, insider stuff. You don’t have to read it if you don’t want to, but if you want to become a physics pro (and who doesn’t?), take a look.

**Where to Go from Here**

In this book, you can jump in anywhere you want. You can start with electricity or light waves or even relativity. But if you want the full story, start with Chapter 1. It’s just around the corner from here. Happy reading!

If you don’t feel comfortable with the level of physics taken for granted from Physics I, check out a Physics I text. I can recommend *Physics For Dummies* wholeheartedly.
Part I
Understanding Physics Fundamentals

The 5th Wave
By Rich Tennant

“This is my old physics teacher, Mr. Wendt, his wife Doris, and their two children, Quark and Wormhole.”
In this part . . .

In this part, you make sure you’re up to speed on the skills you need for Physics II. You start with an overview of the topics I cover in this book. You also review Physics I briefly, making sure you have a good foundation in the math, measurements, and main ideas of basic physics.
Chapter 1

Understanding Your World: Physics II, the Sequel

In This Chapter

▶ Looking at electricity and magnetism
▶ Studying sound and light waves
▶ Exploring relativity, radioactivity, and other modern physics

Physics is not really some esoteric study presided over by guardians who make you take exams for no apparent reason other than cruelty, although it may seem like it at times. Physics is the human study of your world. So don’t think of physics as something just in books and the heads of professors, locking everybody else out.

Physics is just the result of a questioning mind facing nature. And that’s something everyone can share. These questions — what is light? Why do magnets attract iron? Is the speed of light the fastest anything can go? — concern everybody equally. So don’t let physics scare you. Step up and claim your ownership of the topic. If you don’t understand something, demand that it be explained to you better — don’t assume the fault is with you. This is the human study of the natural world, and you own a piece of that.

Physics II takes up where Physics I leaves off. This book is meant to cover — and unravel — the topics normally covered in a second-semester intro physics class. You get the goods on topics such as electricity and magnetism, light waves, relativity (the special kind), radioactivity, matter waves, and more. This chapter gives you a sneak preview.
Getting Acquainted with Electricity and Magnetism

Electricity and magnetism are intertwined. Electric charges in motion (not static, nonmoving charges) give rise to magnetism. Even in bar magnets, the tiny charges inside the atoms of the metal cause the magnetism. That’s why you always see these two topics connected in Physics II discussions. In this section, I introduce electricity, magnetism, and AC circuits.

Looking at static charges and electric field

Electricity is a very big part of your world — and not just in lightning and light bulbs. The configuration of the electric charges in every atom is the foundation of chemistry. As I note in Chapter 14, the arrangement of electrons gives rise to the chemical properties of matter, giving you everything from metals that shine to plastics that bend. That electron setup even gives you the very color that materials reflect when you shine light on them.

Electricity studies usually start with electric charges, particularly the force between two charges. The fact that charges can attract or repel each other is central to the workings of electricity and to the structure of the atoms that make up the matter around you. In Chapter 3, you see how to predict the exact force involved and how that force varies with the distance separating the two charges.

Electric charges also fill the space around them with electric field — a fact familiar to you if you’ve ever felt the hairs on your arm stir when you’ve unloaded clothes from a dryer. Physicists measure electric field as the force per unit charge, and I show you how to calculate the electric field from arrangements of charges.

Next up is the idea of electric potential, which you know as voltage. Voltage is the work done per unit charge, taking that charge between two points. And yes, this is exactly the kind of voltage you see stamped on batteries.

With those three quantities — force, electric field, and voltage, you nail down static electric charges.
### Moving on to magnetism

What happens when electric charges start to move? You get magnetism, that’s what. Magnetism is an effect of electric charge that’s related to but distinct from the electric field; it exists only when charges are in motion. Give an electron a push, send it sailing, and presto! You’ve got magnetic field. The idea that moving electric charges cause magnetic field was big news in physics — that fact’s not obvious when you simply work with magnets.

Electric charges in motion form a current, and various arrangements of electric current create different magnetic fields. That is, the magnetic field you see from a single current-bearing wire is different from what you see from a loop of current — let alone a whole bunch of loops of current, an arrangement known as a solenoid. I show you how to predict magnetic field in Chapter 4.

Not only do moving electric charges give rise to magnetic fields, but magnetic fields also affect moving electric charges. When an electric charge moves through a magnetic field, that charge feels a force on it at right angles to the magnetic field and the direction of motion. The upshot is that left to themselves, moving charges in uniform magnetic fields travel in circles (an idea chemists appreciate, because that’s what allows a mass spectrometer to sort out the chemical makeup of a sample). How big is the circle? How does the radius of the circle correlate with the speed of the charge? Or with the magnitude of the charge? Or with the strength of the magnetic field? Stay tuned. The answers to all these questions are coming up in Chapter 4.

### AC circuits: Regenerating current with electric and magnetic fields

Students often meet electrical circuits in Physics I (you can read about simple direct current [DC] circuits in Physics For Dummies). In Chapter 5, you get the Physics II version: You take a look at what happens when the voltage and current in a circuit fluctuate in time in a periodic way, giving you alternating voltage and currents. You also encounter some new circuit elements, the inductor and capacitor, and see how they behave in AC circuits. Many of the electrical devices that people use every day depend on such elements in alternating currents.
In reading about the inductor, you also encounter one of the fundamental laws that relates electric and magnetic fields: Faraday’s law, which explains how a changing magnetic field induces a voltage that generates its own magnetic field. This law doesn’t just apply to inductors; it applies to all electric and magnetic fields, wherever they occur in the universe!

**Riding the Waves**

Waves are a huge topic in Physics II. A *wave* is a traveling disturbance that carries energy. If the disturbance is *periodic*, the amount of disturbance repeats in space and time over a distance called the *wavelength* and a time called the *period*. Chapter 6 delves into the workings of waves so you can see the relationships among the wave’s speed, wavelength, and *frequency* (the rate at which cycles pass a particular point). In the rest of the chapters in Part III of this book, you explore particular types of waves, including electromagnetic waves (such as light and radio waves) and sound.

**Getting along with sound waves**

Sound is just a wave in air, and the various interactions of sound waves are just a result of the behaviors shared by all waves. For instance, sound waves can reflect off a surface — just let sound waves collide with walls and listen for the echo. Sound waves also interfere with other waves, and you can hear the effects — or silence, as the case may be. These two kinds of interaction form the basis for understanding the harmonic tones in music.

The qualities of a sound, such as pitch and loudness, depend on the properties of the wave. As you may have noticed by hearing the change of pitch of a siren on a police car as it passes by, pitch changes when the source or the listener moves. This is called the *Doppler effect*. You can take this to the extreme by examining the shock wave that happens when objects move very quickly through the air, breaking the sound barrier. This is the origin of the sonic boom. I cover all this and more in Chapter 7.

**Figuring out what light is**

You focus on light a good deal in Physics II. How light works is now well-known, but that wasn’t always the case. Imagine the excitement James Clerk Maxwell must’ve felt when the speed of light suddenly jumped out of his
equations and he realized that by combining electricity and magnetism, he’d come up with light waves. Before that, light waves were a mystery — what made them up? How could they carry energy?

After Maxwell, all that changed, because physicists now knew that light was made up of electrical and magnetic oscillations. In Chapter 8, you follow in Maxwell’s footsteps to come up with his amazing result. There, you see how to calculate the speed of light using two entirely different constants having to do with how well electric and magnetic fields can penetrate empty space.

As a wave, light carries energy as it travels, and physicists know how to calculate just how much energy it can carry. That amount of energy is tied to the magnitude of the wave’s electric and magnetic components. You get a handle on how much power that light of a certain intensity can carry in Chapter 8.

Of course, light is only the visible portion of the electromagnetic spectrum — and it’s a small part at that. All kinds of electromagnetic radiation exist, classified by the frequency of the waves: X-rays, infrared light, ultraviolet light, radio waves, microwaves, even ultra-powerful gamma waves.

Reflection and refraction: Bouncing and bending light

Light’s interaction with matter makes it interesting. For instance, when light interacts with materials, some light is absorbed and some reflected. This process gives rise to everything you see around you in the daily world.

Reflected light obeys certain rules. Primarily, the angle of incidence of a light ray — that is, the angle at which the light strikes the surface (measured from a line pointing straight out of that surface) — must equal the angle of reflection — the angle at which the light leaves the surface. Knowing how light is going to bounce off objects is essential to all kinds of devices, from the periscopes in submarines to telescopes, fiber optics, and even the reflectors that the Apollo astronauts placed on the moon. Chapter 10 covers the rules of reflection.

Light can also travel through materials, of course (or people wouldn’t have windows, sunglasses, stained glass, and a lot more). When light enters one material from another, it bends, a process known as refraction — which is a big topic in Chapter 9. The amount the light bends depends on the materials involved, as measured by their indexes of refraction. That’s useful to know in all kinds of situations. For example, when lens-makers understand how light
bends when it enters and leaves a piece of glass, they can shape the glass to produce images. You take a look through lenses next.

**Searching for images: Lenses and mirrors**

If you’re eager to look at the practical applications of Physics II topics, you’ll probably enjoy optics. Here, you work with lenses and mirrors, allowing you to explore the workings of telescopes, cameras, and more.

**Focusing on lenses**

Lenses can focus light, or they can diverge it. In either case, you can get an image (sometimes upright, sometimes upside down, sometimes bigger than the object, sometimes smaller). The image is either virtual or real. In a *real image*, the light rays converge, so you can put a screen at the image location and see the image on the screen (like at the movies). A *virtual image* is an image from which the light appears to diverge, such as an image in a magnifying glass.

Armed with a little physics, you have the lens situation completely under control. If you’re visually inclined, you can find info on the image using your drawing skills. I explain how to draw ray diagrams, which show how light passes through a lens, in Chapter 9.

You can also get numeric on light passing through lenses. The thin-lens equation gives you all you need to know here about the object and image, and you can even derive the magnification of lenses from that equation. So given a certain lens and an object a certain distance away, you can predict exactly where the image will appear and how big it will be (and whether it’ll be upside down or not).

If one lens is good, why not try two? Or more? After all, that’s the idea behind microscopes and telescopes. You get the goods on such optical instruments in Chapter 9, and if you want, you can be designing microscopes and telescopes in no time.

**All about mirrors**

You can get numeric on the way mirrors reflect light, whether a mirror is flat or curved. For instance, if you know just how much a mirror curves and where an object is with reference to the mirror, you can predict just where the image of the object will appear.

In fact, you can do more than that — you can calculate whether the image will be upright or upside down. You can calculate just how high it will be compared to the original object. You can even calculate whether the image will be real (in front of the mirror) or virtual (behind the mirror). I discuss mirrors in Chapter 10.
Calling interference: When light collides with light

Not only can light rays interact with matter; they can also interact with other light rays. That shouldn’t sound too wild — after all, light is made up of electric and magnetic components, and those components are what interact with the electric fields in matter. So why shouldn’t those components also interact with similar electric and magnetic components from other light rays?

When the electric component of a light ray is at its maximum and it encounters a light ray with its electric component at a minimum, the two components cancel out. Conversely, if the two light rays happen to hit just where the electric components are at a maximum, they add together. The result is that when light collides with light, you can get diffraction patterns — arrangements of light and dark bands, depending on whether the net result is at a maximum or minimum. In Chapter 11, you see how to calculate what the diffraction patterns look like for an assortment of different light sources, all of which has been borne out by experiment.

Branching Out with Modern Physics

The 20th century saw an explosion of physics topics, and collectively, those topics are called modern physics. Some revolutionary ideas — such as quantum mechanics and Einstein’s theory of special relativity — changed the foundations of how physicists saw the universe; Isaac Newton’s mechanics didn’t always apply. As physicists delved deeper into the workings of the world, they found more and more powerful ideas, which allowed them to describe exponentially more about the world. This led to developments in technology, which meant that experiments could probe the universe ever more minutely (or expansively).

Most people have heard of relativity and radioactivity, but you may not be familiar with other topics, such as matter waves (the fact that when matter travels, it exhibits many wave-like properties, just like light) or blackbody radiation (the study of how warm objects emit light). I introduce you to some of these modern-physics ideas in this section.

Shedding light on blackbodies: Warm bodies make their own light

If you’ve ever seen an incandescent light bulb at work (or you’ve glanced at the sun), you know that hot things emit light. In fact, any body with any warmth at all emits electromagnetic waves, such as light.
In particular, physicists can calculate the wavelength of the electromagnetic waves where the emitted spectrum peaks, given an object’s temperature. This topic is intimately tied up with photons — that is, particles of light — and you can predict how much energy a photon carries, given its wavelength. Details are in Chapter 13.

**Speeding up with relativity: Yes, \( E = mc^2 \)**

Here it is at last: special relativity and Einstein. What, exactly, does \( E = mc^2 \) mean? It means that matter and energy can be considered interchangeable, and it gives the energy-equivalent of a mass \( m \) at rest. That is, if you have a tomato that suddenly blows up, converting all its mass into energy (not a likely event), you can calculate how much energy would be released. (*Note:* Converting 100 percent of a tomato’s mass into pure energy would create a huge explosion; a nuclear explosion converts only a small percentage of the matter involved into energy.)

Besides \( E = mc^2 \), Einstein also predicted that at high speeds, time stretches and length contracts. That is, if a rocket ship passes you traveling at 99 percent of the speed of light, it’ll appear contracted along the direction of travel. And time on the rocket ship passes more slowly than you’d expect, using a clock at rest with respect to you. So if you watch a rocket ship pass by at high speed, do clocks tick more slowly on the rocket ship than they do for you, or is that some kind of trick? No trick — in fact, the people on the rocket ship age more slowly than you do, too.

Airplanes travel at much slower speeds, but the same effect applies to them — and you can calculate just how much younger a jet passenger is than you (but here’s a disappointing tip to people searching for the fountain of youth: It’s an immeasurably small amount of time). You explore special relativity in Chapter 12.

**Assuming a dual identity: Matter travels in waves, too**

Light travels in waves — that much doesn’t take too many people by surprise. But the fact that matter travels in waves can be a shocker. For example, take your average electron, happily speeding on its way. In addition to exhibiting particle-like qualities, that electron also exhibits wave-like qualities — even so much so that it can interfere with other electrons in flight, just as two light rays can, and produce actual diffraction patterns.
And electrons aren’t the only type of matter that has a wavelength. Everything does — pizza pies, baseballs, even tomatoes on the move. You wrap your mind around this when I discuss matter waves in Chapter 13.

Meltdown! Knowing the αβγ’s of radioactivity

Nuclear physics has to do with, not surprisingly, the nucleus at the center of atoms. And when you have nuclear physics, you have radioactivity.

In Chapter 15, you find out what makes up the nucleus of an atom. You see what happens when nuclei divide (nuclear fission) or combine (nuclear fusion) — and in particular, you see what happens when nuclei decay by themselves, a process known as radioactivity.

Not all radioactive materials are equally radioactive, of course, and half-life — the time it takes for half of a sample to decay — is one good measure of radioactivity. The shorter the half-life, the more intensely radioactive the sample is.

You encounter all the different types of radioactivity — alpha, beta, and gamma — in the tour of the subject in Chapter 15.
Chapter 2

Gearing Up for Physics II

In This Chapter
▶ Mastering units and math conventions
▶ Reviewing foundational Physics I concepts

This chapter prepares you to jump into Physics II. If you’re already a physics ace, there’s no need to get bogged down here — just fly into the physics topics themselves, starting with the next chapter. But if you’re not fast-tracked for the physics Nobel Prize, it wouldn’t hurt to scan the topics here, at least briefly. Doing so can save you a lot of time and frustration in the chapters coming up.

Math and Measurements: Reviewing Those Basic Skills

Physics excels at measuring and predicting the real world, and those predictions often come through math. So to be a physicsmeister, you have to have certain skills down cold. And because this is Physics II, I assume that you’re somewhat familiar with the world of physics and some of those basics already. You look at those skills here in refresher form (if you’re unclear about anything, check out a book like Physics For Dummies (Wiley) to get up to speed).

The following skills are pretty basic; you can’t get through Physics I without them. But make sure you have at least a passing acquaintance with the topics in this section — especially if it’s been quite some time since Physics I.
Using the MKS and CGS systems of measurement

The most common measurement systems in physics are the centimeter-gram-second (CGS) and meter-kilogram-second (MKS) systems. The MKS system is more common. For reference, Table 2-1 lists the primary units of measurement, along with their abbreviations in parentheses, for both systems.

<table>
<thead>
<tr>
<th>Type of Measurement</th>
<th>CGS Unit</th>
<th>MKS Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>Centimeters (cm)</td>
<td>Meters (m)</td>
</tr>
<tr>
<td>Mass</td>
<td>Grams (g)</td>
<td>Kilograms (kg)</td>
</tr>
<tr>
<td>Time</td>
<td>Seconds (s)</td>
<td>Seconds (s)</td>
</tr>
<tr>
<td>Force</td>
<td>Dynes (dyn)</td>
<td>Newtons (N)</td>
</tr>
<tr>
<td>Energy (or work)</td>
<td>Ergs (erg)</td>
<td>Joules (J)</td>
</tr>
<tr>
<td>Power</td>
<td>Ergs/second (erg/s)</td>
<td>Watts (W) or joules/second (J/s)</td>
</tr>
<tr>
<td>Pressure</td>
<td>Baryes (Ba)</td>
<td>Pascals (Pa) or newtons/square meter (N/m²)</td>
</tr>
<tr>
<td>Electric current</td>
<td>Biots (Bi)</td>
<td>Amperes (A)</td>
</tr>
<tr>
<td>Magnetic field</td>
<td>Gausses (G)</td>
<td>Teslas (T)</td>
</tr>
<tr>
<td>Electric charge</td>
<td>Franklins (Fr)</td>
<td>Coulombs (C)</td>
</tr>
</tbody>
</table>

These are the primary measuring sticks that physicists use to measure the world with, and that measuring process is where physics starts. Other measuring systems, such as the foot-pound-second (FPS) system, are around as well, but the CGS and MKS systems are the main ones you see in physics problems.

Making common conversions

Measurements don’t always come in the units you need them in, so doing physics can involve a lot of conversions. For instance, if you’re using the meter-kilogram-second system (see the preceding section), you can’t plug measurements in centimeters or feet into your formula — you need to get them in the right units first. In this section, I show you some values that are equal to each other and an easy way to know whether to multiply or divide when doing conversions.
Looking at equal units
Converting between CGS (centimeter-gram-second) and MKS (meter-kilogram-second) units happens a lot in physics, so here's a list of equal values of MKS and CGS units for reference — come back to this as needed:

- **Length:** 1 meter = 100 centimeters
- **Mass:** 1 kilogram = 1,000 grams
- **Force:** 1 newton = $10^5$ dynes
- **Energy (or work):** 1 joule = $10^7$ ergs
- **Pressure:** 1 pascal = 10 barye
- **Electric current:** 1 ampere = 0.1 biot
- **Magnetism:** 1 tesla = $10^4$ gauss
- **Electric charge:** 1 coulomb = $2.9979 \times 10^9$ franklins

Converting back and forth between MKS and CGS systems is easy, but what about other conversions? Here are a some handy conversions that you can come back to as needed. First, for length:

- 1 meter = 1,000 millimeters
- 1 inch = 2.54 centimeters
- 1 meter = 39.37 inches
- 1 mile = 5,280 feet = 1.609 kilometers
- 1 kilometer = 0.62 miles
- 1 angstrom (Å) = $10^{-10}$ meters

Here are some conversions for mass:

- 1 slug (foot-pound-second system) = 14.59 kilogram
- 1 atomic mass unit (amu) = $1.6605 \times 10^{-27}$ kilograms

These are for force:

- 1 pound = 4.448 newtons
- 1 newton = 0.2248 pounds

Here are some conversions for energy:

- 1 joule = 0.7376 foot-pounds
- 1 British thermal unit (BTU) = 1,055 joules
Part I: Understanding Physics Fundamentals

- 1 kilowatt-hour (kWh) = 3.600 \times 10^6 \text{ joules}
- 1 electron-volt = 1.602 \times 10^{-19} \text{ joules}

And here are conversions for power:

- 1 horsepower = 550 \text{ foot-pounds/second}
- 1 watt = 0.7376 \text{ foot-pounds/second}

Using conversion factors: From one unit to another

If you know that two values are equal to each other (see the preceding section), you easily use them to convert from one unit of measurement to another. Here’s how it works.

First note that when two values are equal, you can write them as a fraction that’s equal to 1. For instance, suppose you know that there are 0.62 miles in a kilometer:

\[ 1 \text{ km} = 0.62 \text{ miles} \]

You can write this as

\[ \frac{1 \text{ km}}{0.62 \text{ mi.}} = 1 \quad \text{or} \quad \frac{0.62 \text{ mi.}}{1 \text{ km}} = 1 \]

Each of these fractions is a conversion factor. If you need to go from miles to kilometers or kilometers to miles, you can multiply by a conversion factor so that the appropriate units cancel out — without changing the value of the measurement, because you’re multiplying by something equal to 1.

For instance, suppose you want to convert 30 miles to kilometers. First, write 30 miles as a fraction:

\[ \frac{30 \text{ miles}}{1} \]

Now you need to multiply by a conversion factor. But which version of the fraction do you use? Here, miles is in the numerator, so to get the miles to cancel out, you want to multiply a fraction that has miles in the denominator. Because \( \frac{1 \text{ km}}{0.62 \text{ mi.}} = 1 \), you can multiply 30 miles by that fraction without changing the measurement. Then the miles on the bottom cancels the miles on the top:

\[ \frac{30 \text{ miles}}{1} \times \frac{1 \text{ km}}{0.62 \text{ miles}} \approx 48 \text{ km} \]
Always arrange your conversion factors so that you cancel out the part of the unit you want to swap out for something else. Each unit that you don’t want in your final answer has to appear in both a numerator and a denominator.

Sometimes you can’t do a conversion in one step, but you can string together a series of conversion factors. For instance, here’s how you can set up a problem to convert 30 miles per hour to meters per second. Notice how I multiply by a series of fractions, making sure that every unit I want to cancel out appears in the numerator of one fraction and the denominator of another.

\[
\frac{30 \text{ mi}}{1 \text{ hour}} \times \frac{1 \text{ km}}{0.62 \text{ mi}} \times \frac{1,000 \text{ m}}{1 \text{ km}} \times \frac{1 \text{ hour}}{60 \text{ min.}} \times \frac{1 \text{ min.}}{60 \text{ s}} = 13 \text{ m/s}
\]

**Doing speedy metric conversions**

In the metric system, one unit can be used as a basis for a broad range of units by adding a prefix (Table 2-2 shows some of the most common prefixes). Each prefix multiplies the base unit by a power of 10. For example, *kilo-* says that the unit is 1,000 times \((10^3)\) larger than the base unit, so a kilometer is 1,000 meters. And *milli-* means the unit is 0.001 times \((10^{-3})\) smaller than the base unit. This means that converting from one metric unit to another is usually a matter of moving the decimal point.

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Symbol</th>
<th>Meaning (Decimal)</th>
<th>Meaning (Power of Ten)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nano-</td>
<td>n</td>
<td>0.000000001</td>
<td>(10^{-9})</td>
</tr>
<tr>
<td>Micro-</td>
<td>μ</td>
<td>0.000001</td>
<td>(10^{-6})</td>
</tr>
<tr>
<td>Milli-</td>
<td>m</td>
<td>0.001</td>
<td>(10^{-3})</td>
</tr>
<tr>
<td>Centi-</td>
<td>c</td>
<td>0.01</td>
<td>(10^{-2})</td>
</tr>
<tr>
<td>Kilo-</td>
<td>k</td>
<td>1,000</td>
<td>(10^3)</td>
</tr>
</tbody>
</table>

By finding the difference in exponents on the power of 10 of your original units and the units you want to convert to, you can figure out how many places to move the decimal point.

For instance, say you have a distance of 20.0 millimeters, and you’d prefer to express it in centimeters. You know that 1 millimeter is \(10^{-3}\) meters, and 1 centimeter is \(10^{-2}\) meters (as Table 2-2 shows). If you find the difference in exponents, you see that \(-3 - (-2) = -1\). The answer is negative, so you just have to move the decimal point one place to the left (for a positive answer, you move it to the right). Thus, 20.0 millimeters is equal to 2.00 centimeters.
Using temperature-conversion equations

You can use the following equations to convert between the different units of temperature:

✓ Kelvin temperature = Celsius temperature + 273.15
✓ Celsius temperature = \( \frac{5}{9} \) (Fahrenheit temperature – 32°)

Keeping it short with scientific notation

Physicists often delve into the realms of the very small and the very large. Fortunately, they also have a very neat way of writing very large and very small numbers: Scientific notation. Essentially, you write each number as a decimal (with only one digit to the left of the decimal point) multiplied by 10 raised to a power.

Say you want to write down the speed of light in a vacuum, which is about three hundred million meters per second. This is a three followed by eight zeros, but you can write it as just a 3.0 multiplied by \( 10^8 \):

\[
300,000,000 \text{ m/s} = 3.0 \times 10^8 \text{ m/s}
\]

You can write small numbers by using a negative power to shift the decimal point to the left. So if you have a distance of 4.2 billionths of a meter, you could write it as

\[
0.0000000042 \text{ m} = 4.2 \times 10^{-9} \text{ m}
\]

Note how the \( 10^{-9} \) moves the decimal point of the 4.2 nine places to the left.

Brushing up on basic algebra

To do physics, you need to know basic algebra. You’re going to be slinging some equations around, so you should be able to work with variables and move them from one side of an equation to the other as needed, no problem.

You don’t need to be bogged down or daunted by the formulas in physics — they’re only there to help describe what’s going on in the real world. When you see a new formula, consider how the different parts of the equation relate to the physical situation it describes.

Take a simple example — the equation for the speed, \( v \), of an object that covers a distance \( \Delta x \) in a time \( \Delta t \) (Note: The symbol \( \Delta \) means “change in”):

\[
v = \frac{\Delta x}{\Delta t}
\]
Before you go any further, try relating the parts of this equation to what you intuitively understand about speed. You can see in the equation that if $\Delta x$ increases, then $v$ increases — if you cover a greater distance in a given time, then you’re traveling faster. You can also see that if $\Delta t$ (in the denominator of the fraction) increases, then $v$ decreases — if it takes you longer to cover a given distance, then you’re moving more slowly.

If you need to, you can rearrange an equation algebraically to isolate the part you’re interested in. That way, you can get a feel for how other variables affect each other. For instance, see what the equation means for travel time by rearranging it to isolate $\Delta t$:

$$\Delta t = \frac{\Delta x}{v}$$

Now you can see that $\Delta t$ increases as $\Delta x$ increases, and $\Delta t$ decreases as $v$ increases. This just means that travel time increases if you have to travel farther and decreases if you travel faster.

**Using some trig**

You work with some angles to this book — such as those you have to figure out when light bounces off mirrors or bends in lenses. To handle angles and related distances, you need some trigonometry.

 Pretty much everything in trig comes down to the right triangle. For example, take a look at the right triangle in Figure 2-1. The two shorter sides, or legs, are called $x$ and $y$ (because they lie along the $x$- and $y$-axes respectively), and the longest side, across from the $90^\circ$ angle, is the hypotenuse. One of the other internal angles is marked $\theta$.

![Figure 2-1: The two legs ($x$ and $y$) and hypotenuse ($h$) of a right triangle.](image)

Here’s one important formula to know: the Pythagorean theorem. It relates the lengths of $x$, $y$, and $h$, so given the lengths of two sides, you can find the length of the third:

$$x^2 + y^2 = h^2$$
To work with angles (such as $\theta$), you need the trig functions sine, cosine, and tangent. To find the values of trig functions, just divide one side of the triangle by another. Here’s what it looks like:

\[ \text{Sine: } \sin \theta = \frac{y}{h} = \frac{\text{opposite side}}{\text{hypotenuse}} \]
\[ \text{Cosine: } \cos \theta = \frac{x}{h} = \frac{\text{adjacent side}}{\text{hypotenuse}} \]
\[ \text{Tangent: } \tan \theta = \frac{y}{x} = \frac{\text{opposite side}}{\text{adjacent side}} \]

Note that these equations relate any two sides of a right triangle to the angle that’s between the hypotenuse and one of the other sides. So if you know $\theta$ and one of the other sides, you can use some algebra (and your calculator) to find the length of any other side.

To find the angle $\theta$, you can go backward with inverse sines, cosines, and tangents, which are written like this: $\sin^{-1}$, $\cos^{-1}$, and $\tan^{-1}$. If you make the appropriate fraction out of two known sides of a triangle and take the inverse sine of that ($\sin^{-1}$ on your calculator), it’ll give you back the angle itself. Here’s how the inverse trig functions work (see Figure 2-1 for which sides are $x$, $y$, and $h$):

\[ \text{Inverse sine: } \sin^{-1} \left( \frac{y}{h} \right) = \theta \]
\[ \text{Inverse cosine: } \cos^{-1} \left( \frac{x}{h} \right) = \theta \]
\[ \text{Inverse tangent: } \tan^{-1} \left( \frac{y}{x} \right) = \theta \]

Physicists use sine and cosine functions to describe real-world waves and alternating current and voltage. I introduce waves in Chapter 6, and I cover alternating current (AC circuits) in Chapter 5.

**Using significant digits**

You may be surprised to hear that physics isn’t an exact science! It can be pretty accurate, but nothing is ever measured perfectly. The more accurately the quantity is measured, the more digits you know. The digits you know are the *significant figures*. For instance, a stopwatch measurement of 11.26 seconds has four significant figures. Here are a few guidelines for figuring out what’s significant:
For a decimal less than 1, everything that follows the first nonzero digit is significant. For example, 0.0040 has two significant digits.

For a decimal greater than 1, all digits, including zeros after the decimal point, are significant. For instance, 20.10 has four significant digits.

For a whole number, the non-zero digits are significant. Any number of trailing zeros also may be significant.

So how do you show the accuracy of a measurement such as 1,000 meters, which ends in zeros? You may know anywhere between one and four digits from your measurements. The best way to clear this up is to use scientific notation. For instance, if you write 1,000 as $1.000 \times 10^3$, with three zeroes after the decimal point, the number has four significant figures — you measured to the nearest meter. If you write it as $1.00 \times 10^3$, with two zeroes after the decimal point, the number has three significant figures — you measured to the nearest ten meters. (For info on scientific notation, see the earlier section “Keeping it short with scientific notation.”)

When you do calculations with numbers that are known only to a particular accuracy, then your answer is also only of a particular accuracy. After you do all your calculations, you need to round the answer. Here are some simple rules you can apply:

- **If you multiply or divide two numbers:** The answer has the same number of significant figures as the least-accurate of the two numbers being multiplied or divided. For example, consider the following calculation:

  \[ 12.45 \times 0.050 = 0.6225 \]

  Because 0.050 has two significant figures, you round the answer to 0.62.

- **If you add or subtract two numbers:** The answer has the same number of decimal places as the least-accurate of the two numbers you’re adding or subtracting. For example, consider

  \[ 11.432 + 1.3 = 12.732 \]

  Because the least-accurate number, 1.3, has only one decimal place, write the answer as 12.7.

**Refreshing Your Physics Memory**

To make progress, physics often builds on previous physics advances. For example, knowing about vectors is important not just to handle problems with acceleration (that’s Physics I) but also to help you track charged particles in magnetic fields (that’s Physics II).
In this section, you take a down-memory-lane tour of some Physics I concepts that pop up again in Physics II. If you don’t feel comfortable with these topics, check out a physics text to make sure you’re up to speed in Physics I before proceeding.

**Pointing the way with vectors**

Vectors are the physics way of pointing a direction. A vector has a direction and a magnitude (size) associated with it — the magnitude is the vector’s length.

You usually see the names of vectors in bold in physics. Figure 2-2 shows vector \( \mathbf{A} \). That’s just a standard vector, and it may stand for, say, the direction an electron is traveling in. The length of the vector may indicate the speed of the electron — the faster the electron is going, the longer the vector.

![Figure 2-2: The components of a vector.](image)

You don’t see lots of vectors in this book (did I just hear a sigh of relief?), but you should know how to break a vector like \( \mathbf{A} \) up into its components along the \( x \) and \( y \)-axes (you need to do this in Chapter 4 for the magnetic field and in Chapter 5 for alternating currents and voltages).

If you’re given the length of the vector (its magnitude, labeled \( A \) in Figure 2-2) and the angle \( \theta \) (its direction), breaking a vector into its components works like this:

\[
\begin{align*}
    A_x &= A \cos \theta \\
    A_y &= A \sin \theta
\end{align*}
\]
where \( A_x \) is the \( x \) component of vector \( A \) and \( A_y \) is the vector’s \( y \) component. (This is really just a bit of trig, where \( A_x \) and \( A_y \) are the legs of the triangle and \( A \) is the hypotenuse — see the earlier section “Using some trig” for info on the sine and cosine functions.)

Resolving vectors into components is particularly useful if you have to add two vectors, \( \mathbf{A} + \mathbf{B} \). You break them up into their separate components and then add those components to get the components of the vector sum, which is a new vector you can call \( \mathbf{C} \):

\[
\begin{align*}
\mathbf{C}_x &= A_x + B_x \\
\mathbf{C}_y &= A_y + B_y
\end{align*}
\]

When you have the components of a vector like \( \mathbf{C} \), you can covert them into a length (magnitude) for \( \mathbf{C} \) (written as \( |\mathbf{C}| \)) and an angle for \( \mathbf{C} \) this way:

**Magnitude of \( \mathbf{C} \):**

\[
|\mathbf{C}| = \sqrt{C_x^2 + C_y^2}
\]

*Note:* This is just the Pythagorean theorem solved for the hypotenuse \( |\mathbf{C}| \).

**Direction of \( \mathbf{C} \):**

\[
\theta = \tan^{-1}\left(\frac{C_y}{C_x}\right)
\]

See the earlier section “Using some trig” for info on inverse trig functions.

So now you’re able to go from representing a vector in terms of its length and angle to its components and then back again — a very handy skill to have.

### Moving along with velocity and acceleration

This book has a little to say about velocity and acceleration. For example, you work with them when a magnetic field diverts electrically charged particles from the direction in which they’re traveling.

Both velocity and acceleration are vectors, \( \mathbf{v} \) and \( \mathbf{a} \) respectively. *Velocity* is the change in the position-vector divided by the time that change took. For example, if the position of a ping-pong ball is given by the position-vector \( \mathbf{x} \), then the change in the position (\( \Delta \mathbf{x} \)) divided by the amount of time that change took (\( \Delta t \)) is the velocity:

\[
\mathbf{v} = \frac{\Delta \mathbf{x}}{\Delta t}
\]

As a vector, velocity has a direction. The magnitude of the velocity vector is the *speed*, which has a size but not a direction. That is, velocity is a vector, but speed isn’t.
If the velocity isn’t staying constant, the ping-pong ball is undergoing acceleration. Acceleration is defined as the change in velocity divided by the time that change takes, or

$$ a = \frac{\Delta v}{\Delta t} $$

Note that a change in direction is considered a change in velocity, so something can be accelerating even if its speed doesn’t change.

Velocity is commonly measured in meters per second (m/s) — which means that acceleration’s units are commonly meters per second squared (m/s²).

**Strong-arm tactics: Applying some force**

When an electron enters an electric field, it gets pushed one way or another — that is, it experiences a force. Physics I has a lot to say about force — for example, here’s the famous equation that relates total force ($F$), mass ($m$), and acceleration ($a$) (note that acceleration and force are both vectors):

$$ F = ma $$

So to find out how much force is acting on the electron to push it along (and you don’t need much, because electrons don’t weigh very much), you’d put in the electron’s acceleration and its mass, and you’d get the total force acting on it. The formula also shows that applying a force to something can make it accelerate, and you see that idea used every now and then in this book.

The units of force you see most commonly are newtons (in the meter-kilogram-second system), symbol N, named for Sir Isaac Newton (the fellow with the falling apple acted on by the force of gravity).

**Getting around to circular motion**

Charged particles in magnetic fields travel in circles, so you need to know something about circular motion in Physics II. Physics I has plenty to say about circular motion. For example, take a look at Figure 2-3, where an object is traveling in circular motion.

The velocity of an object moving in a circle points along the circle of its path — this is called the tangential direction. The force that keeps the object moving in a circle points toward the center of the circle — in a direction that’s at right angles to the velocity. For instance, when you spin a ball on a string, the string can exert a force on the ball only in the direction that’s along its length, perpendicular to the path of the ball; this is what causes the ball to move in a circular path.
The angle that an object moving in circular motion covers in so many seconds is its **angular velocity**, $\omega$:

$$\omega = \frac{\Delta \theta}{\Delta t}$$

Here, the angle $\theta$ is measured in radians, so the units of angular velocity are radians/second. *(Note: Exactly $2\pi$ radians are in a complete circle, which means that $2\pi$ radians equals 360°, or each radian is $360^\circ \div 2\pi$ degrees.)*

If the object is speeding up or slowing down, it’s undergoing **angular acceleration**, which is given the symbol $\alpha$. **Angular acceleration** is defined as the change in angular velocity ($\Delta \omega$) divided by the time that change took ($\Delta t$):

$$\alpha = \frac{\Delta \omega}{\Delta t}$$

The units of angular acceleration are radians/second$^2$.

In circular terms, force becomes **torque**, with the symbol $\tau$ (also a vector, of course), where the magnitude of torque equals force multiplied by distance and the sine of the angle between them:

$$\tau = Fr \sin \theta$$

And the counterpart of mass in circular terms is the **moment of inertia**, $I$. Newton’s law, force = mass $\times$ acceleration, becomes this in circular terms:

$$\tau = I\alpha$$

That is, torque = moment of inertia $\times$ angular acceleration.
Even linear kinetic energy has an alter ego in the circular world, like this:

\[ KE = \frac{I\omega^2}{2} \]

You can have angular momentum, \( L \), as well:

\[ L = I\omega \]

**Getting electrical with circuits**

Physics I introduces the idea of circuits, at least simple circuits with batteries. The rules of resistance and Kirchhoff’s rules, which I review in this section, form the basis for describing the currents and voltages in circuits. You need these rules whenever you work out the various currents and voltages. For example, in Chapter 5, you use them for a simple circuit with three elements in series. You can find a more thorough description of these rules in *Physics For Dummies*.

According to Ohm’s law, you can determine the current going through any resistor with the following equation, where \( I \) is the current measured in amperes, \( V \) is the voltage across the resistor measured in volts, and \( R \) is the resistance of the resistor measured in ohms (Ω):

\[ I = \frac{V}{R} \]

That helps with individual resistors, but what about when they’re assembled into a circuit as Figure 2-4 shows? There, you can see three resistors with resistances of 2 Ω, 4 Ω, and 6 Ω. The currents in each wire, \( I_1 \), \( I_2 \), and \( I_3 \), are driven by the two batteries, which generate voltages of 12 volts and 6 volts.

![Figure 2-4: A circuit with two loops.](image)
To solve for the currents and voltages, you use Kirchoff’s rules:

✓ **The loop rule:** The sum of voltages ($\Sigma V$) around a loop — any loop in the circuit — is zero:

$$\Sigma V = 0 \text{ around a loop}$$

✓ **The junction rule:** The sum of all currents ($\Sigma I$) into any point in the circuit must equal the sum of all currents out of that point (that is, the net sum of all currents into and out of any point in the circuit must be zero):

$$\Sigma I = 0 \text{ at any point in the circuit}$$
Part I: Understanding Physics Fundamentals
Part II

Doing Some Field Work: Electricity and Magnetism

The 5th Wave

By Rich Tennant

Hey—I just found my first real world application of studying Physics.
In this part . . .

Physicists have long been friends with electricity and magnetism. In this part, you see all about electric field, charges, the force between charges, electric potential, and more. You also explore magnetism, such as the magnetic field from a wire, the force between two wires, how charged particles orbit in magnetic fields, and the like. And you look at AC circuits, in which magnetic and electric fields work together to regenerate the current.
Chapter 3
Getting All Charged Up with Electricity

In This Chapter
▶ Understanding charges
▶ Examining electric forces and Coulomb’s law
▶ Finding electric field
▶ Finding electric potential
▶ Understanding electric potential and capacitors

This chapter is dedicated to things that go zap. Chances are that your day-to-day life would be very different without electrical appliances from the computer to the light bulb. But electricity is even more important than that; it’s a physical interaction that’s fundamental to how the whole universe works. For example, chemical reactions are all basically of an electrical nature, and without electricity, atomic matter — the world as you know it — could not exist.

Even though electricity is integral to the existence of all the complicated constructions of matter and chemistry, it has a simple and beautiful nature. In this chapter, you see what makes static electricity, electric fields, and electric potential work.

Understanding Electric Charges

Where does electric charge come from? It turns out to be built into all matter. An atom is made up of a nucleus and electrons in orbit around the nucleus, and the nucleus is made up of protons and neutrons. The protons have a positive charge (+), and the electrons a negative charge (−). So you have electric charges inside any piece of matter you care to name.
What can you do with that charge? You can separate charges from each other and so charge up objects with an excess of one charge or another. Those separated charges exert forces on each other. In this section, I discuss all these concepts — and more — about electric charges.

**Can’t lose it: Charge is conserved**

Here’s an important fact about charge: Just as you can’t destroy or create matter, you can’t create or destroy charge. You have to work with the charge you have.

Because charge can’t be created or destroyed, physicists say that charge is **conserved**. That is, if you have an isolated system (that is, no charge moves in or out of the system), the net charge of the system stays constant.

Notice that the conservation of charge says the net charge remains constant — that is, the sum of all charges stays the same. The actual distribution of charges can change, such as when one corner of a system becomes strongly negatively charged and another corner becomes positively charged. But the sum total of all charge — whether you’re talking about the whole universe or a smaller system — remains the same. No charge in or out means the charge of the system remains constant.

**Measuring electric charges**

Electric charges are measured in **coulombs** (C) in the MKS system, and each electron’s charge — or each proton’s charge — is a tiny amount of coulombs. The proton’s charge is exactly as positive as the electron’s charge is negative.

The electric charge of a proton is named $e$, and the electric charge of an electron is $-e$. How big is $e$? It turns out that

\[ e = 1.60 \times 10^{-19} \text{ C} \]

That’s really tiny. Here’s how many electrons make up 1 coulomb:

\[
\frac{1 \text{ coulomb}}{1.60 \times 10^{-19} \text{ coulombs/electron}} = 6.25 \times 10^{18} \text{ electrons}
\]

So there are $6.25 \times 10^{18}$ electrons in 1 coulomb of (negative) charge.
**Opposites attract: Repelling and attracting forces**

An uncharged atom is composed of as many electrons as protons. These stay together because of the mutual attraction of these positively and negatively charged components. That’s why all ordinary matter doesn’t instantly disintegrate before your very eyes.

Objects that have the same charge (– and – or + and +) exert a repelling force on each other, and objects that have opposite charges (+ and –) attract each other (you’ve always heard that opposites attract, right?). For example, take a look at Figure 3-1, where two suspended ping-pong balls are being charged in various ways.

In Figure 3-1a, the two ping-pong balls have the same, negative charge. They’re exerting a repelling force on each other, forcing them apart. And the funny thing is that they’ll stay that way indefinitely, no additional action or batteries needed — the static charge on each ping-pong ball just stays there.

Actually, it’s not totally true that the charge on each ping-pong ball just stays there. In fact, charge is continually being transferred from charged objects to the water molecules in the air, which carry it off. On humid days, charged objects actually lose their charge faster.

In Figure 3-1b, the two ping-pong balls have opposite charges, – and +, so they attract each other. Note that if they were to touch, charge would flow and the two ping-pong balls would each end up with the same charge. If the + charge is the same magnitude as the – charge, that means that the balls would end up electrically neutral and would just hang down straight.

In Figure 3-1c, the two ping-pong balls have the same positive charge, so once again they repel each other. The force repelling the two ping-pong balls is the same as that in Figure 3-1a if the positive and negative charges have the same magnitude (just different signs).
Getting All Charged Up

In this section, you see how to deliver charge to objects. I let it sit there a while, and you experience charge in a way that really makes your hair stand on end: static electricity. I also transfer charge from object to object, and I let it flow nice and smoothly through wires.

Static electricity: Building up excess charge

You may not be able to create or destroy charges, but you can move them around and create imbalances within a system. When you charge up an object, you keep adding more and more charges to the object, and if that charge has nowhere to go, it just accumulates. Static electricity is the kind of electricity that comes from this excess charge.

Everyone’s familiar with the unpleasant experience of walking across a rug and then getting zapped by a doorknob. What’s actually happening there? Turns out that you’re actually picking up spare electrons from the rug. Your
body collects an excess of electrons, and when you touch the doorknob, they flow out of you. Ouch!

Protons, being bound inside the nucleus, aren’t really free to flow through matter, so when you charge something, it’s usually the electrons that are moving around and redistributing themselves. When something gets charged negatively, electrons are added to it. When it gets charged positively, electrons are taken away, leaving the protons where they are, and the net surplus of protons makes a positive charge.

Before you’re zapped, that excess of charge is static electricity: It’s *electricity* because it’s made up of electrical charges, and it’s called *static electricity* when it isn’t flowing anywhere. When you’re charged with static electricity, each of the hairs on your head carries a share of this excess of electrons. You may develop a spiky new hairstyle as each hair repels its neighbor (which has the same charge). Your hair quickly returns to normal if you touch something that the excess electrons can flow into. They quickly rush through the contact point, giving you the shock.

Though charge can flow through your fingertip, you usually find it flowing through wires in a circuit, where it doesn’t build up. In circuits, charge doesn’t collect and remain stationary, because it’s always free to flow (however, I show you an exception to this idea in the later section “Storing Charge: Capacitors and Dielectrics”).

But when electricity gets blocked and yet still piles up — it can’t go anywhere — then you have static electricity. If the electricity in a circuit is like a river of electricity that keeps flowing around and around (kept in motion by, say, a battery), then static electricity is like a river of electricity that’s dammed up — but charges keep getting added. So although charges don’t build up in circuits, they do build up when you have static electricity.

**Checking out charging methods**

In this section, I cover two ways to charge objects: by contact and by induction. These are simple physical mechanisms that can help you understand how charge behaves and how it can be redistributed.

**Charging by contact**

Charging by contact is the simplest way of charging objects — you just touch the object with something charged and *zap!* The object becomes charged. No big mystery here.

For example, take a look at Figure 3-2 — a negatively charged rod is brought into contact with a ball that’s originally neutral. The result? The ball is left with a negative charge. That’s because the electrons in the rod are always
pushing each other (because like charges repel), so they’re always looking for ways to redistribute themselves farther apart. When the rod comes in contact with the ball, some of the electrons take the opportunity to slip off the rod and onto the ball. Presto! The ball gets charged. (Note: For this to happen, the electrons need to be free to flow through the materials, which can happen if the materials are conductors. Materials that don’t allow electrons to flow through them are insulators. I discuss both types of materials later in “Considering the medium: Conductors and insulators.”)

Touch a negatively charged glass rod to a neutral ping-pong ball, and the ball acquires a negative charge by contact. But you may wonder how to charge the glass rod in the first place. You can do this in many ways, but the simplest and oldest way is to take a glass rod and some silk and rub the two together. A transfer of electrons from one material to the other occurs due to molecular forces between the two types of material. Different materials have different propensities to exchange electrons — you may have noticed that a balloon and a wool sweater work well.

**Charging by induction**

You can deliver charge to an object indirectly using induction. Here’s how charging by induction works: You bring a charged rod close to a neutral object. Say the rod is charged negatively — the negative charges (electrons) in the neutral object are repelled to the opposite side of the object, leaving a net positive charge close to the rod, as Figure 3-3 shows at the top.

Now comes the clever part: You connect the far side of the object to the ground. Just connect a wire from the far side of the object to the actual Earth, which acts as a huge reservoir of charge. The negative charges — the electrons — that are being forced to the far side of the object are frantic...
to get off the object, because the charge on the rod is repelling them. By connecting the far side of the object to the ground with a wire, you provide those electrons with an escape route. And the electrons take that escape route by the millions and trillions.

Then you cleverly disconnect the wire from the far side of the object. The electrons that wanted to get away have fled — and now there’s nowhere else for any other charges to go. The result is that the object is left with a positive charge, because you’ve drained off much of the negative charge that was being pushed by the rod. And you haven’t lost any of the charge on the glass rod.

The upshot is that the object is left with the opposite charge of the rod. And that’s charging by induction. Pretty cool, eh?

![Figure 3-3: Charging by induction.](image)

Lightning rods work through induction. In thunderclouds, charges get separated from the top to the bottom of the cloud, so the top and the bottom of the cloud become strongly charged. When lightning strikes, the charge on the bottom of the cloud is zapping the Earth. If you have a lightning rod, the strong charge on the bottom of the cloud induces the opposite charge on the lightning rod (which is connected to the ground). When lightning strikes, it’s attracted to that opposite charge and hits the lightning rod.

**Considering the medium: Conductors and insulators**

You’re probably familiar with the concepts of electrical conductors — like the copper wire in an extension cord — and electrical insulators — like the plastic that coats the electrical wire and prevents the electricity in the wire
from delivering a nasty shock. In this section, you take a closer look at conductors and insulators in physics terms.

Say you have two charged objects some distance apart from each other. They’re not losing charge; they’re just sitting there. Then you bring a piece of rubber and touch both of them with the rubber. What happens? Nothing, because rubber is an electrical insulator; electricity is conducted through rubber only with difficulty.

Now say that you bring a copper bar in contact with the two objects — immediately, charge flows from one to the other, because copper is an electrical conductor.

Good electrical conductors consist of atoms for which the outermost electrons are not very tightly bound, so they can easily hop from atom to atom and take part in an electric current. The electrons in the very outer orbit around the nucleus are called valence electrons, and those are the ones that can detach from atoms and roam freely through the conductor. (Interestingly, materials that are good electrical conductors, like most metals, are usually also good thermal conductors.)

Current is always defined as the direction of the flow of positive charges, but in reality, it’s the electrons that do the moving and therefore transport electrical charge. In this case, electrons move from the negatively charged object to the positively charged object. But if you want to draw the direction of the current, that goes from the positively charged object to the negatively charged one. Historically, this convention was adopted before people knew that electrons, not positive charges, carry the current.

**Coulomb’s Law: Calculating the Force between Charges**

Coulomb’s law is one of the physics biggies. That’s the same Coulomb (Charles-Augustin de Coulomb) that the unit of charge, the coulomb, is named after, so you know Coulomb’s law has to be some serious stuff.

And serious stuff it is: Charges can attract or repel each other, and Coulomb’s law lets you calculate the exact force that two point charges a certain distance away will exert on each other. A point charge just has all its charge concentrated in a single point, with no surface area that the charge can be distributed over. Point charges are particularly beloved by physicists because they’re easy to work with.
Say you have two point charges, with opposite signs, attracting each other from a certain distance \( r \) apart. What’s the force between the two charges? Coulomb has the answer: His law says that if the charge of one charge is \( q_1 \) and the charge of the other charge is \( q_2 \), then the force between the two charges is

\[
F = \frac{kq_1 q_2}{r^2}
\]

In this equation, \( k \) is a constant, and its value is \( 8.99 \times 10^9 \) N-m\(^2\)/C\(^2\); \( q_1 \) and \( q_2 \) are the charges, in coulombs (C), of the charged objects doing the attracting or repelling; \( r \) is the distance between the charges; and \( F \) is the electrostatic force between the charges.

Force is a vector, so when you’re looking at point charges, the direction of the force is always along a line between the two charges (assuming there are only two) and

- Toward each other if the charges have opposite signs (that is, the force has a negative sign)
- Away from each other if the charges have opposite signs (that is, the force has a positive sign)

## Introducing Electric Fields

Electric field is the field in space created by electric charges. When two charges attract or repel each other, their electric fields are interacting.

Charge can be distributed in many ways. You can have point charges, sheets of charge, cylinders of charge, and many more configurations. Coulomb’s law, which I discuss earlier in “Coulomb’s Law: Calculating the Force between Charges,” works only for point charges. So how do you handle force for other distributions of charge? You often use electric fields, which I discuss in this section.

## Sheets of charge: Presenting basic fields

How do you calculate force for a sheet of charge? Instead of modifying Coulomb’s law to handle sheets of charge, you can simply measure the force that the sheet of charge exerts on a small positive test charge. From there, you know how much force per coulomb the sheet of charge is capable of exerting, and when you have your own charge, which may be positive or negative, you can simply multiply the force per coulomb by the size of your charge.
The idea of measuring force per coulomb to handle non-point charges got to be very popular and became known as the electric field. Here’s the definition: Electric field \( E \) is the force \( F \) that a small test charge would feel due to the presence of other charges, divided by the test charge \( q \):

\[
E = \frac{F}{q}
\]

Electric field’s units are newtons per coulomb (N/C), and electric field is a vector. The direction of the electric field at any point is the force that would be felt by a positive test charge.

What’s that in plain English? Electric field is just the force per coulomb that a charge would feel at any point in space. You divide out the test charge to leave you just newtons per coulomb, which you can multiply by your own charge to determine the force that charge would feel.

For example, say you’re walking on a wool carpet and pick up a static electricity charge of \(-1.0 \times 10^{-6}\) coulombs. You suddenly encounter a \(5.0 \times 10^6\) N/C electric field in the opposite direction in which you’re walking, as Figure 3-4 shows.

![Figure 3-4: The force on a charge in an electric field.](image)

How big of a force do you feel? Well, the electric field is \(5.0 \times 10^6\) newtons per coulomb and you have a charge of \(-1.0 \times 10^{-6}\) coulombs, so you get the following:

\[
F = qE = (-1.0 \times 10^{-6}\text{ C})(5.0 \times 10^6\text{ N/C}) = -5.0\text{ N}
\]

That is, you feel a force of 5.0 N, and the minus sign means that the force is in the direction opposite of the electric field. That’s a little more than 1 pound of force.

So that’s what electric field tells you — how much electric influence is in a given region, ready to cause a force on any charge you bring into the electric field.
Note that the electric field has a direction. How can you tell which way the force that an electric field causes will push the charge you bring into the electric field? You can do this the hard way, with formal definitions, or you can use the easy way. I prefer the easy way. Just think of electric field as coming from positive charges — that is, electric field arrows always do the following:

- Point away from any positive charges that create the electric field
- Go into negative charges

So you can always think of a bunch of positive charges as sitting at the base of the electric field arrows, and that tells you which way the force will act on the charge you bring into the electric field. For example, you have a negative charge in Figure 3-4, and you can think of the electric field arrows as coming from a bunch of positive charges — and because like charges repel, the force on your charge is away from the base of the arrows.

**Looking at electric fields from charged objects**

Not all electric fields are going to be as polite and evenly spaced as the electric field associated with the sheet of charge you see in Figure 3-4. For example, what’s the electric field from a point charge?

Say that you have a point charge $Q$ and a small test charge $q$. How do you find the force per coulomb? Coulomb’s law to the rescue here — just plug in the charges $Q$ and $q$ (for $q_1$ and $q_2$) and the distance between them to get the size of the force (see the earlier section “Coulomb’s Law: Calculating the Force between Charges” for more on this formula):

$$F = \frac{kQq}{r^2}$$

So what’s the electric field? You know that $E = F/q$, so all you have to do is to divide by your test charge, $q$, to get the following:

$$E = \frac{F}{q} = \frac{kQ}{r^2}$$

So the electric field at a given place falls off by a magnitude of $r^2$, the square of the distance away from a point charge.

And what about the direction of the electric field? Well, the force exerted by a point charge is radial (that is, toward or away from the point charge). And electric field emanates from positive charges and goes into negative charges, so Figure 3-5 shows you what the electric field looks like for a positive point charge and a negative point charge.
Figure 3-5: The electric field from two point charges.

**Uniform electric fields: Taking it easy with parallel plate capacitors**

The electric field between multiple point charges isn’t the easiest thing to come to grips with in terms of vectors. So to make life easier, physicists came up with the parallel plate capacitor, which you see in Figure 3-6.

A *parallel plate capacitor* consists of two parallel conducting plates separated by a (usually small) distance. A charge $+q$ is spread evenly over one plate and a charge $-q$ is spread evenly over the other. That’s great for physicists’ purposes, because the electric field from all the point charges on these plates cancels out all components except the ones pointing between the plates, as you see in Figure 3-6.
So by being clever, physicists arrange to get a constant electric field, all in the same direction, which is a heck of a lot easier to work with than the field from point charges.

So what is the electric field between the plates? You can determine that the electric field \( E \) between the plates is constant (as long as the plates are close enough together), and in magnitude, it’s equal to

\[
E = \frac{q}{\varepsilon_0 A}
\]

where \( \varepsilon_0 \), the so-called permittivity of free space, is \( 8.854 \times 10^{-12} \text{ C}^2/(\text{N} \cdot \text{m}^2) \); \( q \) is the total charge on either of the plates (one plate has charge \( +q \) and the other is \( -q \)); and \( A \) is the area of each plate in square meters.

The equation for the electric field \( E \) between the plates of a parallel plate capacitor is often written in terms of the charge density, \( \sigma \), on each plate, where \( \sigma = \frac{q}{A} \) (the charge per square meter), and here’s what that makes the equation look like:

\[
E = \frac{q}{\varepsilon_0 A} = \frac{\sigma}{\varepsilon_0}
\]

When you work with a parallel plate capacitor, life becomes a little easier because the electric field has a constant value and a constant direction (from the + plate to the − plate), so you don’t have to worry about where you are between those plates to find the electric field.

Take a look at an example. Say, for instance, that you put a positive charge of +1.0 coulombs inside the plates of a parallel plate capacitor, as in Figure 3-7.
And also assume that the charge of the plates is $1.77 \times 10^{-11}$ coulombs and that the area of each plate is 1.0 square meters. That would give you the following result for the electric field between the plates:

$$E = \frac{1.77 \times 10^{-11} \text{C}}{(8.854 \times 10^{-12} \text{C}^2 / \text{N} \cdot \text{m}^2)(1.0 \text{ m}^2)} \approx 2.0 \text{ N/C}$$

The electric field is a constant 2.0 newtons per coulomb. To find the force on a 1.0-coulomb charge, you know that

$$F = qE$$

And in this case, that’s $1.0 \text{ C} \times 2.0 \text{ N/C}$, for a total of 2.0 N (or about 0.45 pounds). That calculation is pretty simple, because the electric field between the parallel plates is constant — unlike what the electric field would be between two point charges.

**Shielding: The electric field inside conductors**

This section examines how, in any electric field at all, people can make a little safe haven — a region of zero electric field — with only the aid of a hollow conductor!

Figure 3-8 shows an internal cross-section of a spherical conducting ball. Say that some charges were planted inside the solid metal ball. Now, electric charges always generate an electric field, and conducting materials let charges flow freely in response to electric field, so what happens?

The electric field generated by the charges implanted in the conducting material push other, similar charges away. As a result, charges move around until like
charges are as far away from each other as possible. You can see the result in Figure 3-9 — charges immediately appear on the surface of the conductor. In fact, all the charge appears immediately on the surface — no net charge is left inside the conductor. (Note: Although this example shows positive charge moving, electrons really do the moving, so a reduction of electrons that appear on the surface is what creates a positive charge.)

This kind of behavior — the free motion of charges in conductors — is very useful. For example, if you're in the middle of a region of electric field and you want to have no electric field present, you can shield yourself from the electric field.

To shield yourself from the electric field, you place a conducting container in the field, as in Figure 3-10. The electric field outside the container induces a charge on the container. But the nature of conductors is to let charges flow if there’s any net charge — and any way for it to move around — so the electric field from the induced charges and the preexisting electric field cancel each other out. The result is that there’s no electric field inside the conducting container. You’ve shielded yourself from the external electric field. Nice.
Voltage: Realizing Potential

This section discusses an electrical concept that’s surely familiar: the idea of voltage. Yes, the kind of voltage you get when you plug something into a wall socket. The kind you get when you put a battery into a flashlight or when you rely on a car battery to start your car.

Here, you see how voltage relates to electrical energy. As charges move in electric fields, they can swap some of their energy of motion (kinetic energy) for electrical energy, and vice versa. For example, if you have two opposite charges close together, you have to do some work to separate them due to their mutual attraction. After you’ve separated the two charges, the work you did doesn’t just disappear — it’s stored in the electrical potential energy between them. This section explains how this idea relates to voltage and how voltage relates to the electric field in the special cases of a uniform field and the field around a point charge.

Getting the lowdown on electric potential

If you have a mass in a gravitational field, it has potential energy. As you throw a ball upward, for example, the kinetic energy of its motion is converted to gravitational potential energy as it reaches its peak, and then the potential energy changes back to kinetic energy as the ball falls back to you. The gravitational forces on the ball do work and exchange potential and kinetic energy. Because a force likewise acts on charges in an electric field, you can speak about potential energy here, too. That potential energy is electric potential energy.

What makes all forms of energy essentially the same is that they can all be converted to mechanical work. As you may remember from Physics I, work done (W) is the result of a force (F) moving a body through a distance (s), and they’re related thus: \( W = Fs \). The energy in the interaction of a charge with an electric field is converted to work when the electrical forces move the charge. Moving twice as much charge takes twice as much work for the electrical forces. The work that’s done for every unit of charge is the voltage.

In physics, voltage is called electric potential (not electric potential energy, which isn’t per unit of charge); sometimes, it’s just shortened to potential. Instead of using the term voltage, it’s more correct to say that electric potential is measured in volts, whose symbol is V.
Chapter 3: Getting All Charged Up with Electricity

In the case of a gravitational field, the gravitational force moves a mass in the direction of lower potential — things fall to the ground because they have lower gravitational potential there. In the same way, electrical forces move charges in the direction of lower electrical potential. The faster that the potential energy drops in that direction, the greater the force.

Now remember that electric field is force per unit charge, and electric potential is potential energy per unit charge. Therefore, the electric field is directed down the gradient (slope) of the electrical potential and has a strength proportional to the steepness of the slope.

The electric potential \( V \) at a particular point in space is the electric potential energy of a test charge located at the point of interest divided by the magnitude of that test charge, like this:

\[
V = \frac{PE}{q}
\]

So you can think of electric potential as electric potential energy per coulomb.

So by how many volts does one plate of a charged parallel plate capacitor differ from the other plate? It differs by the energy needed to move 1 coulomb of charge from one plate to the other. (Note that volts are the same as joules/coulomb.)

**Finding the work to move charges**

Say that you’re sitting around, dismantling the smoke alarm in your apartment (which won’t make your landlord very happy) and you find a 9.0-volt battery.

---

Looking at lightning volts

In a thunderstorm, the clouds are at a different electric potential from the ground. The electric potential becomes too great for the air and the air breaks down, conducting electric charge, so every now and then, lightning strikes between the Earth and the clouds.

How many volts are between clouds and the Earth in a thunderstorm? Plenty. It takes 11,000 volts to make a spark across 1 centimeter of air — and there are 100,000 centimeters in a kilometer (about 0.62 miles), the typical height of a cloud during a thunderstorm. You do the math.

Okay, I’ll do the math: that’s 11,000 volts/centimeter \( \times 100,000 \) centimeters = \( 1.1 \times 10^9 \) volts — which is indeed plenty of volts when compared to, say, a wall socket that has 110 volts.
Whipping out the voltmeter you always carry, you measure the voltage between the terminals as exactly 9.0 volts. Hmm, you think. How much energy does it take to move one electron between the two terminals of the battery — a difference in electrical potential between the terminals of 9.0 volts?

Well, you realize that 9.0 volts is the change in potential energy per coulomb between the terminals. And change in potential energy is equal to work. So how much work does it take to move one electron between the terminals? You start by noting that the electric potential is

\[ \Delta V = \frac{W}{q} \]

which means that

\[ W = q\Delta V \]

Here, \( W \) is the work needed to move charge \( q \) across potential difference \( \Delta V \).

Now plug in some numbers. The charge of an electron is a miniscule \(-1.6 \times 10^{-19}\) coulombs, and the potential difference between the negative and positive battery terminals is 9.0 volts, so

\[ W = q\Delta V = (-1.6 \times 10^{-19})(9.0) = -1.4 \times 10^{-18} \text{ J} \]

Therefore, \(-1.4 \times 10^{-18}\) joules of work is done as the electron moves between the two terminals of a 9-volt battery.

You may remember the meaning of negative work from Physics I. The work is negative because the electron’s potential energy falls — that is, the electric force does the work of moving the electron. Moving the electron in the other direction would require an equal quantity of positive work from you, because you’d have to move it up the potential difference, against the electric field.

**Finding the electric potential from charges**

Say you have a point charge \( Q \). What’s the electric potential due to \( Q \) at some distance \( r \) from the charge? You know that the size of the force on a test charge \( q \) due to the point charge is equal to the following (see the earlier section “Looking at electric fields from charged objects” for details):

\[ F = \frac{kQq}{r^2} \]

where \( k \) is a constant equal to \( 8.99 \times 10^9 \text{ N-m}^2/\text{C}^2 \), \( Q \) is the point charge measured in coulombs, \( q \) is the charge of the test charge, and \( r \) is the distance between the point charge and the test charge.
You also know that the electric field at any point around a point charge $Q$ is equal to the following:

$$E = \frac{kQ}{r^2}$$

Thus, close to the point (when $r$ is small), $E$ is large and the field is therefore strong. As you move away from the point charge, $r$ increases and the electric field quickly becomes weak.

Suppose you place a small test charge, $q$, in this field and try moving it around. The test charge feels a strong force close to the point charge, which quickly falls away as you move it to a greater distance. If the test charge is of the opposite sign to the point charge, you have to do work to pull it away from the point charge. This means that the test charge has a lower potential energy closer to the point charge (this is reversed if the charges are the same sign).

So what’s the electric potential from the point charge? At an infinite distance from the point charge, you can’t see it or be affected by it, so set the potential from the point charge to be zero there. As you bring a test charge closer, to a point $r$ away from the point charge, you have to add up all the work you do and then divide by the size of the test charge. And the result after you do turns out to be gratifyingly simple. Here’s what that it looks like:

$$V = \frac{kQ}{r}$$

So the electrical potential is large close to the point charge (when $r$ is small) and falls away at greater distances. This idea applies to all point charges, so what does it mean for the electrons orbiting the protons in an atom? How hard do you have to work to pull an electron away from an atom?

First find the electrical potential. The size of the charge of the electron and the proton is $1.6 \times 10^{-19}$ coulombs. The electron and proton are typically $5.29 \times 10^{-11}$ meters apart, so the electric potential is

$$V = \frac{kQ}{r} = \left(\frac{8.99 \times 10^9 \text{ N} \cdot \text{m/C}^2}{5.29 \times 10^{-11} \text{ m}}\right) \left(1.60 \times 10^{-19} \text{ C}\right) = 27.2 \text{ volts}$$

That close to the proton, the electric potential is a full 27.2 volts. That’s quite something for such a tiny charge!

The amount of energy needed to move an electron through 1 volt is called an **electron-volt (eV)**. So you may expect that you’d need 27.2 eV of energy to pull an electron out of the atom. But the electron is moving quickly, so it already has some energy to contribute. In fact, the electron has so much kinetic energy that you need only half that, 13.6 eV of energy, to win the electron from the atom!
**Illustrating equipotential surfaces for point charges and plates**

To illustrate electric potential, you can draw *equipotential surfaces*; that is, surfaces that have the same potential at every point. Drawing an equipotential surface gives you an idea of what the electric potential from a charge or charge distribution looks like. For example, on one equipotential surface, the potential could always be 5.0 volts or 10.0 volts.

Because the potential from a point charge depends on the distance you are from the point charge, the equipotential surfaces for a point charge are a set of concentric spheres. You can see what they’d look like in Figure 3-11.

![Figure 3-11: Dashed lines showing equipotential surfaces from a point charge.](image)

Now consider equipotential surfaces between the plates of a parallel plate capacitor (see the earlier section “Uniform electric fields: Taking it easy with parallel plate capacitors” for more on these devices). If you start at the negatively charged plate and move a distance \(s\) toward the positively charged plate, you know that

\[
V = \frac{qs}{\varepsilon_0 A}
\]
In other words, the equipotential surfaces here depend only on how far you are between the two plates. You can see this in Figure 3-12, which shows two equipotential surfaces between the plates of the parallel plate capacitor. **Tip:** This is analogous to the gravitational potential close to the ground, which increases in proportion to height. They’re both cases of uniform fields in which the field is the same at every point.

**Storing Charge: Capacitors and Dielectrics**

A capacitor, generally speaking, is something that stores charge. I discuss parallel plate capacitors earlier in this chapter, but a capacitor need not be shaped like two parallel plates — any two conductors separated by an insulator form a capacitor, regardless of shape. A dielectric increases how much charge a capacitor can hold. This section covers how capacitors and dielectrics work together.

**Figuring out how much capacitors hold**

How much charge is actually stored in a capacitor? That depends on its capacitance, $C$. The amount of charge stored in a capacitor is equal to its capacitance multiplied by the voltage across the capacitor:

$$q = CV$$
The MKS unit for capacitance $C$ is coulombs per volt (C/V), also called the farad, $F$. For a parallel plate capacitor, the following is true (see the preceding section for details):

$$V = \frac{qs}{\varepsilon_0 A}$$

Because $q = CV$, you know that $C = q/V$, and you can solve the preceding equation to get the following:

$$C = \frac{q}{V} = \frac{\varepsilon_0 A}{s}$$

So that’s what the capacitance is for a parallel plate capacitor whose plates each have area $A$ and are distance $s$ apart.

**Getting extra storage with dielectrics**

A dielectric is a semi-insulator that lets a capacitor store even more charge, and this substance works by reducing the electric field between the plates. Take a look at Figure 3-13. When you apply an electric field between the two plates, that induces some opposite charge on the dielectric, which opposes the applied electric field. The net result is that the electric field inside the dielectric (which fills the area between the plates) is reduced, allowing the capacitor to store more charge.

![Figure 3-13: Using a dielectric between the plates of a parallel plate capacitor.](image)

If you fill the space between the plates in a parallel plate capacitor with a dielectric, the capacitance of the parallel plate capacitor becomes the following:

$$C = \frac{\kappa \varepsilon_0 A}{s}$$

The dielectric increases the capacitance of the capacitor by the dielectric constant, $\kappa$, which differs for every dielectric. For example, the dielectric
constant of mica, a commonly used mineral, is 5.4, so capacitors that use mica increase their capacitance by a factor of 5.4. The dielectric constant of a vacuum is 1.0.

**Calculating the energy of capacitors with dielectrics**

Because a capacitor stores charge, it can act as a source of electric current, like a battery. So in terms of capacitance and voltage, how much energy is stored in a capacitor?

Well, when you charge a capacitor, you assemble the final charge $q$ in an average potential $V_{avg}$ (you use the average potential because the potential increases as you add more charge). So here’s the energy stored:

$$\text{Energy} = qV_{avg}$$

This raises the question, of course, of just what $V_{avg}$ is. Because the voltage is proportional to the amount of charge on the capacitor (because $q = CV$), $V_{avg}$ is one-half of the final charge. Or looked at another way, as you charge up a capacitor from 0 to its final voltage, the voltage increases linearly, so the average voltage is half the final voltage:

$$V_{avg} = \frac{1}{2}V$$

Plugging this value of $V_{avg}$ into the energy equation and making the substitution that $q = CV$ gives you the following equation:

$$\text{Energy} = \frac{1}{2}CV^2$$

For a parallel plate capacitor with a dielectric, the capacitance is

$$C = \frac{\kappa \varepsilon_0 A}{s}$$

So the energy in a parallel plate capacitor with a dielectric in it is

$$\text{Energy} = \frac{1}{2} \frac{\kappa \varepsilon_0 A}{s} V^2$$

And there you have it — that’s the energy stored in a capacitor with a dielectric (the energy ends up in joules from this equation). You can see why dielectrics are considered a good idea when it comes to capacitors — they multiply the capacitance and the energy stored in a capacitor many-fold.
Chapter 4

The Attraction of Magnetism

In This Chapter
▶ Understanding magnetism
▶ Looking at magnetic forces on moving charges
▶ Finding magnetic forces on wires
▶ Generating magnetic fields with current in wires

Legend has it that more than 2,000 years ago, Magnes, a Greek shepherd, was walking with his flock when he found the nails that were holding his shoes together became inexplicably stuck to a rock — and that’s how magnetism was found. Thousands of years of mystery came together into a scientific understanding only in the last few hundred years.

In this chapter, you explore the physicists’ understanding of magnetism. You see why permanent magnets (like the one stuck to Magnes’s shoe) attract some apparently nonmagnetic materials (like the iron nail in Magnes’s shoe). You see how magnetism is not really a strange new thing but only a different aspect of electricity. And you see how to work out exactly how big a magnetic force is and in which direction it goes. You find the magnetic influence of electrical currents, and you see that electrical currents are the source of magnetism.

With all this knowledge come all kinds of useful devices, like electric motors, speakers, doorbells, and even sophisticated medical imaging machines. That’s why I also discuss some practical uses of magnets and magnetism. For instance, I explain how a compass works by moving under the magnetic influence of the molten iron swirling at the center of the Earth. And at last you’ll know what makes those things stick to the fridge door.
All About Magnetism: Linking Magnetism and Electricity

Here's the most important thing you need to know about magnetism: It’s closely related to electricity (which I cover in Chapter 3). Magnetism and electricity are just different aspects of the same thing — two sides of the same coin. In fact, where you have electricity flowing, you have magnetism, because magnetism comes from electrical current.

In this section, I discuss how current flows even in permanent magnets, because the electrons are in motion in the magnet’s atoms. I also explain the repelling and attractive forces at work in magnets, just like the forces between electrical charges. Finally, I give you a formal definition of magnetic field that ties magnetism and electric charge together.

Electron loops: Understanding permanent magnets and magnetic materials

Even in bar magnets or refrigerator magnets, electricity is indeed flowing. Every atom in that refrigerator magnet has a bunch of electrons circling a nucleus, and those electrons form an electric current. This is on the level — magnetism in permanent magnets comes from those tiny orbits that electrons are continually zooming around in. Loops of current form a magnetic field, so the atoms behave like tiny magnets.

Most of the time, those tiny magnetic fields are pointing in all different directions, as you see in Figure 4-1a, so they add up to zero. However, in a permanent magnet, those tiny, atomic-level magnets are much more aligned, as in Figure 4-1b. When the micro-magnets in matter align, you end up with magnetism you can measure — and the magnet stays magnetized. That’s why you call such magnets (which take no external electrical power) permanent magnets. That alignment is the only difference between a permanent magnet and a permanent non-magnet.

Some materials are in between these opposites. For instance, all the atoms, acting like little magnets, may all start out pointing in random directions — imagine that the electrons are orbiting in planes that are randomly orientated. But when you put this material close to a magnet, the atoms are forced to rotate and align all together with the magnet. Then the magnetic influence of each atom adds together, so the material becomes magnetic. But when you take the material away from the magnet, the atoms relax back to their random orientations, and the
material stops being magnetic. Such a material is called paramagnetic. Two paramagnets won’t stick to each other, but they will stick to a permanent magnet.

Figure 4-1: Micro-magnets in matter, unaligned and aligned.

In another type of material, such as iron, the atoms are organized into little aligned regions, called domains. Each domain is magnetic, but the material has many domains that are randomly orientated. Now if you place this material near a magnet, the domains are forced to align and the material becomes a magnet. But this time, when you take the material away, the domains stay aligned, and the piece of iron remains magnetic! This type of material is said to be ferromagnetic.

Electromagnets are nonpermanent magnets that work only when you have electricity flowing. I discuss these magnets later in “Going to the Source: Getting Magnetic Field from Electric Current.”

North to south: Going polar

Electricity has two sides to it: positive and negative. Electric field goes from the plus to the minus side (see Chapter 3 for details). Similarly, magnetism involves magnetic poles. And just as electric field goes from + charges to – charges, magnetic field goes from one pole to the other — from north to south.

The names of the magnetic poles come from the popular use of magnets in compasses — the North and South Poles are used in navigation. The north pole of a permanent magnet automatically points toward magnetic north of the Earth.

Magnetic field is often drawn as a set of lines — that is, magnetic field lines, much as electric field is drawn as electric field lines. Figure 4-2 shows the magnetic field lines going from north pole to south of a permanent magnet.
The Earth is a huge magnet, as anyone with a compass knows — just watch your compass needle point unerringly toward magnetic north. When you change locations, the compass needle finds magnetic north for you again. Unfortunately, magnetic north doesn’t match the true North Pole of the Earth — that is, geographic north, where the Earth’s axis of rotation pierces the surface of the Earth. The following figure shows how the Earth’s geographic north pole is offset from the magnetic north pole.

![Diagram of the Earth's magnetic field]

Notice how the magnetic pole in the figure is labeled S. That’s no typo — opposite magnetic poles attract, so to attract the north needle in your compass, the pole that lies just under the surface at the Earth’s “north magnetic pole” is really a south pole, not a north pole. But unlike labeling a bar magnet, the Earth’s actual South Pole is always called its North Pole — for which we have compasses to thank.

The distance from the geographic North Pole to the Earth’s magnetic north pole is fairly large — the magnetic north pole currently lies near Ellesmere Island in Northern Canada. Actually, the position of the Earth’s magnetic north pole wanders over the years. The magnetic pole’s yearly wanderings are appreciable: It’s currently moving at a rate of more than 40 kilometers per year! The Earth’s magnetic field is maintained by the swirling movements of the molten iron deep inside the Earth, in the planet’s liquid outer core, allowing the magnetic pole to wander.

So how far away is the magnetic pole from the geographic pole? That’s measured by the angle of declination. That angle varies, depending on your position on the Earth, but it can be sizeable. For example, in New York City, the angle of declination is about 12°. You can find out much more on the Earth’s magnetic field, with current data and many interesting links, on the Web sites of the Geological Survey of Canada (gsc.nrcan.gc.ca/geomag) and the United States Geological Survey (geomag.usgs.gov).
Note that the magnetic field from a magnet like the one in Figure 4-2 isn’t very constant or uniform — just like the electric field from two point charges wouldn’t be uniform.

If you want a uniform magnetic field, you usually select a location between the two poles of a strong magnet, as Figure 4-3 shows. You can also create a uniform magnetic field using coils of current, as I explain later in “Adding loops together: Making uniform fields with solenoids.”

**Defining magnetic field**

Magnetism and electricity are so interconnected that magnetic field is defined in terms of the strength of the force it exerts on a positive test charge. The symbol $M$ was already taken (it stands for the magnetization of a material), so magnetic field ended up with the symbol $B$. Here’s the formal definition of magnetic field, from a physics point of view:

$$B = \frac{F}{qv \sin \theta}$$
Here, $B$ is the magnitude of the magnetic field and $F$ is the magnitude of the force on the charge $q$, which is moving with speed $v$ at an angle of $\theta$ to the direction of the magnetic field.

In the MKS system, the unit of magnetic field is the tesla, whose symbol is T. In the CGS system, you use the gauss, whose symbol is G. You can relate the two like this:

$$1.0 \text{ G} = 1.0 \times 10^{-4} \text{ T}$$

**Moving Along: Magnetic Forces on Charges**

Electric currents and magnetic fields are linked very closely. Not only do electric currents give rise to magnetic fields, but magnetic fields also exert forces on the electric charges moving in currents.

Note that a charge has to be moving in order for a magnetic field to exert a force on it: No motion, no force on the charge.

In this section, I show you how to figure out the strength and direction of the magnetic force on a moving charge. I also explain how the direction of that force can ensure that magnetic fields don’t do any work. To finish, you see why the direction of the force causes charged particles to travel in circles in a magnetic field.

**Finding the magnitude of magnetic force**

To get numerical with magnetism, you have to start thinking in terms of vectors. Suppose you have an electric charge moving with a velocity $v$. That charge is subject to a magnetic field, $B$. And of course, you need the $F$ vector for the resulting force.

How can you determine the actual force, in newtons, on a charged particle moving though a magnetic field? That force is proportional to both the magnitude of the charge and the magnitude of the magnetic field. It’s also proportional to the component of the charge’s velocity that’s perpendicular to the magnetic field. In other words, if the charge is moving along the direction of the
magnetic field, parallel to it, there will be no force on that charge. If the charge is moving at right angles to the magnetic field, the force is at its highest.

Putting all this together gives you the equation for the magnitude of the force on a moving charge, where θ is the angle (between 0° and 180°) between the \( \mathbf{v} \) and \( \mathbf{B} \) vectors:

\[
F = qvB \sin \theta
\]

For example, suppose you’re carrying around a 1.0-coulomb charge, and you experience a force from the Earth’s magnetic field. The Earth’s magnetic field on the surface is about 0.6 gausses, or \( 6.0 \times 10^{-5} \) teslas. The faster you move with your charge, the more force you’ll feel, so suppose you take it for a spin in a race car. Head off down the track straight at about 224 miles per hour, or 100 meters per second. What force do you feel on your charge at this speed in the direction perpendicular to the field? You know that the magnitude of the force is given by

\[
F = qvB \sin \theta
\]

So plug in the numbers. Here’s what that looks like when you do:

\[
F = qvB \sin \theta
\]

\[
= (1.0 \text{ C})(100 \text{ m/s})(6.0 \times 10^{-5} \text{ T}) \sin 90°
\]

\[
= 6.0 \times 10^{-3} \text{ N}
\]

The force on the charge is \( 6.0 \times 10^{-3} \) newtons, which is less than the weight of a paperclip.

**Finding direction with the right-hand rule**

Say you have a charge, \( q \), traveling along with velocity \( \mathbf{v} \), minding its own business. If that charge travels in a magnetic field, \( \mathbf{B} \), there’s going to be a force on the charge. You can see the direction of the magnetic force on the moving charge in Figure 4-4.

A right-hand rule operates when you’re finding the force on a charge, and there are two versions of it — use whichever one you find easier:
If you place the fingers of your open right hand along the magnetic field, the vector $\mathbf{B}$ in the figure, and your right thumb in the direction of the charge’s velocity, $\mathbf{v}$, then the force on a positive charge extends out of your palm (see Figure 4-4a). For a negative charge, reverse the direction of the force.

Place the fingers of your right hand in the direction of velocity of the charge, $\mathbf{v}$, and then wrap those fingers by closing your hand through the smallest possible angle (less than 180°) until your fingers are along the direction of the magnetic field, $\mathbf{B}$. Your right thumb points in the direction of the force (see Figure 4-4b). For a negative charge, reverse the direction of the force.

Give the two methods a try to make sure you get the direction of the force correct.

Figure 4-4:
The force on a charge on a magnetic field and the associated right-hand rules.

A lazy direction: Seeing how magnetic fields avoid work

Magnetic fields are lazy: They do no work on charged particles that travel through them — at least, not by the physics definition of work. So a charged particle in a magnetic field doesn’t gain or lose kinetic or potential energy.

When you have an electric field, the situation is very different. There, the electric field pushes a charge along or against the direction of travel. And that’s the physics definition of work:

$$W = Fs \cos \theta$$

where $F$ is the force applied, $s$ is the distance over which it’s applied, and $\theta$ is the angle between the force and the direction of travel. In fact, that’s where
the whole idea of electric potential, *voltage*, comes from — the amount of work done on a charge divided by the size of the charge:

\[ V = \frac{W}{q} \]

For the work done by a magnetic field, the trouble is the definition of work:

\[ W = Fs \cos \theta \]

The issue here is that in a magnetic field, the force and the direction of travel are always perpendicular to each other — that is, \( \theta = 90^\circ \) (see the preceding section). And \( \cos 90^\circ = 0 \), so the work done by a magnetic field on a moving charge, \( W = Fs \cos \theta \), is automatically zero.

That’s it — the work done by a magnetic field on a moving charge is zero. That’s why there’s no such thing as magnetic potential (would that be *magnetic volts*?) to correspond to electric potential.

That’s all due to the physics definition of work — work changes the kinetic or potential energy of a system (or the energy is lost to heat), and nothing of the kind happens with magnetic fields. However, the *direction* of the charged particle does change. That’s what changes — not the particle’s speed but its direction.

**Going orbital: Following charged particles in magnetic fields**

The direction and magnitude of the force in a magnetic field affects the path that an electric charge takes. The direction of the force causes the charge to move in circles, and the force’s magnitude affects how big of a radius that circle has. In this section, I discuss the orbital motion of charges in magnetic fields.

**Getting the curve**

If you have an electric field (see Chapter 3), you know which way electric charges will move in such a field — along the electric field lines. For example, if you have a parallel plate capacitor, electrons will travel between the plates along the electric field lines, toward the positive plate. Protons will do the same, except they’ll move toward the negative plate.

The situation changes when you have a magnetic field, not an electric field. Now the force is perpendicular to the direction of travel, which can take a little getting used to. To better show the path of the charge, physicists often draw the magnetic field as though you were looking at it straight on. How
can you tell which way the magnetic field is going? Here’s the physics way of showing direction:

- **Away from you**: If you see a bunch of X’s, the magnetic field goes into the page. Those X’s are intended to be the end of vector arrows, seen tail-on (imagine looking down the end of a real arrow, tail toward you).
- **Toward you**: Dots with circles around them are supposed to represent arrows coming at you, so in that case, the field is coming toward you.

Take a look at Figure 4-5, which shows the path a positive charge moving in a magnetic field will take. The positive charge travels along a straight line, undeflected, until it enters the magnetic field that goes into the page (represented by the X’s). Then a force appears on the charge at right angles, bending its path, as you can see in the figure.

**Note**: This is a good place to test your understanding of the right-hand rule of magnetic force (see the earlier section “Finding direction with the right-hand rule”). Apply it to the velocity and magnetic field you see in Figure 4-5 — do you agree with the direction of the resulting force?

**Figure 4-5**: A positive charge being pushed in a magnetic field.

**Going in circles**

Here’s an interesting point: Which way do you get pushed if you’re a charged particle moving in a magnetic field? The magnetic field is always perpendicular to the direction of travel (as Figure 4-4 shows earlier in this chapter). And no matter which way the charged particle turns, the force on it is perpendicular to its motion.

That’s the hallmark of circular motion: The force is always perpendicular to the direction of travel. Therefore, charged particles moving in magnetic fields travel in circles.

See Figure 4-6 to get the full picture. There, a positive charge is moving to the left in a magnetic field. The dots with circles around them tell you that
magnetic field $B$ is coming straight at you, out of the page. Using the right-hand rule, you can tell which way the resulting force goes — upward when the positive charge is at the location in Figure 4-6.

What happens? The charge responds to that upward force by moving upward. And because the force due to the magnetic field is always perpendicular to the direction of travel, the force changes direction, too.

![Figure 4-6: In a magnetic field, a positive charge goes in circles.](image)

### Finding the radius of orbit

Suppose you want to know the radius of the orbit of the charged particle moving in a magnetic field. Because the force is always perpendicular to the direction of travel, you end up with circular motion. And from Physics I, you have the following equation for the force needed to keep an object in circular motion:

$$F = \frac{mv^2}{r}$$

Here's the magnitude of the force on a charged particle moving in a magnetic field:

$$F = qvB \sin \theta$$

Because $v$ is perpendicular to $B$ in this case, $\theta$ equals 90°; therefore, $\sin \theta$ equals 1, which means you get this:

$$F = qvB$$
So set the two force equations — for circular motion and for the charged particle in the magnetic field — equal to each other:

\[ qvB = \frac{mv^2}{r} \]

Rearranging this equation gives you this new version, solved for the radius:

\[ r = \frac{mv}{qB} \]

That’s great — that gives you the radius of the charged particle’s path in a magnetic field, given its mass, charge, and velocity. This is one of the magnetism equations you should remember.

Note the following relationships between the radius and the magnetic field, mass, and velocity:

- **Magnetic field** \( B \): The stronger the magnetic field, the stronger the force — and therefore the smaller the radius of the charged particle.
- **Velocity** \( v \): The more speed a charged particle has, the harder it is for the magnetic field to corral the particle, and so it travels in a circle with a bigger radius.
- **Mass** \( m \): The more mass the charged particle has, the harder it’ll be to bend its path, so the more mass, the bigger the radius of the circle it travels in.

Notice how the equation reflects all these ideas: The magnetic field \( B \) is in the denominator of the fraction, so increasing \( B \) would give you a smaller answer for \( r \); \( m \) and \( v \) are on top, so increasing either one of those would give you a larger \( r \).

How about seeing this in action? Try some numbers. Say, for example, that you have a bunch of electrons going at \( 1.0 \times 10^6 \) meters per second. You don’t want to disturb the neighbors, so you decide to build a magnetic containment vessel to contain the electrons, sending them around in a circular orbit. Checking your bank account, you see you have only enough money to create a magnetic confinement vessel of \( r = 1.0 \) centimeters (even that may make your landlady suspicious, but she’s learned that physicists sometimes need unusual equipment).

So what magnetic field do you need to limit your electrons to an orbit where \( r = 1.0 \) centimeters? You know that

\[ r = \frac{mv}{qB} \]
An electron has a mass of $9.11 \times 10^{-31}$ kilograms and a charge of $1.6 \times 10^{-19}$ coulombs. Plugging in the numbers for electrons moving at $1.0 \times 10^6$ meters per second, you get

$$0.010 \text{ m} = \frac{\left( 9.11 \times 10^{-31} \text{ kg} \right) \left( 1.0 \times 10^6 \text{ m/s} \right)}{\left( 1.6 \times 10^{-19} \text{ C} \right) B}$$

Rearranging this equation and solving for $B$ gives you the answer:

$$B = \frac{\left( 9.11 \times 10^{-31} \text{ kg} \right) \left( 1.0 \times 10^6 \text{ m/s} \right)}{\left( 1.6 \times 10^{-19} \text{ C} \right) \left( 0.010 \text{ m} \right)} \approx 5.7 \times 10^{-4} \text{ T}$$

That’s a very modest magnetic field — it’s not much more than the Earth’s magnetic field. It really doesn’t take much to push an electron around. (That’s a relief, because if you’d needed a magnetic field of several teslas, the landlady’s silverware, which is silver-plated steel, might’ve ended up stuck to her ceiling.)

The equation

$$r = \frac{mv}{qB}$$

doesn’t apply if the charged particle is traveling near the speed of light, $v \approx 3.0 \times 10^8$ meters per second, because relativistic effects take over, which affect the mass and orbital radius of the charged particle. I discuss special relativity in Chapter 12.

---

**Selecting your atoms with a mass spectroscope**

*Mass spectrometers*, which are machines that determine which chemical elements go into a sample you’re analyzing, rely on orbits in magnetic fields. A mass spectrometer heats the sample you want to analyze, ionizing some of the atoms. The singly ionized atoms get a net charge, $e$ (the same magnitude as an electron’s charge), and those atoms are accelerated through an electric potential, $V$, which gives them kinetic energy. The ionized atoms then enter a magnetic field, $B$, and they turn with radius $r$. You can pick them up with a detector, and positioning the detector tells you the radius for the ionized atoms — and from knowing the radius, electric potential, and magnetic field, you can determine the mass of the atoms and therefore identify them.

So given $r$, $e$, and $B$, you solve for $m$, the mass. The accelerating electric potential, $V$, gives

(continued)
Down to the Wire: Magnetic Forces on Electrical Currents

You may be one of those rare physicists who doesn’t have a bunch of electrons hurtling around at home at $1.0 \times 10^6$ meters per second. You may think that the preceding discussion, about charged particles in magnetic fields, doesn’t really apply to you. However, you surely have electric cables around the house — and what are electric cables but wires through which charges move? In this section, you look at the forces that magnetic fields exert on charges moving in electric wires.

From speed to current: Getting current in the magnetic-force formula

To find the magnetic force on an electric wire in a magnetic field, you can start with the formula for individual charges. Take a look at the equation for the force on a moving electric charge in a magnetic field:

$$F = qvB \sin \theta$$
You want to translate this equation so that instead of using the speed of charged particles, \( v \), it uses electrical current, \( I \). How do you get electrical current out of this? Current is charge, \( q \), divided by the amount of time, \( t \), that a charge takes to pass a particular point:

\[
I = \frac{q}{t}
\]

Divide the equation for force by time and multiply it by time, which doesn’t actually change the equation. Here’s what you get:

\[
F = \frac{q}{t}(vt)B \sin \theta
\]

Note that you now have \( q/t \), or current, here. So here’s the force equation in terms of current:

\[
F = I(vt)B \sin \theta
\]

Okay, so what’s \( I(vt)B \sin \theta \)? Something of a mixed bag here — current and speed together. But if you think about it, the term \( vt \) is just the speed of the charged particles making up the current multiplied by the measured time — and speed times time equals a distance. So replace \( vt \) with \( L \), the distance the charged particles go in time \( t \).

So here’s the force on a wire of length \( L \) carrying current \( I \) in a magnetic field of strength \( B \), where the \( L \) is at angle \( \theta \) with respect to \( B \):

\[
F = ILB \sin \theta
\]

Cool. How about an example? Look at Figure 4-7, which shows an electric current \( I \) in a magnetic field \( B \). In physics, current goes in the direction a stream of positive charges would take (that convention was defined before scientists knew that it was negative charges — electrons — that really flowed to make current go). That means you can apply the right-hand rule to the situation you see in Figure 4-7 — just treat the direction of \( I \) as the direction a positive charge is traveling in. (For the right-hand rule, see the earlier section “Finding direction with the right-hand rule.”)

All right, what if the current, \( I \), equals 1.0 amp, and the magnetic field, \( B \), is equal to 5 teslas? The magnetic force on a wire carrying this current increases in proportion to its length. For every meter of wire that you have, what would be the resulting force? You start with the formula for force:

\[
F = ILB \sin \theta
\]

and then divide by the length, \( L \), to find the force per meter:

\[
\text{Force/meter} = IB \sin \theta
\]
In Figure 4-7, $\theta$ is $90^\circ$, and because $\sin 90^\circ$ equals 1, you get this case:

\[
\text{Force/meter} = IB = (1.0 \text{ A})(5.0 \text{ T}) = 5.0 \text{ N/m}
\]

So you get a result of 5.0 newtons per meter, which works out to be about a third of a pound per foot — something to keep in mind if you have electric cables running through a 5.0-tesla magnetic field (which, admittedly, is pretty rare).

**Torque: Giving current a twist in electric motors**

Scientists saw that magnetic fields exerted forces on electric wires and came up with electric motors. From there came electric washers and dryers, windshield wipers on cars, elevators, automatic doors at grocery stores, refrigerators, and much more (not in that order, of course). As you can see, life without electric motors would be inconvenient. This section helps you see what makes electric motors work, electricity and magnetism-wise — at least in basic terms.

**Big-time currents**

In the big physics labs, where cables can hold huge current (direct current, not alternating current), you can see something curious: When a cable is made up of individual strands of wire, those wires create magnetic field, and that magnetic field acts on the other wires in the cable. The net result is that the cable contracts before your very eyes, getting thinner as the magnetic fields act on the currents.
Seeing how motors work

Figure 4-8 shows an electric motor, stripped down to its basic components. Two permanent magnets of opposite polarity are on either side of the motor. This generates a uniform magnetic field in the space between the poles, from the north pole to the south pole. In this magnetic field, you place a loop of wire, which is free to rotate about the axis in the figure. A battery is connected to the loop, so a current is flowing through the wire in the direction shown by the arrows labeled $I$.

The wire loop is connected to the battery by a strange connection called a commutator. This clever little device is a vital part of the motor because it ensures that the current always flows in the direction shown in the diagram, even when the loop has rotated half a turn. It always connects the side of the loop that's closest to the north pole of the magnet to the positive terminal of the battery and vice versa, while leaving the loop free to turn.

Because the loop is carrying current, the loop experiences a force in the magnetic field. I've shaped the loop as a rectangle to make the calculation of the force it experiences a little easier.

Two sides of the loop are parallel to the axis of rotation, and two are perpendicular to it. The perpendicular wires don't play a part here because the force they experience is directed along the axis of rotation, so they don't produce any torque. Also, they're equal and opposite in size, so you don't get a net force from them.

Most interesting are the two parts of the loop that run parallel to the axis of rotation, which are always at 90° to the magnetic field. The left side of the wire loop is forced down, and the right side is forced up (you can use the right-hand rule to confirm that the directions of the forces in Figure 4-8 are correct). This results in a turning force — that is, a torque — that rotates the loop of wire. If you connect the loop to an axle, then the loop will force the axle to turn — and you can use this turning force for all sorts of things.
**Figuring out the turning force**

So how much turning force does an electric motor give you? Torque, as you may recall from Physics I, is a twisting force, with the symbol \( \tau \). Here's the formula for it:

\[
\tau = Fr \sin \theta
\]

where \( F \) is the applied force, \( r \) is the distance the force acts from the turning point, or *pivot*, and \( \theta \) is the angle between \( F \) and \( r \).

In an electric motor, a loop of current is embedded in a magnetic field, \( B \), and that field creates forces, \( F \), on each wire running parallel to the axis of rotation (as you see in Figure 4-8). The torque on each wire is the force \( (F = ILB) \), multiplied by the distance, \( d \), the force acts from the pivot multiplied by the sine of the angle. Because there are two torques, corresponding to the two sides of the loop, the total torque, \( \tau \), is equal to the following:

\[
\tau = ILB \left( \frac{1}{2} d \sin \theta + \frac{1}{2} d \sin \theta \right) = ILB d \sin \theta
\]

The product \( dL \) is the height multiplied by the width of the loop of wire — that is, the *area* of the loop. So if you write the area as \( A \), the equation for the torque on a loop of wire becomes

\[
\tau = IAB \sin \theta
\]

In fact, electric motors aren’t really made of a single loop of wire — they’re made of coils of wire. So instead of one loop, you actually have \( N \) loops, where \( N \) is the number of coils of wire. That makes the torque into

\[
\tau = NIAB \sin \theta
\]

That’s the total torque on a coil of \( N \) loops of wire, each carrying current \( I \), of cross-sectional area \( A \), in a magnetic field \( B \), at angle \( \theta \) as shown in Figure 4-8.

In physics class, you’re usually asked what the maximum torque would be for such-and-such a coil in such-and-such a magnetic field. If you come across a situation like that and need to find the maximum torque, that occurs when the coil is at right angles to the magnetic field: \( \theta = 90^\circ \), so \( \sin \theta = 1 \), or

\[
\tau = NIAB
\]

Try some numbers here. If you have a coil with 200 turns, a current of 3.0 amps, an area of 1.0 square meters, and a magnetic field of 10.0 teslas, what’s the maximum possible torque? Just plug this into the equation:

\[
\tau = NIAB = (200)(3.0 \, \text{A})(1.0 \, \text{m}^2)(10.0 \, \text{T}) = 6.0 \times 10^3 \, \text{N-m}
\]

So you have a maximum torque of 6,000 newton-meters, which is very large — a car typically generates only about 150 newton-meters.
**Going to the Source: Getting Magnetic Field from Electric Current**

The earlier sections in this chapter concentrate on how magnetic fields exert forces on moving charges, or currents, without worrying too much about where the magnetic field came from in the first place. In this section, you discover the source of that magnetic field. Here, you see the relationship between electricity and magnetism become complete.

Simply put, just as electric charges are the source of electric fields, which exert forces on other electric charges, electric currents are the source of the magnetic fields, which exert forces on other electric currents.

In Chapter 3, I take a couple of simple arrangements of charge (the point charge and the parallel plate capacitor) and examine the resulting electric fields. Now, in this section, I take a few simple arrangements of current (a straight wire, a loop, and a tube of current called a *solenoid*) and examine the resulting magnetic fields. Here, you also see how you can use this idea to make *electromagnets*, magnets that you can switch on and off with a switch.

**Producing a magnetic field with a straight wire**

To understand how electric current produces a magnetic field, first take a look at the magnetic field from a single wire, as Figure 4-9 shows. Why start here? When you know what the magnetic field is from a single wire of current, you’re home free in many problems. You can often break down more-complex distributions of current into many single wires — and then add the magnetic fields from the wires as vectors to get the overall result.

**Figure 4-9:** The magnetic field from a single wire of current.
Assembling the formula for magnetic field from a single wire

When you make physical measurements of the magnetic field from a single wire, you find that the magnetic field, $B$, diminishes as $1$ over the distance, $r$. Therefore, you get this relation (where $\propto$ means proportional to):

$$B \propto \frac{1}{r}$$

What else can the magnetic field depend on? Well, how about the current itself, $I$? Surely if you double the current, you get twice the magnetic field, right? Yep, that’s the way it works, as borne out by measurement, so now you have the following:

$$B \propto \frac{I}{r}$$

That’s all you need.

The constant of proportionality, for historical reasons, is written as $\mu_0/(2\pi)$, which means you finally get this result for the magnetic field from a single wire:

$$B = \frac{\mu_0 I}{2\pi r}$$

Note that the constant $\mu_0 = 4\pi \times 10^{-7}$ T-m/A.

A right-hand rule: Finding field direction from a wire

Magnetic field, $B$, is a vector. If you have a magnetic field from the current in a single wire, which way does the $B$ field go? There’s another right-hand rule for just this occasion. If you put the thumb of your right hand in the direction of the current, the fingers of that hand will wrap around in the direction of the magnetic field. At any one point, the direction your fingers point is the direction of the magnetic field, as Figure 4-10 shows.
Give this a try: Suppose you have two parallel wires. Verify that the force between two wires is toward each other if the current in both is in the same direction and away from each other if the current in each wire is in opposite directions.

**Calculating magnetic field from straight wires**

How about some numbers? Say you have a current of some 10 amps and you want to measure the magnetic field 2.0 centimeters from the center of the wire. What is the strength of the $B$ field you’ll get? Here’s your formula:

$$B = \frac{\mu_0 I}{2\pi r}$$

Plugging in the numbers gives you:

$$B = \frac{\mu_0 I}{2\pi r} = \left(\frac{4\pi \times 10^{-7} \text{T} \cdot \text{m/A}}{}\right) \frac{(10 \text{ A})}{2\pi (0.020 \text{ m})} = 1.0 \times 10^{-4} \text{T} = 1.0 \text{ G}$$

So you get 1.0 gauss, a little more than the Earth’s magnetic field, which is about 0.6 gauss.

That was a quick example — how about one that’s a little tougher? Say you have two wires, parallel to each other, with current $I$ in each going the same way. The wires are a distance $r$ apart. What’s the force on Wire 1 from the magnetic field coming from Wire 2?

You know that the force on Wire 1, which is carrying current $I$ in magnetic field $B$, is the following (to see where this formula comes from, check out the earlier section “From speed to current: Getting current in the magnetic-force formula”):

$$F = ILB$$

All right, but what’s $B$? That’s the magnetic field from Wire 2, measured at the position of Wire 1. Because the wires are $r$ distance apart and Wire 2 is carrying a current $I$, its magnetic field is this at the location of Wire 1:

$$B = \frac{\mu_0 I}{2\pi r}$$

Substituting this expression for $B$ into the $F = ILB$ equation, you get this result:

$$F = \frac{\mu_0 I^2 L}{2\pi r}$$
How about getting the force per unit length? That’s \( F/L \), which is

\[
\frac{F}{L} = \frac{\mu_0 I^2}{2\pi r}
\]

Now try some numbers. Say you have two parallel wires with current \( I \) going in the same direction — current, \( I \), is 10 amps, and the distance between the wires, \( r \), is 2.0 centimeters. Putting in those numbers, you get

\[
\frac{F}{L} = \frac{\mu_0 I^2}{2\pi r} = \frac{(4\pi \times 10^{-7}\text{T} \cdot \text{m/A})(10\text{ A}^2)}{2\pi (0.020\text{ m})} = 1.0 \times 10^{-3}\text{ N/m}
\]

So the force on Wire 1 from Wire 2 is \( 1.0 \times 10^{-3} \) newtons per meter. Note that the force on Wire 2 from Wire 1 is the same magnitude.

**Getting centered: Finding magnetic field from current loops**

Suppose you have a loop of current, such as you see in Figure 4-11. The magnetic field from a single loop of wire (even if it has many turns) is not constant over the various points in space.

That variation in the magnetic field is a bit of a problem, because the actual equation for the magnetic field from a loop of current is very complicated. So physicists do what they always like to do — they simplify. Here, simplifying takes the form of measuring the magnetic field at the very center of the loop. (In the next section, you see that putting multiple loops together to form a tube of current also smoothes out the magnetic field.)

Here, start by noting that the magnetic field at the center of a loop of current is equal to the following:

\[
B = N\frac{\mu_0 I}{2R}
\]

where \( N \) is the number of turns in the loop, \( I \) is the current in the loop, and \( R \) is the radius of the loop.

What’s the direction of the \( B \) field at the center of the loop of wire? You guessed it — there’s a right-hand rule for that. To apply this rule, just wrap the fingers of your right hand around the loop in the direction the current is going — your right thumb points in the direction that the \( B \) field points in the center of the loop.
Try some numbers. Say that you have a loop of 200 turns of wire and a radius of 10 centimeters. What current would you need to get the equivalent of the Earth’s magnetic field, 0.6 gauss, in the center?

Plug in the numbers, making sure you first convert from gauss to teslas (1.0 G = 1.0 \times 10^{-4} \ T) and from centimeters to meters. Here’s what you get:

\[
B = N \frac{\mu_0 I}{2R}
\]

\[
6.0 \times 10^{-5} \ T = \frac{(200)(4\pi \times 10^{-7} \ T\cdot m/A)I}{2(0.10 \ m)}
\]
Solve for \( I \) to find the answer:

\[
I = \frac{2(0.10 \text{ m})(6.0 \times 10^{-5} \text{T})}{(200)(4\pi \times 10^{-7} \text{T} \cdot \text{m/A})} = 4.8 \times 10^{-2} \text{ A}
\]

You’d need \( 4.8 \times 10^{-2} \) amps in this current loop to mimic the Earth’s magnetic field.

Armed with this knowledge, you can understand how an electromagnet works. An electromagnet is simply made of a loop of wire with many turns, usually wound around a piece of iron to concentrate the field. When the current flows, this device produces a magnetic field. So you don’t have to dig in the Earth to find magnetic rocks anymore — you can make a magnet that works at the flick of a switch.

### Adding loops together: Making uniform fields with solenoids

One of the major problems with loops of current is that the magnetic field isn’t constant over various points in space, which is why physicists talk in terms of the magnetic field at the center of a loop.

To get around that problem, you can assemble many loops of current next to each other, just a little distance apart, to create a solenoid. This gives you a uniform magnetic field — just as parallel plate capacitors give you a uniform electric field (see Chapter 3 for info on parallel plate capacitors). How does the magnetic field become constant inside a solenoid? When you put multiple loops next to each other, as in Figure 4-12a, the edge effects of the loops cancel, and you get a uniform magnetic field, as in Figure 4-12b.

What is the magnitude of a solenoid’s magnetic field? If the length of the solenoid is large compared to its radius, you get this equation for the magnetic field:

\[
B = \mu_0 nI
\]

where \( n \) is the number of wire loops in the solenoid per meter — that is, the number of turns per meter — and \( I \) is the current in each turn.

How about the direction of the magnetic field? That’s easy enough: You can use the right-hand rule for current loops (see the preceding section) to figure out which direction the magnetic field goes in for a solenoid. Just take a look at Figure 4-12 to confirm you’re getting it right.
Here's an example. Say that you're conducting some crucial lab experiments and need a 1.395-tesla magnetic field. How much current would you need to run through a solenoid of some 3,000 loops, 1.00 centimeters in length, to get that magnetic field?

Start by solving for the current, $I$:

$$I = \frac{B}{\mu_0 R}$$
Then just plug in the numbers; note that because you have 3,000 1-centimeter loops, you use 300,000 — or $3.0 \times 10^5$ loops per meter — as your value for $n$:

$$I = \frac{B}{\mu_o n} = \frac{1.395 \text{ T}}{(4\pi \times 10^{-7} \text{T}\cdot\text{m/A})(3.00 \times 10^5)} = 3.70 \text{ A}$$

In other words, you need about 3.70 amps, which isn’t too much.
In Physics I, you work with direct-current (DC) circuits, where the current is driven by a battery. Here, you take a look at alternating current (AC) in circuits. Things get more active, because you’re dealing with *alternating voltages*, which means that the voltage in any wire changes from positive to negative and then back again regularly.

You may wonder why alternating current is such a big deal. Many types of circuits, including those you tune to receive signals from airborne waves, would be impossible without it. But alternating current got its start in a big way when people first started sending electricity through power lines. Direct current, which doesn’t alternate, could travel only a short distance before the resistance of the wires overcame the current. Alternating current, however, can travel much farther with no problem (it actually helps regenerate itself through alternating magnetic and electric fields). That’s why power lines always carry alternating current.

You see three types of circuit elements in this chapter: resistors, capacitors, and inductors. They all react differently to alternating current. It’s going to be quite a ride, but I offer a guided tour of the whole shebang, so sit back and relax.

**AC Circuits and Resistors:**

**Opposing the Flow**

Resistors are the easiest components to handle when dealing with AC circuits, perhaps because resistors don’t care a bit if the current through them is alternating or not — they react in exactly the same way to alternating and direct voltages.
A resistor is a circuit element that literally resists current to some degree. Here’s how it works: As I explain in Chapter 3, when there’s a potential difference across a metal, for instance, the electric field induces a current by making the negatively charged electrons flow. As the electrons flow and barge their way past the atoms, they jostle the atoms, so the electrons encounter a resistance to their progress. Here are a couple of important points:

✓ The larger the potential difference you put across the metal, the stronger the electric field and the greater the current.
✓ The greater the resistance of the resistor, the less current you get for a given potential difference across it.

In an ideal resistor, the current is proportional to the potential difference. The size of the potential difference you need to make 1 unit of current flow is called the resistance.

In this section, you see how current, voltage, and resistance relate through Ohm’s law for AC circuits. I also show you how voltage and current relate graphically when you have a resistor in an AC circuit.

### Finding Ohm’s law for alternating voltage

The current through a resistor is related to the voltage across the resistor by Ohm’s law:

\[ I = \frac{V}{R} \]

where \( I \) is current, \( V \) is voltage, and \( R \) is resistance measured in ohms (Ω). So if you know the voltage across a resistor, you can find the current through it. Simple.

Now take this picture from direct current to alternating current. To do that, upgrade from batteries — that is, sources of constant voltage — to alternating voltage sources.

The voltage from an alternating voltage source is not constant — it usually varies like a sine wave. You can see what the voltage from an alternating voltage source looks like in Figure 5-1. The peak voltage — that is, the maximum voltage — is equal to \( V_o \).

Here’s the formula for the voltage from an alternating voltage source as a function of time:

\[ V = V_o \sin(2\pi ft) \]

where \( V_o \) is the maximum voltage and \( f \) is the frequency of the alternating voltage source. Frequency is measured in hertz, whose symbol is Hz. Frequency is the number of complete cycles (from peak to peak along the sine wave, for example) that occur per second.
How does this alternating voltage affect Ohm’s law? Not too badly — Ohm’s law just becomes

$$I = \frac{V_o}{R} \sin(2\pi ft)$$

You can rewrite Ohm’s law in terms of the maximum current, $I_o$, like this (because $V_o = I_o/R$). So here’s Ohm’s law for an AC circuit:

$$I = \frac{V_o}{R} \sin(2\pi ft) = I_o \sin(2\pi ft)$$

**Averaging out: Using root-mean-square current and voltage**

When discussing AC circuits, you usually don’t work in terms of maximum voltages and currents, $V_o$ and $I_o$; instead, you speak in terms of the root-mean-square voltages and currents, $V_{\text{rms}}$ and $I_{\text{rms}}$. What’s the difference?

*Root-mean-square* is a way of treating circuits with alternating voltages much as you’d treat circuits with direct, nonalternating voltages. For example, here’s what the power dissipated as heat in a circuit with nonalternating voltage looks like:

$$P = IV$$

And here’s what the dissipated power looks like in a circuit with alternating voltage:

$$P = I_o V_o \sin^2(2\pi ft)$$

Not exactly the same, are they? So physicists talk in terms of the *average power* dissipated by a circuit with an alternating current source — that is, averaged over time. That’s a way of looking at alternating-voltage circuits much as you’d look at battery-driven circuits. The time average of $\sin^2(2\pi ft)$
works out to be $\frac{1}{2}$, which is nice, so the average power dissipated by an alternating voltage circuit is

$$P_{\text{avg}} = \frac{I_o V_o}{2}$$

You can also write this as

$$P_{\text{avg}} = \frac{I_o}{\sqrt{2}} \frac{V_o}{\sqrt{2}}$$

And that’s where $I_{\text{rms}}$ and $V_{\text{rms}}$ come from, because you can also write this as

$$P_{\text{avg}} = I_{\text{rms}} V_{\text{rms}}$$

where $I_{\text{rms}} = \frac{I_o}{\sqrt{2}}$ and $V_{\text{rms}} = \frac{V_o}{\sqrt{2}}$.

So $I_{\text{rms}}$ and $V_{\text{rms}}$ are each the maximum current or voltage, divided by the square root of 2.

**Staying in phase: Connecting resistors to alternating voltage sources**

Say that you connect an alternating voltage source to a resistor, as Figure 5-2 shows. The circle around the ~ symbol represents an alternating voltage source, and the zigzag represents the resistor.
The voltage across the resistor is just the voltage supplied by the alternating voltage source, so the current through the resistor, at time $t$, is given by Ohm’s law:

$$I = \frac{V_0}{R} \sin(2\pi ft) = I_o \sin(2\pi ft)$$

Note that if you square both sides of this current-voltage relationship and then take the average (remember that the average of $\sin^2(2\pi ft)$ works out to be $\frac{1}{2}$), then you have a relation between the mean-squared voltage and current. If you take the square root, you get the following relation between the root-mean-square voltage and current across a resistor:

$$V_{\text{rms}} = I_{\text{rms}} R$$

This is the root-mean-square equivalent of Ohm’s law in an AC circuit.

You can see a graph of the current and voltage across the resistor in Figure 5-3. Note that the current through the resistor and the voltage across the resistor rise and fall at the same time. That means that the current and the voltage in a resistor are in phase. (However, the current and voltages through and across capacitors and inductors do not mirror each other — that is, they’re not in phase, as you see later in this chapter.)

**AC Circuits and Capacitors: Storing Charge in Electric Field**

A capacitor is a device that stores charge when you apply a voltage across it. You may have already met the capacitor in Chapter 3, in the form of two parallel plates. The more charge you put on the plates, the greater the potential difference between them.
Generally, for any type of capacitor, the amount of charge stored for every unit of potential difference is called the capacitance (measured in farads, a unit named after Michael Faraday). The voltage across a capacitor \( V \) that has capacitance \( C \) is related to the amount of charge stored on it \( Q \) by the following equation:

\[
V = \frac{Q}{C}
\]

How does a capacitor react to alternating voltage? That’s what you look at in this section.

**Introducing capacitive reactance**

Suppose you connect a capacitor to an alternating voltage source, as Figure 5-4 shows (the symbol for a capacitor is two upright bars, meant to represent the plates of a parallel plate capacitor).

Here’s how voltage relates to current when you have a capacitor and an alternating voltage source:

\[
V_{\text{rms}} = I_{\text{rms}} X_c
\]

where \( V_{\text{rms}} \) and \( I_{\text{rms}} \) are the root-mean-square voltage and current (the maximum voltage and current divided by the square root of 2 — see the earlier section “Averaging out: Using root-mean-square current and voltage” for details). Here, \( X_c \) is called the capacitive reactance, and it’s equivalent to the resistance, \( R \), in the root-mean-square voltage and current relation for the resistor (see the earlier section “Staying in phase: Connecting resistors to...” for details).
Chapter 5: Alternating Current and Voltage

alternating voltage sources”). \( X_c \) is measured in ohms (Ω), just as \( R \) is, and it’s equal to the following:

\[
X_c = \frac{1}{2\pi fC}
\]

where \( f \) is the frequency of the alternating voltage source and \( C \) is the capacitor’s capacitance, measured in farads (F).

You can think of the capacitive reactance as the effective resistance the capacitor puts in the way of the alternating voltage source, much like \( R \) for resistors.

Note that the capacitive reactance depends on frequency, which is something that resistance doesn’t do. When the frequency \((f)\) is low, the capacitive reactance \((X_c)\) is large, and when the frequency is high, the capacitive reactance is small. (The equation shows this relationship by putting \( f \) in the bottom of the fraction.)

Why is capacitive reactance high when the frequency is low and low when the frequency is high? Intuitively, you can think of it this way: When the frequency is high, the capacitor doesn’t have much time between voltage reversals to accumulate new charge, so it doesn’t change the voltage across it much. When the frequency is low, the capacitor has more time to accumulate charge during each cycle and so can change its voltage more.

How about seeing some numbers? Say you have a 1.50-μF capacitor (where μF is a microfarad, or \( 10^{-6} \) F), and you connect it across a voltage source whose \( V_{\text{rms}} = 25.0 \) volts. What is \( I_{\text{rms}} \) if the frequency of the voltage source is 100 hertz?

First, find the capacitive reactance:

\[
X_c = \frac{1}{2\pi fC} = \frac{1}{2\pi (100 \text{ Hz}) (1.50 \times 10^{-6} \text{ F})} = 1,060 \Omega
\]

So the capacitive reactance is 1,060 ohms. Now put that to work finding \( I_{\text{rms}} \).

You know that \( V_{\text{rms}} = I_{\text{rms}} X_c \), so

\[
I_{\text{rms}} = \frac{V_{\text{rms}}}{X_c}
\]

Plug in the numbers and solve:

\[
I_{\text{rms}} = \frac{25.0 \text{ V}}{1,060 \Omega} = 2.36 \times 10^{-2} \text{ A}
\]

And there you have the current — \( 2.36 \times 10^{-2} \) amps. If the frequency were higher, the capacitive reactance would be lower, so the current would be higher (because the capacitive reactance is the effective resistance).
Getting out of phase: Current leads the voltage

When you have a capacitor in an AC circuit, the current and voltage sine waves are *out of phase* — that is, they’re shifted in time with respect to each other: One reaches its peak before the other. You can see the applied voltage as a function of time in Figure 5-5 — as well as the actual current that flows in the circuit.

![Figure 5-5: Alternating voltage and current in a capacitor.](image)

Notice that with a capacitor, the current *leads* the voltage — that is, the current reaches its peak before the voltage does. In fact, the current leads the voltage by exactly 90°, or $\pi/2$ radians — that is, one quarter cycle. So when you’re graphing the current and voltage from a capacitor in an alternating voltage circuit, always remember that the current leads.

Why does current reach its peak before voltage does? The answer is simple if you think about how a capacitor works. Current piles charge onto a capacitor, so as long as the current is positive, the capacitor’s voltage increases. When the current is decreasing in magnitude, it will still be positive for a while, so charge is still being added onto the capacitor; thus, the voltage keeps on increasing. Not until the current changes direction and goes negative does the charge start to come off the capacitor, causing the voltage to decrease. Therefore, the voltage reaches its peak after the current does.

In fact, you say that if the applied voltage is $V = V_o \sin(2\pi ft)$, then current, which leads voltage by $\pi/2$ radians, looks like this — note that its argument (in parentheses) reaches a specific value before the voltage does because you’re adding $\pi/2$ to $2\pi ft$:

$$I = I_o \sin\left(2\pi ft + \frac{\pi}{2}\right)$$
Using a little trig, this becomes the following:

\[ I = I_o \cos(2\pi ft) \]

So you can see that if the voltage goes as the sine, current goes as the cosine, so they’re out of phase.

**Preserving power**

Here’s something surprising: The average power dissipated by the capacitor is zero. Why? Well, the power used by an electrical component is \( P = IV \).

Here’s what this looks like for a resistor, where the current and the voltage are in phase:

\[ P = I_o V_o \sin^2(2\pi ft) \]

However, for a capacitor, the power looks like this, because the current and voltage are 90° out of phase:

\[ P = I_o V_o \sin(2\pi ft) \cos(2\pi ft) \]

The time average of \( \sin(2\pi ft) \cos(2\pi ft) \) is zero (because this product spends as much time positive as it does negative), so the average power used by a capacitor is zero:

\[ P_{avg} = 0 \]

This means that no power is lost to the environment as heat (as is the case with a resistor), and in fact, the capacitor spends as much time feeding power back into the circuit as it does getting power from the circuit: The capacitor feeds power back to the circuit when it’s discharging and its voltage is going down, and the capacitor gains energy when it’s being charged up and its voltage is increasing.

**AC Circuits and Inductors: Storing Energy in Magnetic Field**

Just as capacitors store energy in an electric field (that is, charges are separated by some distance, giving rise to an electric field), so inductors store energy — but this time, it’s stored in a magnetic field. For example, a solenoid (see Chapter 4) is an inductor, because when you run current through it, a magnetic field appears — and doing that takes energy. In fact, the electrical symbol for an inductor is just that: a solenoid, or loops of current, as you can see in the circuit in Figure 5-6.
Inductors do the same kind of thing as capacitors: They shift the current relative to voltage from an alternating voltage source. However, instead of leading by $\pi/2$, the current lags by $\pi/2$.

For a capacitor (see the preceding section), the voltage is a function of the capacitance and the charge stored on one plate (the charge stored on one plate is equal in magnitude but opposite in sign to the charge stored on the other plate):

$$V = \frac{Q}{C}$$

A similar relationship exists for inductors, as you see in this section. Here, I show you how inductors produce a voltage based on the concept of Faraday’s law. For that, I introduce the concept of magnetic flux, which you get when a magnetic field passes through a loop of wire. I also explain how inductive reactance, just like capacitive reactance, opposes an alternating current — only this time, voltage comes out ahead of current.

**Faraday’s law: Understanding how inductors work**

Michael Faraday (the same physicist that farads, the units of capacitance, are named after), came up with Faraday’s law, which says the following:

When an inductor suffers a change in magnetic flux, it produces a voltage that tends to resist the change.

What does all that mean? This section explains the physics behind inductors, starting with the idea of magnetic flux.
**Introducing flux: Magnetic field times area**

When a magnetic field goes through a loop of wire, there’s said to be a magnetic flux over the area of the loop. You can see how this works in Figure 5-7. There, a uniform magnetic field \( B \) is going through a wire loop with area \( A \), which is orientated at an angle \( \theta \) to the magnetic field.

![Figure 5-7: Magnetic flux through a loop of wire.](image)

Here’s where it gets strange: In physics, areas are often represented by vectors, and they point directly out of the flat surface whose area they represent. In other words, the vector \( B \) in Figure 5-7 should be familiar — that’s just the magnetic field. But the area vector, \( A \), is new — that’s the vector that’s perpendicular to the wire loop, and its magnitude is the same numerical value as the area of the loop.

**Magnetic flux** is the strength of the magnetic field multiplied by the component of the area vector parallel to the \( B \) field. In other words, magnetic flux is a magnetic field strength multiplied by an area. When the magnetic field is parallel to the area vector, the magnetic flux, whose symbol is \( \Phi \), is \( BA \). On the other hand, when \( B \) is perpendicular to \( A \), no field lines actually go through the wire loop, and the flux is zero. Putting all this together, here’s what magnetic flux is in terms of \( B, A, \) and \( \theta \), the angle between them:

\[
\Phi = BA \cos \theta
\]

**Inducing a voltage to keep the status quo**

Faraday’s law says that if the magnetic flux changes, it induces a voltage around the loop; that voltage creates a current in a way that opposes the change by creating its own magnetic field.

For instance, say that the magnetic field is decreasing in strength. The wire loop wants to keep things the way they are, so it resists change. The wire
loop creates a voltage in itself that causes a current to flow — and that current creates a magnetic field.

The magnetic field is created in such a way as to preserve the status quo; thus, the current flows in a way that adds magnetic field to the applied magnetic field — that is, the applied magnetic field is decreasing, so the current around the loop flows to create more magnetic field to replace what’s being decreased. (The inductor can’t keep the current going forever — it dies away quickly, but while it lasts, it creates a magnetic field to supplement the magnetic field that’s decreasing.)

You can see the result in Figure 5-8, which shows the way that the current would flow if the magnetic field $B$ were decreasing. (Tip: Here’s a chance to show off your right-hand rule prowess from Chapter 4. Verify that the direction of the induced current in Figure 5-8 would flow as shown to add more magnetic field to the decreasing applied magnetic field.)

![Figure 5-8: An induced current in a loop of wire.](image)

**Finding induced voltage using the change in magnetic flux**

How is the voltage induced around the loop of wire related to the change in magnetic flux? The voltage looks like this:

$$V = -N \frac{\Delta \Phi}{\Delta t}$$

That is, the induced voltage is equal to the number of turns in the wire loop ($N$) multiplied by the change in flux ($\Delta \Phi$) divided by the time in which the change in flux takes place ($\Delta t$). The negative sign is there to remind you that the induced voltage acts to oppose the change in flux.
Try some numbers here. Say that you have a solenoid consisting of 40 turns of wire, each with an area of $1.5 \times 10^{-3}$ square meters. A magnetic field of 0.050 teslas is perpendicular to each loop of wire (that is, $\theta = 0^\circ$). A tenth of a second later, $t = 0.10 \text{ s}$, the magnetic field has increased to 0.060 teslas. What is the induced voltage in the solenoid?

Start by finding the change in flux over a tenth of a second. The flux looks like this for each turn of wire:

$$\Phi = BA \cos \theta$$

Therefore, the original flux through each turn of wire is this, bearing in mind that $\theta = 0^\circ$ and the original $B$ field is $B_o$:

$$\Phi_o = B_o A$$

Putting in numbers gives you the following:

$$\Phi_o = (0.050 \text{ T})(1.5 \times 10^{-3} \text{ m}^2) = 7.5 \times 10^{-5} \text{ Wb}$$

Wb stands for weber, the MKS unit of magnetic flux; it’s equal to 1 T-m$^2$.

And the final magnetic flux is equal to this, where $B_f$ is the final magnetic field:

$$\Phi_f = B_f A$$

Plugging in the numbers gives you the following:

$$\Phi_f = (0.060 \text{ T})(1.5 \times 10^{-3} \text{ m}^2) = 9.0 \times 10^{-5} \text{ Wb}$$

So the change in flux is

$$\Delta \Phi = \Phi_f - \Phi_o$$

$$= 9.0 \times 10^{-5} \text{ Wb} - 7.5 \times 10^{-5} \text{ Wb}$$

$$= 1.5 \times 10^{-5} \text{ Wb}$$

This change takes place in 0.10 seconds, and it takes place in all 40 turns of the solenoid, so the equation $V = -N \frac{\Delta \Phi}{\Delta t}$ becomes

$$V = -40 \frac{1.5 \times 10^{-5} \text{ Wb}}{0.10 \text{ s}} = -6.0 \times 10^{-3} \text{ V}.$$  

So there you have it — the voltage the solenoid creates to oppose the change in magnetic flux is 6.0 mV (millivolts). That’s what the induced voltage starts off at — it decays exponentially in time.
**Finding induced voltage using the change in current**

The voltage induced by an inductor looks like this:

\[ V = -N \frac{\Delta \Phi}{\Delta t} \]

However, if you have an *electrical inductor* — that is, a component in a circuit — you don’t typically talk in terms of the change in flux inside that component. Instead, you talk about the change in current through the inductor, because that makes more sense in the context of circuits than speaking of magnetic flux.

How do you relate current through the solenoid and magnetic flux? Plugging in \( \Phi = BA \cos \theta \) gives you the following:

\[ V = -N \frac{\Delta (BA \cos \theta)}{\Delta t} \]

And for a solenoid, \( B = \mu_0 n I \), where \( n \) is the number of wire loops in the solenoid per meter — that is, the number of turns per meter — \( \mu_0 \) is \( 4\pi \times 10^{-7} \) T•m/A, and \( I \) is the current in each turn (see Chapter 4 for details). Also, because you have only one solenoid, with \( n \) turns per meter, then \( N = 1 \). So you can write the voltage as

\[ V = -\frac{\Delta (\mu_0 n I \cos \theta)}{\Delta t} \]

If the current is the only thing changing in an inductor that’s part of an electric circuit, you get this:

\[ V = -\frac{\mu_0 n A \cos \theta \Delta I}{\Delta t} \]

You wrap \( \mu_0 n A \cos \theta \) up into one number — the *inductance* of the inductor, whose symbol is \( L \), and whose units are *henries* (which my friend Henry thinks is a good idea). So you have this equation to tie the induced voltage to the change in current over time:

\[ V = -L \frac{\Delta I}{\Delta t} \]

That’s the result you’re looking for — the inductance connects the change in current over time to the induced voltage. And so all inductors you see in circuits are labeled with their inductance in henries (H).
Introducing inductive reactance

For a resistor, voltage and current relate like this: $V_{\text{rms}} = I_{\text{rms}} R$. And for a capacitor, you have $V_{\text{rms}} = I_{\text{rms}} X_c$, where $X_c$ is the capacitive reactance:

$$X_c = \frac{1}{2\pi f C}$$

So it shouldn’t surprise you that for an inductor, you have another formula that relates root-mean-square voltage and current — the maximum voltage and current divided by the square root of 2 (for more on these terms, see the earlier section “Averaging out: Using root-mean-square current and voltage”):

$$V_{\text{rms}} = I_{\text{rms}} X_L$$

where $X_L$ is the inductive reactance — that is, the effective resistance of the inductor: $X_L = 2\pi f L$.

Note that capacitive reactance gets big when the frequency of the applied voltage gets low, but inductive reactance gets big when the frequency gets big — opposite to capacitors. Why is this? It’s because inductors act to oppose any change in the magnetic fields inside them. And the faster the applied voltage changes, the larger the change in flux divided by time — which means that the induced voltage can get really large when you go to a very high frequency.

Check out an example using inductive reactance. Say that you have an inductor with an inductance of $L = 3.60 \text{ mH}$ (millihenries), and you apply a voltage with a root-mean-square value of 25.0 volts across it at 100 hertz. What’s the induced current in the inductor? Starting with $V_{\text{rms}} = I_{\text{rms}} X_L$, you see that

$$I_{\text{rms}} = \frac{V_{\text{rms}}}{X_L}$$

You know $V_{\text{rms}}$, so you need to figure out $X_L$:

$$X_L = 2\pi f L$$

Putting in the numbers you know gives you the following:

$$X_L = 2\pi(100 \text{ Hz})(3.60 \times 10^{-3} \text{ H}) \approx 2.26 \text{ \Omega}$$

And plugging this into the equation for $I_{\text{rms}}$ gives you the answer:

$$I_{\text{rms}} = \frac{25.0 \text{ V}}{2.26 \text{ \Omega}} = 11.1 \text{ A}$$

So you’d get a pretty hefty 11-amp induced current.
Getting behind: Current lags voltage

How does the current in an inductor behave when you apply an alternating voltage? You can see the result in Figure 5-9, which graphs the current and the voltage in an inductor as a function of time. Note that here, current lags voltage — the opposite behavior from a capacitor, where current leads voltage. When current lags voltage, voltage reaches its peak before current does.

Why does current lag voltage in an inductor? It’s because of the following relation:

\[ V = -L \frac{\Delta I}{\Delta t} \]

Note that this equation means that the voltage is greatest when the current is changing the fastest, because the voltage is directly proportional to the rate of change in current. So when the current is the steepest — when current is changing from negative to positive — it’s changing the fastest, and voltage reaches its highest point. Conversely, when current is flat, it’s not changing much at all, so the voltage goes to zero.

As you may expect, current lags voltage by exactly 90° — that is, \( \pi/2 \) in radians. So if the voltage is \( V = V_o \sin(2\pi ft) \), then current, which lags voltage by 90°, looks like this:

\[ I = I_o \sin \left( 2\pi ft - \frac{\pi}{2} \right) \]

Using a little trig, this becomes the following:

\[ I = -I_o \cos(2\pi ft) \]
So that means that for an inductor, the power looks like this, because the current and voltage are $90^\circ$ out of phase:

$$P = -I_o V_o \sin(2\pi ft) \cos(2\pi ft)$$

Note that, just as for a capacitor, the time average of $\sin(2\pi ft) \cos(2\pi ft)$ is zero, so the average power used by an inductor is zero:

$$P_{avg} = 0$$

**The Current-Voltage Race: Putting It Together in Series RLC Circuits**

Suppose you put a resistor, a capacitor, and an inductor together in the same circuit. The circuit in Figure 5-10 is called a *series RLC circuit* — *series* because all components are connected in series, one after the other (the same current has to flow through all of them) and *RLC* because it’s a resistor-inductor-capacitor circuit (sometimes also called an *RCL circuit*). Note that the behavior of this circuit doesn’t depend on the order of the circuit elements, so RLC or CLR would be just as good of a name.

Why doesn’t the order of a resistor, inductor, and capacitor matter in a circuit? Consider the potential difference across each element: Across the resistor, the potential difference depends only on the current; across the inductor, it depends on only the rate of change of the current; and across the capacitor, it depends only on the sum of the current over time (that is, the charge). So the potential difference across each element depends only on the current, and the same current always flows through each in series, whatever order they’re in.

![Figure 5-10: A series RLC circuit.](image-url)
All the components are fighting each other over whether the voltage leads or lags — the capacitor wants the current to lead the voltage, the inductor wants the current to lag the voltage, and the resistor wants the voltage and the current to be in phase. Who wins? This section tells you where to place your bets.

**Impedance: The combined effects of resistors, inductors, and capacitors**

Earlier in this chapter, you see the relationship between root-mean-square current and voltage for the resistor, the inductor, and the capacitor. When you have a circuit combining various elements like resistors, capacitors, and inductors, then there’s a similar relation for the circuit as a whole. The root-mean-square voltage across the circuit, per unit of root-mean-square current, is called the *impedance*.

**Phasor diagrams: Pointing out alternating voltage and current**

To tackle the problem of alternating voltages in an RLC circuit, you get a new tool: the *phasor diagram*. In this diagram, you represent the various alternating quantities as an arrow that rotates in time — you can see how this works in Figure 5-11:

- The arrow (phasor) on the left represents the alternating voltage $V$, with amplitude $V_0$. The length of the arrow is $V_0$.
- The arrow’s angle from the horizontal, $\theta$, is called the *phase*.

Now if you allow this arrow, initially horizontal, to rotate at a constant frequency $f$, then the phase is $\theta = 2\pi ft$. As you can see in Figure 5-11, if you project horizontally from the phasor at time $t$, then you get the value $V_0 \sin \theta$, which is just an alternating voltage. You can represent the current in the same way with its own arrow. If the current leads the voltage by 90°, for example, then its phasor is rotated 90° further clockwise.
Adding phasors and finding impedance

In Figure 5-12a, you can see three phasors representing the voltages across the resistor ($V_R$), inductor ($V_L$), and capacitor ($V_C$) in a circuit. In this figure, they're shown at time $t$, when the phase of the voltage across the resistor is $\theta = 2\pi ft$. Notice the voltage across the inductor leads the voltage across the resistor by $90^\circ$, and the capacitor lags by $90^\circ$.

The total potential difference across all the circuit elements, $V$, is just the sum of the potential difference across each element. So to find $V$, add the phasors using vector addition (see Chapter 2 for info on adding vectors). Now, because $V_L$ and $V_C$ are always $180^\circ$ out of phase, they simply point in opposite directions, so their sum is a new vector whose length is the difference in the amplitude of these two voltages.

The direction of this new phasor ($V_L + V_C$) is still $90^\circ$ from $V_R$, because you've added two phasors that are both $90^\circ$ from the phasor of $V_R$. To get the total sum, add $V_R$ to this new phasor. You can see the sum of the voltages in Figure 5-12b. Because the phasors are at right angles, you can use the Pythagorean theorem to find the resulting length. The squared length of the sum of the potential differences is

$$V^2_0 = (V_{R,0})^2 + (V_{L,0} - V_{C,0})^2$$

where $V_{0}$, $V_{R,0}$, and $V_{C,0}$ are the amplitudes of the voltages.

Now if you use the relation between the amplitude $V_0$ and the root-mean-square voltage $V_{rms}$, you can use this to write the root-mean-square total voltage as

$$V_{rms}^2 = V_{R,rms}^2 + (V_{L,rms} - V_{C,rms})^2$$
where $V_{R,\text{rms}}$, $V_{L,\text{rms}}$, and $V_{C,\text{rms}}$ are root-mean-square voltages across the resistor, inductor, and capacitor respectively.

To simplify this equation, recognize that the following equations are true (note that because the current flows through all the components in series, only one current, $I_{\text{rms}}$, is in the circuit):

- $V_{R,\text{rms}} = I_{\text{rms}} R$
- $V_{C,\text{rms}} = I_{\text{rms}} X_C$
- $V_{L,\text{rms}} = I_{\text{rms}} X_L$

Therefore, you can put the equations together and solve for $V_{\text{rms}}$:

$$V_{\text{rms}}^2 = I_{\text{rms}}^2 [R^2 + (X_L - X_C)^2]$$
$$V_{\text{rms}} = I_{\text{rms}} [R^2 + (X_L - X_C)^2]^{1/2}$$

Now you’re getting somewhere — you have $V_{\text{rms}}$ in terms of $I_{\text{rms}}$. This equation has the form

$$V_{\text{rms}} = I_{\text{rms}} Z$$

where $Z = [R^2 + (X_L - X_C)^2]^{1/2}$. Very nice. Now you’ve connected $V_{\text{rms}}$ to $I_{\text{rms}}$ with this new quantity, $Z$. $Z$ is called the *impedance* of the whole series RLC circuit, and it functions like the effective resistance of the whole RLC circuit.

**Determining the amount of leading or lagging**

For a series RLC circuit, $V_{\text{rms}} = I_{\text{rms}} Z$ (see the preceding section to find out where this equation comes from). That connects $V_{\text{rms}}$ and $I_{\text{rms}}$ in terms of their magnitude. But which leads — voltage or current? And by how much?

Look at a voltages-as-vectors graph. In Figure 5-13, I’ve added $I$ (which is in phase with the voltage across the resistor, so it overlaps $V_R$) as a thick vector.

The question of whether voltage or current leads becomes, “What’s the angle $\theta$ (as shown in the figure) between $V$ and $I$?” Here’s why:

- If that angle is positive, the net result of all three components is that the voltage leads the current.
- If that angle is negative, voltage lags the current.
According to the figure, the tangent of this angle is

\[
\tan \theta = \frac{V_{L,0} - V_{C,0}}{V_{R,0}} = \frac{V_{L,\text{rms}} - V_{C,\text{rms}}}{V_{R,\text{rms}}} = \frac{I_{\text{rms}}X_L - I_{\text{rms}}X_C}{I_{\text{rms}}R}
\]

Note that in the second line, I’ve used the fact that the root-mean-square voltage is just the amplitude divided by the square root of 2; then I canceled the square root of 2 from the top and bottom of the fraction. Canceling out \(I_{\text{rms}}\) in the last line gives you

\[
\tan \theta = \frac{X_L - X_C}{R}
\]

So take the inverse tangent, \(\tan^{-1}\), to find \(\theta\):

\[
\theta = \tan^{-1}\left(\frac{X_L - X_C}{R}\right)
\]

That’s the angle by which voltage leads or lags the current across all three elements. If \(\theta = 0^\circ\), the voltage is in phase with the current — the effects of the inductor cancel out those of the capacitor. If it’s positive, the inductor is winning; if it’s negative, the capacitor is winning.

**Finding the root-mean-square current**

How about some numbers? Say that you have a circuit consisting of a 148-ohm resistor, a 1.50-microfarad capacitor, and a 35.7-millihenry inductor. The circuit is driven by an alternating voltage source with a root-mean-square voltage of 35.0 volts at 512 hertz. What is the root-mean-square current through the circuit, and by how much does the current lead or lag the voltage?
Start by getting the reactance of the capacitor and the inductor. They look like this:

- **Capacitor**: \( X_C = \frac{1}{2\pi f C} = \frac{1}{2\pi (512 \text{ Hz})(1.50 \times 10^{-6} \text{ F})} = 207 \text{ Ω} \)
- **Inductor**: \( X_L = 2\pi f L = 2\pi (512 \text{ Hz})(35.7 \times 10^{-3} \text{ H}) = 115 \text{ Ω} \)

The impedance is

\[
Z = \left[ R^2 + (X_L - X_C)^2 \right]^{1/2}
\]

Plugging in the numbers for resistance, inductive reactance, and capacitive reactance gives you the following:

\[
Z = \left[ 148 \text{ Ω}^2 + (115 \text{ Ω} - 207 \text{ Ω})^2 \right]^{1/2} = 174 \text{ Ω}
\]

And because \( I_{\text{rms}} = \frac{V_{\text{rms}}}{Z} \), you get the following answer for the current:

\[
I_{\text{rms}} = \frac{35.0 \text{ V}}{174 \text{ Ω}} = 0.201 \text{ A}
\]

**Quantifying the leading or lagging**

Now, does the root-mean-square current lead or lag the voltage? In this example, the capacitive reactance (207 ohms) is greater than the inductive reactance (115 ohms), so you can say that the capacitor wins here and the voltage lags the current. But by how much?

Take the following equation:

\[
\tan \theta = \frac{X_L - X_C}{R}
\]

Just plug in the numbers. Because the resistance is 148 ohms, you find that

\[
\tan \theta = \frac{115 \text{ Ω} - 207 \text{ Ω}}{148 \text{ Ω}} = -0.62
\]

So take the inverse tangent to find the angle:

\[
\theta = \tan^{-1}(-0.62) = -32^\circ
\]

And there you have it — voltage does indeed lag the current, just as in a capacitor.
Peak Experiences: Finding Maximum Current in a Series RLC Circuit

Earlier in this chapter, the resistor, capacitor, and inductor all have fixed values, as does the applied voltage. But all those things can vary: You can use electrical components that let you vary their resistance, their capacitance, their inductance, their voltage — even the frequency of that voltage. If you’re going to vary anything in an RLC circuit, varying the frequency is the most common choice. This section tells you how to find the frequency at which you get the most current.

Canceling out reactance

When you let various quantities vary in an RLC circuit, the amount of current through the circuit changes. Because \( V_{\text{rms}} = I_{\text{rms}} Z \), where \( Z = [R^2 + (X_L - X_c)^2]^{1/2} \), you have the following:

\[
I_{\text{rms}} = \frac{V_{\text{rms}}}{Z}
\]

Note that the current through the circuit (\( I_{\text{rms}} \)) will reach a maximum when \( Z \), the impedance, reaches a minimum — that is, when \( Z \) is at its smallest value. Because \( Z = [R^2 + (X_L - X_c)^2]^{1/2} \), impedance will reach its minimum value when the inductive reactance equals the capacitive reactance:

\[
X_L = X_c
\]

At that point, \( Z = R \).

Note that in this case, when the circuit is in resonance and the effects of the inductor and capacitor cancel each other out, the current and voltage are in phase.

Finding resonance frequency

The frequency at which the current reaches its maximum value is called the *resonance frequency*. At the resonance frequency, the effects of the capacitor and the inductor cancel out, leaving the resistor as the only effective element in the circuit.
Resonance: Getting big vibrations

Resonance is not just a feature of electrical circuits; it’s a general feature of oscillating systems — pendulums and even bridges and skyscrapers can experience it. The oscillating system may wobble with a particular amplitude when driven at a particular frequency, such as when you apply an AC voltage of frequency \( f \) to your circuit or when an earthquake shakes a skyscraper.

If you want to make things wobble the most, it’s not a case of driving at the highest frequency that you can. The system likes to wobble at a certain natural frequency, and if you drive it at this frequency, then you get the biggest response — this is the resonance frequency. There’s a particular frequency in the circuit that gives the greatest amplitude current if you apply the voltage at that frequency. (By the way, anyone designing a skyscraper will make sure that its resonance frequency is different from the frequency at which earthquakes shake!)

What is the resonance frequency for any given RLC circuit? You know that

\[
X_C = \frac{1}{2\pi fC} \quad \text{and} \quad X_L = 2\pi fL.
\]

At the resonance frequency \( f_{res} \), the inductive reactance and capacitive reactance are equal, so the following equation holds:

\[
2\pi f_{res} L = \frac{1}{2\pi f_{res} C}
\]

Rearranging this equation and solving for frequency gives you

\[
f_{res}^2 = \frac{1}{(2\pi)^2 LC}
\]

\[
f_{res} = \frac{1}{2\pi (LC)^{1/2}}
\]

And there you have it — that’s the frequency at which the current reaches its maximum value for any given \( L \) and \( C \) values.

Semiconductors and Diodes: Limiting Current Direction

One of the great leaps of the technological age happened when people started combining the resistor and the capacitor with some new circuit elements made from materials that were semiconductors. The combination was an extremely
powerful one. These circuits, combining resistors, capacitors, and semiconductor devices, eventually became miniaturized into integrated circuits, or microchips, which form the basis for many devices that have changed the way people live — most notably the computer. (So the next time someone complains you’re spending too much time on the computer, tell them you’re doing physics.)

In this section, I first introduce semiconductors so you understand their special properties. Then I introduce an example of a circuit element made from them: the diode. This simple device allows current to pass through it in one direction only — it’s effectively a one-way valve for electrical current.

The straight dope: Making semiconductors

Normal silicon (Si) has a crystalline structure, with four electrons from each atom taking part in bonding each atom to its neighbors. Those electrons are in the outermost orbits of the silicon atom, and because they’re important in creating the crystalline structure, they’re not available to conduct electricity — hence, normal silicon is an insulator.

But by being clever, you can introduce small amounts of impurities (such as one part in a million) that give the silicon conducting properties. Here are two types of semiconductors you can create:

- **N-type semiconductors:** Adding some phosphorus (P) atoms allows the silicon to conduct electricity. Phosphorus has five electrons in its outermost orbit, so when you dope silicon with phosphorus, the phosphorus atoms join the silicon crystal structure, which binds each atom to its neighbors using four electrons. That means that there’s one electron from the phosphorus left over — and that electron is free to roam.

  The resulting doped silicon is called an *n-type semiconductor*, because the charges that carry current in it — the electrons contributed by the phosphorus — are negative.

- **P-type semiconductors:** On the other hand, you can dope silicon with other elements, such as boron (B), which has only three outer electrons per atom. When the boron binds to the silicon-crystal structure, one electron is missing, so there’s a “hole” in the number of electrons.

  That hole can move from atom to atom — and each hole produces a positive charge, because it’s formed from a deficit of electrons. Because the holes (that is, the localized places where you have a missing electron) can move throughout the semiconductor, the charge-carriers in this kind of doped silicon are positive. When you have a material with mobile holes, it’s called a *p-type semiconductor*, because the free charge-carriers are positive.

That’s the whole charm of semiconductors — in addition to negatively charged carriers (electrons), you can also have positively charged carriers (the holes).
**One-way current: Creating diodes**

You can create *diodes* — one-way current valves — by putting some *p*-type semiconductor next to some *n*-type semiconductor (see the preceding section for info on types of semiconductors). In the case at the top of Figure 5-14, voltage is applied with the positive voltage connected to the *p*-type semiconductor, and negative voltage is connected to the *n*-type semiconductor.

In this case, charge flows freely across the junction between the *p*-type and *n*-type semiconductors, because the positive holes on the left are repelled from the positive terminal and travel to the right, and the electrons on the right are repelled by the negative terminal, so they travel to the left. The holes and electrons meet at the junction, and the electrons fill the holes — so current can flow. The negative terminal provides more electrons for this process, and the positive terminal removes them, creating more holes.

On the other hand, if you reverse the terminals of the battery, no current will flow through the diode, as the bottom of Figure 5-14 shows. That’s because in this case, the battery drives the mobile charge-carriers away from the junction. As you can see in the figure, the positive holes travel to the left in the *p*-type semiconductor — away from the junction — and the electrons in the *n*-type semiconductor travel to the right, also away from the junction.

What’s left at the junction are the immobile negative charges in the *p*-type material and the immobile positive charges in the *n*-type material. Those charges don’t move, so they set up an electric field that counteracts the electric field set up by the battery — with the net result that all current stops.
Part III
Catching On to Waves: The Sound and Light Kinds

The 5th Wave
By Rich Tennant

“And that’s the Doppler Effect.”
In this part . . .

In this part, you take a look at waves, specifically sound and light waves. You get the lowdown on sound waves and then spend a few chapters on how light waves work, including what happens when they hit mirrors, bend through lenses and diamonds, and pass through slits. Light wave behavior is one of the favorite topics of physicists, and in this part, you see why.
Waves are all around you — water waves, sound waves, light waves, even waves in jump ropes. (Do the waves in that starlet’s hair count? Not in this chapter.) Waves are such a huge topic in Physics II that I cover them in detail in the next five chapters. In fact, even matter travels in waves and is subject to the same kinds of effects as light waves, including reflection (see Chapter 12 for details on this surprising behavior).

In this chapter, you investigate just what waves are and how they work — and how to describe them mathematically (physicists love describing things mathematically). You work with formulas and get to do a little graphing, too. I wrap up by describing some typical wave behavior. Later, in Chapters 7 through 11, you work with specific types of waves: sound and light.

Energy Travels: Doing the Wave

Understanding waves begins with being able to recognize their characteristics. Here are a few key features of waves that you can discover just from watching water waves:

A wave is a traveling disturbance. Waves don’t occur when a surface such as water is calm. Suppose you and some friends are in a sailboat on a lake when a motorboat roars past, sending your boat bobbing. First, you notice that the surface of the lake is now filled with waves and ripples. The water was disturbed by the motorboat, and that disturbance is being sent all around the lake. When a lake is calm, you don’t have any waves; when a lake is disturbed, you have waves. So something must disturb the water in order to create water waves. The thing that’s disturbed by a wave is called the medium.
A wave transfers energy. All waves transfer energy. In fact, waves are one of the primary means of getting energy from Point A to Point B. Continuing with the earlier example, you realize your sailboat is being lifted up and down in the wake of the motorboat. Lifting the boat takes energy — elevating the boat adds potential energy to it. The humps of water in the waves surrounding you all have potential and kinetic energy.

A wave doesn’t cause bulk transport of the underlying medium (if there is an underlying medium). As a wave travels, the medium wobbles, or oscillates, about its undisturbed position, but it doesn’t shift on the whole — this is what I mean by “no bulk transport.” Each part of the medium oscillates about its resting state without changing on average.

For example, suppose you notice a leaf floating on the lake, going up and down with each passing wave. Even though the waves look like they’re traveling away from your boat, the leaf isn’t moving anywhere except up and down. That’s because the water isn’t really traveling across the lake — the wave is. The wave seems to move on to the next patch of water, then the next, and so on, without making any one part of the water travel across the lake. That is, there’s no bulk movement of the water. No mass of water is moving across the lake; each wave just moves each successive region of water up and down as it passes.

Waves — these traveling disturbances carrying energy — come in two types: transverse and longitudinal. The kind depends on which direction the energy disturbance is traveling. This section takes a look at both wave types.

Up and down: Transverse waves

A transverse wave moves up and down, creating peaks of movement. The motion of this type of wave disturbance is perpendicular to the direction the wave is moving in. If you’ve ever had a vacuum cord get stuck while you were vacuuming and yanked on the cord to dislodge it, you saw a transverse wave in action. When you whipped the cord up and down to free it, waves traveled up and down the cord a little something like Figure 6-1.
Chapter 6: Exploring Waves

Back and forth: Longitudinal waves

In longitudinal waves, the motion of the wave disturbance is parallel to the direction the wave is traveling in. As the different parts of the medium wobble back and forth in the direction of the wave’s travel, they cyclically squash and stretch along the wave. A physicist may call a squashing of the medium compression and the stretching decompression.

This kind of wave can travel only in a medium that’s capable of being stretched and squashed — that is, an elastic medium. For example, a spring can support compression and decompression down its length, but a string can’t. Figure 6-2 depicts a longitudinal wave traveling in repeating cycles of compression and decompression, or pulses.

Most objects are elastic to some extent, so you can send pulses through them. Pulses in the air are referred to as sound, which carries the energy from far-off disturbances to your ears. I discuss sound in Chapter 7.

Figure 6-2: A longitudinal wave.

Wave Properties: Understanding What Makes Waves Tick

All waves, no matter which direction they’re traveling in, have specific parts and properties, such as periods and frequency. In this section, you discover the details of a wave’s basic parts and properties. You also see how all the parts of a wave relate mathematically, as well as what a wave looks like in graph form.

Examining the parts of a wave

To understand waves, you need to have a good grip on the terminology. (How else can you discuss waves with your fellow physicists-in-training?) Take a
Part III: Catching On to Waves: The Sound and Light Kinds

look at Figure 6-3, which lists some important parts of a wave. The subsections that follow delve into these parts in greater detail.

Figure 6-3: The parts of a wave.

**Wavelength**

The distance between one point of a wave and the next equivalent point — such as between neighboring peaks or between consecutive troughs (the lowest points on a wave) — is known as the wave’s *wavelength*. For a longitudinal wave, the wavelength is the distance from one compression to the next.

*Nodes* are specific locations where a wave crosses the axis; there are always two nodes per wavelength. The parts of the medium that are at the nodes of the wave are in their resting, undisturbed positions.

The symbol for wavelength is $\lambda$. You usually measure the distance of a wavelength in meters — unless you’re dealing with light waves, which are typically measured in a much smaller unit called *nanometers* (nm), which are billionths of a meter.

**Amplitude**

A wave is a traveling disturbance, and the wave’s *amplitude* tells you how big that disturbance is. Amplitude represents different things depending on whether you’re working with a transverse wave or a longitudinal wave. The amplitude of a transverse wave is a measure of the distance from the axis to a peak, or from the axis to a trough (that should be the same distance). In other words, amplitude is a measure of how high a wave is (see Figure 6-3). Generally, the amplitude of a wave is half of the peak-to-trough distance.

For longitudinal waves, such as sound waves, amplitude corresponds to the pressure in each pulse. I explain the amplitude of sound waves in Chapter 7.
The symbol for amplitude is \( A \), but the units of measurement for amplitude vary depending on which kind of wave you’re dealing with. For example, the amplitude of a water wave on the surface of a lake is measured in units of distance (such as meters or feet) because you’re trying to find out how high the wave is. The amplitude of a light wave, on the other hand, which alternates between magnetic and electric fields, can be measured in teslas and volts per meter (although the amplitude is truly tiny amounts of both).

**Periods and cycles**

Waves are *periodic*, alternating and repeating in a certain amount of time, as you can see in Figure 6-3. If you go from one part of a wave to the same part again — like from peak to peak in a transverse wave or compression to compression in a longitudinal wave — you’ve gone through one *cycle*. In other words, if you see five peaks or compressions go past, you know that five wave cycles have been completed.

The time it takes to complete a cycle is referred to as the wave’s *period*. So if you see a peak of a transverse wave, wait a moment, and see another peak, you know that one period has passed. You measure periods (symbol \( T \)) in seconds.

**Frequency**

*Frequency* measures the number of times something happens per second. Wave frequency is measured in cycles per second. And because cycles are just numbers, that means the unit for frequency is \( s^{-1} \). Of course, \( s^{-1} \) goes by another, more common name: *hertz* (symbol Hz).

The symbol for frequency is \( f \). To calculate frequency, just take 1 over the period \( T \), like so:

\[
    f = \frac{1}{T}
\]

So, for example, a wave that has a period of \( \frac{1}{100} \) seconds has a frequency of 100 cycles per second, or 100 Hz.

**Relating the parts of a wave mathematically**

Knowing the parts and properties of waves is all well and good, but you also need to be able to do something with them. That’s where the math comes in. By applying a little math to what you know about waves, you’re in a position to say more about them. For instance, you can tell someone how fast a particular wave travels, or you can figure out the wavelength. This section shows you how.
Getting a general formula for wave speed

Speed is the distance traveled divided by the time it took to go that distance, so the speed of a wave is simply the distance that a peak travels divided by the time it took to do so. In other words, you divide the wavelength by the period like this:

\[ v = \frac{\lambda}{T} \]

Because the frequency, \( f \), is \( 1/T \), you can write the basic equation for calculating the wave speed as

\[ v = \lambda f \]

A short message from our sponsors: Calculating wavelength of a radio signal

Try putting some numbers in the general wave speed formula. Say that you’re listening to a radio station, 1230 AM on your dial. What’s the wavelength of that radio signal?

The frequency of the wave is 1230, but 1230 what? AM frequencies are measured in kHz (kilohertz), so that’s a frequency of \( 1,230 \times 10^3 \) Hz, or \( 1.23 \times 10^6 \) Hz.

Because \( v = \lambda f \), you can rearrange the formula to solve for wavelength:

\[ \lambda = \frac{v}{f} \]

All you need now is the speed of the radio signal. Radio signals travel at the speed of light \( (v \approx 3.00 \times 10^8 \text{ meters per second}) \), so plug in the numbers and solve:

\[ \lambda = \frac{3.00 \times 10^8 \text{ m/s}}{1.23 \times 10^6 \text{ Hz}} = 244 \text{ m} \]

So the wavelength is about 244 meters, or 800 feet. The next time you’re listening in on a frequency of 1230 kHz, you can say you’re listing in on a wavelength of 800 feet. Or if you really want to blow your mind, think of the radio signal as a wavelength 800 feet long coming at you 1.23 million times a second. Whoa!

A tense situation: Figuring out the speed of a transverse wave

Sometimes, you can say more than just \( v = \lambda/T \) — you can figure out what the wave speed is for a given setup using properties of the system itself. For example, if you have a string under tension, you can calculate the speed of waves in the string given only the force of tension, the mass of the string, and its length.
Actually, you don’t even need to know the mass and the length of the string — you just need to know the mass per unit length, \( \mu \), which is

\[
\mu = \frac{m}{L}
\]

where \( m \) is the mass in kilograms and \( L \) is the length in meters.

At tension \( F \) (where \( F \) stands for force), the speed of transverse waves in the string turns out to be

\[
v = \left( \frac{F}{\mu} \right)^{1/2}
\]

That makes sense — the stronger the tension (the larger \( F \) is), the faster the waves go, and the heavier the string (the larger \( \mu \) is), the slower the waves go.

Say that you have string that’s 20 grams per meter, and it’s under a tension of 200 newtons. How fast does a transverse wave travel in the string if you pluck it? You know that \( v = (F/\mu)^{1/2} \), so plug in the numbers (after converting to kilograms) and solve:

\[
v = \left( \frac{200 \text{ N}}{0.020 \text{ kg/m}} \right)^{1/2} = \left( 1.0 \times 10^4 \text{ m}^2/\text{s}^2 \right)^{1/2} = 100 \text{ m/s}
\]

So the speed of the transverse wave is 100 meters per second.

**Watching for the sine: Graphs of waves**

Graphing a wave gives you an idea of how a wave changes over time. When you graph a wave, whether it’s transverse or longitudinal, you’re really plotting the magnitude of the disturbance. That may be the magnitude of the string displacement or the magnitude of the pulsing water pressure. Because you’re just graphing magnitude, you can graph both transverse and longitudinal waves as sine waves.

Consider the correlation between sine waves and transverse waves: Transverse waves (the kind you create when you whip a string up and down) look just like sine waves. There’s a good reason for that — they are sine waves!

Longitudinal waves are pulses in the direction of travel, which means they don’t look like sine waves. But if you graph the magnitude of the disturbance along a longitudinal wave — pulses and all — you find that a longitudinal wave from a continuous source looks like a sine wave.
Picture a long succession of longitudinal waves traveling through water. Each pulse corresponds to a peak of a sine wave, and the space between pulses corresponds to a trough. So in this case, when you plot the pressure in the wave due to a passing longitudinal wave, you actually get a sine wave (if, of course, the wave source creates normal longitudinal waves).

In this section, I explain how to graph a sine wave that accurately describes a physical wave.

**Creating a basic sine wave**

So what exactly do you need to graph a real-world wave? First, you have to know what your axes are. Because you’re measuring the magnitude of the disturbance created by the wave, your vertical axis is displacement ($y$). And because you want to know how long that disturbance occurs, your horizontal axis is time ($t$).

You want to complete one cycle of the sine wave you’re drawing in one cycle of the actual wave. A single cycle of a wave takes place in one period, and a single cycle of a sine wave takes place in $2\pi$ radians. That means that in one period, you want the sine wave to go through $2\pi$ radians, as Figure 6-4 shows. You can use this expression for the sine wave:

$$y = \sin\left(\frac{2\pi t}{T}\right)$$

![Figure 6-4: The basic sine wave with period $T$.](image)

Note that when $t = 0$, $y = 0$. And when $t = T$, you have $y = \sin(2\pi)$, which equals 0.

You can get frequency into the equation with a little substitution, because you can relate a wave’s period and frequency like so (as I explain in the earlier section “Frequency”):

$$f = \frac{1}{T}$$
Substitute $f$ into the expression for the sine wave to write the expression as

$$y = \sin(2\pi ft)$$

**Adjusting the equation to represent a real-world wave**

The wave equation $y = \sin(2\pi ft)$ is fine, but you probably need to stretch or shift your graph so it accurately depicts your real-world wave. Otherwise, the graph doesn’t give you any information on the strength of the wave or where it was in its cycle when you started taking measurements.

You want your graphed wave to have its own amplitude, $A$, to show how big the disturbance is — a bit tricky to manage because sine waves oscillate between –1 and 1. Multiply the sine wave by $A$ to get the following:

$$y = A \sin(2\pi ft)$$

Of course, the wave’s displacement doesn’t have to be at 0 when $t = 0$. In Figure 6-5, the wave starts off at a nonzero value when $t = 0$, so you need to adjust your wave expression to take this shift into account. Good news: You can adjust the sine’s argument (the value you’re taking the sine of) by an angle, called the phase angle, to make your graphed wave match the behavior of the actual wave.

![Figure 6-5: An offset wave.](image)

Here’s how to add a phase angle to the expression for a wave (note that $\theta$ can be positive or negative):

$$y = A \sin(2\pi ft + \theta)$$
You can also write this equation in terms of a time shift, \( \Delta t \). Here’s how:

\[
y = A \sin(2\pi f (t + \Delta t))
\]

This means that the original wave is shifted in time so that a peak, originally happening at time \( t \), now happens \( \Delta t \) earlier in the shifted wave.

Note that if \( \theta = \pi/2 \), you get a cosine wave:

\[
y = A \sin \left( 2\pi ft + \frac{\pi}{2} \right) = A \cos(2\pi ft)
\]

If you shift the wave by one whole period, it looks exactly like the original. You can see this is true because you know from basic trig that \( \sin(x + 2\pi) = \sin(x) \). So if you shift the wave by \( \Delta t = T \), the wave becomes

\[
y = A \sin(2\pi f(t + \Delta t)) = A \sin(2\pi ft + 2\pi fT) = A \sin(2\pi ft + 2\pi) = A \sin(2\pi ft)
\]

and you have your original wave back again.

**When Waves Collide: Wave Behavior**

Most waves can’t just travel forever without hitting something — some object, or maybe another wave — and that’s what makes wave behavior interesting in the real world. For example, when light waves travel through a glass lens, the waves bend, so people can create eyeglasses and telescopes and binoculars. Here are some important wave behaviors:

- **Refraction:** When waves enter a new material, they can alter their behavior — change their wavelength, for example, or alter their direction. Light waves do this in lenses and prisms, water waves do this in the shallows, and sound waves do this when traveling from air to glass. This process is called wave refraction, and I cover refraction of light waves in Chapter 9.

- **Reflection:** When waves hit something, such as when light waves hit a mirror, they can bounce off, a process known as reflection. Sound waves can reflect off walls, radio waves can reflect off layers of the
atmosphere, TV signals can reflect off buildings, and so on. You can find lots more on reflection in Chapter 10.

**Interference:** Waves can also hit each other, and when they do, they interfere — and the resulting process is called *interference*. For example, you may have seen the ripples from two stones thrown into a lake overlap — and the result is called an *interference pattern*. The waves’ amplitudes can add to each other or cancel each other out. You can find a great deal on interference in light waves in Chapter 11.
Chapter 7
Now Hear This: The Word on Sound

In This Chapter
▶ Exploring the nature of sound
▶ Determining how quickly sound moves through gases, liquids, and solids
▶ Adding sound intensity and decibels to the picture
▶ Taking a look at the behavior of sound waves

Sound is all around you — the sound of talking, the sound of leaves rustling, the sound of traffic, even *The Sound of Music*. Sound travels in perfect longitudinal waves (that is, the wave’s disturbance travels in the same direction as the wave; see Chapter 6 for details). As such, sound waves are a fit topic for physicists.

You get the lowdown on sound in this chapter — how it works, what it can do, and what it can’t do — starting with a look at sound waves as vibrations. You then explore ideas such as the speed of sound, loudness, echoes, and more.

Vibrating Just to Be Heard: Sound Waves as Vibrations

Sound is a vibration in the medium through which the sound is traveling — air, water, metal, or even stone. But it’s not just any vibration; it’s actually a vibration caused by a vibration. A vibrating object makes the air surrounding it vibrate, too, and those vibrations travel away from the vibrating object through the air.

Say you’re dealing with the diaphragm in a loudspeaker (that’s the part that vibrates) and it’s vibrating furiously, pumping out some loud music. Each time the diaphragm pushes against the air, it compresses the air near it. That creates a condensation in the air. This kind of condensation is a small, high-pressure region in the air — a local pulse. As soon as the speaker diaphragm creates the condensation, that condensation starts traveling off into the air.
Conversely, when the diaphragm springs back, that movement creates a small low-pressure region, known as a rarefaction, in the air around the diaphragm. Just as with the condensation, as soon as a rarefaction is created, it starts traveling away from the loudspeaker through the air. Those alternating condensations and rarefactions travel through the air as a longitudinal wave — much like the pulses you can send through a spring when you rapidly compress and decompress one end of it.

So there you have it: Sound is really a longitudinal wave that travels through the air in a series of condensations and rarefactions — that is, pulses. In Figure 7-1, I’ve magnified the column of air that shoots out from the loudspeaker so the air molecules are actually visible. Notice how the air molecules are close together in the condensations and spread out in the rarefactions.

Normal music is made up of many different sound waves, so the pulses you see coming from the loudspeaker have different amplitudes and different frequencies. When these waves enter your ear, the oscillation of the air causes your eardrum to vibrate, and your brain interprets these sounds as having pitch and loudness. Here’s how a sound wave’s amplitude and frequency affect what you hear:

- **Amplitude:** If a sound wave entering your ear has a large amplitude, then you hear a louder sound.

- **Frequency:** If a sound wave entering your ear has a high frequency, then you hear a high-pitched sound. But this can vary from person to person because the sensitivity of different people’s ears to different frequencies of sounds varies.

The human ear can hear a wide range of sound frequencies. Newborns, for example, can hear from 20 hertz (Hz) up to an astounding 20,000 Hz. As you age, you can’t hear the upper range quite so well. An adult, for example, may hear only up to 14,000 Hz. Sound with a frequency higher than 20,000 Hz is called ultrasonic, and sound with a frequency lower than 20 Hz is called infrasonic.
When you have sound that comes out of a loudspeaker in a pure, unwavering tone, the condensations and rarefactions all have the same strength, and they’re all evenly spaced, as in Figure 7-2. The figure shows the waves of condensation and rarefaction of the molecules (not actual size!). You can see that as the molecules are displaced back and forth, they go through cycles of high and low pressure. Where the molecules are squeezed together in a condensation, the pressure is high, and when they’re stretched apart in a rarefaction, the pressure is low — I’ve plotted the wave’s fluctuation in the following graph.

When you have a single tone coming from a loudspeaker, you can speak of the wavelength of the sound, \( \lambda \), and its frequency, \( f \). Regular sound waves, like the one in Figure 7-2, have so many cycles per second, which is their frequency.

**Figure 7-2:** A constant tone.

---

**Cranking Up the Volume: Pressure, Power, and Intensity**

The loudness, or volume, at which you hear a sound is a direct result of the sound wave’s amplitude — that is, the amount of pressure in each pulse in a sound wave. The greater the pressure amplitude, the greater the volume.

Volume is really a subjective measure; a sound may seem louder to one person than to another, based on how good his or her hearing is. But in physics, you use objective measures, such as pressure amplitude and sound intensity, to talk about the sonic boom that rattled your windows or the rock concert that still has your ears ringing.
Amplitude and sound intensity are related. Here’s how: Making a sound wave takes energy, and making a continuous wave takes a flow of energy over time: power. As a wave propagates and spreads out in the surrounding space, this power is spread over a larger area, so sound can become weaker with distance. The amount of power flowing through a unit area is its intensity. Lower intensity causes less energy to enter your ear every second — and with less sound power entering your ear, the wave has a smaller amplitude, because making a wave with smaller amplitude takes less power. That’s why sounds become quieter with distance.

In this section, I discuss the amplitude, power, and intensity of sound waves. Intensity is related to decibels, a way to compare sounds objectively, so I cover decibels here as well.

**Under pressure: Measuring the amplitude of sound waves**

If you wanted to measure the pressure amplitude of sound waves (or if some crazy professor said you had to), you could start with the setup in Figure 7-3. There, a loudspeaker is sending a pure-tone sound through a tube that has many pressure meters at the top of it. (A pure-tone sound is made up of just one frequency, so it’s a monofrequency sound.) Using this setup, you can measure the amplitude of the traveling sound wave by photographing the settings of all the pressure meters at once. This snapshot can also tell you that the pressure in the whole wave forms a sine wave that’s traveling to the right.

Suppose the loudspeaker in Figure 7-3 is set to create a monofrequency wave at about the volume of human speech, and you need to find the maximum amplitude. Pressure is measured in pascals (Pa), and it takes $1.01 \times 10^5$ Pa to make up the pressure of the atmosphere at sea level. The maximum pressure amplitude of human speech is about $3.0 \times 10^{-2}$ Pa, or an amazing $3.0 \times 10^{-7}$ atmospheres! That’s how sensitive the human ear is. Human speech, which can sound very loud, is actually made up of very weak pulses of air. So the pressure amplitude of a sound wave of human speech is relatively small.
Even though sound is a longitudinal wave, you can graph its pressure amplitude as a sine wave, because you’re measuring the displacement of air (it’s just like the amplitude of a transverse wave in a string, because what you measure there is the actual displacement of the string). For a sound wave, condensations form the peaks of the sine wave and rarefactions form the troughs.

**Introducing sound intensity**

Sounds waves transfer a disturbance in a medium from the source to an observer. That means energy is transferred from the source to some target. Leaves rustling in the street transfer a relatively small amount of energy, but some sounds are powerful enough that they can cause damage. Sonic booms, for example, are strong enough to break windows.

So how much energy is transferred by a sound wave in a given amount of time? That’s a measure of *power*, which is measured in watts (abbreviated W). Power is just energy divided by time:

\[ P = \frac{E}{t} \]

In fact, what’s usually measured is the power per unit area some distance from the sound source, as Figure 7-4 shows. This quantity, power divided by area, is the sound wave’s *intensity*. Sound intensity is measured in watts per meter, and the equation for finding it is as follows:

\[ I = \frac{P}{A} \]

Figure 7-4:

Sound intensity is the power of a sound wave divided by the area.

In this section, you calculate sound intensity and see how it relates to decibels.
**Sound intensity in terms of total power of a sound wave**

The intensity of a sound wave differs depending on how far away you are from a sound source. That’s because sound expands in a sphere from a sound source, and the power of a sound wave is distributed over the whole area of that sphere. The following equation shows how the surface area of a sphere ($A$) grows as you get farther away from the sound source — that is, as the radius increases:

$$A = 4\pi r^2$$

where $r$ is your distance from the sound source.

If you know the total power of a sound wave as it comes out of the source, $P_{\text{total}}$, and you know that sound wave is allowed to expand in a sphere, you can write the intensity as a function of $r$ like so:

$$I = \frac{P_{\text{total}}}{4\pi r^2}$$

Thus, the intensity of a sound wave drops off by a factor of 4 (or $2^2$) every time you double your distance from the sound source.

For example, say you have a sound source that pumps out $3.8 \times 10^{-5}$ watts of sound power. What’s the sound intensity 1 meter from the sound source? Well, the total power of the sound energy the source sends out is $3.8 \times 10^{-5}$ watts. At $r = 1.0$ meters, you have this:

$$I = \frac{P_{\text{total}}}{4\pi r^2}$$

$$= \frac{3.8 \times 10^{-5} \text{ W}}{4\pi (1.0 \text{ m})^2} \approx 3.0 \times 10^{-6} \text{ W/m}^2$$

So the sound intensity at 1 meter is $3.0 \times 10^{-6}$ watts per square meter. That’s the approximate sound intensity of human conversation.

**Measuring sound in decibels**

Decibels are a comparison of one sound intensity to a reference intensity on a logarithmic scale. In plain English, that means decibels tell you how much louder or softer a sound is than a standard sound, such as the threshold of hearing (that’s the reference sound physicists usually use).

Here’s the equation for decibels of a particular sound intensity:

$$\beta = 10 \log \left( \frac{I}{I_o} \right)$$
where \( \log \) refers to the logarithm to the base 10 (it’s on your calculator); \( I_o \) refers to the reference sound you’re measuring against (usually the threshold of hearing, \( 1.0 \times 10^{-12} \text{ W/m}^2 \)); and \( I \) is the sound intensity you’re measuring. The abbreviation for decibels is dB.

How about some representative numbers here? Table 7-1 lists some common decibel measurements from 1 meter away from the source, comparing the sound to the threshold of human hearing.

<table>
<thead>
<tr>
<th>Table 7-1</th>
<th>Intensity and Decibels of Common Sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sound</strong></td>
<td><strong>Intensity</strong></td>
</tr>
<tr>
<td>Threshold of hearing</td>
<td>( 1.0 \times 10^{-12} \text{ W/m}^2 )</td>
</tr>
<tr>
<td>Leaves rustling</td>
<td>( 1.0 \times 10^{-11} \text{ W/m}^2 )</td>
</tr>
<tr>
<td>Whisper</td>
<td>( 1.0 \times 10^{-10} \text{ W/m}^2 )</td>
</tr>
<tr>
<td>Normal conversation</td>
<td>( 3.2 \times 10^{-6} \text{ W/m}^2 )</td>
</tr>
<tr>
<td>Car with no muffler</td>
<td>( 3.2 \times 10^{-2} \text{ W/m}^2 )</td>
</tr>
</tbody>
</table>

Say you have a gasoline-powered lawn mower that sounds especially loud, and you want to find out just how loud it is. You measure the sound intensity at 1 meter from the lawn mower as \( 6.9 \times 10^{-2} \text{ W/m}^2 \). How many decibels is that compared to the threshold of hearing?

The threshold of human hearing has a sound intensity of \( 1.0 \times 10^{-12} \text{ W/m}^2 \), so when you plug that into the \( \beta = 10 \log(I/I_o) \) formula, you have

\[
\beta = 10 \log \left( \frac{I}{I_o} \right) = 10 \log \left( \frac{6.9 \times 10^{-2} \text{ W/m}^2}{1.0 \times 10^{-12} \text{ W/m}^2} \right) = 108 \text{ dB}
\]

Your lawn mower generates about 108 dB at a distance of 1 meter away from the sound source. Hmm. Maybe you should start wearing earplugs when you use it!

### Calculating the Speed of Sound

Sound traveling through the air moves pretty fast, but it can move even faster depending on which medium it’s moving through (another gas, a liquid, or a solid). Of course, the only way to really know how fast it’s traveling is to calculate its speed.
The speed of a wave is frequency multiplied by wavelength, which looks like this:

\[ v = \lambda f \]

However, that basic equation doesn’t help you much because the speed of sound can vary depending on the temperature of the medium. But never fear — in this section, I introduce some speed-of-sound formulas that account for both temperature and medium. (And for some real-world values, check out the nearby sidebar titled “At-a-glance stats for the speed of sound.”)

### Fast: The speed of sound in gases

The speed of sound is lowest when it’s traveling through a gas. To calculate the speed of sound in an ideal gas (which approximates air given the temperature of that gas), you rely on an equation that may look familiar to you from Physics I:

\[ v = \left( \frac{\gamma kT}{m} \right)^{1/2} \]

Here’s what the variables represent:

- \( \gamma \) is the adiabatic constant, and it’s equivalent to \( C_p/C_v \), the ratio of the specific heat capacity at constant pressure to the specific heat capacity at constant volume; for air, \( \gamma \) is 1.40.
- \( k \) is the Boltzmann constant from thermodynamics (1.38 × 10^{-23} kg⋅m^2/s^2K^{-1}, or J/K).
- \( T \) is the temperature of the ideal gas according to the Kelvin scale.
- \( m \) is the mass of a single molecule in kilograms.

### At-a-glance stats for the speed of sound

If you like impressing your friends by spouting random bits of knowledge, file away the following values of the speed of sound:

- **Air at 0°C**: 331 m/s
- **Air at 20°C**: 343 m/s
- **Oxygen at 0°C**: 316 m/s
- **Water at 20°C**: 1,482 m/s
- **Copper (temperature independent)**: 5,010 m/s
- **Steel (temperature independent)**: 5,940 m/s
Okay, time to put this equation to work! Go ahead and assume you have in your hands a camera whose rangefinder uses sound to find the distance to the subject. You’ve just taken a photograph of a fellow physicist, and being a physicist, your friend immediately wants to know the distance between the two of you. Checking your camera, you can see that the rangefinder sent a pulse of sound out that bounced off your friend and came back to the camera in 4.00 \times 10^{-2} \text{ seconds.} 

Your handy pocket thermometer tells you that the temperature of the air is 23°C. So just how far away is your friend, assuming you can treat air as an ideal gas?

First, you need to convert that temperature to kelvins by adding 273 to the Celsius temperature, which looks like this:

\[ 23^\circ \text{C} + 273 \text{ K} = 296 \text{ K} \]

So the ideal temperature is 296 kelvins. Great. Now you can use the speed-of-sound equation for gases:

\[ v = \left( \frac{\gamma kT}{m} \right)^{1/2} \]

Notice that in addition to the temperature, you also need the mass \( m \) of a single molecule of air in kilograms. You just happen to remember that the mass of air is 28.9 \times 10^{-3} \text{ kg/mole} (a \text{ mole} is 22.4 liters of ideal gas). So the mass of one air molecule is the mass of a mole divided by the number of molecules in a mole (Avogadro’s number).

\[ m = \frac{28.9 \times 10^{-3} \text{ kg/mole}}{6.022 \times 10^{23} \text{ molecules/mole}} = 4.80 \times 10^{-26} \text{ kg} \]

Bet you always wanted to know that! Right, moving on. For air, \( \gamma \) is 1.40, so this equation allows you to figure the speed of sound at 23°C:

\[
v = \left( \frac{\gamma kT}{m} \right)^{1/2} = \left( \frac{(1.40)(1.38 \times 10^{-23} \text{ kg m}^2\text{s}^{-2}\text{K}^{-1})(296 \text{ K})}{4.80 \times 10^{-26} \text{ kg}} \right)^{1/2} = 345 \text{ m/s} \]

Tada! The speed of sound where you are is 345 meters per second. You can relate the time the signal took and the speed of sound to the distance this way:

\[ \text{Distance} = \text{speed} \times \text{time} \]

So how much time did it take for the sound to travel from your camera to your friend? Well, the camera recorded 4.00 \times 10^{-2} \text{ seconds, but don’t forget}
that’s the time for a round trip (the sound leaving the camera and then returning after bouncing off your friend). So the time sound takes to reach your friend is $4.00 \times 10^{-2} \text{ seconds} \div 2 = 2.00 \times 10^{-2} \text{ seconds}$, which translates to

$$\text{Distance} = (345 \text{ m/s})(2.00 \times 10^{-2} \text{ s}) = 6.90 \text{ meters}$$

And that’s that. Your friend was standing roughly 6.90 meters away from you when you took the photo.

**Faster: The speed of sound in liquids**

Sound travels faster in liquids than it does in gases. That’s because liquids are less elastic than gases, meaning they “bend” less under the same applied force. When you create a disturbance in a liquid, the force opposing that disturbance is greater in a liquid than in a gas, which means the liquid “snaps back” into place quicker. The end result is that the disturbance is chased through the liquid faster than it is in a gas.

So what’s the expression for the speed of sound in liquids? That depends on two main aspects of the liquid:

- **The resistance to deformation:** The speed of sound is a measure of how fast the medium “snaps back” into place after a disturbance, and the measure of that is closely tied to the medium’s *bulk modulus* (the resistance of a substance to being deformed by pressure). In fact, it’s tied to the adiabatic bulk modulus (*adiabatic* means no heat is exchanged with the environment), whose symbol is $\beta_{\text{ad}}$.

  The larger the adiabatic bulk modulus, the more resistance the liquid puts up against being deformed; so the higher the $\beta_{\text{ad}}$, the higher the speed of sound in that liquid.

- **The density:** The speed of sound in liquids is also tied to the liquid’s density ($\rho$). The higher the density of the liquid, the harder it is to get the liquid to move. Thus, the speed of sound is lower in dense liquids.

Putting it all together, you get this for the speed of sound in a particular liquid:

$$v = \left(\frac{\beta_{\text{ad}}}{\rho}\right)^{1/2}$$

where $\beta_{\text{ad}}$ is the adiabatic bulk modulus and $\rho$ is the liquid’s density.

Here’s your chance to practice calculating the speed of sound in a liquid! (Please contain your excitement.) Suppose you and your Physics II classmate take a trip to the seaside, and you want to document the trip with a photo of your friend. Unfortunately, she’s scuba diving, so you need to go underwater to take the photo. You set your camera to take underwater photos and take...
the snapshot. Your friend sees you doing it and comes over, wanting to know how far apart the two of you were.

When you took the underwater photo of your pal, the camera said the sound signal came back in $4.00 \times 10^{-3}$ seconds. Given that the adiabatic bulk modulus of water is $2.31 \times 10^9$ pascals and the density of water is 1,025 kilograms per cubic meter, how far away was your friend?

Calculate the speed of sound in water with this equation:

$$v = \left( \frac{\beta_{ad}}{\rho} \right)^{1/2}$$

$$v = \left( \frac{2.31 \times 10^9 \text{ Pa}}{1,025 \text{ kg/m}^3} \right)^{1/2}$$

$$v \approx 1,500 \text{ m/s}$$

So the speed of sound in water is about 1,500 meters per second. What does that info buy you? Well, you now know how long it took for the sound pulse from the camera to return to the camera, and you also know that distance = speed $\times$ time.

The camera recorded $4.00 \times 10^{-3}$ seconds for a sound pulse to travel from the camera to your friend and back again, so the sound took $4.00 \times 10^{-3} \div 2 = 2.00 \times 10^{-3}$ seconds to reach your friend. Plugging in the speed of sound and the time the sound pulse took, you discover that the distance between you and your pal was

$$\text{Distance} = (1,500 \text{ m/s})(2.00 \times 10^{-3} \text{ s}) = 3.00 \text{ m}$$

**Fastest: The speed of sound in solids**

If the stiffer the medium is, the faster the speed of sound, then it shouldn’t surprise you that sound travels fastest in solids, which are even less elastic than liquids.

So what’s the expression for the speed of sound in solids? Here, you use a combination of

- Young’s modulus, a measure of the stiffness of uniform materials
- The density of the solid

Here’s how Young’s modulus ($Y$) and density ($\rho$) relate to give you the speed of sound in a solid:

$$v = \left( \frac{Y}{\rho} \right)^{1/2}$$
This equation tells you that the higher Young’s modulus is — in other words, the stiffer the medium — the faster the speed of sound. The greater the density of the material, the slower the speed of sound (because the material is slower to react to a disturbance).

Imagine that you’re on an ocean cruise with your significant other. The two of you are standing on the deck, and because you both happen to be physicists, you naturally decide to measure the length of the deck, which is steel. Your significant other stands on the bow of the ship while you stand at the stern. Borrowing a handy fire axe from a fire control station on deck, you tap your end of the deck. Your significant other then reports that the sound took only $2.00 \times 10^{-2}$ seconds to travel through the deck.

Given that Young’s modulus for steel is $Y = 2.0 \times 10^{11}$ N/m² and that the density of steel is $\rho = 7,860$ kg/m³, you can start determining the length of the deck by plugging numbers into the speed-of-sound expression for solids:

$$v = \left( \frac{Y}{\rho} \right)^{1/2}$$

$$v = \left( \frac{2.0 \times 10^{11} \text{ N/m}^2}{7,860 \text{ kg/m}^3} \right)^{1/2}$$

$$v \approx 5.0 \times 10^3 \text{ m/s}$$

So the speed of sound in steel is about $5.0 \times 10^3$ meters per second — that’s 5 kilometers per second, or about 11,000 miles per hour.

You can find the length of the steel deck by multiplying the speed of sound by the time it took the sound to travel, which gives you

$$\text{Distance} = (5.0 \times 10^3 \text{ m/s})(2.00 \times 10^{-2} \text{ s}) = 100 \text{ m}$$

There you have it: The deck is about 100 meters long.

---

**The solid sounds of the railroad tracks**

As a kid, I was able to verify that sound travels faster in solids than through gases using railroad tracks that had those connector bars attached. By putting my ear to the track and watching a friend hit the tracks with a hammer some distance away, I could hear a definite clank-CLANK-clank coming through the tracks and then through the air. *(Note: I don’t recommend this as an experiment, especially if there are any trains lurking about.)*
Analyzing Sound Wave Behavior

This section considers some of the weird and wonderful things that sound waves can do. You see them bouncing and bending and find out what happens when two sound waves meet. This leads you to discover a new kind of wave — the standing wave, which doesn’t propagate; these are the kinds of waves that come from musical instruments. You see what happens when sources of sounds and listeners move. And finally, you break through the sound barrier to find out what happens when sound sources move faster than the speed of sound.

All the properties of sound that I discuss here are also properties of waves generally. So by understanding these aspects of sound behavior, you’re actually getting a lot more in the bargain. For example, grasping sound waves can take you a long way toward an understanding of light and optics in the next few chapters. These wave properties go right to the heart of lots of the workings of the physical world.

Echoing back: Reflecting sound waves

Reflection occurs when a wave encounters a boundary. You’re familiar with the reflection of sound waves in the form of an echo.

In the case of a sound wave in air, the condensation of the air in a high-pressure peak pushes against the air immediately next to it, which becomes condensed in turn, and so the wave spreads. However, if a high-pressure peak meets a solid surface such as a wall and tries to push against it, the wave doesn’t find the wall so forgiving. The high pressure pushes the wall, but the wall pushes back with a force of resistance. The high-pressure peak now sits against the wall and pushes against the air behind it, which, being more forgiving than the wall, allows the high-pressure peak to propagate back the way it came — and the wave is reflected. Physicists say that the wall provides a boundary condition on the wave.

Sound is a longitudinal wave, in which the air molecules oscillate in the direction of motion of the wave. As the wave approaches the wall, the wall restricts the motion of the air molecules. The molecules right next to the wall can’t oscillate at all. As the wave strikes, the molecules of the air near to the wall continue toward the wall until they effectively bounce off it — redirecting their motion in the opposite direction — thereby reflecting the wave. In these terms, the boundary condition on the wave is that there must be zero oscillation at the wall.
Seeing with sound

As you probably know, a bat “sees” not with reflected light but with reflected sound. When a bat hunts, it uses echolocation; it makes clicking sounds that bounce back from any unfortunate insects that may be flying by, and then it listens to the echo.

Bats had a head start, but as physicists came to understand the reflection of sound waves, people were able to use the same ideas for technologies like sonar and sonograms. These devices enabled people to see in a similar way, right down to the depths of the oceans and even inside the human body.

To illustrate this echoing process, I’ve graphed a pressure wave of sound as it reflects off a solid wall in Figure 7-5. For clarity, I haven’t reflected a whole wave. I’ve sent just a part of a wave — a pulse — as a speaker would generate if its diaphragm moved out and then back just once. In this figure, x measures the distance from the wall and p is the pressure fluctuation.

Figure 7-5: The reflection of a single pulse of pressure.
With a different type of boundary may come different boundary conditions. For example, the wave may fall on a soft wall, which is deformable. In this case, you may see some oscillation immediately next to the wall, and the wall may absorb some of the wave’s energy as the molecules do work on the wall to move it. In that case, the reflected wave has a smaller amplitude, and the echo is quieter.

Sharing spaces: Sound wave interference

Two waves can occupy the same place at the same time. When this happens, they’re said to interfere. The resulting oscillation is incredibly simple to work out: Just add the oscillation from one wave to the other. This idea is called the principle of superposition. So at any point, if one wave would cause a displacement of the medium of $y_1$ and the other wave would cause a displacement of $y_2$, then the actual displacement of the medium at that point is simply $y_1 + y_2$.

To see the principle of superposition in action, check out Figure 7-6. It shows the wave displacements over time for two separate waves. In the same graph, I show what would happen if those two waves were traveling through the medium at the same time.

![Figure 7-6: The interference of two waves.](image)

Adding amplitudes: Constructive and destructive interference

Interference can be constructive or destructive. With constructive interference, the amplitudes of two waves combine to make a wave of larger amplitude. With destructive interference, the amplitudes of the waves cancel each other out.
For instance, suppose you have a stereo with a pair of speakers like those in Figure 7-7. Now put on a CD that plays a pure tone — that is, each speaker makes the same sine-shaped sound wave (see the earlier section “Under pressure: Measuring the amplitude of sound waves” for details on why the graph takes this shape). This would be a very uninteresting piece of music to play, but it has some surprising effects.

Say the speakers play a tone with frequency $f$, wavelength $\lambda$, and amplitude $A$. Now sit in a position equally distant from each speaker (point $a$ in Figure 7-7). You receive two waves — one from each speaker. If you call the displacement you experience from the wave travelling from the left speaker $y_1$, you can say that the displacement is given by the sine wave

$$y_1 = A \sin(2\pi ft)$$

Because the other speaker is the same distance away, the displacement of the wave coming from it, $y_2$, is just the same as $y_1$, so $y_1 = y_2$. Now work out the wave that you experience, $y_T$, as you sit at point $a$. Use the principle of superposition and add the waves coming from each speaker:

$$y_T = y_1 + y_2 = A \sin(2\pi ft) + A \sin(2\pi ft) = 2A \sin(2\pi ft)$$

This is just a sine wave with twice the amplitude of the wave from each speaker. That’s not too surprising — if you sit at point $a$, you just hear a louder sound than you would if you had just one speaker instead of two. The two speakers are combining to make a larger-amplitude wave — this is **constructive interference**.

Now suppose you move — sit just to one side (at point $b$ in Figure 7-7) so that you’re exactly half a wavelength closer to the right speaker than you are to the left one. This means that the wave from the right speaker reaches you
Chapter 7: Now Hear This: The Word on Sound

half a period earlier than the wave from the left speaker — that is, it’s shifted by \( T/2 \). So you can write \( y_2 \) as

\[
y_2 = A \sin \left( 2\pi f \left( t + \frac{T}{2} \right) \right)
\]

\[
= A \sin \left( 2\pi ft + 2\pi f \frac{T}{2} \right)
\]

\[
= A \sin \left( 2\pi ft + \pi \right)
\]

\[
= -A \sin (2\pi ft)
\]

\[
= -y_1
\]

Now if you work out the combined wave from both speakers, you see

\[
y_T = y_1 + y_2 = y_1 + (-y_1) = 0
\]

You receive no sound wave at all — silence! The waves from each speaker are canceling each other at point \( b \) — this is destructive interference. You can read about constructive and destructive interference in light waves in Chapter 10.

**Standing waves: Destructive interference at regular intervals**

A standing wave is a kind of wave that doesn’t travel — the peaks simply oscillate at the same place without propagating. This kind of wave occurs when a propagating wave is confined, such as on a piece of string or, as you see in this section, when sound is contained in a tube. Here, I show you how to construct a setup to contain the sound, and you see how sound reflects inside the tube and interferes to produce a standing wave.

**The setup: Getting identical waves going in opposite directions**

Suppose you take a long tube that’s closed at one end and has a diaphragm at the other end (you stretch an elastic sheet over the end, for example). Place a speaker near the diaphragm. When you turn the speaker on, the sound waves cause the diaphragm to vibrate. The sound waves from the diaphragm travel down the tube (acting as the incident wave), reflect off the closed end, and travel back up the tube to the diaphragm again (the reflected wave).

But remember, the speaker does not produce a single pulse; instead, you have a sine wave of sound. So in this situation, you have two waves in the tube at the same time, one traveling away from the speaker and one traveling toward it. Imagine that the reflection is ideal so that both waves have the same amplitude, frequency, and wavelength; their only difference is that they’re traveling in opposite directions.

Now look at the total wave. The wave from the diaphragm travels down the tube to the closed end, where the boundary condition is that there can be
no displacement of the molecules. The displacement of the reflected wave, at the closed end, is always the opposite of the displacement of the incident wave. So when the waves interfere, there’s no displacement (destructive interference) at all at the closed end, satisfying the boundary condition.

Because both waves are periodic, this destructive interference must happen at regular intervals along the tube. To see this, you move away from the wall by half a wavelength. Because both waves are sine waves, they both have the opposite displacement here that they had at the wall, so they’re still both equal and opposite — you have destructive interference again. In this way, at every point along the tube that’s a whole-number of half-wavelengths from the closed end, there’s destructive interference, so the molecules do not oscillate at all in these places. The total wave in this tube must be different from the sine wave you usually see for sound — you get the full picture of this strange new wave next.

**Graphing a standing wave**

Figure 7-8 shows a graph of incident and reflected sound waves at several different times. They’re two identical waves, with the only difference being that they’re travelling in opposite directions. The incident and reflected waves have amplitude $A$, and the horizontal axis measures distance from the wall.
Chapter 7: Now Hear This: The Word on Sound

The figure also shows the interference between these two waves, which is just the sum of the two graphs (according to the principle of superposition). You can see this wave at three different times. Notice that the wave doesn’t travel anywhere; this is a standing wave, and it just stays where it is. It oscillates, but the peaks and troughs do not propagate; they just move up and down.

The part of the wave that crosses the axis (where there’s no displacement) is called the node. Because the wave doesn’t propagate, its nodes don’t move. These nodes are just points of destructive interference.

Between the nodes are points in the wave that oscillate with the greatest amplitude — these are the antinodes. The amplitude of the standing wave at these antinodes is just equal to twice the amplitude of the incident and reflected waves. So you have a picture of the standing wave as a nonpropagating oscillation, which has points of maximum and zero oscillation at intervals of $\lambda/2$ along its length.

**Harmonics: Putting the standing wave in normal mode**

When the oscillation of a diaphragm coincides with the antinode (greatest amplitude) of a standing wave, the standing wave is in a normal mode. Normal modes occur wherever there are standing waves, such as the ones on a vibrating string or in the pipes of an organ.

Suppose you have the closed-tube setup I describe earlier in “The setup: Getting identical waves going in opposite directions.” You set the speaker to make a pure tone with a wavelength that makes a standing wave in the tube, which has an antinode at the diaphragm. The oscillation of the diaphragm coincides with the antinode of the standing wave, so it’s in a normal mode.

Standing waves are in a normal mode when the speaker makes a sound with wavelength $\lambda_n$, which is given by

$$\lambda_n = \frac{4L}{n} \quad (n = 1, 3, 5, \ldots)$$

where $L$ is the length of the tube and $n$ is a whole number that labels the various normal modes.

The frequencies of normal modes of vibration are called the harmonics. The first harmonic ($n = 1$) is called the fundamental frequency. Musicians often call the frequencies of the higher frequency modes overtones.

Note that $n$ must be an odd number because there are an odd number of quarter-wavelengths from the barrier (the closed end of the tube) to an antinode. As $n$ increases, so does the number of nodes and antinodes in your normal mode. So when your speaker makes a sound with wavelength $4L$, there’s only one antinode, which is the one at the diaphragm.
Figure 7-9 shows some normal modes of your tube — the two lines show the
two positions of maximum displacement of the normal mode. Here, the hori-
zontal axis measures distance from the diaphragm, and you can see the posi-
tion of the closed end of the tube on the right side of the graph.

Now try a concrete example — what are the first few notes that your tube
likes to play? Suppose your tube is 0.983 meters long. Then the wavelength of
its normal modes are

\[ \lambda_n = \frac{4}{n} (0.983 \text{ m}) \]

If the speed of sound in your tube is 343 meters per second, then the frequen-
cies of these modes, \( f_n \), are

\[ f_n = \frac{\nu}{\lambda_n} \quad (n = 1, 3, 5, \ldots) \]

\[ = \frac{343 \text{ m/s}}{\frac{4}{n} (0.983 \text{ m})} \]

\[ = n(87.2 \text{ Hz}) \]

This means that the frequency of the lowest normal mode is 87.2 hertz.
The next normal mode, when \( n = 3 \), has a frequency of 262 hertz, which is about middle C on a piano. Where would you have to place your ear in the tube in order to hear silence when you play middle C on your speaker? You’d just have to listen at a node, which happens every half-wavelength from the closed end of the tube. When \( n = 3 \), your wavelength, \( \lambda_3 \), is given by

\[
\lambda_3 = \frac{4}{3} (0.983 \text{ m}) = 1.31 \text{ m}
\]

So you’d have to place your ear half this distance from the closed end of the tube — that is, 0.655 meters. For this normal mode, there are no other nodes in the tube (except, of course, the one at the closed end of the tube), so this is the only place you’d get silence.

It turns out that any possible vibration of sound in the closed-tube-and-diaphragm setup is simply an interference of normal modes! So even the craziest, most complicated, erratic vibration can be boiled down to a matter of how much of each normal mode you have. This understanding comes from some extremely powerful ideas in mathematics that have pervaded physics. For example, in quantum mechanics, particles (like the electron) are allowed to be only in certain particular states. These states are like the normal modes of your tube. Like your tube, the particles can be in a state that’s an interference of these normal modes — but when you actually measure the state of the electron, say, you can only ever see it in one of the normal modes! This is just a hint of some of the quantum weirdness that lies ahead for you in physics.

**Reaching resonance frequency: The highest amplitude**

You can drive things at a frequency that maximizes the amplitude of vibration. For instance, consider the sound vibration in the speaker-and-tube setup, which is driven by the speaker. As you increase the frequency of the sound wave from the speaker, you find that the amplitude of the sound vibration peaks whenever the speaker drives at one of the harmonics — that is, one of the frequencies of the normal modes. So the tube has an infinite number of resonance frequencies, given by \( f_n \), where \( n \) is an odd number.

It’s testament to the power of the ideas in physics that resonance is also a feature of the RLC electrical circuit that you look at in Chapter 5. You see that in the RLC circuit, there’s likewise a natural frequency at which the current in the circuit oscillates with the greatest amplitude.

**Getting beats from waves of slightly different frequencies**

Anyone who has tuned a guitar has heard the effect of simultaneously playing two very slightly different notes — the sound seems to oscillate in loudness. These oscillations are called *beats*.

Figure 7-10a shows a graph of two waves of slightly different frequencies, and Figure 7-10b shows their sum, which is the wave of interference between the
two. You can see that the interfering wave oscillates with a frequency similar to the frequency of the original two waves, but the amplitude increases and decreases with another frequency, the beat frequency. The beat frequency is simply the difference between the frequencies of the original waves.

\[ t \]

\[ (a) \]

\[ (b) \]

Figure 7-10: The beats formed from two waves with slightly different frequencies.

**Bending rules: Sound wave diffraction**

You can hear a police car approaching with its siren wailing even if it’s around a corner, hidden by tall buildings. And you can talk to a person in another room through an open door, even if you can’t see that person through the doorway. Sound waves travel in straight lines, but when they hit a boundary like the edge of a wall or a lamppost, the sound waves bend around it — this behavior is diffraction.

Diffraction happens in all waves, including sound. Figure 7-11 shows a sound wave approaching two gaps in a wall — the lines represent the position of the wave peaks. One gap is much wider than the wavelength of the sound, and the other is similar in size to the wavelength. You see that as the sound wave travels through the wider gap, it mostly goes straight through, with some bending at each edge. But when the sound goes through the gap that’s of similar width to the wavelength of the sound, the wave spreads over a wider angle. This bending of the wave, to where it wouldn’t go if it traveled in a straight line, is diffraction.

The wider angle you get when a wave spreads by diffraction explains why you can hear around corners but you can’t see around them. Light has a much shorter wavelength than sound, so if you were to shine a light through the gaps in Figure 7-11, both gaps would be much wider than the wavelength of the light; therefore, you’d see hardly any spreading of the light wave — not enough for you to notice, anyway.
Diffraction is really just a manifestation of interference (as I show you in Chapter 10, on light). People use two different terms for essentially the same thing, but the difference is that interference is usually understood to mean the interaction of just a few waves, whereas diffraction is the interference of a very great number of waves.

**Coming and going with the Doppler effect**

The *Doppler effect*, named after Christian Doppler, says that a sound wave’s frequency changes if the source of the sound is moving (or you’re moving toward or away from the source). If you and the source of the sound are getting closer, you hear the sound at a higher pitch. And if you and the source are getting farther apart, you hear the sound at a lower pitch.

For instance, consider a police car with its wailing siren. Because you’re law-abiding, it passes right by you. What do you hear? You’re familiar with the high-pitched *ne-naw-ne-naw* as the car travels toward you, turning into a low-pitched version of the same sound after it has passed and is traveling away. You can understand this effect using the picture of sound as a wave.

**Moving toward the source of the sound**

First consider what happens when the source of sound is stationary but you’re moving toward it. You can see this situation in Figure 7-12a. The source produces a wave with wavelength $\lambda_s$ and frequency $f_s$, and this wave travels with the speed of sound $v$. You walk toward the source with speed $v_s$. 
If you were to remain stationary, your ear would experience a sound wave of frequency $f_s$, which is the frequency of the source. But if you walk toward the source, then your ear experiences a sound wave with a higher frequency. You’re walking into the wave, so each wave peak has to travel a slightly smaller distance to reach you than it would if you were to remain still.

As you move toward the source of the sound, the speed of the wave as it appears to you is $v + v_a$. So when a wave peak is at your ear, the time for the next wave peak to reach you is $\lambda_s/(v + v_a)$ seconds. Therefore, the frequency you hear, $f_a$, is given by

$$f_a = \frac{v + v_a}{\lambda_s}$$

Because $v = \lambda_s f_s$, you can write this as

$$f_a = \frac{v + v_a}{v} f_s = \left(1 + \frac{v_a}{v}\right) f_s$$

So you see that the frequency you hear is a factor of $(1 + v_a/v)$ greater than the frequency from the source.
**Having the source of the sound move**

When the source of a sound moves, the speed of the sound waves remains the same, because the speed of sound is determined only by the air — it has nothing to do with the source. So if the source moves away from you with a speed \( v_b \), as in Figure 7-12b, \( v \) is still the same. But what changes is the wavelength of the sound waves.

To see why the wavelength changes, think about how the waves propagate. Say the source emits a wave peak, which propagates behind the source. In the time before the next peak, the source moves a distance \( v_b T_s \), where \( T_s \) is the period of the source waves (\( T_s = 1/f_s \)). So the wavelength behind the source is

\[
\lambda_s = \frac{v}{f_s} + \frac{v_b}{f_s} = \frac{v + v_b}{f_s}
\]

The wavelength of the wave behind the source is now enlarged. Now plug this new wavelength into the equation relating the frequency that you hear to the frequency of the source (from the preceding section). Replace \( \lambda_s \) with the new enlarged wavelength to find

\[
f_a = \frac{v + v_b}{\lambda} = \frac{v + v_a}{v + v_b} f_s
\]

This is the frequency you hear when you travel toward a moving source of sound, which is in front of you. The source is moving in a particular direction with speed \( v_b \), and you’re behind it traveling in the same direction with speed \( v_a \).

If you’re in front of the source, then you’re in the region where the waves from the source have a shorter wavelength, and you’re walking away from the approaching peaks. The frequency you hear is given by

\[
f_a = \frac{v - v_a}{\lambda} = \frac{v - v_a}{v - v_b} f_s
\]

**Doing the math on the Doppler effect**

Now put some numbers into a police siren example. Suppose that the police car passes very close to you. It’s initially traveling pretty much straight toward you, and after it passes, it travels pretty much straight away from you.

A police siren makes a sound that has a frequency of about 320 hertz. If the car has its siren on, it must be in a hurry, so say it’s going at about 70 miles per hour (31.29 meters per second). Also, you’re walking along the sidewalk
at a speed of 1.5 meters per second. The actual frequency of the sound you hear is

\[ f_a = \left( \frac{343 \text{ m/s} - 1.5 \text{ m/s}}{343 \text{ m/s} - 31.29 \text{ m/s}} \right) 320 \text{ Hz} \]

\[ = \left( \frac{341.5}{311.71} \right) 320 \text{ Hz} \]

\[ = (1.0956) 320 \text{ Hz} \]

\[ = 351 \text{ Hz} \]

This is about a 10 percent increase on the actual frequency of the siren. Now work out the frequency you hear when the police car has passed and is traveling away:

\[ f_a = \left( \frac{343 \text{ m/s} + 1.5 \text{ m/s}}{343 \text{ m/s} + 31.29 \text{ m/s}} \right) 320 \text{ Hz} \]

\[ = \left( \frac{344.5}{374.29} \right) 320 \text{ Hz} \]

\[ = (0.9204) 320 \text{ Hz} \]

\[ = 295 \text{ Hz} \]

This is about an 8 percent decrease on the actual frequency of the siren.

If you have a piano handy, you can play these sounds. The original police siren is roughly the same tone as the E, which is two whole notes above middle C. Then the sound that you hear as the car travels toward you is about the same as playing one note higher, and the sound when the car has passed is about the same as one note lower.

**Breaking the sound barrier: Shock waves**

Sound moves pretty quickly in air, but some things move faster than sound. When Concorde (the French and British supersonic passenger jet) was flying before its retirement in 2003, you could travel across the Atlantic Ocean at about twice the speed of sound. Meteors entering the Earth’s atmosphere travel through the air much faster than this. Objects breaking the sound barrier produce a *sonic boom*, a loud sound that people can hear from the ground. In this section, I discuss what happens when something breaks the sound barrier.

**Producing shock waves**

Because of the Doppler effect, the wavelength of the sound produced by a moving source is stretched behind it and shortened in front (see the earlier section “Coming and going with the Doppler effect” for details). When an
airplane (or another moving object) approaches the speed of sound, it has to work extra hard to compress the air in front of it as it bunches up all the wave peaks; this extra work gave rise to the term *sound barrier*.

Figure 7-13 shows the wave peaks of a source of sound moving faster than the speed of sound (that is, the source is in *supersonic motion*). The wave peaks spread uniformly away from where the source was when it emitted them. At any one time, this makes a series of circles whose centers are evenly spaced along the path of the source (assuming it’s moving at a constant speed), and the radii of the earliest circles are uniformly greater than the most recent ones. The edge of these circles forms a line of constructively interfering waves along their outer edge, which is called a *shock wave*.

The concentration of constructively interfering sound waves along the shock wave makes a very loud sound. Any listener who happens to be at a point on the shock wave hears this sound — the *sonic boom*.

Note that as the plane travels through the air, the shock wave is actually a cone; the figure shows only a cross section. As the plane travels faster than sound and produces a shock wave, the air around the tip of the plane is at a higher pressure than the surrounding air. The speed of sound can vary as the pressure of the air varies, which means that the pressure variation around the tip of the plane causes the shape of the shock wave to curve slightly in this region instead of the straight lines you see in Figure 7-13.

![Figure 7-13: The shock wave of an airplane.](image)

**Finding the angle of a shock wave**

With basic trigonometry, you can easily work out a good approximation of the angle a shock wave makes from the direction of travel. Look at the right triangle in Figure 7-13. Since the source (an airplane) emitted a sound wave at point *a*, the source has been traveling for a time *t* at speed *vₐ*. So the length
of the hypotenuse of this triangle is \( vt \), and the length of the opposite side of the triangle is just the distance that the sound wave has traveled, \( vt \). The sine of the angle of the shock wave, \( \alpha \), is then just the ratio

\[
\sin(\alpha) = \frac{vt}{v_a t} = \frac{v}{v_a}
\]

You may hear that a jet travels at such-and-such a Mach number, such as Mach 3.3 (the SR 71 Blackbird) or Mach 9.6 (NASA’s X-43A). The **Mach number** is just the speed of the jet compared to the speed of sound, \( v/v_a \).

Check out an example — what would be the angle of the shock wave that Concorde would’ve made as it traveled across the Atlantic at twice the speed of sound? This means that the Mach number is 2.0; hence,

\[
\sin(\alpha) = \frac{v}{v_a} = \frac{1.0}{2.0} = 0.50
\]

If you take the inverse sine of this, you find the angle is 30°.
Cracking the secret of light was a major advance for both scientists and the general population. Now physicists know what creates light waves. They can even predict how fast light waves go, how much energy they transfer from Point A to Point B, and more. Consider this chapter to be your guided tour of the nature of light. I’m your friendly guide (minus the name badge), and I start things off by covering what light really is. You then dive into topics such as the electromagnetic spectrum, light intensity, and more.

Let There Be Light! Generating and Receiving Electromagnetic Waves

The big name in the discovery of how light works is James Clerk Maxwell. He’s the lucky physicist who first figured out that light is nothing more than alternating electric and magnetic fields that regenerate each other as light travels, allowing it to keep going forever.

In this section, I explain how electricity and magnetism combine to create electromagnetic waves. I also note how radio receivers work by catching either the electric or the magnetic field of those waves.
**Creating an alternating electric field**

The process of generating an alternating electric field (known as an *E field*) starts with an oscillating charge. To create an oscillating charge, you can connect an alternating voltage source to the top and bottom of a wire. Figure 8-1 shows this setup at four consecutive times. The alternating voltage source causes the electrons in the wire to race up and down its length, creating an alternating electric field in the wire.

Electric fields propagate through space, so as the electric field you created goes up and down the wire, that same electric field moves out into space as well. Because the electric field in the wire is constantly changing directions as the voltage source alternates, you get an alternating electric field in the wire, which leads to an alternating electric field propagating through space, as you see in Figure 8-1.

At first, the electric field starts off small (see Figure 8-1a). Consequently, the electric field that leaves the wire and propagates through space is also small. In time, however, the electric field in the wire becomes larger (Figure 8-1b), and the propagated electric field does the same.
Then, as the voltage source alternates, the electric field begins to switch directions in the wire. The propagated electric field follows, growing smaller even though it’s still pointing in the same direction. At some later time, the voltage across the wire changes polarity completely (the *polarity* of the potential difference between two points just describes which point is of higher potential and which is lower) — and the electric field in the wire changes polarity, too, as you can see in Figure 8-1c. Not surprisingly, the direction of the electric field that’s propagated into space also changes.

As you can see in Figure 8-1d, as the oscillating charge completes its alternating cycle, the wave in the electric field completes its cycle, too.

Note how the oscillating electric field in Figure 8-1 always points in a direction that’s perpendicular to the direction of propagation. The wave propagates to the right, with the electric field always in the vertical orientation — that is, it oscillates up and down. When the electric field oscillates with a constant orientation as the wave propagates, the wave is *linearly polarized*. So you can say that the wave propagating in Figure 8-1 is linearly polarized with the electric field in the vertical orientation. (Why not bring it up at your next party?)

**Getting an alternating magnetic field to match**

How exactly do you pair an alternating electric (*E*) field (see the preceding section) with the magnetic (*B*) field that’s supposed to be the other half of a light wave? Are you supposed to have a spinning bar magnet or something?

Actually, creating the matching *B* field is easier than it seems. In fact, if you have a straight wire where the voltage (and hence the current) alternates up and down, you’ve already done it, because current in wires creates a magnetic field.

Here’s how applying an alternating voltage across a wire creates your matching *E* and *B* fields:

- When you have *voltage* alternating up and down a wire, you create an oscillating *E* field.
- That *E* field causes *current* to race up and down in the wire, and the current generates an alternating *B* field.

Notice the direction of the created *B* field in Figure 8-2. The created *E* field is parallel to the wire (following the electrons as they race madly up and down the wire), but the *B* field is perpendicular to the *E* field. That’s because the *B* field is created perpendicular to the wire.
Putting all this together means that an alternating voltage source applied across a wire creates an alternating $E$ field and an alternating $B$ field, both of which propagate away from the wire, as Figure 8-3a shows. Note that the $E$ and $B$ fields are perpendicular to each other. And both of them, the $E$ field and the $B$ field, are perpendicular to the direction of propagation — you have all three dimensions covered.

When you know the directions of the $E$ and $B$ fields, just follow one of the right-hand rules to find the direction of propagation:

- If you put the fingers of your right hand in the direction of the $E$ field and then bend them toward the $B$ field using the shortest possible arc, then your thumb will point in the direction of propagation.

- Hold out your palm, pointing your fingers in the direction of the electric field and your thumb in the direction of the magnetic field. Then the direction of propagation of the wave is the direction your palm is facing. Figure 8-3b shows this version.

*Electromagnetic waves* are just propagating fluctuations of the electric and magnetic fields. Electromagnetic waves of the lowest frequencies — like the ones from a wire connected to an alternating voltage source — are *radio waves*. A higher range of frequency of electromagnetic waves is even more
familiar: light. That’s right — light and radio are essentially the same thing; the only difference is that your eyes are sensitive to the frequencies of visible light waves.

The wire in my example, believe it or not, is actually an antenna. Perhaps you’ve seen radio towers that soar up to great heights. At their core, they rely on a single wire with an alternating voltage placed across that wire from top to bottom. The wire, by having charges race up and down its length, creates radio waves.

Can you generate visible light with a wire in the same way that you can generate radio waves? Probably not. No alternating voltage source in the world oscillates fast enough to approach the frequencies of visible light. For the scoop on visible light and other parts of the electromagnetic spectrum, check out the later section “Looking at Rainbows: Understanding the Electromagnetic Spectrum.”

**Receiving radio waves**

*Creating* radio waves (see the preceding sections) is only half the story; you still need a way of receiving them. That’s where receiving antennas come in.

As I show you in Figure 8-3, the electric and magnetic fields of a radio wave are perpendicular to each other — there, the $E$ field moves vertically and the $B$ field moves horizontally. Vertical antennas and loop antennas are the two primary ways of receiving radio waves, and they correspond to the $E$ and $B$ field parts of radio waves, respectively.
**Vertical antennas: Catching the E field**

To detect a vertically moving electric field from a sending antenna, you simply use a vertical receiving antenna, which is really just a long wire.

The electric field (E field) from the sending antenna is in the vertical plane, just like the antenna itself, because the E field follows the movement of the electrons in the wire (see the earlier section “Creating an alternating electric field”). When you use a vertical receiving antenna, the E-field component of the radio wave makes the electrons in the receiving antenna race up and down. When the antenna receives the E field, a tiny voltage appears from the top to the bottom of the receiving antenna. Your radio can then amplify that voltage until it becomes a signal that lets you make out words and music.

**Loop antennas: Catching the B field**

To receive a horizontally moving magnetic field (B field) from a sending antenna, you can use a wire loop or coil. *(Note: Receiving radio antennas use a combination of both loops and coils.)* First, set up the loop or coil in the vertical plane to maximize the magnetic flux through it (I cover magnetic flux in detail in Chapter 5). If that sounds counterintuitive to you, consider this: The magnetic field you’re trying to detect is in the horizontal plane, so setting up a loop or coil of wire vertically allows you to make as much of that magnetic field as possible go through your antenna.

So the rapidly oscillating B field is oscillating in your loop or coil of wire. That’s great, but how do you actually measure that B field? A changing magnetic flux in a loop or coil of wire induces a current in that loop or coil in a way that counteracts the applied magnetic field from the radio station. Your radio is able to measure that tiny current and decipher it, just as other radios can decipher the tiny voltages created by the radio station’s electric field.

---

**Making radio waves a hit**

Physicist Heinrich Hertz was the one who first generated and received radio waves in his laboratory in 1886. This was a breakthrough for physics, but he wasn’t sure how to put these waves into use.

Guglielmo Marconi, an Italian physicist, was one of many people who set out to use these new waves to communicate over great distances almost instantaneously. He patented a version of the telegraph that marked one of the first practical advances in “wireless” communication.

Early in radio’s development, radio waves were detected over distances of about a mile. But physicists soon noted that the more charge racing up and down the antenna, the greater the amplitude of the wave and therefore the greater the distance from which it could be received. As transmitters and receivers advanced in technology, the distance increased to hundreds and then thousands of miles.
With a loop antenna, your radio decodes the current that flows through the loop due to the fluctuating magnetic flux. With a straight antenna, your radio decodes the tiny voltages that appear across the wire from the electric field component of the wave.

Looking at Rainbows: Understanding the Electromagnetic Spectrum

Electromagnetic waves have the same general properties shared by all waves — wavelength, frequency, and speed (see Chapter 6 for details). In this section, you see how those properties apply to light waves. You also find out how the continuous range of frequencies is divided up into different wave types within the electromagnetic spectrum.

Perusing the electromagnetic spectrum

Even though all electromagnetic waves are essentially the same — differing only in frequency — they differ in how they interact with matter. For example, the waves with a particular range of frequencies are visible as light, whereas other waves at a higher frequency are invisible but can give you a nasty sunburn. You see this variation because matter is made up of charged particles (electrons and protons) in various configurations, and the way that these particles interact with electromagnetic waves depends on the details of this configuration.

Different frequencies of electromagnetic waves correspond to different parts of the electromagnetic spectrum — that is, the range of all electromagnetic waves, arranged in increasing frequency. Most divisions of the spectrum are made according to how the different parts of the spectrum interact with matter, but the division is sometimes based on how the wave is produced or used.

People sometimes debate which wavelengths go in which category, but Figure 8-4 can give you approximate ranges of the main divisions of the electromagnetic spectrum, with labels for the names of the electromagnetic waves within them.
Starting at the lower frequencies, here are the types of electromagnetic waves in order:

- **Radio waves**: As you can see in Figure 8-4, radio waves include the familiar AM and FM bands. The radio band’s AM frequencies are in the $10^6$ hertz (Hz) region, and the FM frequencies are in the $10^8$ Hz region. Radio waves have long wavelengths and are generally produced with antennas (see the earlier section “Let There Be Light! Generating and Receiving Electromagnetic Waves” for details).

- **Microwaves**: When the frequency of radio waves increases to the point where the wavelength is about the same size of the electrical circuits used to make them, the wave can have a feedback effect on the circuit. The methods of generating waves of this frequency have to take this into account, so these waves have a special name: **microwaves**. Some liquids consist of molecules that absorb microwaves and become heated, which microwave ovens take advantage of.

- **Infrared light**: This kind of light is invisible to the naked eye. Humans have to wear night-vision goggles to pick up this part of the spectrum.

- **Visible light**: The light you can see is actually a very narrow band of the spectrum that exists solely in the $4.0 \times 10^{14}$ Hz to $7.9 \times 10^{14}$ Hz region (this is one of the few frequency ranges pretty much everyone agrees on). The lowest-frequency end of this part of the spectrum corresponds to the red end of the rainbow, and the highest-frequency end corresponds to the violet part of the rainbow. The rest of the rainbow is distributed within this range.

  Why is visible light restricted to such a narrow range? One answer is that much of the rest of the light spectrum is absorbed by water and water vapor — both of which are plentiful on Earth. Infrared light, for example, is absorbed by water vapor, making it an unfavorable option to rely on for your vision.

- **Ultraviolet light**: Higher in the spectrum still, you have ultraviolet light, where the frequency is higher and the wavelength is shorter. This is the region of the so-called black lights that make phosphorescent paints glow. These are also the waves responsible for sunburn.
X-rays: This part of the light spectrum travels easily through the human body, which is why X-rays play such a starring role in medicine to check for broken bones.

Gamma rays: These high-energy rays are created by high-power transitions in the atomic nucleus (as opposed to other kinds of electromagnetic waves, which mostly come from transitions in the electron structure of an atom).

Relating the frequency and wavelength of light

Because light is made up of electromagnetic waves, it must obey the general wave equations (see Chapter 6). In particular, you can relate the frequency \( f \) of a wave to its wavelength \( \lambda \) to find its speed \( v \) like this:

\[
  v = f\lambda
\]

In a vacuum, light travels at the speed \( c \), which is about equal to \( 3.0 \times 10^8 \) m/s (I explain where this number comes from in the next section). So for a vacuum (or air), you can say the following:

\[
  c = f\lambda
\]

So using this formula, what’s the wavelength of red light if its frequency is \( 4.0 \times 10^{14} \) Hz? And at the other end of the visible spectrum, what’s the wavelength of violet light (whose frequency is \( 7.9 \times 10^{14} \) Hz)? You know that \( c = f\lambda \), so the wavelength formula is

\[
  \lambda = \frac{c}{f}
\]

Plugging in the numbers and doing the math for the red-light question gives you this:

\[
  \lambda = \frac{3.0 \times 10^8 \text{ m/s}}{4.0 \times 10^{14} \text{ Hz}} = 7.5 \times 10^{-7} \text{ m}
\]

Now take a look at the calculations for violet light, where the frequency is \( 7.9 \times 10^{14} \) Hz:

\[
  \lambda = \frac{3.0 \times 10^8 \text{ m/s}}{7.9 \times 10^{14} \text{ Hz}} = 3.8 \times 10^{-7} \text{ m}
\]
The nanometer (abbreviated nm), or $10^{-9} \text{ m}$, is often used for wavelengths in the visible region. So you can say the two wavelengths are 750 nanometers and 380 nanometers. What do these numbers actually mean? Turns out that’s up to your eye.

Red light is the longest wavelength your eye can perceive, and 750 nanometers is the longest of the red wavelengths most eyes can see. Violet is the shortest wavelength of light you can see, and 380 nanometers is the shortest of the violet wavelengths your eye can normally pick up. So in between 380 and 750 nanometers — a very short range — lie all the glorious colors of the light spectrum that are visible to the human eye.

**See Ya Later, Alligator: Finding the Top Speed of Light**

Light is fast — nothing can travel faster, *Star Trek* and *Star Wars* gadgetry included, unfortunately. The speed of light in a vacuum is approximately $3.0 \times 10^8 \text{ meters per second}$, or $3.0 \times 10^{10} \text{ centimeters per second}$, or about 186,000 miles per second. (If you’re a stickler for accuracy, try the value 299,792,458 meters per second.)

The distance around the world is about 40,000 kilometers, or $4.0 \times 10^7 \text{ meters}$, so at the speed of light, you could make 7.5 trips around the world in 1 second ($3.0 \times 10^8 \text{ m/s} \div 4.0 \times 10^7 \text{ m} = 7.5 \text{ trips/second}$). You could even go to the moon in that amount of time. So although light is fast, it’s not infinite.

---

**A not-so-illuminating light experiment**

There was a time, of course, when people had no idea how fast light was. Many experiments were tried, and many failed (utterly). Case in point: In a touching show of confidence, two scientists synchronized their pocket watches to within a second and then trooped to opposite ends of a mile-long field. At just the agreed-upon moment, the first scientist opened a lantern. The problem was that from the second scientist’s point of view, as soon as his watch showed the correct time, the beam of light from the first scientist was already shining full force. Neither scientist could believe anything could be so fast; each one thought his watch must’ve been off!

Of course, given the speed of human reflexes, the two scientists could’ve been standing 100,000 miles apart, and the beam of light would’ve arrived within less than a second — that is, less than the accuracy of the watches and the scientists’ ability to open their respective lanterns.
In this section, you discover how physicists figured out how fast light travels in a vacuum. Of course, as with other waves, the speed of light depends on the medium it’s traveling through, if any. I touch on light as it travels through materials such as diamond and glass in Chapter 9.

**Checking out the first speed-of-light experiment that actually worked**

Many people attempted to measure the speed of light, often relying on astronomical phenomena. Armand Fizeau and Léon Foucault were the first to make Earth-bound measurements of the speed of light. Foucault’s method used a rotating mirror to improve upon the space-based estimates.

Albert Michelson, an American who adapted and improved upon Foucault’s method, measured the speed of light in 1926 — and dramatically increased the accuracy of the measurements.

**Setting up the experiment**

Michelson’s apparatus was pretty clever; it involved bouncing light off a mirror 35 kilometers away. However, because light makes the 70-kilometer round trip in about a ten-thousandth of a second, Michelson needed to do more than just bounce light off a mirror some distance away.

His solution made him famous, and you can see a depiction of it in Figure 8-5. To accurately capture the speed of light, Michelson determined that in addition to bouncing off a mirror 35 kilometers away, light had to hit a rotating, eight-sided mirror just right. Specifically, the light needed to bounce off one side of the eight-sided mirror, make a round trip of 70 kilometers, and then hit another part of the eight-sided mirror just right to get into the detector. If the mirror rotated too much or too little, the side that the light signal was meant to bounce off of into the detector just wouldn’t be there (in other words, it wouldn’t have reached its proper position yet). Because Michelson could regulate how fast the mirror rotated, which was pretty darn fast, he was able to make the window for light to hit the eight-sided mirror very small. Pretty clever, eh?

In the 1926 round of experiments, Michelson determined the speed of light to be 299,796 kilometers per second, plus or minus 4 kilometers per second. (However, $3.0 \times 10^8$ meters per second is sufficiently accurate for the calculations in this book.)

**Finding the mirror speed**

Try calculating how quickly Michelson’s mirror must’ve been rotating to capture the speed of light. Say you’re working with the setup in Figure 8-5
and want to reflect a light beam off a mirror 35 kilometers away (that’s a round trip of 70 kilometers). The key to calculating the speed of the mirror’s rotations is to realize that the shortest time the experiment can measure is the amount of time it takes for the eight-sided mirror to make one-eighth of a revolution. That’s the shortest time the light can take to bounce off the far mirror, return, and still enter the detector.

So in the time it takes light to go 70 kilometers, your eight-sided mirror makes one-eighth of a revolution. How fast does the eight-sided mirror turn? First, figure out the amount of time light needs to go 70 kilometers. You already know that

\[
\text{Speed} = \frac{\text{distance}}{\text{time}}
\]

So this equation must also be true:

\[
\text{Time} = \frac{\text{distance}}{\text{speed}}
\]

To travel 35 kilometers to the mirror and 35 kilometers back at the speed of light, you need this much time:

\[
\text{Time} = \frac{2 \left(3.5 \times 10^4 \text{ m}\right)}{3.0 \times 10^8 \text{ m/s}} = 2.3 \times 10^{-4} \text{ s}
\]
That means your eight-sided mirror must make one-eighth of a turn in \(2.3 \times 10^{-4}\) seconds, giving it an angular speed of

\[
\omega = \frac{\frac{1}{8} \text{ revolution}}{2.3 \times 10^{-4} \text{ s}} = 540 \text{ revolutions/s}
\]

So your eight-sided mirror needs to be turning at 540 revolutions per second in order to accurately measure the speed of light.

## Calculating the speed of light theoretically

As James Clark Maxwell discovered, the astonishing fact is that absolutely every property of electric and magnetic fields — every aspect of their behavior — is contained in just four equations. Most of the math goes beyond trig, so you can skip the actual equations for now, but here’s a preview of what they cover (see Chapters 4 and 5 for more on electric and magnetic fields):

- Faraday’s law describes the electric field that comes from a changing magnetic field.
- Ampère’s law describes the magnetic field that results from a current and a changing electric field.
- A third equation simply expresses the fact that there are no magnetic monopoles, so magnetic field lines are all loops.
- Gauss’s law describes the flux of the electric field in terms of the electric charge. (For uniform fields, the electric flux of a field through an area is simply the size of the area multiplied by the size of the component of the field that’s perpendicular to the area.)

Maxwell summarized and organized all the laws of electricity and magnetism because he was trying to solve a great puzzle. Before Maxwell came along, electric charge was thought to be divided — there was one form of charge for static electric fields and another for magnetic fields. But it turned out that if one of these units of charge was divided by the other, then the answer was equal to the speed of light! This was thought to be an incredible coincidence. But Maxwell resolved the puzzle purely by thinking about it and what was known about electricity and magnetism, and in doing so, he revealed the true nature of light.

Maxwell brought together the equations governing electric and magnetic fields and showed that one of their solutions was to have a wave. The real
thunderbolt came when Maxwell showed that these waves must travel at the speed of light. It then didn’t take long to realize that this was no coincidence — the waves were light!

I’m not embarrassed to admit it — I think the theoretical calculation of the speed of light is one of the most spectacular results physics has ever had. And it’s right on. As you know, light is made up of electromagnetic waves. To start calculating the speed of light with that information, you first need to examine the values typically involved with both electric and magnetic fields. The size of the force between two charges, for example, is this:

\[ F = \frac{kq_1q_2}{r^2} \]

That’s actually the modern shorthand version of the following equation:

\[ F = \frac{q_1q_2}{4\pi \varepsilon_0 r^2} \]

where \( \varepsilon_0 \), which equals \( 8.85 \times 10^{-12} \text{ C}^2/(\text{N}\cdot\text{m}^2) \), is a constant called the electric permittivity of free space, a measure of how easily an electric field passes through free space. (Sounds promising for finding the speed of light, doesn’t it?)

Similarly, magnetic fields often involve the constant \( \mu_0 \), the so-called magnetic permeability of free space (again, sounds like something you’d want to include in a calculation of the speed of light through free space). So the force between two current-carrying wires is

\[ F = \frac{\mu_0 I^2 L}{2\pi r} \]

And \( \mu_0 = 4\pi \times 10^{-7} \text{ T}\cdot\text{m}/\text{A} \).

So how do \( \mu_0 \) and \( \varepsilon_0 \) connect to the speed of light? Well, Maxwell derived some famous equations describing how light works, and here’s the payoff: He was actually able to derive the speed of light like this:

\[ c = \frac{1}{\left( \mu_0 \varepsilon_0 \right)^{1/2}} \]

No, your eyes aren’t deceiving you. Maxwell was indeed able to calculate the speed of light by simply connecting it to the two fundamental constants of electric and magnetic fields. This relation is exact, as determined by the laws of electric and magnetic fields. Now if that doesn’t get you excited, I don’t know what will!
Chapter 8: Seeing the Light: When Electricity and Magnetism Combine

You’ve Got the Power: Determining the Energy Density of Light

Like water waves (which I touch on in Chapter 6), electromagnetic waves can carry energy. If they didn’t, everyone would be in trouble, because energy from the sun would never reach Earth, meaning there wouldn’t be any solar power, oil, plant life, or warmth. Not a pretty picture, huh?

To get an idea of how much energy an electromagnetic wave carries, you have to look at the wave’s energy density, the amount of energy that wave carries per cubic meter. The units for an electromagnetic wave’s energy density are joules per cubic meter.

Why not just find total energy? Well, when you’re talking about a light source like the sun, you don’t just switch it on and off, so you can’t really think about it in terms of total energy. This section explains how you find energy density.

Finding instantaneous energy

Light waves are electromagnetic waves, so you can reasonably assume that the energy in a light wave comes from its electric and magnetic components. In an electromagnetic wave, you have an electric field and a magnetic field, which are changing with time (see the earlier section “Let There Be Light!"

Light plus friction equals hot soup

Microwave ovens provide an excellent example of how electromagnetic waves can transfer energy. Here, I use the detailed picture of the physics of the electromagnetic wave to heat a bowl of soup.

Although water molecules have no charge overall, each molecule has a positive end and a negative end because the electrons in the molecule aren’t evenly distributed. Therefore, you say that the water molecule is an electric dipole.

When you place the polar water molecule in an electromagnetic wave, the molecule tries to align itself with the alternating electric field. This causes the molecule to rotate back and forth as the electric field oscillates. The motion causes the water molecule to push and pull on neighboring molecules, causing them to move and vibrate — and this greater vibration of the molecules is exactly what it means for something to be at a higher temperature. The frequency of the waves in a microwave oven transfers energy to water molecules at a rate that’s good for cooking: $2.45 \times 10^9$ Hz.
Generating and Receiving Electromagnetic Waves" for details). The energy in these fields is spread through the space that they occupy.

In this section, I show you how to work out the density of the energy stored in these fields at any point and any time. This gives you the foundation to figure out how much power is in an electromagnetic wave, as you see later in “Averaging light’s energy density.” When you know that, you can figure out things like how much energy Earth’s equator receives from the sun.

**Looking at electric energy density**

Obviously, you need energy to set up an electric field in space. For example, to charge a capacitor (see Chapter 4), which stores energy, you have to do work to put the charges on each plate. After you’ve charged it, the work you did isn’t lost — it’s stored in the electric field between the plates. Because this field is uniform, the energy stored in it is uniformly distributed throughout the space between the two plates.

If you work out how much work you did to charge the capacitor and then divide it by the volume of the space between the plates, you have an expression for the energy density of the electric field. It turns out to be the following:

$$\text{Electric energy density} = \frac{\varepsilon_0 E^2}{2}$$

where $E$ is the magnitude (strength) of the electric field and $\varepsilon_0$ is a constant equal to $8.85 \times 10^{-12}$ C$^2$/N-m$^2$.

Actually, that’s the amount of energy density you need to set up an electric field from any source — a parallel plate capacitor or light waves. Wherever the electric field has magnitude $E$, the energy density at that point is given by the preceding equation. So now you know one component of the total energy density in an electromagnetic wave: the energy density of an electric field.

**Considering magnetic energy density**

You can make a uniform magnetic field in a solenoid by setting up a current in the loops of wire (see Chapter 4). Setting up the current takes work, and this work is stored in the magnetic field inside the solenoid.

You can work out how much work you did to set up the uniform magnetic field and divide it by the volume of the space that it occupies to find the density of the energy stored in this field. The answer turns out to be the following:

$$\text{Magnetic energy density} = \frac{B^2}{2\mu_0}$$
where $B$ is the magnitude (strength) of the magnetic field and $\mu_0$ is a constant equal to $4\pi \times 10^{-7}$ T·m/A.

Guess what — that’s exactly how much energy density (energy per cubic meter) you need to set up a magnetic field from any source, whether that’s wires in a solenoid or an electromagnetic wave. So at any point, where the magnetic field strength is $B$, the density of the energy stored in the magnetic field there is given by the preceding relation.

**Adding the energy densities together**

Because light is made up of an electric field component and a magnetic field component, the total energy density of an electromagnetic wave is simply the sum of the two energy densities. The equation for total energy density ($u$) looks like this:

$$u = \frac{\varepsilon_0 E^2}{2} + \frac{B^2}{2\mu_0}$$

That’s the total energy density, $u$, of an electromagnetic field (electric and magnetic fields together) per cubic meter. You can use this expression to work out the energy density at every point and time of the fluctuating fields of an electromagnetic wave.

Now you’re making progress! So consider this question: How does nature decide which component of an electromagnetic wave to put more energy into — the electric component or the magnetic component? Turns out both components have equal energy. That is, the electric energy component is equal to the magnetic energy component, which means you can say the following:

$$\frac{\varepsilon_0 E^2}{2} = \frac{B^2}{2\mu_0}$$

That’s interesting, because using that equation, along with the formula for the speed of light (from the earlier section “Calculating the speed of light theoretically”), you can do some pretty slick algebra. First, isolate $E$ on one side of the equation:

$$\frac{\varepsilon_0 E^2}{2} = \frac{B^2}{2\mu_0}$$

$$E^2 = \frac{2B^2}{2\mu_0\varepsilon_0}$$

$$E = \frac{B}{(\mu_0\varepsilon_0)^{1/2}}$$
Now, because the speed of light in a vacuum is $c = \frac{1}{\left(\mu_0\varepsilon_0\right)^{1/2}}$, you can just plug $c$ into the equation:

$$E = \frac{B}{\left(\mu_0\varepsilon_0\right)^{1/2}}$$

$$E = cB$$

So the magnitude of the electric component in a light wave is connected to the magnitude of the magnetic component by a factor of $c$. Well, because

$$\frac{\varepsilon_0 E^2}{2} = \frac{B^2}{2\mu_0}$$

and

$$u = \frac{\varepsilon_0 E^2}{2} + \frac{B^2}{2\mu_0},$$

you finally get the following equation for the energy density:

$$u = \varepsilon_0 E^2$$

Or equivalently,

$$u = \frac{B^2}{\mu_0}$$

Nice work! You’ve found the energy density at every point of an electromagnetic wave in terms of the strength of the electric and magnetic fields there.

**Averaging light’s energy density**

Light’s energy density depends only on electric and magnetic fields, as I show you in the preceding section. In an electromagnetic wave, these fields fluctuate.

Assume that the fields are fluctuating in the shape of a sine wave. The frequency of the fluctuations in a light wave is something like hundreds of thousands of billions of times per second ($10^{14}$ hertz) — too fast to measure. So instead, physicists calculate the average energy density in the space occupied by an electromagnetic wave.

To get the average energy density, you work with the *root-mean-square* (rms) of the electric and magnetic fields (the maximum field divided by the square root of 2). The root-mean-square electric field is given in terms of the amplitude of the fluctuation of the sine-shaped electric field, $E_0$, by the following equation:

$$E_{\text{rms}} = \frac{E_0}{\sqrt{2}}$$
And for the magnetic field, it’s given by

\[ B_{\text{rms}} = \frac{B_0}{\sqrt{2}} \]

where \( B_0 \) is the amplitude of the sine-shaped magnetic field fluctuation. So the average energy density \( u_{\text{avg}} \) in a space occupied by an electromagnetic wave is

\[ u_{\text{avg}} = \varepsilon_0 E_{\text{rms}}^2 \quad \text{and} \quad u_{\text{avg}} = \frac{B_{\text{rms}}^2}{\mu_0}. \]

Here’s a fun problem for you: The sun’s light rays arrive with a root-mean-square \( E \) field of roughly 720 newtons per coulomb. What’s their energy density?

Use the \( u_{\text{avg}} = \varepsilon_0 E_{\text{rms}}^2 \) equation and plug in the numbers to get

\[ u_{\text{avg}} = (8.85 \times 10^{-12} \text{ C}^2/\text{N-m}^2)(720 \text{ N/C})^2 \approx 4.6 \times 10^{-6} \text{ J/m}^3 \]

Looks like the time-average energy density of the sun’s light rays at the Earth is \( 4.6 \times 10^{-6} \) joules/meter\(^3\).

Okay, now imagine a plane area, \( A \). Suppose you want to know how much energy falls on this area every second when the electromagnetic wave is traveling straight down onto it. In a time, \( t \), the wave will travel a distance of \( ct \) (speed times time). So all the energy that is in the volume \( Act \) will strike the plane area. You can use the average energy density formula to work out this energy:

\[ u_{\text{avg}} Act = \varepsilon_0 E_{\text{rms}}^2 Act = \frac{B_{\text{rms}}^2 Act}{\mu_0} \]

So the energy falling per unit area per unit time, \( I \), is given by

\[ I = c \varepsilon_0 E_{\text{rms}}^2 = \frac{c B_{\text{rms}}^2}{\mu_0} \]

The power in a wave per unit area is the intensity (as I show you in Chapter 7). So \( I \) in the preceding formula is the intensity of an electromagnetic wave. This is just the electromagnetic equivalent of the sound intensity that you see in Chapter 7.

Using this formula, you can find the intensity of the sun’s light here on Earth,

\[ I = (3.00 \times 10^8 \text{ m/s})(8.85 \times 10^{-12} \text{ C}^2/\text{N-m}^2)(720 \text{ N/C})^2 = 1,380 \text{ J/s-m}^2 \]
This means that every square meter of the surface of the Earth receives 1,380 joules of energy every second from the sun. (Note however, that the preceding equation only applies if the wave strikes straight down on the area, so this really applies only near the equator. Nearer the poles, you’d have to include a factor to account for the surface’s tilting away from the sun at greater latitudes — obviously the North Pole doesn’t receive as much energy per unit area from the sun as a Caribbean island!)
Here’s a cool quality of light: It interacts with matter so that it bends. Instead of just passing through the universe oblivious to everything else, light is affected by the matter through which it passes, whether that matter is dense like a diamond or thin as air.

Why does light bend? It bends because light is made up of electromagnetic waves — that is, tiny electric and magnetic fields — and they actually interact with the tiny electric and magnetic fields you find in matter (coming from charged particles, such as electrons and protons, and their motion).

This chapter first introduces a different way of representing light waves: the ray. Then it begins a discussion of the tricks that light can play as it bends in glass, water, and other such media. You start by getting a handle on the index of refraction, which is all about just how much light bends in any given material. You also see lenses bring images into focus, or even total internal reflection when light can’t make it out of a block of glass like a prism.

Wave Hello to Rays: Drawing Light Waves More Simply

When you’re exploring the various paths that light waves take as they bounce off and bend through various reflective or transparent materials, you’re more
interested in the places that the waves go, their directions and deflections, than the details of the wave fluctuations of electric and magnetic fields. So for simplicity, you can forget about electric and magnetic fields in most of this chapter and deal with rays of light (I point out when you need to take note of light’s wavy nature). A ray just tells you the wave’s direction of travel without showing the wavelength or speed or frequency or the positions of the wave peaks — the kinds of things I cover in Chapter 8.

Rays are not a new thing — you probably already think of light as rays anyway. They’re just a simpler way to refer to the light wave. You can see what I mean by rays in terms of the light-wave picture in Figure 9-1. Here’s how to interpret this figure:

✓ The dotted lines represent the light waves by showing the positions of the wave peaks of the electric field.

✓ Solid lines are some of the rays that represent the same wave. You can see that these are just lines that are always at right angles to the wave peaks — so they always lie in the direction of travel of the waves. The arrows on the rays show this direction.

In Figure 9-1a, the light waves come from a single point (that is, you have a point source), and they’re spreading out in all directions. Because this light is traveling in all directions from the central point, any line drawn radially outward from this point is a ray.

Figure 9-1b shows another example of rays representing waves. This time you have a plane light ray traveling to the right. I’ve drawn three of the rays that represent this wave.
When working with light rays, just remember two basic principles:

✓ **Rays travel in straight lines.** When they meet a surface, they may reflect or deflect, but while they’re traveling through the same medium without boundaries, rays travel in straight lines.

✓ **Rays are reversible.** When a light ray travels between two points (say, A to B) along a path, then the light from B to A follows the same path in the opposite direction.

### Slowing Light Down: The Index of Refraction

As soon as you get past the concept that light consists of alternating $E$ and $B$ fields that regenerate each other, that there’s a maximum speed at which light can travel (which you can calculate theoretically), and a few other items (see Chapter 8), light traveling in a straight line through a vacuum forever isn’t all that interesting. Sure, you could spend some time studying the situation, but when you have it pretty well scoped out, you kind of wish something else would happen.

But when light hits and starts traveling through something else, then light becomes interesting again. When light in a vacuum enters any material and begins to travel through that material instead, the light slows down, because the electric and magnetic fields around the light in the material act as a drag. For example, when light impacts a block of transparent material (this is all theoretical, so make it a 60-pound block of diamond), the light slows down and bends.

In this section, you look at how much that material slows light down and see how much the light bends as a result. I also show you that not all light bends equally, because the index of refraction varies depending on the wavelength of the light.

### Figuring out the slowdown

Light reaches its maximum speed, $c$, in a vacuum. That’s about $3.0 \times 10^8$ meters per second, and it’s all downhill from there, because whenever light travels through anything else — even air — it slows down.
The ratio of the speed of light in a vacuum, \( c \), to the speed of light in a material, \( v \), is a constant for any given material, and that ratio is called the index of refraction, \( n \). Here’s the definition of the index of refraction:

\[
n = \frac{\text{speed of light in a vacuum}}{\text{speed of light in the material}} = \frac{c}{v}
\]

The index of refraction is just a pure number, because it’s the ratio of speeds, so it has no units, like relatively few other quantities in physics.

Generally speaking, the denser the material, the more electric and magnetic fields it has to slow light down. So diamond, for example, has a higher index of refraction than air. Table 9-1 gives a starter list of indexes of refraction for various materials. The table also includes temperature, which can affect the density of the material and therefore its index of refraction.

A material usually contracts as its temperature decreases, so it becomes denser and its index of refraction can rise. However, water is a special case. Ice (at 0°C) has a refractive index of 1.32, and water has the higher value of 1.33. When the water freezes, the molecules form ice crystals, which happen to have a structure that’s less dense than the original water. That’s why ice floats on water.

<table>
<thead>
<tr>
<th>Material</th>
<th>Temperature (°C)</th>
<th>Index of Refraction (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diamond</td>
<td>20°C</td>
<td>2.42</td>
</tr>
<tr>
<td>Window glass</td>
<td>20°C</td>
<td>1.52</td>
</tr>
<tr>
<td>Benzene (liquid)</td>
<td>20°C</td>
<td>1.50</td>
</tr>
<tr>
<td>Water</td>
<td>20°C</td>
<td>1.33</td>
</tr>
<tr>
<td>Ice</td>
<td>0°C</td>
<td>1.32</td>
</tr>
<tr>
<td>Air</td>
<td>20°C</td>
<td>1.00029</td>
</tr>
<tr>
<td>Oxygen</td>
<td>20°C</td>
<td>1.00027</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>20°C</td>
<td>1.00014</td>
</tr>
</tbody>
</table>

So if diamond has a refractive index of 2.42 at 20°C, what’s the speed of light in a diamond? Well, it’s

\[
v_{\text{diamond}} = \frac{c}{n_{\text{diamond}}} = \frac{3.0 \times 10^8 \text{ m/s}}{2.42} = 1.2 \times 10^8 \text{ m/s}
\]

So light travels at only \( 1.2 \times 10^8 \) meters per second in diamond. Positively pokey.
Calculating the bending: Snell’s law

When light slows, it bends. You can put the index of refraction to work with Snell’s law, which tells you exactly how much light bends when it enters a new medium. (See the preceding section for info on the index of refraction.)

The incident (incoming) light comes in at an angle of $\theta_1$, measured with respect to a line perpendicular to the material’s surface — that perpendicular line is called a normal (see Figure 9-2). And when the light bends and travels off into the medium, it goes at a new angle with respect to the normal, $\theta_2$. Here’s how the angles relate:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

where $n_1$ is the index of refraction of the medium the light is coming from (it doesn’t need to be vacuum to have Snell’s law work) and $n_2$ is the index of refraction of the medium the light enters (which could be diamond, or glass, or even vacuum).

Note that if you know the incident angle and the indexes of refraction of the materials involved, you can solve for the angle at which light heads off into the new medium like this:

$$\sin^{-1}\left(\frac{n_1 \sin \theta_1}{n_2}\right) = \theta_2$$

That’s a nice result — it tells you what angle you can expect light to go speeding off at in a new medium.

[Figure 9-2: Snell’s law.]
Here’s one thing that Snell’s law tells you that’s not immediately obvious: If you’re going from less-dense to denser material, light is bent toward the normal; if you’re going from denser to less-dense material, light is bent away from the normal.

For instance, if you look at Figure 9-2, you see the light ray traveling from less-dense to denser material, and the light bends toward the normal. Now if you remember that light rays are reversible, you can imagine the same ray going in the opposite direction (from the denser to the less-dense material) — then the ray bends away from the normal.

Now try some numbers. Say that light goes from air (which you can treat as vacuum for the purposes of this example) to your 60-pound diamond block. And say that light hits the diamond at 65° with respect to the normal. What’s the angle the light bends to inside the diamond?

That is, you have $n_1 = 1.00$, $n_2 = 2.42$, and $\theta_1 = 65°$, and you need to find $\theta_2$. You can use Snell’s law like this:

$$\theta_2 = \sin^{-1} \left( \frac{n_1 \sin \theta_1}{n_2} \right) = \sin^{-1} \left( \frac{1.00 \sin 65°}{2.42} \right) = 22°$$

So the light comes in at 65° and ends up at 22°.

**Rainbows: Separating wavelengths**

Here’s something that you may not like to hear because it complicates things a bit: The index of refraction of materials varies slightly depending on the wavelength of the light. On the other hand, you probably like the results of this fact: rainbows. Because the colors in sunlight (which contains all colors) bend different amounts in water droplets, you get a separation of colors into that familiar display of rainbows.

The index of refraction does vary by light wavelength but not strongly (so physicists often ignore it). Table 9-2 lists some values for the various colors of light and what the corresponding indexes of refraction are in glass.

| **Table 9-2** Indexes of Refraction According to Wavelength |  |
| --- | --- | --- |
| **Color** | **Wavelength (nanometers)** | **Index of Refraction in Glass** |
| Red | 660 | 1.520 |
| Orange | 610 | 1.522 |
| Yellow | 580 | 1.523 |
Say that light is incident at 45° to a sheet of glass. How much does red light ($\lambda = 660$ nm) bend, compared to violet light ($\lambda = 410$ nm)?

Snell’s law tells you that $n_1 \sin \theta_1 = n_2 \sin \theta_2$, so

$$\sin^{-1}\left(\frac{n_1 \sin \theta_1}{n_2}\right) = \theta_2$$

The light is first traveling through the air, so $n_1 = 1.00$. For red light in glass, the index of refraction is 1.520, so you get

$$\theta_2 = \sin^{-1}\left(\frac{1.00 \sin 45°}{1.520}\right) = 27.7°$$

For violet light, the index of refraction in glass is 1.536, so you get

$$\theta_2 = \sin^{-1}\left(\frac{1.00 \sin 45°}{1.536}\right) = 27.4°$$

So as you can see, you get different amounts of bending depending on the color of light. Note that the angle calculated is with respect to the normal, so violet light is bent slightly more than red light here.

Because of the differing indexes of refraction for different wavelengths, light splits in a prism (see Figure 9-3). Here’s how it works: when light enters the prism, it’s going from air into a medium with a higher index of refraction — typically glass — so it bends in the glass toward the normal (a line perpendicular to the surface). Because the index of refraction is stronger for shorter wavelengths, red light (with a longer wavelength) bends less than violet light (with a shorter wavelength). When the light emerges from the prism, it bends away from the normal, and how much it bends depends on the index of refraction — so red light is further separated from violet light.

In actual rainbows, the light not only refracts when it enters the water droplet but also reflects inside the water droplet. You can read more about this phenomenon in the sidebar “Reflecting on rainbows,” later in this chapter.
Bending Light to Get Internal Reflection

When light enters a material with a lower index of refraction, that light bends away from the normal (an imaginary line perpendicular to the material’s surface). If the incident light comes in at a large enough angle, the light may bend away so much that it doesn’t refract at all — it gets reflected instead.

In this section, I discuss two cases in which you get reflection. In the first, all the incident light is reflected. In the second, only polarized light — light with aligned electric and magnetic fields — is reflected, and the rest of the light is refracted as it enters the less-dense material.

Right back at you: Total internal reflection

Sometimes light doesn’t make it out of a material, and it ends up bouncing around inside. Perhaps you’ve noticed that when you turn a glass paper-weight, one of the internal edges sometimes looks like a mirror, reflecting with a silvery appearance. That’s total internal reflection.

To see how this works, take a look at Figure 9-4. Light is going from a dense medium such as glass to air. That means that the light bends away from the normal when it gets into the air, as you see in ray 1. If you keep increasing $\theta_1$, ultimately, $\theta_2$ reaches 90° — that is, the light just skirts along the glass surface, as in ray 2. If you increase $\theta_2$ any more, the light will be reflected back into the glass, as you see in ray 3. This is what’s known as total internal reflection. When light goes from a dense medium to a less-dense medium, it bends away from the normal, and if the incident angle becomes large enough, the light will be reflected back into the denser medium where the two materials meet.
Total internal reflection happens when the angle at which the light tries to exit the dense medium, \( \theta_2 \), becomes so large that it reaches \( 90^\circ \). Right at that point, when the light ends up skimming the glass/air interface, you have total internal reflection. The incident angle at which this happens is called the critical angle — \( \theta_c \). At the critical angle, light ends up with an exit angle of \( 90^\circ \) with respect to the normal. In other words, when \( \theta_1 = \theta_c \), \( \theta_2 = 90^\circ \).

What value does \( \theta_c \) have? You can use Snell’s law to find out:

\[
n_1 \sin \theta_1 = n_2 \sin \theta_2
\]

Plugging in the values, you get

\[
n_1 \sin \theta_c = n_2 \sin 90^\circ
\]

\[
\sin \theta_c = \frac{n_2 \sin 90^\circ}{n_1}
\]

Because \( \sin 90^\circ = 1 \), you have the following for the critical angle for total internal reflection:

\[
\sin \theta_c = \frac{n_2}{n_1}
\]

\[
\theta_c = \sin^{-1} \left( \frac{n_2}{n_1} \right)
\]

Note that this equation requires that \( n_2 < n_1 \) (otherwise you can’t have total internal reflection).
For example, say you have a diamond ring and that light is bouncing around inside the gem. What’s the critical angle beyond which light already inside the diamond gets totally internally reflected back into the diamond? You can use the equation for total internal reflection. The index of refraction of air is near 1.00, and the index of refraction for diamonds is 2.42, so you have the following:

\[ \theta_c = \sin^{-1} \left( \frac{n_1}{n_2} \right) = \sin^{-1} \left( \frac{1.00}{2.42} \right) = \sin^{-1} (0.413) = 24.4^\circ \]

So if light hits the diamond-air interface at an angle of more than 24.4°, it’ll bounce back into the diamond, and that facet of the diamond acts like a mirror — that’s one of the reasons properly cut diamonds seem to exhibit so much fire.

**Polarized light: Getting a partial reflection**

Here’s a peculiar fact about light — when it bounces off a nonmetallic surface, it gets polarized. That means that the light rays’ electric fields are lined up and its magnetic fields are lined up.

When you talk about polarization, you normally discuss the direction of the electric (E) field in the light ray. The E field oscillates in a direction that’s perpendicular to the light ray’s direction of travel, and the plane that the light ray and the E vector form is called its polarization. So if you have a light ray coming toward you and its E vector is oscillating horizontally, the light is polarized horizontally.
With normal light, the $E$ vector can oscillate in any direction perpendicular to the direction of travel. But when you bounce light off a nonmetallic surface, the reflected light ends up being polarized to some extent in the plane of the surface. For example, if light bounces off a pool of water, the reflected light ends up being chiefly polarized in the horizontal direction.

In Figure 9-5, you see the incident ray coming in from the left. A number of arrows in all directions (perpendicular to the direction of travel, of course) represent the unpolarized electric-field component of this ray. The reflected ray on the right is totally polarized, so it has electric field oscillations only in the horizontal direction. The refracted ray is partially polarized, because the reflected wave preferentially carried away electric field oscillations in the horizontal direction, leaving the refracted wave with relatively few. You can see this in Figure 9-5 — in the refracted ray, the arrows representing the horizontal electric field oscillations are diminished compared to the others.

### Reflecting polarized light at Brewster’s angle

When light bounces off a nonmetallic surface, the amount of polarization depends on the angle of incidence (with respect to the normal, an imaginary line perpendicular to the surface). And at an angle of incidence called **Brewster’s angle**, $\theta_B$, the polarization is total. So when light reflects off a pool of water at Brewster’s angle, the reflected light is completely polarized in the horizontal direction. Here’s the formula for Brewster’s angle:

$$\tan \theta_B = \frac{n_2}{n_1}$$
where \( n_1 \) and \( n_2 \) are the indexes of refraction. So what is Brewster’s angle for water? Well, the index of refraction for water is about 1.33 and the index of refraction for air is about 1.00, so you have the following:

\[
\tan \theta_B = \frac{1.33}{1.00}
\]

\[
\theta_B = \tan^{-1}1.33 = 53^\circ
\]

So Brewster’s angle for water is 53°.

**Noting the angle between the reflected and the refracted rays**

You can prove that the refracted ray, which enters the water, is at 90° with respect to the reflected ray if the incident light enters the water at Brewster’s angle (the angle at which polarization is total). To prove this, start with Snell’s law:

\[
n_1 \sin \theta_1 = n_2 \sin \theta_2
\]

where \( \theta_1 \) and \( \theta_2 \) are shown in Figure 9-5. You can write this as the following, using Brewster’s angle, \( \theta_B \), for \( \theta_1 \):

\[
\sin \theta_B = \frac{n_2 \sin \theta_2}{n_1}
\]

Using Brewster’s equation, you know that

\[
\tan \theta_B = \frac{n_2}{n_1}
\]

You have \( \tan \theta_B = \sin \theta_B / \cos \theta_B \), which follows from the trig definition of the tangent, so do the following calculations:

\[
\sin \theta_B = \frac{\sin \theta_B \sin \theta_2}{\cos \theta_B}
\]

\[
1 = \frac{\sin \theta_2}{\cos \theta_B}
\]

\[
\cos \theta_B = \sin \theta_2
\]

And because \( \sin \theta = \cos(90° - \theta) \), you can say that

\[
\cos \theta_B = \cos(90° - \theta_2)
\]

\[
\theta_B = 90° - \theta_2
\]

So Brewster’s angle is 90° from the refracted angle. Cool.
Getting Visual: Creating Images with Lenses

Many weird and wonderful things happen to light when it hits curved surfaces, so in this section, you step into a world of images. You find out what the field of optics considers to be an object and how images are made by the curved surfaces of lenses. This is the physics that led to telescopes and microscopes, which opened up new doors of perception of the universe.

Here, I demonstrate how you work out where and how big an image will be simply by drawing a few lines. Such ray drawings can give you a good mental picture of what lenses do to the light that passes through them. After this, I show you a couple of equations, which tell you exactly where the images are and how big they are, without your having to take out your ruler and pencil.

Defining objects and images

As far as the field of optics is concerned, an object is simply a source of light rays. It doesn’t have to glow with light; it can just reflect light from another source. The important point is that light rays should radiate away from the object. For example, this book would be considered an object in physics. The book isn’t generating the light, only reflecting it from your lamp or the sun or whatever the light source is where you’re reading.

For simplicity, this book considers only very simple objects like point sources, which are simply points that radiate rays, or line sources. A line source is simply a line that radiates rays in all directions from every part of it — physicists draw these as arrows, because sometimes you find them upside down, and the arrowhead emphasizes direction.
So much for objects — now what about the images that are made from them? Well, the simplest example of an image is one you probably see every day — your own image in the mirror. In this case, you’re the object (no offense!), and the mirror reflects the rays coming from you and makes an image behind the mirror. The image is the point from which the rays leaving your face seem to be coming from. You know that you — the object — are not behind the mirror, but your image appears behind. This kind of image is called a virtual image.

Another example of an image is the one projected in a cinema. Each frame of the movie is turned into an image, which is projected onto the screen. This is a different type of image from the one you see in the mirror, because the rays don’t just appear to be coming from the image — they actually converge onto the image. That is, the rays make the picture all come together on the movie screen. Because this type of image can be made to fall on a screen, it’s called a real image.

**Now it’s coming into focus:**

**Concave and convex lenses**

You can find lenses everywhere — in standard digital cameras, in TV cameras, perched on the end of peoples’ noses, in flashlights, even sometimes in peoples’ watches when the really tiny date has to be magnified. A lens is simply a transparent object (usually a disk of glass), that takes an object and makes an image. It can do this because it has two curved surfaces.

Here are two types of lenses:

- **Convex (converging):** A convex lens curves so that it bulges in the middle (see Figure 9-6). When you put a point object in front of this lens, some of the rays that radiate away travel through the lens. When they meet the first surface, they refract, or bend, and when they leave the lens, they refract some more. The effect of the convex lens in Figure 9-6a is to gather all these rays together back to a point — this is your image. Because all the rays converge there, it’s a real image.

  If a number of parallel rays strike the lens, then they all converge on a point called the focal point, as you see in Figure 9-6b. You may have already discovered the focal point for yourself if you’ve ever tried to focus sunlight into a single bright spot with a magnifying glass. The sun is so far away that the rays coming from it are pretty much parallel, so when they pass through your magnifying glass, they converge at the focal point. If you place a piece of paper there, then all the converging rays can make the paper burn.
Concave (diverging): A concave lens is narrower in the middle. This time, when the rays from the object strike the lens, they diverge. All the rays appear to diverge from a certain point, and this is the virtual image. See Figure 9-7a.

Now if you send a bunch of parallel rays into a concave lens, they all diverge, but they all appear to be diverging from the focal point (see Figure 9-7b).

The way I remember the difference between convex lenses and concave lenses is that a diverging lens forms a sort of a cave (because its middle is all hollowed out), as in concave.

The distance between the lens and the focal point is called the focal length, \( f \). The strength of a lens is measured solely by this length — if it’s shorter, then the rays are bent through a larger angle, and the lens is stronger.

In many lenses, one of the lens’s surfaces may curve more than the other. Even in those cases, there’s still only one focal length, so the focal point on each side of the lens is at the same distance, \( f \). Also, however the sides are curved, if the lens is thicker in the middle than at the edges, it’s convex; otherwise, it’s concave.
Another special point for a lens is the center of curvature, which is a distance, \( C \), from the lens. The distance between this point and the lens is called the radius of curvature.

The radius of curvature is not related in a simple way to the amount of curvature of the lenses; instead, just think of it in terms of the focal length of the lens. The radius of curvature is simply twice the focal length, \( C = 2f \).

The dotted line in Figures 9-6 and 9-7 is called the optical axis of the lens. It’s just the line that passes through the lens at its widest part (or thinnest part if the lens is concave) and that’s normal (perpendicular) to the surface of the lens there.

**Drawing ray diagrams**

You can draw three special lines to find the image that a lens makes, based on where the object is. These lines tell you where the image appears, whether it’s upside down or right-side-up, whether it’s larger or smaller than the original object, and whether the image is real (made of converging light rays) or virtual. In this section, you discover how to draw ray diagrams for both convex (converging) and concave (diverging) lenses.
The object I use in each section is a line object, which looks like an arrow. The arrowhead makes sure you always know whether the image is upside down or the right way up. You can think of this type of object as just being a lot of point objects all in a line. (For more on objects and point and line sources, see the earlier section “Defining objects and images.”)

**X marks the spot: Finding images from convex lenses**

The position and size of the image depend on the position and size of the object. For a convex (converging) lens, here’s how to draw three special lines that help you figure out the position and size of the image (see Figure 9-8):

- **Ray 1:** One ray leaves the object, travels toward the center of the lens, and travels straight through without being deflected at all.
- **Ray 2:** Another ray travels from the object, parallel to the axis of the lens, and is deflected so that it passes through the lens’s focal point.
- **Ray 3:** A third ray travels from the object and passes through the focal point on the near side before reaching the lens. The lens deflects this ray so that it then travels parallel to the axis of the lens.

The image is located where these three lines cross.

You can draw these ray lines for any point on your object and find every point of the image, but for simplicity, most people draw only the lines from the tip of the object (the point of the arrow). And although I draw three rays in Figure 9-8, you really need only two rays to locate the image. Three is better for safety — as a check on the other two — but if you know what you’re doing and are under time pressure (such as when you’re taking a test), two is enough.

If you go through the line-drawing process for convex lenses, you encounter three special cases for what the image looks like. Here are these cases, which are all based on the position of the object (if you need more info on the focus and radius of curvature, see the earlier section “Now it’s coming into focus: Concave and convex lenses”):
✓ The object is beyond the radius of curvature, \( C \): If the object is this far out, the image is real, upside down, and smaller than the object (see Figure 9-9a).

✓ The object is between the radius of curvature, \( C \), and the focal length, \( f \): Here the image is still real and upside down, but now it’s larger than the object (see Figure 9-9b).

When the object is at the center of curvature, then its image is the same size as the object, and the closer the object gets to the focal point without crossing, the larger the image becomes.

✓ The object is closer to the lens than the focal length, \( f \): This is an interesting case, because for the first time, you don’t get a real image (see Figure 9-9c). You can’t even draw the third ray, because it doesn’t go through the lens. There’s no place in space that the three rays come together, no place that you can bring an actual physical screen and get an in-focus image — the image is virtual. It’s also larger than the object and right-side-up.

Figure 9-9: Ray diagrams for three special cases of images from a convex lens.
Recognize the situation in which the object is between the focal point and the lens? That’s the case where a converging lens forms an upright image that’s larger than the object you’re looking at, on the same side of the lens as the object itself (the image is virtual, so no light rays actually come together to form the image, but looking through the lens bends the rays so that they appear to be coming from the image). That’s a magnifying glass — and the fact that you don’t get an enlarged upright image until the object is between the focal point and the lens shows why you have to hold the lens up close to whatever you’re trying to magnify. Cool, eh? Sherlock Holmes would be proud.

**Going virtual with concave lenses**

With a concave (diverging) lens, light bends away from the horizontal after passing through the lens. You can see a diverging lens in Figure 9-10.

So when you have a diverging lens, can you work with ray diagrams to find the image? Absolutely — but this time you use only two rays:

**Ray 1:** This ray goes from the tip of the object and through the center of the lens. Figure 9-10 shows that this ray goes through the center of the lens and isn’t deflected at all. Easy.

**Ray 2:** This ray goes horizontally from the tip of the object to the lens, parallel to the axis of the lens; then the ray deflects away from the axis along a line that passes through the nearest focal point — that is, the ray travels as though it came from the nearest focal point. Figure 9-10 shows that the second ray bends away from the horizontal on the other side of the lens.

So if the second ray is bent away from the horizontal, where does it intersect the first ray? The answer is that they don’t intersect on the side of the lens that the observer (who is looking through the lens) is on. Instead, you extend the rays back through the lens until they intersect. Because you’re extending these rays in a straight line back to somewhere they don’t actually exist, the image is virtual. In other words, the image forms on the same side of the lens as the object, as you see in Figure 9-10.
Regardless of where the object is, the virtual image from a concave lens is always right-side-up, and it’s no farther from the lens than the focal length. The image is also upright and smaller than the object.

**Getting Numeric: Finding Distances and Magnification**

With a few lens equations, you can find out where images appear and how big they are. Drawing ray diagrams (as I show you in a preceding section) is a good way to get a strong picture of what lenses do, but when you have that in mind, you’ll find these equations a much quicker way of finding out what your lenses are doing.

There’s really nothing mystical about the equations in this section — they just come from the laws of refraction (which you can examine in “Slowing Light Down: The Index of Refraction,” earlier in this chapter). People derived these equations by applying the law of refraction to the curved surfaces of the lenses, but you don’t have to bother doing that — here, I just show you how the equations work.

**Going the distance with the thin-lens equation**

Using the thin-lens equation, you can relate the distance an object is from a lens, the distance from the lens to the image, and the lens’s focal length. The equation is called the *thin-lens equation* because it’s actually an approximation, and that approximation really holds only for “thin” lenses — that is, lenses whose bending power isn’t too great (stronger lenses have a shorter focal length, so strong lenses have to be more curved and therefore thicker). This section gives you the equation, shows you how it works, and provides a couple of example calculations.

**Introducing the thin-lens equation**

Here’s the thin-lens equation:

\[
\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}
\]

This equation relates the distance the object is from the lens \((d_o)\) and the distance the image is from the lens \((d_i)\) with the focal length, \(f\).
The signs of $d_o$, $d_i$, and $f$ are important. For instance, you give converging lenses a positive focal length $f$, but diverging lenses get a negative focal length (that’s because the image forms on the other side of the lens from the observer). And if you get a negative distance for the image distance, $d_i$, that means the image is virtual (which, for a lens, means that the image forms on the same side as the object).

The best way to state the rules for the signs of $d_o$, $d_i$, and $f$ in the thin-lens equation is in terms of the incoming and outgoing sides of the lens (see Figure 9-11). When light from an object travels through a lens, I call the side that the light enters the incoming side; the side of the lens through which the light leaves is the outgoing side. Then the rules are simple to state:

- **Object distance, $d_o$:** When the object is on the incoming side of the lens, then the distance from the object to the lens is positive (in this book, this is always the case).
- **Image distance, $d_i$:** When the image is on the outgoing side of the lens, then the image distance is positive; otherwise, it’s negative.
- **Focal length, $f$:** When the lens is convex, its focal length is positive; otherwise, it’s negative.

**Note:** This book always pictures the object to the left of the lens, so the incoming side is on the left in the figures (including Figure 9-11). But these rules still apply in exactly the same way if this situation is reversed, because then the incoming and outgoing sides would also reverse.
Doing calculations with the thin-lens equation

Now try some numbers with the thin-lens equation. Say that you have a camera with a converging lens that has a focal length of 5.0 centimeters, and the flower you’re taking a picture of is 2.00 meters in front of the lens. How far on the other side of the lens does the image form? You can put the thin-lens equation to work right away:

\[
\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}
\]

Rearranging this gives you

\[
\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}
\]

Combining the fractions and solving for \(d_i\) gives you

\[
d_i = \frac{1}{\frac{d_o}{f} - f}
\]

\[
d_i = \frac{fd_o}{d_o - f}
\]

So plugging in the numbers gives you the following:

\[
d_i = \frac{(0.050 \text{ m}) (2.00 \text{ m})}{2.00 \text{ m} - 0.05 \text{ m}} = \frac{0.10 \text{ m}}{1.95} = 0.051 \text{ m} = 5.1 \text{ cm}
\]

So the image forms at 5.1 centimeters behind the lens of the camera (on the side opposite to the flower).

Now try one with a diverging lens. Say you have a diverging lens with a focal length of –5.0 centimeters (See? I told you diverging lenses get negative focal lengths), and you place an object 7.0 cm in front of it. Where does the image appear to form?

You can use the thin-lens equation like this:

\[
\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}
\]

And I’ve already solved for \(d_i\) this way:

\[
d_i = \frac{fd_o}{d_o - f}
\]
Chapter 9: Bending and Focusing Light: Refraction and Lenses

Plugging in the numbers gives you the answer:

\[
d_i = \frac{(-0.050 \text{ m})(0.070 \text{ m})}{0.070 \text{ m} - (-0.050 \text{ m})} = \frac{-0.0035 \text{ m}^2}{0.12 \text{ m}} = -0.029 \text{ m} = -2.9 \text{ cm}
\]

So the image forms at \(-2.9\) centimeters — that is, between the focal length and the lens. Note that this result is negative. That means that the image is visible only by looking through the lens — not on the same side of the lens as the observer. In other words, it’s a virtual image, as you’d expect from the type of lens and the placement of the object (to see why, check out the ray diagram in Figure 9-10).

**Sizing up the magnification equation**

The thin-lens equation tells you where an image will form, but it doesn’t tell you very much about the image itself. Sometimes, you want to know whether that image is bigger or smaller than the object and whether it’s upright or upside down with respect to the object. That’s where the magnification equation comes in.

**Finding the magnification equation**

Say that \(h_i\) is the height of the image and \(h_o\) is the height of the object. You can see that the magnification of the image compared to the object would be \(m\), like this:

\[
m = \frac{h_i}{h_o}
\]

Figure 9-12 shows two of the rays that go to make the image of an object from a convex lens. The figure shows the size of the object and the image, along with the distances of the object and image from the lens. The ray that travels straight through the lens makes an angle \(\theta\) with the axis. By using geometry and similar triangles (shaded gray in Figure 9-12), you can show that the magnification is equal to the ratio of the image distance to the object distance like this:

\[
m = \frac{h_i}{h_o} = -\frac{d_i}{d_o}
\]

Note that a negative value for the magnifications means the image is upside down with respect to the object, and positive magnification (the kind magnifying glasses and most optical telescopes produce) means that the image is right-side-up compared to the object.
Why does the magnification equation have a minus sign in it? That’s because the magnification of a magnifying glass — that is, a converging lens — is considered positive when the image is virtual, because you look through the lens to see the image and it’s right-side-up. But because the image is virtual, the distance to the image from the lens, \( d_i \), is negative. The minus sign in the magnification equation corrects that — so even though the image comes out virtual (which means negative in the thin-lens equation), it’s still upright (which by convention means positive in the magnification equation).

**Plugging in some numbers**

Now you can figure out the distance from a lens to an image using the thin-lens equation, and because you already know the distance from the object to the lens, you can use the magnification equation to figure out the magnification.

Try some numbers. Start by taking a look at the converging-lens problem from the earlier section “Going the distance with the thin-lens equation”: The camera has a converging lens that has a focal length of 5.0 centimeters, and the object you’re taking a picture of is 2.0 meters in front of the lens. In this problem, the distance from the lens to the image turns out to be 5.1 centimeters.

What’s the magnification of the converging lens in this setup? You can use the magnification equation:

\[
m = -\frac{d_i}{d_o}
\]

Plugging in the numbers gives you the answer:

\[
m = -\frac{5.1 \text{ cm}}{200 \text{ cm}} = -2.6 \times 10^{-2}
\]

So the magnification is \(-2.6 \times 10^{-2}\). That tells you a few things. First, note that magnification is another of those relatively few quantities in physics that has no units — it’s just a multiplier. Second, the negative sign tells you that the image formed is upside down with respect to the object (did you know that camera images are inverted?). And third, it tells you that the magnification
is very small, so you can capture big objects on small film surfaces or pixel arrays (of digital cameras).

Now take a look at the diverging lens. In the second problem I solve in “Going the distance with the thin-lens equation,” the focal length is –5.0 centimeters, and the object is placed 7.0 centimeters in front of it. The image forms at –2.9 centimeters. So what’s the magnification of the lens with this setup? You can use the magnification equation:

\[ m = -\frac{d_i}{d_o} \]

Plugging in the numbers here gives you this:

\[ m = -\frac{-2.9 \text{ cm}}{7.0 \text{ cm}} \approx 0.41 \]

So although the image is upright, it’s still smaller than the object (putting it closer to the lens will result in magnification greater than 1).

**Combining Lenses for More Magnification Power**

You can use lenses together — in fact, that’s one of their most popular uses, in microscopes and telescopes and such. Lenses used in combination are almost always converging lenses. In this section, you look at how combining two lenses gives you more magnification power, and you see how such combinations usually work.

**Understanding how microscopes and telescopes work**

When you combine two converging lenses, the first lens is nearest the object, so it’s called the *objective lens*. As you can see in Figure 9-13, the object is farther away from the objective lens than the focal length \( f_o \) of the objective lens. In a microscope, the object you’re looking at may not be much beyond the focal length, but in a telescope, the object is always much farther from the objective lens.

As you can see in the figure, rays from the tip of the object form an image past the focal length of the objective lens on the right side of the lens. This image is larger than the object, and it’s inverted — it’s also a real image.
Two common vision problems are nearsightedness, where people can focus only on near objects, and farsightedness, where people can focus only on objects farther away. Corrective lenses help with both — diverging lenses for nearsightedness and converging lenses for farsightedness.

The diagram at the top of the following figure shows an uncorrected nearsighted eye. The problem here is that the lens of the eye tends to focus objects some distance in front of the retina. As a result, objects look blurry. The bottom diagram shows the same eye corrected with a diverging lens. Now the diverging lens makes the rays from objects diverge slightly to counteract the overly strong converging powers of the lens of the eye. As a result, the image comes into focus directly on the retina, as it should.
The diagram at the top of the next figure shows an uncorrected farsighted eye. Here, the problem is that the lens of the eye doesn’t converge light passing through it enough. As a result, the image forms past the retina. The solution is to use a converging lens, as in the bottom diagram. The converging lens makes the light rays from the object converge a little, which helps the lens of the eye focus the image, which appears on the retina.

Now here’s the clever part — the image from the objective lens becomes the object for the second lens. That is, the second lens looks at the image (which is real) just as though it were an actual object. Because light rays come together at the image, this works very well.

The second lens is called the eyepiece (not surprisingly, because that’s the lens that is closest to the eye). Everything is set up so that the first image, the one created by the objective lens, falls just inside the eyepiece’s focal length, $f_e$. That ensures that the second lens magnifies the image to a very large size. You can see this in Figure 9-14.

This time, the eyepiece creates a virtual image (so you can see it by looking through the eyepiece, as with microscopes and telescopes). Thus, the final image ends up large and inverted, as you can see in Figure 9-14. But how large? What’s the magnification of such a combination of lenses?

Well, the image from the objective lens is magnified. Then this magnified image is the object for the eyepiece lens, which magnifies again. The total resulting magnification is the product of the magnification from each lens.
Here’s why it’s best to place the object not too far beyond the focal length of the objective lens in a microscope: If the object is between the center of curvature and the focal length of a convex lens (as I explain earlier in “X marks the spot: Finding images from convex lenses”), then the image is real and has a greater size than the object — and this size increases as the object approaches the focal length. In its normal operation, a microscope needs a real image from the object, and the larger this image, the greater the magnification.

**Getting a new angle on magnification**

Magnification for microscopes and for telescopes is often figured in terms of *angular magnification*. That is, the object you want to look at takes up a certain angle of your vision (for example, the moon takes up about half of a degree of the 360° you can see by turning completely around), and if you use a telescope, the object looks larger (the moon may take up what seems like three times the same angle). The symbol for angular magnification is \( M \). In this section, you use some formulas to find the magnification of microscopes and telescopes.

**Getting up close and personal with microscopes**

Microscopes are typically made from two converging lenses in combination. Here, the object is between one and two focal lengths from the objective lens. If the distance between the objective lens and the eyepiece lens is \( L \), then the angular magnification for a microscope turns out to be the following:

\[
M = \left( \frac{L - f_e}{f_o f_e} \right) N
\]
$N$ is the distance to the near point for the eye. The near point is the closest you can hold, for example, some text and still read it. For a normal eye, $N$ equals 25 centimeters.

Suppose you have an eyepiece with focal length of 5.0 centimeters and an objective lens with a focal length of 0.40 centimeters. The length between the two lenses, $L$, is 25.0 centimeters. What is the angular magnification of the microscope?

Plugging in the numbers — using $N = 25.0$ centimeters — gives you the answer:

$$M = \left( \frac{L - f_e}{f_o f_e} \right) N = \left( \frac{25.0 \text{ cm} - 5.0 \text{ cm}}{(0.40 \text{ cm})(5.0 \text{ cm})} \right)(25.0 \text{ cm}) = 250$$

So the angular magnification of the microscope is 250.

**Bringing the heavens near with telescopes**

Like microscopes, optical telescopes are frequently made with two converging lenses. With telescopes, the object you’re looking at is a far distance away compared with the distance to the eye’s near point, $N$, and the focal length of the objective lens.

In this case, you can make some approximations, and the angular magnification of a telescope is about equal to the following:

$$M = -\frac{f_o}{f_e}$$

Say, for example, that you have a telescope whose objective lens has a focal length of 100 centimeters and an eyepiece with the focal length of 0.5 centimeters. What angular magnification will the telescope give you?

You can use the angular-magnification equation and plug in numbers like this:

$$M = -\frac{f_o}{f_e} = -\frac{100 \text{ cm}}{0.5 \text{ cm}} = -200$$

So the angular magnification of the telescope is about −200, where the negative sign simply means that the image is inverted.
You can say a lot about mirrors — both plane (straight) mirrors and trickier spherical mirrors. As for spherical mirrors, you can get images in both concave mirrors (mirrors that look like the inside of a bowl) and convex mirrors (mirrors that look like the outside of the mirrored bowl). You can predict where images will form and whether they’ll be upright or upside down — no mean feat, given that those mirrors can act pretty wacky when you hold them in your hand (or put them on a funhouse wall).

In this chapter, you work with some basic properties of reflection, and you see how light bounces off both flat and curved surfaces. After you see some ray diagrams, I present a couple of equations so you can get down with the math.

The Plane Truth: Reflecting on Mirror Basics

Even people with the most casual disregard for their appearance probably see themselves in a mirror every day. The flat plane mirror you use so frequently is also extremely important to optics. The basic law of how light reflects is expressed in terms of how light bounces off a plane mirror. Then,
if you take any curved reflecting surface — like the ones in a carnival’s hall of mirrors — and look at it closely enough, it appears flat at every point (just as the Earth is curved, but because you see it so close up, it appears flat wherever you’re on it). So if you know how light reflects off a flat surface, you also know how light reflects off every part of any curved surface — bargain!

This idea applies wherever reflection occurs from a flat surface, even if it’s not a mirror. So without further ado, here’s your introduction to mirrors and other reflective surfaces.

Mirrors were often made of polished metal in the ancient world. These days, they’re commonly made of metal electroplated onto glass. And glass itself can form a partial mirror — if you stand next to a window and look out, you often see a ghostly image of yourself looking out in the window glass. You see that image because glass commonly reflects about 7 percent of the light that hits it instead of transmitting it through the glass.

**Getting the angles on plane mirrors**

Figure 10-1 shows a *plane mirror* — that is, a straight mirror — lying on its back. A light ray comes in from upper left in the figure, hits the mirror, bounces off the mirror, and leaves to the upper right.

The angle at which the light ray comes in to the mirror is called the *angle of incidence*, \( \theta_i \), and the angle at which light is reflected is called the *angle of reflection*, \( \theta_r \). Note that these angles are with respect to the normal — a *normal* is a line perpendicular to the mirror’s surface.

The angle of incidence is equal to the angle of reflection:

\[ \theta_i = \theta_r \]
That’s why when you’re driving and you see an approaching car’s image in your rear-view mirror, you know just which way to turn to see the actual car.

**Forming images in plane mirrors**

Plane mirrors are especially good at forming images. This section takes a look at image formation in a little more depth.

Mirrors form virtual images of objects, as you see in Figure 10-2. The image is virtual because no actual light rays meet to form that image (see Chapter 9 for more on virtual versus real images). In other words, you can’t focus the image on a screen at the place the image appears to come from.

Here’s a set of observations you can make about an image formed in a plane mirror, besides the fact that it’s virtual:

- The image is upright.
- The image is the same size as the object.
- The image is located as far behind the mirror as the object is in front of the mirror.
- The image is flipped back to front (see the nearby sidebar “Reversing a mirror myth: the left-right flip”).
Part III: Catching On to Waves: The Sound and Light Kinds

Finding the mirror size

Many stores sell full-length mirrors, but that may be more about making profit than about image-formation necessity. In this section, I show that a plane mirror only needs to be one-half your actual height to let you see yourself fully in a mirror. To see this, start with a picture — Figure 10-3:

✓ A person (represented by a thick black line) is standing to the left of a plane mirror. The line representing the person includes points labeling the position of the top of the head \(T\), the eyes \(E\), and the feet \(F\). (Note: To make the diagram clearer, the position of the eyes is shown much lower than it actually is — unless the person is wearing a top hat!)

✓ The vertical gray shaded line in the center represents a full-length plane mirror.

✓ The light rays leaving the person reflect from the mirror, creating an image, which is shown on the right as another thick, black line. The image has corresponding points labeled, showing the position of the image’s top of the head \(T’\), eyes \(E’\), and feet \(F’\).

The points A and B show where on the mirror’s surface the person sees the top of his or her head and feet, respectively. You can already tell that you don’t need the whole length of this mirror to see all of yourself, because the distance \(AB\) is much less than the length of the mirror, \(CD\). With a little geometry, you can work out exactly how big \(AB\) is — that is, how much of the full-length mirror you really need.

Reversing a mirror myth: The left-right flip

If you hold your right hand up to a mirror, you find that its image looks like a left hand (so you wouldn’t be able to shake hands with this image — even if it were real and the mirror weren’t in the way!). You may wonder why the image flips left and right without also flipping up and down. But it actually does neither of these. To see why, try the following experiment:

✓ Stand in front of a mirror and point your finger to the left; you see the image of your hand also pointing to the left, parallel to your pointing direction. Therefore, left and right are not flipped.

✓ Now point your finger straight up, and you see the image of your hand pointing straight up, parallel to your pointing finger. So up and down aren’t flipped, either.

✓ Now try pointing away from yourself, straight into the mirror, and you see your image pointing straight toward you, out of the mirror — the complete opposite direction! So you can say that the mirror flips back to front.
Look again at Figure 10-3. You can see that the mirror is obeying the law of reflection: The angles of incidence and reflection from the rays from the top of your head, $\alpha$, are equal. Then, because $T'E$ is a straight line, the angle that it makes with your image must also be $\alpha$. This means that the triangle $T'ET'$ is similar to triangle $T'AD$ — because they’re both right triangles that also share the angle $\alpha$.

You also know that the image is as far behind the mirror as you are in front of it, so $T'A$ is half the length of $TE$. Because you have similar triangles, that means triangle $T'AD$ is half the size of $TET'$ — which means that $AD$ is half the length of $ET$:

$$AD = \frac{ET}{2}$$

You can do exactly the same thing for triangles $F'EF$ and $F'BC$, because you can see that they’re similar for the same reason. So $BC$ is half the length of $EF$:

$$BC = \frac{EF}{2}$$

Now all you have to do to find the length of mirror you actually need is to subtract these two unused lengths ($AD$ and $BC$) from the total ($CD$):

$$AB = CD – AD – BC$$
Part III: Catching On to Waves: The Sound and Light Kinds

Now try some numbers. If your height ($TF$) is 1.66 meters, how big does the mirror have to be so you can see your full length in the mirror? The full-length mirror, $CD$, is likewise 1.66 meters. Suppose your eyes are 0.06 meters below the top of your head ($ET$). That means the distance from eyes to feet is then $TF - ET = 1.60$ meters. You can now work out the length of the part of the mirror that you use:

$$AB = 1.66 \text{ m} - \frac{0.06 \text{ m}}{2} - \frac{1.60 \text{ m}}{2}$$
$$= 1.66 \text{ m} - 0.03 \text{ m} - 0.80 \text{ m}$$
$$= 0.83 \text{ m}$$

This is half your height — you don’t need a full-length mirror; you only need a half-length mirror.

Note that you may have seen this much more quickly by noticing that the triangle $T'F'E$ is similar to triangle $ABE$ and that $ABE$ must be half the size of $T'F'E$ (because the image is as far behind the mirror as the object is in front). Therefore, $AB$ must be half the length of $TF$. Then, because your image is the same size as you are, $AB$ is just half your height!

Working with Spherical Mirrors

A plane mirror makes an image that’s the same size as the original object, at a position that’s as far behind the mirror as the object is in front. When the mirror is curved, then the position, size, and orientation of the image can be very different. The inside or outside surface of a sphere creates such images. This is a convenient shape of mirror to study because it’s simple, though you usually use only part of the surface of a sphere rather than the whole.

There are only two ways of looking at spherical mirrors — as convex and concave. Remember, if you’re looking into a mirrored “cave,” that’s a concave mirror. Otherwise, it’s convex.

Like lenses (see Chapter 9), spherical mirrors have a center of curvature. That’s the center of the sphere that the mirror was cut from, and it’s marked $C$ in Figure 10-4. The distance from the center of curvature, $C$, to the mirror is called the radius of curvature, $R$. There’s also a focal point, marked $F$. The focal point is where light rays that come in horizontally from the left end up being focused. The focal length is half of the radius of curvature, or looked at another way, $R = 2f$.

How do you handle spherical mirrors? You can draw ray diagrams that trace how several light rays travel from an object, bounce off the mirror, and end up forming an image (just as for lenses, which I cover in Chapter 9). In this section, I show you how to draw ray diagrams for both concave and convex mirrors.
Finding practical uses for curved mirrors

Spherical mirrors are used in many everyday devices, such as magnifying makeup mirrors and the security mirrors in shops. They’re also used to turn the light from a light bulb into a beam in car headlights and flashlights. There’s even a legend that Archimedes (the famous Greek mathematician who ran down the street shouting “Eureka!”) had an idea to use curved mirrors as a weapon of war, focusing the sun’s rays onto enemy ships and setting them on fire!

Curved mirrors are also used to make the largest telescopes in the world. Why mirrors? Because it’s easier to build a large mirror than a large lens. Not only do you have to shape only one side, but also you can support the large mirror all along its unsilvered side to stop it from curving further under its own weight.

Mirrors in telescopes aren’t quite spherical. Anyone who has looked at his or her reflection in the back of a spoon or visited a hall of mirrors knows how the curves of a mirror can create very distorted images. When objects are very far from a spherical mirror, then the distortion is very small, but for the very fine level of precision required for astronomy, these distortions are too large, and corrections to the spherical curve are made to improve the image.

Figure 10-4: A spherical mirror.
Getting the inside scoop on concave mirrors

For a concave mirror, the part of the mirror that does the reflecting is on the inside of the spherical mirror. For concave mirrors, three different cases yield different types of images:

✓ The object is out farther than the center of curvature.
✓ The object is between the center of curvature and the focal point.
✓ The object is located between the focal point and the mirror itself.

This section takes a look at the various possibilities, starting by placing the object beyond the center of curvature and finding where the image forms.

Object farther out than the center of curvature

Figure 10-5 shows an object (represented by the thick arrow) being reflected in a concave mirror. Look at the ray diagram to see where the image will form in this situation and whether it’s upright or upside down. Here’s how the rays work:

✓ Ray 1: The first ray goes from the tip of the object to the mirror, where it bounces off and then goes through the center of curvature. Obviously, the center of curvature of a sphere is the center of the sphere, and any straight line passing through the center of a sphere is normal (perpendicular) to its surface, so this light ray strikes the mirror with an angle of incidence of zero. The angle of reflection is the same, so the ray is just sent back the way it came.

✓ Ray 2: The second ray goes from the tip of the object through the focal point, and then it gets reflected in a horizontal direction — that’s the key for Rays 2 and 3: These rays alternate between going through the focal point and going horizontally.

✓ Ray 3: The third ray starts off from the tip of the object in a horizontal direction, bounces off the mirror, and ends up going through the focal point.

The rays meet to form an image that’s inverted with respect to the object, between the radius of curvature and the focal point.

Is this image real or virtual? It’s real, because it forms on the side of the mirror where the object is — that’s where the rays are present physically (virtual images form on the other side of the mirror, where no light rays from the object are actually present). If you bring a screen up to the location of the image, you see the image focused there — that’s what makes it a real image.
**Object between the center of curvature and the focal point**

Figure 10-6 shows an object being reflected in a concave mirror when the object is placed between the center of curvature and the focal point. Here’s how to draw the three rays in the figure:

- **Ray 1:** The first ray goes from the tip of the object through the center of curvature to the mirror, where it’s reflected back on its same path.

- **Ray 2:** The second ray travels from the tip of the object horizontally until it hits the mirror. Then it’s reflected — and as is usual for rays that hit the mirror horizontally, it gets reflected through the focal point.

- **Ray 3:** The third ray travels from the tip of the object through the focal point, then to the mirror. When it’s reflected from the mirror, the ray is traveling horizontally.

What’s the net result? As you can see in Figure 10-6, the image is real (on the same side of the mirror as the object), inverted with respect to the object, and out past the center of curvature.

**Object between the focal point and the mirror**

Now for something really different — you end up with a virtual image in this case. If you place an object between the focal point of a spherical mirror and the mirror itself, all the rules change, because rays from the object can’t pass through the focal point and then bounce off the mirror anymore.
Figure 10-6: An object between the center of curvature and the focal point.

As you can see in Figure 10-7, you’re dealing with three rays:

- **Ray 1**: The first ray goes from the tip of the object to the center of curvature, and you extend the ray back to the mirror to complete this ray.

- **Ray 2**: The second ray goes from the tip of the object horizontally to the mirror — then it reflects from the mirror and goes through the focal point.

- **Ray 3**: The third ray is the tricky one. Normally, this ray goes from the tip of the object through the focal point and ends up going horizontally, but that’s not going to work here, because if you send this ray through the focal point, it’ll never hit the mirror. Instead, you send this ray from the tip of the object to the mirror as though it were coming from the focal point, as you can see in Figure 10-7. That does the trick.

Where do these rays come together? That’s a trick question, because they don’t come together at all — you have to extend the reflected rays behind the mirror itself. And with mirrors, that’s the mark of a virtual image (that is, no light rays from the object penetrate behind the mirror, so the image that forms there is not actually caused by light rays that meet there — you can’t bring a screen there and focus the image). So the image is virtual — and upright and magnified — as you see in the figure.

So the next time you’re creating a salad in a mirrored metal bowl, take a look at what happens when you bring the lettuce close to the metal. As you pass the focal point, the image of the lettuce suddenly snaps into focus as upright and enlarged — and you get an image of some mega-sized lettuce.