MICHAEL A. SEEDS
DANA E. BACKMAN

HORIZONS

EXPLORING THE UNIVERSE | 11E

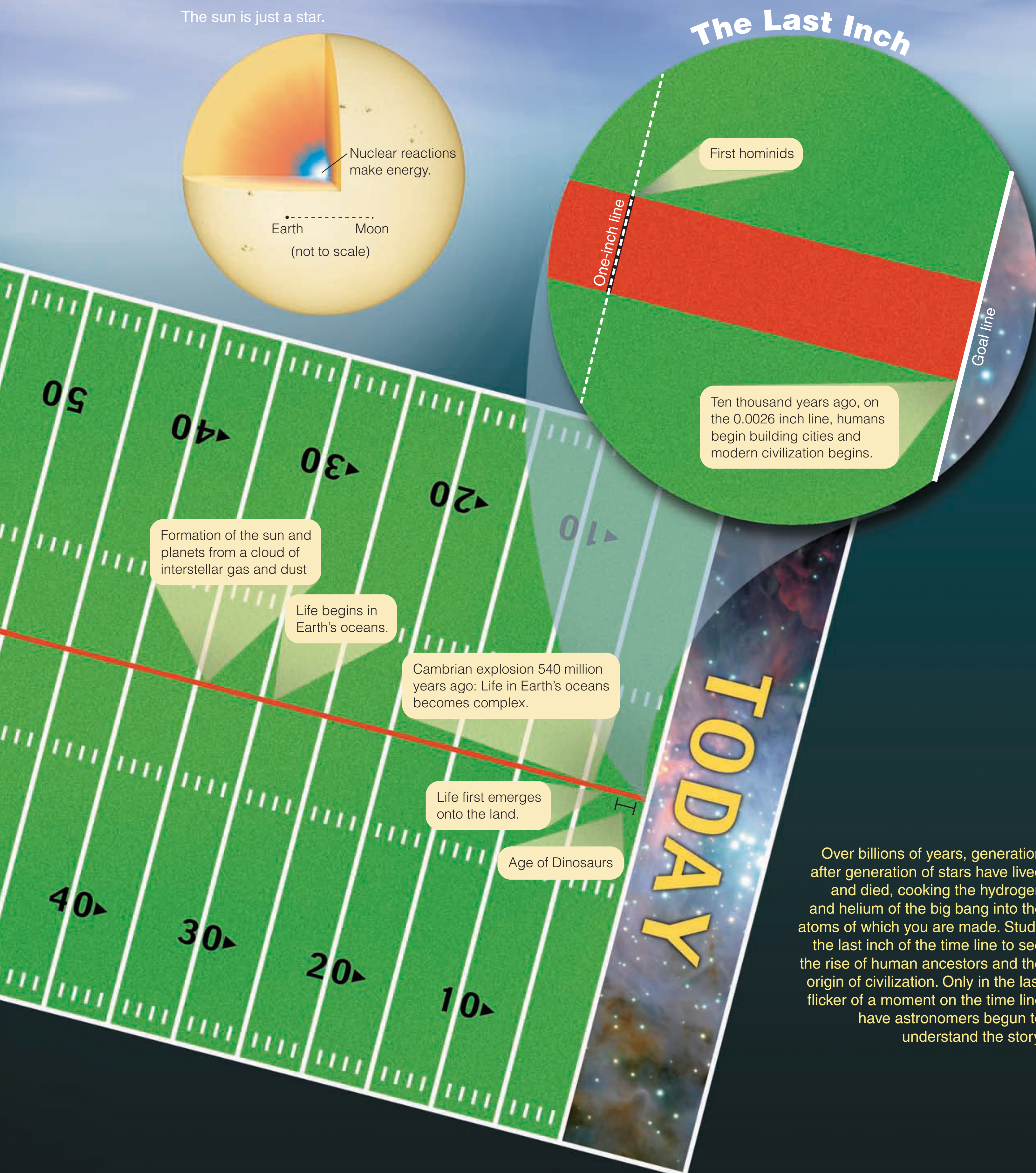
INTERSTELLAR CLOUDS AND ORGANIC MOLECULES

GAMMA RAY BURSTS FROM THE FIRST STARS

DIGGING FOR WATER ON MARS

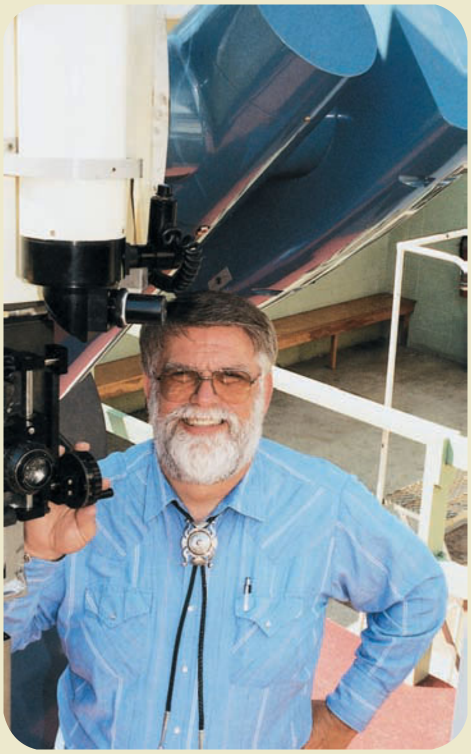
Imagine the history of the universe as a time line down the middle of a football field. The story begins on one goal line as the big bang fills the universe with energy and a fantastically hot gas of hydrogen and helium. Follow the history from the first inch of the time line as the expansion of the universe cools the gas and it begins to form galaxies and stars.

A typical galaxy contains 100 billion stars.



Over billions of years, generation after generation of stars have lived and died, cooking the hydrogen and helium of the big bang into the atoms of which you are made. Study the last inch of the time line to see the rise of human ancestors and the origin of civilization. Only in the last flicker of a moment on the time line have astronomers begun to understand the story.

About the Authors



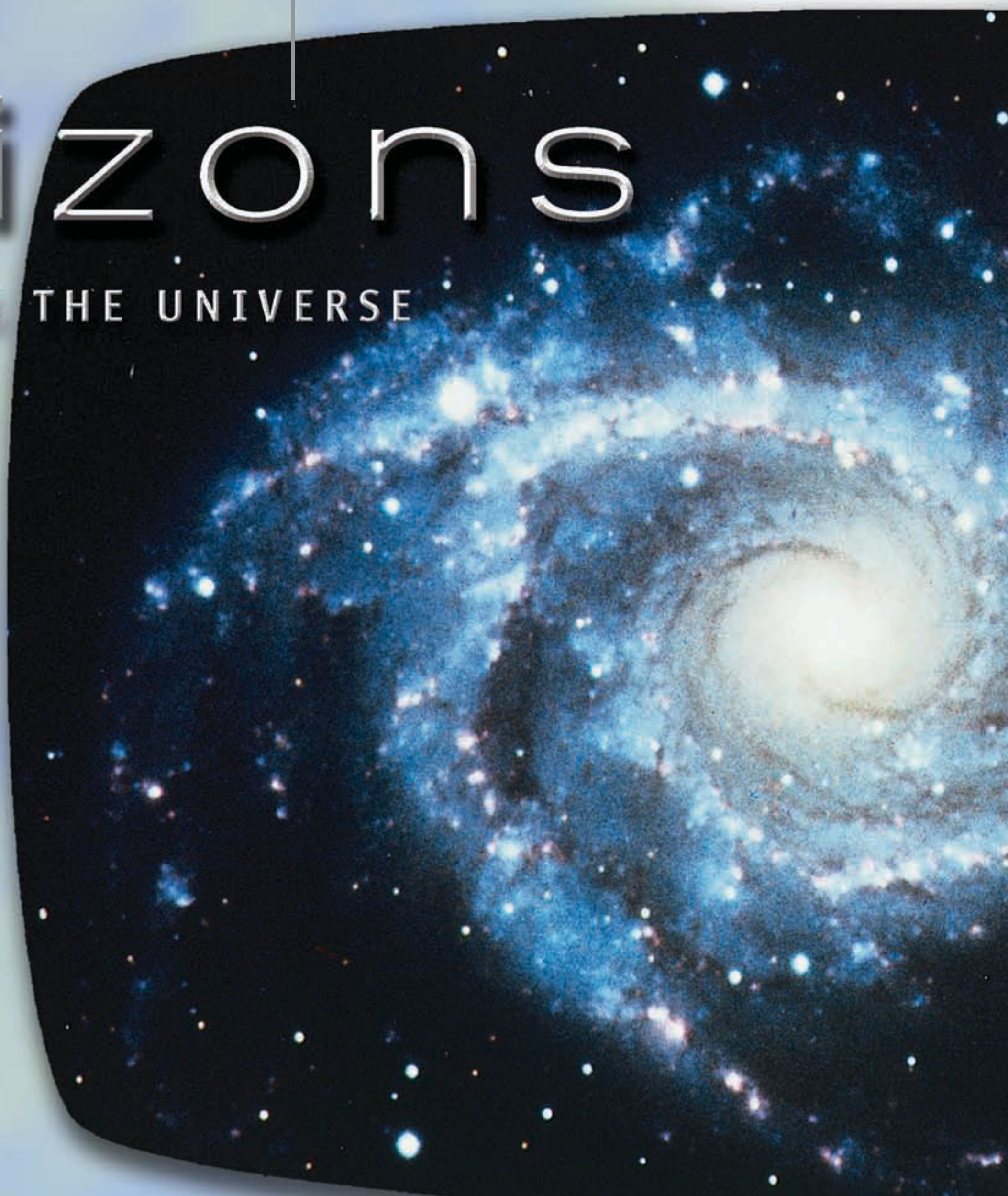
Mike Seeds has been a Professor of Physics and Astronomy at Franklin and Marshall College in Lancaster, Pennsylvania, since 1970. In 1989 he received F&M College's Lindback Award for Distinguished Teaching. Mike's love for the history of astronomy led him to create upper-level courses on "Archaeoastronomy" and "Changing Concepts of the Universe." His research interests focus on variable stars and the automation of astronomical telescopes. Mike is the author of *Horizons: Exploring the Universe,* Eleventh Edition (2010); *Astronomy: The Solar System and Beyond,* Sixth Edition (2010); *Foundations of Astronomy,* Tenth Edition (2008); and *Perspectives on Astronomy* (2008), all published by Brooks/Cole. He was Senior Consultant for creation of the 20-episode telecourse accompanying his book *Horizons: Exploring the Universe.*

Dana Backman taught in the physics and astronomy department at Franklin and Marshall College in Lancaster, Pennsylvania, from 1991 until 2003. He invented and taught a course titled "Life in the Universe" in F&M's interdisciplinary Foundations program. Dana now teaches introductory astronomy, astrobiology, and cosmology courses in Stanford University's Continuing Studies Program. His research interests focus on infrared observations of planet formation, models of debris disks around nearby stars, and evolution of the solar system's Kuiper Belt. Dana is the author of the first edition of *Perspectives on Astronomy* (2008); *Horizons: Exploring the Universe,* Eleventh Edition (2010); and *Astronomy: The Solar System and Beyond,* Sixth Edition (2010), all published by Brooks/Cole. He is with the SETI Institute in Mountain View, California, in charge of the education and public outreach program for SOFIA (Stratospheric Observatory for Infrared Astronomy) at NASA's Ames Research Center.

11 | ELEVENTH EDITION

Horizons

EXPLORING THE UNIVERSE



Michael A. Seeds

Joseph R. Grundy Observatory
Franklin and Marshall College

Dana E. Backman

Stratospheric Observatory for Infrared
Astronomy (SOFIA)
SETI Institute

BROOKS/COLE
CENGAGE Learning™

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

BROOKS/COLE
CENGAGE Learning™

Horizons: Exploring the Universe,
Eleventh Edition
Michael A. Seeds, Dana E. Backman

Astronomy Editor: Kilian Kennedy

Publisher: Mary Finch

Development Editor: Teri Hyde

Editorial Assistant: Joshua Duncan

Media Editor: Rebecca Berardy Schwartz

Marketing Manager: Nicole Mollica

Marketing Assistant: Elizabeth Wong

Marketing Communications Manager: Belinda Krohmer

Project Manager, Editorial Production: Hal Humphrey

Art Director: John Walker

Print Buyer: Karen Hunt

Permissions Editor: Timothy Sisler

Production Service: Graphic World Inc.

Text Designer: Liz Harasymczuk, Linda Beaupré

Photo Researcher: Kathleen Olson

Copy Editor: Margaret Pinette

Illustrator: Precision Graphics

Cover Designer: Irene Morris

Cover Images: *Background:* Young stars in the Rho Ophiuchi Cloud (NASA/JPL-Caltech/Harvard-Smithsonian CfA). *Top inset:* Nebula in the Large Magellanic Cloud (NASA, ESA, and the Hubble Heritage Team STScI/Aura). *Middle:* Gamma-ray burst (NASA/D. Berry). *Bottom:* Phoenix Mars Lander (NASA/JPL-Caltech/University of Arizona).

Compositor: Graphic World Inc.

Library of Congress Control Number: 2008937384

Student Edition:

ISBN-13: 978-0-495-55973-3

ISBN-10: 0-495-55973-3

Brooks/Cole
10 Davis Drive
Belmont, CA 94002-3098
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at **international.cengage.com/region.**

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

For your course and learning solutions, visit **academic.cengage.com.**

Purchase any of our products at your local college store or at our preferred online store **www.ichapters.com.**

Printed in the United States of America
1 2 3 4 5 6 7 12 11 10 09

For Janet and Jamie

Brief Contents

Part 1: The Sky

CHAPTER 1 HERE AND NOW 1
CHAPTER 2 THE SKY 10
CHAPTER 3 CYCLES OF THE SKY 21
CHAPTER 4 THE ORIGIN OF MODERN ASTRONOMY 42
CHAPTER 5 LIGHT AND TELESCOPES 69

Part 2: The Stars

CHAPTER 6 ATOMS AND STARLIGHT 94
CHAPTER 7 THE SUN 112
CHAPTER 8 THE FAMILY OF STARS 134
CHAPTER 9 THE FORMATION AND STRUCTURE OF STARS 158
CHAPTER 10 THE DEATHS OF STARS 184
CHAPTER 11 NEUTRON STARS AND BLACK HOLES 210

Part 3: The Universe of Galaxies

CHAPTER 12 THE MILKY WAY GALAXY 233
CHAPTER 13 GALAXIES 259
CHAPTER 14 ACTIVE GALAXIES AND SUPERMASSIVE BLACK HOLES 282
CHAPTER 15 MODERN COSMOLOGY 295

Part 4: The Solar System

CHAPTER 16 THE ORIGIN OF THE SOLAR SYSTEM 320
CHAPTER 17 THE TERRESTRIAL PLANETS 343
CHAPTER 18 THE JOVIAN PLANETS, PLUTO, AND THE KUIPER BELT 378
CHAPTER 19 METEORITES, ASTEROIDS AND COMETS 407

Part 5: Life

CHAPTER 20 LIFE ON OTHER WORLDS 425

Contents

Part 1: The Sky

Chapter 1 | Here and Now 1

1-1 WHERE ARE WE? 2
1-2 WHEN IS NOW? 6
1-3 WHY STUDY ASTRONOMY? 7

Chapter 2 | The Sky 10

2-1 THE STARS 11
2-2 THE SKY AND ITS MOTION 14

Chapter 3 | Cycles of the Sky 21

3-1 CYCLES OF THE SUN 22
3-2 ASTRONOMICAL INFLUENCES ON EARTH'S CLIMATE 26
3-3 THE CYCLES OF THE MOON 29

Chapter 4 | The Origin of Modern Astronomy 42

4-1 CLASSICAL ASTRONOMY 43
4-2 COPERNICUS 46
4-3 PLANETARY MOTION 49
4-4 GALILEO GALILEI 55
4-5 ISAAC NEWTON AND ORBITAL MOTION 58

Chapter 5 | Light and Telescopes 69

5-1 RADIATION: INFORMATION FROM SPACE 70
5-2 OPTICAL TELESCOPES 72
5-3 SPECIAL INSTRUMENTS 84
5-4 RADIO TELESCOPES 86
5-5 ASTRONOMY FROM SPACE 89

Reasoning with Numbers

2-1 Magnitudes 15
3-1 The Small-Angle Formula 34
4-1 Circular Velocity 61
5-1 The Powers of a Telescope 77

How Do We Know?

1-1 The So-Called Scientific Method 7
2-1 Scientific Models 18
3-1 Pseudoscience 26
3-2 Evidence as the Foundation of Science 28
3-3 Scientific Arguments 29
4-1 Scientific Revolutions 49
4-2 Hypothesis, Theory, and Law 53
4-3 Cause and Effect 60
4-4 Testing a Theory by Prediction 65
5-1 Resolution and Precision 76

Concept Art Portfolios

The Sky Around You 16–17
The Cycle of the Seasons 24–25
The Phases of the Moon 32–33
The Ancient Universe 44–45
Orbiting Earth 62–63
Modern Astronomical Telescopes 80–81

Part 2: The Stars

Chapter 6 | Atoms and Starlight 94

6-1 ATOMS 95
6-2 THE INTERACTION OF LIGHT AND MATTER 97
6-3 STELLAR SPECTRA 101

Chapter 7 | The Sun 112

7-1 THE SOLAR ATMOSPHERE 113
7-2 NUCLEAR FUSION IN THE SUN 119
7-3 SOLAR ACTIVITY 123

Chapter 8 | The Family of Stars 134

8-1 MEASURING THE DISTANCES TO STARS 135
8-2 INTRINSIC BRIGHTNESS 136
8-3 THE DIAMETERS OF STARS 139
8-4 THE MASSES OF STARS 145
8-5 A SURVEY OF THE STARS 151

Chapter 9 | The Formation and Structure of Stars 158

9-1 THE BIRTH OF STARS 159
9-2 FUSION IN STARS 171
9-3 STELLAR STRUCTURE 174
9-4 MAIN-SEQUENCE STARS 178

Chapter 10 | The Deaths of Stars 184

10-1 GIANT STARS 185
10-2 THE DEATHS OF LOWER-MAIN-SEQUENCE STARS 190
10-3 THE EVOLUTION OF BINARY SYSTEMS 197
10-4 THE DEATHS OF MASSIVE STARS 201

Chapter 11 | Neutron Stars and Black Holes 210

11-1 NEUTRON STARS 211
11-2 BLACK HOLES 222
11-3 COMPACT OBJECTS WITH DISKS AND JETS 227

Reasoning with Numbers

6-1 Blackbody Radiation 100
6-2 The Doppler Formula 109
7-1 Hydrogen Fusion 120
8-1 Parallax and Distance 137
8-2 Absolute Magnitude 139
8-3 Luminosity, Radius, and Temperature 140
8-4 The Masses of Binary Stars 146
8-5 The Mass–Luminosity Relation 153
9-1 The Life Expectancies of Stars 181

How Do We Know?

6-1 Quantum Mechanics 97
7-1 Scientific Confidence 123
7-2 Confirmation and Consolidation 129
8-1 Chains of Inference 146
8-2 Basic Scientific Data 152
9-1 Separating Facts from Theories 167
9-2 Mathematical Models 177
10-1 Toward Ultimate Causes 188
11-1 Theories and Proof 221
11-2 Checks on Fraud in Science 225

Concept Art Portfolios

Atomic Spectra 102–103
Sunspots and the Sunspot Cycle 126–127
Magnetic Solar Phenomena 130–131
The Family of Stars 154–155
Three Kinds of Nebulae 160–161
Observational Evidence of Star Formation 168–169
Star Formation in the Orion Nebula 172–173
Star Cluster H–R Diagrams 192–193
The Formation of Planetary Nebulae 194–195
The Lighthouse Model of a Pulsar 214–215

Celestial Profile 1 | The Sun 113

Part 3: The Universe of Galaxies

Chapter 12 | The Milky Way Galaxy 233

12-1 THE DISCOVERY OF THE GALAXY 234
12-2 THE ORIGIN OF THE MILKY WAY 243
12-3 THE NUCLEUS 248
12-4 SPIRAL ARMS AND STAR FORMATION 248

Chapter 13 | Galaxies 259

13-1 THE FAMILY OF GALAXIES 260
13-2 MEASURING THE PROPERTIES OF GALAXIES 265
13-3 THE EVOLUTION OF GALAXIES 273

Chapter 14 | Active Galaxies and Supermassive Black Holes 282

14-1 ACTIVE GALACTIC NUCLEI 283
14-2 SUPERMASSIVE BLACK HOLES 289

Chapter 15 | Modern Cosmology 295

15-1 INTRODUCTION TO THE UNIVERSE 296
15-2 THE BIG BANG THEORY 300
15-3 SPACE, TIME, MATTER, AND ENERGY 306
15-4 21ST-CENTURY COSMOLOGY 310

Reasoning with Numbers

13-1 The Hubble Law 269
15-1 The Age of the Universe 300

How Do We Know?

12-1 Calibration 238
12-2 Nature as Processes 245
13-1 Classification in Science 261
13-2 Selection Effects 264
14-1 Statistical Evidence 284
15-1 Reasoning by Analogy 298
15-2 Science: A System of Knowledge 305

Concept Art Portfolios

Sagittarius A 250–251*
Galaxy Classification 262–263
Interacting Galaxies 274–275
Cosmic Jets and Radio Lobes 286–287

Part 4: The Solar System

Chapter 16 | The Origin of the Solar System 320

16-1 THE GREAT CHAIN OF ORIGINS 321
16-2 A SURVEY OF THE SOLAR SYSTEM 323
16-3 THE STORY OF PLANET BUILDING 329
16-4 PLANETS ORBITING OTHER STARS 336

Chapter 17 | The Terrestrial Planets 343

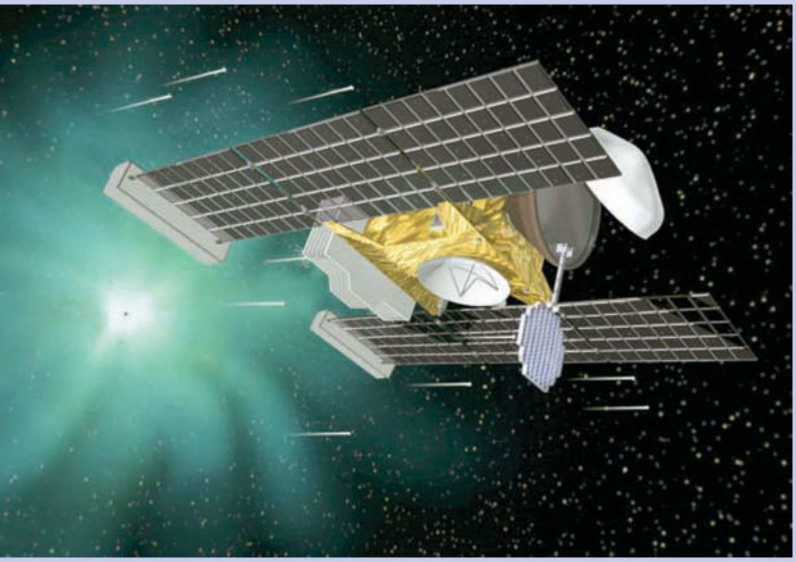
17-1 A TRAVEL GUIDE TO THE TERRESTRIAL PLANETS 344
17-2 EARTH: THE ACTIVE PLANET 345
17-3 THE MOON 352
17-4 MERCURY 358
17-5 VENUS 361
17-6 MARS 367

Chapter 18 | The Jovian Planets, Pluto, and the Kuiper Belt 378

18-1 A TRAVEL GUIDE TO THE OUTER SOLAR SYSTEM 379
18-2 JUPITER 381
18-3 SATURN 388
18-4 URANUS 394
18-5 NEPTUNE 397
18-6 PLUTO—THE FIRST DWARF PLANET 402

Chapter 19 | Meteorites, Asteroids, and Comets 407

19-1 METEOROIDS, METEORS, AND METEORITES 409
19-2 ASTEROIDS 412
19-3 COMETS 413
19-4 IMPACTS ON EARTH 420



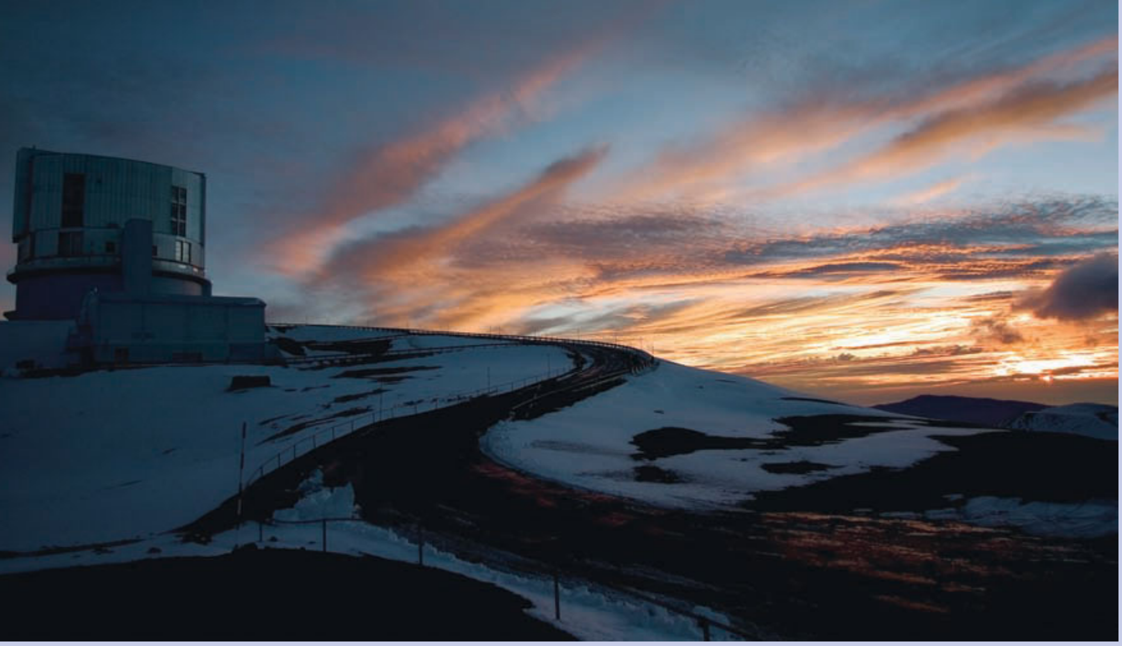
How Do We Know?

16-1 Two Kinds Of Theories: Catastrophic and Evolutionary 322
16-2 Reconstructing the Past from Evidence and Hypothesis 330
16-3 Scientists: Courteous Skeptics 340
17-1 Understanding Planets: Follow the Energy 346
17-2 Hypotheses and Theories Unify the Details 361
17-3 The Present Is the Key to the Past 374
18-1 Basic Science and Practical Technology 379
18-2 Funding for Basic Research 389
19-1 Enjoying the Natural World 410

Concept Art Portfolios

Terrestrial and Jovian Planets 326–327
The Active Earth 350–351
Impact Cratering 354–355
Volcanoes 364–365
Jupiter's Atmosphere 384–385
The Ice Rings of Saturn 392–393
The Rings of Uranus and Neptune 398–399
Observations of Asteroids 414–415
Observations of Comets 418–419

Celestial Profile 2 | Earth 347
Celestial Profile 3 | The Moon 353
Celestial Profile 4 | Mercury 359
Celestial Profile 5 | Venus 363
Celestial Profile 6 | Mars 369
Celestial Profile 7 | Jupiter 381
Celestial Profile 8 | Saturn 391
Celestial Profile 9 | Uranus 395
Celestial Profile 10 | Neptune 401



Part 5: Life

Chapter 20 | Life on Other Worlds 425

20-1 THE NATURE OF LIFE 426

20-2 THE ORIGIN OF LIFE 430

20-3 COMMUNICATION WITH DISTANT CIVILIZATIONS 437

How Do We Know?

20-1 The Nature of Scientific Explanation 426

20-2 UFOs and Space Aliens 438

Concept Art Portfolios

DNA: The Code of Life 428–429

AFTERWORD 443

APPENDIX A UNITS AND ASTRONOMICAL DATA 445

APPENDIX B OBSERVING THE SKY 453

GLOSSARY 466

ANSWERS TO EVEN-NUMBERED PROBLEMS 475

INDEX 476

A Note to the Student

From Mike and Dana

We are excited that you are taking an astronomy course and using our book. You are going to see some amazing things, from the icy rings of Saturn to monster black holes. We are proud to be your guides as you explore.

We have developed this book to help you expand your knowledge of astronomy, from recognizing the moon and a few stars in the evening sky, to a deeper understanding of the extent, power, and diversity of the universe. You will meet worlds where it rains methane, stars so dense their atoms are crushed, colliding galaxies that are ripping each other apart, and a universe that is expanding faster and faster.

Two Goals

This book is designed to help you answer two important questions:

- What are we?
- How do we know?

By the question *What are we?* we mean: How do we fit into the universe and its history? The atoms you are made of had their first birthday in the big bang when the universe began, but those atoms were cooked and remade inside stars, and now they are inside you. Where will they be in a billion years? Astronomy is the only course on campus that can tell you that story, and it is a story that everyone should know.

By the question *How do we know?* we mean: How does science work? What is the evidence, and how do you know it is true? For instance, how can anyone know there was a big bang? In today's world, you need to think carefully about the things so-called experts say. You should demand explanations. Scientists have a special way of knowing based on evidence that makes scientific knowledge much more powerful than just opinion, policy, marketing, or public relations. It is the human race's best understanding of nature. To comprehend the world around you, you need to understand how science works. Throughout this book, you will find boxes called ***How Do We Know?*** They will help you understand how scientists use the methods of science to know what the universe is like.

Expect to Be Astonished

One reason astronomy is exciting is that astronomers discover new things every day. Astronomers expect to be astonished. You can share in the excitement because we have worked hard to include new images, new discoveries, and new insights that will take you, in an introductory course, to the frontier of human knowledge. Huge telescopes on remote mountaintops and in space provide a daily dose of excitement that goes far beyond entertainment. These new discoveries in astronomy are exciting because they are about us. They tell us more and more about what we are.

As you read this book, notice that it is not organized as lists of facts for you to memorize. That could make even astronomy boring. Rather, this book is organized to show you how scientists use evidence and theory to create logical arguments that show how nature works. Look at the list of special features that follows this note. Those features were carefully designed to help you understand astronomy as evidence and theory. Once you see science as logical arguments, you hold the key to the universe.

Do Not Be Humble

As teachers, our quest is simple. We want you to understand your place in the universe—not just your location in space, but your location in the unfolding history of the physical universe. Not only do we want you to know where you are and what you are in the universe, but we want you to understand how scientists know. By the end of this book, we want you to know that the universe is very big, but that it is described and governed by a small set of rules and that we humans have found a way to figure out the rules—a method called science.

To appreciate your role in this beautiful universe, you must learn more than just the facts of astronomy. You must understand what we are and how we know. Every page of this book reflects that ideal.

Mike Seeds
mseeds@fandm.edu
Dana Backman
dbackman@sofia.usra.edu

Key Content and Pedagogical Changes for the Eleventh Edition

- Every chapter has been reorganized to focus on the two main themes of the book. The **What Are We?** boxes at the end of each chapter provide a personal link between human life and the astronomy in that chapter, including, for example, the origin of the elements, the future of exploration in the solar system, and the astronomically short span of our civilization.
- The **How Do We Know?** boxes have been rewritten to be more focused on helping you understand how science works and how scientists think about nature.
- Every chapter has been rewritten to place the "new terms" in context for you rather than as a vocabulary list. New terms are boldfaced where they are first defined in the text of the chapter and reappear in context as boldface terms in each chapter summary. Those new terms that appear in Concept Art portfolios are boldfaced in the art and are previewed in italics as the portfolios are introduced.
- Guideposts have been rewritten, shortened, and focused on a short list of essential questions that guide you to the key objectives of the chapter.
- Every chapter has been updated to include new research, images, and the latest understanding, ranging from discoveries of how planets form in dust disks around young stars to the latest insights into the nature of dark energy.

Special Features

- **What Are We?** items are short summaries at the end of each chapter to help you see how you fit in to the cosmos.
- **How Do We Know?** items are short boxes that help you understand how science works. For example, the How Do We Know? boxes discuss the difference between a hypothesis and a theory, the use of statistical evidence, and the construction of scientific models.
- **Concept Art Portfolios** cover topics that are strongly graphic and provide an opportunity for you to create your own understanding and share in the satisfaction that scientists feel as they uncover the secrets of nature. Color and numerical keys in the introduction to the portfolios guide you to the main concepts.
- **Guideposts** on the opening page of each chapter help you see the organization of the book. The Guidepost connects the chapter with the preceding and following chapters and provides you with a short list of essential questions as guides to the objectives of the chapter.
- **Scientific Arguments** at the end of many text sections are carefully designed questions to help you review and synthesize concepts from the section. An initial question and a short answer show how scientists construct scientific arguments from observations, evidence, theories, and natural laws that lead to a conclusion. A further question then gives you a chance to construct your own argument on a related issue.
- **Celestial Profiles** of objects in our solar system directly compare and contrast planets with each other. This is the way planetary scientists understand the planets, not as isolated unrelated bodies but as siblings with noticeable differences but many characteristics and a family history in common.
- **End-of-Chapter Review Questions** are designed to help you review and test your understanding of the material.
- **End-of-Chapter Discussion Questions** go beyond the text and invite you to think critically and creatively about scientific questions.

This book also offers the following online study aids as optional bundle items or for separate purchase:

- **Enhanced WebAssign.** Assign, collect, grade, and record homework via the Web with this proven system, using more than 1,000 questions both from the text and written specifically for WebAssign. Questions include animated activities, ranking tasks, multiple-choice, and fill-in-the-blank exercises.
- **Virtual Astronomy Labs.** These online labs give you an exciting, interactive way to learn, putting some of astronomy's most useful instruments into your hands—precise telescope controls to measure angular size, a photometer to measure light intensity, and a spectrograph to measure Doppler-shifted spectral lines.

Acknowledgments

Over the years we have had the guidance of a great many people who care about astronomy and teaching. We would like to thank all of the students and teachers who have contributed to this book. Their comments and suggestions have been very helpful in shaping this book.

Many observatories, research institutes, laboratories, and individual astronomers have supplied figures and diagrams for this edition. They are listed on the credits page, and we would like to thank them specifically for their generosity.

Special thanks goes to Kathryn Coolidge, who has reviewed most of the chapters word by word and has been a tremendous help with issues of organization, presentation, and writing. Jamie Backman has also been a careful reader, contributing many insights to the way the text should best be organized and presented.

We are happy to acknowledge the use of images and data from a number of important programs. In preparing materials for this book we used NASA's Sky View facility located at NASA Goddard Space Flight Center. We have used atlas images and mosaics obtained as part of the Two Micron All Sky Survey (2MASS), a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. A number of solar images are used by the courtesy of the SOHO consortium, a project of international cooperation between ESA and NASA.

It is always a pleasure to work with the Cengage team, including Hal Humphrey, Alex Brady, and photo researcher Kathleen Olson. Special thanks go to all of the people who have contributed to this project. We have enjoyed working with Margaret Pinette and Bill Heckman of Heckman & Pinette, and we appreciate their understanding and goodwill. We would like to thank Carol O'Connell of Graphic World for her help in keeping everything on track. We would especially like to thank editors Marcus Boggs and Teri Hyde for their help and guidance throughout this project.

Most of all, we would like to thank our families for putting up with "the books." They know all too well that textbooks are made of time.

Reviewers

We would especially like to thank the following reviewers, whose careful analysis and thoughtful suggestions have been invaluable in completing this new edition:

Scott Hildreth, *Chabot College*
Andrea N Lommen, *Franklin and Marshall College*
Chris McKay, *NASA Ames*
Scott Miller, *Pennsylvania State University*
Luisa Rebull, *California Institute of Technology*
Ata Sarajedini, *University of Florida*
Larry C. Sessions, *Metropolitan State College of Denver*

This page intentionally left blank

1 Here and Now

Guided by detailed observations and calculations, an artist interprets the birth of a cluster of stars deep inside the nebula known as the Lynx Arc. Light from these stars traveled through space for 12 billion years before reaching Earth. (ESA/Space Telescope—European Coordinating Facility, Germany)

Guidepost

As you study astronomy, you will learn about yourself. You are a planetwalker, and this chapter will give you a preview of what it means to live on a planet that whirls around a star that drifts through a universe of other stars and galaxies. You owe it to yourself to know where you are. That is the first step to knowing what you are.

In this chapter, you will meet three essential questions about astronomy:

Where are you in the universe?

How does human history fit on the time scale of the universe?

Why should you study astronomy?

As you study astronomy, you will see how science gives you a way to know how nature works. In this chapter, you can begin by thinking about science in a general way. Later chapters will give you more specific insights into how scientists work and think and know about nature.

This chapter is just a jumping-off place. From here onward you will be exploring deep space and deep time. The next chapter begins your journey by looking at the night sky as seen from Earth.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

The longest journey begins with a single step.

LAO TSE

1-1 Where Are We?

As you study astronomy, you are learning about yourself, and knowing where you are in space and time is a critical part of the story of astronomy. To find yourself among the stars, you can take a cosmic zoom, a ride out through the universe to preview the kinds of objects you are about to study.

You can begin with something familiar. ■ Figure 1-1 shows a region about 50 feet across occupied by a human being, a sidewalk, and a few trees—all objects whose size you can understand. Each successive picture in this cosmic zoom will show you a region of the universe that is 100 times wider than the preceding picture. That is, each step will widen your **field of view,** the region you can see in the image, by a factor of 100.

Widening your field of view by a factor of 100 allows you to see an area 1 mile in diameter (■ Figure 1-2). People, trees, and sidewalks have become too small to see, but now you see a college campus and surrounding streets and houses. The dimensions of houses and streets are familiar. This is still the world you know.

Before leaving this familiar territory, you should make a change in the units you use to measure sizes. Astronomers, as do all scientists, use the metric system of units because it is well understood worldwide and, more importantly, because it simplifies calculations. If you are not already familiar with the metric system, or if you need a review, study Appendix A before reading on.

The photo in Figure 1-2 is 1 mile across, which equals 1.609 kilometers. You can see that a kilometer (abbreviated km) is a bit under two-thirds of a mile—a short walk across a neighborhood. But when you expand your field of view by a factor of 100, the neighborhood you saw in the previous photo has vanished (■ Figure 1-3). Now your field of view is 160 km wide, and you see cities and towns as patches of gray. Wilmington, Delaware, is visible at the lower right. At this scale, you can see the natural features of Earth's surface. The Allegheny Mountains of southern Pennsylvania cross the image in the upper left, and the Susquehanna River flows southeast into Chesapeake Bay. What look like white bumps are a few puffs of clouds.

Figure 1-3 is an infrared photograph, which is why healthy green leaves and crops show up as red. Human eyes are sensitive to only a narrow range of colors. As you explore the universe, you will learn to use a wide range of other "colors," from X-rays to radio waves, to reveal sights invisible to unaided human eyes. You will learn much more about infrared, X-rays, and radio energy in later chapters.

■ **Figure 1-1**

Michael A. Seeds

■ **Figure 1-2**

This box ■ represents the relative size of the previous frame. (USGS)

■ **Figure 1-3**

NASA

■ **Figure 1-4**

NASA

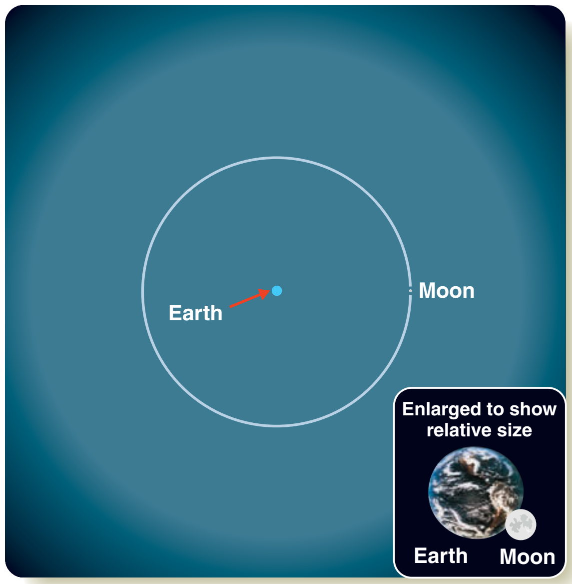
When you once again enlarge your field of view by a factor of 100, Earth, the moon, and the moon's orbit all lie in the small red box at lower left of ■ Figure 1-6. Now you can see the sun and two other planets that are part of our solar system. Our **solar system** consists of the sun, its family of planets, and some smaller bodies such as moons and comets.

Like Earth, Venus and Mercury are **planets,** small, spherical, nonluminous bodies that orbit a star and shine by reflected light. Venus is about the size of Earth, and Mercury is just over a third of Earth's diameter. On this diagram, they are both too small to be seen as anything but tiny dots. The sun is a **star,** a self-luminous ball of hot gas that generates its own energy. Even though the sun is 109 times larger in diameter than Earth (inset), it too is nothing more than a dot in this diagram.

This diagram represents an area with a diameter of 1.6×10^8 km. One way astronomers simplify calculations using large numbers is to define larger units of measurement. The average distance from Earth to the sun is a unit of distance called the **astronomical unit (AU),** a distance of 1.5×10^8 km. Now you can see that the

■ **Figure 1-5**

NASA



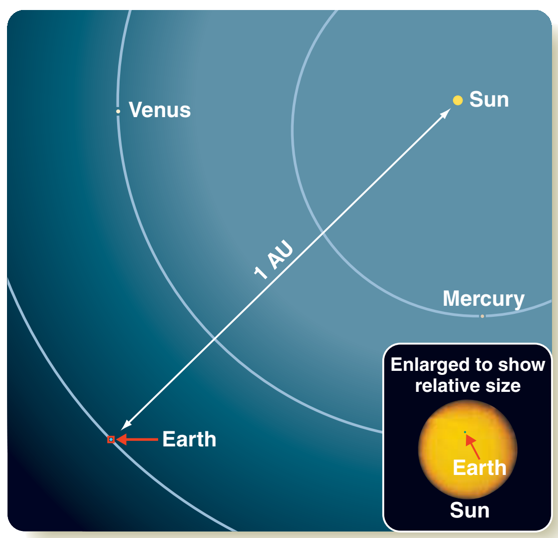
At the next step in your journey, you can see your entire planet, which is nearly 13,000 km in diameter (■ Figure 1-4). The photo shows most of the daylight side of the planet. Earth rotates on its axis once a day, exposing half of its surface to daylight at any particular moment. It is the rotation of the planet that causes the cycle of day and night. The rotation of Earth carries you eastward, and as you cross into darkness, you see the sun set in the west. The blurriness you see at the extreme right of the photo is the boundary between day and night—the sunset line. This is a good example of how a photo can give you visual clues to understanding a concept. Special questions called "Learning to Look" at the end of each chapter give you a chance to use your own imagination to connect images with the theories that describe astronomical objects.

Enlarge your field of view by a factor of 100, and you see a region 1,600,000 km wide (■ Figure 1-5). Earth is the small blue dot in the center, and the moon, whose diameter is only one-fourth that of Earth, is an even smaller dot along its orbit 380,000 km away.

These numbers are so large that it is inconvenient to write them out. Astronomy is sometimes known as the science of big numbers, and soon you will use numbers much larger than these to discuss the universe. Rather than writing out these numbers as in the previous paragraph, it is convenient to write them in **scientific notation.** This is nothing more than a simple way to write very big or very small numbers without using lots of zeros. In scientific notation, 380,000 becomes 3.8×10^5. If you are not familiar with scientific notation, read the section on powers of 10 notation in the Appendix. The universe is too big to discuss without using scientific notation.

■ **Figure 1-6**

NOAO



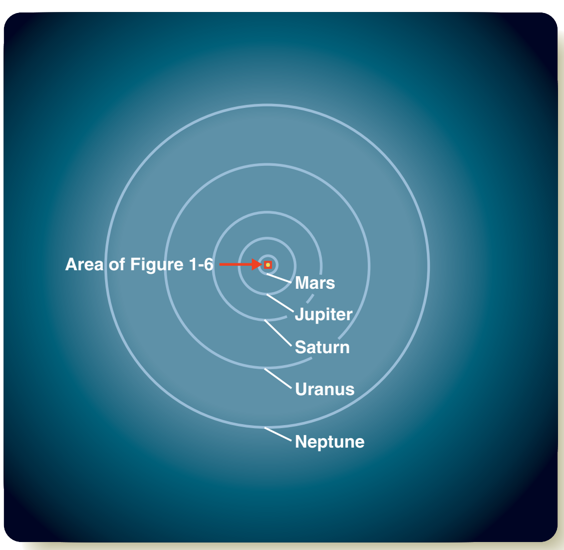
average distance from Venus to the sun is about 0.72 AU, and the average distance from Mercury to the sun is about 0.39 AU.

These distances are averages because the orbits of the planets are not perfect circles. This is particularly apparent in the case of Mercury. Its orbit carries it as close to the sun as 0.307 AU and as far away as 0.467 AU. You can see the variation in the distance from Mercury to the sun in Figure 1-6. Earth's orbit is more circular, and its distance from the sun varies by only a few percent.

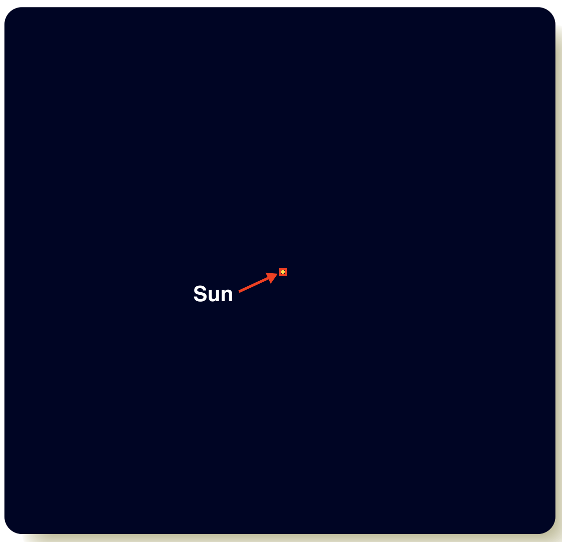
Enlarge your field of view again, and you can see the entire solar system (■ Figure 1-7). The details of the preceding figure are now lost in the red square at the center of this diagram. You see only the brighter, more widely separated objects. The sun, Mercury, Venus, and Earth lie so close together that you cannot see them separately at this scale. Mars, the next planet outward, lies only 1.5 AU from the sun. In contrast, Jupiter, Saturn, Uranus, and Neptune are farther away and so are easier to place in this diagram. They are cold worlds far from the sun's warmth. Light from the sun reaches Earth in only 8 minutes, but it takes over 4 hours to reach Neptune.

When you again enlarge your field of view by a factor of 100, the solar system vanishes (■ Figure 1-8). The sun is only a point of light, and all the planets and their orbits are now crowded into the small red square at the center. The planets are too small and too faint to be visible so near the brilliance of the sun.

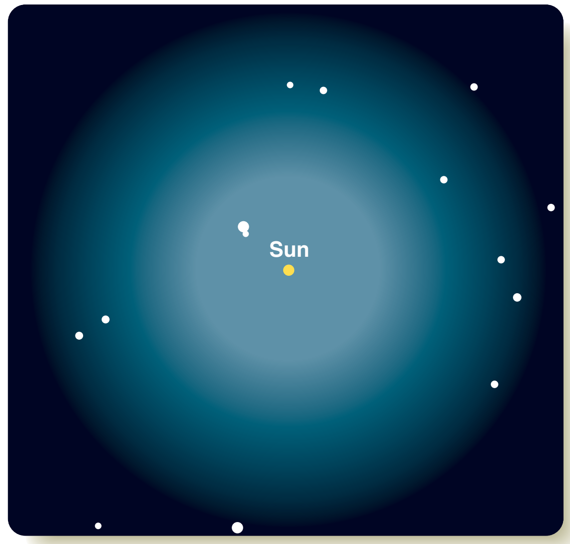
■ **Figure 1-7**



■ **Figure 1-8**



■ **Figure 1-9**



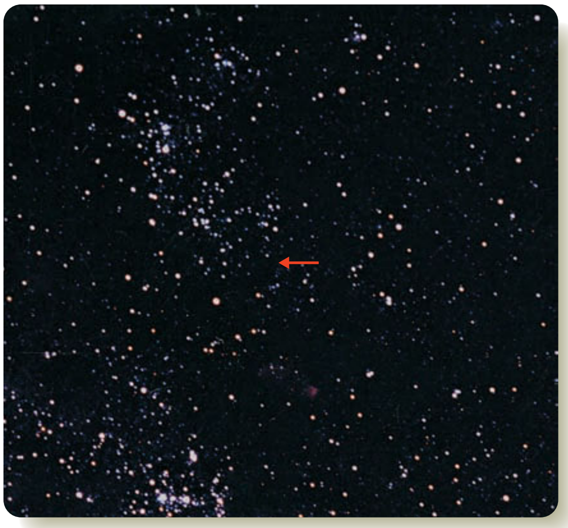
Nor are any stars visible except for the sun. The sun is a fairly typical star, and it seems to be located in a fairly average neighborhood in the universe. Although there are many billions of stars like the sun, none are close enough to be visible in this diagram, which shows a region only 11,000 AU in diameter. The stars are typically separated by distances about 10 times larger than the distance represented by the diameter of this diagram.

In ■ Figure 1-9, your field of view has expanded to a diameter of a bit over 1 million AU. The sun is at the center, and at this scale you can see a few of the nearest stars. These stars are so distant that it is not reasonable to give their distances in astronomical units. To express distances so large, astronomers define a new unit of distance, the light-year. One **light-year (ly)** is the distance that light travels in one year, roughly 10^{13} km or 63,000 AU. It is a **Common Misconception** that a light-year is a unit of time, and you can sometimes hear the term misused in science fiction movies and TV shows. The next time you hear someone say, "It will take me light-years to finish my history paper," you can tell that person that a light-year is a distance, not a time. The diameter of your field of view in Figure 1-9 is 17 ly.

Another **Common Misconception** is that stars look like disks when seen through a telescope. Although stars are roughly the same size as the sun, they are so far away that astronomers cannot see them as anything but points of light. Even the closest star to the sun—Alpha Centauri, only 4.2 ly from Earth—looks like a point of light through even the biggest telescopes on Earth. Furthermore, any planets that might circle other stars are much too small, too

Figure 1-10

NOAO



faint, and too close to the glare of their star to be visible directly. Astronomers have used indirect methods to detect over 200 planets orbiting other stars, but you can't see them by just looking through a telescope.

In Figure 1-9, the sizes of the dots represent not the sizes of the stars but their brightnesses. This is the custom in astronomical diagrams, and it is also how star images are recorded on photographs. Bright stars make larger spots on a photograph than faint stars, so the size of a star image in a photograph tells you not how big the star is but only how bright it looks.

In ■ Figure 1-10, you expand your field of view by another factor of 100, and the sun and its neighboring stars vanish into the background of thousands of other stars. The field of view is now 1700 ly in diameter. Of course, no one has ever journeyed thousands of light-years from Earth to look back and photograph the solar neighborhood, so this is a representative photograph of the sky. The sun is a relatively faint star that would not be easily located in a photo at this scale.

If you again expand your field of view by a factor of 100, you see our galaxy, a disk of stars about 80,000 ly in diameter (■ Figure 1-11). A **galaxy** is a great cloud of stars, gas, and dust held together by the combined gravity of all the matter. Galaxies range from 1500 to over 300,000 ly in diameter and can contain over 100 billion stars. In the night sky, you see our galaxy as a great, cloudy wheel of stars ringing the sky. This band of stars is known as the **Milky Way,** and our galaxy is called the **Milky Way Galaxy.**

How does anyone know what our galaxy looks like if no one can leave it and look back? Astronomers use evidence and theory as guides and can imagine what the Milky Way looks like, and then artists can use those scientific conceptions to create a painting. Many images in this book are artists' renderings of objects and events that are too big or too dim to see clearly, emit energy your eyes cannot detect, or happen too slowly or too rapidly for humans to sense. These images are not just guesses; they are guided by the best information astronomers can gather. As you explore, notice how astronomers use their scientific imaginations understand cosmic events.

The artist's conception of the Milky Way reproduced in Figure 1-11 shows that our galaxy, like many others, has graceful **spiral arms** winding outward through its disk. In a later chapter, you will learn that stars are born in great clouds of gas and dust when they pass through the spiral arms. Our own sun was born in one of these spiral arms, and if you could see it in this picture, it would be in the disk of the galaxy about two-thirds of the way out from the center.

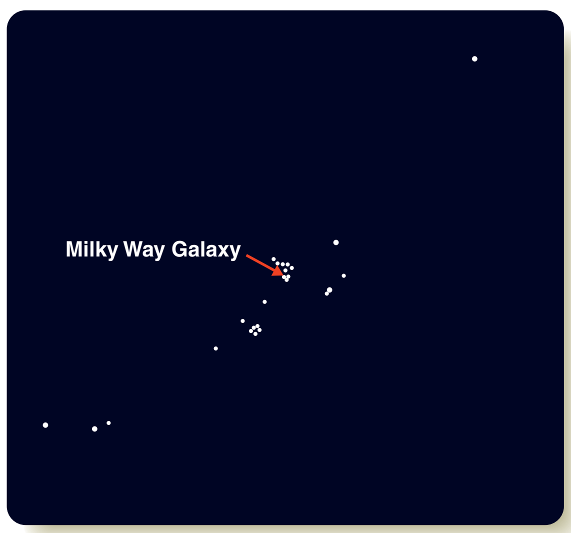
Figure 1-11

© Mark Garlick/space-art.com

Ours is a fairly large galaxy. Only a century ago astronomers thought it was the entire universe—an island cloud of stars in an otherwise empty vastness. Now they know that our galaxy is not unique; it is only one of many billions of galaxies scattered throughout the universe.

When you expand your field of view by another factor of 100, our galaxy appears as a tiny luminous speck surrounded by other specks (■ Figure 1-12).

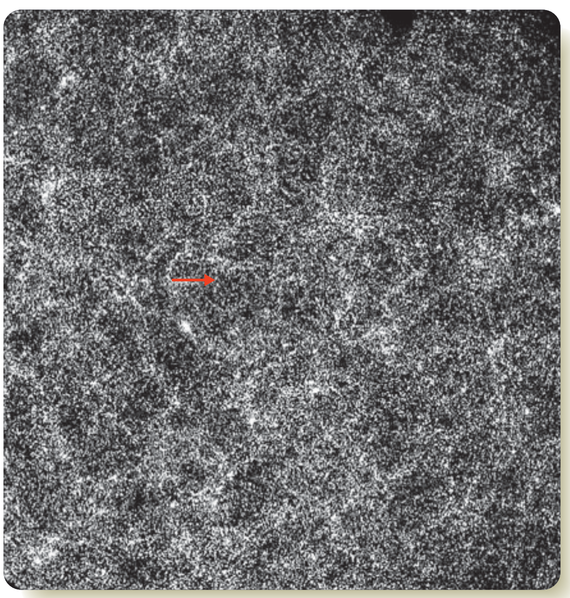
Figure 1-12



This diagram includes a region 17 million ly in diameter, and each of the dots represents a galaxy. Notice that our galaxy is part of a cluster of a few dozen galaxies. Galaxies are commonly grouped together in such clusters. Some galaxies have beautiful spiral patterns like our own galaxy, but others do not. Some are strangely distorted. One of the mysteries of modern astronomy is what produces these differences among the galaxies.

Now is a chance for you to correct another **Common Misconception**. People often say "galaxy" when they mean "solar system," and they sometimes confuse those terms with "universe." Your cosmic zoom has shown you the difference. The solar system is the sun and its planets. Our galaxy contains our solar system plus billions of other stars and whatever planets orbit around them. The universe includes everything, all of the galaxies, stars, and planets, including our own galaxy and our solar system.

If you again expand your field of view, you can see that galaxies tend to occur in clusters and that the clusters of galaxies are connected in a vast network (■ Figure 1-13). Clusters are grouped into superclusters—clusters of clusters—and the superclusters are linked to form long filaments and walls outlining nearly empty voids. These filaments and walls appear to be the largest structures in the universe. Were you to expand your field of view another time, you would probably see a uniform fog of filaments and walls. When you puzzle over the origin of these structures, you are at the frontier of human knowledge.



■ **Figure 1-13**

(Based on data from M. Seldner, B. L. Siebers, E. J. Groth, and P. J. E. Peebles, *Astronomical Journal 82* [1977].)

1-2 When Is Now?

ONCE YOU HAVE an idea where you are in space, you need to know where you are in time. The stars have shone for billions of years before the first human looked up and wondered what they were. To get a sense of your place in time, all you need is a long red ribbon.

Imagine stretching a ribbon from goal line to goal line down the center of a football field as shown on the inside front cover of this book. Imagine that one end of the ribbon is *Today* and that the other end represents the beginning of the universe—the moment of beginning that astronomers call the *big bang*. In the chapter "Modern Cosmology," you will learn all about the big bang, and you will see evidence that the universe is about 14 billion years old. Your long red ribbon represents 14 billion years, the entire history of the universe.

Imagine beginning at the goal line labeled *Big Bang.* You could replay the entire history of the universe by walking along your ribbon toward the goal line labeled *Today*. Observations tell astronomers that the big bang filled the entire universe with hot, dense gas, but as the gas cooled the universe went dark. All that happened in the first half inch on the ribbon. There was no light for the first 400 million years, until gravity was able to pull some of the gas together to form the first stars. That seems like a lot of years, but if you stick a little flag beside the ribbon to mark the birth of the first stars it would be not quite 3 yards from the goal line where the universe began.

You would go only about 5 yards before galaxies formed in large numbers. Our home galaxy would be one of those taking shape. By the time you crossed the 50-yard line, the universe would be full of galaxies, but the sun and Earth would not have formed yet. You would have to walk past the 50-yard line down to the 35-yard line before you could finally stick a flag to mark the formation of the sun and planets—our solar system.

You would have to carry your flags a few yards further to the 29-yard line to mark the appearance of the first life on Earth—microscopic creatures in the oceans. You would have to walk all the way to the 3-yard line before you could mark the emergence of life on land, and your dinosaur flag would go just inside the 2-yard line. Dinosaurs would go extinct as you passed the one-half-yard line.

What about people? You could put a little flag for the first humanlike creatures only about an inch from the goal line labeled *Today*. Civilization, the building of cities, began about 10,000 years ago. You have to try to fit that flag in only 0.0026 inches from the goal line. That's half the thickness of a sheet of paper. Compare the history of human civilization with the history of the universe. Every war you have ever heard of, every person whose name is recorded, every structure ever built from Stonehenge to the building you are in right now fits into that 0.0026 inches.

Humanity is very new to the universe. Our civilization on Earth has existed for only a flicker of an eyeblink in the history of the universe. As you will discover in the chapters that follow, only in the last hundred years or so have astronomers began to understand where we are in space and in time.

1-3 Why Study Astronomy?

Your exploration of the universe will help you answer two fundamental questions:

What are we?

How do we know?

What are we? That is the first organizing theme of this book. Astronomy is important to you because it will tell you what you are. Notice that the question is not "*Who* are we?" If you want to know who we are, you may want to talk to a sociologist, theologian, paleontologist, artist, or poet. "*What* are we?" is a fundamentally different question.

As you study astronomy, you will learn how you fit into the history of the universe. You will learn that the atoms in your body had their first birthday in the big bang when the universe began. Those atoms have been cooked and remade inside stars, and now, after billions of years, they are inside you. Where will they be in another billion years? This is a story everyone should know, and astronomy is the only course on campus that can tell you that story.

Every chapter in this book ends with a short segment titled "What Are We?" This summary shows how the astronomy in the chapter relates to your role in the story of the universe.

"How do we know?" That is the second organizing theme of this book. It is a question you should ask yourself whenever you encounter statements made by so-called experts in any field. Should you swallow a diet supplement recommended by a TV star? Should you vote for a candidate who warns of a climate crisis? To understand the world around you and to make wise decisions for yourself, for your family, and for your nation, you need to understand how science works.

You can use astronomy as a case study in science. In every chapter of this book, you will find short essays titled "How Do We Know?" They are designed to help you think not about *what* is known but about *how* it is known. That is, they will explain different aspects of scientific reasoning and in that way help you understand how scientists know about the natural world.

Over the last four centuries, scientists have developed a way to understand nature that is called the scientific method (■ How Do We Know? 1-1). You will see this process applied over and over as you read about exploding stars, colliding galaxies, and whirling planets. The universe is very big, but it is described by a small set of rules, and we humans have found a way to figure out the rules—a method called *science*.

How Do We Know? 1-1

The So-Called Scientific Method

How do scientists learn about nature? You have probably heard of the **scientific method** as the process by which scientists form hypotheses and test them against evidence gathered by experiment or observation. Scientists use the scientific method all the time, and it is critically important, but they rarely think of it. It is such an ingrained way of thinking about nature that it is almost invisible.

Scientists try to form hypotheses that explain how nature works. If a hypothesis is contradicted by experiments or observations, it must be revised or discarded. If a hypothesis is confirmed, it must be tested further. In that very general way, the scientific method is a way of testing and refining ideas to better describe how nature works.

For example, Gregor Mendel (1822–1884) was an Austrian abbot who liked plants. He formed a hypothesis that offspring usually inherited traits from their parents not as a smooth blend, as most scientists of the time believed, but according to strict mathematical rules. Mendel cultivated and tested over 28,000 pea plants, noting which produced smooth peas and which wrinkled peas and how that trait was inherited by successive generations. His study of pea plants and others confirmed his hypothesis and allowed the development of a series of laws of inheritance. Although the importance of his work was not recognized in his lifetime, it was combined with the discovery of chromosomes in 1915, and Mendel is now called the "father of modern genetics."

The scientific method is not a simple, mechanical way of grinding facts into understanding. It is, in fact, a combination of many ways of analyzing information, finding relationships, and creating new ideas. A scientist needs insight and ingenuity to form and test a good hypothesis. Scientists use the scientific method almost automatically, forming, testing, revising, and discarding hypotheses almost minute by minute as they discuss a new idea. Sometimes, however, a scientist will spend years studying a single important hypothesis. The so-called scientific method is a way of thinking and a way of knowing about nature. The "How Do We Know?" essays in the chapters that follow will introduce you to some of those methods.

Whether peas are wrinkled or smooth is an inherited trait. (Inspirestock/jupiterimages)

What Are We? Part of the Story

Astronomy will give you perspective on what it means to be here on Earth. This chapter used astronomy to locate you in space and time. Once you realize how vast our universe is, Earth seems quite small. People on the other side of the world seem like neighbors. And in the entire history of the universe, the human story is only the blink of an eye. This may seem humbling at first, but you can be proud of how much we humans have understood in such a short time.

Not only does astronomy locate you in space and time, it places you in the physical processes that govern the universe. Gravity and atoms work together to make stars, light the universe, generate energy, and create the chemical elements in your body. Astronomy locates you in that cosmic process.

Although you are very small and your kind have existed in the universe for only a short time, you are an important part of something very large and very beautiful.

Study and Review

Summary

- You surveyed the universe by taking a cosmic zoom in which each **field of view (p. 2)** was 100 times wider than the previous field of view.
- Astronomers use the metric system because it simplifies calculations and use **scientific notation (p. 3)** for very large or very small numbers.
- You live on a **planet (p. 3),** Earth, which orbits our **star (p. 3),** the sun, once a year. As Earth rotates once a day, you see the sun rise and set.
- The moon is only one-fourth the diameter of Earth, but the sun is 109 times larger in diameter than Earth—a typical size for a star.
- The **solar system (p. 3)** includes the sun at the center and all of the planets that orbit around it—Mercury, Venus, Mars, Jupiter, Saturn, Uranus, and Neptune.
- The **astronomical unit (AU) (p. 3)** is the average distance from Earth to the sun. Mars, for example, orbits 1.5 AU from the sun. The **light-year (ly) (p. 4)** is the distance light can travel in one year. The nearest star is 4.2 ly from the sun.
- Many stars seem to have planets, but such small, distant worlds are difficult to detect. Only a few hundred have been found so far, but planets seem to be common, so you can probably trust that there are lots of planets in the universe including some like Earth.
- The **Milky Way (p. 5),** the hazy band of light that encircles the sky, is the **Milky Way Galaxy (p. 5)** seen from inside. The sun is just one out of the billions of stars that fill the Milky Way Galaxy.
- **Galaxies (p. 5)** contain many billions of stars. Our galaxy is about 80,000 ly in diameter and contains over 100 billion stars.
- Some galaxies, including our own, have graceful **spiral arms (p. 5)** bright with stars, but some galaxies are plain clouds of stars.
- Our galaxy is just one of billions of galaxies that fill the universe in great clusters, clouds, filaments, and walls—the largest things in the universe.
- The universe began about 14 billion years ago in an event called the big bang, which filled the universe with hot gas.
- The hot gas cooled, the first galaxies began to form, and stars began to shine only about 400 million years after the big bang.
- The sun and planets of our solar system formed about 4.6 billion years ago.
- Life began in Earth's oceans soon after Earth formed but did not emerge onto land until only 400 million years ago. Dinosaurs evolved not long ago and went extinct only 65 million years ago.
- Human-like creatures appeared on Earth only about 4 million years ago, and human civilizations developed only about 10,000 years ago.
- Although astronomy seems to be about stars and planets, it describes the universe in which you live, so it is really about you. Astronomy helps you answer the question, "What are we?"
- As you study astronomy, you should ask "How do we know?" and that will help you understand how science gives us a way to understand nature.
- In its simplest outline, science follows the **scientific method (p. 7),** in which scientists expect statements to be supported by evidence compared with theory. In fact, science is a complex and powerful way to think about nature.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds.**

1. What is the largest dimension of which you have personal knowledge? Have you run a mile? Hiked 10 miles? Run a marathon?
2. What is the difference between our solar system, our galaxy, and the universe?
3. Why are light-years more convenient than miles, kilometers, or astronomical units for measuring certain distances?
4. Why is it difficult to detect planets orbiting other stars?
5. What does the size of the star image in a photograph tell you?
6. What is the difference between the Milky Way and the Milky Way Galaxy?
7. What are the largest known structures in the universe?
8. How does astronomy help answer the question, "What are we?"
9. **How Do We Know?** How does the scientific method give scientists a way to know about nature?

Discussion Questions

1. Do you think you have a right to know the astronomy described in this chapter? Do you think you have a duty to know it? Can you think of ways this knowledge helps you enjoy a richer life and be a better citizen?
2. How is a statement in a political campaign speech different from a statement in a scientific discussion? Find examples in newspapers, magazines, and this book.

Problems

1. The diameter of Earth is 7928 miles. What is its diameter in inches? In yards? If the diameter of Earth is expressed as 12,756 km, what is its diameter in meters? In centimeters?
2. If a mile equals 1.609 km and the moon is 2160 miles in diameter, what is its diameter in kilometers?
3. One astronomical unit is about 1.5×10^8 km. Explain why this is the same as 150×10^6 km.
4. Venus orbits 0.72 AU from the sun. What is that distance in kilometers?
5. Light from the sun takes 8 minutes to reach Earth. How long does it take to reach Mars?
6. The sun is almost 400 times farther from Earth than is the moon. How long does light from the moon take to reach Earth?
7. If the speed of light is 3×10^5 km/s, how many kilometers are in a light-year? How many meters?
8. How long does it take light to cross the diameter of our Milky Way Galaxy?
9. The nearest galaxy to our own is about 2 million light-years away. How many meters is that?
10. How many galaxies like our own would it take laid edge-to-edge to reach the nearest galaxy? (*Hint:* See Problem 9.)

Learning to Look

1. In Figure 1-4, the division between daylight and darkness is at the right on the globe of Earth. How do you know this is the sunset line and not the sunrise line?
2. Look at Figure 1-6. How can you tell that Mercury follows an elliptical orbit?
3. Of the objects listed here, which would be contained inside the object shown in the photograph at the right? Which would contain the object in the photo?

   stars
   planets
   galaxy clusters
   filaments
   spiral arms

Bill Schoening/NOAO/AURA/NSF

4. In the photograph shown here, which stars are brightest, and which are faintest? How can you tell? Why can't you tell which stars in this photograph are biggest or which have planets?

NOAO

2 The Sky

The sky above mountaintop observatories far from city lights is the same sky you see from your window. The stars above you are other suns scattered through the universe.
(Kris Koenig/Coast Learning Systems)

Guidepost

The previous chapter took your on a cosmic zoom through space and time. That quick preview only sets the stage for the drama to come. Now it is time to look closely at the sky, and answer three essential questions:

How do astronomers refer to stars?

How can you compare the brightness of the stars?

How does the sky appear to move as Earth rotates?

As you study the sky and its motions, you will be learning to think of Earth as a planet rotating on its axis. The next chapter will introduce you to the orbital motion of Earth and to a family of objects in the sky that move against the background of stars.

Animated! This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.

The Southern Cross I saw every night abeam. The sun every morning came up astern; every evening it went down ahead. I wished for no other compass to guide me, for these were true.

CAPTAIN JOSHUA SLOCUM
SAILING ALONE AROUND THE WORLD

■ **Figure 2-1**

The constellations are an ancient heritage handed down for thousands of years as celebrations of great heroes and mythical creatures. Here Sagittarius and Scorpius hang above the southern horizon.

THE NIGHT SKY is the rest of the universe as seen from our planet. When you look up at the stars, you are looking out through a layer of air only a little more than a hundred kilometers deep. Beyond that, space is nearly empty, and the stars are spread light-years apart.

As you read this chapter, keep in mind that you live on a planet in the midst of these scattered stars. Because our planet rotates on its axis once a day, the sky appears to revolve around you in a daily cycle. Not only does the sun rise in the east and set in the west, but so do the stars.

2-1 The Stars

ON A DARK night far from city lights, you can see a few thousand stars in the sky. The ancients organized what they saw by naming stars and groups of stars. Some of those names survive today.

Constellations

All around the world, ancient cultures celebrated heroes, gods, and mythical beasts by naming groups of stars—**constellations** (■ Figure 2-1). You should not be surprised that the star patterns do not look like the creatures they represent any more than Columbus, Ohio, looks like Christopher Columbus. The constellations simply celebrate the most important mythical figures in each culture. The constellations named by Western cultures originated in Mesopotamia over 5000 years ago, with other constellations added by Babylonian, Egyptian, and Greek astronomers during the classical age. Of these ancient constellations, 48 are still used today.

To the ancients, a constellation was a loose grouping of stars. Many of the fainter stars were not included in any constellation, and the stars of the southern sky not visible to the ancient astronomers of northern latitudes were not grouped into constellations. Constellation boundaries, when they were defined at all, were only approximate (■ Figure 2-2a), so a star like Alpheratz could be thought of as part of Pegasus or part of Andromeda. To correct these gaps and ambiguities, astronomers have added 40 modern constellations, and in 1928 the International Astronomical Union established 88 official constellations with clearly defined boundaries (Figure 2-2b). Consequently, a constellation now represents not a group of stars but an area of the sky, and any star within the region belongs to one and only one constellation. Alpheratz belongs to Andromeda.

In addition to the 88 official constellations, the sky contains a number of less formally defined groupings called **asterisms.** The Big Dipper, for example, is a well-known asterism that is part of the constellation Ursa Major (the Great Bear). Another asterism is the Great Square of Pegasus (Figure 2-2b), which includes three stars from Pegasus plus Alpheratz from Andromeda. The star charts at the end of this book will introduce you to the brighter constellations and asterisms.

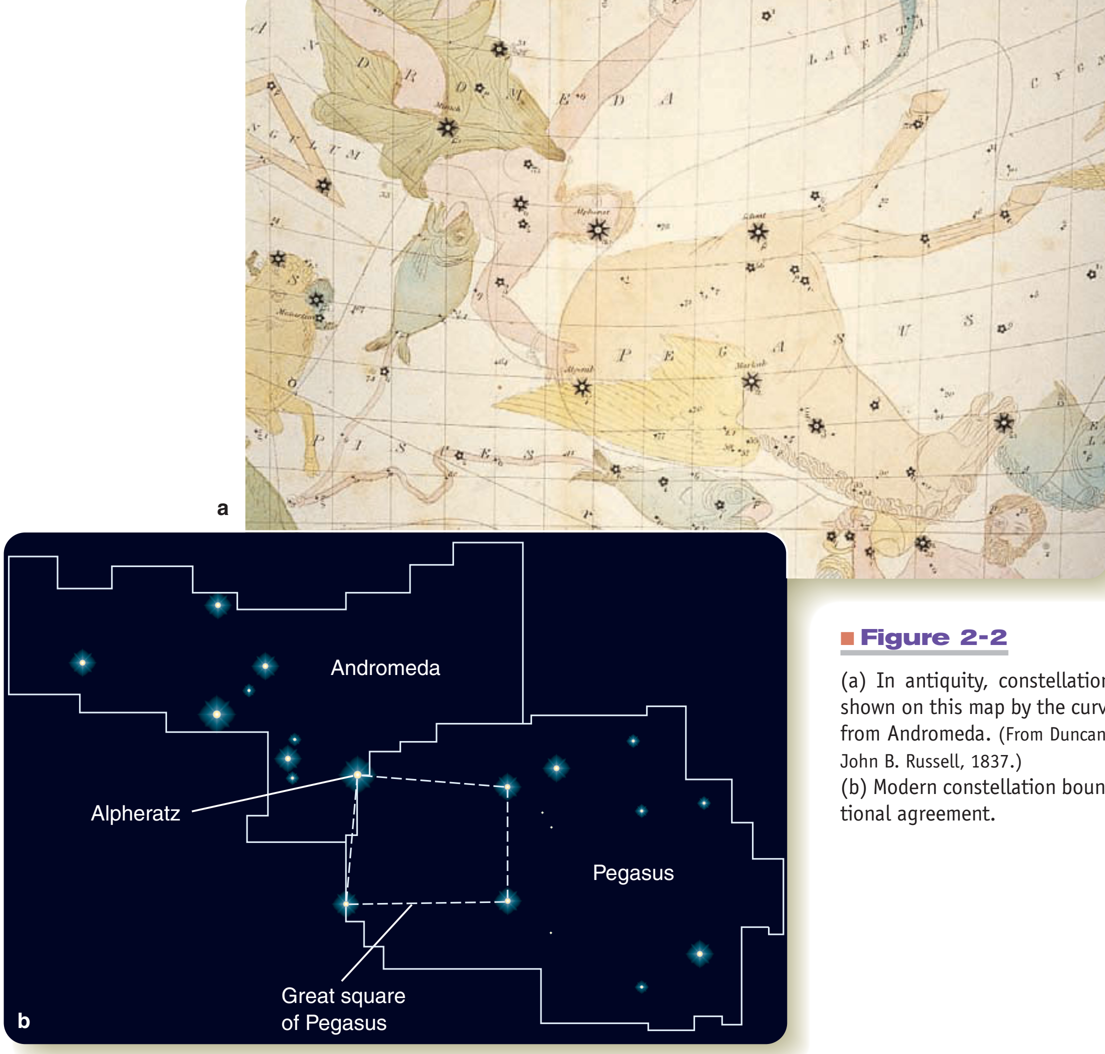
Although constellations and asterisms refer to stars grouped together in the sky, it is important to remember that most are made up of stars that are not physically associated with one another. Some stars may be many times farther away than others and moving through space in different directions. The only thing they have in common is that they lie in approximately the same direction from Earth (■ Figure 2-3).

The Names of the Stars

In addition to naming groups of stars, ancient astronomers gave individual names to the brighter stars. Modern astronomers still use many of those names. The constellation names come from Greek translated into Latin—the language of science from the fall of Rome to the 19th century—but most star names come from ancient Arabic, though much altered by the passing centuries. The name of Betelgeuse, the bright red star in Orion, for example, comes from the Arabic *yad al jawza,* meaning "shoulder of Jawza [Orion]." Names such as Sirius (the Scorched One), and Aldebaran (the Follower of the Pleiades) are beautiful additions to the mythology of the sky.

Naming individual stars is not very helpful because you can see thousands of them. How many names could you remember? A more useful way to identify stars is to assign Greek letters to the bright stars in a constellation in approximate order of brightness. Thus the brightest star is usually designated alpha, the second brightest beta, and so on. Often the name of the Greek letter is spelled out, as in "alpha," but sometimes the actual Greek letter is used. You will find the Greek alphabet in Appendix A. For many constellations, the letters follow the order of brightness, but some constellations, by tradition, mistake, or the personal preferences of early chart makers, are exceptions (■ Figure 2-4).

To identify a star by its Greek-letter designation, you give the Greek letter followed by the possessive (genitive) form of the constellation name; for example, the brightest star in the constellation Canis Major is alpha Canis Majoris, which can also be written α Canis Majoris. This both identifies the star and the constellation and gives a clue to the relative brightness of the star.

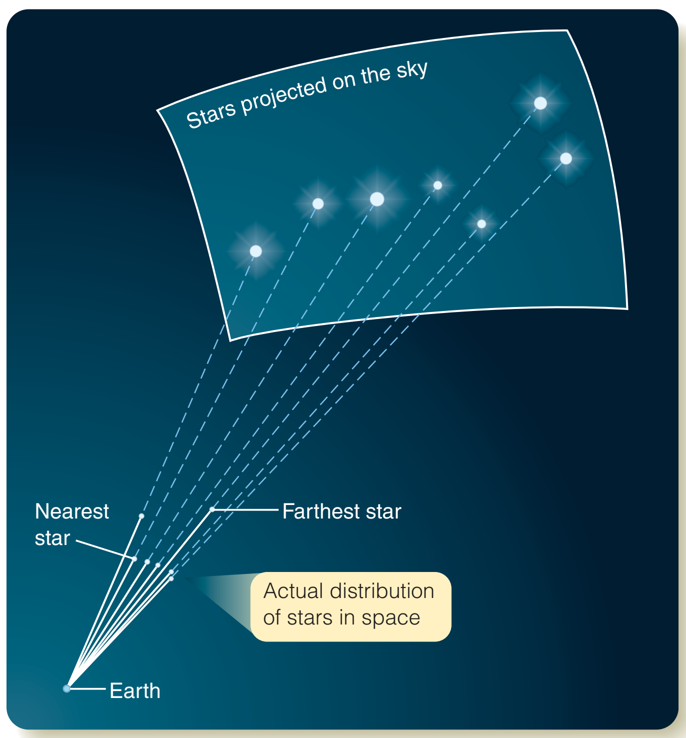


■ **Figure 2-2**

(a) In antiquity, constellation boundaries were poorly defined, as shown on this map by the curving dotted lines that separate Pegasus from Andromeda. (From Duncan Bradford, *Wonders of the Heavens,* Boston: John B. Russell, 1837.)
(b) Modern constellation boundaries are precisely defined by international agreement.

■ **Figure 2-3**

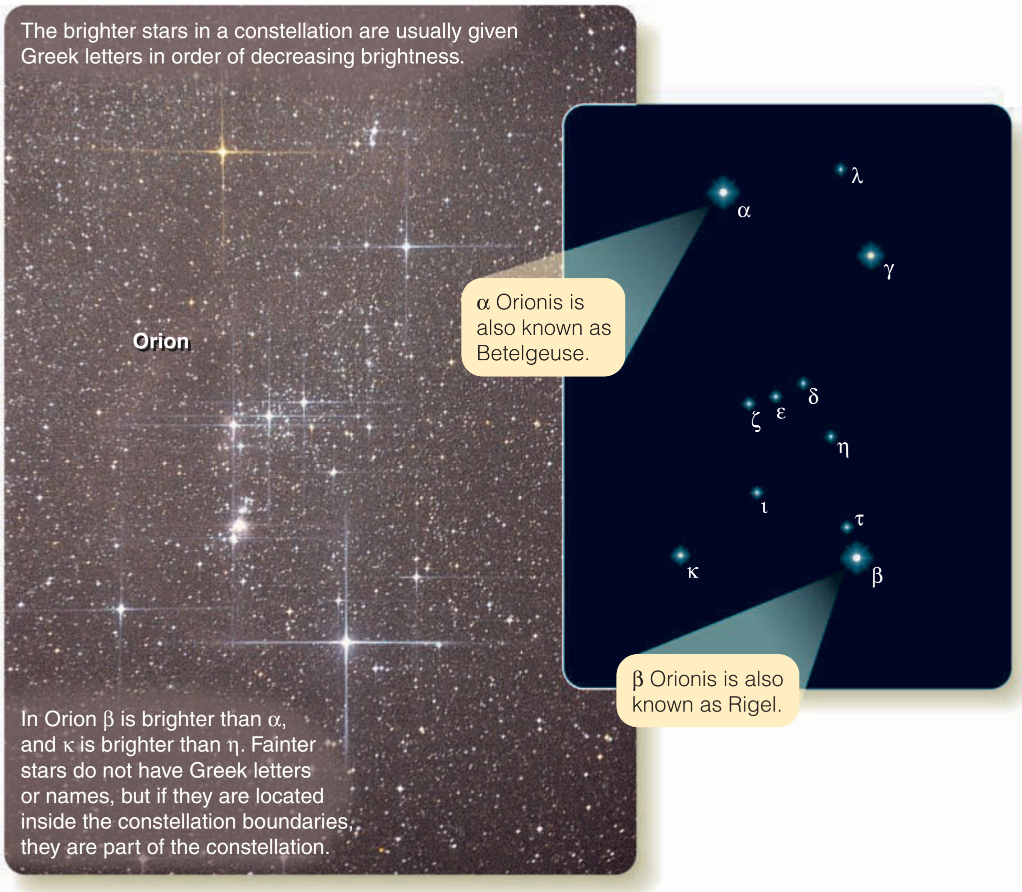
You see the Big Dipper in the sky because you are looking through a group of stars scattered through space at different distances from Earth. You see them as if they were projected on a screen, and they form the shape of the Dipper.



The Brightness of Stars

Astronomers measure the brightness of stars using the **magnitude scale,** a system that first appeared in the writings of the ancient astronomer Claudius Ptolemy about AD 140. The system probably originated earlier than Ptolemy, and most astronomers attribute it to the Greek astronomer Hipparchus (about 190–120 BC). Hipparchus compiled the first known star catalog, and he may have used the magnitude system in that catalog. Almost 300 years later, Ptolemy used the magnitude system in his own catalog, and successive generations of astronomers have continued to use the system.

The ancient astronomers divided the stars into six classes. The brightest were called first-magnitude stars and those that were fainter, second-magnitude. The scale continued downward to sixth-magnitude stars, the faintest visible to the human eye. Thus, the larger the magnitude number, the fainter the star. This makes sense if you think of the bright stars as first-class stars and the faintest stars visible as sixth-class stars.



■ **Figure 2-4**

Stars in a constellation can be identified by Greek letters and by names derived from Arabic. The spikes on the star images in the photograph were produced by the optics in the camera. (William Hartmann)

Compare this with the ancient name for this star, Sirius, which tells you nothing about location or brightness.

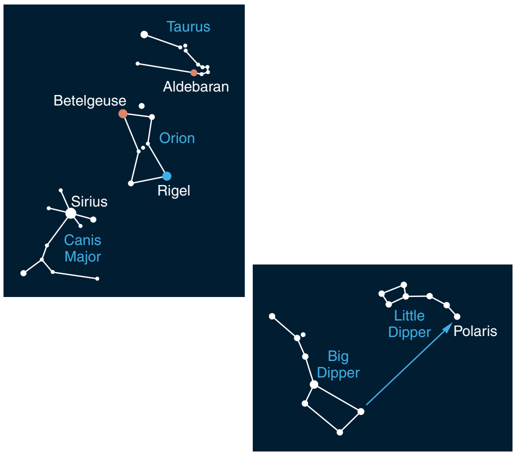
It is fun to know the names of the brighter stars, but they are more than points of light in the sky. They are glowing spheres of gas much like the sun, each with its unique characteristics. ■ Figure 2-5 identifies eight bright stars that you can adopt as Favorite Stars. As you study astronomy you will discover their peculiar personalities and enjoy finding them in the evening sky.

You can use the star charts at the end of this book to help locate these Favorite Stars. You can see Polaris year round, but Sirius, Betelgeuse, Rigel, and Aldebaran are in the winter sky. Spica is a summer star, and Vega is visible evenings in later summer or fall. Alpha Centauri is a special star, and you will have to travel as far south as southern Florida to glimpse it above the southern horizon.

Naming stars is helpful, but to discuss the sky with precision, you must have an accurate way of referring to the brightness of stars, and for that you must consult two of the first great astronomers.

Modern astronomers can measure the brightness of stars to high precision, so they have made adjustments to the ancient scale of magnitudes. Instead of saying that the star known by the charming name Chort (Theta Leonis) is third magnitude, they can say its magnitude is 3.34. Accurate measurements show that some stars are brighter than magnitude 1.0. For example, Favorite Star Vega (alpha Lyrae) is so bright that its magnitude, 0.04, is almost zero. A few are so bright the magnitude scale must extend into negative numbers (■ Figure 2-6). On this scale, our Favorite Star Sirius, the brightest star in the sky, has a magnitude of −1.47. Modern astronomers have had to extend the faint end of the magnitude scale as well. The faintest stars you can see with your unaided eyes are about sixth magnitude, but if you use a telescope, you will see stars much fainter. Astronomers must use magnitude numbers larger than 6 to describe these faint stars.

These numbers are known as **apparent visual magnitudes** **(m_V),** and they describe how the stars look to human eyes observing from Earth. Although some stars emit large amounts of infrared or ultraviolet light, human eyes can't see it, and it is not included in the apparent visual magnitude. The subscript "V" stands for "visual" and reminds you that you are including only



Sirius	**Brightest star in the sky**	**Winter**
Betelgeuse	**Bright red star in Orion**	**Winter**
Rigel	**Bright blue star in Orion**	**Winter**
Aldebaran	**Red eye of Taurus the Bull**	**Winter**
Polaris	**The North Star**	**Year round**
Vega	**Bright star overhead**	**Summer**
Spica	**Bright southern star**	**Summer**
Alpha Centauri	**Nearest star to the sun**	**Spring, far south**

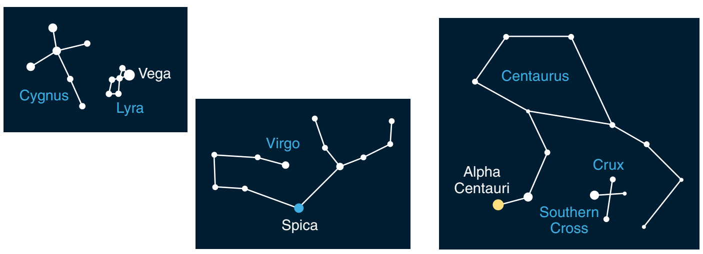


Figure 2-5

Favorite Stars: Locate these bright stars in the sky and learn why they are interesting.

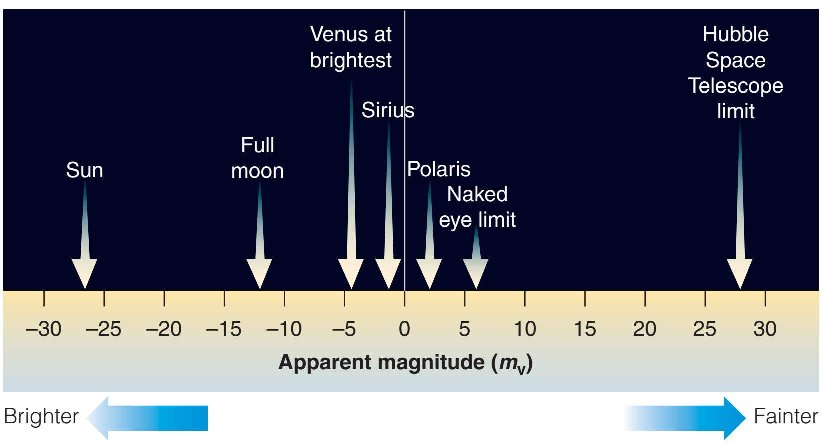


Figure 2-6

The scale of apparent visual magnitudes extends into negative numbers to represent the brightest objects and to positive numbers larger than 6 to represent objects fainter than the human eye can see.

light you can see. Apparent visual magnitude also does not take into account the distance to the stars. Very distant stars look fainter, and nearby stars look brighter. Apparent visual magnitude ignores the effect of distance and tells you only how bright the star looks as seen from Earth.

Your interpretation of brightness is quite subjective, depending on both the physiology of human eyes and the psychology of perception. To be accurate you should refer to **flux**—a measure of the light energy from a star that hits one square meter in one second. Such measurements precisely define the intensity of starlight, and a simple relationship connects apparent visual magnitudes and intensity (■ Reasoning with Numbers 2-1). In this way, modern astronomers can measure the brightness of stars to high precision while still making comparisons to observations of apparent visual magnitude that go back to the time of Hipparchus.

2-2 The Sky and Its Motion

THE SKY ABOVE seems to be a great blue dome in the daytime and a sparkling ceiling at night.

The Celestial Sphere

Ancient astronomers believed the sky was a great sphere surrounding Earth with the stars stuck on the inside like thumbtacks in a ceiling. Modern astronomers know that the stars are scattered through space at different distances, but it is still convenient to think of the sky as a great starry sphere enclosing Earth.

The Concept Art Portfolio ■ **The Sky Around You** on pages 16–17 takes you on an illustrated tour of the sky. Throughout this book, these two-page art spreads introduce new concepts and new terms through photos and diagrams. These concepts and new terms are not discussed elsewhere, so examine the art spreads carefully. Notice that The Sky Around You introduces you to three important principles and 16 new terms that will help you understand the sky:

1. The sky appears to rotate westward around Earth each day, but that is a consequence of the eastward rotation of Earth. That rotation produces day and night. Notice how reference points on the *celestial sphere* such as the *zenith, nadir, horizon, celestial equator,* and *north and south celestial poles* define the four directions, *north point, south point, east point,* and *west point.*
2. Astronomers measure *angular distance* across the sky as angles and express them as degrees, *minutes,* and *seconds of arc*. The same units are used to measure the *angular diameter* of an object.
3. What you can see of the sky depends on where you are on Earth. If you lived in Australia, you would see

Reasoning with Numbers | 2-1

Magnitudes

Astronomers use a simple formula to convert between magnitudes and intensities. If two stars have intensities I_A and I_B, then the ratio of their intensities is I_A/I_B. Modern astronomers have defined the magnitude scale so that two stars that differ by five magnitudes have an intensity ratio of exactly 100. Then two stars that differ by one magnitude must have an intensity ratio that equals the fifth root of 100, $\sqrt[5]{100}$, which equals 2.512 . . . That is, the light of one star must be 2.512 times more intense. Two stars that differ by two magnitudes will have an intensity ratio of 2.512×2.512, or about 6.3, and so on (■ Table 2-1).

Example A: Suppose star C is third magnitude, and star D is ninth magnitude. What is the intensity ratio?

Solution: The magnitude difference is six magnitudes, and the table shows the intensity ratio is 250. Therefore light from star C is 250 times more intense than light from star D.

A table is convenient, but for more precision you can express the relationship as a simple formula. The intensity ratio I_A/I_B is equal to 2.512 raised to the power of the magnitude difference $m_B - m_A$:

$$\frac{I_A}{I_B} = (2.512)^{(m_B - m_A)}$$

Example B: If the magnitude difference is 6.32 magnitudes, what is the intensity ratio?

Solution: The intensity ratio must be $2.512^{6.32}$. A pocket calculator tells you the answer: 337.

When you know the intensity ratio and want to find the magnitude difference, it is convenient to solve the formula for the magnitude difference:

$$m_B - m_A = 2.5 \log(I_A/I_B)$$

Example C: The light from Sirius is 24.2 times more intense than light from Polaris. What is the magnitude difference?

Solution: The magnitude difference is 2.5 log(24.2). Your pocket calculator tells you the logarithm of 24.2 is 1.38, so the magnitude difference is 2.5×1.38, which equals 3.4 magnitudes.

■ **Table 2-1 | Magnitude and Intensity**

Magnitude Difference	Intensity Ratio
0	1
1	2.5
2	6.3
3	16
4	40
5	100
6	250
7	630
8	1600
9	4000
10	10,000
⋮	⋮
15	1,000,000
20	100,000,000
25	10,000,000,000
⋮	⋮

many constellations and asterisms invisible from North America, but you would never see the Big Dipper. How many *circumpolar constellations* you see depends on where you are. Remember your Favorite Star Alpha Centauri? It is in the southern sky and isn't visible from most of the United States. You could just glimpse it above the southern horizon if you were in Miami, but you could see it easily from Australia.

Pay special attention to the new terms on pages 16–17. You need to know these terms to describe the sky and its motions, but don't fall into the trap of memorizing new terms. The goal of science is to understand nature, not to memorize definitions. Study the diagrams and see how the geometry of the celestial sphere and its motions produce the sky you see above you.

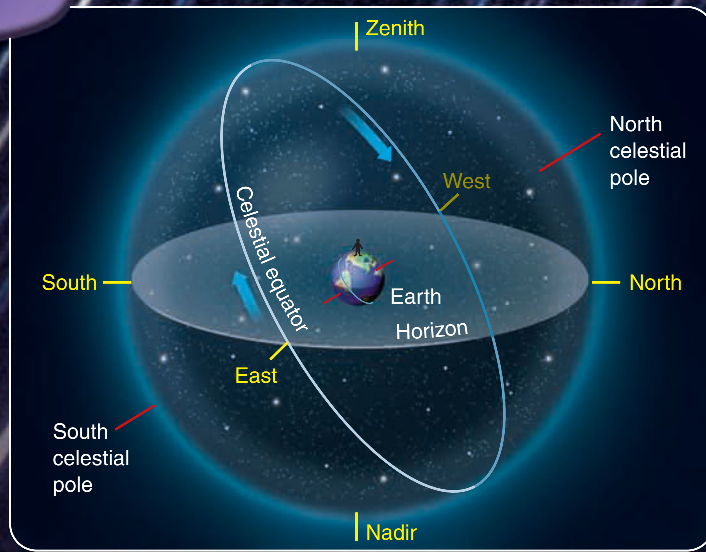
The celestial sphere is an example of a **scientific model,** a common feature of scientific thought (■ How Do We Know? 2-1). Notice that a scientific model does not have to be true to be useful. You will encounter many scientific models in the chapters that follow, and you will discover that some of the most useful models are highly simplified descriptions of the true facts.

This is a good time to eliminate a couple of **Common Misconceptions.** Lots of people, without thinking about it much, assume that the stars are not in the sky during the daytime. The stars are actually there day and night; they are just invisible during the day because the sky is lit up by sunlight. Also, many people insist that Favorite Star Polaris is the brightest star in the sky. You now know that Polaris is important because of its position, not because of its brightness.

The Sky Around You

1 The eastward rotation of Earth causes the sun, moon, and stars to move westward in the sky as if the **celestial sphere** were rotating westward around Earth. From any location on Earth you see only half of the celestial sphere, the half above the **horizon.** The **zenith** marks the top of the sky above your head, and the **nadir** marks the bottom of the sky directly under your feet. The drawing at right shows the view for an observer in North America. An observer in South America would have a dramatically different horizon, zenith, and nadir.

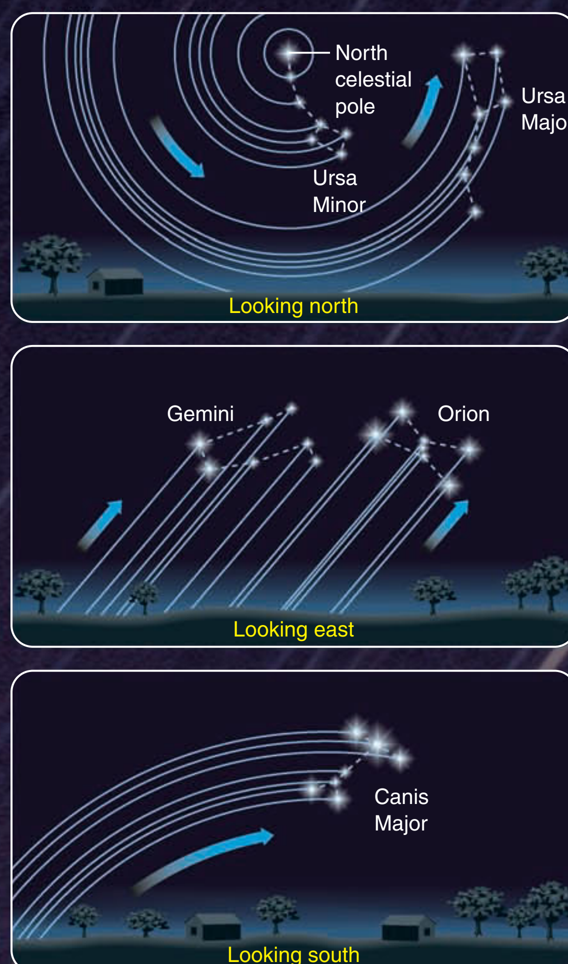
The apparent pivot points are the **north celestial pole** and the **south celestial pole** located directly above Earth's north and south poles. Halfway between the celestial poles lies the **celestial equator.** Earth's rotation defines the directions you use every day. The **north point** and **south point** are the points on the horizon closest to the celestial poles. The **east point** and the **west point** lie halfway between the north and south points. The celestial equator always meets the horizon at the east and west points.



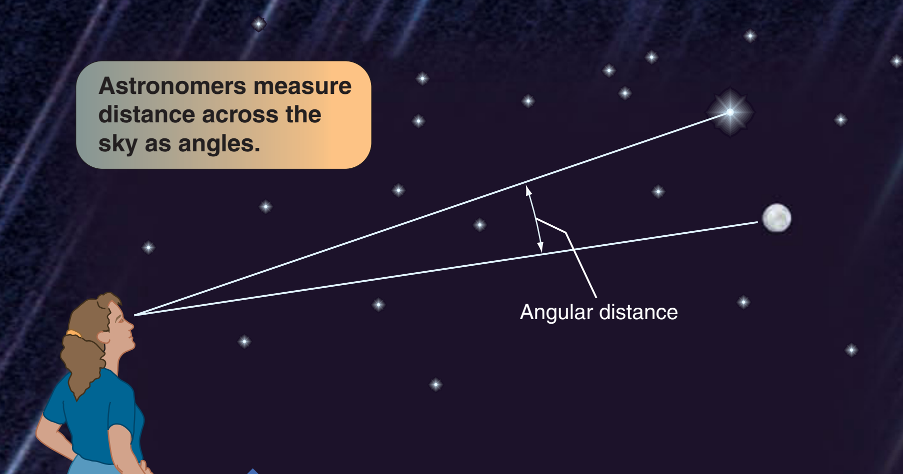
CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Celestial Sphere." Notice how each location on Earth has its unique horizon.

AURA/NOAO/NSF

1a This time exposure of about 30 minutes shows stars as streaks, called star trails, rising behind an observatory dome. The camera was facing northeast to take this photo. The motion you see in the sky depends on which direction you look, as shown at right. Looking north, you see the Favorite Star Polaris, the North Star, located near the north celestial pole. As the sky appears to rotate westward, Polaris hardly moves, but other stars circle the celestial pole. Looking south from a location in North America, you can see stars circling the south celestial pole, which is invisible below the southern horizon.

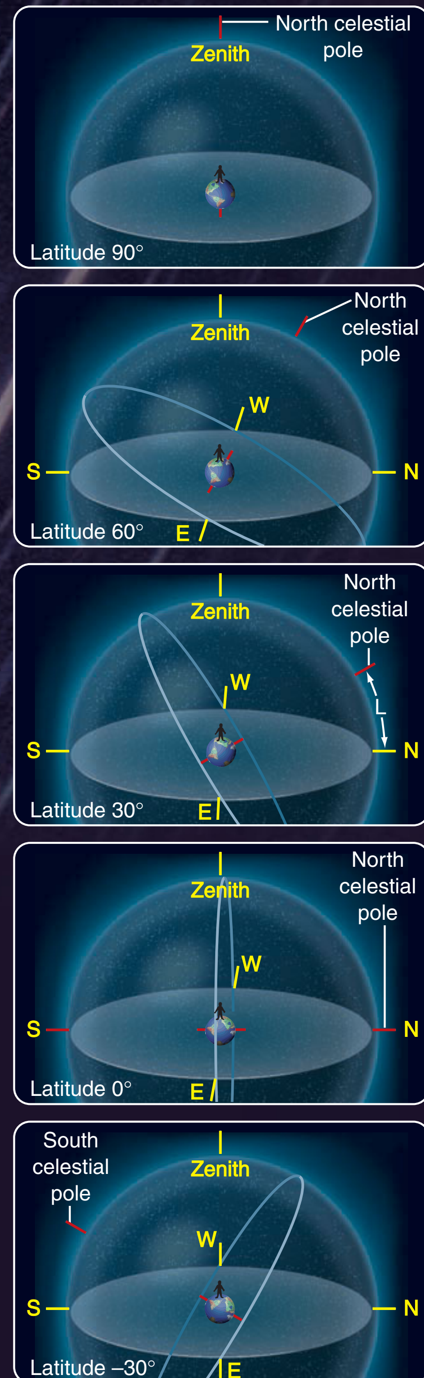


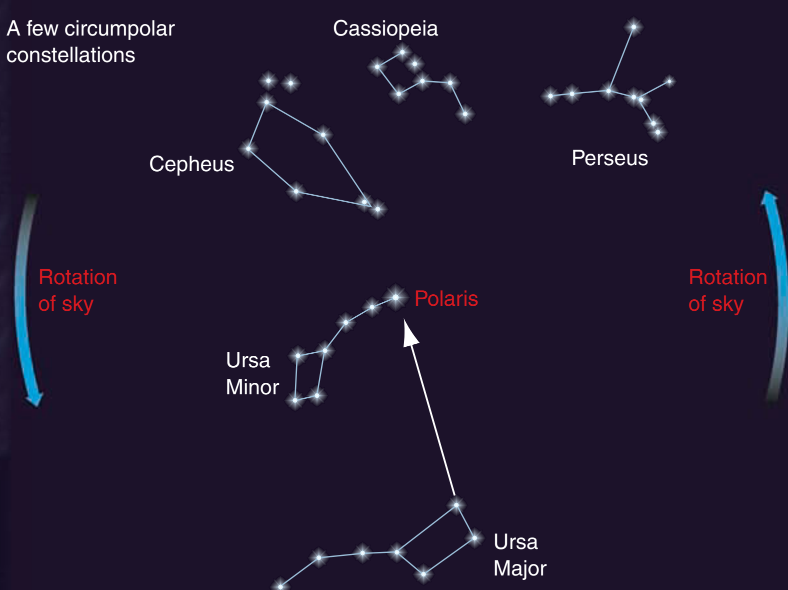
CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Rotation of the Sky." Look in different directions and compare the motions of the stars.



2 Astronomers might say, “The star was only 2 degrees from the moon.” Of course, the stars are much farther away than the moon, but when you think of the celestial sphere, you can measure distances *on the sky* as **angular distances** in degrees, minutes of arc, and seconds of arc. A **minute of arc** is 1/60th of a degree, and a **second of arc** is 1/60th of a minute of arc. Then the **angular diameter** of an object is the angular distance from one edge to the other. The sun and moon are each about half a degree in diameter, and the bowl of the Big Dipper is about 10° wide.

3 What you see in the sky depends on your latitude as shown at right. Imagine that you begin a journey in the ice and snow at Earth’s North Pole with the north celestial pole directly overhead. As you walk southward, the celestial pole moves toward the horizon, and you can see further into the southern sky. The angular distance from the horizon to the north celestial pole always equals your latitude (L)—the basis for celestial navigation. As you cross Earth’s equator, the celestial equator would pass through your zenith, and the north celestial pole would sink below your northern horizon.





3a **Circumpolar constellations** are those that never rise or set. From mid-northern latitudes, as shown at left, you see a number of familiar constellations circling Polaris and never dipping below the horizon. As the sky rotates, the pointer stars at the front of the Big Dipper always point toward Polaris. Circumpolar constellations near the south celestial pole never rise as seen from mid-northern latitudes. From a high latitude such as Norway, you would have more circumpolar constellations, and from Quito, Ecuador, located on Earth’s equator, you would have no circumpolar constellations at all.

CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure “Constellations from Different Latitudes.”

How Do We Know? | 2-1

Scientific Models

How can a scientific model be useful if it isn't entirely true? A scientific model is a carefully devised conception of how something works, a framework that helps scientists think about some aspect of nature, just as the celestial sphere helps astronomers think about the motions of the sky.

Chemists, for example, use colored balls to represent atoms and sticks to represent the bonds between them, kind of like Tinkertoys. Using these molecular models, chemists can see the three-dimensional shape of molecules and understand how the atoms interconnect. The molecular model of DNA proposed by Watson and Crick in 1953 led to our modern understanding of the mechanisms of genetics. You have probably seen elaborate ball-and-stick models of DNA, but does the molecule really look like Tinkertoys? No, but the model is both simple enough and accurate enough help scientists think about their theories.

A scientific model is not a statement of truth; it does not have to be precisely true to be useful. In an idealized model, some complex aspects of nature can be simplified or omitted. The ball-and-stick model of a molecule doesn't show the relative strength of the chemical bonds, for instance. A model gives scientists a way to think about some aspect of nature but need not be true in every detail.

When you use a scientific model, it is important to remember the limitations of that model. If you begin to think of a model as true, it can be misleading instead of helpful. The celestial sphere, for instance, can help you think about the sky, but you must remember that it is only a model. The universe is much larger and much more interesting than this ancient scientific model of the heavens.

Balls represent atoms and rods represent chemical bonds in this model of a DNA molecule. (Digital Vision/Getty Images)

In addition to the obvious daily motion of the sky, Earth's daily rotation conceals a very slow celestial motion that can be detected only over centuries.

Precession

Over 2000 years ago, Hipparchus compared a few of his star positions with those recorded nearly two centuries earlier and realized that the celestial poles and equator were slowly moving across the sky. Later astronomers understood that this motion is caused by the toplike motion of Earth.

If you have ever played with a gyroscope or top, you have seen how the spinning mass resists any sudden change in the direction of its axis of rotation. The more massive the top and the more rapidly it spins, the more it resists your efforts to twist it out of position. But you probably recall that even the most rapidly spinning top slowly sweeps its axis around in a conical motion. That is, the axis of the top pivots so the axis sweeps out the surface of a cone. The weight of the top tends to make it tip, and this combines with its rapid rotation to make its axis sweep around in a conical motion called **precession** (■ Figure 2-7a).

Earth spins like a giant top, but it does not spin upright in its orbit; it is tipped 23.5° from vertical. Earth's large mass and rapid rotation keep its axis of rotation pointed toward a spot near the star Polaris, and the axis would not wander if Earth were a perfect sphere. However, because of its rotation, Earth has a slight bulge around its middle. The gravity of the sun and moon pull on this bulge, tending to twist Earth upright in its orbit. The combination of these forces and Earth's rotation causes Earth's axis to precess in a conical motion, taking about 26,000 years for one cycle (Figure 2-7b).

Because the locations of the celestial poles and equator are defined by Earth's rotational axis, precession slowly moves these reference marks. You would notice no change at all from night to

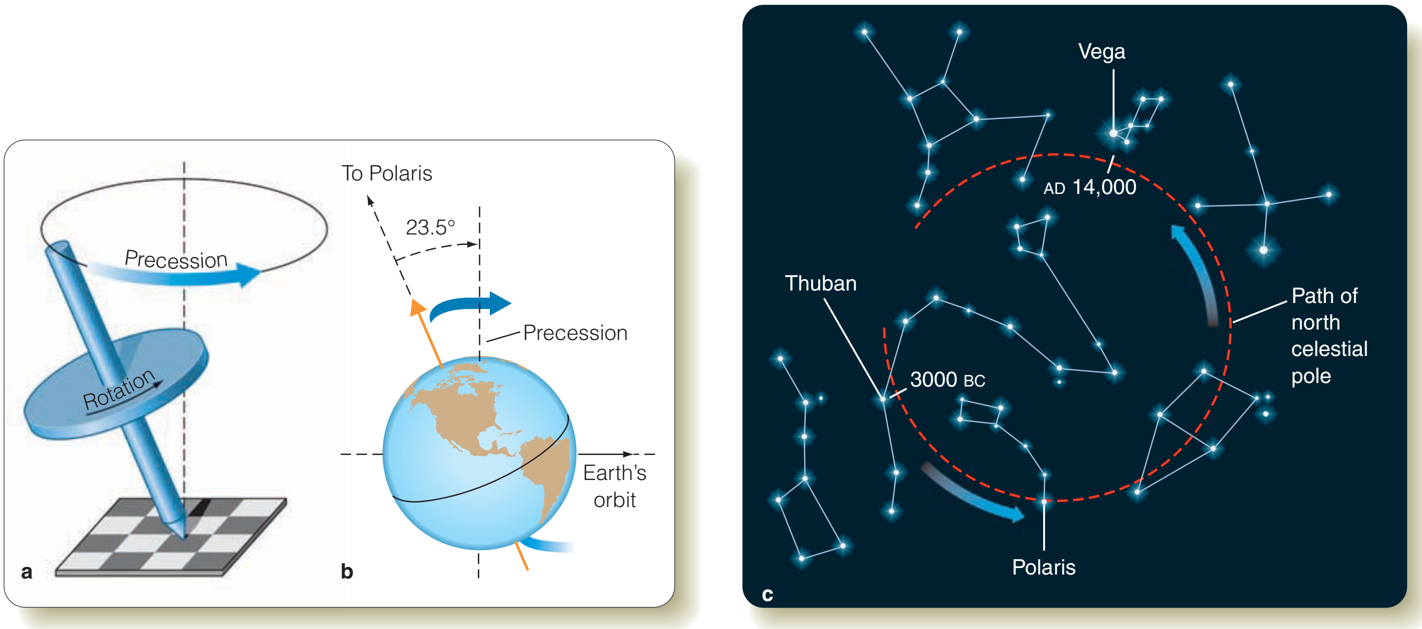


Figure 2-7

Precession. (a) A spinning top precesses in a conical motion around the perpendicular to the floor because its weight tends to make it fall over. (b) Earth precesses around the perpendicular to its orbit because the gravity of the sun and moon tend to twist it upright. (c) Precession causes the north celestial pole to move slowly among the stars, completing a circle in 26,000 years.

night or year to year, but precise measurements can reveal the slow precession of the celestial poles and equator.

Over centuries, precession has dramatic effects. Egyptian records show that 4800 years ago the north celestial pole was near the star Thuban (alpha Draconis). The pole is now approaching Polaris and will be closest to it in about 2100. In about 12,000 years, the pole will have moved to within 5° of Vega (alpha Lyrae). Next time you glance at Favorite Star Vega, remind yourself that it will someday be a very impressive north star. Figure 2-7c shows the path followed by the north celestial pole. You will discover in later chapters that precession is common among rotating astronomical bodies.

What Are We? Along for the Ride

We humans are planetwalkers. We live on the surface of a whirling planet, and as we look out into the depths of the universe we see the scattered stars near us. Because our planet spins, the stars appear to move westward across the sky in continuous procession.

The sky is a symbol of remoteness, order, and power, and that may be why so many cultures worship the sky in one way or another. Every culture divides the star patterns up to represent their heroes, gods, and symbolic creatures. Hercules looked down on the ancient Greeks, and the same stars represent the protector Båakkaataxpitchee (Bear Above) to the Crow people of North America. Among the hundreds of religions around the world, nearly all locate their gods and goddesses in the heavens. The gods watch over us from their remote and powerful thrones among the stars.

Our days are filled with necessary trivia, but astronomy enriches our lives by fitting us into the continuity of life on Earth. As you rush to an evening meeting, a glance at the sky will remind you that the sky carries our human heritage. Jesus, Moses, and Muhammad saw the same stars that you see. Aristotle watched the stars of Hercules rise in the east and set in the west just as you do. Astronomy helps us understand what we are by linking us to the past of human experience on this planet.

Study and Review

Summary

- Astronomers divide the sky into 88 **constellations (p. 11).** Although the constellations originated in Greek and Middle Eastern mythology, the names are Latin. Even the modern constellations, added to fill in the spaces between the ancient figures, have Latin names.
- Named groups of stars that are not constellations are called **asterisms (p. 11).**
- The names of stars usually come from ancient Arabic, though modern astronomers often refer to a star by its constellation and a Greek letter assigned according to its brightness within the constellation.
- Astronomers refer to the brightness of stars using the **magnitude scale (p. 13).** First-magnitude stars are brighter than second-magnitude stars, which are brighter than third-magnitude stars, and so on. The magnitude you see when you look at a star in the sky is its **apparent visual magnitude (p. 13),** which does not take into account its distance from Earth.
- **Flux (p. 14)** is a measure of light energy related to intensity. The magnitude of a star can be related directly to the flux of light received on Earth and so to its intensity.
- The **celestial sphere (p. 16)** is a **scientific model (p. 15)** of the sky, to which the stars appear to be attached. Because Earth rotates eastward, the celestial sphere appears to rotate westward on its axis.
- The **north** and **south celestial poles (p. 16)** are the pivots on which the sky appears to rotate, and they define the four directions around the **horizon (p. 16):** the **north, south, east,** and **west points (p. 16).** The point directly over head is the **zenith (p. 16),** and the point on the sky directly underfoot is the **nadir (p. 16).**
- The **celestial equator (p. 16),** an imaginary line around the sky above Earth's equator, divides the sky in half.
- Astronomers often refer to distances "on" the sky as if the stars, sun, moon, and planets were equivalent to spots painted on a plaster ceiling. These **angular distances (p. 17),** measured in degrees, **minutes of arc (p. 17),** and **seconds of arc (p. 17),** are unrelated to the true distance between the objects in light-years. The angular distance across an object is its **angular diameter (p. 17).**
- What you see of the celestial sphere depends on your latitude. Much of the southern hemisphere of the sky is not visible from northern latitudes. To see that part of the sky, you would have to travel southward over Earth's surface. **Circumpolar constellations (p. 17)** are those close enough to a celestial pole that they do not rise or set.
- The angular distance from the horizon to the north celestial pole always equals your latitude. This is the basis for celestial navigation.
- **Precession (p. 18)** is caused by the gravitational forces of the moon and sun acting on the spinning Earth and causing its axis to sweep around like that of a top. Earth's axis of rotation precesses with a period of 26,000 years, and consequently the celestial poles and celestial equator move slowly against the background of the stars.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why have astronomers added modern constellations to the sky?
2. What is the difference between an asterism and a constellation? Give some examples.
3. What characteristic do stars in a constellation or asterism share?
4. Do people from other cultures on Earth see the same stars, constellations, and asterisms that you see?
5. How does the Greek-letter designation of a star give you a clue to its brightness?
6. How did the magnitude system originate in a classification of stars by brightness?
7. What does the word *apparent* mean in *apparent visual magnitude?*
8. In what ways is the celestial sphere a scientific model?
9. Why do astronomers use the word *on* to describe angles *on* the sky rather than angles *in* the sky?
10. If Earth did not rotate, could you define the celestial poles and celestial equator?
11. Where would you go on Earth if you wanted to be able to see both the north celestial pole and the south celestial pole at the same time?
12. Where would you go on Earth to place a celestial pole at your zenith?
13. Explain how to make a simple astronomical observation that would determine your latitude.
14. Why does the number of circumpolar constellations depend on the latitude of the observer?
15. How could you detect Earth's precession by examining star charts from ancient Egypt?
16. **How Do We Know?** How can a scientific model be useful if it isn't a correct description of nature?

Discussion Questions

1. All cultures on Earth named constellations. Why do you suppose this was such a common practice?
2. If you were lost at sea, you could find your approximate latitude by measuring the altitude of Polaris. But Polaris isn't exactly at the celestial pole. What else would you need to know to measure your latitude more accurately?

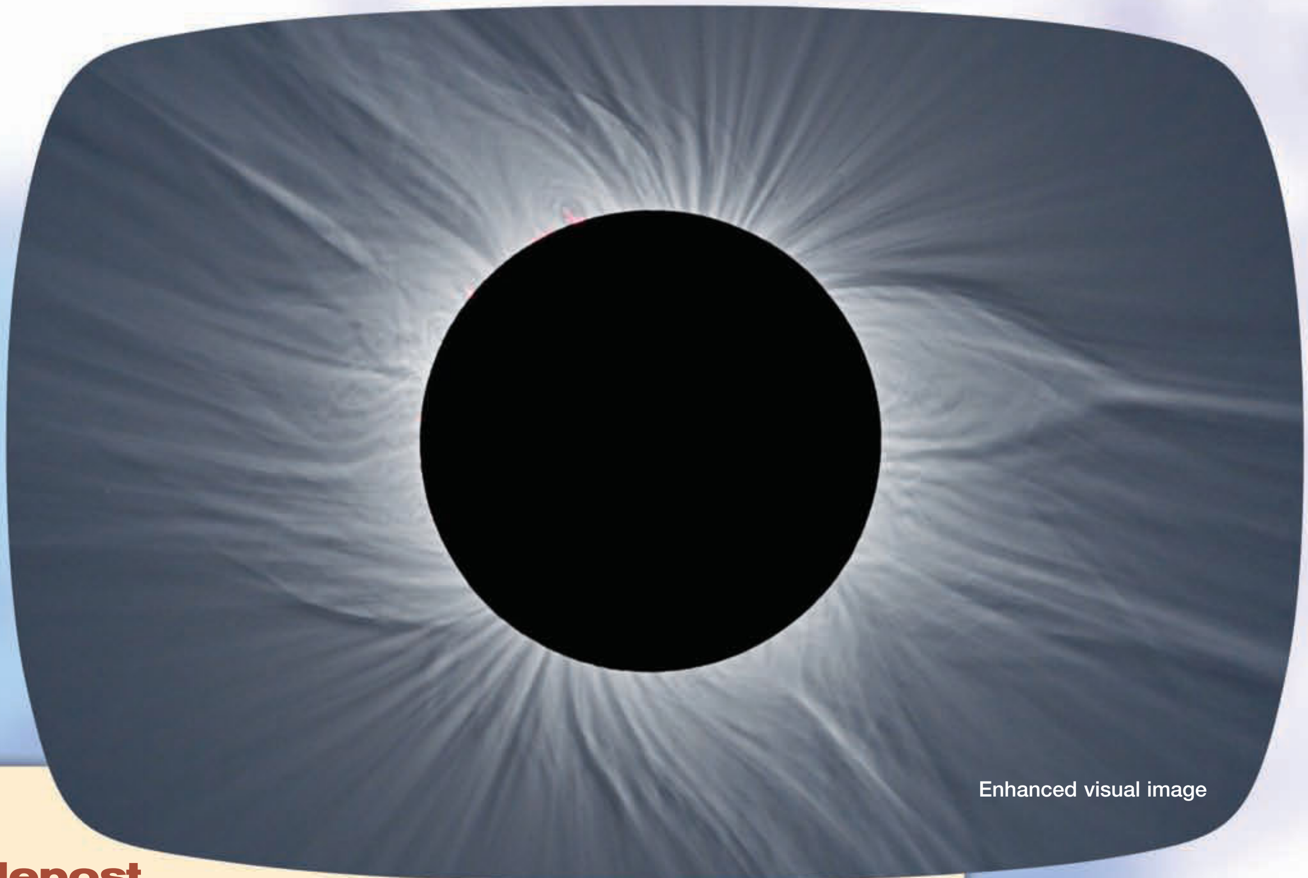
Problems

1. If light from one star is 40 times more intense than light from another star, what is their difference in magnitudes?
2. If two stars differ by 8.6 magnitudes, what is their intensity ratio?
3. Star A has a magnitude of 2.5; Star B, 5.5; and Star C, 9.5. Which is brightest? Which are visible to the unaided eye? Which pair of stars has an intensity ratio of 16?
4. By what factor is sunlight more intense than moonlight? (*Hint:* See Figure 2-6)
5. If you are at a latitude of 35 degrees north of Earth's equator, what is the angular distance from the northern horizon up to the north celestial pole? From the southern horizon down to the south celestial pole?

Learning to Look

1. Find Sagittarius and Scorpius in the photograph that opens this chapter.
2. The stamp at right shows the constellation Orion. Explain why this looks odd to residents of the northern hemisphere.

3 Cycles of the Sky



A total solar eclipse occurs when the moon crosses in front of the sun and hides its brilliant surface. Then you can see the sun's extended atmosphere. (©2001 F. Espenak, www.MrEclipse.com)

Guidepost

In the previous chapter you looked at the sky and saw how its motion is produced by the daily rotation of Earth. In this chapter, you will discover that the sun, moon, and planets move against the background of stars. Some of those motions have direct influences on your life and produce dramatic sights in the sky. As you explore, you will find answers to four essential questions:

What causes the seasons?

How can astronomical cycles affect Earth's climate?

Why does the moon go through phases?

What causes lunar and solar eclipses?

The cycles of the sky are elegant and dramatic, and you can understand them because you understand that Earth is a moving planet. That was not always so. How humanity first understood that Earth is a planet is the subject of the next chapter.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Even a man who is pure in heart and says
his prayers by night
May become a wolf when the wolfbane
blooms
and the moon shines full and bright.

PROVERB FROM OLD WOLFMAN MOVIES

YOUR ALARM CLOCK and your calendar are astronomical instruments that track the motion of the sun in the sky. Furthermore, your calendar is divided into months, and that recognizes the monthly orbital motion of the moon. Your life is regulated by the cycles of the sky, and the most obvious cycle is that of the sun.

3-1 Cycles of the Sun

THE SUN RISES and sets because Earth rotates on its axis, and that defines the day. In addition, Earth revolves around the sun in its orbit, and that defines the year. Notice an important distinction. **Rotation** is the turning of a body on its axis, but **revolution** means the motion of a body around a point outside the body. Consequently, astronomers are careful to say Earth rotates once a day on its axis and revolves once a year around the sun.

The Annual Motion of the Sun

Even in the daytime, the sky is filled with stars, but the glare of sunlight fills Earth's atmosphere with scattered light, and you can see only the brilliant sun. If the sun were fainter, you would be able to see it rise in the morning in front of the stars. During the day, you would see the sun and the stars moving westward, and the sun would eventually set in front of the same stars. If you watched carefully as the day passed, you would notice that the sun was creeping slowly eastward against the background of stars. It would move a distance roughly equal to its own diameter between sunrise and sunset. This motion is caused by the motion of Earth in its nearly circular orbit around the sun.

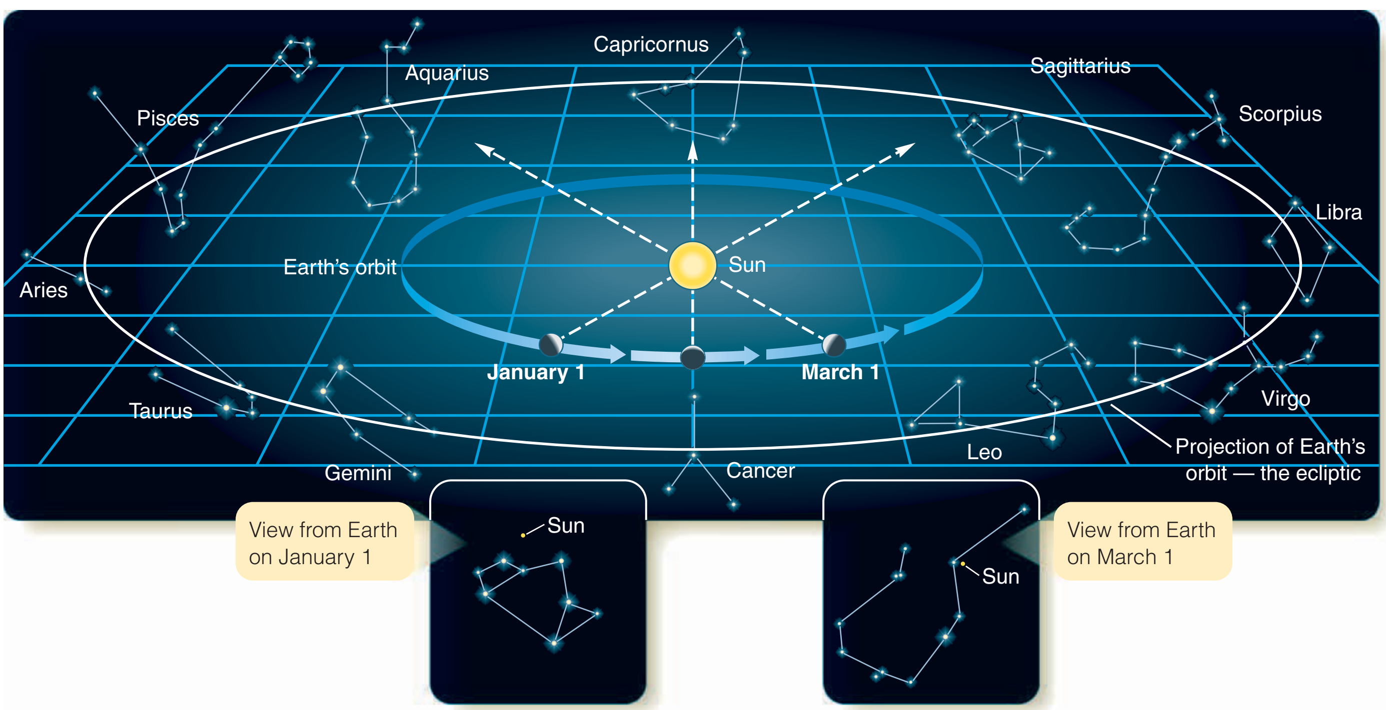
For example, in January, you would see the sun in front of the constellation Sagittarius (■ Figure 3-1). As Earth moves along its circular orbit, the sun appears to move eastward among the stars. By March, you would see it in front of Aquarius.

The apparent path of the sun against the background of stars is called the **ecliptic.** If the sky were a great screen, the ecliptic would be the shadow cast by Earth's orbit. That is why the ecliptic is often called the projection of Earth's orbit on the sky.

Earth circles the sun in 365.25 days, and consequently the sun appears to circle the sky in the same period. That means the sun, traveling 360° around the ecliptic in 365.25 days, travels about 1° eastward in 24 hours, about twice its angular diameter. You don't notice this apparent motion of the sun because you

■ **Figure 3-1**

Earth's orbit is a nearly perfect circle, but it is inclined in this diagram. Earth's motion around the sun makes the sun appear to move against the background of the stars. Earth's circular orbit is thus projected on the sky as the circular path of the sun, the ecliptic. If you could see the stars in the daytime, you would notice the sun crossing in front of the distant constellations as Earth moves along its orbit. **Animated!**



cannot see the stars in the daytime, but it does have an important consequence that you do notice—the seasons.

The Seasons

The seasons arise because of a simple fact: Earth's axis of rotation is tipped 23.5° from the perpendicular to its orbit. As you study ■ **The Cycle of the Seasons** on pages 24–25, notice two important principles and six new terms:

1. Because Earth's axis of rotation is inclined 23.5°, the sun moves into the northern sky in the spring and into the southern sky in the fall. That causes the cycle of the seasons. Notice how the *vernal equinox,* the *summer solstice,* the *autumnal equinox,* and the *winter solstice* mark the beginning of the seasons. Further, notice the very minor effects of Earth's slightly elliptical orbit as marked by the two terms *perihelion* and *aphelion.*

2. Earth goes through a cycle of seasons because of the changes in solar energy that Earth's northern and southern hemispheres receive at different times of the year. Because of circulation patterns in Earth's atmosphere, the northern and southern hemispheres are mostly isolated from each other and exchange little heat. When one hemisphere receives more solar energy than the other, it grows rapidly warmer.

Now you can set your friends straight if they mention two of the most **Common Misconceptions** about the seasons. First, the seasons don't occur because Earth moves closer to or farther from the sun. Earth's orbit is nearly circular. Its distance from the sun varies by less than 4 percent, and that doesn't cause the seasons. Second, it is not easier to stand a raw egg on end on the day of the vernal equinox! Have you heard that one? Radio and TV personalities love to talk about it, but it just isn't true. It is one of the silliest misconceptions in science. You can stand a raw egg on end any day of the year if you have steady hands. (*Hint:* It helps to shake the egg really hard to break the yolk inside so it can settle to the bottom.)

Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercises "Sunrise through the Seasons" and "The Seasons."

The Motion of the Planets

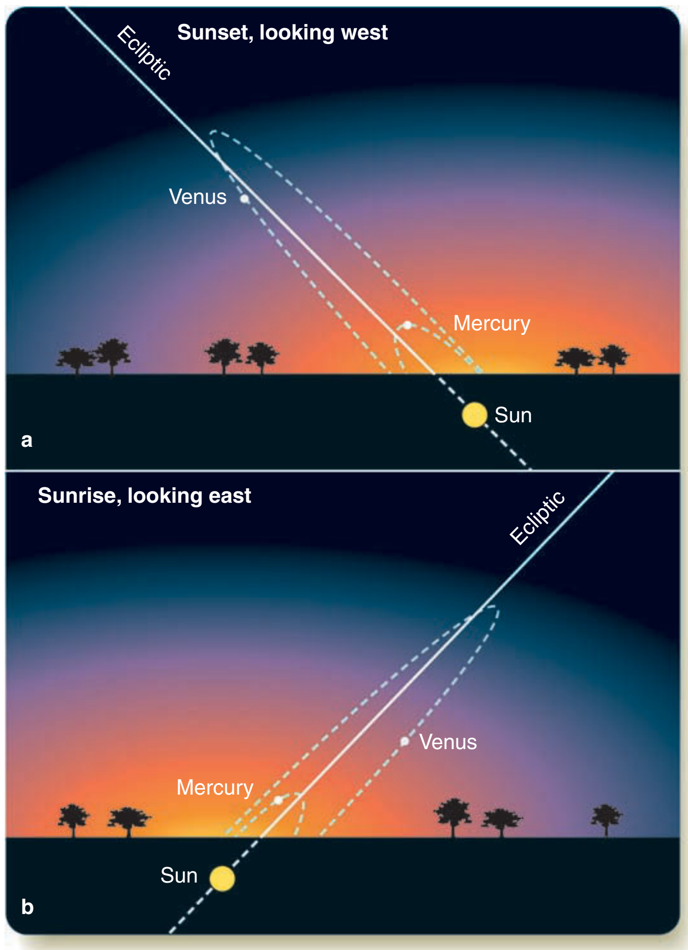
The planets of our solar system produce no visible light of their own; they are visible only by reflected sunlight. Mercury, Venus, Mars, Jupiter, and Saturn are all easily visible to the naked eye and look like stars, but Uranus is usually too faint to be seen, and Neptune is never bright enough.

All the planets of the solar system move in nearly circular orbits around the sun. If you were looking down on the solar system from the north celestial pole, you would see the planets moving in the same counterclockwise direction around their orbits, with the planets farthest from the sun moving the slowest.

When you look for planets in the sky, you always find them near the ecliptic because their orbits lie in nearly the same plane as Earth's orbit. The planets whose orbits lie outside Earth's orbit move slowly eastward along the ecliptic as they orbit the sun.* Mars moves completely around the ecliptic in slightly less than 2 years, but Saturn, being farther from the sun, takes nearly 30 years.

Mercury and Venus also stay near the ecliptic, but they move differently from the other planets. They have orbits inside Earth's orbit, and that means they can never move far from the sun in the sky. As seen from Earth, they move eastward away from the sun and then back toward the sun, crossing the near part of their orbit. Then they continue moving westward away from the sun and then move back crossing the far part of their orbit before they move out east of the sun again. To find one of these planets, you need to look above the western horizon just after sunset or above the eastern horizon just before sunrise. Venus is easier to locate because it is brighter and because its larger orbit carries it higher above the horizon than does Mercury's (■ Figure 3-2).

*You will discover occasional exceptions to this eastward motion in Chapter 4.



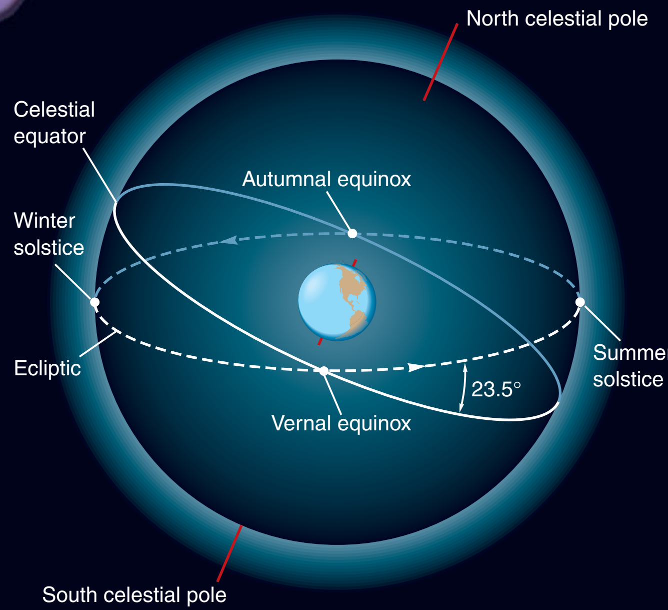
■ **Figure 3-2**

Mercury and Venus follow orbits that keep them near the sun, and they are visible only soon after sunset or before sunrise when the brilliance of the sun is hidden below the horizon. Venus takes 584 days to move from the morning sky to the evening sky and back again, but Mercury zips around in only 116 days.

The Cycle of the Seasons

1 You can use the celestial sphere to help you think about the seasons. The celestial equator is the projection of Earth's equator on the sky, and the ecliptic is the projection of Earth's orbit on the sky. Because Earth is tipped in its orbit, the ecliptic and equator are inclined to each other by 23.5° as shown at right. As the sun moves eastward around the sky, it spends half the year in the southern half of the sky and half of the year in the northern half. That causes the seasons.

The sun crosses the celestial equator going northward at the point called the **vernal equinox.** The sun is at its farthest north at the point called the **summer solstice.** It crosses the celestial equator going southward at the **autumnal equinox** and reaches its most southern point at the **winter solstice.**



1a The seasons are defined by the dates when the sun crosses these four points, as shown in the table at the right. *Equinox* comes from the word for "equal"; the day of an equinox has equal amounts of daylight and darkness. *Solstice* comes from the words meaning "sun" and "stationary." *Vernal* comes from the word for "green." The "green" equinox marks the beginning of spring.

Event	Date*	Season
Vernal equinox	March 20	Spring begins
Summer solstice	June 22	Summer begins
Autumnal equinox	September 22	Autumn begins
Winter solstice	December 22	Winter begins

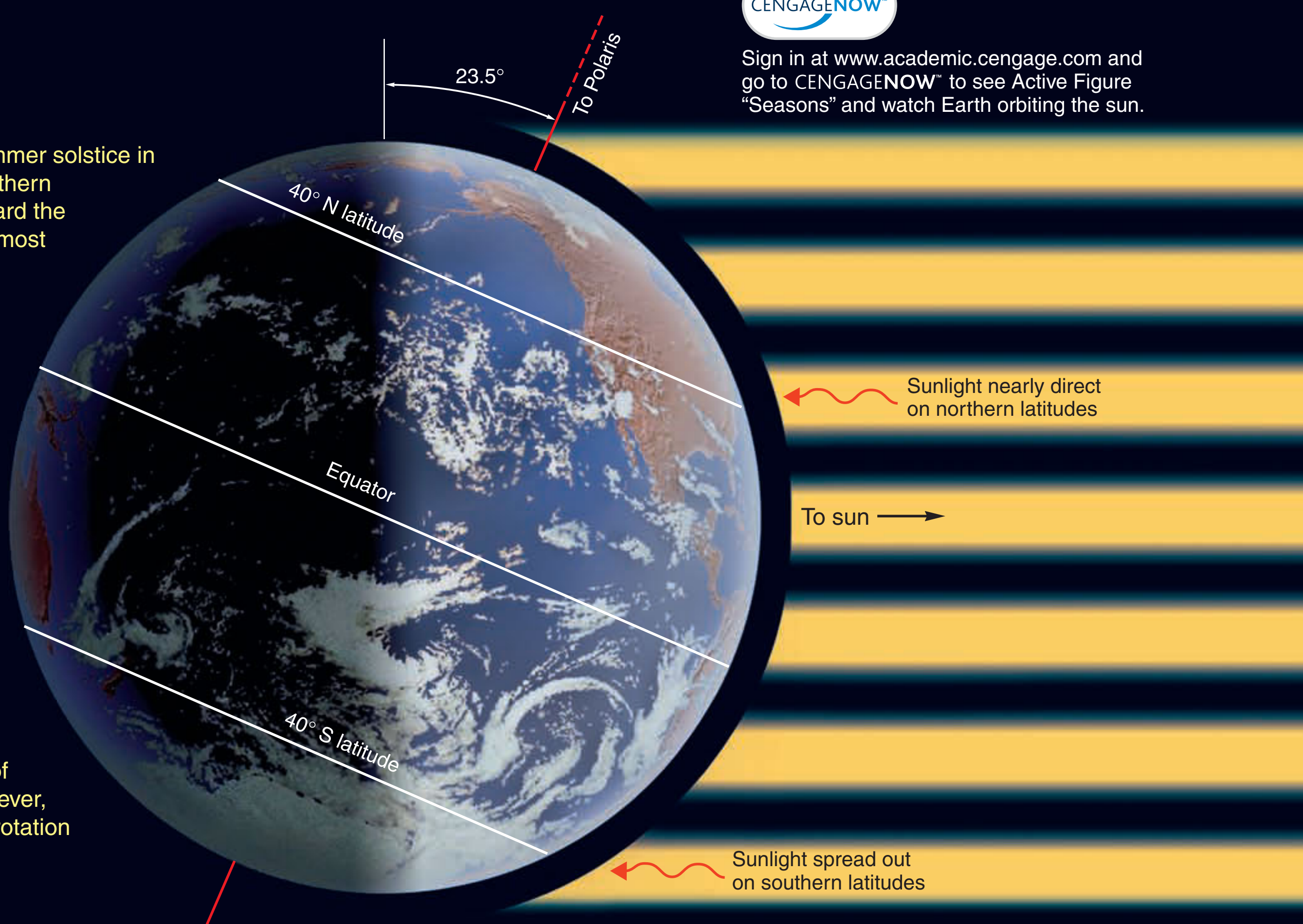
* Give or take a day due to leap year and other factors.

CENGAGENOW™

Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Seasons" and watch Earth orbiting the sun.

1b On the day of the summer solstice in late June, Earth's northern hemisphere is inclined toward the sun, and sunlight shines almost straight down at northern latitudes. At southern latitudes, sunlight strikes the ground at an angle and spreads out. North America has warm weather, and South America has cool weather.

Earth's axis of rotation points toward Polaris, and, like a top, the spinning Earth holds its axis fixed as it orbits the sun. On one side of the sun, Earth's northern hemisphere leans toward the sun; on the other side of its orbit, it leans away. However, the direction of the axis of rotation does not change.

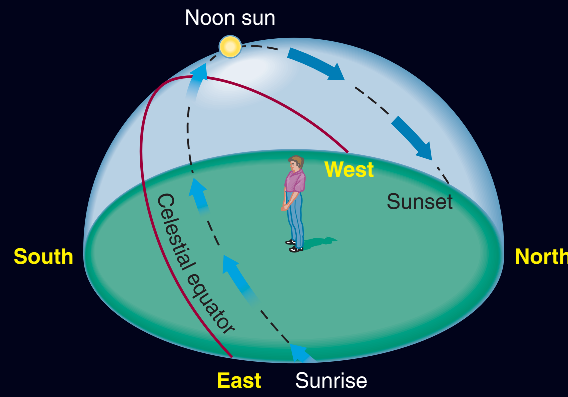


Earth at summer solstice

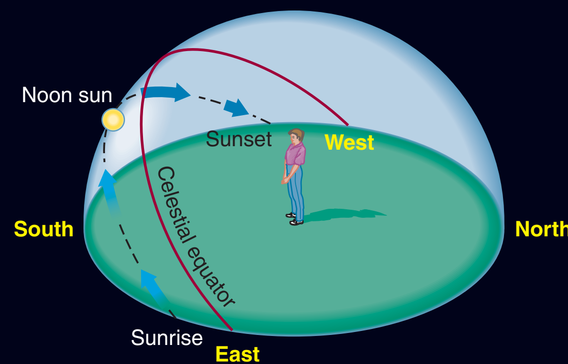
1c Light striking the ground at a steep angle spreads out less than light striking the ground at a shallow angle. Light from the summer-solstice sun strikes northern latitudes from nearly overhead and is concentrated.

Light from the winter-solstice sun strikes northern latitudes at a much shallower angle and spreads out. The same amount of energy is spread over a larger area, so the ground receives less energy from the winter sun.

2 The two causes of the seasons are shown at right for someone in the northern hemisphere. First, the noon summer sun is higher in the sky and the winter sun is lower, as shown by the longer winter shadows. Thus winter sunlight is more spread out. Second, the summer sun rises in the northeast and sets in the northwest, spending more than 12 hours in the sky. The winter sun rises in the southeast and sets in the southwest, spending less than 12 hours in the sky. Both of these effects mean that northern latitudes receive more energy from the summer sun, and summer days are warmer than winter days.



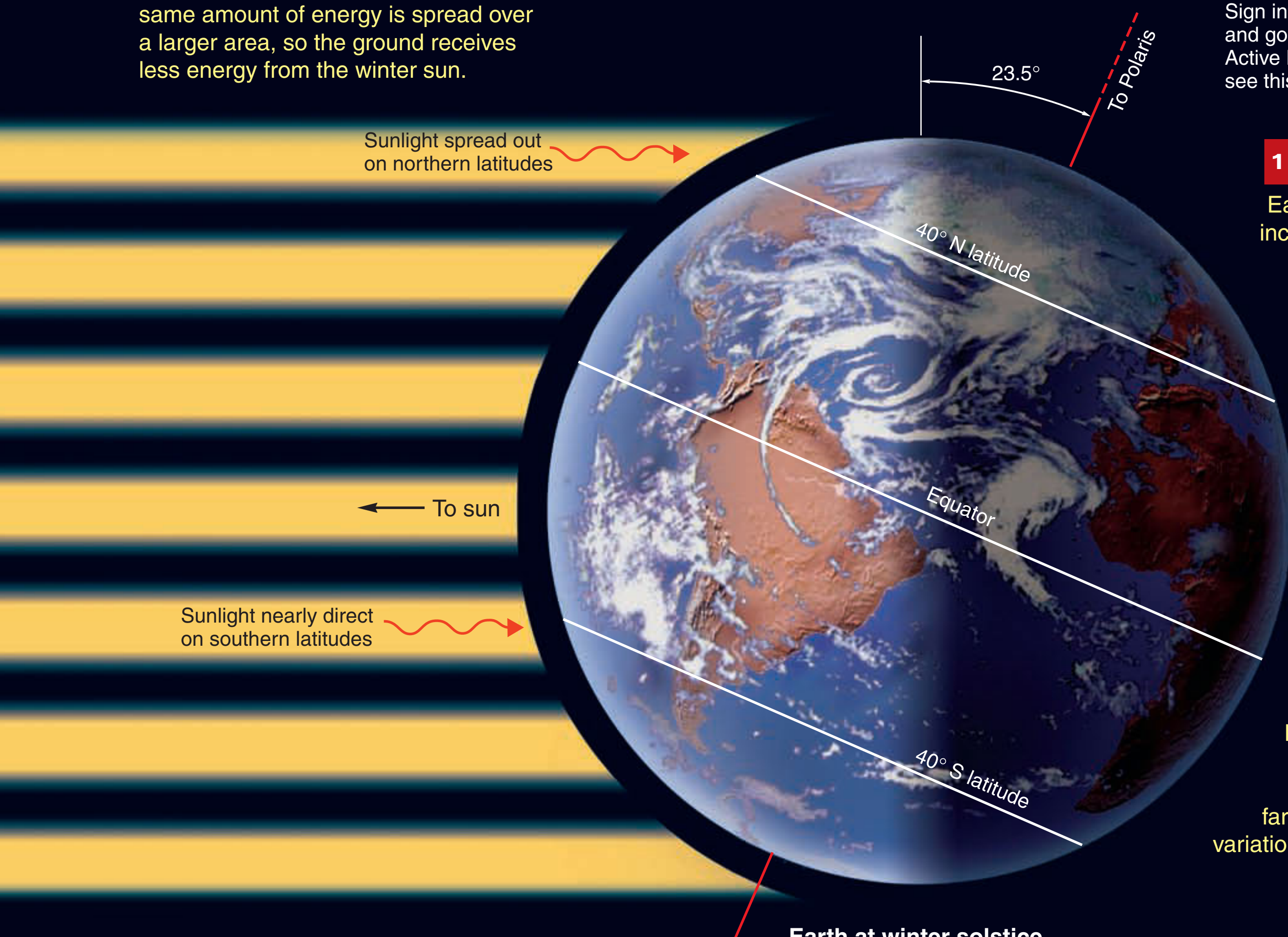
At summer solstice



At winter solstice

CENGAGENOW™

Sign in at www.academic.cengage.com and go to CENGAGE**NOW**™ to see Active Figure "Path of the Sun" and see this figure from the inside.



Earth at winter solstice

1d On the day of the winter solstice in late December, Earth's northern hemisphere is inclined away from the sun, and sunlight strikes the ground at an angle and spreads out. At southern latitudes, sunlight shines almost straight down and does not spread out. North America has cool weather and South America has warm weather.

Earth's orbit is only very slightly elliptical. About January 3, Earth is at **perihelion,** its closest point to the sun, when it is only 1.7 percent closer than average. About July 5, Earth is at **aphelion,** its most distant point from the sun, when it is only 1.7 percent farther than average. This small variation does not significantly affect the seasons.

How Do We Know? 3-1

Pseudoscience

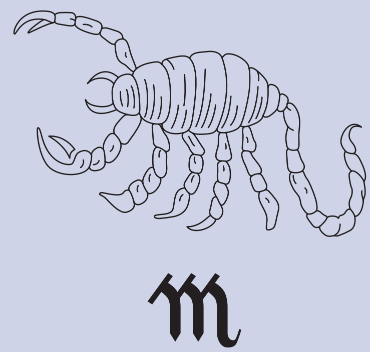
What is the difference between a science and a pseudoscience? Astronomers have a low opinion of beliefs such as astrology, not so much because they are groundless but because they *pretend* to be a sciences. They are **pseudosciences,** from the Greek *pseudo,* meaning false.

A pseudoscience is a set of beliefs that appear to be based on scientific ideas but that fail to obey the most basic rules of science. For example, in the 1970s a claim was made that pyramidal shapes focus cosmic forces on anything underneath and might even have healing properties. For example, it was claimed that a pyramid made of paper, plastic, or other materials would preserve fruit, sharpen razor blades, and do other miraculous things. Many books promoted the idea of the special power of pyramids, and this idea led to a popular fad.

A key characteristic of science is that its claims can be tested and verified. In this case, simple experiments showed that any shape, not just a pyramid, protects a piece of fruit from airborne spores and allows it to dry without rotting. Likewise, any shape allows oxidation to improve the cutting edge of a razor blade. Because experimental evidence contradicted the claim and because supporters of the theory declined to abandon or revise their claims, you can recognize pyramid power as a pseudoscience. Disregard of contradictory evidence and alternate theories is a sure sign of a pseudoscience.

Pseudoscientific claims can be self-fulfilling. For example, some believers in pyramid power slept under pyramidal tents to improve their rest. There is no logical mechanism by which such a tent could affect a sleeper, but because people wanted and expected the claim to be true they reported that they slept more soundly. Vague claims based on personal testimony that cannot be tested are another sign of a pseudoscience.

Astrology is a pseudoscience. It has been tested over and over for centuries, and it doesn't work. Nevertheless, many people believe in astrology despite contradictory evidence. Many pseudosciences appeal to our need to understand and control the world around us. Some such claims involve medical cures, ranging from using magnetic bracelets and crystals to focus mystical power to astonishingly expensive, illegal, and dangerous treatments for cancer. Logic is a stranger to pseudoscience, but human fears and needs are not.



Astrology may be the oldest pseudoscience.

Mercury's orbit is so small that it can never get farther than 28° from the sun. Consequently, it is hard to see against the sun's glare and is often hidden in the clouds and haze near the horizon.

By tradition, any planet visible in the evening sky is called an **evening star,** even though planets are not stars. Similarly, any planet visible in the sky shortly before sunrise is called a **morning star.** Perhaps the most beautiful is Venus, which can become as bright as magnitude −4.7. As Venus moves around its orbit, it can dominate the western sky each evening for many weeks, but eventually its orbit carries it back toward the sun, and it is lost in the haze near the horizon. In a few weeks, it reappears in the dawn sky, a brilliant morning star.

The cycles of the sky are so impressive that it is not surprising that people have strong feelings about them. Ancient peoples saw the motion of the sun around the ecliptic as a powerful influence on their daily lives, and the motion of the planets along the ecliptic seemed similarly meaningful. The ancient superstition of astrology is based on the cycle of the sun and planets around the sky. You have probably heard of the **zodiac,** a band around the sky extending 9 degrees above and below the ecliptic. The signs of the zodiac take their names from the 12 principal constellations along the ecliptic. Centuries ago astrology was an important part of astronomy, but the two are now almost exact opposites—astronomy is a science that depends on evidence, and astrology is a superstition that survives in spite of evidence (■ How Do We Know? 3-1). The signs of the zodiac are no longer important in astronomy.

3-2 Astronomical Influences on Earth's Climate

The seasons are produced by the annual motion of Earth around the sun, but subtle changes in that motion can have dramatic effects on climate. You don't notice these changes during your lifetime, but over thousands of years, they can bury continents under glaciers.

Earth has gone through ice ages, when the worldwide climate was cooler and dryer and thick layers of ice covered northern latitudes. One major ice age occurred about 570 million years ago, and the next about 280 million years ago. The most recent ice age began only about 3 million years ago and is still going on. You are living during one of the periodic episodes during an ice age when the glaciers melt back and Earth grows slightly warmer. The current warm period began about 12,000 years ago.

Ice ages seem to occur with a period of roughly 250 to 300 million years, and cycles of glaciation within ice ages occur with periods of 40,000 to 100,000 years. (These cycles have no connection with global warming, which can produce changes in Earth's climate over just a few decades. Global warming is discussed in Chapter 17.) Evidence shows that these slow cycles of the ice ages have an astronomical origin.

The Hypothesis

Sometimes a theory or hypothesis is proposed long before scientists can find the critical evidence to test it. That happened in 1920 when Yugoslavian meteorologist Milutin Milankovitch proposed what became known as the **Milankovitch hypothesis**—that small changes in Earth's orbit, precession, and inclination affect Earth's climate and can trigger ice ages. You should examine each of these motions separately.

First, Earth's orbit is only very slightly elliptical, but astronomers know that the elliptical shape varies slightly over a period of about 100,000 years. At present, Earth's orbit carries it 1.7 percent closer than average to the sun during northern hemisphere winters and 1.7 percent farther away in northern hemisphere summers. This makes the northern climate very slightly warmer, and that is critical—most of the landmass where ice can accumulate is in the northern hemisphere. If Earth's orbit became more elliptical, for example, northern summers might be too cool to melt all of the snow and ice from the previous winter. That would make glaciers grow larger.

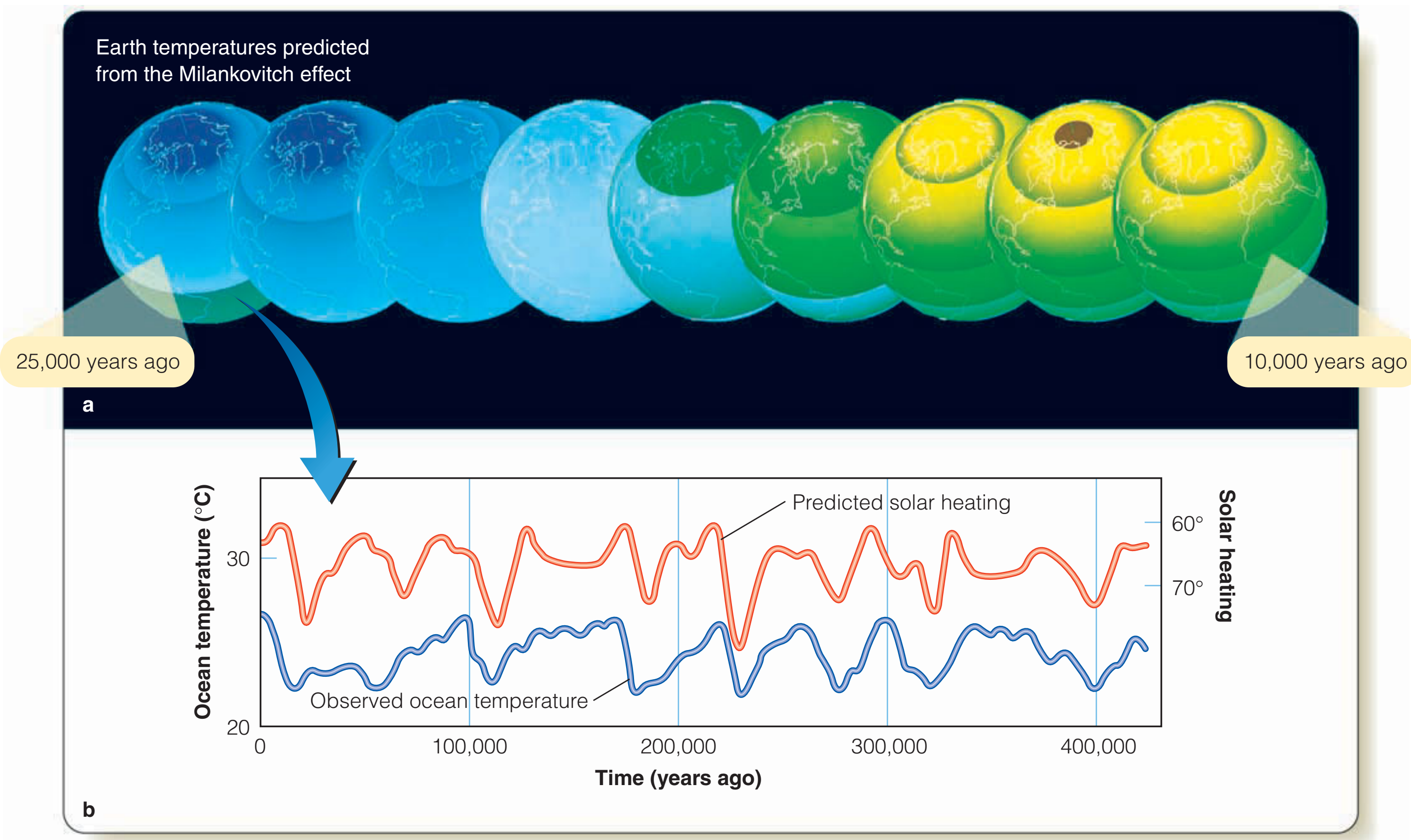
A second factor is also at work. Precession causes Earth's axis to sweep around a cone with a period of about 26,000 years, and that gradually changes the points in Earth's orbit where a given hemisphere experiences the seasons. Northern hemisphere summers now occur when Earth is 1.7 percent farther from the sun, but in 13,000 years northern summers will occur on the other side of Earth's orbit where Earth is 1.7 percent closer to the sun. Northern summers will be warmer, which could melt all of the previous winter's snow and ice and prevent the growth of glaciers.

The third factor is the inclination of Earth's equator to its orbit. Currently at 23.5°, this angle varies from 22° to 24°, with a period of roughly 41,000 years. When the inclination is greater, seasons are more severe.

In 1920, Milankovitch proposed that these three factors cycle against each other to produce complex periodic variations in Earth's climate and the advance and retreat of glaciers (■ Figure 3-3a). But no evidence was available to test the theory in 1920, and scientists treated it with skepticism. Many thought it was laughable.

The Evidence

By the middle 1970s, Earth scientists could collect the data that Milankovitch had lacked. Oceanographers could drill deep into the seafloor and collect long cores of sediment. In the laboratory,



■ **Figure 3-3**

(a) Mathematical models of the Milankovitch effect can be used to predict temperatures on Earth over time. In these Earth globes, cool temperatures are represented by violet and blue and warm temperatures by yellow and red. These globes show the warming that occurred beginning 25,000 years ago, which ended the last ice age. (Courtesy Arizona State University, Computer Science and Geography Departments) (b) Over the last 400,000 years, changes in ocean temperatures measured from fossils found in sediment layers from the seabed match calculated changes in solar heating. (Adapted from Cesare Emiliani)

How Do We Know? 3-2

Evidence as the Foundation of Science

Why is evidence critical in science? From colliding galaxies to the inner workings of atoms, scientists love to speculate and devise theories, but all scientific knowledge is ultimately based on evidence from observations and experiments. Evidence is reality, and scientists constantly check their ideas against reality.

When you think of evidence, you probably think of criminal investigations in which detectives collect fingerprints and eyewitness accounts. In court, that evidence is used to try to understand the crime, but there is a key difference in how lawyers and scientists use evidence. A defense attorney can call a witness and intentionally fail to ask a question that would reveal evidence harmful to the defendant. In contrast, the scientist must be objective and not ignore any known evidence. The attorney is presenting only one side of the case, but the scientist is searching for the truth. In a sense, the scientist must deal with the evidence as both the prosecution and the defense.

It is a characteristic of scientific knowledge that it is supported by evidence. A scientific statement is more than an opinion or a speculation because it has been tested objectively against reality.

As you read about any science, look for the evidence in the form of observations and experiments. Every theory or conclusion should have supporting evidence. If you can find and understand the evidence, the science will make sense. All scientists, from astronomers to zoologists, demand evidence. You should, too.

Fingerprints are evidence to past events. (Dorling Kindersley/Getty Images)

geologists could take samples from different depths in the cores and determine the age of the samples and the temperature of the oceans when they were deposited on the sea floor. From this, scientists constructed a history of ocean temperatures that convincingly matched the predictions of the Milankovitch hypothesis (Figure 3-3b).

The evidence seemed very strong, and by the 1980s the Milankovitch hypothesis was widely considered the leading hypothesis. But science follows a mostly unstated set of rules that holds that a hypothesis must be tested over and over against all available evidence (■ How Do We Know? 3-2). In 1988, scientists discovered contradictory evidence.

For 500,000 years rainwater has collected in a deep crack in Nevada called Devil's Hole. That water has deposited the mineral calcite in layer on layer on the walls of the crack. It isn't easy to get to, and scientists had to dive with scuba gear to drill out samples of the calcite, but it was worth the effort. Back in the laboratory, they could determine the age of each layer in their core samples and the temperature of the rainwater that had formed the calcite in each layer. That gave them a history of temperatures at Devil's Hole that spanned many thousands of years, and the results were a surprise. The evidence seemed to show that Earth had begun warming up thousands of years too early for the last ice age to have been caused by the Milankovitch cycles.

These contradictory findings are irritating because we humans naturally prefer certainty, but such circumstances are common in science. The disagreement between ocean floor samples and Devil's Hole samples triggered a scramble to understand the problem. Were the ages of one or the other set of samples wrong? Were the ancient temperatures wrong? Or were scientists misunderstanding the significance of the evidence?

In 1997, a new study of the ages of the samples confirmed that those from the ocean floor are correctly dated. But the same study found that the ages of the Devil's Hole samples are also correct. Evidently the temperatures at Devil's Hole record local climate changes in the region that became the southwestern United States. The ocean floor samples record global climate changes, and they fit well with the Milankovitch hypothesis. This gave scientists renewed confidence in the Milankovitch hypothesis, and although it is widely accepted today, it is still being tested whenever scientists can find more evidence.

As you review this section, notice that it is a **scientific argument,** a careful presentation of theory and evidence in a logical discussion. ■ How Do We Know? 3-3 expands on the ways scientists organize their ideas in logical arguments. Throughout this book, many chapter sections end with short reviews called "Scientific Argument." These feature a review question, which is then analyzed in a scientific argument. A second question gives you a chance to build your own scientific argument. You can use

How Do We Know? 3-3

Scientific Arguments

How is a scientific argument different from an advertisement? Advertisements sometimes sound scientific, but they are fundamentally different from scientific arguments. An advertisement is designed to convince you to buy a product. "Our shampoo promises 85 percent shinier hair." The statement may sound like science, but it isn't a complete, honest discussion. "Shinier than what?" you might ask. An advertiser's only goal is a sale.

Scientists construct arguments because they want to test their own ideas and give an accurate explanation of some aspect of nature. For example, in the 1960s, biologist E. O. Wilson presented a scientific argument to show that ants communicate by smells. The argument included a description of his careful observations and the ingenious experiments he had conducted to test his theory. He also considered other evidence and other theories for ant communication. Scientists can include any evidence or theory that supports their claim, but they must observe one fundamental rule of science: They must be totally honest—they must include all of the evidence and all of the theories.

Scientists publish their work in scientific arguments, but they also think in scientific arguments. If, in thinking through his argument, Wilson had found a contradiction, he would have known he was on the wrong track. That is why scientific arguments must be complete and honest. Scientists who ignore inconvenient evidence or brush aside other theories are only fooling themselves.

A good scientific argument gives you all the information you need to decide for yourself whether the argument is correct. Wilson's study of ant communication is now widely understood and is being applied to other fields such as pest control and telecommunications networks.

Scientists have discovered that ants communicate with a large vocabulary of smells. (Eye of Science/Photo Researchers, Inc.)

these "Scientific Argument" features to review chapter material but also to practice thinking like a scientist.

◄ SCIENTIFIC ARGUMENT ►

Why should precession affect Earth's climate?

Here exaggeration is a useful analytical tool in your argument. If you exaggerate the elliptical shape of Earth's orbit, you can see dramatically the influence of precession. At present, Earth reaches perihelion (closest to the sun) during winter in the northern hemisphere and aphelion (farthest from the sun) during summer. The variation in distance is only 1.7 percent, and that difference doesn't cause much change in the severity of the seasons. But if Earth's orbit were much more elliptical, then winter in the northern hemisphere would be much warmer, and summer would be much cooler.

Now you can see the importance of precession. As Earth's axis precesses, the seasons occur at different places around Earth's orbit. In 13,000 years, northern winter will occur at aphelion, and, if Earth's orbit were highly elliptical, northern winter would be terribly cold. Similarly, summer would occur at perihelion, and the heat would be awful. Such extremes might deposit large amounts of ice in the winter but then melt it away in the hot summer, thus preventing the accumulation of glaciers.

Continue this analysis by modifying in your scientific argument further. **What effect would precession have if Earth's orbit were more circular?**

◄ ►

3-3 The Cycles of the Moon

You have no doubt seen the moon in the sky and noticed that its shape changes from night to night. The cycle of the moon is one of the most obvious phenomena in the sky, and that cycle has been a natural timekeeper since before the dawn of human civilization.

The Motion of the Moon

Just as the planets revolve counterclockwise around the sun, the moon revolves counterclockwise around Earth. Because the moon's orbit is tipped a few degrees from the plane of Earth's orbit, the moon's path takes it slightly north and then slightly south of the ecliptic, but it is always somewhere along the band of the zodiac.

The moon moves rapidly against the background of the constellations. If you watch the moon for just an hour, you can see it move eastward by slightly more than its angular diameter. In the previous chapter, you learned that the moon is about 0.5° in angular diameter, so it moves eastward a bit more than 0.5° per hour. In 24 hours, it moves 13°. Each night you see the moon about 13° eastward of its location the night before.

As the moon orbits around Earth, its shape changes from night to night in a month-long cycle.

The Cycle of Phases

The changing shape of the moon as it revolves around Earth is one of the most easily observed phenomena in astronomy. Study ■ **The Phases of the Moon** on pages 32–33 and notice three important points and two new terms:

1. The moon always keeps the same side facing Earth. "The man in the moon" is produced by the familiar features on the moon's near side, but you never see the far side of the moon.

2 The changing shape of the moon as it passes through its cycle of phases is produced by sunlight illuminating different parts of the side of the moon you can see.

3 Notice the difference between the orbital period of the moon around Earth *(sidereal period)* and the length of the lunar phase cycle *(synodic period)*. That difference is a good illustration of how your view from Earth is produced by the combined motions of Earth and other heavenly bodies such as the sun and moon.

You can make a moon-phase dial from the middle diagram on page 32 by covering the lower half of the moon's orbit with a sheet of paper and aligning the edge of the paper to pass through the word "Full" at the left and the word "New" at the right. Push a pin through the edge of the paper at Earth's North Pole to make a pivot and, under the word "Full," write on the paper "Eastern Horizon." Under the word "New," write "Western Horizon." The paper now represents the horizon you see when you stand facing south. You can set your moon-phase dial for a given time by rotating the diagram behind the horizon-paper. For example, set the dial to sunset by turning the diagram until the human figure labeled "sunset" is standing at the top of the Earth globe; the dial shows, for example, that the full moon at sunset would be at the eastern horizon.

The phases of the moon are dramatic, and they have attracted a number of peculiar ideas. You have probably heard a number of **Common Misconceptions** about the moon. Sometimes people are surprised to see the moon in the daytime sky, and they think something has gone wrong! No, the gibbous moon is often visible in the daytime, although quarter moons and especially crescent moons are harder to see when the sun is above the horizon. You may hear people mention "the dark side of the moon," but you will be able to assure them that there is no dark side. Any location on the moon is sunlit for two weeks and is in darkness for two weeks as the moon rotates in sunlight. Also, you may have heard people say the moon is larger when it is on the horizon. Certainly the rising full moon looks big when you see it on the horizon, but that is an optical illusion. In reality, the moon is the same angular diameter on the horizon as when it is high overhead. Finally, you have probably heard one of the strangest misconceptions about the moon: that people tend to act up at full moon. Actual statistical studies of records from schools, prisons, hospitals, and so on show that it isn't true. There are always a few people who misbehave; the moon has nothing to do with it.

For billions of years, the man in the moon has looked down on Earth. Ancient civilizations saw the same cycle of phases that you see (■ Figure 3-4), and even the dinosaurs may have noticed the changing phases of the moon. Occasionally, however, the moon displays more complicated moods when it turns copper-red in a lunar eclipse.

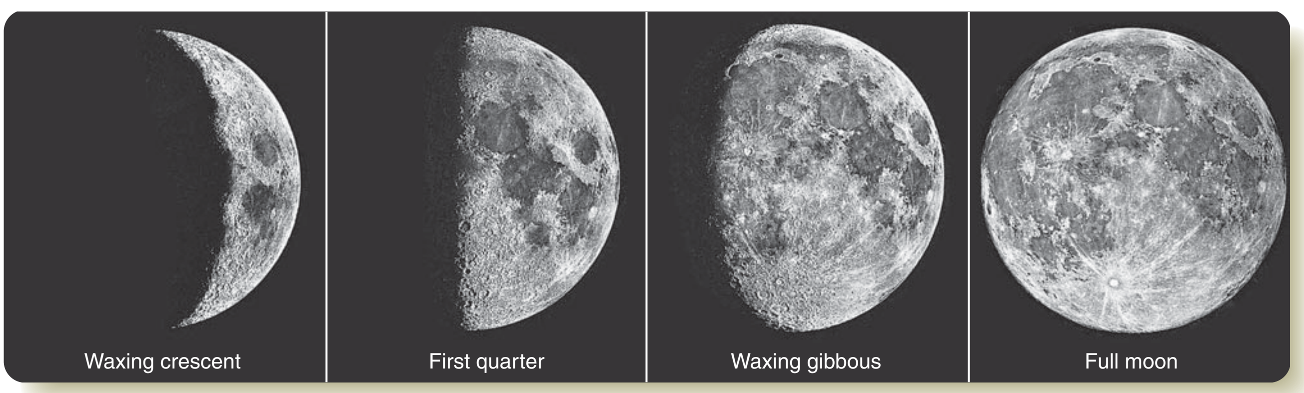
Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercises "Phases of the Moon" and "Moon Calendar."

Lunar Eclipses

A **lunar eclipse** can occur at full moon if the moon moves through the shadow of Earth. Because the moon shines only by reflected sunlight, the moon grows dark while it is crossing through the shadow.

Earth's shadow consists of two parts (■ Figure 3-5). The **umbra** is the region of total shadow. If you were drifting in your spacesuit in the umbra of Earth's shadow, the sun would be completely hidden behind Earth, and you would see no portion of the sun's bright disk. If you drifted into the **penumbra,** however, you would see part of the sun peeking around the edge of Earth, so you would be in partial shadow. In the penumbra, sunlight is dimmed but not extinguished.

Once or twice a year, the orbit of the moon carries it through the umbra of Earth's shadow, and you see a total lunar eclipse (■ Figure 3-6). As you watch the eclipse begin, the moon first moves into the penumbra and dims slightly; the deeper it moves



■ **Figure 3-4**

In this sequence of the waxing moon, you see the same face of the moon, the same mountains, craters, and plains, but the changing direction of sunlight produces the lunar phases. (©UC Regents/Lick Observatory)

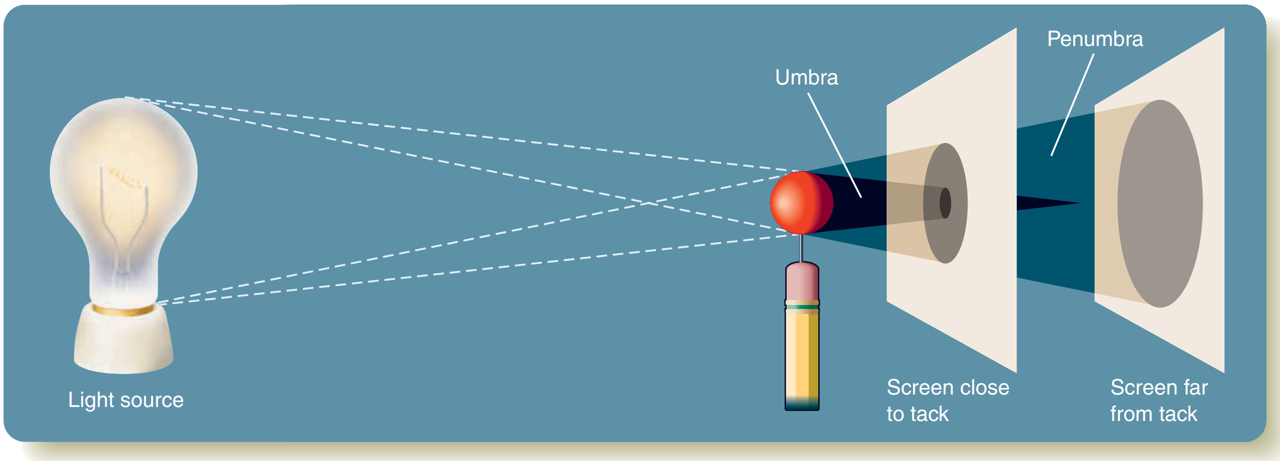


Figure 3-5

The shadows cast by a map tack resemble those of Earth and the moon. The umbra is the region of total shadow; the penumbra is the region of partial shadow.

into the penumbra, the more it dims. After about an hour, the moon reaches the umbra, and you see the umbral shadow darken part of the moon. It takes about an hour for the moon to enter the umbra completely and become totally eclipsed. **Totality,** the period of total eclipse, may last as long as 1 hour 45 minutes, though the length of totality depends on where the moon crosses the shadow.

When the moon is totally eclipsed, it does not disappear completely. Although it receives no direct sunlight, the moon in the umbra does receive some sunlight that is refracted (bent) through Earth's atmosphere. If you were on the moon during totality, you would not see any part of the sun because it would be entirely hidden behind Earth. However, you would see Earth's atmosphere illuminated from behind by the sun. The red glow from this ring of "sunsets" and "sunrises" illuminates the moon during totality and makes it glow coppery red, as shown in Figure 3-6.

Lunar eclipses are not always total. If the moon passes a bit too far north or south, it may only partially enter the umbra, and you see a partial lunar eclipse. The part of the moon that remains in the penumbra receives some direct sunlight, and the glare is usually great enough to prevent your seeing the faint coppery glow of the part of the moon in the umbra.

A penumbral lunar eclipse occurs when the moon passes through the penumbra but misses the umbra entirely. Because the penumbra is a region of partial shadow, the moon is only partially dimmed. A penumbral eclipse is not very impressive.

Although there are usually no more than one or two lunar eclipses each year, it is not difficult to see one. You need only be on the dark side of Earth when the moon passes through Earth's shadow. That is, the eclipse must occur between sunset and sunrise at

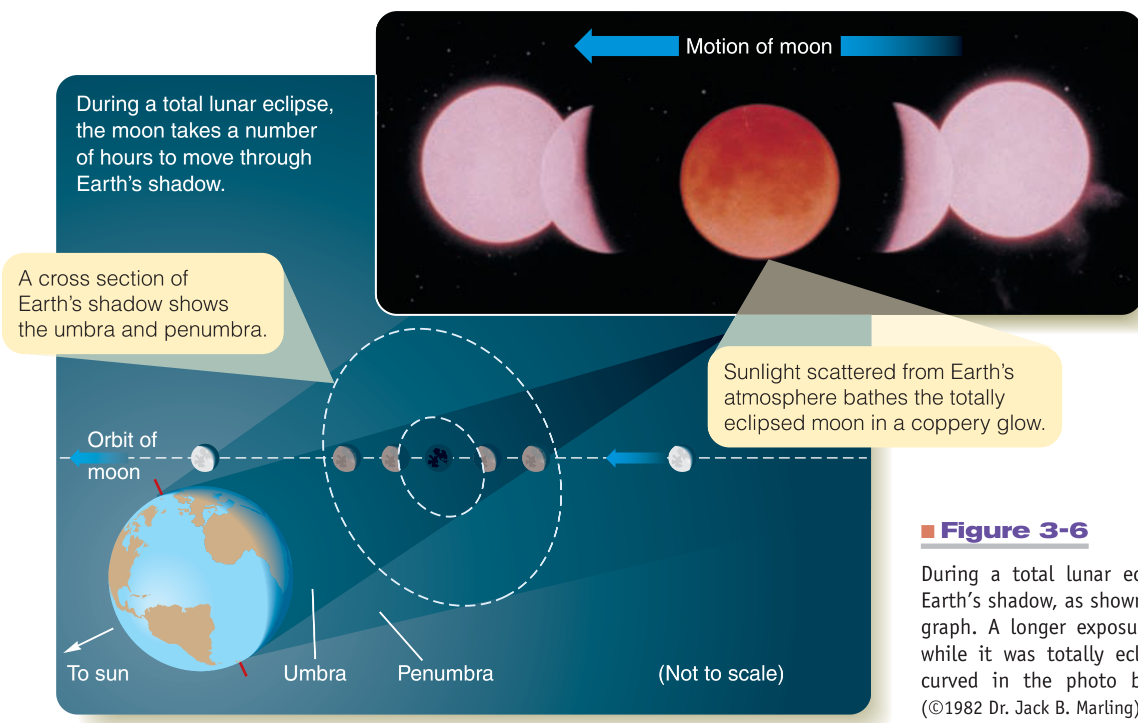
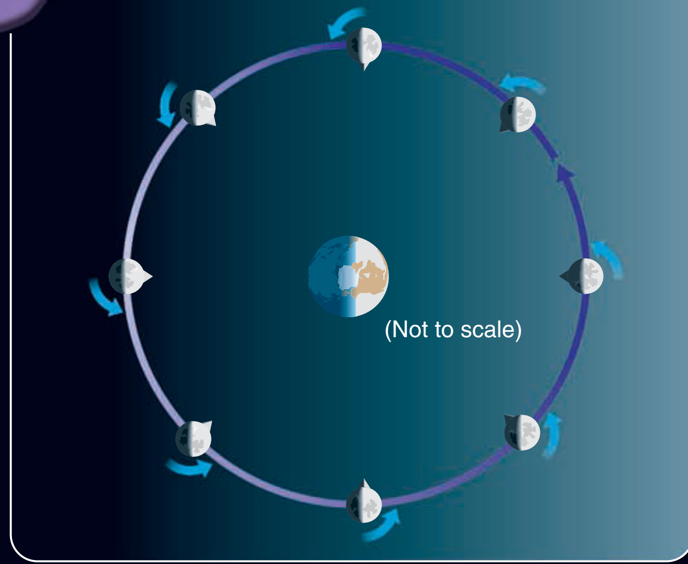


Figure 3-6

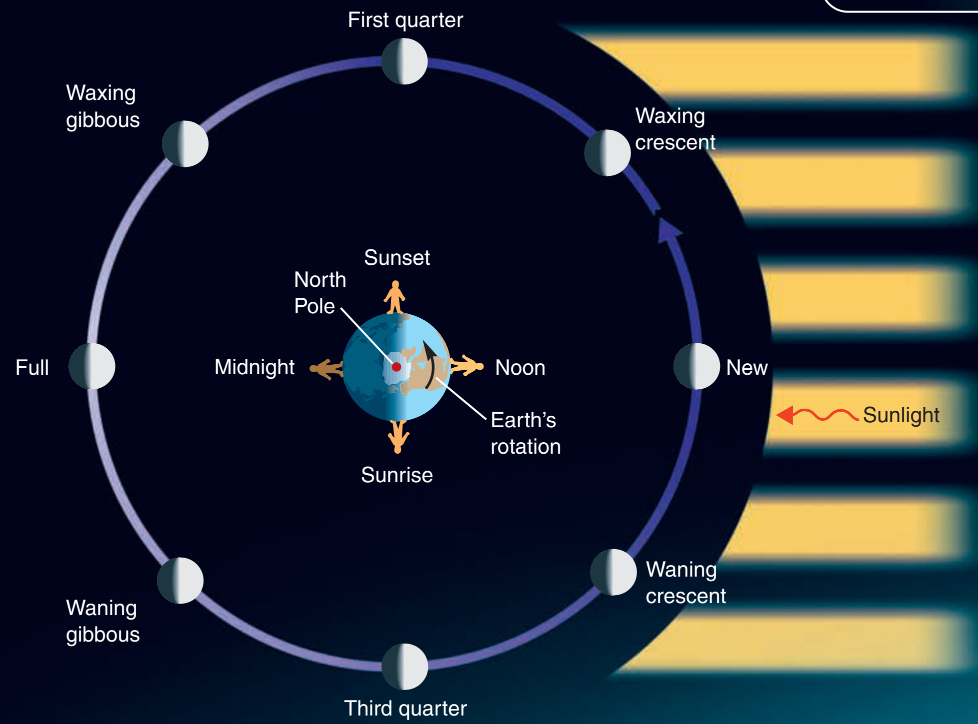
During a total lunar eclipse, the moon passes through Earth's shadow, as shown in this multiple-exposure photograph. A longer exposure was used to record the moon while it was totally eclipsed. The moon's path appears curved in the photo because of photographic effects. (©1982 Dr. Jack B. Marling)

The Phases of the Moon

1 As the moon orbits Earth, it rotates to keep the same side facing Earth as shown at right. Consequently you always see the same features on the moon, and you never see the far side of the moon. A mountain on the moon that points at Earth will always point at Earth as the moon revolves and rotates.



CENGAGENOW Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Lunar Phases" and take control of this diagram.

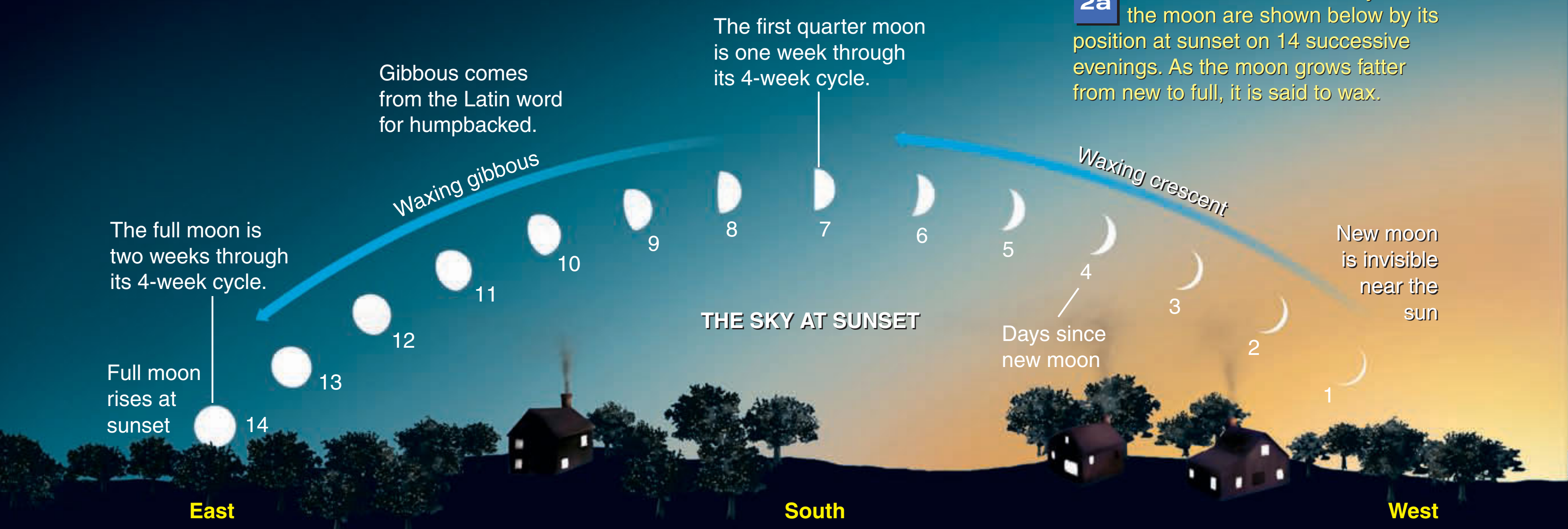


2 As seen at left, sunlight always illuminates half of the moon. Because you see different amounts of this sunlit side, you see the moon cycle through phases. At the phase called "new moon," sunlight illuminates the far side of the moon, and the side you see is in darkness. At new moon you see no moon at all. At full moon, the side you see is fully lit, and the far side is in darkness. How much you see depends on where the moon is in its orbit.

Notice that there is no such thing as the "dark side of the moon." All parts of the moon experience day and night in a month-long cycle.

In the diagram at the left, you see that the new moon is close to the sun in the sky, and the full moon is opposite the sun. The time of day depends on the observer's location on Earth.

2a The first 2 weeks of the cycle of the moon are shown below by its position at sunset on 14 successive evenings. As the moon grows fatter from new to full, it is said to wax.

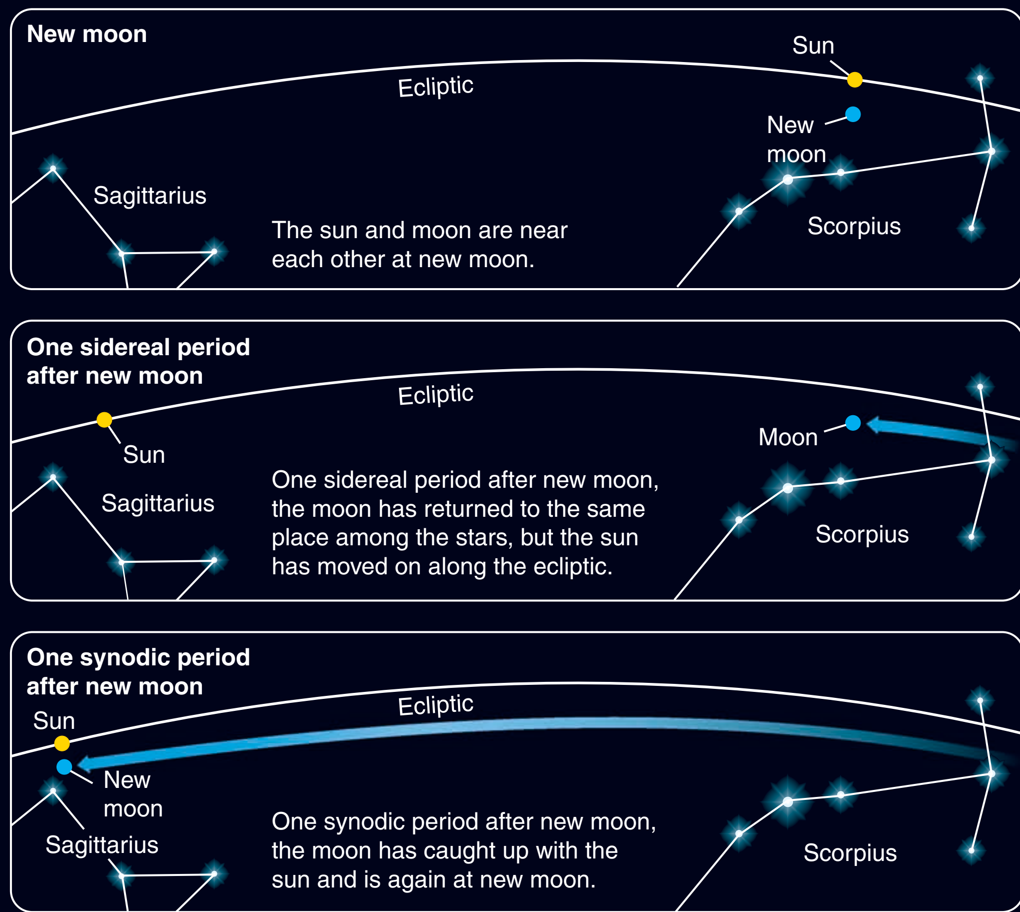


3 The moon orbits eastward around Earth in 27.32 days, its **sidereal period.** This is how long the moon takes to circle the sky once and return to the same position among the stars.

A complete cycle of lunar phases takes 29.53 days, the moon's **synodic period.** (Synodic comes from the Greek words for "together" and "path.")

To see why the synodic period is longer than the sidereal period, study the star charts at the right.

Although you think of the lunar cycle as being about 4 weeks long, it is actually 1.53 days longer than 4 weeks. The calendar divides the year into 30-day periods called months (literally "moonths") in recognition of the 29.53 day synodic cycle of the moon.

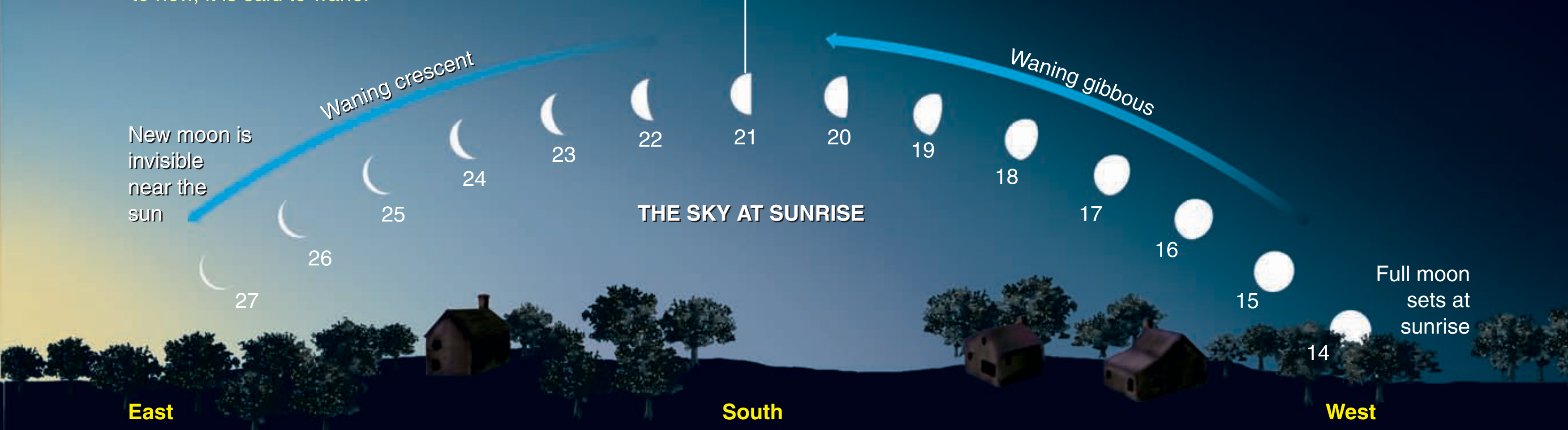


You can use the diagram on the opposite page to determine when the moon rises and sets at different phases.

TIMES OF MOONRISE AND MOONSET

Phase	Moonrise	Moonset
New	Dawn	Sunset
First quarter	Noon	Midnight
Full	Sunset	Dawn
Third quarter	Midnight	Noon

2b The last two weeks of the cycle of the moon are shown below by its position at sunrise on 14 successive mornings. As the moon shrinks from full to new, it is said to wane.



your location. ■ Table 3-1 will allow you to determine which upcoming total and partial lunar eclipses will be visible from your location.

Solar Eclipses

From Earth you can see a phenomenon that is not visible from most planets. It happens that the sun is 400 times larger than our moon but, on the average, nearly 400 times farther away, so the sun and moon have nearly equal angular diameters of about 0.5°. (See ■ Reasoning with Numbers 3-1.) This means that the moon is just the right size to cover the bright disk of the sun and cause a **solar eclipse.** If the moon covers the entire disk of the sun, you see a total eclipse. If it covers only part of the sun, you see a partial eclipse.

Every new moon, the shadow of the moon points toward Earth, but it usually misses. When the moon's shadow does sweep over Earth, the umbra barely reaches Earth and produces a small spot of darkness. The penumbra produces a larger circle of dimmed sunlight (■ Figure 3-8). What you see of the resulting eclipse depends on where you are in those shadows. Standing in that umbral spot, you would be in total shadow, unable to see any part of the sun's bright surface, and the eclipse would be total. But if you were located outside the umbra, in the penumbra, you would see part of the sun peeking around the edge of the moon, and the eclipse would be partial. Of course, if you are outside the penumbra, you would see no eclipse at all.

Because of the orbital motion of the moon and the rotation of Earth, the moon's shadow sweeps rapidly across Earth in a

■ **Table 3-1 | Total and Partial Eclipses of the Moon, 2009–2017***

Date	Time** of Mideclipse (GMT)	Length of Totality (Min)	Length of Eclipse (Hr:Min)
2009 Dec. 31	19:24	Partial	1:00
2010 June 26	11:40	Partial	2:42
2010 Dec. 21	8:18	72	3:28
2011 June 15	20:13	100	3:38
2011 Dec. 10	14:33	50	3:32
2012 June 4	11:04	Partial	2:06
2013 April 25	20:07	Partial	0:32
2014 April 15	7:46	78	3:38
2014 Oct. 8	10:55	60	3:20
2015 April 4	12:02	Partial	3:28
2015 Sept. 28	2:48	72	3:20
2017 Aug. 7	18:22	Partial	1:54

*There are no total or partial lunar eclipses during 2016.

**Times are Greenwich Mean Time. Subtract 5 hours for Eastern Standard Time, 6 hours for Central Standard Time, 7 hours for Mountain Standard Time, and 8 hours for Pacific Standard Time. For your time zone, lunar eclipses that occur between sunset and sunrise will be visible, and those at midnight will be best placed.

Reasoning with Numbers | 3-1

The Small-Angle Formula

■ Figure 3-7 shows the angular diameter of an object, its linear diameter, and its distance. Linear diameter is the distance between an object's opposite sides. The linear diameter of the moon, for instance, is 3476 km. Recall that the angular diameter of an object is the angle formed by two lines extending from opposite sides of the object and meeting at your eye. Clearly, the farther away an object is, the smaller its angular diameter. The small-angle formula allows you to find any of these three quantities if you know the other two.

In the small-angle formula, you always express angular diameter in seconds of arc,* and you always use the same units for distance and linear diameter:

$$\frac{\text{angular diameter}}{206{,}265} = \frac{\text{linear diameter}}{\text{distance}}$$

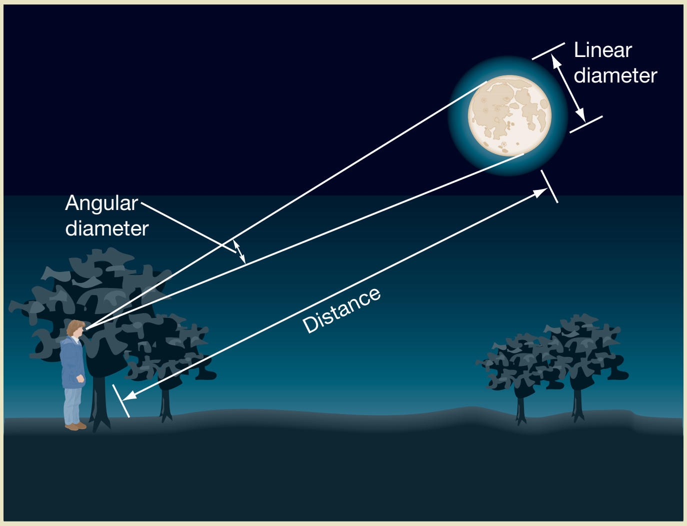
Example: The moon has a linear diameter of 3476 km and is about 384,000 km away. What is its angular diameter?

Solution: You can leave linear diameter and distance in kilometers and find the angular diameter in seconds of arc:

$$\frac{\text{angular diameter}}{206{,}265} = \frac{3476\text{ km}}{384{,}000\text{ km}}$$

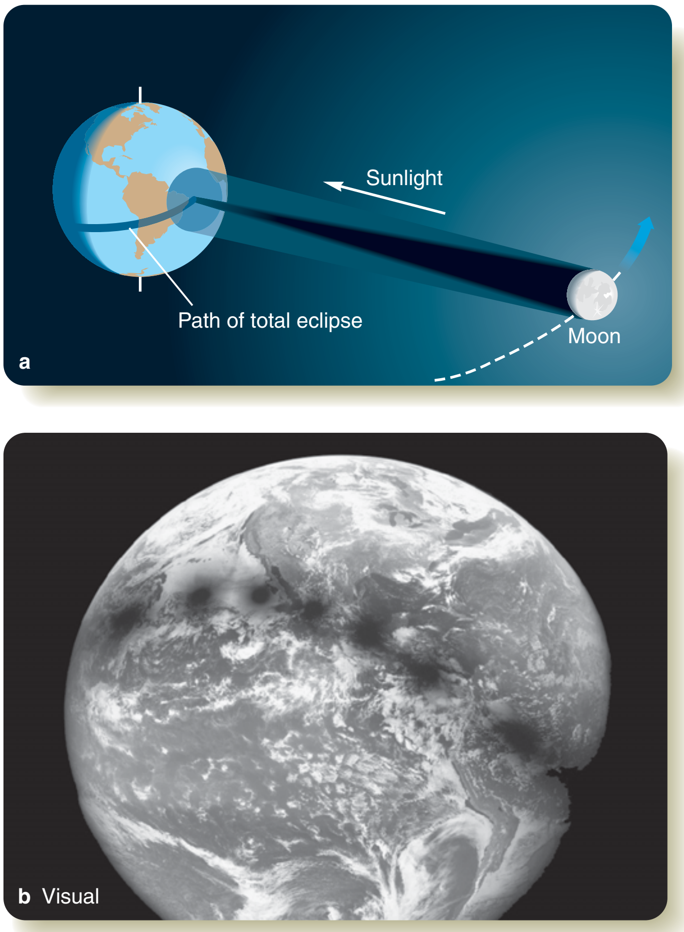
The angular diameter is 1870 seconds, which equals 31 minutes, of arc—about 0.5°.

*The number 206,265 is the number of seconds of arc in a radian. When you divide by 206,265, you convert the angle from seconds of arc into radians.



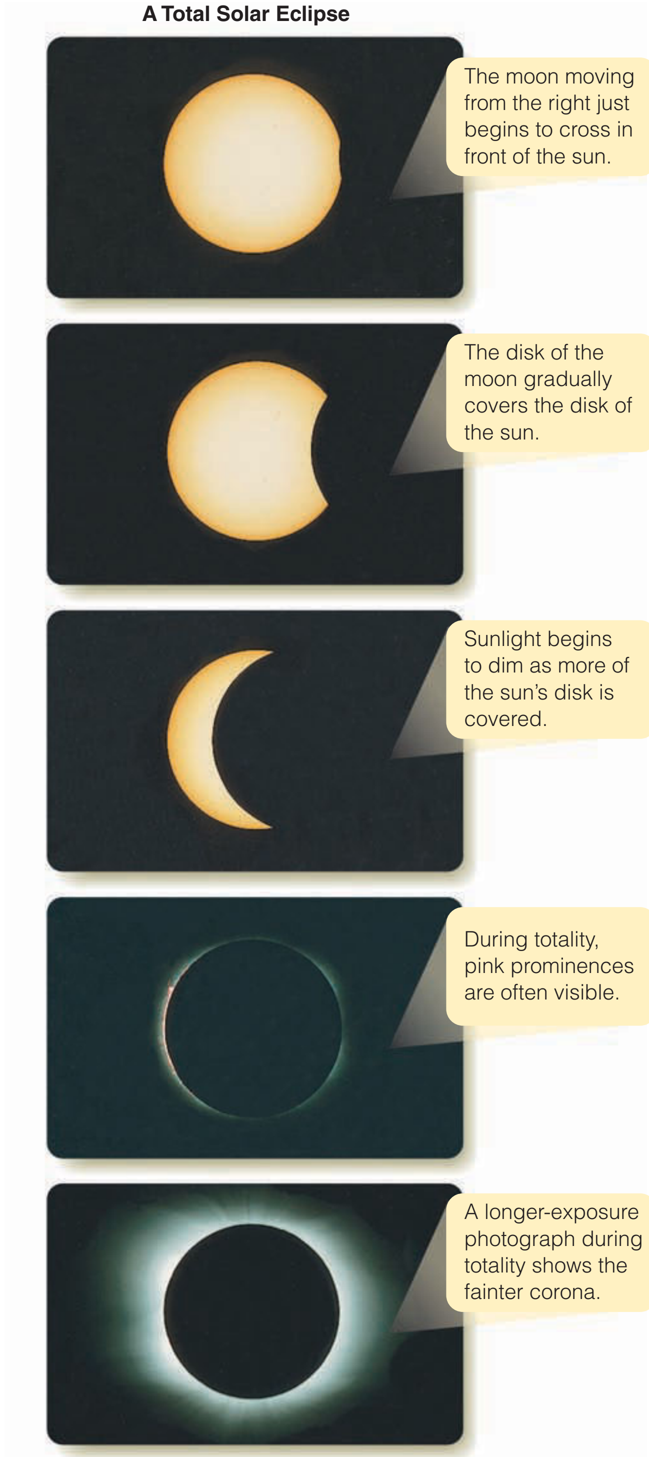
■ **Figure 3-7**

The three quantities related by the small-angle formula. Angular diameter is given in seconds of arc in the formula. Distance and linear diameter must be expressed in the same units—both in meters, both in light-years, and so on. **Animated!**



■ **Figure 3-8**

(a) The umbra of the moon's shadow sweeps from west to east across Earth, and observers in the path of totality see a total solar eclipse. Those outside the umbra but inside the penumbra see a partial eclipse. (b) Eight photos made by a weather satellite have been combined to show the moon's shadow moving across Mexico, Central America, and Brazil. (NASA GOES images courtesy of MrEclipse.com)



■ **Figure 3-9**

This sequence of photos shows the first half of a total solar eclipse. (Daniel Good)

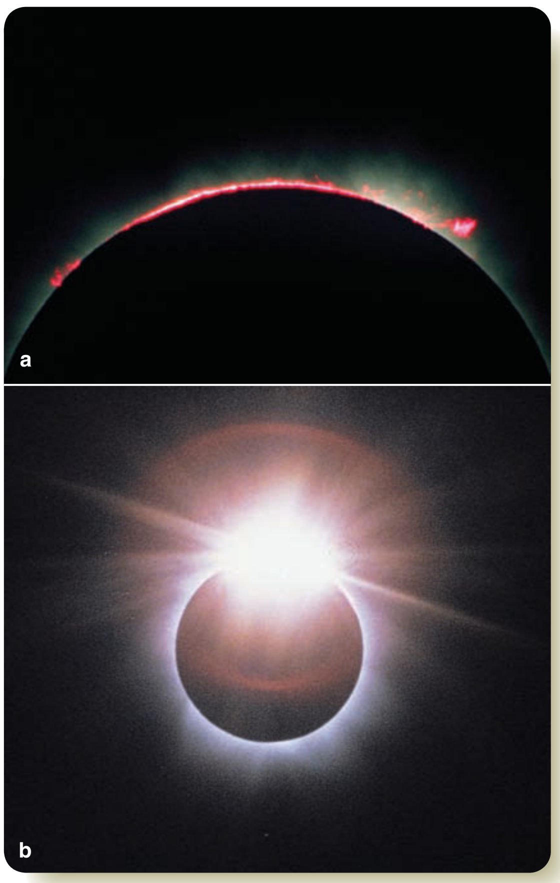
long, narrow path of totality. If you want to see a total solar eclipse, you must be in the path of totality. When the umbra of the moon's shadow sweeps over you, you see one of the most dramatic sights in the sky—a total eclipse of the sun.

The eclipse begins as the moon slowly crosses in front of the sun. It takes about an hour for the moon to cover the solar disk, but as the last sliver of sun disappears behind the moon, only the glow of the sun's outer atmosphere is visible (■ Figure 3-9) and darkness falls in a few seconds. Automatic streetlights come on, drivers of cars turn on their headlights, and birds go to roost. The sky becomes so dark you can even see the brighter stars.

The darkness lasts only a few minutes because the umbra is never more than 270 km (168 miles) in diameter and sweeps across Earth's surface at over 1600 km/hr (1000 mph). The sun cannot remain totally eclipsed for more than 7.5 minutes, and the average period of totality lasts only 2 or 3 minutes.

The brilliant surface of the sun is called the **photosphere,** and when the moon covers the photosphere, you can see the fainter **chromosphere,** the higher layers of the sun's atmosphere, glowing a bright pink. Above the chromosphere you see the **corona,** the sun's outer atmosphere. The corona is a low-density, hot gas that glows with a pale white color. Streamers caused by the solar magnetic field streak the corona, as may be seen in the last frame of Figure 3-9. The chromosphere is often marked by eruptions on the solar surface called **prominences** (■ Figure 3-10a). The corona, chromosphere, and prominences are visible only when the brilliant photosphere is covered. As soon as part of the photosphere reappears, the fainter corona, chromosphere, and prominences vanish in the glare, and totality is over. The moon moves on in its orbit, and in an hour the sun is completely visible again.

Just as totality begins or ends, a small part of the photosphere can peek out from behind the moon through a valley at the edge of the lunar disk. Although it is intensely bright, such a tiny bit of the photosphere does not completely drown out the fainter corona, which forms a silvery ring of light with the brilliant spot of photosphere gleaming like a diamond (Figure 3-10b). This **diamond-ring effect** is one of the most spectacular of astronomical sights, but it is not visible during every solar eclipse. Its occurrence depends on the exact orientation and motion of the moon.



■ **Figure 3-10**

(a) During a total solar eclipse, the moon covers the photosphere, and the ruby-red chromosphere and prominences are visible. Only the lower corona is visible in this image. (©2005 Fred Espenak, www.MrEclipse.com) (b) The diamond ring effect can sometimes occur momentarily at the beginning or end of totality if a small segment of the photosphere peeks out through a valley at the edge of the lunar disk. (National Optical Astronomy Observatory)

The moon's angular diameter changes depending on where it is around its slightly elliptical orbit. When it is near **perigee,** its point of closest approach to Earth, it looks a little bit larger than when it is near **apogee,** the most distant point in its orbit. Furthermore, Earth's orbit is also slightly elliptical, so the Earth–sun distance varies, and that changes the angular diameter of the solar disk by a few percent (■ Figure 3-11). If the moon is in the farther part of its orbit during totality, its angular diameter will be less than the angular diameter of the sun, and when that happens, you see an **annular eclipse,** a solar eclipse in which a ring (or annulus) of the photosphere is visible around the disk of the moon. Because a portion of the brilliant photosphere remains visible, it never quite gets dark, and you can't see the prominences, chromosphere, and corona (Figure 3-11).

A list of future total and annular eclipses of the sun is given in ■ Table 3-2. If you plan to observe a solar eclipse, remember that the sun is bright enough to burn your eyes and cause permanent damage if you look at it directly. It is a **Common Misconception** that sunlight during an eclipse is somehow extra dangerous. Sunlight is bright enough to burn your eyes any day, whether there is an eclipse or not. Only during totality, while the brilliant photosphere is entirely hidden, is it safe to look directly at the eclipse. See ■ Figure 3-12 for a safe way to observe the partially eclipsed sun.

Predicting Eclipses

Predicting lunar or solar eclipses is quite complex, and if you wanted to make precise predictions, you would have to do some sophisticated calculations. But you can make general eclipse predictions by thinking about the geometry of an eclipse and the cyclic motions of the sun and moon.

Solar eclipses occur when the moon passes between Earth and the sun, that is, when the lunar phase is new moon. Lunar eclipses occur at full moon. However, you don't see eclipses at every new moon or full moon. Why not? That's the key question. The answer is that the moon's orbit is tipped a few degrees to the plane of Earth's orbit, so at most new or full moons, the shadows miss as you can see in the lower part of ■ Figure 3-13. If the shadows miss, there are no eclipses.

For an eclipse to occur, the moon must be passing through the plane of Earth's orbit. The points where it passes through the plane of Earth's orbit are called the **nodes** of the moon's orbit, and the line connecting these is called the line of nodes. In other words, the planes of the two orbits intersect along the line of nodes. The moon crosses its nodes every month, but eclipses can occur only if the moon is also new or full. That can happen twice a year when the line of nodes points toward the sun, and for a few weeks eclipses are possible at new moons and full moons (Figure 3-13). These intervals when eclipses are possible are called eclipse seasons, and they occur about six months apart.

If the moon's orbit were fixed in space, the eclipse seasons would always occur at the same times each year. The moon's orbit

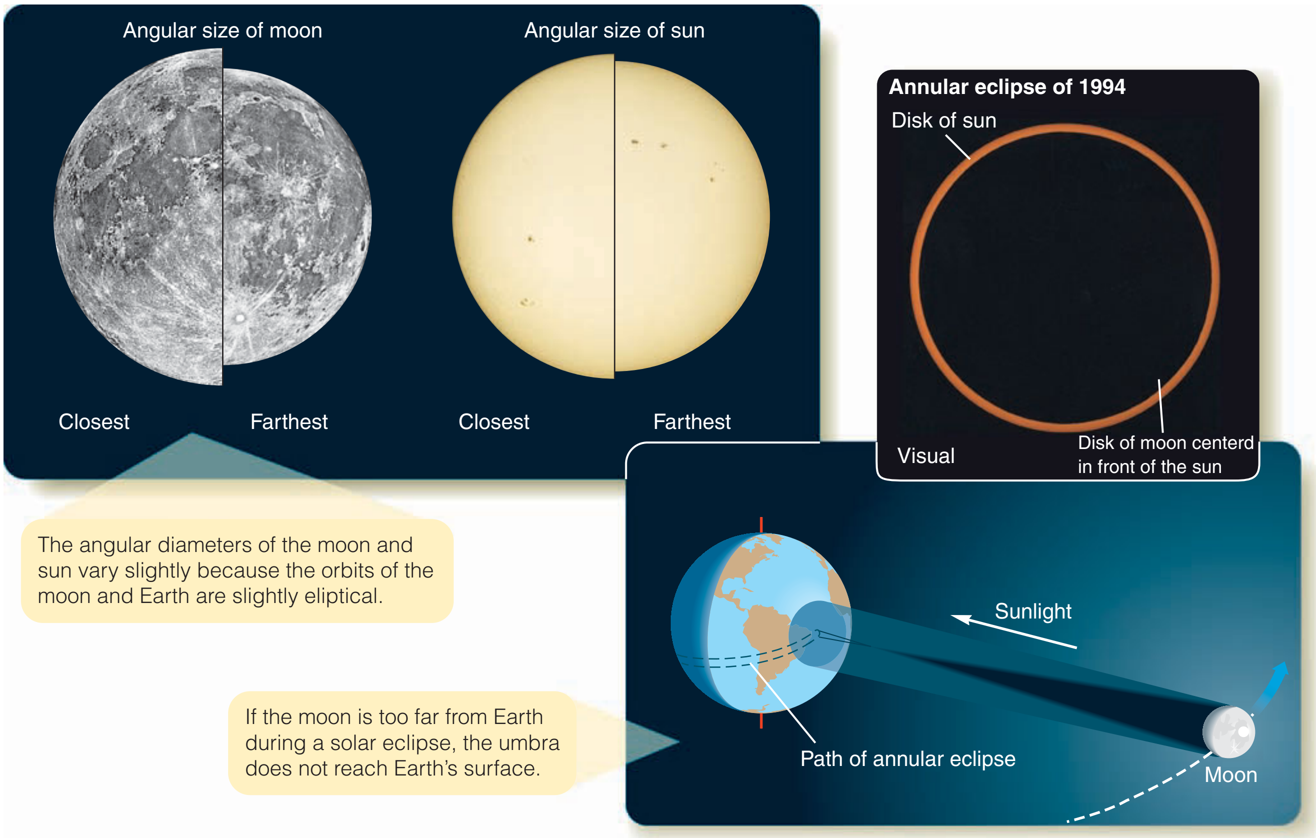


Figure 3-11

An annular eclipse occurs when the moon is far enough from Earth that its umbral shadow does not reach Earth's surface. From Earth, you see an annular eclipse because the moon's angular diameter is smaller than the angular diameter of the sun. In the photograph of the annular eclipse of 1994, the dark disk of the moon is almost exactly centered on the bright disk of the sun. (Daniel Good)

Table 3-2 | Total and Annular Eclipses of the Sun, 2009–2019*

Date	Total/Annular (T/A)	Time of Mideclipse‡ (GMT)	Maximum Length of Total or Annular Phase (Min:Sec)	Area of Visibility
2009 Jan. 26	A	8^h	7:56	S. Atlantic, Indian Oc
2009 July 22	T	3^h	6:40	Asia, Pacific
2010 Jan. 15	A	7^h	11:10	Africa, Indian Ocean
2010 July 11	T	20^h	5:20	Pacific, S. America
2012 May 20	A	23^h	5:46	Japan, N. Pacific, W. US
2012 Nov. 13	T	22^h	4:02	Australia, S. Pacific
2013 May 10	A	0^h	6:04	Australia, Pacific
2013 Nov. 3	AT	13^h	1:40	Atlantic, Africa
2015 March 20	T	10^h	2:47	N. Atlantic, Arctic
2016 March 9	T	2^h	4:10	Borneo, Pacific
2016 Sept. 1	A	9^h	3:06	Atlantic, Africa, Indian Oc
2017 Feb 26	A	15^h	1:22	S. Pacific to Africa
2017 Aug 21	T	18^h	2:40	United States
2019 July 2	T	19^h	4:32	Pacific, S. America
2019 Dec. 26	A	5^h	3:40	S. E. Asia, Pacific

The next major total solar eclipse visible from the United States will occur on August 21, 2017, when the path of totality will cross the United States from Oregon to South Carolina.

*There are no total or partial solar eclipses in 2011, 2014, or 2018.

‡Times are Greenwich Mean Time. Subtract 5 hours for Eastern Standard Time, 6 hours for Central Standard Time, 7 hours for Mountain Standard Time, and 8 hours for Pacific Standard Time.

^{h}hours.

Figure 3-12

A safe way to view the partial phases of a solar eclipse. Use a pinhole in a card to project an image of the sun on a second card. The greater the distance between the cards, the larger (and fainter) the image will be.

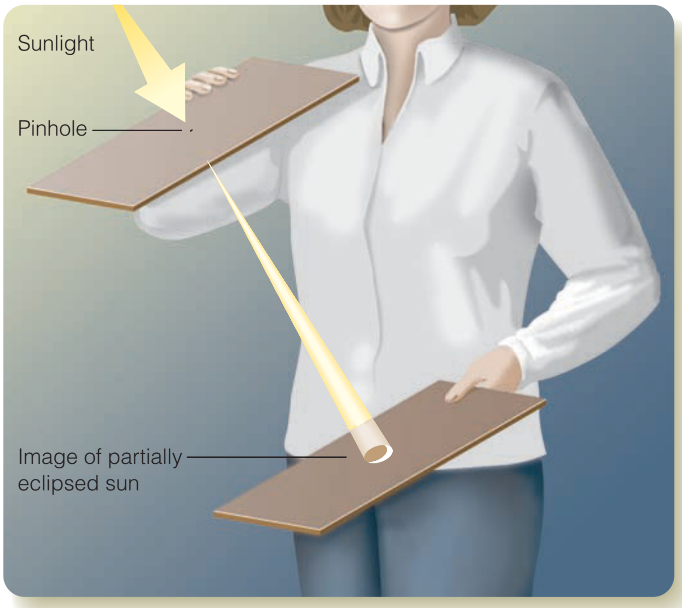
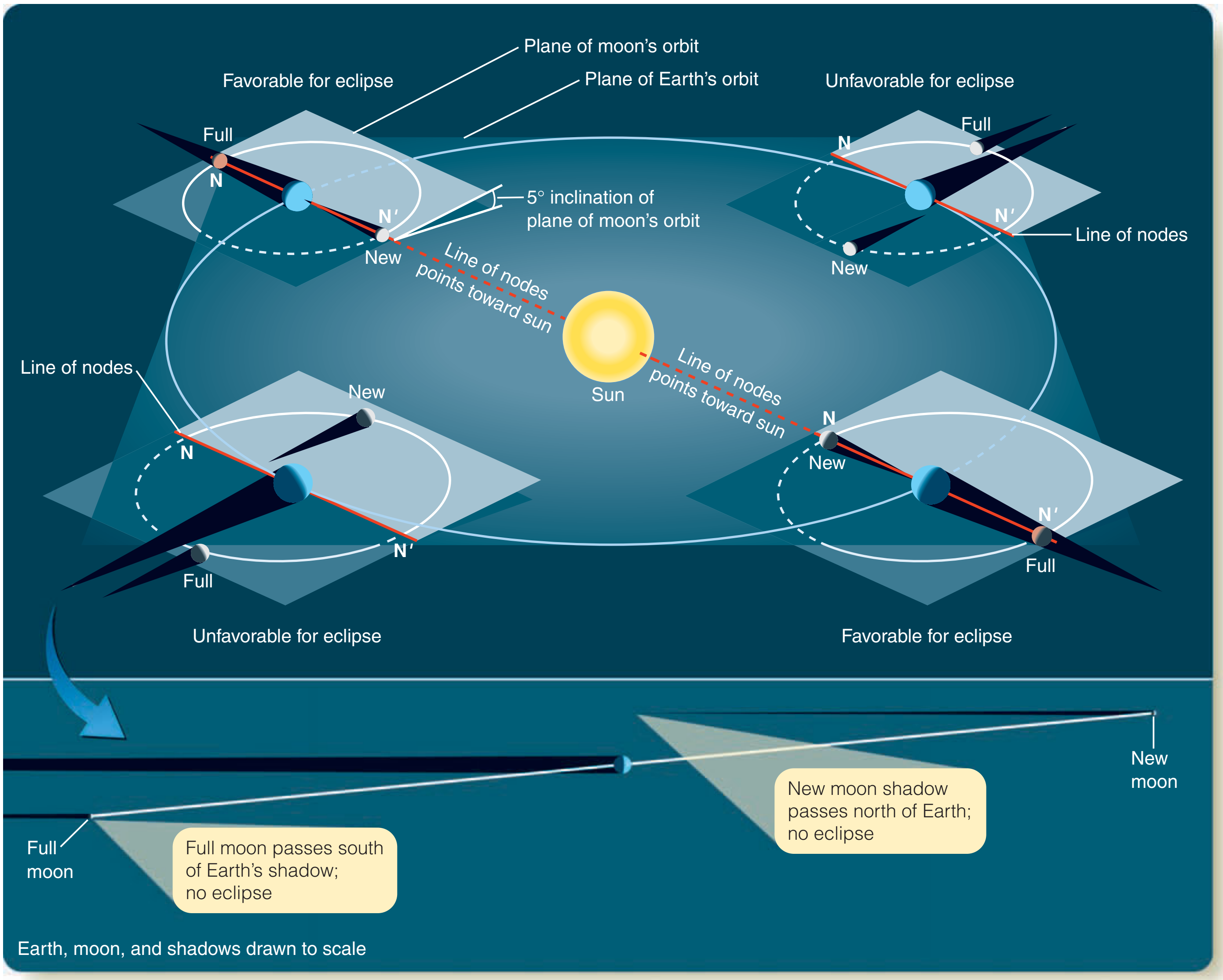


Figure 3-13

The moon's orbit is tipped about 5° to Earth's orbit. The nodes N and N′ are the points where the moon passes through the plane of Earth's orbit. If the line of nodes does not point at the sun, the long narrow shadows miss, and there are no eclipses at new moon and full moon. At those parts of Earth's orbit where the line of nodes points toward the sun, eclipses are possible at new moon and full moon.



precesses, however, because of the gravitational pull of the sun on the moon, and the precession slowly changes the direction of the line of nodes. The line turns gradually westward, making one complete rotation in 18.61 years. As a result, the eclipse seasons occur about three weeks earlier each year. Many ancient peoples noticed this pattern and could guess which full and new moons were likely to produce eclipses.

Another way the ancients predicted eclipses was to notice that the pattern of eclipses repeats every 6585.3 days — the **Saros cycle.** After one Saros, the sun, moon, and nodes have circled the sky many times and finally returned to the same arrangement they occupied when the Saros began. Then the cycle of eclipses begins to repeat. One Saros equals 18 years $11^1/_3$ days. Because of the extra third of a day, an eclipse visible in North America will recur after one Saros, but it will be visible one-third of the way around the world in the North Pacific. Once ancient astronomers recognized the Saros cycle, they could predict eclipses from records of previous eclipses.

◄ SCIENTIFIC ARGUMENT ►

What would astronauts on the moon observe while people on Earth were seeing a total lunar eclipse?

This scientific argument requires that you change your point of view and imagine seeing an event from a new location. Remember that when you see a total lunar eclipse, the full moon is passing through Earth's shadow. Astronauts standing on the moon would look up and see Earth crossing in front of the brilliant sun. The lunar day would begin to grow dim as the moon entered Earth's penumbra. The visible part of the sun would grow narrower and narrower until it vanished entirely behind Earth, and the astronauts would be left standing in the dark as the moon carried them through the umbra of Earth's shadow. Except for faint starlight, their only light would come from the glow of Earth's atmosphere lit from behind, a red ring around the dark disk of Earth made up of every sunset and sunrise. The red light from Earth's atmosphere would bathe the dusty plains and mountains of the moon in a copper-red glow. The astronauts would have a cold and tedious wait for the sun to reemerge from behind Earth, but they would see a lunar eclipse from a new and dramatic vantage point.

Imagining the same event from different points of view can help you sort out complex geometries. Now change your argument slightly and imagine the eclipse once again. **If Earth had no atmosphere, how would this eclipse look different as viewed from Earth and from the moon?**

◄ ►

What Are We? Scorekeepers

The rotation and revolution of Earth produce the cycles of day and night and winter and summer, and we have evolved to live within those cycles. One theory holds that we sleep at night because dozing in the back of a cave (or in a comfortable bed) is safer than wandering around in the dark. The night is filled with predators, so sleeping may keep us safe. Our bodies depend on that cycle of light and dark: People who live and work in the Arctic or Antarctic where the cycle of day and night does not occur can suffer psychological problems from the lack of the daily cycle.

The cycle of the seasons controls the migration of game and the growth of crops, so cultures throughout history have followed the motions of the sun along the ecliptic with special reverence. The people who built Stonehenge were marking the summer solstice sunrise because it was a moment of power, order, and promise in the cycle of their lives.

The moon's cycles mark the passing days and divide our lives into weeks and months. In a Native American story, Coyote gambles with the sun to see if the sun will continue to warm Earth, and the moon keeps score. The moon is a symbol of regularity, reliability, and dependability. It is the scorekeeper counting out your days and months.

Like the ticking of a cosmic clock, the passing weeks, months, and seasons mark the passage of time on Earth, but, as you have seen, the cycle of the seasons is also affected by longer period changes in the motion of Earth. Ice ages come and go, and Earth's climate cycles in ways we do not entirely understand. If you don't feel quite as secure as you did when you started this chapter, then you are catching on. Astronomy tells us that Earth is a beautiful world, but it is also a complicated, spinning planet. Our clocks, calendars, and lives count the passing cycles in the sky.

Study and Review

Summary

- The **rotation (p. 22)** of Earth on its axis produces the cycle of day and night, and the **revolution (p. 22)** of Earth around the sun produces the cycle of the year.
- Because Earth orbits the sun, the sun appears to move eastward along the **ecliptic (p. 22)** through the constellations completing a circuit of the sky in a year.
- Because the ecliptic is tipped 23.5° to the celestial equator, the sun spends half the year in the northern celestial hemisphere and half in the southern celestial hemisphere.
- In the summer, the sun is above the horizon longer and shines more directly down on the ground. Both effects cause warmer weather in the northern hemisphere. In the winter, the sun is in the southern sky, and Earth's northern hemisphere has colder weather.
- The seasons are reversed in Earth's southern hemisphere relative to the northern hemisphere.
- The beginning of spring, summer, winter, and fall are marked by the **vernal equinox (p. 24)**, the **summer solstice (p. 24)**, the **autumnal equinox (p. 24)**, and the **winter solstice (p. 24)**.
- Earth is slightly closer to the sun at **perihelion (p. 25)** in January and slightly farther away from the sun at **aphelion (p. 25)** in July. This has almost no effect on the seasons.
- The planets move generally eastward along the ecliptic, and all but Uranus and Neptune are visible to the unaided eye looking like stars. Mercury and Venus never wander far from the sun and are sometimes visible in the evening sky after sunset or in the dawn sky before sunrise.
- Planets visible in the sky at sunset are traditionally called **evening stars (p. 26)**, and planets visible in the dawn sky are called **morning stars (p. 26)**.
- The locations of the sun and planets along the **zodiac (p. 26)** are the bases for the ancient **pseudoscience (p. 26)** known as astrology.
- According to the **Milankovitch hypothesis (p. 27)**, changes in the shape of Earth's orbit, in its precession, and in its axial tilt can alter the planet's heat balance and cause the cycle of ice ages. Evidence found in sea floor samples support the hypothesis and it is widely accepted today.
- Scientists routinely test their own ideas by organizing theory and evidence into a **scientific argument (p. 28)**.
- The moon orbits eastward around Earth once a month and rotates on its axis, keeping the same side facing Earth throughout the month.
- Because you see the moon by reflected sunlight, its shape appears to change as it orbits Earth and sunlight illuminates different amounts of the side you can see.
- The lunar phases wax from new moon to first quarter to full moon and wane from full moon to third quarter to new moon.
- A complete cycle of lunar phases takes 29.53 days, which is known as the moon's **synodic period (p. 33)**. The **sidereal period (p. 33)** of the moon—its orbital period with respect to the stars—is a bit over 2 days shorter.
- If a full moon passes through Earth's shadow, sunlight is cut off, and the moon darkens in a **lunar eclipse (p. 30)**. If the moon fully enters the dark **umbra (p. 30)** of Earth's shadow, the eclipse is total; but if it only grazes the umbra, the eclipse is partial. If the moon enters the partial shadow of the **penumbra (p. 30)** but not the umbra, the eclipse is penumbral.
- During **totality (p. 31)**, the eclipsed moon looks copper-red because of sunlight refracted through Earth's atmosphere.
- A **solar eclipse (p. 34)** occurs if a new moon passes between the sun and Earth and the moon's shadow sweeps over Earth's surface. Observers inside the path of totality see a total eclipse, and those just outside the path of totality see a partial eclipse as the penumbra sweeps over their location.
- During a total eclipse, the bright **photosphere (p. 35)** of the sun is covered, and the fainter **corona (p. 35)**, **chromosphere (p. 35)**, and **prominences (p. 36)** become visible.
- Sometimes just as totality begins or ends, the bright photosphere peeks out through a valley at the edge of the lunar disk and produces the **diamond-ring effect (p. 36)**.
- When the moon is near **perigee (p. 36)**, the closest point in its orbit, its angular diameter is large enough to cover the sun's photosphere and produce a total eclipse. But if the moon is near **apogee (p. 36)**, the farthest point in its orbit, it looks too small and can't entirely cover the photosphere. A solar eclipse occurring then would be an **annular eclipse (p. 36)**.
- Because the moon's orbit is tipped a few degrees from the plane of Earth's orbit, most full moons pass north or south of Earth's shadow, and no lunar eclipse occurs. Also, most new moons cross north or south of the sun, and there is no solar eclipse.
- Eclipses can only occur when a full moon or a new moon occurs near one of the two **nodes (p. 36)** of its orbit, where it crosses the ecliptic. These two eclipse seasons occur about 6 months apart, but move slightly earlier each year. By keeping track of the location of the nodes of the moon's orbit, you could predict which full and new moons were most likely to be eclipsed.
- Eclipses follow a pattern lasting 18 years $11^1/_3$ days called the **Saros cycle (p. 39)**. If ancient astronomers understood that pattern, they could predict eclipses.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What is the difference between the daily and annual motions of the sun?
2. If Earth did not rotate, could you still define the ecliptic? Why or why not?
3. What would the seasons be like if Earth were tipped 35° instead of 23.5°? What would they be like if Earth's axis were perpendicular to its orbit?
4. Why are the seasons reversed in the southern hemisphere relative to the northern hemisphere?
5. How could small changes in the inclination of Earth's axis affect world climate?
6. Do the phases of the moon look the same from every place on Earth, or is the moon full at different times as seen from different locations?
7. What phase would Earth be in if you were on the moon when the moon was full? At first quarter? At waning crescent?
8. Why have most people seen a total lunar eclipse, while few have seen a total solar eclipse?
9. Why isn't there an eclipse at every new moon and at every full moon?
10. Why is the moon red during a total lunar eclipse?
11. Why should the eccentricity of Earth's orbit make winter in the northern hemisphere different from winter in the southern hemisphere?
12. **How Do We Know?** What are the main characteristics of a pseudoscience? Can you suggest other examples?

13. **How Do We Know?** Why would it be appropriate to refer to evidence as the reality checks in science?
14. **How Do We Know?** Why must a scientific argument dealing with some aspect of nature include all of the evidence?

Discussion Questions

1. Do planets orbiting other stars have ecliptics? Could they have seasons?
2. Why would it be difficult to see prominences if you were on the moon during a total lunar eclipse?

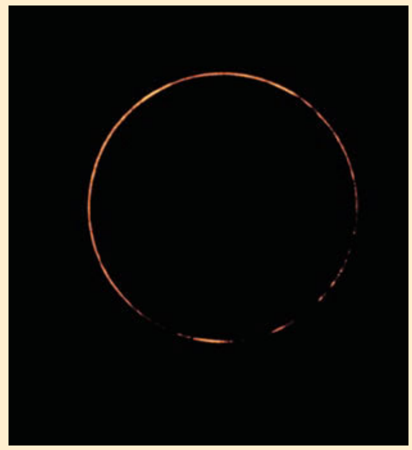
Problems

1. If Earth is about 4.6 billion (4.6×10^9) years old, how many precessional cycles have occurred?
2. Identify the phases of the moon if on March 20 the moon were located at (a) the vernal equinox, (b) the autumnal equinox, (c) the summer solstice, (d) the winter solstice.
3. Identify the phases of the moon if at sunset the moon were (a) near the eastern horizon, (b) high in the south, (c) in the southeast, (d) in the southwest.
4. About how many days must elapse between first-quarter moon and third-quarter moon?
5. Draw a diagram showing Earth, the moon, and shadows during (a) a total solar eclipse, (b) a total lunar eclipse, (c) a partial lunar eclipse, (d) an annular eclipse.
6. Phobos, one of the moons of Mars, is 20 km in diameter and orbits 5982 km above the surface of the planet. What is the angular diameter of Phobos as seen from Mars? (*Hint:* See Reasoning with Numbers 3-1.)
7. A total eclipse of the sun was visible from Canada on July 10, 1972. When did this eclipse occur next? From what part of Earth was it total?
8. When will the eclipse described in Problem 7 next be total as seen from Canada?

Learning to Look

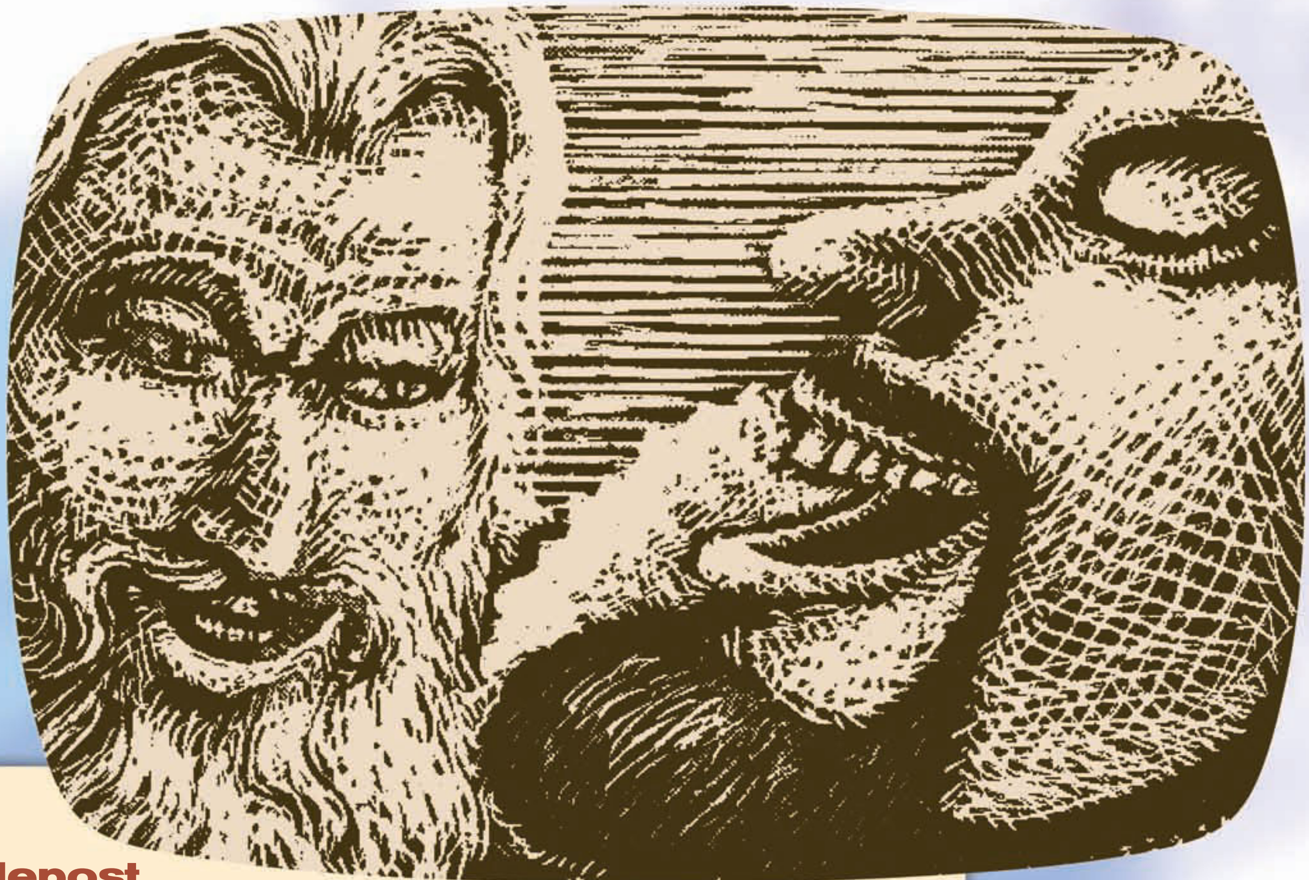
1. Look at the chapter opening photo for Chapter 2 and notice the glow on the horizon at lower right. Is that sunset glow or sunrise glow? About what time of day or night was this photo taken? About what season of the year was it taken? You may want to consult the star charts at the back of this book.
2. The stamp at right shows a crescent moon. Explain why the moon could never look this way.

3. The photo at right shows the annular eclipse of May 30, 1984. How is it different from the annular eclipse shown in Figure 3-11? Why do you suppose it is different?



Laurence Marschall

4 The Origin of Modern Astronomy



Astronomers like Galileo Galilei and Johannes Kepler struggled against 2000 years of tradition as they tried to understand the place of Earth and the motion of the planets.

Guidepost

The preceding chapters gave you a modern view of Earth. You can now imagine how Earth, the moon, and the sun move through space and how that produces the sights you see in the sky. But how did humanity first realize that we live on a planet moving through space? That required the revolutionary overthrow of an ancient and honored theory of Earth's place.

By the 16th century, many astronomers were uncomfortable with the theory that Earth sat at the center of a spherical universe. In this chapter, you will discover how an astronomer named Copernicus changed the old theory, how Galileo Galilei changed the rules of debate, and how Isaac Newton changed humanity's concept of nature. Here you will find answers to four essential questions:

How did classical philosophers describe Earth's place in the universe?

How did Copernicus revise that ancient theory?

Why was Galileo condemned by the Inquisition?

How did Isaac Newton change humanity's view of nature?

This chapter is not just about the history of astronomy. As they struggled to understand Earth and the heavens, the astronomers of the Renaissance invented a new way of understanding nature—a way of thinking that is now called science. Every chapter that follows will use the methods that were invented when Copernicus tried to repair that ancient theory that Earth was the center of the universe.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

How you would burst out laughing, my dear Kepler, if you would hear what the greatest philosopher of the Gymnasium told the Grand Duke about me . . .

FROM A LETTER BY GALILEO GALILEI

NEXT TIME YOU look at the sky, imagine how prehistoric families felt as they huddled around the safety of their fires and looked up at the stars. Astronomy had its beginnings in simple human curiosity about the lights in the sky. As early civilizations developed, great philosophers struggled to understand the movements of the sun, moon, and planets. Later, mathematical astronomers made precise measurements and computed detailed models in their attempts to describe celestial motions. It took hard work and years of effort, but the passions of astronomy gripped some of the greatest minds in history and drove them to try to understand the sky. As you study the history of astronomy, notice that two themes twist through the story.

One theme is the struggle to understand the place of Earth in the universe. It seemed obvious to the ancients that Earth was the center of everything, but today you know that's not true. The debate over the place of Earth involved deep theological questions and eventually led Galileo before the Inquisition.

The second theme is the long and difficult quest to understand planetary motion. Astronomers built more and more elaborate mathematical models, but they still could not predict precisely the motion of the visible planets along the ecliptic. That mystery was finally solved when Isaac Newton described gravity and orbital motion in the late 1600s.

Only a few centuries ago, as astronomers were struggling to understand the sky, they invented a new way of understanding nature—a new way of knowing about the physical world. That new way of knowing is based on the comparison of theories and evidence. Today, that new way of knowing is called science.

4-1 Classical Astronomy

THE GREAT PHILOSOPHERS of ancient Greece wrote about many different subjects, including what they saw in the sky. Those writings became the foundation on which later astronomers built modern astronomy.

The Aristotelian Universe

You have probably heard of the two greatest philosophers of ancient Greece—Plato and Aristotle. Their writings shaped the history of astronomy. Plato (427?–347 BC) wrote about moral responsibility, ethics, the nature of reality, and the ideals of civil government. His student Aristotle (384–322 BC) wrote on almost every area of knowledge and is probably the most famous philosopher in history. These two philosophers established the first widely accepted ideas about the structure of the universe.

Science and its methods of investigation did not exist in ancient Greece, so when Plato and Aristotle turned their minds to the problem of the structure of the universe, they made use of a process common to their times—reasoning from first principles. A first principle is something that is held to be obviously true. Once a principle is recognized as true, whatever can be logically derived from it must also be true.

But what was obviously true to the ancients is not so obvious to us today. Study ■ **The Ancient Universe** on pages 44–45 and notice three important ideas and seven new terms that show how first principles influenced early descriptions of the universe and its motions:

1. Ancient philosophers and astronomers accepted as first principles that the universe was *geocentric* with Earth located at the center and that the heavens moved in *uniform circular motion*. They thought it was obvious that Earth did not move because they did not see the shifting of the stars called *parallax*.
2. Notice how the observed motion of the planets, the evidence, did not fit the theory very well. The *retrograde motion* of the planets was very difficult to explain using geocentrism and uniform circular motion.
3. Finally, notice how Claudius Ptolemy attempted to explain the motion of the planets mathematically by devising a small circle, the *epicycle,* rotating along the edge of a larger circle, the *deferent,* that enclosed Earth. He even allowed the speed of the planets to vary slightly as they circled a slightly off-center point called the *equant*. In these ways he weakened the principles of geocentrism and uniform circular motion.

Ptolemy lived roughly five centuries after Aristotle in the Greek colony in Egypt, and although Ptolemy believed in the Aristotelian universe, he was interested in a different problem—the motion of the planets. He was a brilliant mathematician, and he used his talents to create a mathematical description of the motions he saw in the heavens. For him, first principles took second place to mathematical precision.

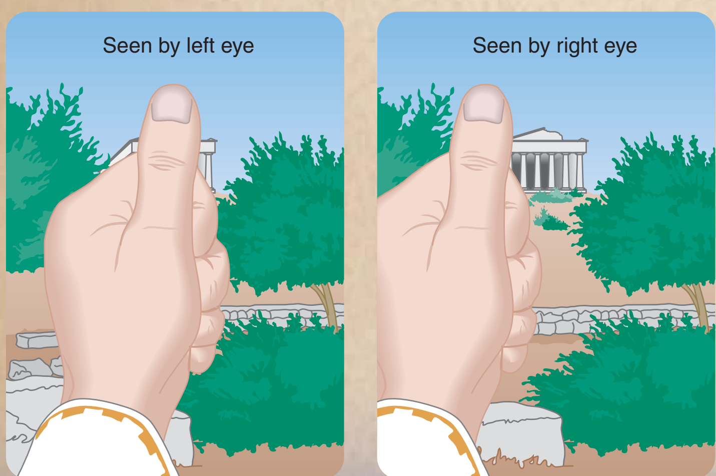
Aristotle's universe, as embodied in Ptolemy's mathematical model, dominated ancient astronomy, but it was wrong. The universe is not geocentric, and the planets don't follow circles at uniform speeds. At first the Ptolemaic system predicted the positions of the planets well; but, as centuries passed, errors accumulated. Astronomers tried to update the system, computing new constants and adjusting epicycles. In the middle of the 13th century, a team of astronomers supported by King Alfonso X of Castile studied the *Almagest* for 10 years. Although they did not revise the theory very much, they simplified the calculation of the positions of the planets using the Ptolemaic system and published the result as *The Alfonsine Tables,* the last great attempt to make the Ptolemaic system of practical use.

1 For 2000 years, the minds of astronomers were shackled by a pair of ideas. The Greek philosopher Plato argued that the heavens were perfect. Because the only perfect geometrical shape is a sphere, which carries a point on its surface around in a circle, and because the only perfect motion is uniform motion, Plato concluded that all motion in the heavens must be made up of combinations of circles turning at uniform rates. This idea was called **uniform circular motion.**

Plato's student Aristotle argued that Earth was imperfect and lay at the center of the universe. Such a model is known as a **geocentric universe.** His model contained 55 spheres turning at different rates and at different angles to carry the seven known planets (the moon, Mercury, Venus, the sun, Mars, Jupiter, and Saturn) across the sky.

Aristotle was known as the greatest philosopher in the ancient world, and for 2000 years his authority chained the minds of astronomers with uniform circular motion and geocentrism. See the model at right.

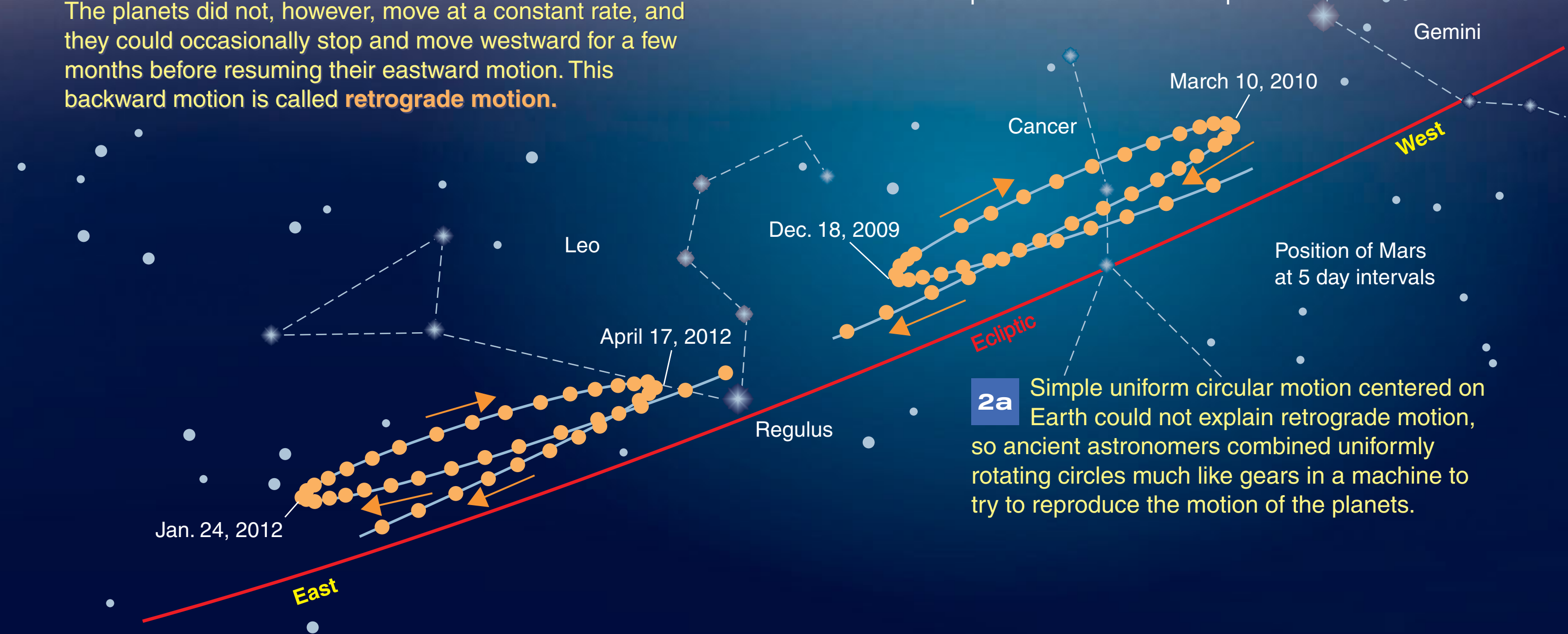
From *Cosmographica* by Peter Apian (1539).



1a Ancient astronomers believed that Earth did not move because they saw no **parallax,** the apparent motion of an object because of the motion of the observer. To demonstrate parallax, close one eye and cover a distant object with your thumb held at arm's length. Switch eyes, and your thumb appears to shift position as shown at left. If Earth moves, ancient astronomers reasoned, you should see the sky from different locations at different times of the year, and you should see parallax distorting the shapes of the constellations. They saw no parallax, so they concluded Earth could not move. Actually, the parallax of the stars is too small to see with the unaided eye.

2 Planetary motion was a big problem for ancient astronomers. In fact, the word *planet* comes from the Greek word for "wanderer," referring to the eastward motion of the planets against the background of the fixed stars. The planets did not, however, move at a constant rate, and they could occasionally stop and move westward for a few months before resuming their eastward motion. This backward motion is called **retrograde motion.**

Every 2.14 years, Mars passes through a retrograde loop. Two successive loops are shown here. Each loop occurs further east along the ecliptic and has its own shape.

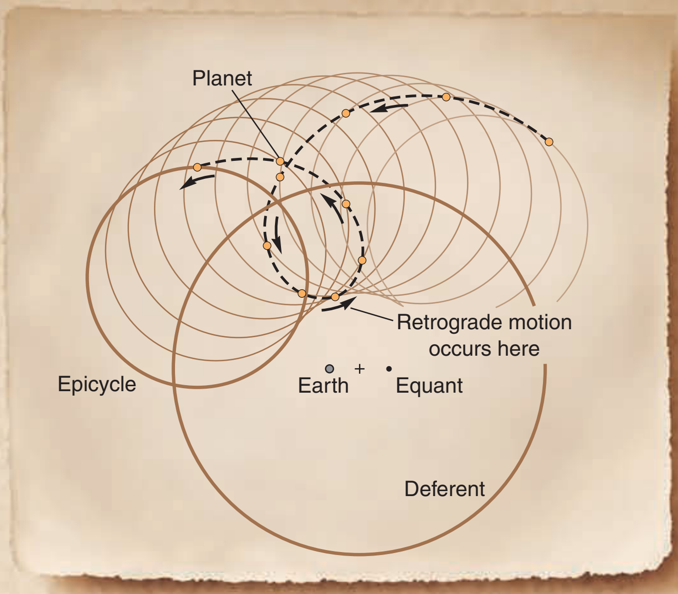


2a Simple uniform circular motion centered on Earth could not explain retrograde motion, so ancient astronomers combined uniformly rotating circles much like gears in a machine to try to reproduce the motion of the planets.

3 Uniformly rotating circles were key elements of ancient astronomy. Claudius Ptolemy created a mathematical model of the Aristotelian universe in which the planet followed a small circle called the **epicycle** that slid around a larger circle called the **deferent.** By adjusting the size and rate of rotation of the circles, he could approximate the retrograde motion of a planet. See illustration at right.

To adjust the speed of the planet, Ptolemy supposed that Earth was slightly off center and that the center of the epicycle moved such that it appeared to move at a constant rate as seen from the point called the **equant.**

To further adjust his model, Ptolemy added small epicycles (not shown here) riding on top of larger epicycles, producing a highly complex model.



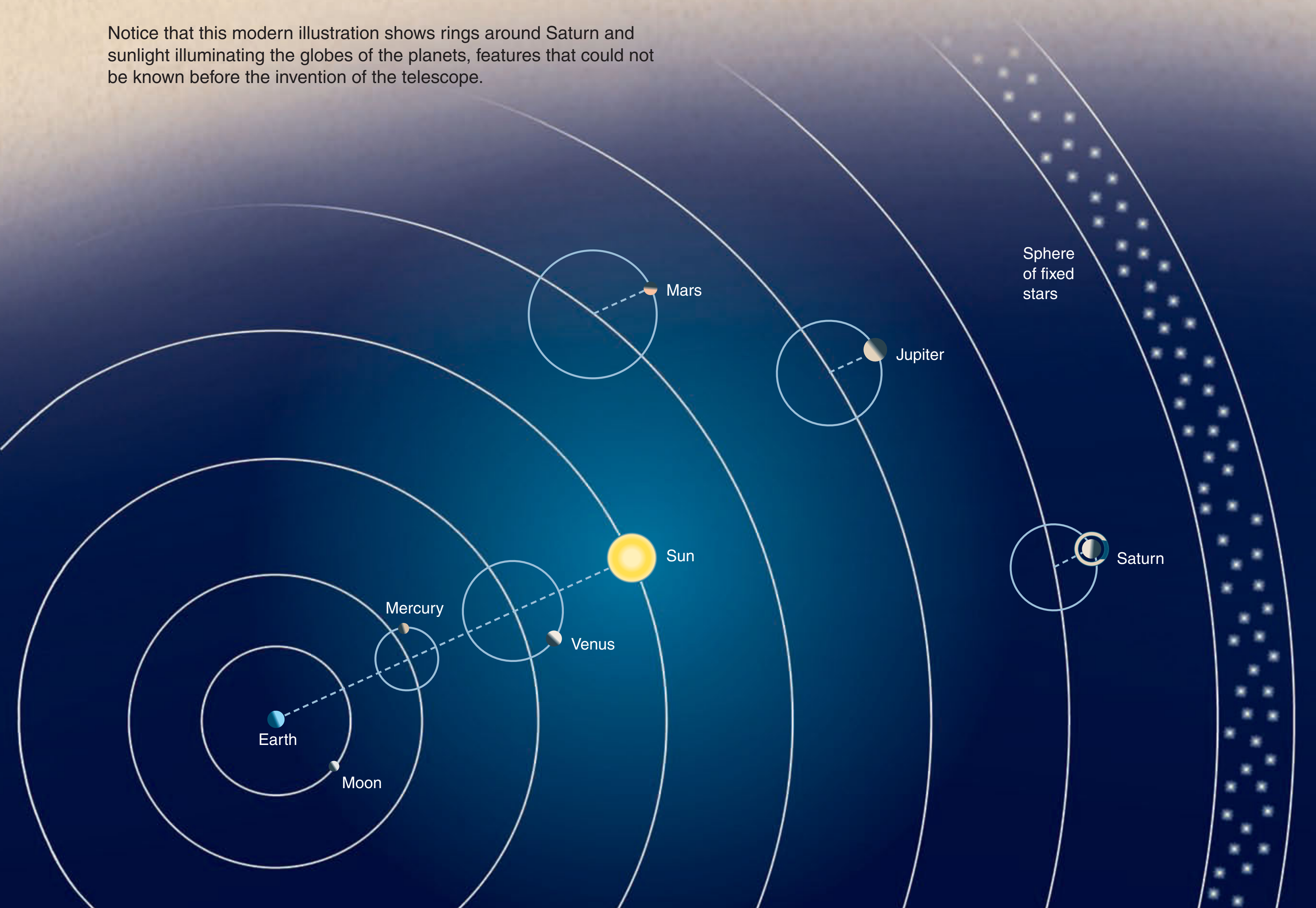
Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Epicycles." Notice how the counterclockwise rotation of the epicycle produces retrograde motion.

3a Ptolemy's great book *Mathematical Syntaxis* (c. AD 140) contained the details of his model. Islamic astronomers preserved and studied the book through the Middle Ages, and they called it *Al Magisti* (The Greatest). When the book was found and translated from Arabic to Latin in the 12th century, it became known as *Almagest*.

3b The Ptolemaic model of the universe shown below was geocentric and based on uniform circular motion. Note that Mercury and Venus were treated differently from the rest of the planets. The centers of the epicycles of Mercury and Venus had to remain on the Earth–Sun line as the sun circled Earth through the year.

Equants and smaller epicycles are not shown here. Some versions contained nearly 100 epicycles as generations of astronomers tried to fine-tune the model to better reproduce the motion of the planets.

Notice that this modern illustration shows rings around Saturn and sunlight illuminating the globes of the planets, features that could not be known before the invention of the telescope.



In Chapter 1, the cosmic zoom gave you a preview of the scale of the universe as you expanded your field of view from Earth to include our solar system, our galaxy, and finally billions of other galaxies. To the ancients, the universe was much smaller. They didn't know about stars and galaxies. Earth lay at the center of their universe surrounded by crystalline shells carrying the planets, and the starry sphere lay just beyond the outermost shell.

Scholars and educated people knew Aristotle's astronomy well. You may have heard the **Common Misconception** that Christopher Columbus had to convince Queen Isabella of Spain that the world was round and not flat. Not so. Like all educated people of her time, the Queen knew the world was round. Aristotle said so. Columbus had to convince the Queen that the world was *small*—so small he could sail to the Orient by heading west. In making his sales pitch, he underestimated the size of Earth and overestimated the eastward extent of Asia, so he thought China and Japan were within a few days' sailing distance of Spain. If North America had not been in his way, he and his crew would have starved to death long before they reached Japan.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Parallax."

◄ **SCIENTIFIC ARGUMENT** ►

Why did classical astronomers conclude the heavens were made up of spheres?

Today, scientific arguments depend on evidence and theory; but, in classical times, philosophers reasoned from first principles. Plato argued that the perfect geometrical figure was a sphere. Then the heavens, which everyone agreed were perfect, must be made up of spheres. The natural motion of a sphere is rotation, and the only perfect motion is uniform motion, so the heavenly spheres were thought to move in uniform circular motion. In this way, classical philosophers argued that the daily motion of the heavens around Earth and the motions of the seven planets (the sun and moon were counted as planets) against the background of the stars had to be produced by the combination of uniformly rotating spheres carrying objects around in perfect circles.

Now build a new argument. Although ancient astronomers didn't use evidence as modern scientists do, they did observe the world around them. **What observations led them to conclude that Earth didn't move?**

◄ ►

4-2 Copernicus

You would not have expected Nicolaus Copernicus to trigger a revolution in astronomy and science. He was born in 1473 to a merchant family in Poland. Orphaned at the age of 10, he was raised by his uncle, an important bishop, who sent him to the University of Cracow and then to the best universities in Italy. There he studied law and medicine before pursuing a lifelong career as an important administrator in the Church. Nevertheless, he had a passion for astronomy (■ Figure 4-1).

The Copernican Model

If you had sat beside Copernicus in his astronomy classes, you would have studied the Ptolemaic universe. The central location of Earth was widely accepted, and everyone knew that the heavens moved by the combination uniform circular motion. For most scholars, questioning these principles was not an option because, over the course of centuries, Aristotle's proposed geometry had become linked with Christian teachings. According to the Aristotelian universe, the most perfect region was in the heavens and the most imperfect at Earth's center. This classical geocentric universe matched the commonly held Christian geometry of heaven and hell, and anyone who criticized the Ptolemaic model was questioning Aristotle's geometry and indirectly challenging belief in heaven and hell.

Copernicus studied the Ptolemaic universe and probably found it difficult at first to consider alternatives. Throughout his life, he was associated with the Catholic Church, which had adopted many of Aristotle's ideas. His uncle was an important bishop in Poland, and, through his uncle's influence, Copernicus was appointed a canon at the cathedral in Frauenberg at the unusually young age of 24. (A canon was not a priest but a Church administrator.) This gave Copernicus an income, although he continued his studies at the universities in Italy. When he left the universities, he joined his uncle and served as his secretary and personal physician until his uncle died in 1512. At that point, Copernicus moved into quarters adjoining the cathedral in Frauenberg, where he served as canon for the rest of his life.

His close connection with the Church notwithstanding, Copernicus began to consider an alternative to the Ptolemaic universe, probably while he was still at university. Sometime before 1514, he wrote an essay proposing a **heliocentric** model in which the sun, not Earth, was the center of the universe. To explain the daily and annual cycles of the sky, he proposed that Earth rotated on its axis and revolved around the sun. He distributed this commentary in handwritten form, without a title, and in some cases anonymously, to friends and astronomical correspondents. He may have been cautious out of modesty, out of respect for the Church, or out of fear that his revolutionary ideas would be attacked unfairly. After all, the place of Earth was a controversial theological subject. Although this early essay discusses every major aspect of his later work, it did not include observations and calculations to add support. His ideas needed supporting evidence, and he began gathering observations and making detailed calculations to be published as a book that would demonstrate the truth of his revolutionary idea.

De Revolutionibus

Copernicus worked on his book *De Revolutionibus Orbium Coelestium (The Revolutions of the Celestial Spheres)* over a period of many years and was essentially finished by about 1529; yet he

Figure 4-1

Nicolaus Copernicus (1473–1543) pursued a lifetime career in the Church, but he was also a talented mathematician and astronomer. His work triggered a revolution in human thought. These stamps were issued in 1973 to mark the 500th anniversary of his birth.

hesitated to publish it even though other astronomers already knew of his theories. Even Church officials, concerned about the reform of the calendar, sought his advice and looked forward to the publication of his book.

One reason he hesitated was that the idea of a heliocentric universe was highly controversial. This was a time of rebellion in the Church—Martin Luther (1483–1546) was speaking harshly about fundamental Church teachings, and others, both scholars and scoundrels, were questioning the authority of the Church. Even matters as abstract as astronomy could stir controversy. Remember, too, that Earth's place in astronomical theory was linked to the geometry of heaven and hell, so moving Earth from its central place was a controversial and perhaps heretical idea.

Another reason Copernicus may have hesitated to publish was that his work was incomplete. His model could not accurately predict planetary positions, so he continued to refine it. Finally in 1540 he allowed the visiting astronomer Joachim Rheticus (1514–1576) to publish an account of the Copernican universe in Rheticus's book *Prima Narratio (First Narrative).* In 1542, Copernicus sent the manuscript for *De Revolutionibus* off to be printed. He died in the spring of 1543 before the printing was completed.

The most important idea in the book was the location of the sun at the center of the universe. That single innovation had an astonishing consequence—the retrograde motion of the planets was immediately explained in a straightforward way without the large epicycles that Ptolemy had used.

In the Copernican system, Earth moves faster along its orbit than the planets that lie farther from the sun. Consequently, Earth periodically overtakes and passes these planets. Imagine that you are in a race car, driving rapidly along the inside lane of a circular racetrack. As you pass slower cars driving in the outer lanes, they fall behind, and if you did not realize you were moving, it would look as if the cars in the outer lanes occasionally slowed to a stop and then backed up for a short interval. ■ Figure 4-2 shows how the same thing happens as Earth passes a planet such as Mars. Although Mars moves steadily along its orbit, as seen from Earth it appears to slow to a stop and move westward (retrograde) as Earth passes it. This happens to any planet whose orbit lies outside Earth's orbit, so the ancient astronomers saw Mars, Jupiter, and Saturn occasionally move retrograde along the ecliptic. Because the planetary orbits do not lie in precisely the same plane, a planet does not resume its eastward motion in precisely the same path it followed earlier. Consequently, it describes a loop whose shape depends on the angle between the orbital planes.

Copernicus could explain retrograde motion without epicycles, and that was impressive. The Copernican system was elegant and simple compared with the whirling epicycles and off-center equants of the Ptolemaic system. You can see Copernicus's own diagram for his heliocentric system in the top stamp in Figure 4-1. However, *De Revolutionibus* failed in one critical way—the Copernican model could not predict the positions of the planets any more accurately than the Ptolemaic system could. To understand why it failed this critical test, you must understand Copernicus and his world.

Copernicus proposed a revolutionary idea in making the planetary system heliocentric, but he was a classical astronomer with tremendous respect for the old concept of uniform circular motion. In fact, Copernicus objected strongly to Ptolemy's use of the equant. It seemed arbitrary to Copernicus, an obvious violation of the elegance of Aristotle's philosophy of the heavens. Copernicus called equants "monstrous" because they undermined both geocentrism and uniform circular motion. In devis-

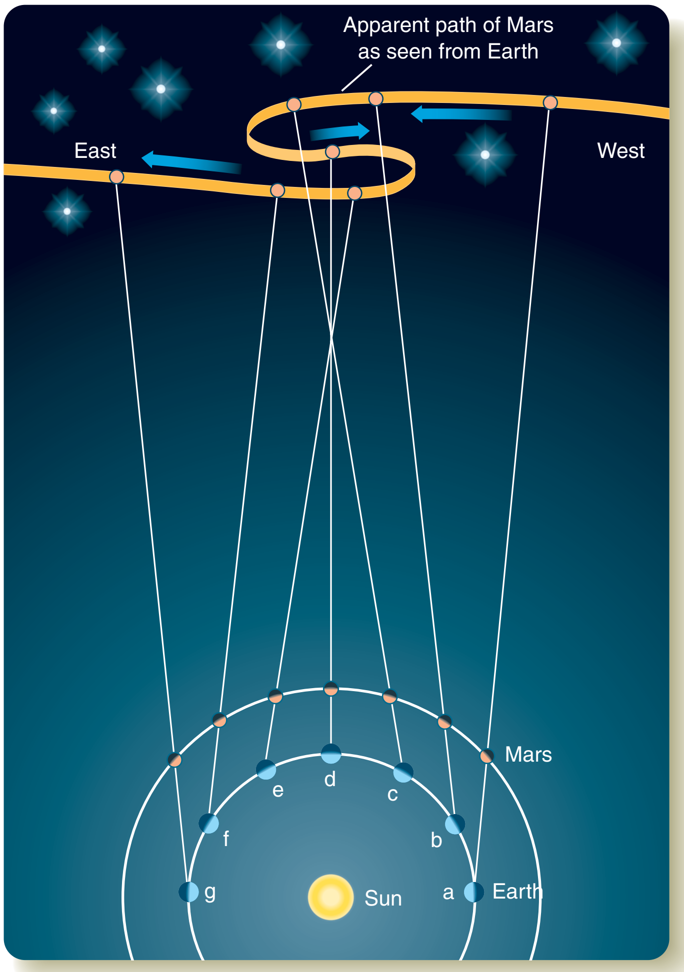


Figure 4-2

The Copernican explanation of retrograde motion. As Earth overtakes Mars (a–c), Mars appears to slow its eastward motion. As Earth passes Mars (d), Mars appears to move westward. As Earth draws ahead of Mars (e–g), Mars resumes its eastward motion against the background stars. The positions of Earth and Mars are shown at equal intervals of 1 month.

ing his model, Copernicus demonstrated a strong belief in uniform circular motion.

Although he did not need epicycles to explain retrograde motion, Copernicus quickly discovered that the sun, moon, and planets suffered other smaller variations in their motions that he could not explain with uniform circular motion centered on the sun. Today astronomers recognize those variations as evidence of elliptical orbits, but because Copernicus held firmly to uniform circular motion, he had to introduce small epicycles to reproduce these minor variations in the motions of the sun, moon, and planets.

Because Copernicus imposed uniform circular motion on his model, it could not accurately predict the motions of the planets. *The Prutenic Tables* (1551) were based on the Copernican model, and they were not significantly more accurate than the 13th century *Alfonsine Tables* that were based on Ptolemy's model. Both could be in error by as much as 2°, which is four times the angular diameter of the full moon.

The Copernican *model* is inaccurate. It includes uniform circular motion and consequently does not precisely describe the motions of the planets. But the Copernican *hypothesis* that the universe is heliocentric is correct, considering how little astronomers of the time knew of other stars and galaxies. The planets circle the sun, not Earth, so the universe that Copernicus knew was heliocentric. Why that hypothesis gradually won acceptance in spite of its inaccuracy is a question historians still debate.

Although astronomers throughout Europe read and admired *De Revolutionibus,* they did not immediately accept the Copernican hypothesis. The mathematics were elegant, and the astro-

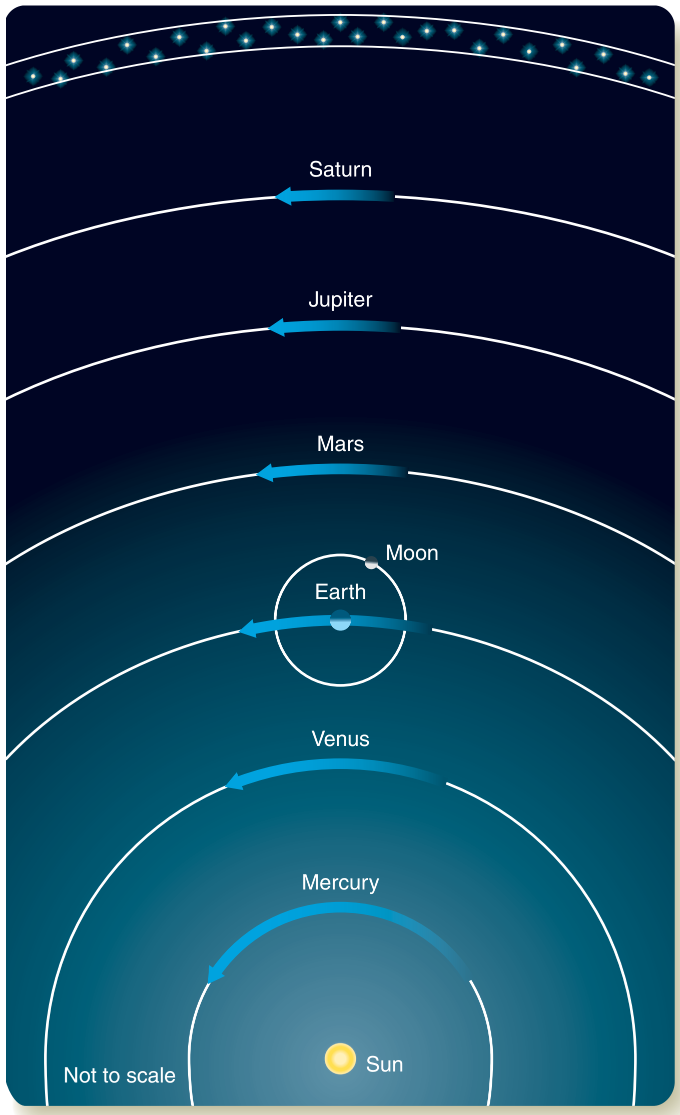


Figure 4-3

The Copernican universe was elegant in its arrangement and its motions. Mercury and Venus are treated just like all the other planets, and orbital velocities (blue arrows) decrease smoothly from that of Mercury, the fastest, to that of Saturn, the slowest. Compare the elegance of this model with the complexity of the Ptolemaic model as shown on page 45.

How Do We Know? 4-1

Scientific Revolutions

How do scientific revolutions occur? You might think from what you know of the scientific method that science grinds forward steadily as new theories are tested against evidence and accepted or rejected. In fact, science sometimes leaps forward in scientific revolutions. The Copernican Revolution is often cited as the perfect example; in a few decades, astronomers rejected the 2000-year-old geocentric model and adopted the heliocentric model. Why does that happen? It's all because scientists are human.

The American philosopher of science Thomas Kuhn has referred to a commonly accepted set of scientific ideas and assumptions as a scientific **paradigm.** The pre-Copernican astronomers shared a geocentric paradigm that included uniform circular motion and the perfection of the heavens. Although they were really smart, they were prisoners of that paradigm. A scientific paradigm is powerful because it shapes your perceptions. It determines what you judge to be important questions and what you judge to be significant evidence. Consequently, the ancient astronomers could not recognize how their geocentric paradigm limited what they understood.

You have seen how the work of Copernicus, Galileo, and Kepler overthrew the geocentric paradigm. Scientific revolutions occur when the deficiencies of the old paradigm build up and finally a scientist has the insight to think "outside the box." Pointing out the failings of the old ideas and proposing a new paradigm with supporting evidence is like poking a hole in a dam; suddenly the pressure is released, and the old paradigm is swept away.

Scientific revolutions are exciting because they give you sudden and dramatic new insights, but they are also times of conflict as new observations and new evidence sweep away old ideas.

The ancients believed the stars were attached to a starry sphere. (NOAO and Nigel Sharp)

nomical observations and calculations were of tremendous value; but few astronomers believed, at first, that the sun actually was the center of the planetary system and that Earth moved. How the Copernican hypothesis was gradually recognized as correct has been called the Copernican Revolution, because it was not just the adoption of a new idea but a total change in the way astronomers thought about the place of the Earth (■ How Do We Know? 4-1).

There are probably a number of reasons why the Copernican hypothesis gradually won support, including the revolutionary temper of the times, but the most important factor may be the elegance of the idea. Placing the sun at the center of the universe produced a symmetry among the motions of the planets that is pleasing to the eye as well as to the intellect (■ Figure 4-3). In the Ptolemaic model, Mercury and Venus were treated differently from the rest of the planets; their epicycles had to remain centered on the Earth–sun line. In the Copernican model, all of the planets were treated the same. They all followed orbits that circled the sun at the center. Furthermore, their speed depended in an orderly way on their distance from the sun, with those closest moving fastest.

The most astonishing consequence of the Copernican hypothesis was not what it said about the sun but what it said about Earth. By placing the sun at the center, Copernicus made Earth move along an orbit like the other planets. By making Earth a planet, Copernicus revolutionized humanity's view of its place in the universe and triggered a controversy that would eventually bring the astronomer Galileo Galilei before the Inquisition. This controversy over the apparent conflict between scientific knowledge and philosophical and theological ideals continues even today.

◄ SCIENTIFIC ARGUMENT ►

Why would you say the Copernican hypothesis was correct but the model was inaccurate?

To build this argument, you must distinguish carefully between a hypothesis and a model. The Copernican hypothesis was that the sun and not Earth was the center of the universe. Given the limited knowledge of the Renaissance astronomers about distant stars and galaxies, that hypothesis was correct.

The Copernican model, however, included not only the heliocentric hypothesis but also uniform circular motion. The model is inaccurate because the planets don't really follow circular orbits, and the small epicycles that Copernicus added to his model never quite reproduced the motions of the planets.

Now build a new argument. The Copernican hypothesis won converts because it is elegant and can explain retrograde motion. **How does its explanation of retrograde motion work, and how is it more elegant than the Ptolemaic explanation?**

◄ ►

4-3 Planetary Motion

The Copernican hypothesis solved the problem of the place of Earth, but it didn't explain planetary motion. If planets don't move in uniform circular motion, how do they move? The puzzle

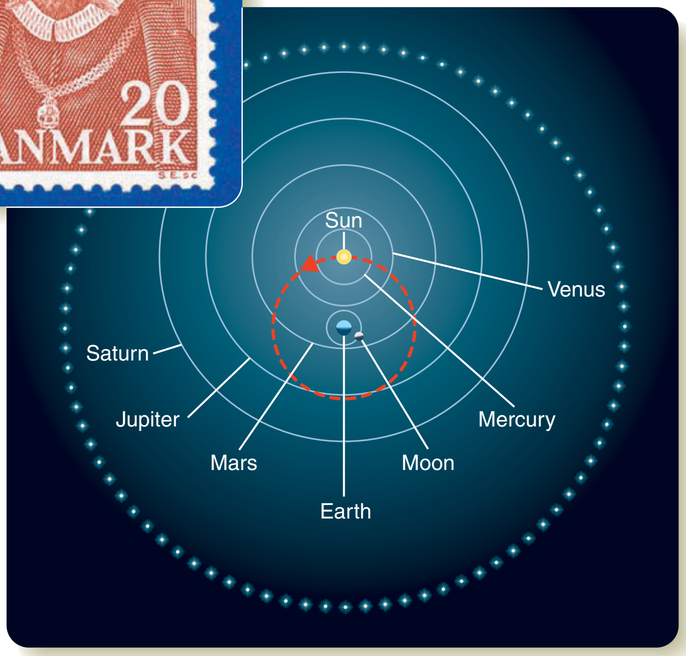


Figure 4-4

Tycho Brahe (1546–1601) was, during his lifetime, the most famous astronomer in the world. Proud of his noble rank, he wears the elephant medal awarded him by the king of Denmark. His artificial nose is suggested in this engraving. Tycho Brahe's model of the universe retained the first principles of classical astronomy; it was geocentric with the sun and moon revolving around Earth, and the planets revolving around the sun. All motion was along circular paths.

of planetary motion was solved during the century following the death of Copernicus through the work of two men. One compiled the observations, and the other did the analysis.

Tycho Brahe

Tycho Brahe (1546–1601) was not a churchman like Copernicus but rather a nobleman from an important family educated at the finest universities. He was well known for his vanity and his lordly manners, and by all accounts he was a proud and haughty nobleman. Tycho's disposition was not improved by a dueling injury from his university days. His nose was badly disfigured and for the rest of his life he wore false noses made of gold and silver and stuck on with wax (■ Figure 4-4).

Although Tycho officially studied law at the university, his real passions were mathematics and astronomy, and early in his university days he began measuring the positions of the planets in the sky. In 1563, Jupiter and Saturn passed very near each other in the sky, nearly merging into a single point on the night of August 24. Tycho found that the *Alfonsine Tables* were a full month in error and that the *Prutenic Tables* were in error by a number of days.

In 1572, a "new star" (now called Tycho's supernova) appeared in the sky, shining more brightly than Venus, and Tycho carefully measured its position. According to classical astronomy, the new star represented a change in the heavens and therefore had to lie below the sphere of the moon. In that case, the new star should show parallax, meaning that it would appear slightly too far east as it rose and slightly too far west as it set. But Tycho saw no parallax in the position of the new star, so he concluded that it must lie above the sphere of the moon and was probably on the starry sphere itself. This contradicted Aristotle's conception of the starry sphere as perfect and unchanging.

No one before Tycho could have made this discovery because no one had ever measured the positions of celestial objects so accurately. Tycho had great confidence in the precision of his measurements, and he had studied astronomy thoroughly, so when he failed to detect parallax for the new star, he knew it was important evidence against the Ptolemaic theory. He announced his discovery in a small book, *De Stella Nova (The New Star)*, published in 1573.

The book attracted the attention of astronomers throughout Europe, and soon Tycho's family introduced him to the court of the Danish King Frederik II, where he was offered funds to build an observatory on the island of Hveen just off the Danish coast. Tycho also received a steady income as lord of a coastal district from which he collected rents. (He was not a popular landlord.) On Hveen, Tycho constructed a luxurious home with six towers especially equipped for astronomy and populated it with servants, assistants, and a dwarf to act as jester. Soon Hveen was an international center of astronomical study.

Tycho Brahe's Legacy

Tycho made no direct contribution to astronomical theory. Because he could measure no parallax for the stars, he concluded that Earth had to be stationary, thus rejecting the Copernican hypothesis. However, he also rejected the Ptolemaic model because of its inaccuracy. Instead he devised a complex model in which Earth was the immobile center of the universe around which the sun and moon moved. The other planets circled the sun (Figure 4-4). The model thus incorporated part of the Copernican model, but in it Earth—not the sun—was stationary. In this way, Tycho preserved the central immobile Earth. Although Tycho's model was very popular at first, the Copernican model replaced it within a century.

The true value of Tycho's work was observational. Because he was able to devise new and better instruments, he was able to make highly accurate observations of the position of the stars, sun, moon, and planets. Tycho had no telescopes—they were

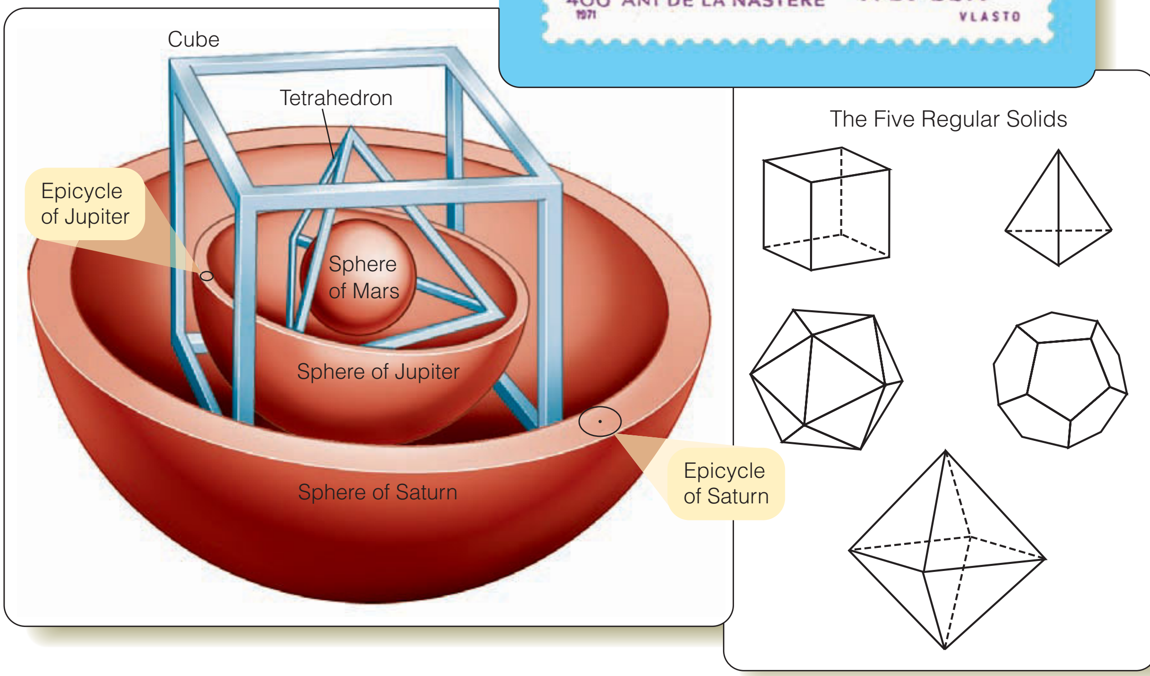


Figure 4-5

Johannes Kepler (1571–1630) was Tycho Brahe's successor. This diagram, based on one drawn by Kepler, shows how he believed the sizes of the celestial spheres carrying the outer three planets—Saturn, Jupiter, and Mars—are determined by spacers (blue) consisting of two of the five regular solids. Inside the sphere of Mars, the remaining regular solids separate the spheres of Earth, Venus, and Mercury. The sun lay at the very center of this Copernican universe based on geometrical spacers.

not invented until the next century—so his observations were made by the naked eye peering along sights. He and his assistants made precise observations for 20 years at Hveen.

Unhappily for Tycho, King Fredrik II died in 1588, and his young son took the throne. Suddenly, Tycho's temper, vanity, and noble presumptions threw him out of favor. In 1596, taking most of his instruments and books of observations, he went to Prague, the capital of Bohemia, and became imperial mathematician to the Holy Roman Emperor Rudolph II. His goal was to revise the *Alfonsine Tables* and publish the result as a monument to his new patron. It would be called the *Rudolphine Tables*.

Tycho did not intend to base the *Rudolphine Tables* on the Ptolemaic system but rather on his own Tyconic system, proving once and for all the validity of his hypothesis. To assist him, he hired a few mathematicians and astronomers, including one Johannes Kepler. Then, in November 1601, Tycho collapsed at a nobleman's home. Before he died, 11 days later, he asked Rudolph II to make Kepler imperial mathematician. The newcomer became Tycho's replacement (though at one-sixth Tycho's salary).

Kepler: An Astronomer of Humble Origins

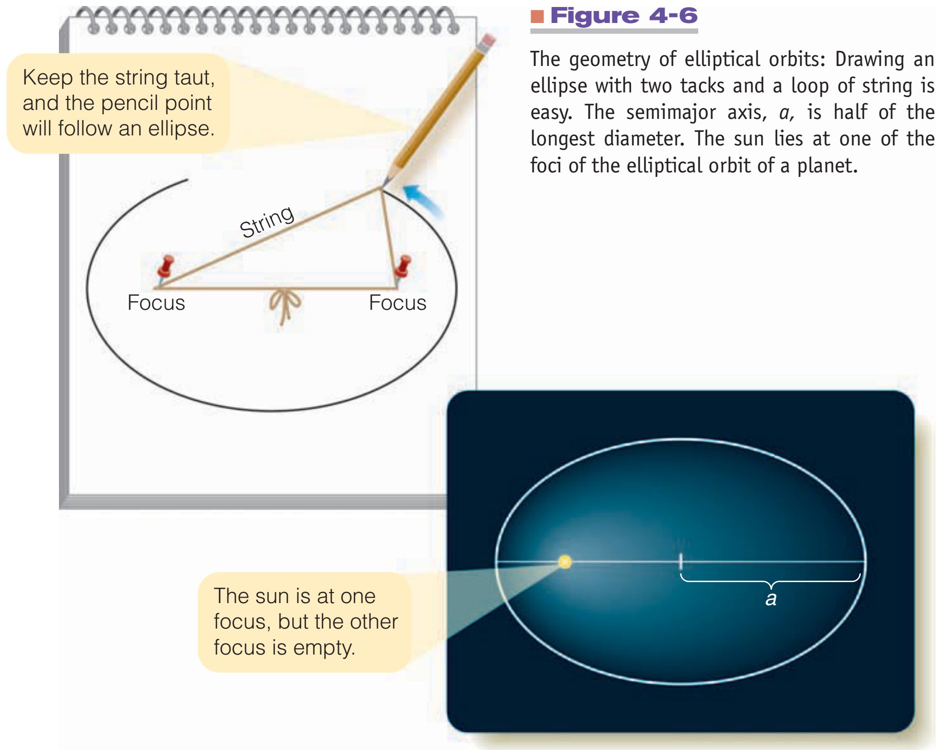
No one could have been more different from Tycho Brahe than Johannes Kepler (■ Figure 4-5). Kepler was born in 1571 to a poor family in a region that is now part of southwest Germany. His father was unreliable and shiftless, principally employed as a mercenary soldier fighting for whoever paid enough. He was often absent for long periods and finally failed to return from a military expedition. Kepler's mother was apparently an unpleasant and unpopular woman. She was accused of witchcraft in later years, and Kepler had to defend her in a trial that dragged on for three years. She was finally acquitted but died the following year.

In spite of family disadvantages and chronic poor health, Kepler did well in school, winning promotion to a Latin school and eventually a scholarship to the university at Tübingen, where he studied to become a Lutheran pastor. During his last year of study, Kepler accepted a job in Graz teaching mathematics and astronomy, a job he resented because he knew little about the subjects. Evidently he was not a good teacher—he had few students his first year and none at all his second. His superiors put him to work teaching a few introductory courses and preparing an annual almanac that contained astronomical, astrological, and weather predictions. Through good luck, in 1595 some of his weather predictions were fulfilled, and he gained a reputation as an astrologer and seer. Even in later life he earned money from his almanacs.

While still a college student, Kepler had become a believer in the Copernican hypothesis, and at Graz he used his extensive spare time to study astronomy. By 1596, the same year Tycho arrived in Prague, Kepler was sure he had solved the mystery of the universe. That year he published a book called *The Forerun-*

ner of Dissertations on the Universe, Containing the Mystery of the Universe*. The book, like nearly all scientific works of that age, was written in Latin and is now known as *Mysterium Cosmographicum*.

By modern standards, the book contains almost nothing of value. It begins with a long appreciation of Copernicanism and then goes on to speculate on the reasons for the spacing of the planetary orbits. Kepler assumed that the heavens could be described by only the most perfect of shapes. Therefore he felt that he had found the underlying architecture of the universe in the sphere plus the five regular solids.* In Kepler's model, the five regular solids became spacers for the orbits of the six planets which were represented by nested spheres (■ Figure 4-5). In fact, Kepler concluded that there could be only six planets (Mercury, Venus, Earth, Mars, Jupiter, and Saturn) because there were only five regular solids to act as spacers between their spheres. He advanced astrological, numerological, and even musical arguments for his theory.



■ **Figure 4-6**

The geometry of elliptical orbits: Drawing an ellipse with two tacks and a loop of string is easy. The semimajor axis, *a,* is half of the longest diameter. The sun lies at one of the foci of the elliptical orbit of a planet.

The second half of the book is no better than the first, but it has one virtue—as Kepler tried to fit the five solids to the planetary orbits, he demonstrated that he was a talented mathematician and that he was well versed in astronomy. He sent copies of his book to Tycho on Hveen and to Galileo in Rome.

Joining Tycho

Life was unsettled for Kepler because of the persecution of Protestants in the region, so when Tycho Brahe invited him to Prague in 1600, Kepler went readily, eager to work with the famous Danish astronomer. Tycho's sudden death in 1601 left Kepler in a position to use the observations from Hveen to analyze the motions of the planets and complete *The Rudolphine Tables.* Tycho's family, recognizing that Kepler was a Copernican and guessing that he would not follow the Tychonic system in completing *The Rudolphine Tables,* sued to recover the instruments and books of observations. The legal wrangle went on for years. Tycho's family did get back the instruments Tycho had brought to Prague, but Kepler had the books, and he kept them.

Whether Kepler had any legal right to Tycho's records is debatable, but he put them to good use. He began by studying the motion of Mars, trying to deduce from the observations how the planet moved. By 1606, he had solved the mystery, this time correctly. The orbit of Mars is an ellipse and not a circle, he said, and with that he abandoned the 2000-year-old belief in the circular motion of the planets. But even this insight was not enough to explain the observations. The planets do not move at uniform speeds along their elliptical orbits. Kepler's analysis showed that they move faster when close to the sun and slower when farther away. With those two brilliant discoveries, Kepler abandoned both circular motion and uniform motion and finally solved the puzzle of planetary motion. He published his results in 1609 in a book called *Astronomia Nova (New Astronomy).*

In spite of the abdication of Rudolph II in 1611, Kepler continued his astronomical studies. He wrote about a supernova that had appeared in 1604 (now known as Kepler's supernova) and about comets, and he wrote a textbook about Copernican astronomy. In 1619, he published *Harmonice Mundi (The Harmony of the World),* in which he returned to the cosmic mysteries of *Mysterium Cosmographicum*. The only thing of note in *Harmonice Mundi* is his discovery that the radii of the planetary orbits are related to the planets' orbital periods. That and his two previous discoveries are so important that they have become known as the three most fundamental rules of orbital motion.

Kepler's Three Laws of Planetary Motion

Although Kepler dabbled in the philosophical arguments of his day, he was at heart a mathematician, and his triumph was his explanation of the motion of the planets. The key to his solution was the ellipse.

*The five regular solids, also known as the Platonic solids, are the tetrahedron, cube, octahedron, dodecahedron, and icosahedron. They were considered perfect because the faces and the angles between the faces are the same at every corner.

How Do We Know? 4-2

Hypothesis, Theory, and Law

Why is a theory much more than just a guess? Scientists study nature by devising new hypotheses and then developing those ideas into theories and laws that describe how nature works. A good example is the connection between sour milk and the spread of disease.

A scientist's first step in solving a natural mystery is to propose a reasonable explanation based on what is known so far. This proposal, called a **hypothesis,** is a single assertion or statement that must then be tested through observation and experimentation. From the time of Aristotle, philosophers believed that food spoils as a result of the spontaneous generation of life — mold out of drying bread. French chemist Louis Pasteur (1822–1895) hypothesized that microorganisms were not spontaneously generated but were carried through the air. To test his hypothesis he sealed an uncontaminated nutrient broth in glass completely protecting it from the mold spores and dust particles in the air; no mold grew, effectively disproving spontaneous generation. Although others had argued against spontaneous generation before Pasteur, it was Pasteur's meticulous testing of his hypothesis through experimentation that finally convinced the scientific community.

A **theory** generalizes the specific results of well-confirmed hypotheses to give a broader description of nature, which can be applied to a wide variety of circumstances. For instance, Pasteur's specific hypothesis about mold growing in broth contributed to a broader theory that disease is caused by microorganisms transmitted from sick people to well people. This theory, called the germ theory of disease, is a cornerstone of modern medicine.

Sometimes when a theory has been refined, tested, and confirmed so often that scientists have great confidence in it, it is called a **natural law.** Natural laws are the most fundamental principles of scientific knowledge. Newton's laws of motion are good examples.

In general, scientists have more confidence in a theory than in a hypothesis and the most confidence in a natural law. However, there is no precise distinction between a theory and a law, and use of these terms is sometimes a matter of tradition. For instance, some textbooks refer to the Copernican "theory" of heliocentrism, but it had not been well tested when Copernicus proposed it, and it is more rightly called the Copernican hypothesis. At the other extreme, Darwin's "theory" of evolution, containing many hypotheses that have been tested and confirmed over and over for nearly 150 years, might more rightly be called a natural law.

A fossil of a 500-million-year-old trilobite: Darwin's theory of evolution has been tested many times and is universally accepted in the life sciences, but by custom it is called Darwin's theory and not Darwin's law. (From the collection of John Coolidge III)

An **ellipse** is a figure drawn around two points, called the *foci,* in such a way that the distance from one focus to any point on the ellipse and back to the other focus equals a constant. This makes it easy to draw ellipses with two thumbtacks and a loop of string. Press the thumbtacks into a board, loop the string about the tacks, and place a pencil in the loop. If you keep the string taut as you move the pencil, it traces out an ellipse (■ Figure 4-6).

The geometry of an ellipse is described by two simple numbers. The **semimajor axis, *a*,** is half of the longest diameter, as you can see in Figure 4-6. The **eccentricity, *e*,** of an ellipse is half the distance between the foci divided by the semimajor axis. The eccentricity of an ellipse tells you its shape; if e is nearly equal to one, the ellipse is very elongated. If e is closer to zero, the ellipse is more circular. To draw a circle with the string and tacks shown in Figure 4-6, you would have to move the two thumbtacks together because a circle is really just an ellipse with eccentricity equal to zero. Try fiddling with real thumbtacks and string, and you'll be surprised how easy it is to draw graceful, smooth ellipses with various eccentricities.

Ellipses are a prominent part of Kepler's three fundamental rules of planetary motion. They have been tested and confirmed so many times that astronomers now refer to them as natural laws (■ How Do We Know? 4-2). They are commonly called Kepler's laws of planetary motion (■ Table 4-1).

Kepler's first law says that the orbits of the planets around the sun are ellipses with the sun at one focus. Thanks to the

■ Table 4-1 | Kepler's Laws of Planetary Motion

I. The orbits of the planets are ellipses with the sun at one focus.
II. A line from a planet to the sun sweeps over equal areas in equal intervals of time.
III. A planet's orbital period squared is proportional to its average distance from the sun cubed:

$$P^2_{\text{y}} = a^3_{\text{AU}}$$

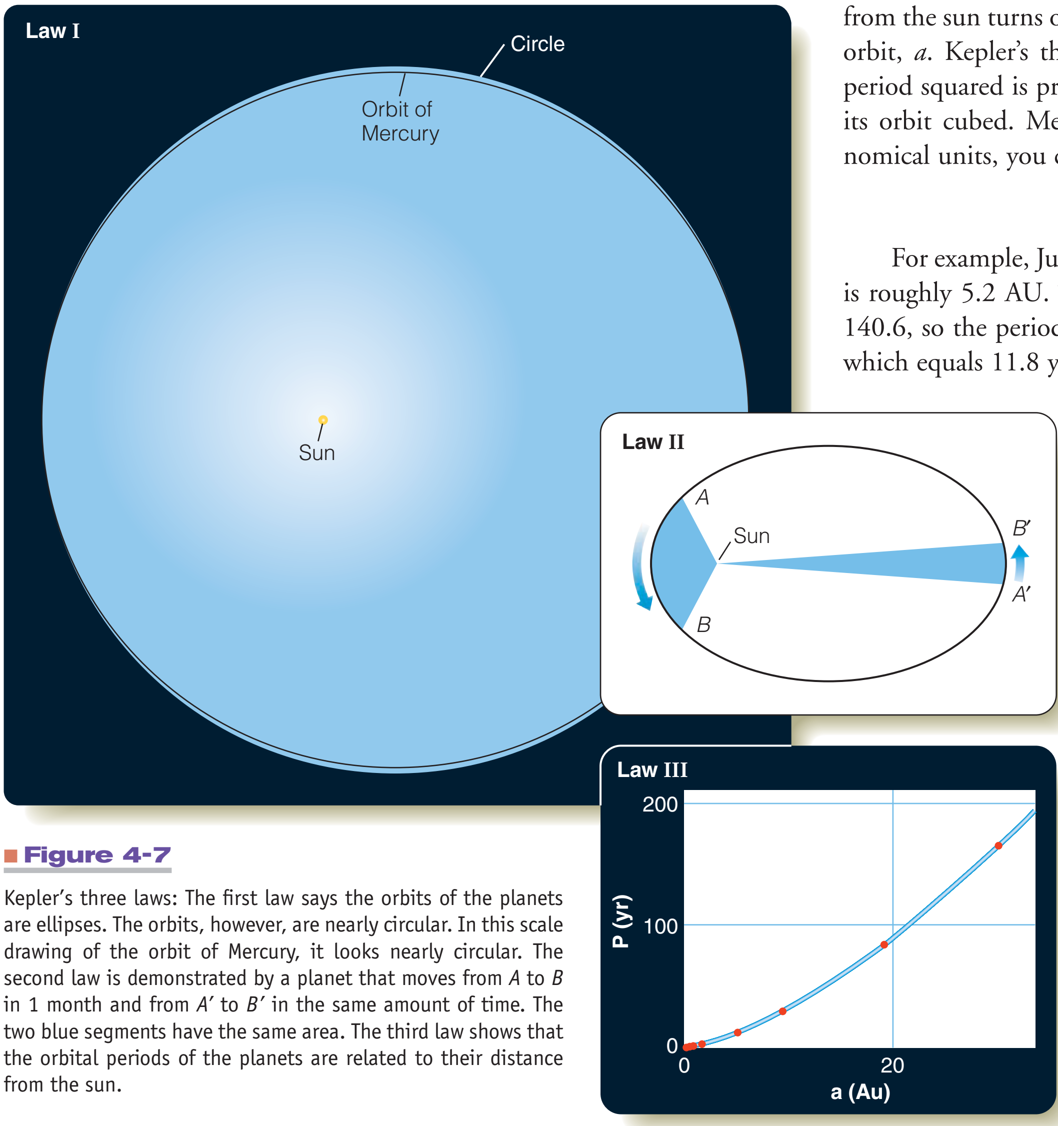


Figure 4-7

Kepler's three laws: The first law says the orbits of the planets are ellipses. The orbits, however, are nearly circular. In this scale drawing of the orbit of Mercury, it looks nearly circular. The second law is demonstrated by a planet that moves from *A* to *B* in 1 month and from *A′* to *B′* in the same amount of time. The two blue segments have the same area. The third law shows that the orbital periods of the planets are related to their distance from the sun.

precision of Tycho's observations and the sophistication of Kepler's mathematics, Kepler was able to recognize the elliptical shape of the orbits even though they are nearly circular. Mercury has the most elliptical orbit, but even it deviates only slightly from a circle (■ Figure 4-7).

Kepler's second law says that an imaginary line drawn from the planet to the sun always sweeps over equal areas in equal intervals of time. This means that when the planet is closer to the sun and the line connecting it to the sun is shorter, the planet moves more rapidly, and the line sweeps over the same area that is swept over when the planet is farther from the sun. You can see how the planet in Figure 4-7 would move from point *A* to point *B* in one month, sweeping over the area shown. But when the planet is farther from the sun, one month's motion would be shorter, from *A′* to *B′*. But the area swept out would be the same.

Kepler's third law relates a planet's orbital period to its average distance from the sun. The orbital period, *P*, is the time a planet takes to travel around the sun once. Its average distance from the sun turns out to equal the semimajor axis of its orbit, *a*. Kepler's third law says that a planet's orbital period squared is proportional to the semimajor axis of its orbit cubed. Measuring *P* in years and *a* in astronomical units, you can summarize the third law as

$$P^2_y = a^3_{AU}$$

For example, Jupiter's average distance from the sun is roughly 5.2 AU. The semimajor axis cubed is about 140.6, so the period must be the square root of 140.6, which equals 11.8 years.

Notice that Kepler's three laws are empirical. That is, they describe a phenomenon without explaining why it occurs. Kepler derived the laws from Tycho's extensive observations, not from any first principle, fundamental assumption, or theory. In fact, Kepler never knew what held the planets in their orbits or why they continued to move around the sun.

The Rudolphine Tables

Kepler continued his mathematical work on *The Rudolphine Tables,* and at last, in 1627, they were ready. He financed their printing himself, dedicating them to the memory of Tycho Brahe. In fact, Tycho's name appears in larger type on the title page than Kepler's own. This is especially surprising because the tables were not based on the Tyconic system but on the heliocentric model of Copernicus and the elliptical orbits of Kepler. The reason for Kepler's evident deference was Tycho's family, still powerful and still intent on protecting Tycho's reputation. They even demanded a share of the profits and the right to censor the book before publication, though they changed nothing but a few words on the title page and added an elaborate dedication to the emperor.

The Rudolphine Tables was Kepler's masterpiece. It could predict the positions of the planets 10 to 100 times more accurately than previous tables. Kepler's tables were the precise model of planetary motion that Copernicus had sought but failed to find. The accuracy of *The Rudolphine Tables* was strong evidence that both Kepler's laws of planetary motion and the Copernican hypothesis for the place of Earth were correct. Copernicus would have been pleased.

Kepler died in 1630. He had solved the problem of planetary motion, and his *Rudolphine Tables* demonstrated his solution. Although he did not understand why the planets moved or why they followed ellipses, insights that had to wait half a century for Isaac Newton, Kepler's three laws worked. In science the only test of a theory is, "Does it describe reality?" Kepler's laws have been used for almost four centuries as a true description of orbital motion.

◄ SCIENTIFIC ARGUMENT ►

How was Kepler's model with regular solids based on first principles? How were his three laws based on evidence?

When he was younger, Kepler argued that the five regular solids were perfect geometrical figures. Along with the sphere, he reasoned, those perfect figures should be part of the perfect heavens. He then arranged the figures to produce the approximate spacing among the spheres that carried the planets in the Copernican model. Kepler's model was based on a belief in the perfection of the heavens.

In contrast, Kepler derived his three laws of motion from the years of observations made by Tycho Brahe during 20 years on Hveen. The observations were the evidence, and they gave Kepler a reality check each time he tried a new calculation. He chose ellipses, for example, because they fit the data and not because he thought ellipses had any special significance.

The Copernican model was a poor predictor of planetary motion, but the *Rudolphine Tables* were much more accurate. **What first principle did Copernicus follow that was abandoned when Kepler looked at the evidence?**

◄ ►

4-4 Galileo Galilei

Most people think they know two facts about Galileo, but both facts are wrong; they are **Common Misconceptions,** so you have probably heard them. Galileo did not invent the telescope, and he was not condemned by the Inquisition for believing that Earth moved around the sun. Then why is Galileo so famous? Why did the Vatican reopen his case in 1979, almost 400 years after his trial? As you learn about Galileo, you will discover that his trial concerned not just the place of Earth and the motion of the planets but also a new and powerful method of understanding nature, a method called science.

Telescopic Observations

Galileo Galilei (■ Figure 4-8) was born in 1564 in Pisa, a city in what is now Italy, and he studied medicine at the university there. His true love, however, was mathematics; and, although he had to leave school early for financial reasons, he returned only four years later as a professor of mathematics. Three years after that he became professor of mathematics at the university at Padua, where he remained for 18 years.

During this time, Galileo seems to have adopted the Copernican model, although he admitted in a 1597 letter to Kepler that he did not support Copernicanism publicly. At that time, the Copernican hypothesis was not considered heretical, but it was hotly debated among astronomers, and Galileo, living in a region controlled by the Church, cautiously avoided trouble. It was the telescope that drove Galileo to publicly defend the heliocentric model.

Galileo did not invent the telescope. It was apparently invented around 1608 by lens makers in Holland. Galileo, hearing descriptions in the fall of 1609, was able to build telescopes in his workshop. In fact, Galileo was not the first person to look at the sky through a telescope, but he was the first person to apply telescopic observations to the theoretical problem of the day—the place of Earth.

What Galileo saw through his telescopes was so amazing that he rushed a small book into print. *Sidereus Nuncius (The Sidereal Messenger)* reported three major discoveries. First, the moon was not perfect. It had mountains and valleys on its surface, and Galileo used the mountain's shadows to calculate their height. Aristotle's philosophy held that the moon was perfect, but Galileo showed that it was not only imperfect but was a world with features like Earth's.

The second discovery reported in the book was that the Milky Way was made up of myriad stars too faint to see with the unaided eye. While intriguing, this could not match Galileo's third discovery. Galileo's telescope revealed four new "planets" circling Jupiter, satellites known today as the Galilean moons of Jupiter (■ Figure 4-9).

■ Figure 4-8

Galileo Galilei (1564–1642), remembered as the great defender of Copernicanism, also made important discoveries in the physics of motion. He is honored here on an old Italian 2000-lira note.

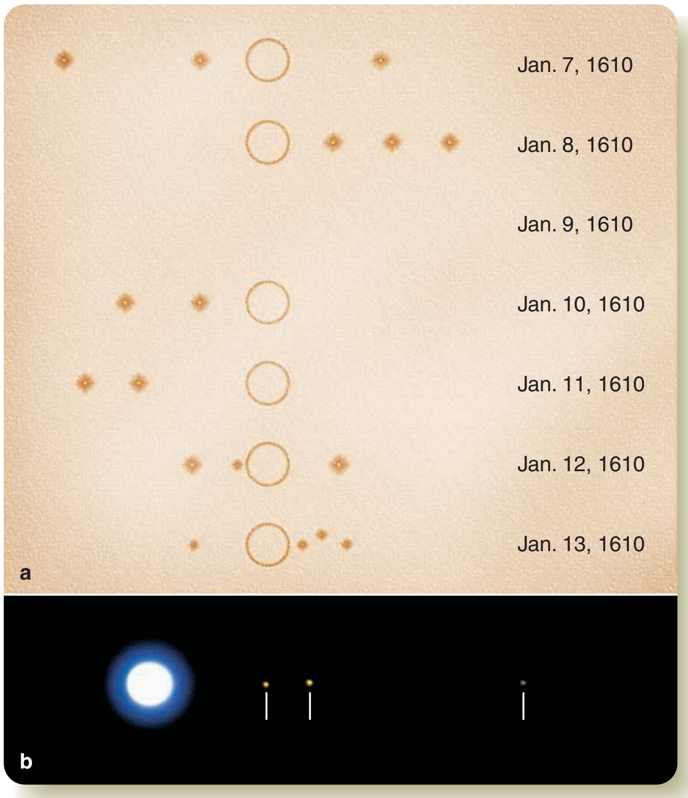


Figure 4-9

(a) On the night of January 7, 1610, Galileo saw three small "stars" near the bright disk of Jupiter and sketched them in his notebook. On subsequent nights (excepting January 9, which was cloudy), he saw that the stars were actually four moons orbiting Jupiter. (b) This photograph taken through a modern telescope shows the overexposed disk of Jupiter and three of the four Galilean moons. (Grundy Observatory)

The moons of Jupiter were strong evidence for the Copernican model. Critics of Copernicus had said Earth could not move because the moon would be left behind; but Galileo's discovery showed that Jupiter, which everyone agreed was moving, was able to keep its satellites. That suggested that Earth, too, could move and keep its moon. Aristotle's philosophy also included the belief that all heavenly motion was centered on Earth. Galileo's observations showed that Jupiter's moons revolve around Jupiter, suggesting that there could be other centers of motion besides Earth.

Some time after *Sidereus Nuncius* was published, Galileo noticed something else that made Jupiter's moons even stronger evidence for the Copernican model. When he measured the orbital periods of the four moons, he found that the innermost moon had the shortest period and that the moons farther from Jupiter had proportionally longer periods. Jupiter's moons made up a harmonious system ruled by Jupiter, just as the planets in the Copernican universe were a harmonious system ruled by the sun. (See Figure 4-3.) The similarity isn't proof, but Galileo saw it as an argument that the solar system was sun centered and not Earth centered.

In the years following publication of *Sidereus Nuncius,* Galileo made two additional discoveries. When he observed the sun, he discovered sunspots, raising the suspicion that the sun was less than perfect. Further, by noting the movement of the spots, he concluded that the sun was a sphere and that it rotated on its axis.

His most dramatic discovery came when he observed Venus. Galileo saw that it was going through phases like those of the moon. In the Ptolemaic model, Venus moves around an epicycle centered on a line between Earth and the sun. That means it would always be seen as a crescent (■ Figure 4-10a). But Galileo saw Venus go through a complete set of phases, which proved that it did indeed revolve around the sun (Figure 4-10b). There is no way the Ptolemaic model could produce those phases. This was the strongest evidence that came from Galileo's telescope, but when controversy erupted, it focused more on the perfection of the sun and moon and the motion of the satellites of Jupiter.

Sidereus Nuncius was very popular and made Galileo famous. He became chief mathematician and philosopher to the Grand Duke of Tuscany in Florence. In 1611, Galileo visited Rome and was treated with great respect. He had long, friendly discussions with the powerful Cardinal Barberini, but he also made enemies. Personally, Galileo was outspoken, forceful, and sometimes tactless. He enjoyed debate, but most of all he enjoyed being right. In lectures, debates, and letters he offended important people who questioned his telescopic discoveries.

By 1616, Galileo was the center of a storm of controversy. Some critics said he was wrong, and others said he was lying. Some refused to look through a telescope lest it mislead them, and others looked and claimed to see nothing (hardly surprising, given the awkwardness of those first telescopes). Pope Paul V decided to end the disruption, so when Galileo visited Rome in 1616 Cardinal Bellarmine interviewed him privately and ordered him to cease debate. There is some controversy today about the nature of Galileo's instructions, but he did not pursue astronomy for some years after the interview. Books relevant to Copernicanism were banned in all Catholic lands, although *De Revolutionibus,* recognized as an important and useful book in astronomy, was only suspended pending revision. Everyone who owned a copy of the book was required to cross out certain statements and add handwritten corrections stating that Earth's motion and the central location of the sun were only theories and not facts.

Dialogo and Trial

In 1621 Pope Paul V died, and his successor, Pope Gregory XV, died in 1623. The next pope was Galileo's friend Cardinal Barberini, who took the name Urban VIII. Galileo rushed to Rome hoping to have the prohibition of 1616 lifted; and, although the new pope did not revoke the orders, he did apparently encourage Galileo. Soon after returning home, Galileo began to write his great defense of the Copernican model, finally completing it on December 24, 1629. After some delay, the book

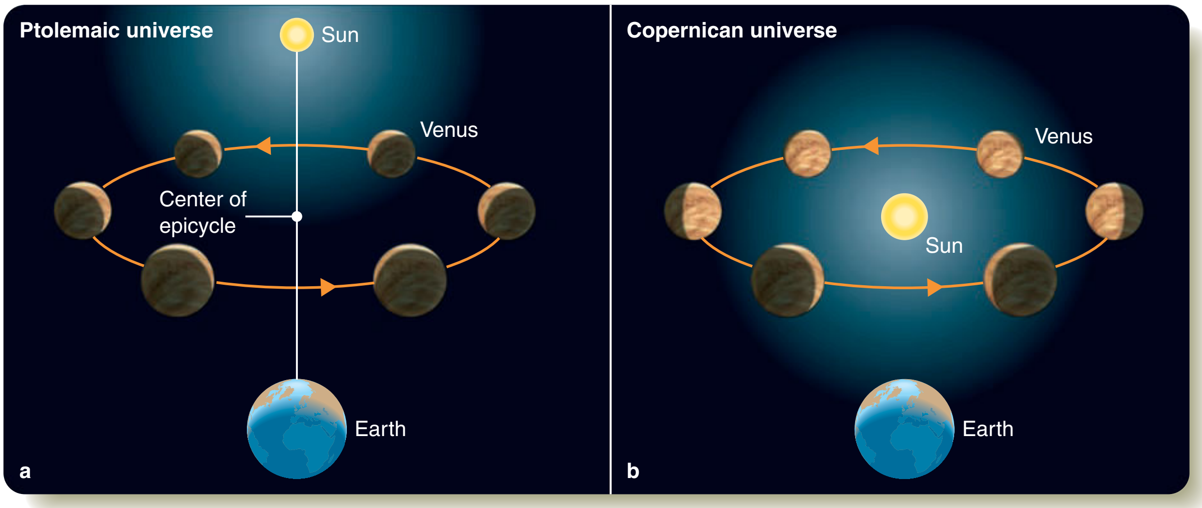


Figure 4-10

(a) If Venus moved in an epicycle centered on the Earth–sun line, it would always appear as a crescent. (b) Galileo's telescope showed that Venus goes through a full set of phases, proving that it must orbit the sun.

was approved by both the local censor in Florence and the head censor of the Vatican in Rome. It was printed in February 1632.

Called *Dialogo Sopra i Due Massimi Sistemi del Mondo (Dialogue Concerning the Two Chief World Systems),* it confronts the ancient astronomy of Aristotle and Ptolemy with the Copernican model and with telescopic observations as evidence. Galileo wrote the book as a debate among three friends. Salviati, a swift-tongued defender of Copernicus, dominates the book; Sagredo is intelligent but largely uninformed. Simplicio, the dismal defender of Ptolemy, makes all the old arguments and sometimes doesn't seem very bright.

The publication of *Dialogo* created a storm of controversy, and it was sold out by August 1632, when the Inquisition ordered sales stopped. The book was a clear defense of Copernicus, and, probably unintentionally, Galileo exposed the pope's authority to ridicule. Urban VIII was fond of arguing that, as God was omnipotent, He could construct the universe in any form while making it appear to humans to have a different form, and thus its true nature could not be deduced by mere observation. Galileo placed the pope's argument in the mouth of Simplicio, and Galileo's enemies showed the passage to the pope as an example of Galileo's disrespect. The pope thereupon ordered Galileo to face the Inquisition.

Galileo was interrogated by the Inquisition four times and was threatened with torture. He must have thought often of Giordano Bruno, a philosopher, poet, and member of the Dominican order, who was tried, condemned, and burned at the stake in Rome in 1600. One of Bruno's offenses had been Copernicanism. However, Galileo's trial did not center on his belief in Copernicanism. *Dialogo* had been approved by two censors. Rather, the trial centered on the instructions given Galileo in 1616. From his file in the Vatican, his accusers produced a record of the meeting between Galileo and Cardinal Bellarmine that included the statement that Galileo was "not to hold, teach, or defend in any way" the principles of Copernicus. Some historians believe that this document, which was signed neither by Galileo nor by Bellarmine nor by a legal secretary, was a forgery. Others suspect it may be a draft that was never used. It is quite possible that Galileo's actual instructions were much less restrictive; but, in any case, Bellarmine was dead and could not testify at Galileo's trial.

The Inquisition condemned Galileo not for heresy but for disobeying the orders given him in 1616. On June 22, 1633, at the age of 70, kneeling before the Inquisition, Galileo read a recantation admitting his errors. Tradition has it that as he rose he whispered *"E pur si muove"* ("Still it moves"), referring to Earth.

Although he was sentenced to life imprisonment, he was actually confined at his villa for the next ten years, perhaps through the intervention of the pope. He died there on January 8, 1642, 99 years after the death of Copernicus.

Galileo was not condemned for heresy, nor was the Inquisition interested when he tried to defend Copernicanism. He was tried and condemned on a charge you might call a technicality. Then why is his trial so important that historians have studied it for almost four centuries? Why have some of the world's greatest authors, including Bertolt Brecht, written about Galileo's trial? Why in 1979 did Pope John Paul II create a commission to reexamine the case against Galileo?

To understand the trial, you must recognize that it was the result of a conflict between two ways of understanding the universe. Since the Middle Ages, scholars had taught that the only path to true understanding was through religious faith. St. Augustine (AD 354–430) wrote *"Credo ut intelligame,"* which can be translated as "Believe in order to understand." Galileo and other scientists of the Renaissance, however, used their own observations as evidence to try to understand the universe. When their observations contradicted Scripture, they assumed that it was their observations that truly represented reality. Galileo paraphrased Cardinal Baronius in saying, "The Bible tells us how to go to heaven, not how the heavens go." The trial of Galileo was not about the place of Earth in the universe. It was not about Copernicanism. It wasn't really about the instructions Galileo received in 1616. It was, in a larger sense, about the birth of modern science as a rational way to understand the universe (■ Figure 4-11). The commission appointed by John Paul II in 1979, reporting its conclusions in October 1992, said of Galileo's inquisitors, "This subjective error of judgment, so clear to us today, led them to a disciplinary measure from which Galileo 'had much to suffer.'" Galileo was not found innocent in 1992 so much as the Inquisition was forgiven for having charged him in the first place.

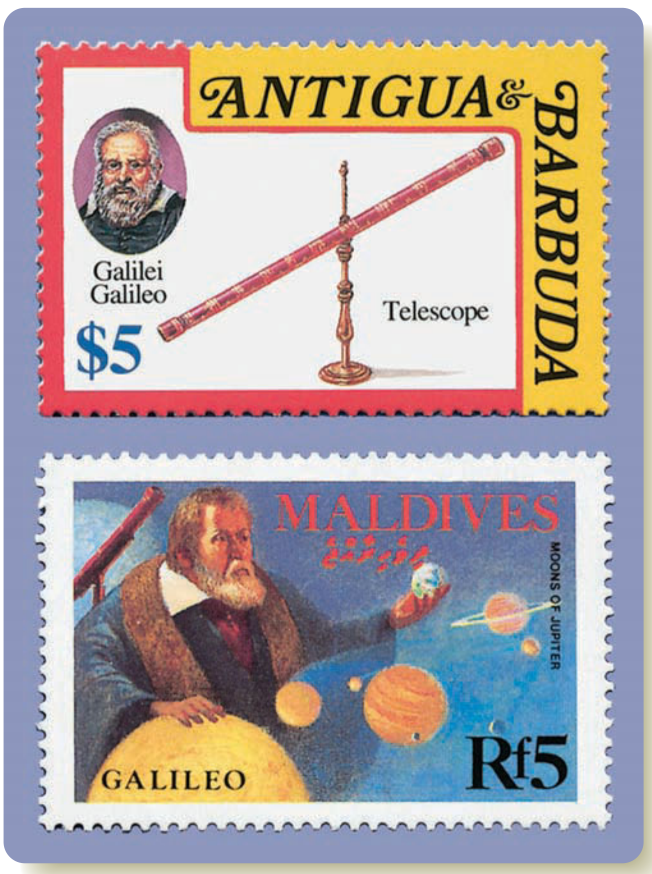
◄ SCIENTIFIC ARGUMENT ►

How were Galileo's observations of the moons of Jupiter evidence against the Ptolemaic model?

Scientific arguments are based on evidence, and reasoning from evidence was Galileo's fundamental way of knowing about the heavens. Galileo presented his arguments in the form of evidence and conclusions, and the moons of Jupiter were key evidence. Ptolemaic astronomers argued that Earth could not move or it would lose its moon, but even in the Ptolemaic universe Jupiter moved, and the telescope showed that it had moons and kept them. Evidently, Earth could move and not leave its moon behind. Furthermore, moons circling Jupiter did not fit the classical belief that all motion was centered on Earth. Obviously there could be other centers of motion. Finally, the orbital periods of the moons were related to their distance from Jupiter, just as the orbital periods of the planets were, in the Copernican system, related to their distance from the sun. This similarity suggested that the sun rules its harmonious family of planets just as Jupiter rules its harmonious family of moons.

Of all of Galileo's telescopic observations, the moons of Jupiter caused the most debate, but the craters on the moon and the phases of Venus were also critical evidence. Build an argument to discuss that evidence. **How did craters on the moon and the phases of Venus argue against the Ptolemaic model?**

◄ ►



■ Figure 4-11

Although he did not invent it, Galileo will always be remembered along with the telescope because it was the source of the evidence from which he reasoned. By depending on direct observation of reality instead of the first principles of philosophy and theology, Galileo led the way to the invention of modern astronomy and modern science as a way to know about the natural world.

4-5 Isaac Newton and Orbital Motion

THE BIRTH OF modern astronomy and of modern science date from the 99 years between the deaths of Copernicus and Galileo. The Renaissance is commonly taken to be the period between 1350 and 1600, and that places the 99 years of this story at the culmination of the reawakening of learning in all fields (■ Figure 4-12). Not only did the world adopt a new model of the universe, but it also adopted a new way of understanding humanity's place in nature.

The problem of the place of Earth was resolved by the Copernican Revolution, but the problem of planetary motion was only partly solved by Kepler's laws. For the last 10 years of his life, Galileo studied the nature of motion, especially the accelerated motion of falling bodies. Although he made some important progress, he was not able to relate his discoveries about motion on Earth to that in the heavens. That final step fell to Isaac Newton.

Isaac Newton

Galileo died in January 1642. Some 11 months later, on Christmas day 1642,* a baby was born in the English village of Woolsthorpe. His name was Isaac Newton (■ Figure 4-13), and his life represented the first flower of the seeds planted by the four astronomers in this story. Newton was a quiet child from a farming family, but his work at school was so impressive that his uncle financed his education at Trinity College, where he studied mathematics and physics. In 1665, plague swept through England, and the colleges were closed. During 1665 and 1666, Newton spent his time at home in Woolsthorpe, thinking and studying. It was during these years that he made most of his discoveries in optics, mechanics, and mathematics. Among other things, he studied optics, developed three laws of motion, divined the nature of gravity, and invented differential calculus. The publication of his work in his book *Principia* in 1687 placed science on a firm analytical base.

*Because England had not yet reformed its calendar, December 25, 1642, in England was January 4, 1643, in Europe. It is only a small deception to use the English date and thus include Newton's birth in our 99-year history.

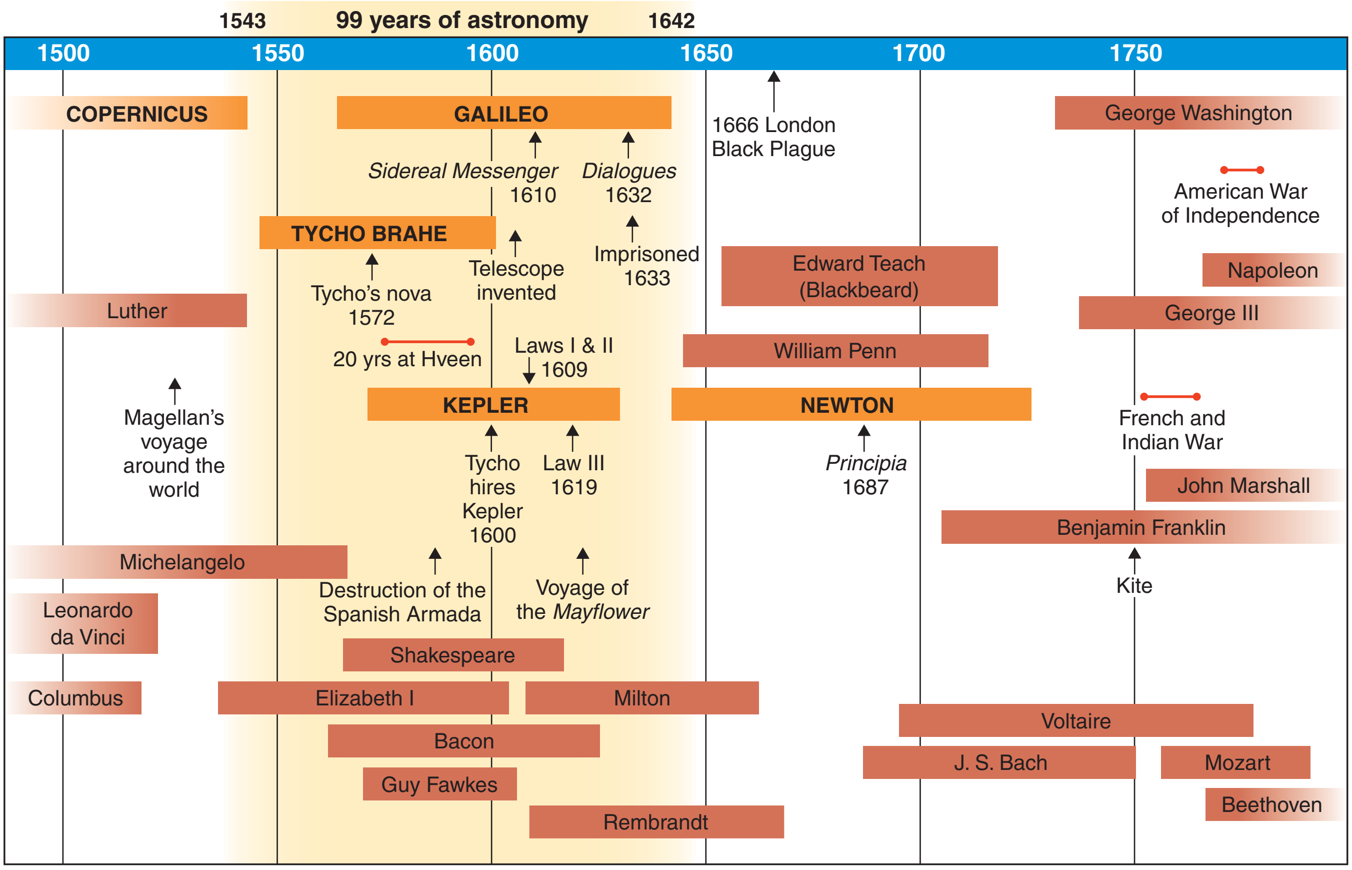


Figure 4-12

The 99 years between the death of Copernicus in 1543 and the birth of Newton in 1642 marked the transition from Aristotle's ancient astronomy as modeled by Ptolemy to Newton's modern understanding of motion and gravity. This period saw the birth of modern science as a way to understand the universe.

It is beyond the scope of this book to analyze all of Newton's work, but his laws of motion and gravity shaped the future of astronomy. From his study of the work of Galileo, Kepler, and others, Newton extracted three laws that relate the motion of a body to the forces acting on it (■ Table 4-2). These laws made it possible to predict exactly how a body would move if the forces were known (■ How Do We Know? 4-3).

When Newton thought carefully about motion, he realized that some force must pull the moon toward Earth's center. If there were no such force altering the moon's motion, it would continue moving in a straight line and leave Earth forever. It can circle Earth only if Earth attracts it. Newton's insight was to recognize that the force that holds the moon in its orbit is the same as the force that makes apples fall from trees—gravity.

Figure 4-13

Isaac Newton (1642–1727) worked from the discoveries of Galileo and Kepler to study motion and gravitation. He and some of his discoveries were honored on this old English 1-pound note.

How Do We Know? | 4-3

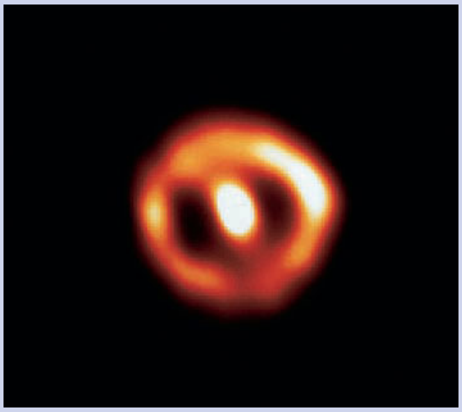
Cause and Effect

Why is cause and effect so important to scientists? One of the most often used and least often stated principles of science is cause and effect. Ancient philosophers such as Aristotle argued that objects moved because of tendencies. They said that air and fire had a natural tendency to move away from the center of the universe, and thus they rise. This natural motion had no cause but was inherent in the nature of the objects. Modern scientists all believe that events have causes and, for example, that things move because of forces.

Newton's second law of motion *($F = ma$)* was the first clear statement of the principle of cause and effect. If an object (of mass m) changes its motion (a in the equation), then it must be acted on by a force (F in the equation). Any effect *(a)* must be the result of a cause *(F)*.

The principle of cause and effect goes far beyond motion. It gives scientists confidence that every effect has a cause. The struggle against disease is an example. Cholera is a horrible disease that can kill its victims in hours. Long ago it was probably blamed on bad magic or the will of the gods, and only two centuries ago it was blamed on "bad air." When an epidemic of cholera struck England in 1854, Dr. John Snow carefully mapped cases in London showing that the victims had drunk water from a small number of wells contaminated by sewage. In 1876, the German Dr. Robert Koch traced cholera to an even more specific cause when he identified the microscopic bacillus that causes the disease. Step by step, scientists tracked down the cause of cholera.

If the universe did not depend on cause and effect, then you could never expect to understand how nature works. Newton's second law of motion was arguably the first clear statement that the behavior of the universe depends rationally on causes.



Cause and effect: Why did this star explode in 1992? There must have been a cause. (ESA/STScI and NASA)

Table 4-2 | Newton's Three Laws of Motion

I. A body continues at rest or in uniform motion in a straight line unless acted upon by some force.

II. The change of motion *(a)* of a body of mass m is proportional to the force *(F)* acting on it and is in the direction of the force.

$$F = ma$$

III. When one body exerts a force on a second body, the second body exerts an equal and opposite force back on the first body.

Newtonian gravitation is sometimes called universal mutual gravitation. Newton's third law points out that forces occur in pairs, so if one body attracts another, the second body must also attract the first. Thus gravitation must be mutual. Furthermore, gravity must be universal. That is, all masses must attract all other masses in the universe. The force between two bodies depends on the masses of the bodies and the distance between them.

The **mass** of an object is a measure of the amount of matter in the object, usually expressed in kilograms. Mass is not the same as weight. An object's weight is the force that Earth's gravity exerts on the object. An object in space far from Earth would have no weight, but it would contain the same amount of matter and would thus have the same mass that it had on Earth.

Newton realized that, in addition to mass, the distance between two objects affects the gravitational attraction between them. He recognized that the force of gravity decreases as the square of the distance between the objects increases. Specifically, if the distance from, say, Earth to the moon were doubled, the gravitational force between them would decrease by a factor of 2^2, which equals 4. If the distance were tripled, the force would decrease by a factor of 3^2, which equals 9. This relationship is known as the **inverse square relation.** (This relation is discussed in more detail in Chapter 8, where it is applied to the intensity of light.)

With these definitions of mass and the inverse square relation, you can describe Newton's law of gravity in a simple equation:

$$F = -G\frac{Mm}{r^2}$$

Here F is the force of gravity acting between two objects of mass M and m, and r is the distance between their centers. G is the gravitational constant, just a number that depends on the units used for mass, distance, and force. The minus sign reminds you that the force is attractive, tending to make r decrease. To summarize, the force of gravity attracting two objects to each other equals the gravitational constant times the product of their masses divided by the square of the distance between the objects.

Orbital Motion

Newton's laws of motion and gravitation make it possible to understand why the moon orbits Earth and how the planets move along their orbits around the sun. You can even discover why Kepler's laws work.

To understand how an object can orbit another object, it helps to describe orbital motion as Newton did—as a form of falling. Study ■ **Orbiting Earth** on pages 62–63 and notice three important ideas and six new terms:

1. An object orbiting Earth is actually falling (being accelerated) toward Earth's center. The object continuously misses Earth because of its motion. To maintain a circular orbit, the object must move with *circular velocity,* which, for example, explains how *geosynchronous satellites* can remain fixed above one spot on Earth.

2. Also, two objects orbiting each other actually revolve around their *center of mass.*

3. Finally, notice the difference between *closed orbits* and *open orbits*. If you want to leave Earth never to return, you must give your spaceship a high enough velocity *(escape velocity)* so it will follow an open orbit.

When the captain of a spaceship says, "Put us into a circular orbit," the ship's computers must quickly calculate the velocity needed to achieve a circular orbit. That circular velocity depends only on the mass of the planet and the distance from the center of the planet (■ Reasoning with Numbers 4-1). Once the engines fire and the ship reaches circular velocity, the engines can shut down. The spaceship is then in orbit and will fall around the planet forever so long as it is above the atmosphere where there is no friction. No further effort is needed to maintain orbit, thanks to Newton's laws.

You have probably seen a **Common Misconception** if you watch science fiction movies. People in spaceships are usually shown walking around as if they had gravity holding them to the floor. Of course, they should be floating in free fall in their spaceships, unless the rockets are firing, in which case the crew should be strapped into their seats. Authors invent artificial gravity to explain this problem away, but no physicist has ever found a way to generate artificial gravity.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Falling Bodies," "Orbital Motion," and "Escape Velocity."

Reasoning with Numbers | 4-1

Circular Velocity

Circular velocity is the velocity a satellite must have to remain in a circular orbit around a larger body. If the mass of the satellite is small compared with the central body, then the circular velocity is given by

$$V_c = \sqrt{\frac{GM}{r}}$$

In this formula, M is the mass of the central body in kilograms, r is the radius of the orbit in meters, and G is the gravitational constant, 6.67×10^{-11} m^3/s^2kg. This formula is all you need to calculate how fast an object must travel to stay in a circular orbit.

For example, how fast does the moon travel in its orbit? The mass of Earth is 5.98×10^{24} kg, and the radius of the moon's orbit is 3.84×10^8 m. The moon's velocity is

$$V_c = \sqrt{\frac{6.67 \times 10^{-11} \times 5.98 \times 10^{24}}{3.84 \times 10^8}}$$

$$V_c = \sqrt{\frac{39.9 \times 10^{13}}{3.84 \times 10^8}}$$

$$V_c = \sqrt{1.04 \times 10^6} = 1020 \text{ m/s}$$

This calculation shows that the moon travels 1.02 km along its orbit each second.

Tides

Newton understood that gravity is mutual—Earth attracts the moon, and the moon attracts Earth—and that means the moon's gravity can explain the ocean tides. But Newton also realized that gravitation is universal, and that means there is much more to tides than just Earth's oceans.

Tides are caused by small differences in gravitational forces. For example, Earth's gravity attracts your body downward with a force equal to your weight. The moon is less massive and more distant, so it attracts your body with a force that is a tiny percent of your weight. You don't notice that little force, but Earth's oceans respond dramatically.

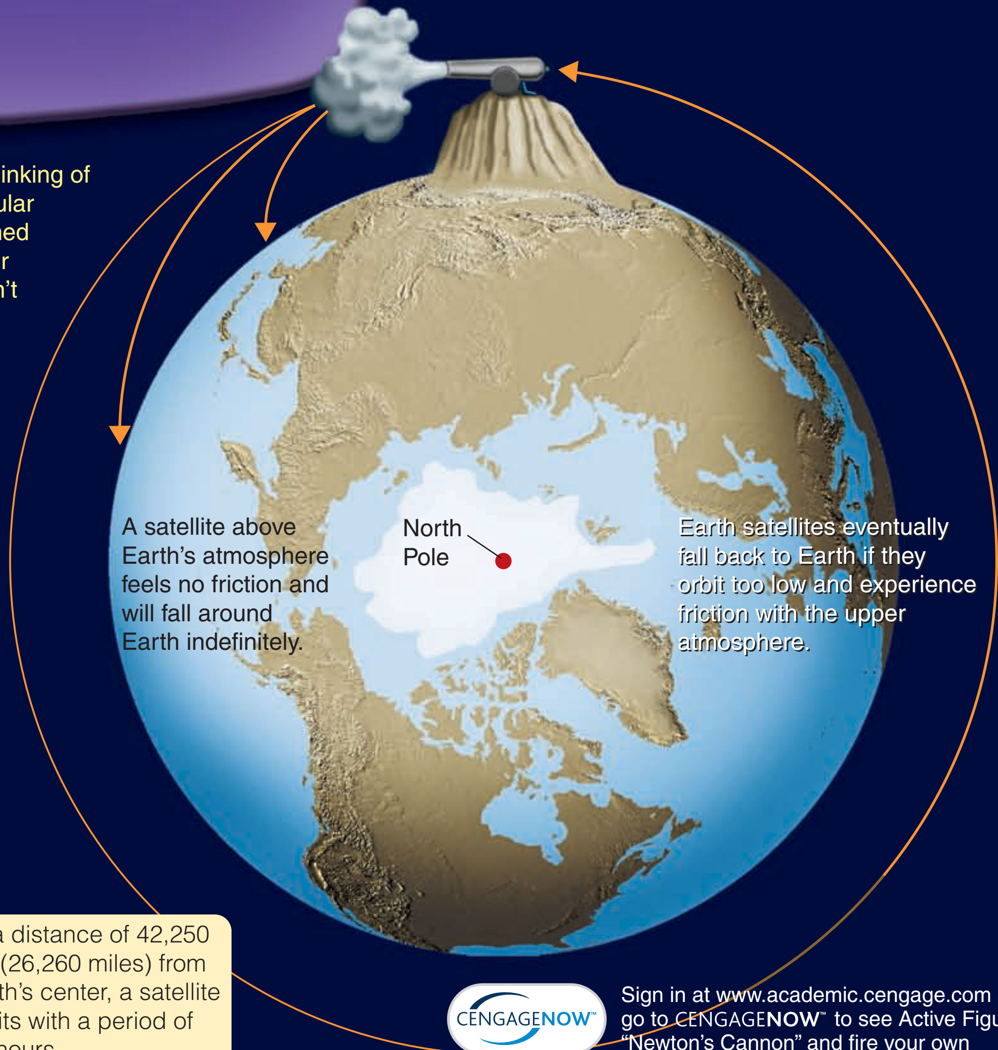
The side of Earth that faces the moon is about 4000 miles closer to the moon than is the center of Earth. Consequently, the moon's gravity, tiny though it is at the distance of Earth, is just a bit stronger when it acts on the near side of Earth than on the center. It pulls on the oceans on the near side of Earth a bit more strongly than on Earth's center, and the oceans respond by flowing into a bulge of water on the side of Earth facing the moon. There is also a bulge on the side of Earth that faces away from the moon because the moon pulls more strongly on Earth's center than on the far side. Thus the moon pulls Earth away from the oceans, which flow into a bulge away from the moon as shown at the top of ■ Figure 4-14.

You might wonder: If Earth and moon accelerate toward each other, why don't they smash together? The answer is that they would collide in about two weeks except that they are orbiting around their common center of mass. The ocean tides are caused by the accelerations Earth and its oceans feel as they move around that center of mass.

A **Common Misconception** holds that the moon's effect on tides means that the moon has an affinity for water—including the water in your body—and, according to some people, that's how the moon makes you behave in weird ways. That's not true.

Orbiting Earth

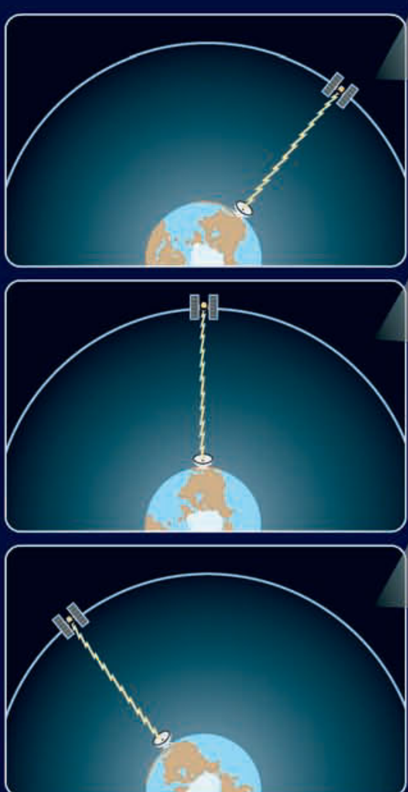
1 You can understand orbital motion by thinking of a cannonball falling around Earth in a circular path. Imagine a cannon on a high mountain aimed horizontally as shown at right. A little gunpowder gives the cannonball a low velocity, and it doesn't travel very far before falling to Earth. More gunpowder gives the cannonball a higher velocity, and it travels farther. With enough gunpowder, the cannonball travels so fast it never strikes the ground. Earth's gravity pulls it toward Earth's center, but Earth's surface curves away from it at the same rate it falls. It is in orbit. The velocity needed to stay in a circular orbit is called the **circular velocity.** Just above Earth's atmosphere, circular velocity is 7790 m/s or about 17,400 miles per hour, and the orbital period is about 90 minutes.



Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Newton's Cannon" and fire your own version of Newton's cannon.

1a A **geosynchronous satellite** orbits eastward with the rotation of Earth and remains above a fixed spot — ideal for communications and weather satellites.

A Geosynchronous Satellite



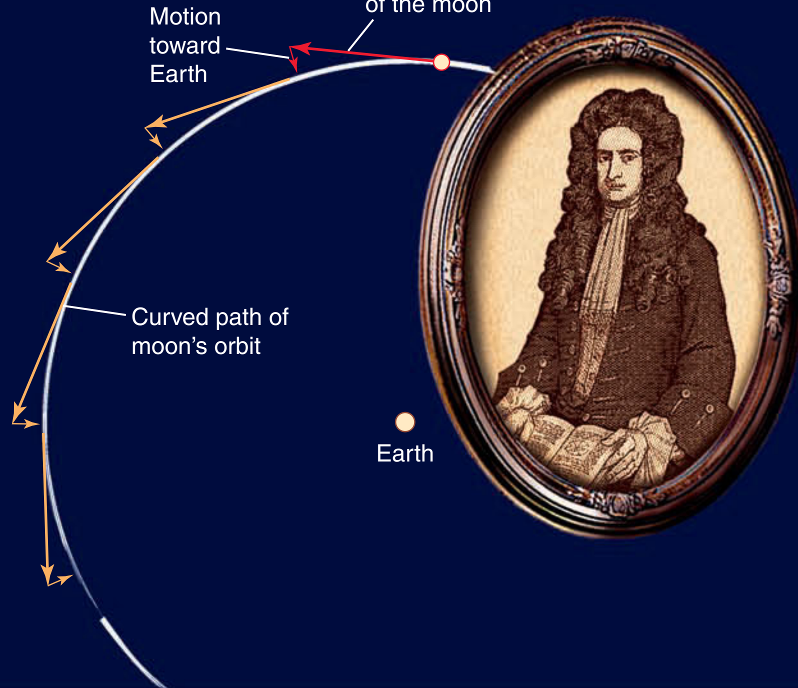
At a distance of 42,250 km (26,260 miles) from Earth's center, a satellite orbits with a period of 24 hours.

The satellite orbits eastward, and Earth rotates eastward under the moving satellite.

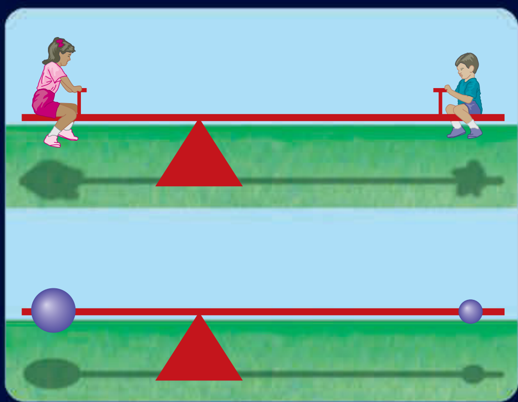
The satellite remains fixed above a spot on Earth's equator.

Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Geosynchronous Orbit" and place your own satellite into geosynchronous orbit.

1b According to Newton's first law of motion, the moon should follow a straight line and leave Earth forever. Because it follows a curve, Newton knew that some force must continuously accelerate it toward Earth — gravity. Each second the moon moves 1020 m (3350 ft) eastward and falls about 1.6 mm (about 1/16 inch) toward Earth. The combination of these motions produces the moon's curved orbit. The moon is falling.

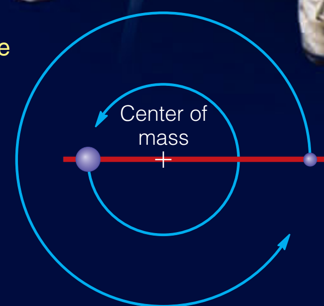


1c Astronauts in orbit around Earth feel weightless, but they are not "beyond Earth's gravity," to use a term from old science fiction movies. Like the moon, the astronauts are accelerated toward Earth by Earth's gravity, but they travel fast enough along their orbits that they continually "miss the Earth." They are literally falling around Earth. Inside or outside a spacecraft, astronauts feel weightless because they and their spacecraft are falling at the same rate. Rather than saying they are weightless, you should more accurately say they are in free fall.



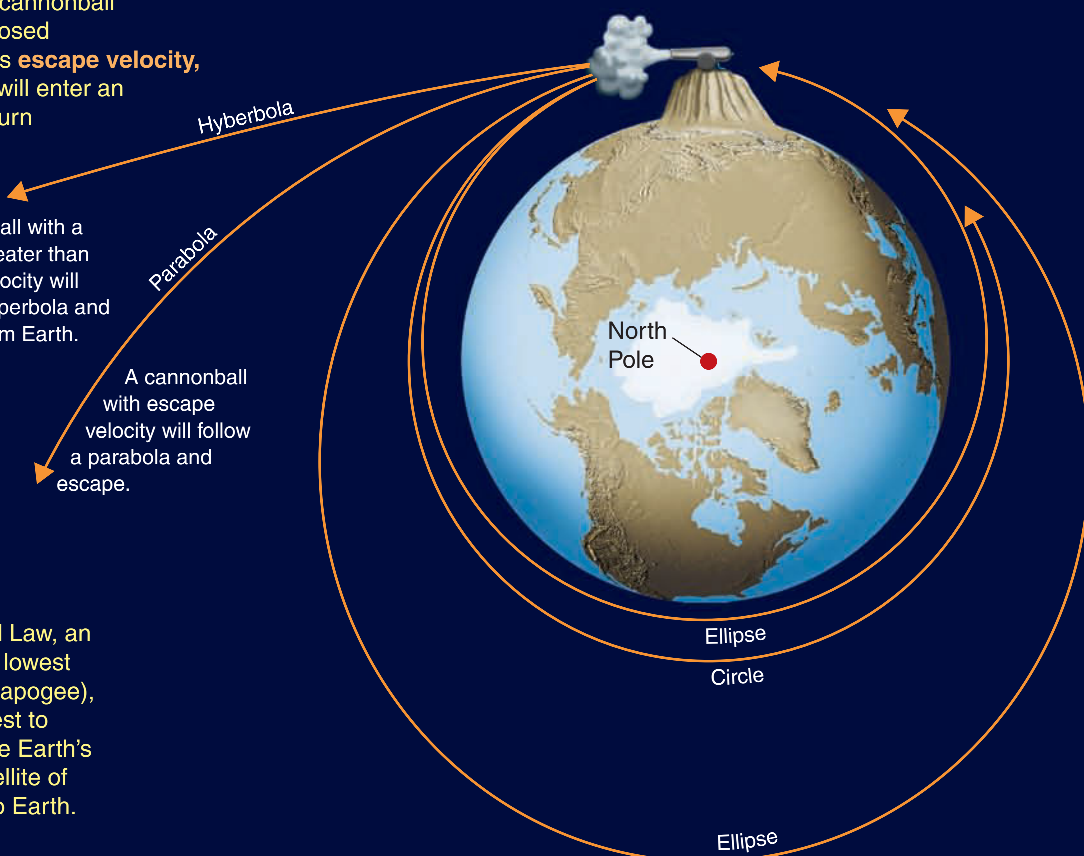
2 To be precise you should not say that an object orbits Earth. Rather the two objects orbit each other. Gravitation is mutual, and if Earth pulls on the moon, the moon pulls on Earth. The two bodies revolve around their common **center of mass,** the balance point of the system.

2a Two bodies of different mass balance at the center of mass, which is located closer to the more massive object. As the two objects orbit each other, they revolve around their common center of mass as shown at right. The center of mass of the Earth–moon system lies only 4708 km (2926 miles) from the center of Earth — inside the Earth. As the moon orbits the center of mass on one side, the Earth swings around the center of mass on the opposite side.



3 **Closed orbits** return the orbiting object to its starting point. The moon and artificial satellites orbit Earth in closed orbits. Below, the cannonball could follow an elliptical or a circular closed orbit. If the cannonball travels as fast as **escape velocity,** the velocity needed to leave a body, it will enter an open orbit. An **open orbit** does not return the cannonball to Earth. It will escape.

CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGE**NOW**™ to see Active Figure "Center of Mass." Change the mass ratio to move the center of mass.



3a As described by Kepler's Second Law, an object in an elliptical orbit has its lowest velocity when it is farthest from Earth (apogee), and its highest velocity when it is closest to Earth (perigee). Perigee must be above Earth's atmosphere, or friction will rob the satellite of energy and it will eventually fall back to Earth.

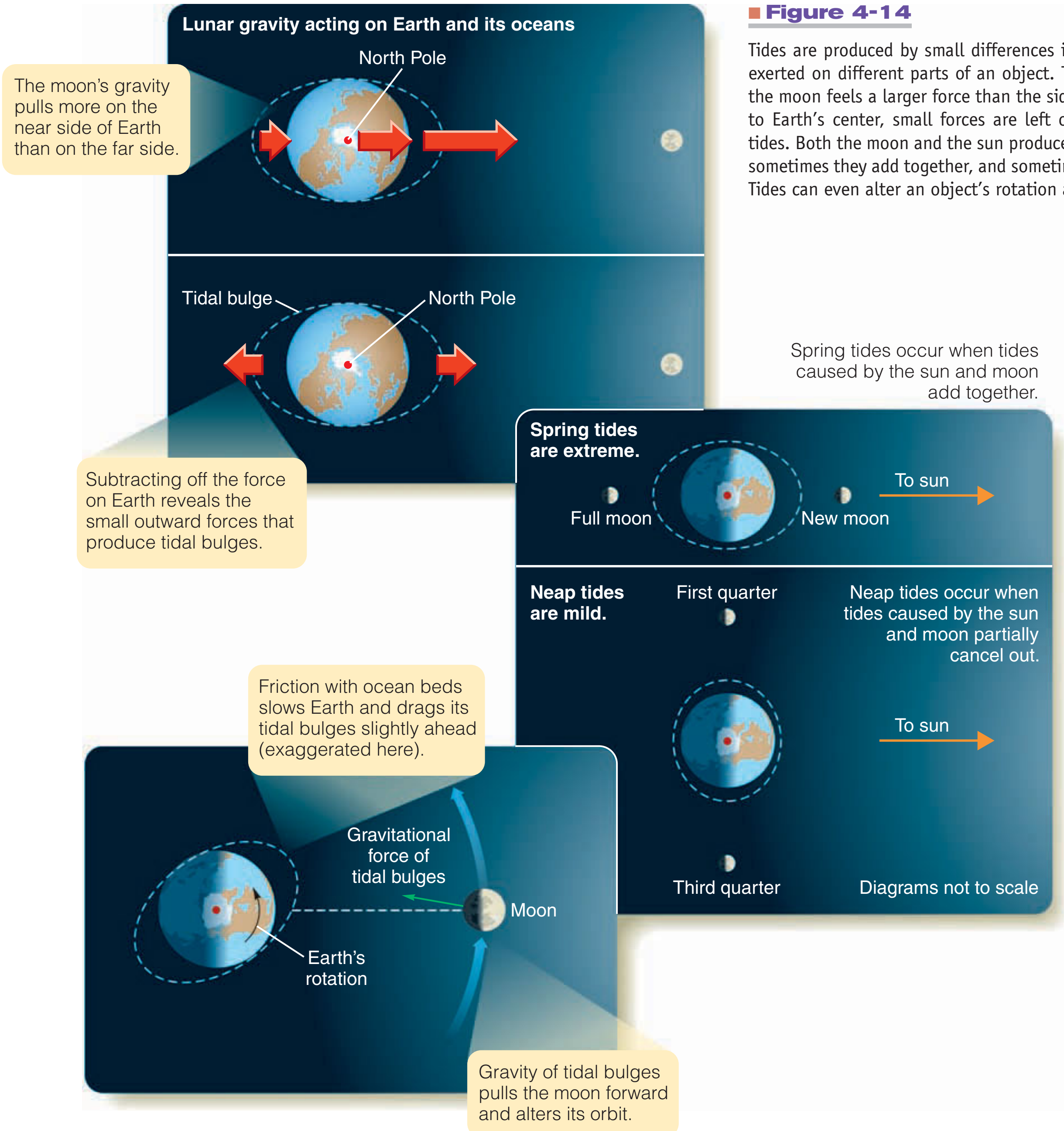


Figure 4-14

Tides are produced by small differences in the gravitational force exerted on different parts of an object. The side of Earth nearest the moon feels a larger force than the side farthest away. Relative to Earth's center, small forces are left over, and they cause the tides. Both the moon and the sun produce tides on Earth's oceans; sometimes they add together, and sometimes they partially cancel. Tides can even alter an object's rotation and orbital motion.

If the moon's gravity only affected water, then there would be only one tidal bulge, the one facing the moon. As you know, the moon's gravity acts on the rock of Earth as well as on water, and that produces the tidal bulge on the far side of Earth. The rocky bulk of Earth responds to these tidal forces, and although you don't notice, Earth flexes, with the mountains and plains rising and falling by a few centimeters in response to the moon's gravitational pull. The moon has no special affinity for water, and, because your body is so much smaller than Earth, any tides the moon raises in your body are immeasurably small. Ocean tides are large because oceans are large.

You can see dramatic evidence of tides if you watch the ocean shore for a few hours. Though Earth rotates on its axis, the tidal bulges remain fixed with respect to the moon. As the turning Earth carries you and your beach into a tidal bulge, the ocean water deepens, and the tide crawls up the sand. The tide does not so much "come in" as you are carried into the tidal bulge. Later, when Earth's rotation carries you out of the bulge, the ocean becomes shallower, and the tide falls. Because there are two bulges on opposite sides of Earth, the tides rise and fall twice a day on an ideal coast.

In reality, the tidal cycle at any given location can be quite complex because of the latitude of the site, shape of the shore, winds, and so on. Tides in the Bay of Fundy (New Brunswick, Canada), for example, occur twice a day and can exceed 40 feet. In contrast, the northern coast of the Gulf of Mexico has only one tidal cycle a day of roughly 1 foot.

Gravity is universal, so the sun also produces tides on Earth. The sun is roughly 27 million times more massive than the moon, but it lies almost 400 times farther from Earth. Consequently, tides on Earth caused by the sun are less than half as high as those caused by the moon. Twice a month, at new moon and at full moon, the moon and sun produce tidal bulges that add together and produce extreme tidal changes; high tide is very high, and low tide is very low. Such tides are called **spring tides.** Here the word *spring* does not refer to the season of the year but to the rapid welling up of water. At first- and third-quarter moons, the sun and moon pull at right angles to each other, and the sun's tides cancel out some of the moon's tides. These less-extreme tides are called **neap tides,** and they do not rise very high or fall very low. The word *neap* comes from an obscure Old English word, *nep,* that seems to have meant "lacking power to advance." Spring tides and neap tides are illustrated in Figure 4-14.

Galileo tried to understand tides, but it was not until Newton described gravity that astronomers could analyze tidal forces and recognize their surprising effects. For example, the friction of the tidal bulges with the ocean beds slows Earth's rotation and makes the length of a day grow by 0.0023 seconds per century. Fossils of ancient tide markings confirm that only 900 million years ago Earth's day was 18 hours long. Tidal forces can also affect orbital motion. Earth rotates eastward, and friction with the ocean beds drags the tidal bulges slightly eastward out of a direct Earth–moon line. These tidal bulges are massive, and their gravitational field pulls the moon forward in its orbit, as shown at the bottom of Figure 4-14. As a result, the moon's orbit is growing larger by about 3.8 cm a year, an effect that astronomers can measure by bouncing laser beams off reflectors left on the lunar surface by the Apollo astronauts.

Earth's gravitation exerts tidal forces on the moon, and although there are no bodies of water on the moon, friction within the flexing rock has slowed the moon's rotation to the point that it now keeps the same face toward Earth.

Newton's gravitation is much more than just the force that makes apples fall. In later chapters, you will see how tides can pull gas away from stars, rip galaxies apart, and melt the interiors of small moons orbiting near massive planets. Tidal forces produce some of the most surprising and dramatic processes in the universe.

The Newtonian Universe

Newton's insight gave the world a new conception of nature. His laws of motion and gravity were general laws that described the motions of all bodies under the action of external forces. In addition, the laws were productive because they made possible specific calculations that could be tested by observation. For example, Newton's laws of motion can be used to derive Kepler's third law from the law of gravity.

How Do We Know? 4-4

Testing a Theory by Prediction

How are a theory's predictions useful in science? Scientific theories look back into the past and explain phenomena previously observed. But theories also look forward in that they make predictions about what you should find as you explore further. In this way, Newton's laws explained past observations, but they also allowed astronomers to predict the motions of comets and eventually understand their origin.

Scientific predictions are important in two ways. First, if a theory's prediction is confirmed, scientists gain confidence that the theory is a true description of nature. But second, predictions can reveal unexplored avenues of knowledge.

Particle physics is a field in which predictions have played a key role in directing research. In the early 1970s physicists proposed a theory of the fundamental forces and particles in atoms called the Standard Model. This theory was supported by what scientists had already observed in experiments, but it also predicted the existence of particles that hadn't yet been observed. In the interest of testing the theory, scientists focused their efforts on building more and more powerful particle accelerators in the hopes of detecting the predicted particles.

A number of these particles have since been discovered, and they do match the characteristics predicted by the Standard Model, further confirming the theory. One predicted particle, the Higgs boson, has not yet been found, as of this writing, but an even larger accelerator soon to begin operation may allow its detection. Will the Higgs boson be found? If it exists, the Standard Model is confirmed, but if it can't be found, the prediction will be a warning to physicists that nature is even more interesting than the Standard Model supposes.

As you read about any scientific theory, think about both what it can explain and what it can predict.

Physicists build huge accelerators to search for the subatomic particles predicted by their theories. (Brookhaven National Laboratory)

Newton's discoveries remade astronomy into an analytical science in which astronomers could measure the positions and motions of celestial bodies, calculate the gravitational forces acting on them, and predict their future motion (■ How Do We Know? 4-4).

Were you to trace the history of astronomy after Newton, you would find scientists predicting the motion of comets, the gravitational interaction of the planets, the orbits of double stars, and so on. Astronomers built on the discoveries of Newton, just as he had built on the discoveries of Copernicus, Tycho, Kepler, and Galileo. It is the nature of science to build on the discoveries of the past, and Newton was thinking of that when he wrote, "If I have seen farther than other men, it is because I stood upon the shoulders of giants."

What Are We? Participants

The scientific revolution began when Copernicus made humanity part of the universe. Before Copernicus, people thought of Earth as a special place different from any of the objects in the sky; but, in trying to explain the motions in the sky, Copernicus made Earth one of the planets. Galileo and those who brought him to trial understood the significance of making Earth just a planet. It made humanity part of nature, part of the universe.

Kepler showed that the planets move, not at the whim of ancient gods, but according to simple rules, and Newton found simple rules that account for the fall of an apple, orbital motion, and the ocean tides. We are not in a special place ruled by mysterious planetary forces. Earth, the sun, and all of humanity are part of a universe whose motions can be described by a few fundamental laws. If simple laws describe the motions of the planets, then the universe is not ruled by mysterious influences as in astrology or by the whim of the gods atop Mount Olympus. And if the universe can be described by simple rules, then it is open to scientific study.

Before Copernicus, people felt they were special because they thought they were at the center of the universe. Copernicus, Kepler, and Newton showed that we are not at the center but are part of an elegant and complex universe. Astronomy tells us that we are special because we can study the universe and eventually understand what we are. But it also tells us that we are not just observers; we are participants.

Study and Review

Summary

- Ancient philosophers accepted as a first principle that the heavens were perfect, so philosophers such as Plato argued that, because the sphere was the only perfect geometrical form and carried a point on its surface around in a circle, the heavens must move in **uniform circular motion (p. 44).**
- They also accepted that Earth was the unmoving center of all motion, and that **geocentric (p. 44)** universe became part of the teachings of the great philosopher Aristotle, who argued that the sun, moon, and stars were carried around Earth on rotating crystalline spheres.
- The lack of any **parallax (p. 44)** in the positions of the stars gave astronomers confidence that Earth could not move.
- About AD 140, Ptolemy gave mathematical form to Aristotle's model in the *Almagest.* Ptolemy preserved the principles of geocentrism and uniform circular motion, but he added **epicycles (p. 45), deferents (p. 45),** and **equants (p. 45)** to better predict the motions of the planets. To account for **retrograde motion (p. 44),** his epicycles had to be quite large. Even so, his model was not very accurate in predicting the positions of the planets.
- The problem of the place of Earth was solved when Copernicus devised a model that was a **heliocentric universe (p. 46).** He preserved the principle of uniform circular motion, but he put the sun at the center and argued that Earth rotates on its axis and circles the sun once a year. His theory was controversial in part because it contradicted Church teaching.
- Copernicus published his theory in his book *De Revolutionibus* in 1543, the same year he died.
- In Copernicus's model, retrograde motion was explained without epicycles, but because he kept uniform circular motion, he had to include small epicycles, and his model did not predict the motions of the planets well.
- One reason the Copernican model won gradual acceptance was that it was more elegant. Venus and Mercury were treated the same as all the other planets, and the velocity of each planet was related to its distance from the sun. The shift from the geocentric **paradigm (p. 49)** to the heliocentric paradigm is an example of a scientific revolution.
- The problem of planetary motion was finally solved through the work of two astronomers, Tycho Brahe and Johannes Kepler.
- Tycho developed his own model in which the sun and moon circled Earth and the planets circled the sun. His great contribution was to compile detailed observations over a period of 20 years, observations that were later used by Kepler.
- Johannes Kepler inherited Tycho's books of observations in 1601 and used them to uncover three laws of planetary motion. The first law says that the planets follow **ellipses (p. 53)** with the sun at one focus. According to the second law, planets move faster when nearer the sun and slower when farther away. The third law says that a planet's orbital period squared is proportional to the **semimajor axis (p. 53)** of its orbit cubed.
- The **eccentricity (p. 53)** of an ellipse equals zero for a circle and grows closer and closer to one as the ellipse becomes more and more elongated.

- A **hypothesis (p. 53)** is a statement about nature that needs further testing, but a **theory (p. 53)** is usually a description of nature that has been tested. Some theories are very well understood and widely accepted. A **natural law (p. 53)** is a fundamental principle in which scientists have great confidence.
- Kepler's final book, *The Rudolphine Tables* (1627), combined heliocentrism with elliptical orbits and predicted the positions of the planets well.
- Galileo used the newly invented telescope to observe the heavens, and he recognized the significance of what he saw there. His discoveries of the phases of Venus, the satellites of Jupiter, the mountains of the moon, and other phenomena helped undermine the Ptolemaic universe.
- Galileo based his analysis on observational evidence rather than on first principles or on scripture. In 1633, he was condemned before the Inquisition for refusing to halt his defense of Copernicanism.
- Newton used the work of Kepler and Galileo to discover three laws of motion and the law of gravity. These laws made it possible to understand such phenomena as orbital motion and the tides.
- Newton showed that gravity was mutual and universal. It depends on the **mass (p. 60)** of the bodies and the distance between them according to the **inverse square relation (p. 60).**
- Newton used the image of a cannon on a mountaintop to explain that an object in orbit is falling toward Earth's center and simultaneously moving fast enough to continually miss hitting Earth's surface. To maintain a circular orbit, the object must have **circular velocity (p. 62).** Circular and elliptical orbits are **closed orbits (p. 63)**, but if the object's velocity equals or exceeds **escape velocity (p. 63)** it will follow an **open orbit (p. 63)** and never return.
- **Geosynchronous satellites (p. 62)** orbit far enough from Earth that their orbital period is 24 hours, and they remain above a single spot on Earth as Earth turns.
- Two objects that orbit each other actually orbit their common **center of mass (p. 63).**
- Newton's laws gave scientists a unified way to think about nature—cause and effect. Every effect has a cause, and science is the search for those causes.
- Newton's laws also explain that tides are caused by small differences in the moon's gravity acting on different parts of a body. Ocean tides occur because the moon's gravity pulls more strongly on the near side of Earth than on the center. A tidal bulge occurs on the far side of Earth because the moon's gravity is slightly weaker there than on the center of Earth.
- Tides produced by the moon combine with tides produced by the sun to cause extreme tides, called **spring tides (p. 65)**, at new and full moons. The moon and sun work against each other to produce less extreme tides, called **neap tides (p. 65)**, at quarter moons.
- Friction from tides can slow the rotation of a rotating world, and the gravitational pull of tidal bulges can make orbits change slowly.
- The 99 years from the death of Copernicus to the birth of Newton marked the beginning of modern science. From that time on, science depended on evidence to test theories and relied on the analytic methods first demonstrated by Kepler and Newton.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why did Greek astronomers conclude that the heavens were made up of perfect crystalline spheres moving at constant speeds?
2. Why did classical astronomers conclude that Earth had to be motionless?
3. How did the Ptolemaic model explain retrograde motion?
4. In what ways were the models of Ptolemy and Copernicus similar?
5. Why did the Copernican hypothesis win gradual acceptance?
6. Why is it difficult for scientists to replace an old paradigm with a new paradigm?
7. Why did Tycho Brahe expect the new star of 1572 to show parallax? Why was the lack of parallax evidence against the Ptolemaic model?
8. How was Tycho's model of the universe similar to the Ptolemaic model? How did it resemble the Copernican model?
9. Explain how Kepler's laws contradict uniform circular motion.
10. What is the difference between a hypothesis, a theory, and a law?
11. How did *The Alfonsine Tables, The Prutenic Tables,* and *The Rudolphine Tables* differ?
12. Review Galileo's telescopic discoveries and explain why they supported the Copernican model and contradicted the Ptolemaic model.
13. Galileo was condemned by the Inquisition, but Kepler, also a Copernican, was not. Why not?
14. How do Newton's laws lead you to conclude that gravitation has to be universal?
15. Explain why you might describe the orbital motion of the moon with the statement, "The moon is falling."
16. **How Do We Know?** Why is it fair to say that a paradigm affects the questions you ask and the answers you find acceptable?
17. **How Do We Know?** How would you respond to someone who said, "Oh, that's only a theory."
18. **How Do We Know?** Why is consideration of cause and effect necessary if you expect to learn about nature using the scientific method?
19. **How Do We Know?** *The Rudolphine Tables* could predict the position of the planets on future dates. Why was the accuracy of those predictions confirmation of Kepler's theories of orbital motion?

Discussion Questions

1. Science historian Thomas Kuhn has said that *De Revolutionibus* was a revolution-making book but not a revolutionary book. How was it classical and conservative?
2. Why might Tycho Brahe have hesitated to hire Kepler? Why do you suppose he finally decided to appoint Kepler his scientific heir?
3. How does the modern controversy over creationism and evolution reflect two ways of knowing about the physical world?

Problems

1. If you lived on Mars, which planets would describe retrograde loops? Which would never be visible as crescent phases?
2. Galileo's telescope showed him that Venus has a large angular diameter (61 seconds of arc) when it is a crescent and a small angular diameter (10 seconds of arc) when it is nearly full. Use the small-angle formula to find the ratio of its maximum distance to its minimum distance. Is this ratio compatible with the Ptolemaic universe shown on page 45?
3. Galileo's telescopes were not of high quality by modern standards. He was able to see the moons of Jupiter, but he never reported seeing features on Mars. Use the small-angle formula to find the maximum angular diameter of Mars when it is closest to Earth. How does that compare with the maximum diameter of Jupiter?
4. If a planet had an average distance from the sun of 10 AU, what would its orbital period be?
5. If a space probe were sent into an orbit around the sun that brought it as close as 0.5 AU to the sun and as far away as 5.5 AU, what would its orbital period be?
6. Neptune orbits the sun with a period of 164.8 years. What is its average distance from the sun?

7. Venus's average distance from the sun is 0.72 AU and Saturn's is 9.54 AU. Calculate the circular orbital velocities of Venus and Saturn around the sun. (*Hints:* The mass of the sun is 2.0×10^{30} kg. An AU is 1.50×10^{11} m.)
8. The circular velocity of Earth around the sun is about 30 km/s. Are the arrows for Venus and Saturn correct in Figure 4-3? (*Hint:* See Problem 7.)
9. What is the orbital velocity of an Earth satellite 42,250 km from Earth? How long does it take to circle its orbit once?

Learning to Look

1. What three astronomical objects are represented here? What are the two rings?

2. Why can the object shown at the right be bolted in place and used 24 hours a day without adjustment?

Larry Mulvehill/The Image Works

3. Why is it a little bit misleading to say that this astronaut is weightless?

NASA/JSC

5 Light and Telescopes

At night, inside the dome of a major observatory, only the hum of motors breaks the silence as the huge telescope peers out at the sky and gathers starlight.
(Gemini Observatory/AURA)

Guidepost

In the early chapters of this book, you looked at the sky the way ancient astronomers did, with the unaided eye. In the last chapter, you got a glimpse through Galileo's telescope, and it revealed astonishing things about the moon, Jupiter, and Venus. Now it is time to examine the instruments of the modern astronomer.

You can begin by studying telescopes that gather and focus visible light, so you need to be sure you understand what light is and how it behaves. But you will quickly meet telescopes that gather invisible forms of radiation such as X-rays and radio waves. Astronomers cannot overlook any clues, so they must use all forms of light. This chapter will help you answer five essential questions:

What is light?

How do telescopes work, and how are they limited?

What kind of instruments do astronomers use to record and analyze light?

Why do astronomers use radio telescopes?

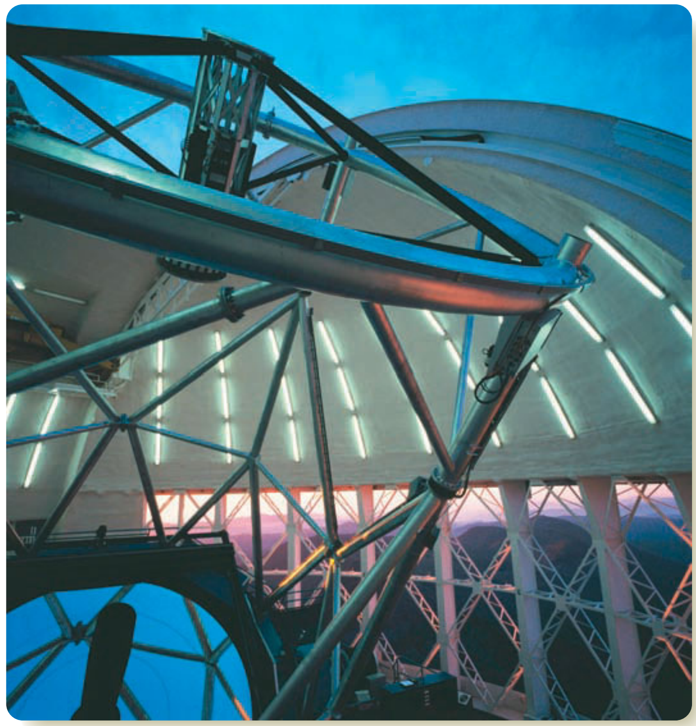
Why must some telescopes go into orbit?

Astronomy is almost entirely an observational science. Astronomers cannot visit distant galaxies and far-off worlds, so they must observe using astronomical telescopes. Fifteen chapters remain in your exploration, and every one will discuss information gathered by telescopes.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

The strongest thing that's given us to see with's
A telescope. Someone in every town
Seems to me owes it to the town to keep one.

ROBERT FROST, *"THE STAR-SPLITTER"*



■ **Figure 5-1**

Astronomical telescopes are often very large to gather large amounts of starlight. The Southern Gemini telescope stands over 19 m (60 ft) high when pointed straight up, and its main mirror, shown at lower left, is 8.1 m (26.5 ft) in diameter—larger than some classrooms. The sides of the telescope dome open to allow quick equalization of inside and outside temperatures at sunset. (Gemini Observatory/AURA)

STARLIGHT IS GOING to waste. Every night it falls on trees, oceans, and parking lots, and it is all wasted. To an astronomer, nothing is so precious as starlight. It is the only link to the sky, so the astronomer's quest is to gather as much of it as possible and extract from it the secrets of the stars.

The telescope is the symbol of the astronomer because it gathers and concentrates light for analysis. Most of the interesting objects in the sky are faint, so astronomers are driven to build huge telescopes to gather the maximum amount of light (■ Figure 5-1). Some telescopes collect radio waves or X-rays and some go into space, but they all gather information about our universe.

In the quote that opens this chapter, Robert Frost suggests that someone in every town should own a telescope. Astronomy is more than technology and scientific analysis. It tells us what we are, and every town should have a telescope to keep us looking upward.

5-1 Radiation: Information from Space

JUST AS A book on baking bread might begin with a discussion of flour, this chapter on telescopes begins with a discussion of light—not just visible light, but the entire range of radiation from the sky.

Light as a Wave and a Particle

When you admire the colors of a rainbow, you are seeing light behave as a wave. But when you use a digital camera to take a picture of the same rainbow, the light hitting the camera's detector acts like a particle. Light is peculiar in that it is both wave and particle, and how it acts depends on how you observe it.

Light is a form of **electromagnetic radiation** and carries energy through space as electric and magnetic waves**.** We use the word *light* to refer to electromagnetic radiation that we can see, but visible light is only a small part of a range that also includes x-rays and radio waves. Electromagnetic radiation travels through space at 300,000 km/s (186,000 mi/s). This is commonly referred to as the speed of light, *c,* but it is in fact the speed of all electromagnetic radiation.

Some people flinch at the word *radiation,* but that reflects a **Common Misconception**. *Radiation* refers to anything that radiates from a source. High-energy particles emitted from radioactive atoms are called radiation, and you have learned to be a little bit concerned when you see this word. But light, like all electromagnetic radiation, spreads outward from a source, so you can correctly refer to light as a form of radiation.

Electromagnetic radiation can act as a wave phenomenon—that is, it is associated with a periodically repeating disturbance, a wave. You are familiar with waves in water: If you disturb a pool of water, waves spread across the surface. Imagine that you use a meter stick to measure the distance between the successive peaks of a wave. This distance is the **wavelength,** usually represented by the Greek letter lambda (λ).

Sound is also a wave, a mechanical disturbance that travels through air from source to ear. Sound requires a medium; so, on the moon, where there is no air, there can be no sound. In contrast, light is made up of electric and magnetic fields that can travel through empty space. Unlike sound, light does not require a medium, and so it can travel through a perfect vacuum. There is no sound on the moon, but there is plenty of sunlight.

Although electromagnetic radiation can behave as a wave, it can also behave as a flood of particles. A particle of electromagnetic radiation is called a **photon,** and you can think of a photon as a bundle of waves.

The amount of energy a photon carries depends inversely on its wavelength. That is, shorter-wavelength photons carry more energy, and longer-wavelength photons carry less. A simple formula expresses this relationship:

$$E = \frac{hc}{\lambda}$$

Here h is Planck's constant (6.6262×10^{-34} joule s), c is the speed of light (3×10^8 m/s), and λ is the wavelength in meters. This book will not use this formula for calculations; the important point is the inverse relationship between the energy E and the wavelength λ. As λ gets smaller, E gets larger. A photon of long wavelength carries a very small amount of energy, but a photon with a very short wavelength can carry much more energy.

The Electromagnetic Spectrum

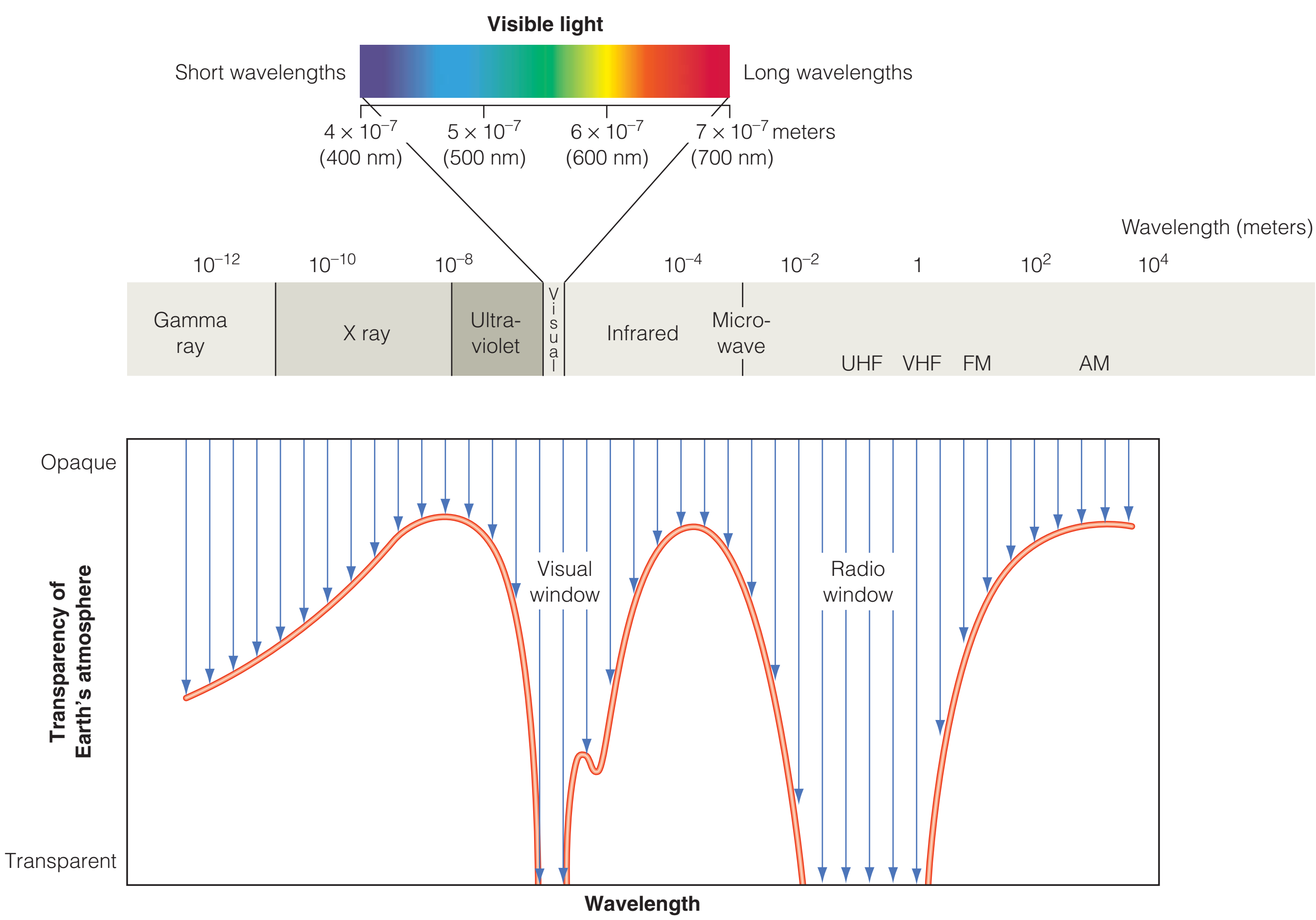
A **spectrum** is an array of electromagnetic radiation displayed in order of wavelength. You are most familiar with the spectrum of visible light, which you see in rainbows. The colors of the visible spectrum differ in wavelength, with red having the longest wavelength and violet the shortest. The visible spectrum is shown at the top of ■ Figure 5-2.

The average wavelength of visible light is about 0.0005 mm. You could put 50 light waves end to end across the thickness of a sheet of household plastic wrap. It is too awkward to measure such short distances in millimeters, so scientists measure the wavelength of light using the **nanometer (nm),** one-billionth of a meter (10^{-9} m). Another unit that astronomers commonly use is called the **angstrom (Å)** (named after the Swedish astronomer Anders Jonas Ångström). One angstrom is 10^{-10} m, one tenth of a nanometer. The wavelength of visible light ranges from about 400 to 700 nm. Just as you sense the wavelength of sound as pitch, you sense the wavelength of light as color. Light near the short-wavelength end of the visible spectrum (400 nm) looks violet to your eyes, and light near the long-wavelength end (700 nm) looks red.

Figure 5-2 shows that the visible spectrum makes up only a small part of the entire electromagnetic spectrum. Beyond the red

■ **Figure 5-2**

The spectrum of visible light, extending from red to violet, is only part of the electromagnetic spectrum. Most radiation is absorbed in Earth's atmosphere, and only radiation in the visual window and the radio window can reach Earth's surface.



end of the visible spectrum lies **infrared radiation,** where wavelengths range from 700 nm to about 1 mm. Your eyes are not sensitive to this radiation, but your skin senses it as heat. For example, a "heat lamp" warms you by giving off infrared radiation.

Beyond the infrared part of the electromagnetic spectrum lie radio waves. The radio radiation used for AM radio transmissions has wavelengths of a few kilometers down to a few hundred meters, while FM, television, military, government, cell phone, and ham radio transmissions have wavelengths that range down to a few centimeters. Microwave transmission, used for radar and some long-distance telephone communications, for instance, has wavelengths from a few centimeters down to about 1 mm.

You may not think of radio waves in terms of wavelength because radio dials are marked in units of **frequency,** the number of waves that pass a stationary point in 1 second. Wavelength and frequency are related; to calculate the wavelength of a radio wave, divide the speed of light by the frequency. When you tune in your favorite FM station at 89.5 MHz (million cycles per second), you are adjusting your radio to detect radio photons with a wavelength of 3.35 m.

The boundaries between the wavelength ranges are not sharp. Long-wavelength infrared radiation blends smoothly into the shortest microwave radio waves. Similarly, there is no natural division between the short-wavelength infrared and the long-wavelength part of the visible spectrum.

Look at the other end of the electromagnetic spectrum in Figure 5-2 and notice that electromagnetic waves shorter than violet are called **ultraviolet.** Electromagnetic waves that are even shorter are called X-rays, and the shortest are gamma rays. Again, the boundaries between these wavelength ranges are not clearly defined.

Recall the formula for the energy of a photon. Extremely short-wavelength photons such as X-rays and gamma rays have high energies and can be dangerous. Even ultraviolet photons have enough energy to do harm. Small doses of ultraviolet produce a suntan, and larger doses cause sunburn and skin cancers. Contrast this to the lower-energy infrared photons. Individually they have too little energy to affect skin pigment, a fact that explains why you can't get a tan from a heat lamp. Only by concentrating many low-energy photons in a small area, as in a microwave oven, can you transfer significant amounts of energy.

Astronomers are interested in electromagnetic radiation because it carries clues to the nature of stars, planets, and other celestial objects. Earth's atmosphere is opaque to most electromagnetic radiation, as shown by the graph at the bottom of Figure 5-2. Gamma rays, X-rays, and some radio waves are absorbed high in Earth's atmosphere, and a layer of ozone (O_3) at an altitude of about 30 km absorbs ultraviolet radiation. Water vapor in the lower atmosphere absorbs the longer-wavelength infrared radiation. Only visible light, some shorter-wavelength infrared, and some radio waves reach Earth's surface through two wavelength regions called **atmospheric windows.** Obviously, if you wish to study the sky from Earth's surface, you must look out through one of these windows.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "The Electromagnetic Spectrum."

◄ SCIENTIFIC ARGUMENT ►

What would you see if your eyes were sensitive only to X-rays?

As you build this scientific argument, you must imagine a totally new situation. That is sometimes a powerful tool in the critical analysis of an idea. In this case, you might at first expect to be able to see through walls, but remember that your eyes detect only light that already exists. There are almost no X-rays bouncing around at Earth's surface, so if you had X-ray eyes, you would be in the dark and would be unable to see anything. Even when you looked up at the sky, you would see nothing, because Earth's atmosphere is not transparent to X-rays. If Superman can see through walls, it is not because his eyes can detect X-rays.

But now imagine a slightly different situation and modify your argument. **Would you be in the dark if your eyes were sensitive only to radio wavelengths?**

◄ ►

5-2 Optical Telescopes

Earth has two atmospheric windows, so there are two main types of ground-based telescopes—optical telescopes and radio telescopes. You can start with optical telescopes, which gather light and focus it into sharp images. This requires sophisticated optical and mechanical designs, and it leads astronomers to build gigantic telescopes on the tops of high mountains.

Two Kinds of Optical Telescopes

Optical telescopes can focus light into an image by using either a lens or a mirror, as shown in ■ Figure 5-3. In a **refracting telescope,** the **primary** (or **objective**) **lens** bends (refracts) the light as it passes through the glass and brings it to a focus to form a small inverted image. In a **reflecting telescope,** the **primary** (or **objective**) **mirror**—a concave piece of glass with a reflective surface—forms an image by reflecting the light. In either case, the **focal length** is the distance from the lens or mirror to the image of a distant light source such as a star. Short-focal-length lenses and mirrors must be strongly curved, and long-focal-length lenses and mirrors are less strongly curved. Grinding the proper shape on a lens or mirror is a delicate, time-consuming, and expensive process.

The image formed by the primary lens or primary mirror of a telescope is small, inverted, and difficult to view directly. Astronomers use a small lens called the **eyepiece** to magnify the image and make it convenient to view (■ Figure 5-4).

Refracting telescopes suffer from a serious optical distortion that limits their usefulness. When light is refracted through glass, shorter wavelengths bend more than longer wavelengths, so blue light, having shorter wavelengths, comes to a focus closer to the

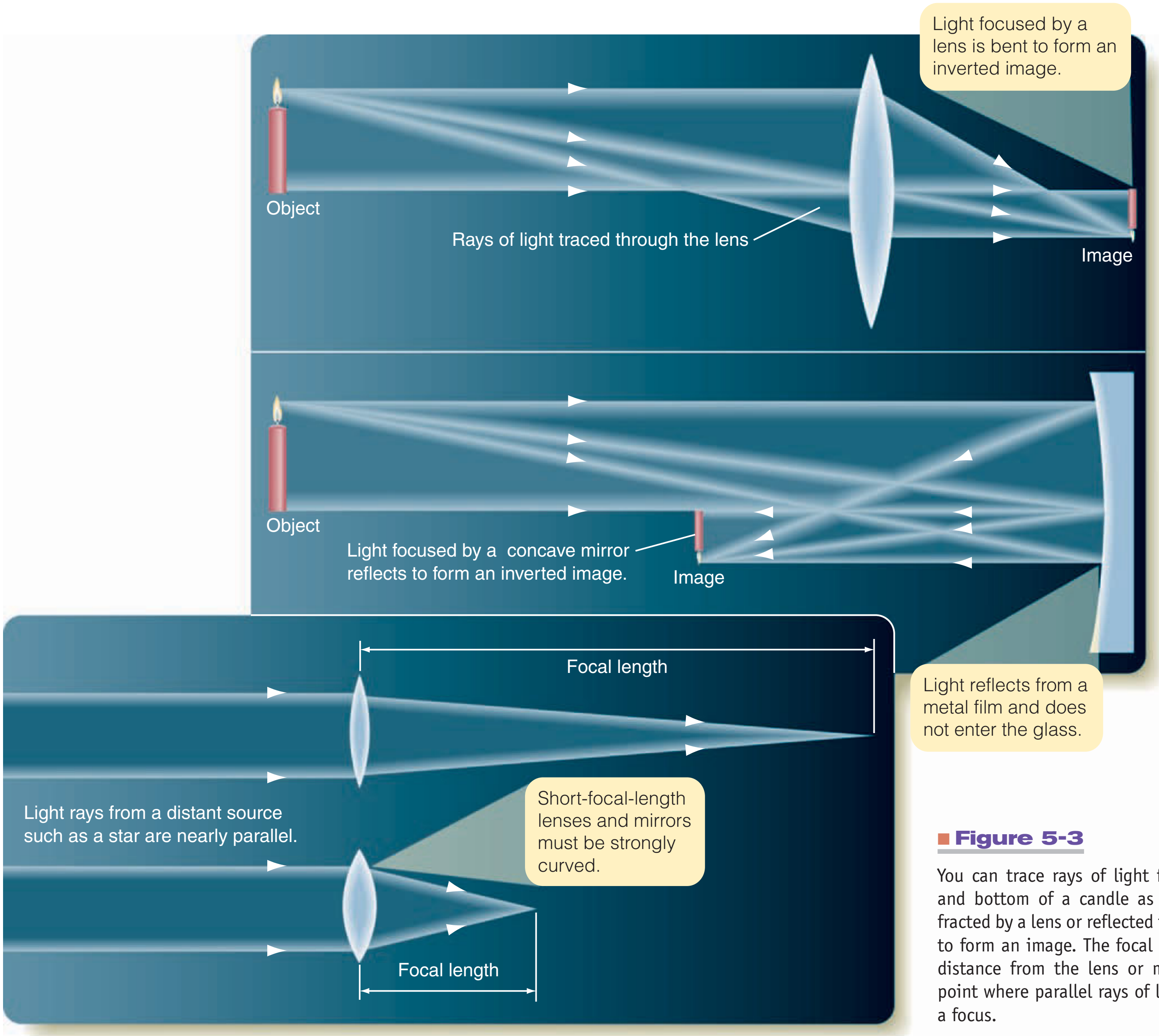


Figure 5-3

You can trace rays of light from the top and bottom of a candle as they are refracted by a lens or reflected from a mirror to form an image. The focal length is the distance from the lens or mirror to the point where parallel rays of light come to a focus.

lens than does red light (■ Figure 5-5a). If you focus the eyepiece on the blue image, the other colors are out of focus, and you see a colored blur around the image. If you focus on the red image, all the other colors blur. This color separation is called **chromatic aberration.** Telescope designers can grind a telescope lens of two components made of different kinds of glass and so bring two different wavelengths to the same focus (Figure 5-5b). This does improve the image, but these **achromatic lenses** are not totally free of chromatic aberration, because other wavelengths still blur. Telescopes made with acromatic lenses were popular until the end of the 19th century.

The primary lens of a refracting telescope is more expensive than a mirror of the same size. The lens must be achromatic, so it must be made of two different kinds of glass with four precisely ground surfaces. Also, the glass must be pure and flawless because the light passes through it. The largest refracting telescope in the world was completed in 1897 at Yerkes Observatory in Wisconsin. Its lens is 1 m (40 in.) in diameter and weighs half a ton. Larger refracting telescopes are prohibitively expensive.

The primary mirrors of reflecting telescopes are much less expensive because the light reflects off the front surface of the mirror. This means that only the front surface needs to be ground to precise shape. This front surface is coated with a highly reflective surface of an aluminum alloy, and the light reflects from this front surface without entering the glass. Consequently, the glass of the mirror need not be perfectly transparent, and the mirror can be supported over its back surface to reduce sagging. Most important, reflecting telescopes do not suffer from chromatic aberration because the light is reflected before it enters the glass. For these reasons, every large astronomical telescope built since the beginning of the 20th century has been a reflecting telescope.

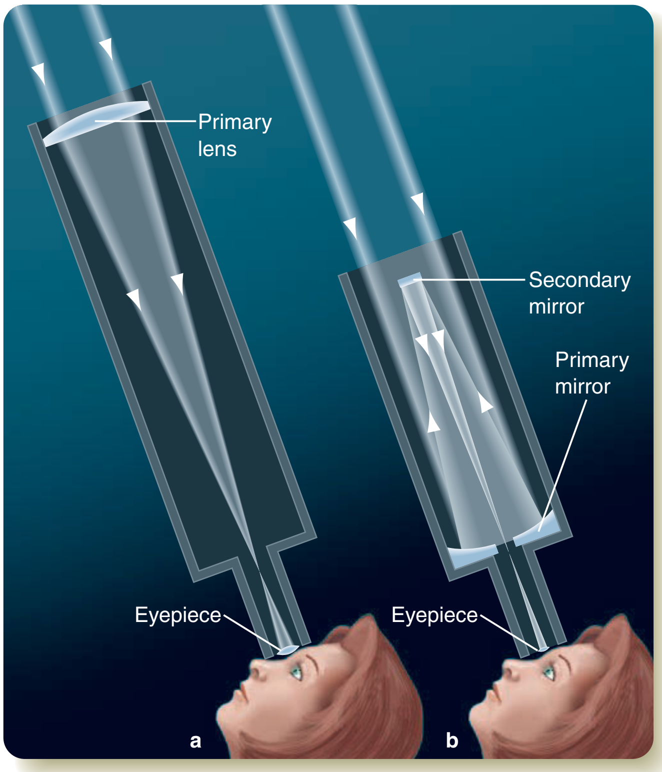


Figure 5-4

(a) A refracting telescope uses a primary lens to focus starlight into an image that is magnified by a lens called an eyepiece. The primary lens has a long focal length, and the eyepiece has a short focal length. (b) A reflecting telescope uses a primary mirror to focus the light by reflection. A small secondary mirror reflects the starlight back down through a hole in the middle of the primary mirror to the eyepiece. **Animated!**

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Lenses: Focal Length" and "Telescopes: Objective Lens and Eyepiece."

The Powers of a Telescope

Astronomers build large telescopes because a telescope can aid your eyes in three ways—the three powers of a telescope—and the two most important of these powers depend on the diameter of the telescope.

Nearly all of the interesting objects in the sky are faint sources of light, so astronomers need telescopes that can gather large amounts of light to produce bright images. **Light-gathering power** refers to the ability of a telescope to collect light. Catching light in a telescope is like catching rain in a bucket—the bigger the bucket, the more rain it catches (■ Figure 5-6). Light-gathering power is proportional to the area of the telescope objective. A lens or mirror with a large area gathers a large amount of light. Even a small increase in diameter produces a large increase in light-gathering power and allows astronomers to study much fainter objects.

The second power, **resolving power,** refers to the ability of the telescope to reveal fine detail. Because light acts as a wave, it produces a small **diffraction fringe** around every point of light in the image, and you cannot see any detail smaller than the fringe (■ Figure 5-7). Astronomers can't eliminate diffraction fringes, but the larger a telescope is in diameter, the smaller the diffraction fringes are. That means the larger the telescope, the better its resolving power.

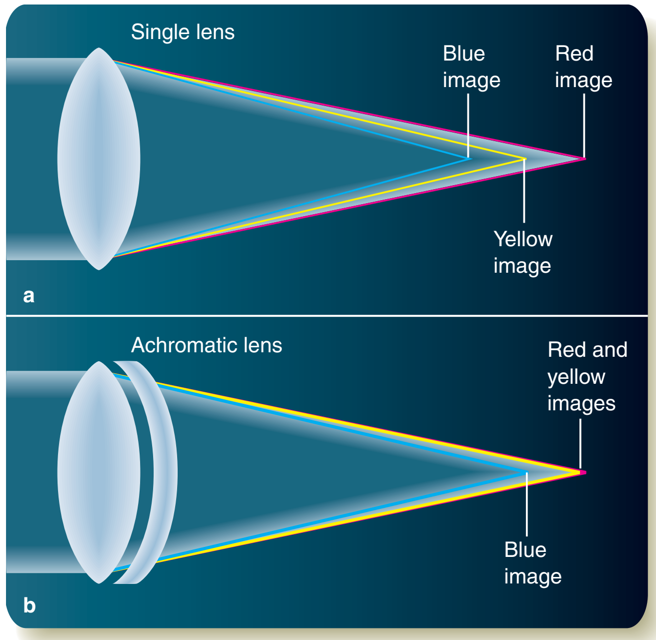


Figure 5-5

(a) A normal lens suffers from chromatic aberration because short wavelengths bend more than long wavelengths. (b) An achromatic lens, made in two pieces of two different kinds of glass, can bring any two colors to the same focus, but other colors remain slightly out of focus.

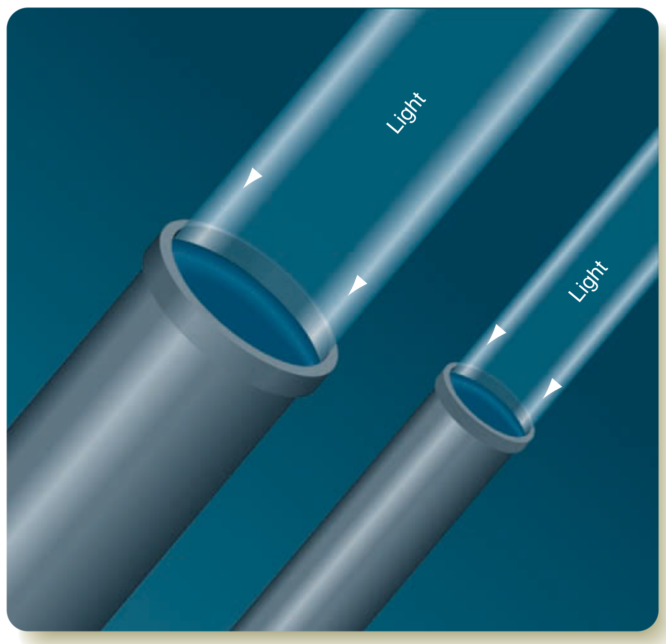
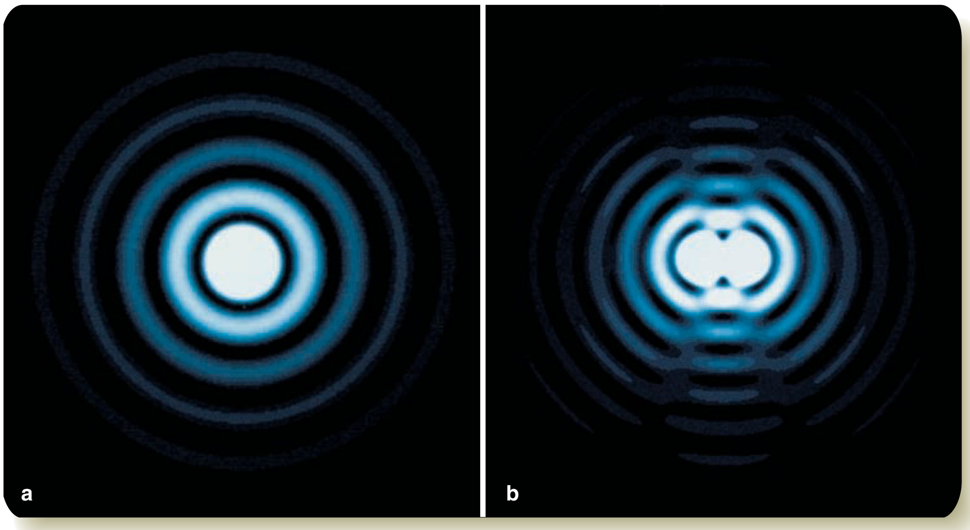


Figure 5-6

Gathering light is like catching rain in a bucket. A large-diameter telescope gathers more light and has a brighter image than a smaller telescope of the same focal length.



■ **Figure 5-7**

(a) Stars are so far away that their images are points, but the wave nature of light surrounds each star image with diffraction fringes (much magnified in this computer model). (b) Two stars close to each other have overlapping diffraction fringes and become impossible to detect separately. (Computer model by M. A. Seeds) **Animated!**

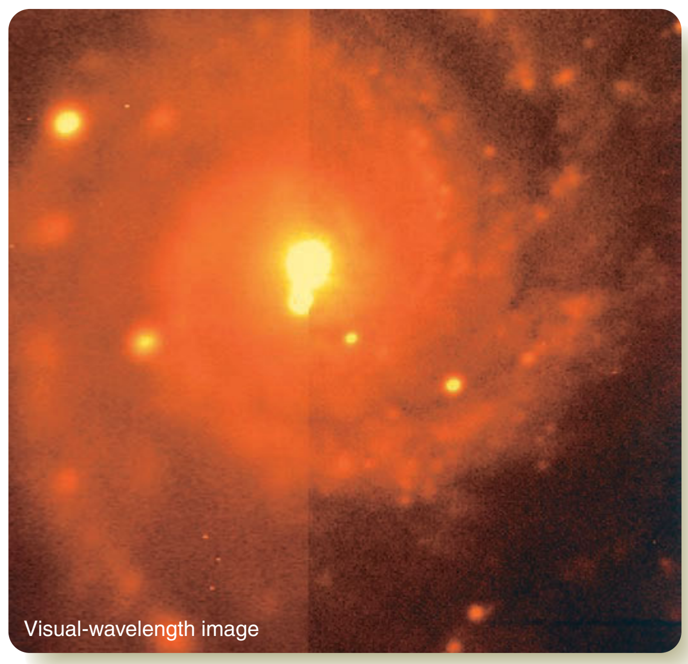
In addition to resolving power, two other factors—lens quality and atmospheric conditions—limit the detail you can see through a telescope. A telescope must contain high-quality optics to achieve its full potential resolving power. Even a large telescope reveals little detail if its optics are marred with imperfections. Also, when you look through a telescope, you are looking up through miles of turbulent air in Earth's atmosphere, which makes the image dance and blur, a condition called **seeing.** A related phenomenon is the twinkling of stars. The twinkles are caused by turbulence in Earth's atmosphere, and a star near the horizon, where you look through more air, will twinkle more than a star overhead.

On a night when the atmosphere is unsteady, the images are blurred, and the seeing is bad (■ Figure 5-8). Even under good seeing conditions, the detail visible through a large telescope is limited, not by its diffraction fringes, but by the air through which the telescope must look. A telescope performs better on a high mountaintop where the air is thin and steady, but even there Earth's atmosphere limits the detail the best telescopes can reveal to about 0.5 second of arc. You will learn later in this chapter about telescopes that orbit above Earth's atmosphere and are not limited by seeing.

Seeing and diffraction limit the amount of information in an image, and that limits the accuracy of a measurement made based on that image. Have you ever tried to magnify a newspaper photo in order to distinguish some detail? Newspaper photos are made up of tiny dots of ink, and no detail smaller than a single dot will be visible no matter how much you magnify the photo. In an astronomical image, the resolution is often set by seeing. You can't see a detail in the image that is smaller than the resolution. That's why stars look like fuzzy points of light no matter how big your telescope. All measurements have some built-in uncertainty (■ How Do We Know? 5-1), and scientists must learn to work within those limitations.

It is a **Common Misconception** that the purpose of an astronomical telescope is to magnify the image. In fact, the **magnifying power** of a telescope, its ability to make the image bigger, is actually the least significant of the three powers. Because the amount of detail you can see is limited by the seeing conditions and the resolving power, very high magnification does not necessarily show more detail. Also, you can change the magnification by changing the eyepiece, but you cannot alter the telescope's light-gathering power or resolving power without changing the diameter of the objective lens or mirror, and that would be so expensive that you might as well build a whole new telescope.

Notice that the two most important powers of the telescope, light-gathering power and resolving power, depend on the diameter of the telescope. This explains why astronomers refer to telescopes by diameter and not by magnification. Astronomers will refer to a telescope as an 8-meter telescope or a 10-meter



■ **Figure 5-8**

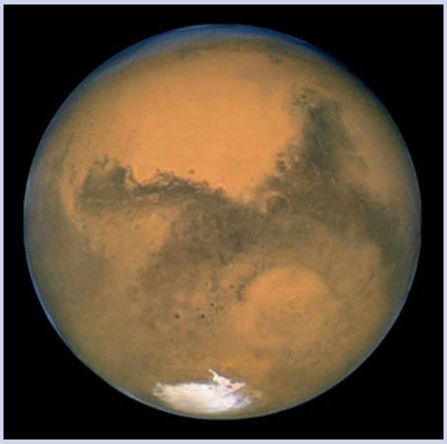
The left half of this photograph of a galaxy is from an image recorded on a night of poor seeing. Small details are blurred. The right half of the photo is from an image recorded on a night when Earth's atmosphere above the telescope was steady and the seeing was better. Much more detail is visible under good seeing conditions. (Courtesy William Keel)

Resolution and Precision

What limits the detail you can see in an image? All images have limited resolution. You see this on your computer screen because images there are made up of picture elements, pixels. If your screen has large pixels, the resolution is low, and you can't see much detail. In an astronomical image, the size of a picture element is set by seeing and by diffraction in the telescope. You can't see detail smaller than that resolution limit.

This limitation on the detail in an image is related to the limited precision of a measurement. Imagine a zoologist trying to measure the length of a live snake by holding it along a meter stick. The wriggling snake is hard to hold, so it is hard to measure accurately. Also, meter sticks are usually not marked finer than millimeters. Both factors limit the precision of the measurement. If the zoologist said her snake was 43.28932 cm long, you might be suspicious. The resolution of the measurement technique does not justify the accuracy implied by all those digits.

Whenever you make a measurement you should ask yourself how accurate that measurement can be. The accuracy of the measurement is limited by the resolution of the measurement technique, just as the amount of detail in a photograph is limited by its resolution. If you photographed a star, you would not be able to see details on its surface for the same reason the zoologist can't measure the snake to high precision.



A high-resolution image of Mars reveals details such as mountains, craters, and the southern polar cap. (NASA)

telescope, but they would never identify a telescope as a 200-power telescope.

The quest for light-gathering power and high resolution explains why nearly all major observatories are located far from big cities and usually on high mountains. Astronomers avoid cities because **light pollution,** the brightening of the night sky by light scattered from artificial outdoor lighting, can make it impossible to see faint objects (■ Figure 5-9). In fact, many residents of cities are unfamiliar with the beauty of the night sky because they can see only the brightest stars. Even far from cities, nature's own light pollution, the moon, is sometimes so bright it drowns out fainter objects, and astronomers are often unable to observe on the nights near full moon when faint objects cannot be detected even with the largest telescopes on high mountains.

Astronomers prefer to place their telescopes on carefully selected high mountains. The air there is thin, very dry, and more transparent. For the best seeing, astronomers select mountains where the air flows smoothly and is not turbulent. Building an observatory on top of a

■ Figure 5-9

(a) This satellite view of the continental United States at night shows the light pollution and energy waste produced by outdoor lighting. Observatories cannot be located near large cities. (NOAA) (b) The domes of four giant telescopes are visible at upper left at Paranal Observatory, built by the European Southern Observatory. The Atacama Desert is believed to be the driest place on Earth. (ESO)

Reasoning with Numbers | 5-1

The Powers of a Telescope

Light-gathering power is proportional to the area of the telescope objective. A lens or mirror with a large area gathers a large amount of light. The area of a circular lens or mirror of diameter D is $\pi(D/2)^2$. To compare the relative light-gathering powers *(LGP)* of two telescopes A and B, you can calculate the ratio of the areas of their objectives, which reduces to the ratio of their diameters *(D)* squared:

$$\frac{LGP_A}{LGP_B} = \left(\frac{D_A}{D_B}\right)^2$$

Example A: Suppose you compare a 4-cm telescope with a 24-cm telescope. How much more light will the large telescope gather?

Solution:

$$\frac{LGP_{24}}{LGP_4} = \left(\frac{24}{4}\right)^2 = 6^2 = 36 \text{ times more light}$$

Example B: Your eye acts like a telescope with a diameter of about 0.8 cm, the maximum diameter of the pupil. How much more light can you gather if you use a 24-cm telescope?

Solution:

$$\frac{LGP_{24}}{LGP_{eye}} = \left(\frac{24}{0.8}\right)^2 = 30^2 = 900 \text{ times more light}$$

The resolving power of a telescope is the angular distance between two stars that are just barely visible through the telescope as two separate images. The resolving power α, in seconds of arc, equals 11.6 divided by the diameter of the telescope in centimeters:

$$\alpha = \left(\frac{11.6}{D}\right)$$

Example C: What is the resolving power of a 10.0-cm telescope?

Solution:

$$\alpha = \frac{11.6}{10} = 1.16 \text{ seconds of arc}$$

If the lenses are of good quality, and if the seeing is good, you should be able to distinguish as separate points of light any pair of stars farther apart than 1.16 seconds of arc. If the stars are any closer together, diffraction fringes blur the stars together into a single image.

The magnification M of a telescope is the ratio of the focal length of the primary lens or mirror F_P divided by the focal length of the eyepiece F_e:

$$M = \left(\frac{F_P}{F_e}\right)$$

Example D: What is the magnification of a telescope whose primary mirror has a focal length of 80 cm if it is used with an eyepiece whose focal length is 0.5 cm?

Solution: The magnification is 80 divided by 0.5, or 160 times.

high mountain far from civilization is difficult and expensive, as you can imagine from the photo in Figure 5-9b, but the dark sky and steady seeing make it worth the effort.

When you compare telescopes, you should consider their powers. ■ Reasoning with Numbers 5-1 shows how to calculate the powers of a telescope.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises “Telescopes and Resolution I,” “Telescopes and Resolution II,” and “Particulate, Heat, and Light Pollution.”

Observing at the Ends of the Visible Spectrum

Just beyond the red end of the visible spectrum some near-infrared radiation leaks through the atmosphere in narrow, partially open atmospheric windows scattered from 1200 nm to about 30,000 nm. Infrared astronomers usually measure wavelength in micrometers (10^{-6} meters), so they refer to this wavelength range as 1.2 to 30 micrometers (or microns for short). Even in this range, much of the radiation is absorbed by water vapor, plus carbon dioxide and oxygen molecules, which also absorb infrared. Nevertheless, some infrared observations can be made by telescopes on mountaintops where the air is thin and dry. For example, a number of important infrared telescopes observe from the 4200-m (13,800-ft) summit of Mauna Kea in Hawaii. At this altitude, the telescopes are above much of the water vapor in Earth’s atmosphere (■ Figure 5-10).

Infrared telescopes have flown to high altitudes under balloons and in airplanes. NASA is now testing the Stratospheric Observatory for Infrared Astronomy (SOFIA), a Boeing 747 that will carry a 2.5-m telescope, control systems, and a team of technicians and astronomers to the fringes of the atmosphere. Once at that altitude, they can open a door above the telescope and make infrared observations for hours as the plane flies a precisely calculated path. You can see the door in the photo in Figure 5-10.

To reduce internal noise, the light-sensitive detectors in astronomical telescopes are cooled to very low temperatures, usually with liquid nitrogen, as shown in Figure 5-10. This is especially

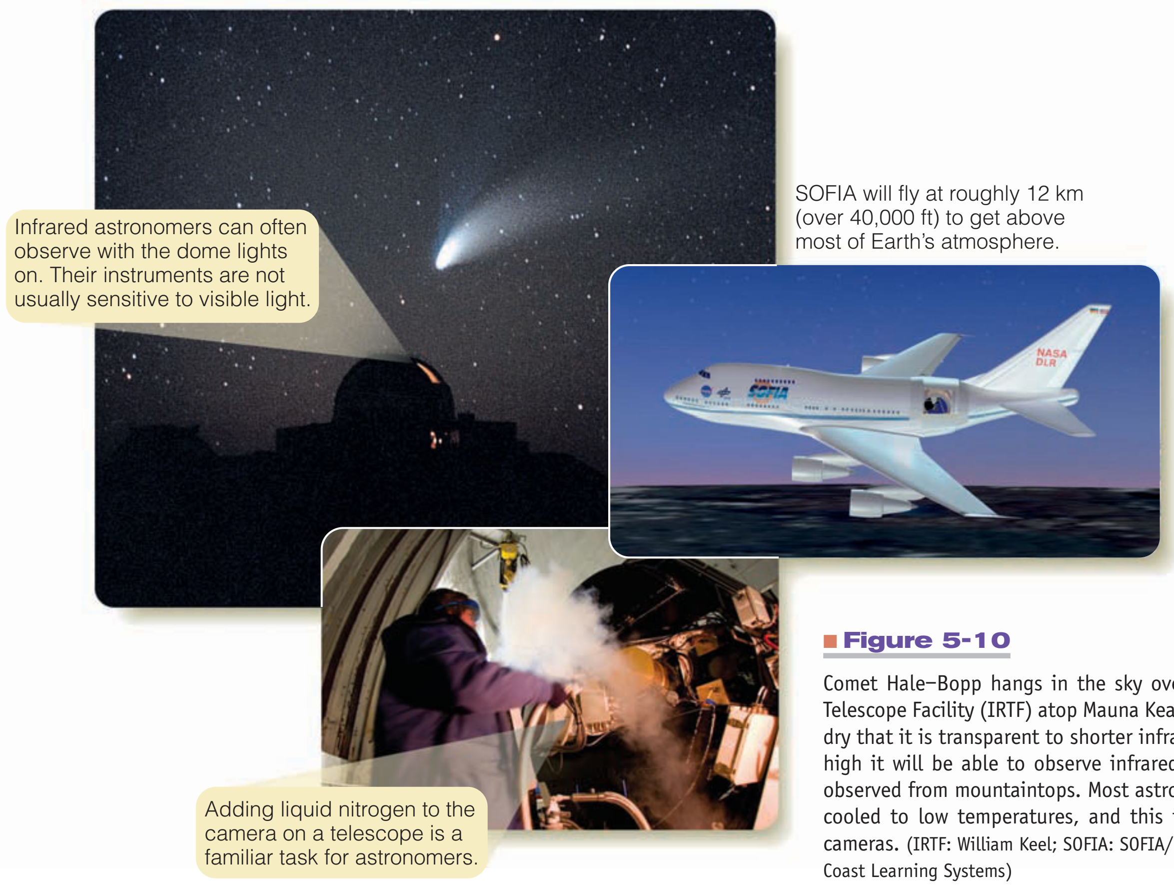


Figure 5-10

Comet Hale–Bopp hangs in the sky over the 3-meter NASA Infrared Telescope Facility (IRTF) atop Mauna Kea. The air at high altitudes is so dry that it is transparent to shorter infrared photons. SOFIA will fly so high it will be able to observe infrared wavelengths that cannot be observed from mountaintops. Most astronomical CCD cameras must be cooled to low temperatures, and this is especially true for infrared cameras. (IRTF: William Keel; SOFIA: SOFIA/USRA/NASA; Camera: Kris Koenig/Coast Learning Systems)

necessary for a telescope observing at infrared wavelengths. Infrared radiation is emitted by heated objects, and if the telescope is warm it will emit many times more infrared radiation than that coming from a distant object. Imagine trying to look for rabbits at night through binoculars that are themselves glowing.

At the other end of the spectrum, astronomers can observe in the near-ultraviolet. Your eyes don't detect this radiation, but it can be recorded by specialized detectors. Wavelengths shorter than about 290 nm, the far-ultraviolet, are completely absorbed by the ozone layer extending from 20 km to about 40 km above Earth's surface. No mountaintop is that high, and no airplane can fly to such an altitude. To observe in the far-ultraviolet or beyond at X-ray or gamma-ray wavelengths, telescopes must be in space above the atmosphere.

Buying a Telescope

Thinking about how to shop for a new telescope will not only help you if you decide to buy one but will also illustrate some important points about astronomical telescopes.

Assuming you have a fixed budget, you should buy the highest-quality optics and the largest-diameter telescope you can afford. You can't make the atmosphere less turbulent, but you can choose good optics. If you buy a telescope from a toy store and it has plastic lenses, you shouldn't expect to see very much. Also, you want to maximize the light-gathering power of your telescope, so you want to purchase the largest-diameter telescope you can afford. Given a fixed budget, that means you should buy a reflecting telescope rather than a refracting telescope. Not only will you get more diameter per dollar, but your telescope will not suffer from chromatic aberration.

You can safely ignore magnification. Department stores and camera shops may advertise telescopes by quoting their magnification, but it is not an important number. What you can see is fixed by light-gathering power, optical quality, and Earth's atmosphere. Besides, you can change the magnification by changing eyepieces.

Other things being equal, you should choose a telescope with a solid mounting that will hold the telescope steady and allow you to point it at objects easily. Computer-controlled pointing systems are available for a price on many small telescopes. A good telescope on a poor mounting is almost useless.

You might be buying a telescope to put in your backyard, but you must think about the same issues astronomers consider when they design giant telescopes to go on mountaintops. In fact, some of the new telescopes solve these traditional problems in new ways.

New-Generation Telescopes

For most of the 20th century, astronomers faced a serious limitation on the size of astronomical telescopes. Traditional telescope mirrors were made thick to avoid sagging that would distort the reflecting surface, but those thick mirrors were heavy. The 5-m (200-in.) mirror on Mount Palomar weighs 14.5 tons. These traditional telescopes were big, heavy, and expensive.

Modern astronomers have solved these problems in a number of ways. Read ■ **Modern Astronomical Telescopes** on pages 80–81 and notice three important points about telescope design and ten new terms that describe astronomical telescopes and their operation:

1. Traditional telescopes use large, solid, heavy mirrors to focus starlight to a *prime focus*, or, by using a *secondary mirror,* to a *Cassegrain focus*. Some small telescopes have a *Newtonian focus* or a *Schmidt-Cassebrain focus.*

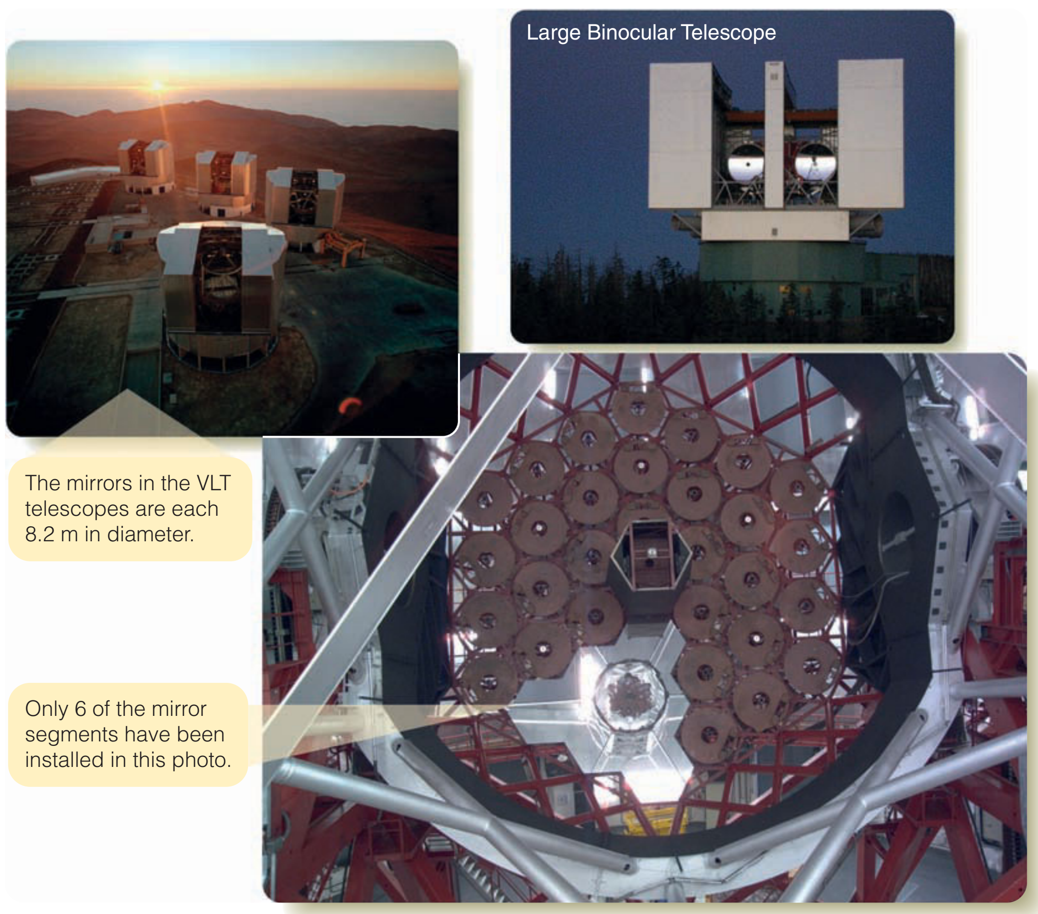
2. Telescopes must have a *sidereal drive* to follow the stars, and an *equatorial mounting* with easy motion around a *polar axis* is the traditional way to provide that motion. Today, astronomers can build simpler, lighter-weight telescopes on *alt-azimuth mountings* and depend on computers to move the telescope and follow the westward motion of the stars as Earth rotates.

3. *Active optics,* computer control of the shape of telescope mirrors, allows the use of thin, lightweight mirrors—either "floppy" mirrors or segmented mirrors. Lowering the weight of the mirror lowers the weight of the rest of the telescope and makes it stronger and less expensive. Also, thin mirrors cool faster at nightfall and produce better images.

High-speed computers have allowed astronomers to build new, giant telescopes with unique designs. A few are shown in ■ Figure 5-11. The European Southern Observatory has built the Very Large Telescope (VLT) high in the remote Andes Mountains of northern Chile. The VLT consists of four telescopes with computer-controlled mirrors 8.2 m in diameter and only 17.5 cm (6.9 in.) thick. The four telescopes can work singly or can combine their light to work as one large telescope. Italian and American astronomers have built the Large Binocular Telescope, which carries a pair of 8.4-m mirrors on a single mounting. The Gran Telescopio Canarias, located atop a volcanic peak in the Canary Islands, carries a segmented mirror 10.4 m in diameter and holds, for the moment, the record as the largest single telescope in the world.

Other giant telescopes are being planned with segmented mirrors or with multiple mirrors (■ Figure 5-12). The Giant Magellan Telescope will carry seven thin mirrors, each 8.4 m in diameter, on a single mounting. It will be located in the Chilean Andes and will have the light-gathering power of a 22-m telescope. The Thirty Meter Telescope, now under development by American astronomers, will have a mirror 30 m in diameter comprised of 492 hexagonal segments. The European Extremely Large Telescope is being planned by an international team. It will carry 906 segments making up a mirror 42 m in diameter. Other very large telescopes are being proposed with completion dates of 2016 or later.

Modern computers have revolutionized telescope design and operation. Nearly all large telescopes are operated by astronomers



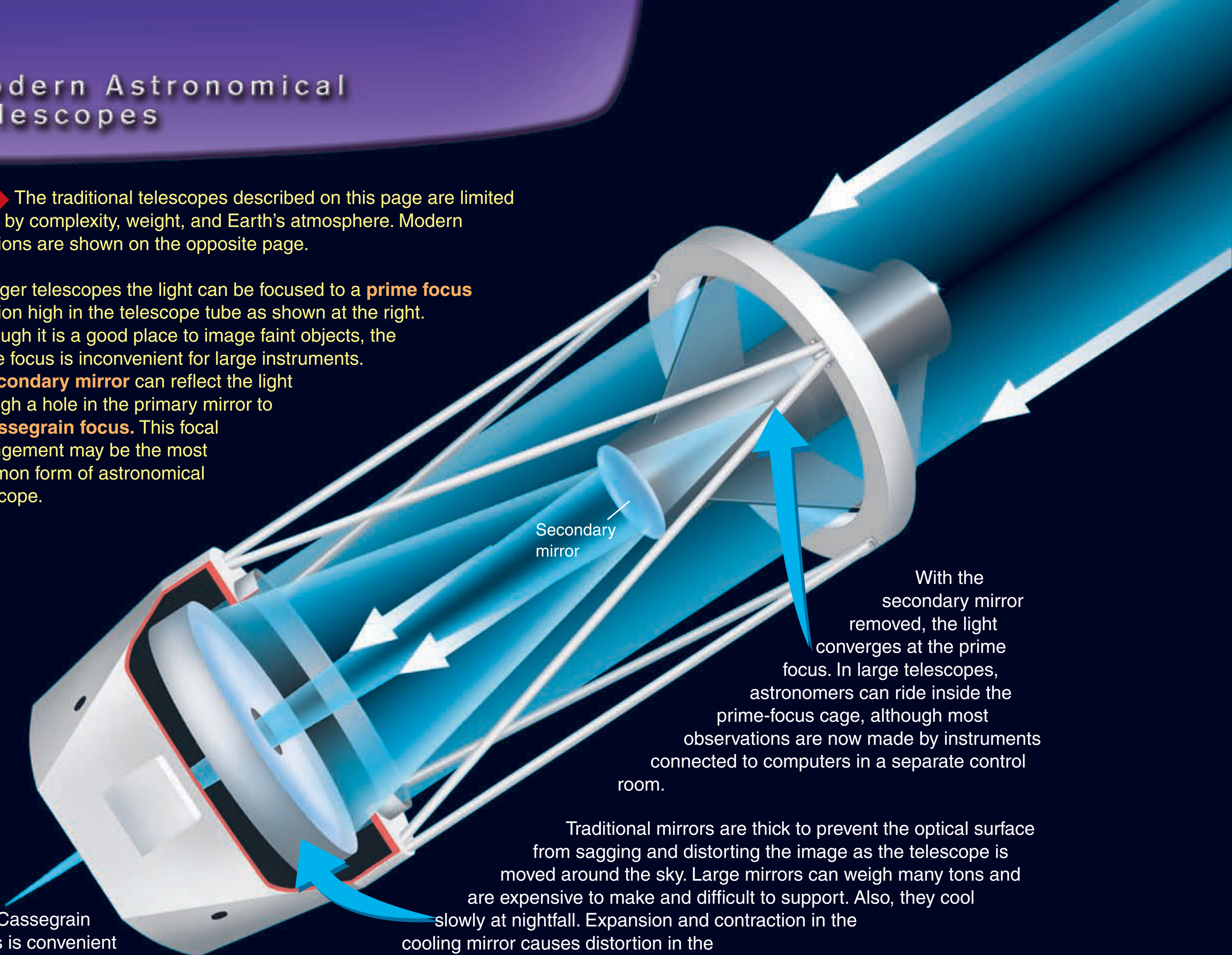
■ **Figure 5-11**

The four telescopes of the VLT are housed in separate domes at Paranal Observatory in Chile (Figure 5-9). The Large Binocular Telescope (LBT) carries two 8.4-m mirrors that combine their light. The entire building rotates as the telescope moves. The Gran Telescopio Canarias contains 36 hexagonal mirror segments in its 10.4-m primary mirror. (VLT: ESO; LBT: Large Binocular Telescope Project and European Industrial Engineer; GMT: ESO; Gran Telescopio CANARIAS: Instituto de Astrofisica de Canarias)

Modern Astronomical Telescopes

1 The traditional telescopes described on this page are limited by complexity, weight, and Earth's atmosphere. Modern solutions are shown on the opposite page.

In larger telescopes the light can be focused to a **prime focus** position high in the telescope tube as shown at the right. Although it is a good place to image faint objects, the prime focus is inconvenient for large instruments. A **secondary mirror** can reflect the light through a hole in the primary mirror to a **Cassegrain focus.** This focal arrangement may be the most common form of astronomical telescope.

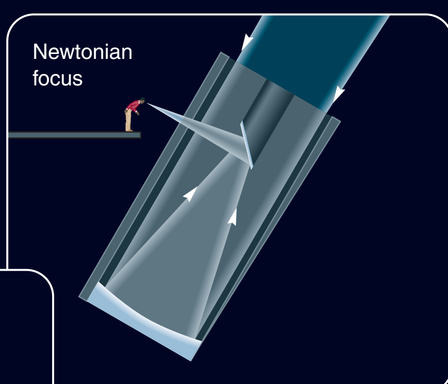


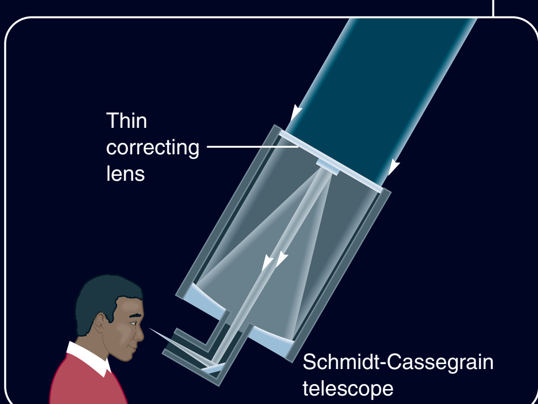
With the secondary mirror removed, the light converges at the prime focus. In large telescopes, astronomers can ride inside the prime-focus cage, although most observations are now made by instruments connected to computers in a separate control room.

Traditional mirrors are thick to prevent the optical surface from sagging and distorting the image as the telescope is moved around the sky. Large mirrors can weigh many tons and are expensive to make and difficult to support. Also, they cool slowly at nightfall. Expansion and contraction in the cooling mirror causes distortion in the images.

The Cassegrain focus is convenient and has room for large instruments.

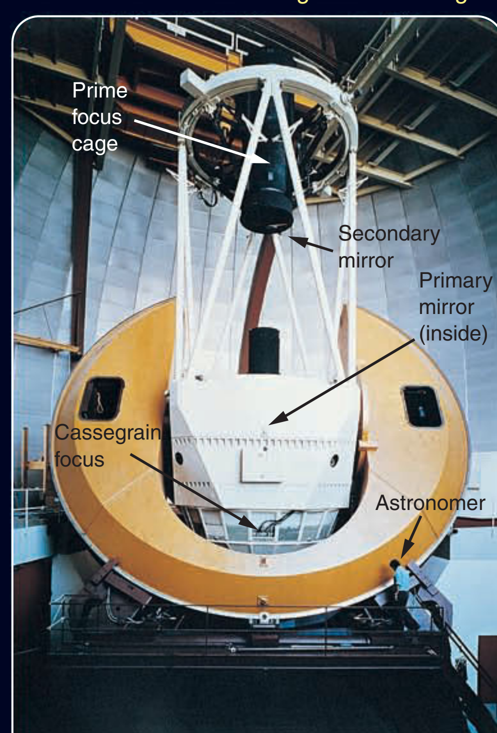
1a Smaller telescopes are often found with a **Newtonian focus,** the arrangement that Isaac Newton used in his first reflecting telescope. The Newtonian focus is inconvenient for large telescopes as shown at right.



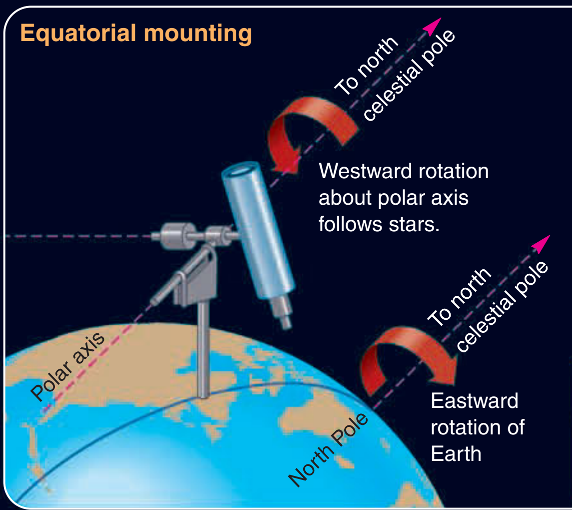


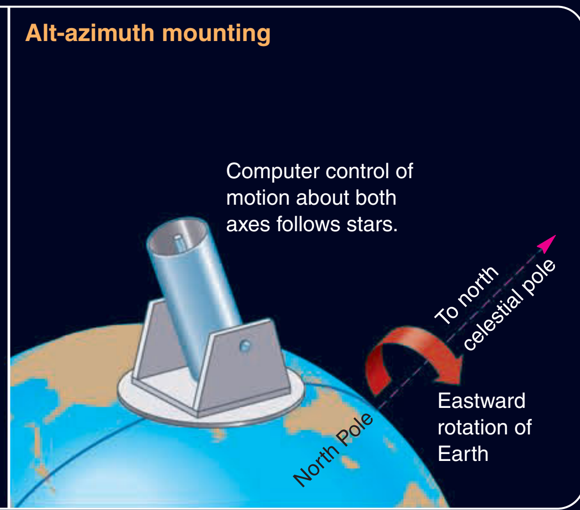
1b Many small telescopes such as the one on your left use a **Schmidt-Cassegrain focus.** A thin correcting plate improves the image but is too slightly curved to introduce serious chromatic aberration.

1c Shown below, the 4-meter Mayall Telescope at Kitt Peak National Observatory in Arizona can be used at either the prime focus or the Cassegrain focus. Note the human figure at lower right.



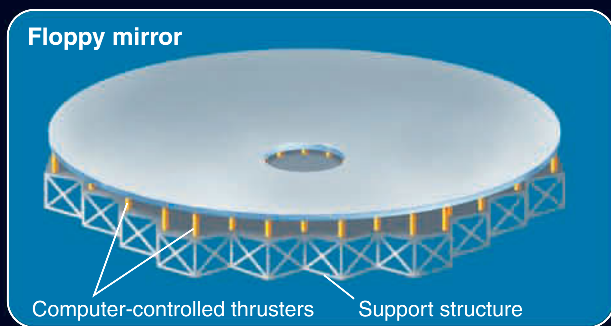
AURA/NOAO/NSF





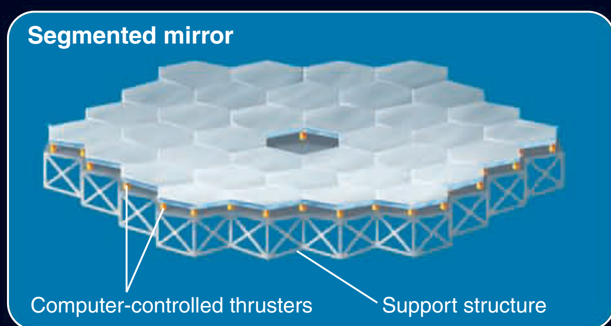
2 Telescope mountings must contain a **sidereal drive** to move smoothly westward and counter the eastward rotation of Earth. The traditional **equatorial mounting** (far left) has a **polar axis** parallel to Earth's axis, but the modern **alt-azimuth mounting** (near left) moves like a cannon — up and down and left to right. Such mountings are simpler to build but need computer control to follow the stars.

3 Unlike traditional thick mirrors, thin mirrors, sometimes called floppy mirrors as shown at right, weigh less and require less massive support structures. Also, they cool rapidly at nightfall and there is less distortion from uneven expansion and contraction.



3a Grinding a large mirror may remove tons of glass and take months, but new techniques speed the process. Some large mirrors are cast in a rotating oven that causes the molten glass to flow to form a concave upper surface. Grinding and polishing such a preformed mirror is much less time consuming.

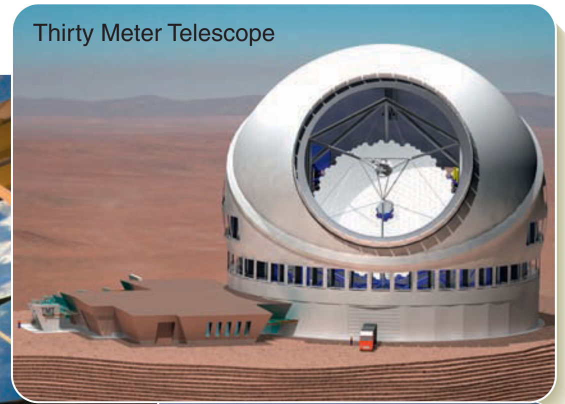
3b Mirrors made of segments are economical because the segments can be made separately. The resulting mirror weighs less and cools rapidly. See image at right.



3c Both floppy mirrors and segmented mirrors sag under their own weight. Their optical shape must be controlled by computer-driven thrusters under the mirror in what is called **active optics.**

W.M. Keck Observatory

3d The two Keck telescopes, each 10 meters in diameter, are located atop the volcano Mauna Kea in Hawaii. The two mirrors are composed of hexagonal mirror segments as shown at right.



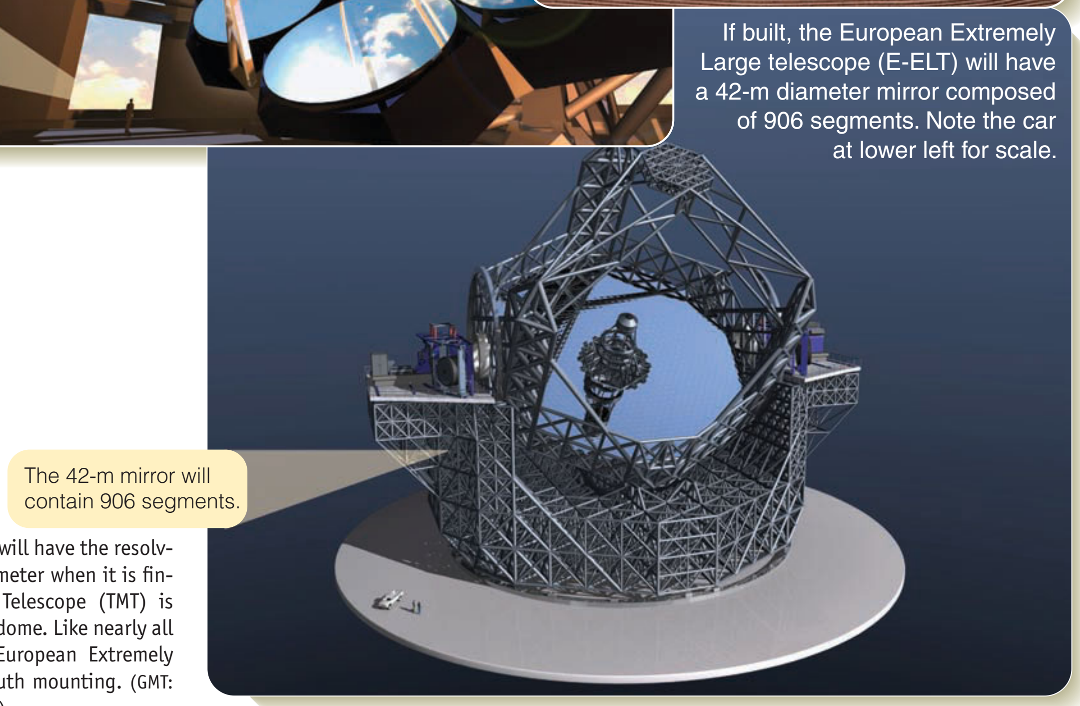


Figure 5-12

The proposed Giant Magellan Telescope will have the resolving power of a telescope 24.5 m in diameter when it is finished about 2016. The Thirty Meter Telescope (TMT) is planned to occupy a specially designed dome. Like nearly all of the newest large telescopes, the European Extremely Large Telescope will be on an alt-azimuth mounting. (GMT: ESO; TMT: Thirty-Meter Telescope; E-ELT: ESO)

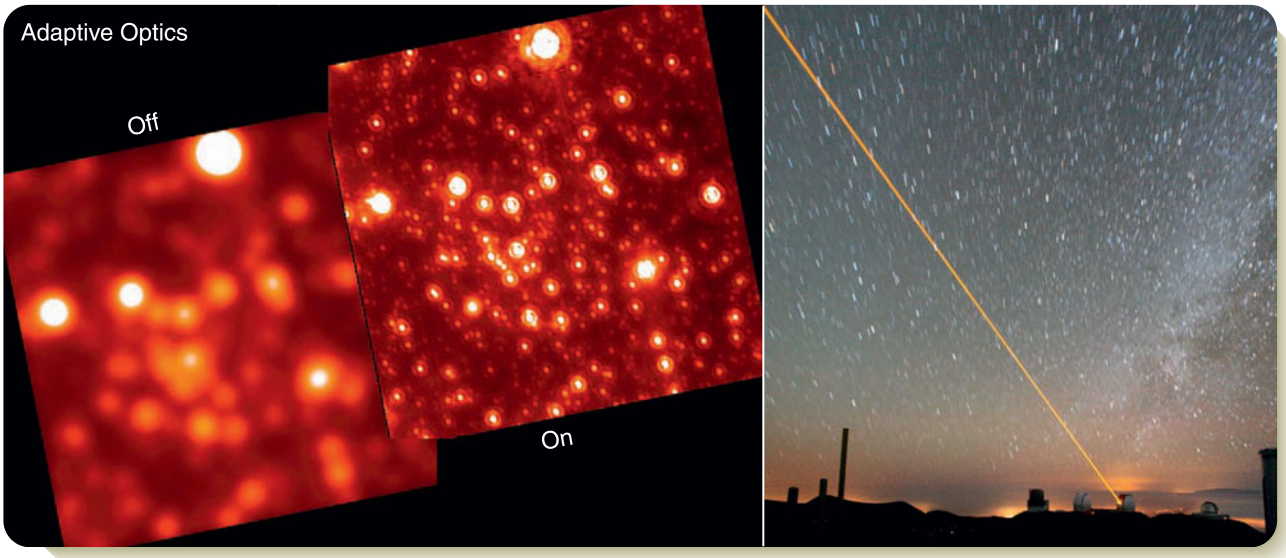
and technicians working at computers in a control room, and some telescopes can be operated by astronomers thousands of miles from the observatory. Some telescopes are fully automated and observe without direct human supervision. This has made possible huge surveys of the sky in which millions of objects are observed. The Sloan Digital Sky Survey, for example, mapped the sky, measuring the position and brightness of 100 million stars and galaxies at a number of wavelengths. The Two-Micron All Sky Survey (2MASS) has mapped the entire sky at three wavelengths in the infrared. Other surveys are being made at other wavelengths. Astronomers will study those data banks for decades to come.

Adaptive Optics

Not too many years ago, astronomers thought it was pointless to build more large telescopes on Earth's surface because of seeing distortion caused by the atmosphere. In the 1990s, computers became fast enough to allow astronomers to correct for some of that distortion, and that has made a new generation of giant telescopes possible.

Adaptive optics uses high-speed computers to monitor the distortion produced by turbulence in Earth's atmosphere and then correct the telescope image to sharpen a fuzzy blob into a crisp picture. The resolution of the image is still limited by diffraction in the telescope, but removing much of the seeing distortion produces a dramatic improvement in the detail visible (■ Figure 5-13). Don't confuse adaptive optics with the slower-speed active optics that controls the overall shape of a telescope mirror.

To monitor the distortion in an image, adaptive optics systems must look at a fairly bright star in the field of view, and there isn't always such a star properly located near a target object such as a faint galaxy. In that case, astronomers can point a laser at a spot in the sky very close to their target object, and where the laser excites Earth's upper atmosphere, it produces an artifi-



■ **Figure 5-13**

In these images of the center of our galaxy, the adaptive optics system was turned off for the left image and on for the right image. Not only are the images of stars sharper, but because the light is focused into smaller images, fainter stars are visible. The laser beam shown leaving one of the Keck Telescopes produces an artificial star in the field of view, and the adaptive optics system uses the laser-produced point of light to reduce seeing distortion in the entire image. (left: CFHT; right: Paul Hirst)

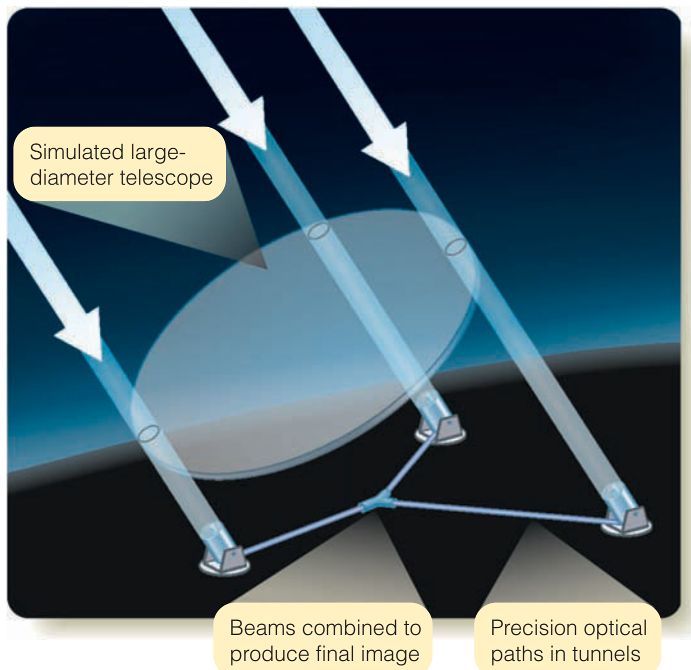
cial star in the field of view. The adaptive optics system can use the artificial star to correct the image of the fainter target.

Today astronomers are planning huge optical telescopes composed of segmented mirrors tens of meters in diameter. Those telescopes would be almost useless without adaptive optics.

Interferometry

One of the reasons astronomers build big telescopes is to increase resolving power, and astronomers have been able to achieve very high resolution by connecting multiple telescopes together to work as if they were a single telescope. This method of synthesizing a larger telescope is known as **interferometry** (■ Figure 5-14). One expert said, "We combine the light from separate telescopes and fool the waves into thinking they were collected by one big 'scope.'" The images from such a virtual telescope are not limited by the diffraction fringes of the individual small telescopes but rather by the diffraction fringes of the much larger virtual telescope.

To work as an interferometer, the separate telescopes must combine their light through a network of mirrors, and the path that each light beam travels must be controlled so that it does not vary more than some small fraction of the wavelength. Turbulence in Earth's atmosphere constantly distorts the light, and high-speed computers must continuously adjust the light paths. Recall that the wavelength of light is very short, roughly 0.0005 mm, so building optical interferometers is one of the most difficult technical problems that astronomers face. Infrared- and radio-wavelength interferometers are slightly easier to build because the wavelengths are longer. In fact, as you will discover later in this chapter, the first astronomical interferometers were built by radio astronomers.



■ **Figure 5-14**

In an astronomical interferometer, smaller telescopes can combine their light through specially designed optical tunnels to simulate a larger telescope with a resolution set by the separation of the smaller telescopes.

The VLT shown in Figure 5-11 consists of four 8.2-m telescopes that can operate separately but can also be linked together through underground tunnels with three 1.8-m telescopes on the same mountaintop. The resulting optical interferometer provides the resolution of a telescope 200 m in diameter. Other telescopes such as the two Keck 10-m telescopes can work as interferometers. The CHARA array on Mt. Wilson combines six 1-m telescopes to create the equivalent of a telescope one-fifth of a mile in diameter. The Large Binocular Telescope shown in Figure 5-11 can be used as an interferometer.

Although turbulence in Earth's atmosphere can be partially averaged out in an interferometer, plans are being made to put interferometers in space to avoid atmospheric turbulence altogether. The Space Interferometry Mission, for example, will work at visual wavelengths and study everything from the cores of erupting galaxies to planets orbiting nearby stars.

◄ SCIENTIFIC ARGUMENT ►

Why do astronomers build observatories at the tops of mountains?

To build this argument you need to think about the powers of a telescope. Astronomers have joked that the hardest part of building a new observatory is constructing the road to the top of the mountain. It certainly isn't easy to build a large, delicate telescope at the top of a high mountain, but it is worth the effort. A telescope on top of a high mountain is above the thickest part of Earth's atmosphere. There is less air to dim the light, and there is less water vapor to absorb infrared radiation. Even more important, the thin air on a mountaintop causes less disturbance to the image, and consequently the seeing is better. A large telescope on Earth's surface has a resolving power much better than the distortion caused by Earth's atmosphere. So, it is limited by seeing, not by its own diffraction. It really is worth the trouble to build telescopes atop high mountains.

Astronomers not only build telescopes on mountaintops, they also build gigantic telescopes many meters in diameter. Revise your argument to focus on telescope design. **What are the problems and advantages in building such giant telescopes?**

◄ ►

5-3 Special Instruments

Just looking through a telescope doesn't tell you much. A star looks like a point of light. A planet looks like a little disk. A galaxy looks like a hazy patch. To use an astronomical telescope to learn about the universe, you must be able to analyze the light the telescope gathers. Special instruments attached to the telescope make that possible.

Imaging Systems

The original imaging device in astronomy was the photographic plate. It could record images of faint objects in long time exposures and could be stored for later analysis. But photographic plates have been almost entirely replaced by electronic imaging systems.

Most modern astronomers use **charge-coupled devices (CCDs)** to record images. A CCD is a specialized computer chip containing roughly a million microscopic light detectors arranged in an array about the size of a postage stamp. Although CCDs for astronomy are extremely sensitive and therefore expensive, less sophisticated CCDs are used in video and digital cameras. Not only can CCD chips replace photographic plates, but they have some dramatic advantages. They can detect both bright and faint objects in a single exposure, are much more sensitive than photographic plates, and can be read directly into computer memory for later analysis.

You can sharpen and enhance images from your digital camera because the image from a CCD is stored as numbers in computer memory. Astronomers can also manipulate images to bring out otherwise invisible details. For example, astronomical images are often reproduced as negatives with the sky white and the stars dark. This makes the faint parts of the image easier to see (■ Figure 5-15). Astronomers can also produce **false-color images** in which the colors represent different levels of intensity and are not related to the true colors of the object. You can see an example in Figure 5-15. In fact, false-color images are common in many fields such as medicine and meteorology.

In the past, measurements of intensity and color were made using specialized light meters attached to a telescope or on photographic plates. Today, nearly all such measurements are made more easily and more accurately with CCD images.

The Spectrograph

To analyze light in detail, astronomers need to spread the light out according to wavelength to form a spectrum, a task performed by a **spectrograph.** You can understand how this works if you imagine reproducing an experiment performed by Isaac Newton in 1666. Newton bored a small hole in the window shutter of his bedroom to admit a thin beam of sunlight. When he placed a prism in the beam, it spread the light into a beautiful spectrum that splashed across his bedroom wall. From this Newton concluded that white light was made of a mixture of all the colors.

Light passing through a prism is bent at an angle that depends on its wavelength. Violet (short wavelength) bends most, red (long wavelength) least, spreading the white light into a spectrum (■ Figure 5-16). You could build a spectrograph with a prism to spread the light and a lens to guide the light into a camera.

Nearly all modern spectrographs use a grating in place of a prism. A **grating** is a piece of glass with thousands of microscopic parallel grooves scribed onto its surface. Different wavelengths of light reflect from the grating at slightly different angles, so white light is spread into a spectrum. You have probably noticed this effect when you look at the closely spaced lines etched onto a compact disk; as you tip the disk, different colors flash across its surface. You could build a modern spectrograph by using a high-quality grating to spread light into a spectrum and a CCD camera to record the spectrum.

The spectrum of an astronomical object can contain hundreds of **spectral lines**—dark or bright lines that cross the spectrum at specific wavelengths. The sun's spectrum, for instance,

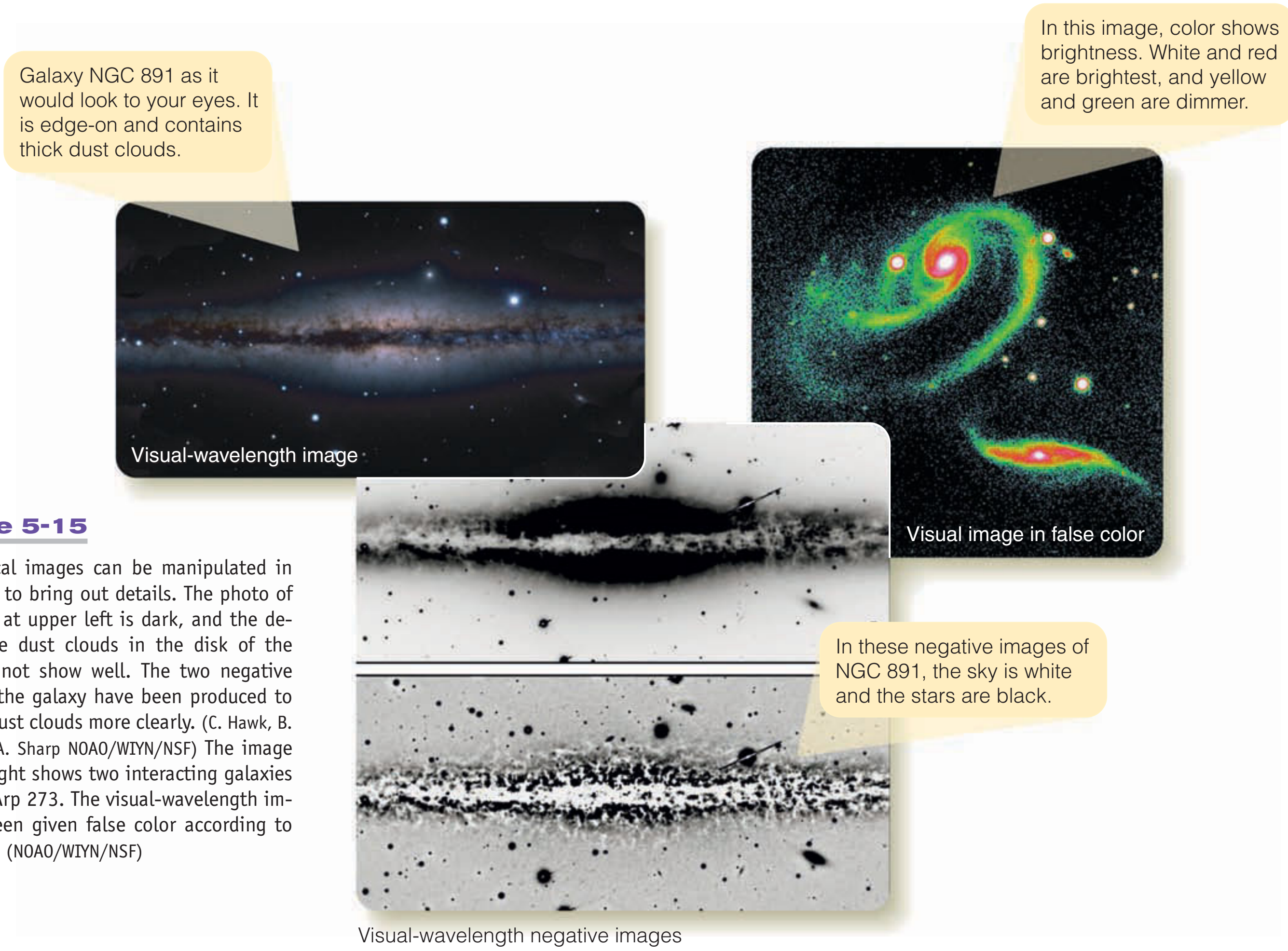


Figure 5-15

Astronomical images can be manipulated in many ways to bring out details. The photo of the galaxy at upper left is dark, and the details of the dust clouds in the disk of the galaxy do not show well. The two negative images of the galaxy have been produced to show the dust clouds more clearly. (C. Hawk, B. Savage, N. A. Sharp NOAO/WIYN/NSF) The image at upper right shows two interacting galaxies known as Arp 273. The visual-wavelength image has been given false color according to brightness. (NOAO/WIYN/NSF)

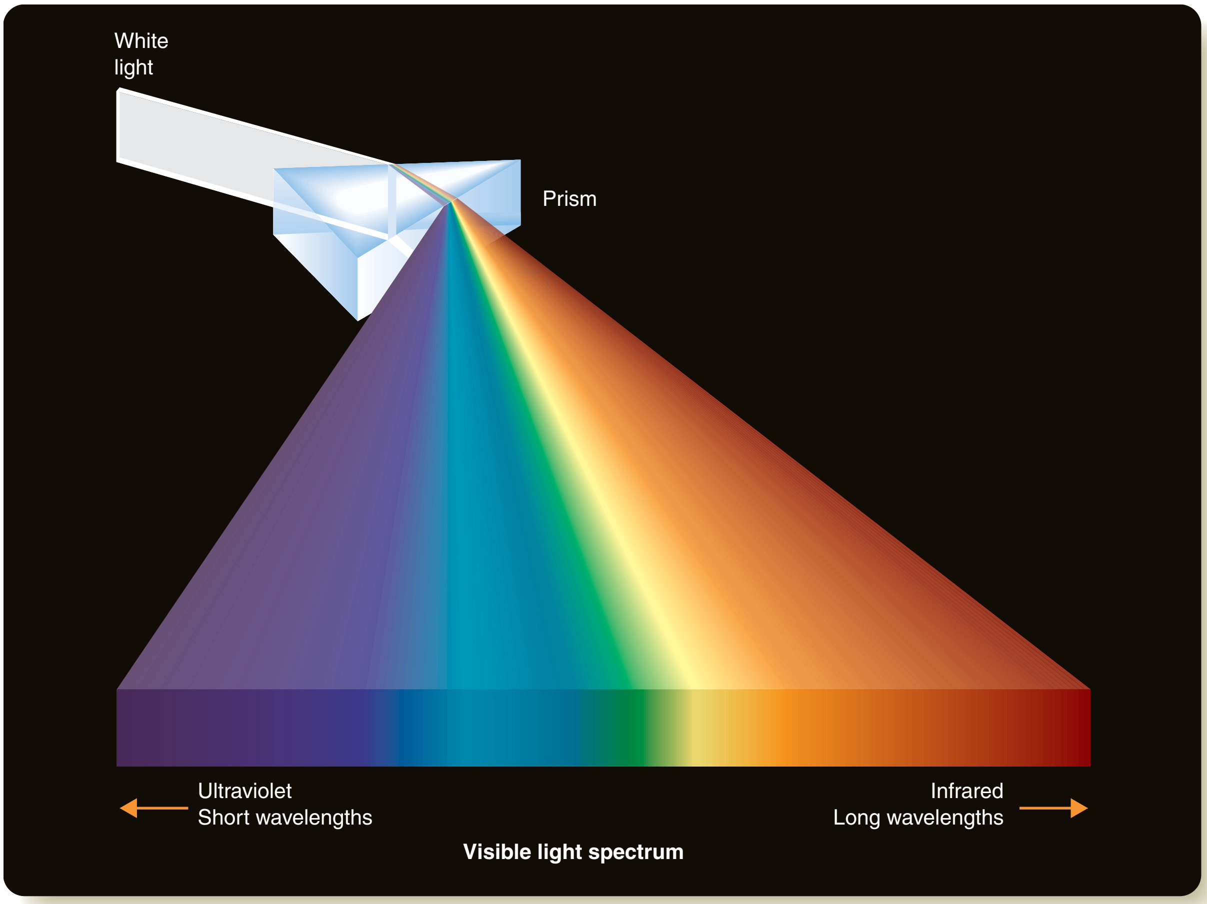


Figure 5-16

A prism bends light by an angle that depends on the wavelength of the light. Short wavelengths bend most and long wavelengths least. Thus, white light passing through a prism is spread into a spectrum.

contains hundreds of dark spectral lines produced by the atoms in the sun's hot gases. To measure the precise wavelengths of individual lines and identify the atoms that produced them, astronomers use a **comparison spectrum** as a calibration. Special bulbs built into the spectrograph produce bright lines given off by such atoms as thorium and argon or neon. The wavelengths of these spectral lines have been measured to high precision in the laboratory, so astronomers can use spectra of these light sources as guides to measure wavelengths and identify spectral lines in the spectrum of a star, galaxy, or planet.

Because astronomers understand how light interacts with matter, a spectrum carries a tremendous amount of information (as you will see in the next chapter), and that makes a spectrograph the astronomer's most powerful instrument. An astronomer once remarked, "We don't know anything about an object till we get a spectrum," and that is only a slight exaggeration.

5-4 Radio Telescopes

Celestial objects such as clouds of gas and erupting stars emit radio energy, and astronomers on Earth can study such objects by observing at wavelengths in the radio window where Earth's atmosphere is transparent to radio waves (see Figure 5-2). You might think an erupting star would produce a strong radio signal, but the signals arriving on Earth are astonishingly weak—a million to a billion times weaker than the signal from an FM radio station. Detecting such weak signals calls for highly sensitive equipment.

The Operation of a Radio Telescope

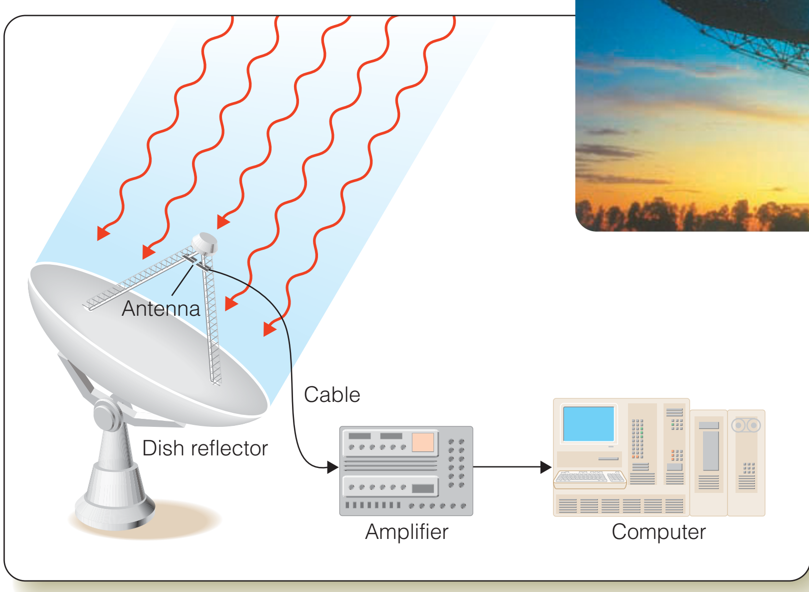
A radio telescope usually consists of four parts: a dish reflector, an antenna, an amplifier, and a recorder (■ Figure 5-17). These components, working together, make it possible for astronomers to detect radio radiation from celestial objects.

The dish reflector of a radio telescope, like the mirror of a reflecting telescope, collects and focuses radiation. Because radio waves are much longer than light waves, the dish need not be as smooth as a mirror; wire mesh will reflect all but the shortest wavelength radio waves.

Though a radio telescope's dish may be many meters in diameter, the antenna may be as small as your hand. Like the antenna on a TV set, its only function is to absorb the radio energy collected by the dish. Because the radio energy from celestial objects is so weak, it must be strongly amplified before it is recorded into computer memory.

A single observation with a radio telescope measures the amount of radio energy coming from a specific point on the sky, but the intensity at one spot doesn't tell you much. So the radio telescope must be scanned over an object to produce a map of the radio intensity at different points.

Because humans can't see radio waves, astronomers draw maps in which contours mark areas of similar radio intensity. You could compare such a radio map to a weather map showing

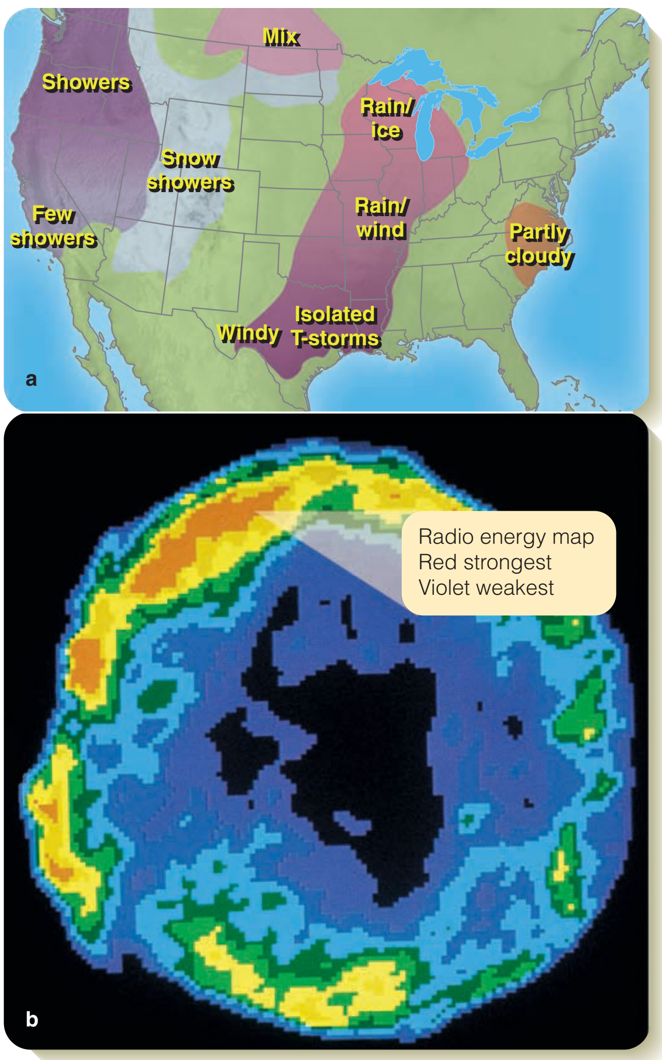


■ **Figure 5-17**

In most radio telescopes, a dish reflector concentrates the radio signal on the antenna. The signal is then amplified and recorded. For all but the shortest radio waves, wire mesh is an adequate reflector (photo). (Courtesy Seth Shostak/SETI Institute) contours filled with color to indicate areas of precipitation (■ Figure 5-18).

Limitations of a Radio Telescope

A radio astronomer works under three handicaps: poor resolution, low intensity, and interference. Recall that the resolving power of an optical telescope depends on the diameter of the objective lens or mirror. It also depends on the wavelength of the radiation. At very long wavelengths, like those of radio waves, the diffraction fringes are very large and the radio maps can't show fine detail. As with an optical telescope, there is no way to improve the resolving power without building a bigger telescope. Consequently, radio telescopes generally have large diameters to minimize the diffraction fringes.



■ **Figure 5-18**

(a) A typical weather map uses contours with added color to show which areas are likely to receive precipitation. (b) A false-color-image radio map of Tycho's supernova remnant, the expanding shell of gas produced by the explosion of a star in 1572. The radio contour map has been color-coded to show intensity. (Courtesy NRAO)

Even so, the resolving power of a radio telescope is not good. A dish 30 m in diameter receiving radiation with a wavelength of 21 cm has a resolving power of about 0.5°. Such a radio telescope would be unable to detect any details in the sky smaller than the moon. Fortunately, radio astronomers can combine two or more radio telescopes to form a **radio interferometer** capable of much higher resolution. Just as in the case of optical interferometers, the radio astronomer combines signals from two or more widely separated dishes and "fools the waves" into behaving as if they were collected by a much bigger radio telescope.

Radio interferometers can be quite large. The Very Large Array (VLA) consists of 27 dish antennas spread in a Y-shape across the New Mexico desert (■ Figure 5-19). In combination, they have the resolving power of a radio telescope 36 km (22 mi) in diameter. The VLA can resolve details smaller than 1 second of arc. Eight new dish antennas being added across New Mexico will give the VLA ten times better resolving power. Another large radio interferometer, the Very Long Baseline Array (VLBA), consists of matched radio dishes spread from Hawaii to the Virgin Islands and has an effective diameter almost as large as Earth. The Allen Telescope Array being built in California will eventually include 350 separate radio dishes. Radio astronomers are now planning the Square Kilometre Array, which will contain a huge number of radio dishes totaling a square kilometer of collecting area and spread to a diameter of at least 6000 km. These huge radio interferometers depend on state-of-the-art, high-speed computers to combine signals and create radio images.

The second handicap radio astronomers face is the low intensity of the radio signals. You learned earlier that the energy of a photon depends on its wavelength. Photons of radio energy have such long wavelengths that their individual energies are quite low. To get detectable signals focused on the antenna, the radio astronomer must build large collecting areas either as single large dishes or arrays of smaller dishes.

The largest fully steerable radio telescope in the world is at the National Radio Astronomy Observatory in Green Bank, West Virginia (■ Figure 5-20a). The telescope has a reflecting surface 100 m in diameter, big enough to hold an entire football field, and can be pointed anywhere in the sky. Its surface consists of 2004 computer-controlled panels that adjust to maintain the shape of the reflecting surface.

The largest radio dish in the world is 300 m (1000 ft) in diameter. So large a dish can't be supported in the usual way, so it is built into a mountain valley in Arecibo, Puerto Rico. The reflecting dish is a thin metallic surface supported above the valley floor by cables attached near the rim, and the antenna hangs above the dish on cables from three towers built on three mountain peaks that surround the valley (Figure 5-20b). By moving the antenna above the dish, radio astronomers can point the telescope at any object that passes within 20 degrees of the zenith as Earth rotates. Since completion in 1963, the telescope has been an international center of radio astronomy research.

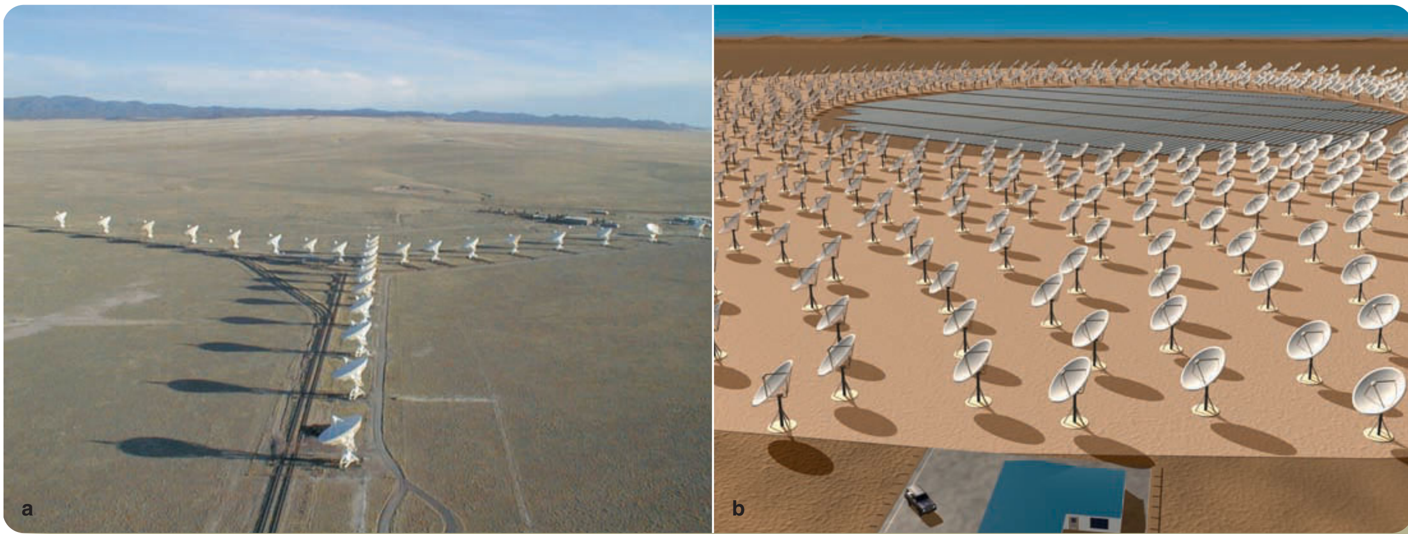
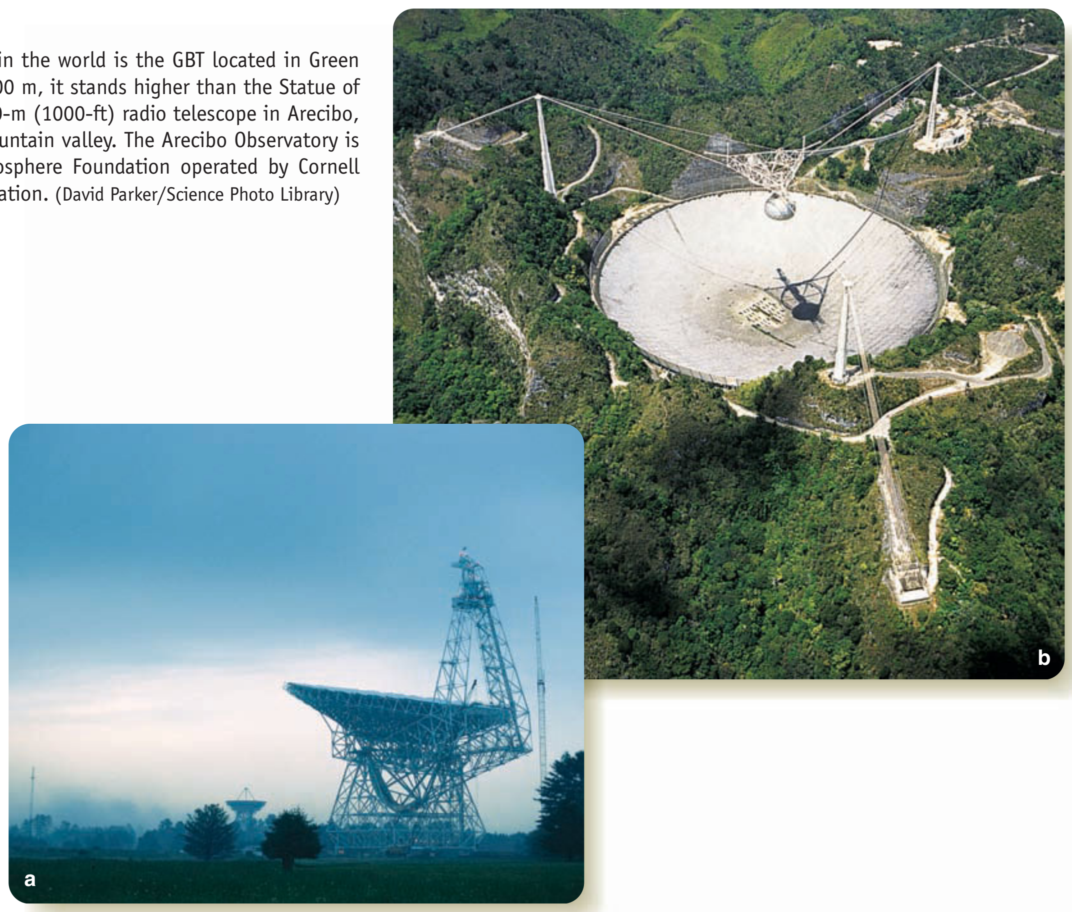


Figure 5-19

(a) The Very Large Array uses 27 radio dishes, which can be moved to different positions along a Y-shaped set of tracks across the New Mexico desert. They are shown here in the most compact arrangement. Signals from the dishes are combined to create very-high-resolution radio maps of celestial objects. (NRAO) (b) The proposed Square Kilometre Array will have a concentration of detectors and radio dishes near the center with more dishes scattered up to 3000 km away. (© Xilostudios/SKA Program Development Office)

Figure 5-20

(a) The largest steerable radio telescope in the world is the GBT located in Green Bank, West Virginia. With a diameter of 100 m, it stands higher than the Statue of Liberty. (Mike Bailey: NRAO/AUI) (b) The 300-m (1000-ft) radio telescope in Arecibo, Puerto Rico, hangs from cables over a mountain valley. The Arecibo Observatory is part of the National Astronomy and Ionosphere Foundation operated by Cornell University and the National Science Foundation. (David Parker/Science Photo Library)



The third handicap the radio astronomer faces is interference. A radio telescope is an extremely sensitive radio receiver listening to faint radio signals. Such weak signals are easily drowned out by interference that includes everything from poorly designed transmitters in Earth satellites to automobiles with faulty ignition systems. A few narrow radio bands in the electromagnetic spectrum are reserved for radio astronomy, but even those are often contaminated by radio noise. To avoid interference, radio astronomers locate their telescopes as far from civilization as possible. Hidden deep in mountain valleys, they are able to listen to the sky protected from human-made radio noise.

Advantages of a Radio Telescope

Building large radio telescopes in isolated locations is expensive, but three factors make it all worthwhile. First, and most important, a radio telescope can reveal where clouds of cool hydrogen, and other atoms and molecules, are located. These hydrogen clouds are important because, for one thing, they are the places where stars are born. Large clouds of cool hydrogen are completely invisible to normal telescopes because they produce no visible light of their own and reflect too little to be detected in photographs. However, cool hydrogen emits a radio signal at the specific wavelength of 21 cm. (You will see how the hydrogen produces this radiation in the discussion of the gas clouds in space in Chapter 12.) The only way astronomers can detect these clouds of hydrogen is with a radio telescope that receives the 21-cm radiation.

Another reason radio telescopes are important is related to dust in space. Astronomers observing at visual wavelengths can't see through the dusty clouds in space. Light waves are so short they are scattered by the tiny dust grains and never get through the dust to reach optical telescopes on Earth. However, radio signals have wavelengths much longer than the diameters of dust grains, so radio waves from far across the galaxy pass unhindered through the dust, giving radio astronomers an unobscured view.

Finally, radio telescopes are important because they can detect objects that are more luminous at radio wavelengths than at visible wavelengths. This includes intensely hot gas orbiting black holes. Some of the most violent events in the universe are detectable at radio wavelengths.

5-5 Astronomy from Space

You have learned about the observations that ground-based telescopes can make through the two atmospheric windows in the visible and radio parts of the electromagnetic spectrum. Most of the rest of the electromagnetic radiation—infrared, ultraviolet, X-ray, and gamma ray—never reaches Earth's surface; it is absorbed high in Earth's atmosphere. To observe at these wavelengths, telescopes must go above the atmosphere.

The Hubble Space Telescope

Named after Edwin Hubble, the astronomer who discovered the expansion of the universe, the Hubble Space Telescope is the most successful telescope ever to orbit Earth (■ Figure 5-21). It was launched in 1990 and contains a 2.4-m (95-in.) mirror with which it can observe from the near-infrared to the near-ultraviolet. It is controlled from a research center on Earth and observes continuously. Nevertheless, there is time to complete only a fraction of the projects proposed by astronomers from around the world.

Most of the observations Hubble makes are at visual wavelengths, so its greatest advantage in being above Earth's atmosphere is the lack of seeing distortion. It can detect fine detail and by concentrating light into sharp images can see faint objects.

The telescope is as big as a large bus and has been visited a number of times by the space shuttle so that astronauts can maintain its equipment and install new cameras and spectrographs. Astronomers hope that it will last until it is replaced by the James Webb Space Telescope expected to launch no sooner than 2013. The Webb telescope will carry a 6.5-m (256-in.) mirror.

Infrared Astronomy from Orbit

Telescopes that observe in the far-infrared must be protected from heat and must get above Earth's absorbing atmosphere. They have limited lifetimes because they must carry coolant to chill their optics. The Infrared Astronomical Satellite (IRAS) was a joint project of the United Kingdom, the United States, and the Netherlands. IRAS was launched in January of 1983 and carried liquid helium coolant to keep its telescope cold. It made 250,000 observations and, for example, discovered disks of dust around stars where planets are now thought to have formed. Its coolant ran out after 300 days of observation.

The most sophisticated of the infrared telescopes put in orbit, the Spitzer Space Telescope is cooled to –269°C (–452°F). Launched in 2003, it observes from behind a sunscreen. In fact, it could not observe from Earth orbit because Earth is so hot, so the telescope was sent into an orbit around the sun that will carry it slowly away from Earth as its coolant is used up. Named after theoretical physicist Lyman Spitzer Jr., it has made important discoveries concerning star formation, planets orbiting other stars, distant galaxies, and more.

High-Energy Astrophysics

High-energy astrophysics refers to the use of X-ray and gamma-ray observations of the sky. Making such observations is difficult but can reveal the secrets of processes such as the collapse of massive stars and eruptions of supermassive black holes.

The first astronomical satellite, Ariel 1, was launched by British astronomers in 1962 and made solar observations in the ultraviolet and X-ray part of the spectrum. Since then many space telescopes have made high-energy observations from orbit.

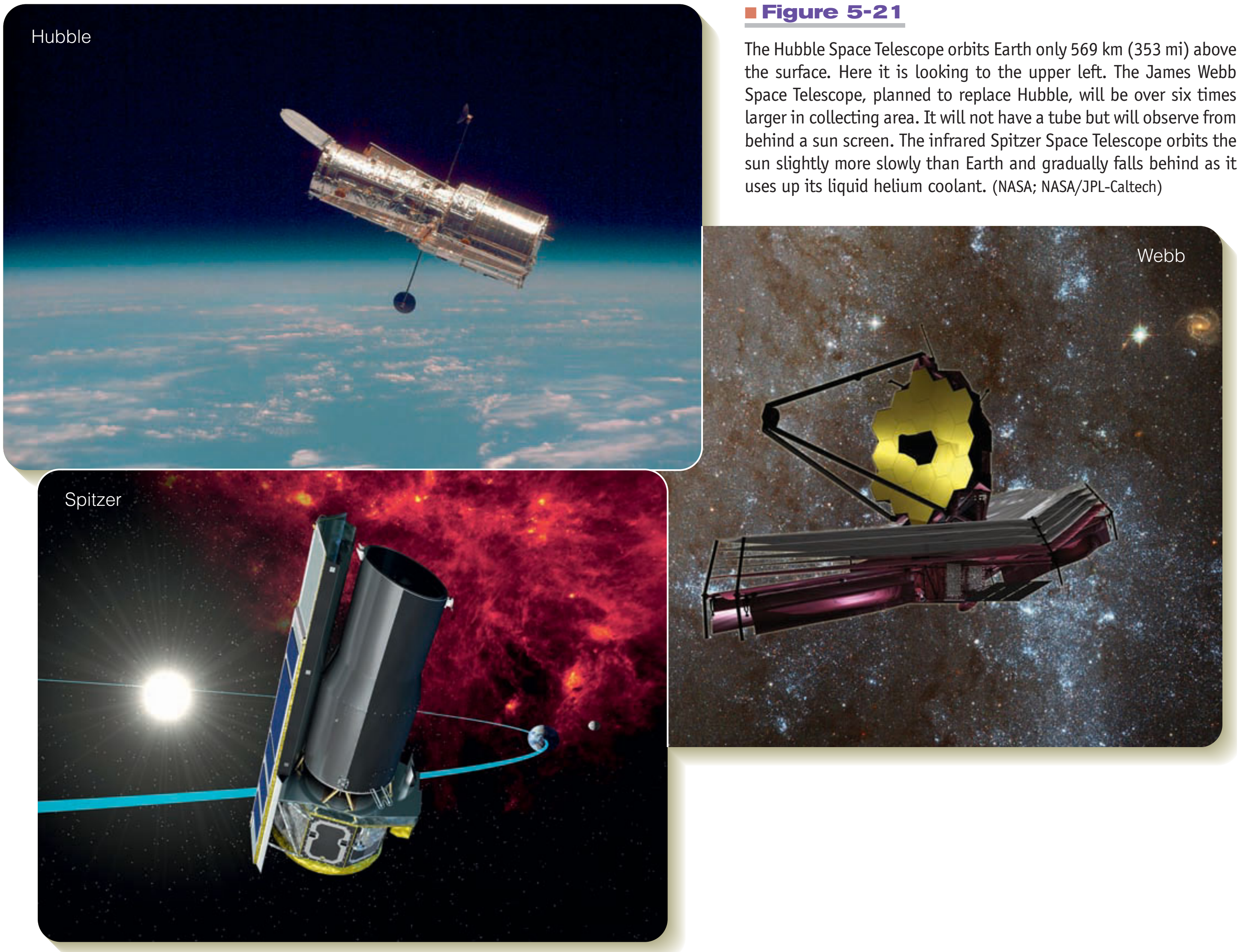


Figure 5-21

The Hubble Space Telescope orbits Earth only 569 km (353 mi) above the surface. Here it is looking to the upper left. The James Webb Space Telescope, planned to replace Hubble, will be over six times larger in collecting area. It will not have a tube but will observe from behind a sun screen. The infrared Spitzer Space Telescope orbits the sun slightly more slowly than Earth and gradually falls behind as it uses up its liquid helium coolant. (NASA; NASA/JPL-Caltech)

Some of these satellites have been general-purpose telescopes that can observe many different kinds of objects. ROSAT, for example, was an X-ray observatory developed by an international consortium of European astronomers. Some space telescopes are designed to study a single problem or a single object. The Japanese satellite Hinode, for example, studies the sun continuously at visual, ultraviolet, and X-ray wavelengths.

The largest X-ray telescope to date was launched in 1999; the Chandra X-Ray Observatory orbits a third of the way to the moon and is named for the late Indian-American Nobel laureate Subrahmanyan Chandrasekhar, who was a pioneer in many branches of theoretical astronomy. Focusing X-rays is difficult because they penetrate into most mirrors, so astronomers devised cylindrical mirrors in which the X-rays reflect from the polished inside of the cylinders and form images on special detectors. The telescope has made important discoveries about everything from star formation to monster black holes in distant galaxies (■ Figure 5-22).

One of the first gamma-ray observatories was the Compton Gamma Ray Observatory, launched in 1991. It mapped the entire sky at gamma-ray wavelengths. The European INTEGRAL satellite was launched in 2002 and has been very productive in the study of violent eruptions of stars and black holes. The GLAST (Gamma-Ray Large Area Space Telescope) launched in 2008 is capable of mapping large areas of the sky to high sensitivity.

Modern astronomy has come to depend on observations that cover the entire electromagnetic spectrum. More orbiting space telescopes are planned that will be more versatile and more sensitive.

Cosmic Rays

All of the radiation you have read about in this chapter has been electromagnetic radiation, but there is another form of energy raining down from space, and scientists aren't sure where it

Figure 5-22

From Earth orbit, the Chandra X-Ray Observatory recorded this X-ray image of the remains of a star that exploded several thousand years ago. Each color represents different-energy X-ray photons. The image is superimposed on a visual wavelength image. (X-ray: NASA/CXC/Penn Sate/S. Park et al.; Optical: Pal. Obs. DSS)

comes from. **Cosmic rays** are subatomic particles traveling through space at tremendous velocities. Almost no cosmic rays reach the ground, but they do smash gas atoms in the upper atmosphere, and fragments of those atoms shower down on you day and night over your entire life. These secondary cosmic rays are passing through you as you read this sentence.

Some cosmic-ray research can be done from high mountains or high-flying aircraft; but, to study cosmic rays in detail, detectors must go into space. A number of cosmic-ray detectors have been carried into orbit, but this area of astronomical research is just beginning to bear fruit.

Astronomers can't be sure what produces cosmic rays. Because they are atomic particles with electric charges, they are deflected by the magnetic fields spread through our galaxy, and that means astronomers can't tell where they are coming from. The space between the stars is a glowing fog of cosmic rays. Some lower-energy cosmic rays come from the sun, and observations show that at least some high-energy cosmic rays are produced by the violent explosions of dying stars and by supermassive black holes at the centers of galaxies.

At present, cosmic rays largely remain an exciting mystery. You will meet them again in future chapters.

What Are We? Curious

Telescopes are creations of curiosity. You look through a telescope to see more and to understand more. The unaided eye is a limited detector, and the history of astronomy is the history of bigger and better telescopes gathering more and more light to search for fainter and more distant objects.

The old saying, "Curiosity killed the cat," is an insult to the cat and to curiosity. We humans are curious, and curiosity is a noble trait—the mark of an active, inquiring mind. At the limits of human curiosity lies the fundamental question, "What are we?" Telescopes extend and amplify our senses, but they also extend and amplify our curiosity about the universe around us.

When people find out how something works, they say their curiosity is *satisfied*. Curiosity is an appetite like hunger or thirst, but it is an appetite for understanding. As astronomy expands our horizons and we learn how distant stars and galaxies work, we feel satisfaction because we are learning about ourselves. We are beginning to understand what we are.

Study and Review

Summary

- Light is the visible form of **electromagnetic radiation (p. 70),** an electric and magnetic disturbance that transports energy at the speed of light. The electromagnetic **spectrum (p. 71)** includes gamma rays, X-rays, ultraviolet radiation, visible light, infrared radiation, and radio waves.
- You can think of a particle of light, a **photon (p. 71),** as a bundle of waves that acts sometimes as a particle and sometimes as a wave.
- The energy a photon carries depends on its **wavelength (p. 70).** The wavelength of visible light, usually measured in **nanometers (p. 71)** (10^{-9} m) or **Ångstroms (p. 71)** (10^{-10} m), ranges from 400 nm to 700 nm. Radio and **infrared radiation (p. 72)** have longer wavelengths and carry less energy. X-ray, gamma ray, and **ultraviolet** radiation **(p. 72)** have shorter wavelengths and more energy.
- **Frequency (p. 72)** is the number of waves that pass a stationary point in 1 second. Wavelength equals the speed of light divided by the frequency.
- Earth's atmosphere is transparent in only two **atmospheric windows (p. 72)**—visible light and radio.
- Astronomical telescopes use a **primary lens** or **mirror (p. 72)** (also called an **objective lens** or **mirror [p. 72]**) to gather light and focus it into a small image, which can be magnified by an **eyepiece (p. 72).** Short-**focal-length (p. 72)** lenses and mirrors must be more strongly curved and are more expensive to grind to shape.
- A **refracting telescope (p. 72)** uses a lens to bend the light and focus it into an image. Because of **chromatic aberration (p. 73),** refracting telescopes cannot bring all colors to the same focus, resulting in color fringes around the images. An **achromatic lens (p. 73)** partially corrects for this, but such lenses are expensive and cannot be made much larger than about 1 m in diameter.
- **Reflecting telescopes (p. 72)** use a mirror to focus the light and are less expensive than refracting telescopes of the same diameter. Also, reflecting telescopes do not suffer from chromatic aberration. Most large telescopes are reflectors.
- **Light-gathering power (p. 74)** refers to the ability of a telescope to produce bright images. **Resolving power (p. 74)** refers to the ability of a telescope to resolve fine detail. **Diffraction fringes (p. 74)** in the image limit the detail visible. **Magnifying power (p. 75),** the ability to make an object look bigger, is less important because it can be changed by changing the eyepiece.
- Astronomers build observatories on remote, high mountains for two reasons. Turbulence in Earth's atmosphere blurs the image of an astronomical telescope, a phenomenon that astronomers refer to as **seeing (p. 75).** Atop a mountain, the air is steady, and the seeing is better. Observatories are located far from cities to avoid **light pollution (p. 76).**
- In reflecting telescopes, light first comes to a focus at the **prime focus (p. 80),** but **secondary mirrors (p. 80)** can direct light to other focus locations such as a **Cassegrain focus (p. 80)** or a **Newtonian focus (p. 80).** The **Schmidt-Cassegrain focus (p. 80)** is popular for small telescopes.
- Because Earth rotates, telescopes must have a **sidereal drive (p. 81)** to follow the stars. An **equatorial mounting (p. 81)** with a **polar axis (p. 81)** makes this possible, but **alt-azimuth mountings (p. 81)** are becoming more popular.
- Very large telescopes can be built with **active optics (p. 81),** maintaining the shape of floppy mirrors that are thin or in segments. Such thin mirrors weigh less, are easier to support, and cool faster at nightfall.
- High-speed **adaptive optics (p. 82)** can monitor distortions caused by turbulence in Earth's atmosphere and partially cancel out the blurring caused by seeing.
- **Interferometry (p. 83)** refers to connecting two or more separate telescopes together to act as a single large telescope that has a resolution equivalent to that of a telescope as large in diameter as the separation between the telescopes.
- For many decades astronomers used photographic plates to record images at the telescope, but modern electronic systems such as **charge-coupled devices (CCDs) (p. 84)** have replaced photographic plates in most applications.
- Astronomical images in digital form can be computer enhanced and reproduced as **false-color images (p. 84)** to bring out subtle details.
- **Spectrographs (p. 84)** using prisms or a **grating (p. 84)** spread starlight out according to wavelength to form a spectrum revealing hundreds of **spectral lines (p. 84)** produced by atoms in the object being studied. A **comparison spectrum (p. 86)** containing lines of known wavelength allows astronomers to measure wavelengths in spectra of astronomical objects.
- Astronomers use radio telescopes for three reasons: They can detect cool hydrogen and other atoms and molecules in space; they can see through dust clouds that block visible light; and they can detect certain objects invisible at other wavelengths.
- Most radio telescopes contain a dish reflector, an antenna, an amplifier, and a data recorder. Such a telescope can record the intensity of the radio energy coming from a spot on the sky. Scans of small regions are used to produce radio maps.
- Because of the long wavelength, radio telescopes have very poor resolution, and astronomers often link separate radio telescopes together to form a **radio interferometer (p. 87)** capable of resolving much finer detail.
- Earth's atmosphere absorbs gamma rays, X-rays, ultraviolet, and far-infrared. To observe at these wavelengths, telescopes must be located in space.
- Earth's atmosphere distorts and blurs images. Telescopes in orbit are above this seeing distortion and are limited only by diffraction in their optics.
- **Cosmic rays (p. 91)** are not electromagnetic radiation; they are subatomic particles such as electrons and protons traveling at nearly the speed of light. They can best be studied from above Earth's atmosphere.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why would you not plot sound waves in the electromagnetic spectrum?
2. If you had limited funds to build a large telescope, which type would you choose, a refractor or a reflector? Why?
3. Why do nocturnal animals usually have large pupils in their eyes? How is that related to astronomical telescopes?
4. Why do optical astronomers often put their telescopes at the tops of mountains, while radio astronomers sometimes put their telescopes in deep valleys?
5. Optical and radio astronomers both try to build large telescopes but for different reasons. How do these goals differ?
6. What are the advantages of making a telescope mirror thin? What problems does this cause?
7. Small telescopes are often advertised as "200 power" or "magnifies 200 times." As someone knowledgeable about astronomical telescopes, how would you improve such advertisements?

8. Not too many years ago an astronomer said, “Some people think I should give up photographic plates.” Why might she change to something else?
9. What purpose do the colors in a false-color image or false-color radio map serve?
10. How is chromatic aberration related to a prism spectrograph?
11. Why would radio astronomers build identical radio telescopes in many different places around the world?
12. Why do radio telescopes have poor resolving power?
13. Why must telescopes observing in the far-infrared be cooled to low temperatures?
14. What might you detect with an X-ray telescope that you could not detect with an infrared telescope?
15. The moon has no atmosphere at all. What advantages would you have if you built an observatory on the lunar surface?
16. **How Do We Know?** How is the resolution of an astronomical image related to the precision of a measurement?

Discussion Questions

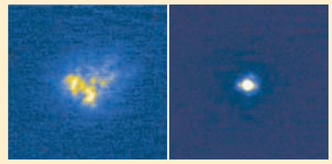
1. Why does the wavelength response of the human eye match so well the visual window of Earth’s atmosphere?
2. Most people like beautiful sunsets with brightly glowing clouds, bright moonlit nights, and twinkling stars. Astronomers don’t. Why?

Problems

1. The thickness of the plastic in plastic bags is about 0.001 mm. How many wavelengths of red light is this?
2. What is the wavelength of radio waves transmitted by a radio station with a frequency of 100 million cycles per second?
3. Compare the light-gathering powers of one of the Keck telescopes and a 0.5-m telescope.
4. How does the light-gathering power of one of the Keck telescopes compare with that of the human eye? (*Hint:* Assume that the pupil of your eye can open to about 0.8 cm.)
5. What is the resolving power of a 25-cm telescope? What do two stars 1.5 seconds of arc apart look like through this telescope?
6. Most of Galileo’s telescopes were only about 2 cm in diameter. Should he have been able to resolve the two stars mentioned in Problem 5?
7. How does the resolving power of a 5-m telescope compare with that of the Hubble Space Telescope? Why does the HST outperform a 5-m telescope?
8. If you build a telescope with a focal length of 1.3 m, what focal length should the eyepiece have to give a magnification of 100 times?
9. Astronauts observing from a space station need a telescope with a light-gathering power 15,000 times that of the human eye, capable of resolving detail as small as 0.1 second of arc and having a magnifying power of 250. Design a telescope to meet their needs. Could you test your design by observing stars from Earth?
10. A spy satellite orbiting 400 km above Earth is supposedly capable of counting individual people in a crowd. Roughly what minimum-diameter telescope must the satellite carry? (*Hint:* Use the small-angle formula.)

Learning to Look

1. The two images at the right show a star before and after an adaptive optics system was switched on. What causes the distortion in the first image, and how does adaptive optics correct the image?

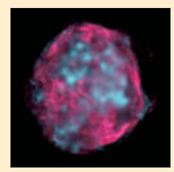


ESO

2. The star images in the photo at the right are tiny disks, but the diameter of these disks is not related to the diameter of the stars. Explain why the telescope can’t resolve the diameter of the stars.

NASA, ESA and G. Meylan

3. The X-ray image at right shows the remains of an exploded star. Explain why images recorded by telescopes in space are often displayed in false color rather than in the “colors” received by the telescope.



NASA/CXC/PSU/S. Park

6 Atoms and Starlight

Clouds of glowing gas illuminated by hot, bright stars lie thousands of light-years away, but clues hidden in starlight tell a story of star birth and star death. (ESO)

Guidepost

In the last chapter you read how telescopes gather starlight and how spectrographs spread the light out into spectra. Now you are ready to see what all the fuss is about. Here you will find answers to four essential questions:

What is an atom?

How do atoms interact with light?

What kinds of spectra do you see when you look at celestial objects?

What can you learn from a star's spectrum?

This chapter marks a change in the way you will look at nature. Up to this point, you have been thinking about what you can see with your eyes alone or aided by telescopes. In this chapter, you will begin using modern astrophysics to search out the secrets of the stars that lie beyond what you can see.

The analysis of spectra is a powerful technique, and in the chapters that follow you will use that method to study stars, galaxies, and planets.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Awake! for Morning in the Bowl of Night
Has flung the Stone that puts the Stars
to Flight:
And Lo! the Hunter of the East has caught
The Sultan's Turret in a Noose of Light.

THE RUBÁIYÁT OF OMAR KHAYYÁM,
TRANS. EDWARD FITZGERALD

THE UNIVERSE IS filled with fabulously beautiful clouds of glowing gas illuminated by brilliant stars, but they are all hopelessly beyond reach. No laboratory jar on Earth holds a sample labeled "star stuff," and no space probe has ever visited the inside of a star. The stars are far away, and the only information you can obtain about them comes hidden in starlight (■ Figure 6-1).

Earthbound humans knew almost nothing about stars until the early 19th century, when the Munich optician Joseph von Fraunhofer studied the solar spectrum and found it interrupted by some 600 dark lines. As scientists realized that the lines were related to the various atoms in the sun and found that the spectra of other stars had similar patterns of lines, the door to an understanding of stars finally opened.

■ **Figure 6-1**

What's going on here? The sky is filled with beautiful and mysterious objects that lie far beyond your reach—in the case of the nebula NGC 6751, about 6500 ly beyond your reach. The only way to understand such objects is by analyzing their light. Such an analysis reveals that this object is a dying star surrounded by the expanding shell of gas it ejected a few thousand years ago. You will learn more about this phenomenon in a later chapter. (NASA Hubble Heritage Team/STScI/AURA)

6-1 Atoms

STARS ARE GREAT balls of hot gas, and the atoms in that gas leave their marks on the light the stars emit. By understanding what atoms are and how they interact with light, you can decode the spectra of the stars and learn their secrets.

A Model Atom

To think about atoms and how they interact with light, you need a working model of an atom. In Chapter 2, you used a working model of the sky, the celestial sphere. In this chapter, you will begin your study of atoms by creating a model of an atom.

Your model atom contains a positively charged **nucleus** at the center, which consists of two kinds of particles. **Protons** carry a positive electrical charge, and **neutrons** have no charge, leaving the nucleus with a net positive charge.

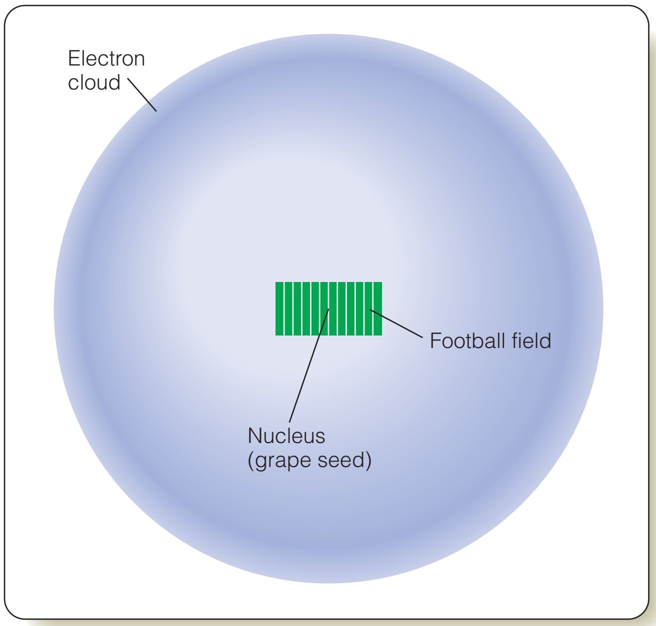
The nucleus of this model atom is surrounded by a whirling cloud of orbiting **electrons,** low-mass particles with negative charges. Normally the number of electrons equals the number of protons, and the positive and negative charges balance to produce a neutral atom. Protons and neutrons have masses about 1840 times that of an electron, so most of the mass of an atom lies in the nucleus. Even so, a single atom is not a massive object. A hydrogen atom, for example, has a mass of only 1.67×10^{-27} kg, about a trillionth of a trillionth of a gram.

An atom is mostly empty space. To see this, imagine constructing a simple scale model of a hydrogen atom. Its nucleus is a proton with a diameter of about 0.0000016 nm, or 1.6×10^{-15} m. If you multiply this by one trillion (10^{12}), you can represent the nucleus of your model atom with something about 0.16 cm in diameter—a grape seed would do. The region of a hydrogen atom that contains the whirling electron has a diameter of about 0.4 nm, or 4×10^{-10} m. Multiplying by a trillion increases the diameter to about 400 m, or about 4.5 football fields laid end to end (■ Figure 6-2). When you imagine a grape seed in the middle of a sphere 4.5 football fields in diameter, you can see that an atom is mostly empty space.

Now you can understand a **Common Misconception**. Most people, without thinking about it much, imagine that matter is solid, but you have seen that atoms are mostly empty space. The chair you sit on, the floor you walk on, are mostly not there. When you study the deaths of stars in a later chapter, you will see what happens to a star when most of the empty space gets squeezed out of its atoms.

Different Kinds of Atoms

There are over a hundred chemical elements. Which element an atom is depends only on the number of protons in the nucleus. For example, a carbon atom has six protons in its nucleus. An atom with one more proton than this is nitrogen, and an atom with one fewer proton is boron.



■ **Figure 6-2**

Magnifying a hydrogen atom by 10^{12} makes the nucleus the size of a grape seed and the diameter of the electron cloud about 4.5 times longer than a football field. The electron itself is still too small to see.

Although the number of protons in an atom of a given element is fixed, the number of neutrons is less restricted. For instance, if you added a neutron to a carbon nucleus, it would still be carbon, but it would be slightly heavier. Atoms that have the same number of protons but a different number of neutrons are **isotopes.** Carbon has two stable isotopes. One contains six protons and six neutrons for a total of 12 particles and is thus called carbon-12. Carbon-13 has six protons and seven neutrons in its nucleus.

The number of electrons in an atom of a given element can vary. Protons and neutrons are bound tightly into the nucleus, but the electrons are held loosely in the electron cloud. Running a comb through your hair creates a static charge by removing a few electrons from their atoms. An atom that has lost one or more electrons is said to be **ionized** and is called an **ion.** A neutral carbon atom has six electrons to balance the positive charge of the six protons in its nucleus. If you ionize the atom by removing one or more electrons, the atom is left with a net positive charge. Under some circumstances, an atom may capture one or more extra electrons, giving it more negative charges than positive. Such a negatively charged atom is also considered an ion.

Atoms that collide may form bonds with each other by exchanging or sharing electrons. Two or more atoms bonded together form a **molecule.** Atoms do collide in stars, but the high temperatures cause violent collisions that are unfavorable for chemical bonding. Only in the coolest stars are the collisions gentle enough to permit the formation of chemical bonds. You will see later that the presence of molecules such as titanium oxide (TiO) in a star is a clue that the star is very cool. In later chapters, you will see that molecules can form in cool gas clouds in space and in the atmospheres of planets.

Electron Shells

So far you have been thinking of the cloud of the whirling electrons in a general way, but now it is time to be more specific as to how the electrons behave within the cloud.

Electrons are bound to the atom by the attraction between their negative charge and the positive charge on the nucleus. This attraction is known as the **Coulomb force,** after the French physicist Charles-Augustin de Coulomb (1736–1806). To ionize an atom, you need a certain amount of energy to pull an electron away from the nucleus. This energy is the electron's **binding energy,** the energy that holds it to the atom.

The size of an electron's orbit is related to the energy that binds it to the atom. If an electron orbits close to the nucleus, it is tightly bound, and a large amount of energy is needed to pull it away. Consequently, its binding energy is large. An electron orbiting farther from the nucleus is held more loosely, and less energy is needed to pull it away. That means it has less binding energy.

Nature permits atoms only certain amounts (quanta) of binding energy, and the laws that describe how atoms behave are called the laws of **quantum mechanics** (■ How Do We Know? 6-1). Much of this discussion of atoms is based on the laws of quantum mechanics.

Because atoms can have only certain amounts of binding energy, your model atom can have orbits of only certain sizes, called **permitted orbits.** These are like steps in a staircase: You can stand on the number-one step or the number-two step, but not on the number-one-and-one-quarter step. The electron can occupy any permitted orbit but not orbits in between.

The arrangement of permitted orbits depends primarily on the charge of the nucleus, which in turn depends on the number of protons. Consequently, each kind of element has its own pattern of permitted orbits (■ Figure 6-3). Isotopes of the same elements have nearly the same pattern because they have the same number of protons. However, ionized atoms have orbital patterns that differ from their un-ionized forms. Thus the arrangement of permitted orbits differs for every kind of atom and ion.

◄ SCIENTIFIC ARGUMENT ►

How many hydrogen atoms would it take to cross the head of a pin?
This is not a frivolous question. In answering it, you will discover how small atoms really are, and you will see how powerful physics and mathematics can be as a way to understand nature. Many scientific arguments are convincing because they have the precision of mathematics. To begin, assume that the head of a pin is about 1 mm in diameter. That is 0.001 m. The size of a hydrogen atom is represented by the diameter of the electron cloud, roughly 0.4 nm. Because 1 nm equals 10^{-9} m, you can multiply and discover that 0.4 nm equals 4×10^{-10} m. To find out how many atoms would stretch 0.001 m, you can divide the diameter of the pinhead by the

How Do We Know? 6-1

Quantum Mechanics

How can you understand nature if it depends on the atomic world you cannot see? You can see objects such as stars, planets, aircraft carriers, and hummingbirds, but you can't see individual atoms. As scientists apply the principle of cause and effect, they study the natural effects they can see and work backward to find the causes. Invariably that quest for causes leads back to the invisible world of atoms.

Quantum mechanics is the set of rules that describe how atoms and subatomic particles behave. On the atomic scale, particles behave in ways that seem unfamiliar. One of the principles of quantum mechanics specifies that you cannot know simultaneously the exact location and motion of a particle. This is why physicists prefer to describe the electrons in an atom as if they were a cloud of negative charge surrounding the nucleus rather than small particles following individual orbits.

This raises some serious questions about reality. Is an electron really a particle at all? If you can't know simultaneously the position and motion of a specific particle, how can you know how it will react to a collision with a photon or another particle? The answer is that you can't know, and that seems to violate the principle of cause and effect.

Many of the phenomena you can see depend on the behavior of huge numbers of atoms, and quantum mechanical uncertainties average out. Nevertheless, the ultimate causes that scientists seek lie at the level of atoms, and modern physicists are trying to understand the nature of the particles that make up atoms. That is one of the most exciting frontiers of science.

The world you see, including these neon signs, is animated by the properties of atoms and subatomic particles. (Jeff Greenberg/PhotoEdit)

diameter of an atom. That is, divide 0.001 m by 4×10^{-10} m, and you get 2.5×10^6. It would take 2.5 million hydrogen atoms lined up side by side to cross the head of a pin.

Now you can see how tiny an atom is and also how powerful a bit of physics and mathematics can be. It reveals a view of nature beyond the capability of your eyes. Now build an argument using another bit of arithmetic. **How many hydrogen atoms would you need to add up to the mass of a paper clip (1 g)?**

◄ ►

6-2 The Interaction of Light and Matter

If light did not interact with matter, you would not be able to see these words. In fact, you would not exist, because, among other problems, photosynthesis would be impossible, and there would be no grass, wheat, bread, beef, cheeseburgers, or any other kind of food. The interaction of light and matter makes life possible, and it also makes it possible for you to understand the universe.

You should begin your study of light and matter by considering the hydrogen atom. It is both simple and common. Roughly 90 percent of all atoms in the universe are hydrogen.

The Excitation of Atoms

Each electron orbit in an atom represents a specific amount of binding energy, so physicists commonly refer to the orbits as **energy levels.** Using this terminology, you can say that an electron in its smallest and most tightly bound orbit is in its lowest permitted energy level, which is called the atom's **ground state.** You could move the electron from one energy level to another by supplying enough energy to make up the

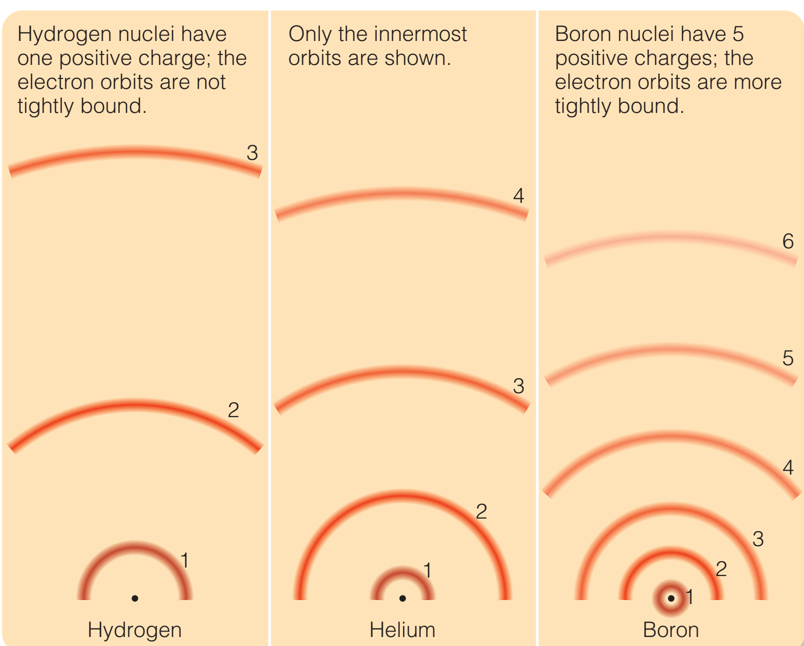


Figure 6-3

The electron in an atom may occupy only certain permitted orbits. Because different elements have different charges on their nuclei, the elements have different, unique patterns of permitted orbits.

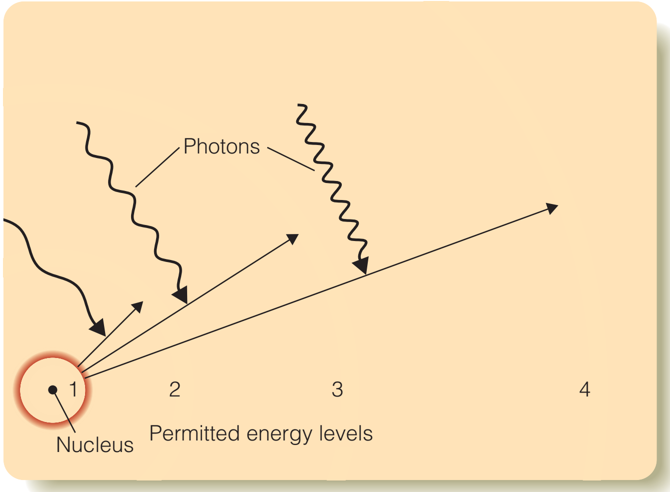


Figure 6-4

A hydrogen atom can absorb only those photons that move the atom's electron to one of the higher-energy orbits. Here three different photons are shown along with the change they would produce if they were absorbed.

difference between the two energy levels. It would be like moving a flowerpot from a low shelf to a high shelf; the greater the distance between the shelves, the more energy you would need to raise the pot. The amount of energy needed to move the electron is the energy difference between the two energy levels.

If you move the electron from a low energy level to a higher energy level, the atom becomes an **excited atom.** That is, you have added energy to the atom by moving its electron. An atom can become excited by collision. If two atoms collide, one or both may have electrons knocked into a higher energy level. This happens very commonly in hot gas, where the atoms move rapidly and collide often.

Another way an atom can become excited is to absorb a photon. Only a photon with exactly the right amount of energy can move the electron from one level to another. If the photon has too much or too little energy, the atom cannot absorb it. Because the energy of a photon depends on its wavelength, only photons of certain wavelengths can be absorbed by a given kind of atom. ■ Figure 6-4 shows the lowest four energy levels of the hydrogen atom, along with three photons the atom could absorb. The longest-wavelength photon has only enough energy to excite the electron to the second energy level, but the shorter-wavelength photons can excite the electron to higher levels. A photon with too much or too little energy cannot be absorbed. Because the hydrogen atom has many more energy levels than shown in Figure 6-4, it can absorb photons of many different wavelengths.

Atoms, like humans, cannot exist in an excited state forever. An excited atom is unstable and must eventually (usually within 10^{-6} to 10^{-9} seconds) give up the energy it has absorbed and return its electron to a lower energy level. Thus the electron in an excited atom tends to tumble down to its lowest energy level, its ground state.

When an electron drops from a higher to a lower energy level, it moves from a loosely bound level to one that is more tightly bound. The atom then has a surplus of energy—the energy difference between the levels—that it can emit as a photon. Study the sequence of events in ■ Figure 6-5 to see how an atom can absorb and emit photons. Because each type of atom or ion has its unique set of energy levels, each type absorbs and emits photons with a unique set of wavelengths. As a result, you can identify the elements in a gas by studying the characteristic wavelengths of light that are absorbed or emitted.

The process of excitation and emission is a common sight in urban areas at night. A neon sign glows when atoms of neon gas in a glass tube are excited by electricity flowing through the tube. As the electrons in the electric current flow through the gas, they collide with the neon atoms and excite them. Almost immediately after a neon atom is excited, its electron drops back to a lower energy level, emitting the surplus energy as a photon of a certain wavelength. The photons emitted by excited neon blend to produce a reddish-orange glow. Signs of other colors, erroneously called "neon," contain other gases or mixtures of gases instead of pure neon. Whenever you look at a neon sign, you are seeing atoms absorbing and emitting energy.

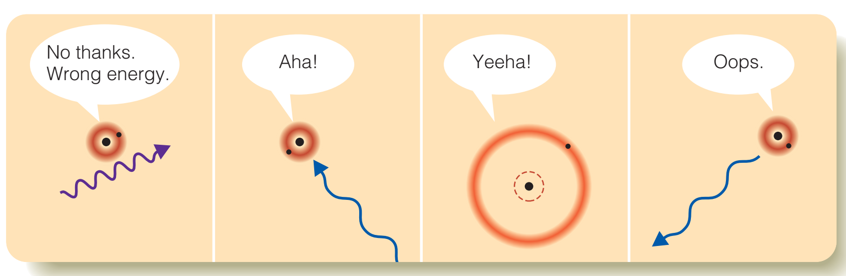
Radiation from a Heated Object

If you look closely at the stars in the constellation Orion, you will notice that they are not all the same color (see Figure 2-4). One of your Favorite Stars, Betelgeuse, in the upper left corner of Orion, is quite red; another Favorite Star, Rigel, in the lower right corner, is blue. These differences in color arise from the way the stars emit light, and as you learn why Betelgeuse is red and Rigel is blue, you will begin to see how astronomers can learn about stars by analyzing starlight.

The starlight that you see comes from gases that make up the visible surface of the star, its photosphere. (Recall that you met the photosphere of the sun in Chapter 3.) Layers of gas deeper in the star also emit light, but that light is reabsorbed before it can

Figure 6-5

An atom can absorb a photon only if the photon has the correct amount of energy. The excited atom is unstable and within a fraction of a second returns to a lower energy level, reradiating the photon in a random direction.



reach the surface. The gas above the photosphere is too thin to emit much light. The photosphere is the visible surface of a star because it is dense enough to emit lots of light but thin enough to allow that light to escape.

Stars produce their light for the same reason a heated horseshoe glows in a blacksmith's forge. If it is not too hot, the horseshoe is ruddy red, but as it heats up it grows brighter and yellower. Yellow-hot is hotter than red-hot but not as hot as white-hot. Stars produce their light the same way.

The light from stars and horseshoes is produced by moving electrons. An electron is surrounded by an electric field, and if you disturb an electron, the change in its electric field spreads outward at the speed of light as electromagnetic radiation. Whenever you change the motion of an electron, you generate electromagnetic waves. If you run a comb through your hair, you disturb electrons in both hair and comb, producing static electricity. That produces electromagnetic radiation, which you can hear as snaps and crackles if you are standing near an AM radio. Stars don't comb their hair, of course, but they are hot, and they are made up of ionized gases, so there are plenty of electrons zipping around.

The molecules and atoms in any object are in constant motion, and in a hot object they are more agitated than in a cool object. You can refer to this agitation as **thermal energy.** If you touch an object that contains lots of thermal energy it will feel hot as the thermal energy flows into your fingers. The flow of thermal energy is called **heat.** In contrast, **temperature** refers to the average speed of the particles. Hot cheese and hot green beans can have the same temperature, but the cheese can contain more thermal energy and can burn your tongue. Thus, *heat* refers to the flow of thermal energy, and *temperature* refers to the intensity of the agitation among the particles.

When astronomers refer to the temperature of a star, they are talking about the temperature of the gases in the photosphere, and they express those temperatures on the **Kelvin temperature scale.** On this scale, zero degrees Kelvin (written 0 K) is **absolute zero** (−459.7°F), the temperature at which an object contains no thermal energy that can be extracted. Water freezes at 273 K and boils at 373 K. The Kelvin temperature scale is useful in astronomy because it is based on absolute zero and consequently is related directly to the motion of the particles in an object.

Now you can understand why a hot object glows. The hotter an object is, the more motion among its particles. The agitated particles collide with electrons, and when electrons are accelerated, part of the energy is carried away as electromagnetic radiation. The radiation emitted by a heated object is called **blackbody radiation,** a name that refers to the way a perfect emitter of radiation would behave. A perfect emitter would also be a perfect absorber and at room temperature would look black. You will often see the term *black body radiation* referring to objects that glow brightly.

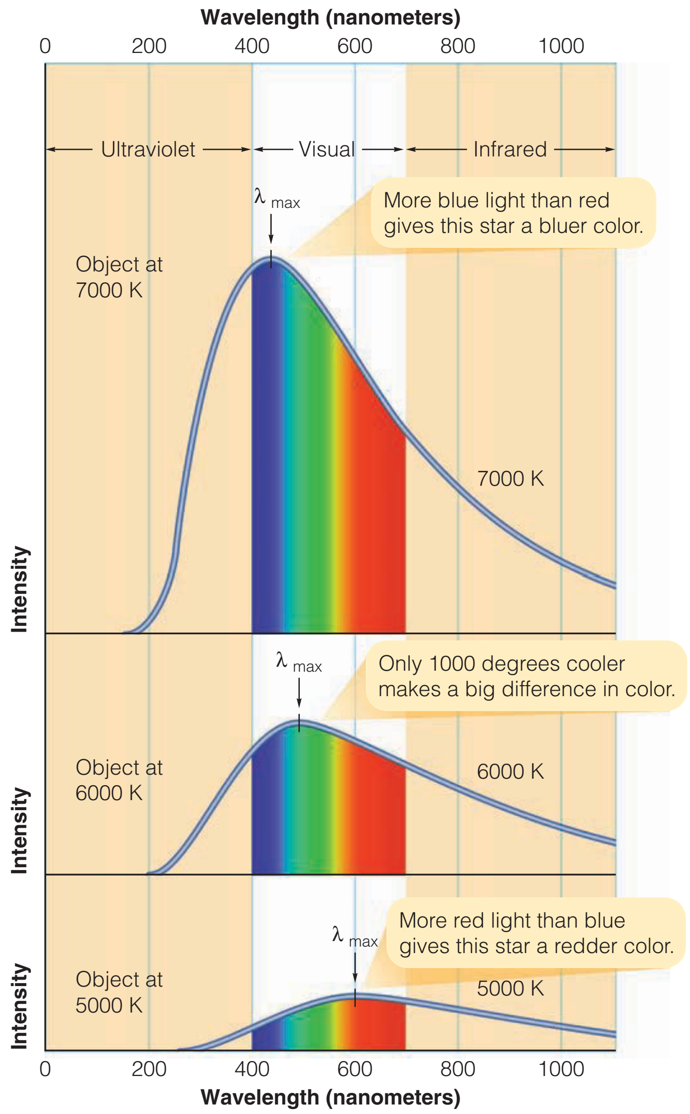
Blackbody radiation is quite common. In fact, it is responsible for the light emitted by an incandescent lightbulb. Electricity flowing through the filament of the lightbulb heats it to high temperature, and it glows. You can also recognize the light emitted by a heated horseshoe as blackbody radiation. Many objects in astronomy, including stars, emit radiation approximately as if they were blackbodies.

Hot objects emit blackbody radiation, but so do cold objects. Ice cubes are cold, but their temperature is higher than absolute zero, so they contain some thermal energy and must emit some blackbody radiation. The coldest gas drifting in space has a temperature only a few degrees above absolute zero, but it too emits blackbody radiation.

Two features of blackbody radiation are important. First, the hotter an object is, the more blackbody radiation it emits. Hot objects emit more radiation because their agitated particles collide more often and more violently with electrons. That's why a glowing coal from a fire emits more total energy than an ice cube of the same size.

The second feature is the relationship between the temperature of the object and the wavelengths of the photons it emits. The wavelength of the photon emitted when a particle collides with an electron depends on the violence of the collision. Only a violent collision can produce a short-wavelength (high-energy) photon. The electrons in an object have a distribution of speeds; a few travel very fast, and a few travel very slowly, but most travel at intermediate speeds. The hotter the object is, the faster, on average, the electrons travel. Because high-velocity electrons are rare, extremely violent collisions don't occur very often, and short-wavelength photons are rare. Similarly, most collisions are not extremely gentle, so long-wavelength (low-energy) photons are also rare. Consequently, blackbody radiation is made up of photons with a distribution of wavelengths, with medium wavelengths most common. The **wavelength of maximum intensity ($\boldsymbol{\lambda_{max}}$)** is the wavelength at which the object emits the most intense radiation and occurs at some intermediate wavelength. (Make special note that λ_{max} does not refer to the maximum wavelength but to the wavelength of maximum.)

■ Figure 6-6 shows the intensity of radiation versus wavelength for three objects of different temperatures. The curves are high in the middle and low at either end, because the objects emit most intensely at intermediate wavelengths. The total area under each curve is proportional to the total energy emitted, and you can see that the hotter object emits more total energy than the cooler objects. Look closely at the curves, and you will see that that the wavelength of maximum intensity depends on temperature. The hotter the object, the shorter the wavelength of maximum intensity. The figure shows how temperature determines the color of a glowing blackbody. The hotter object emits more blue light than red and thus looks blue, and the cooler object emits more red than blue and consequently looks red. Now you can understand why two of your Favorite Stars, Betelgeuse and Rigel, have such



■ **Figure 6-6**

Blackbody radiation from three bodies at different temperatures demonstrates that a hot body radiates more total energy and that the wavelength of maximum intensity is shorter for hotter objects. The hotter object here will look blue to your eyes, while the cooler object will look red.

Reasoning with Numbers | 6-1

Blackbody Radiation

Blackbody radiation is described by two simple laws. So many objects in astronomy behave like blackbodies that these two laws are important principles in the analysis of light from the sky.

Wien's law expresses quantitatively the relation between temperature and the wavelength of maximum. According to this law, for conventional intensity units, the wavelength of maximum intensity in nanometers, λ_{max}, equals 3,000,000 divided by the temperature in degrees Kelvin:

$$\lambda_{max} = \frac{3{,}000{,}000}{T}$$

For example, a cool star with a temperature of 3000 K will emit most intensely at a wavelength of 1000 nm, which is in the infrared part of the spectrum. A hot star with a temperature of 30,000 K will emit most intensely at a wavelength of 100 nm, which is in the ultraviolet.

The Stefan–Boltzmann law relates temperature to the total radiated energy. According to this law, the total energy radiated in 1 second from 1 square meter of an object equals a constant times the temperature raised to the fourth power:*

$$E = \sigma T^4 \ (\text{J/s/m}^2)$$

Here the temperature is expressed in degrees Kelvin and the energy in units called *joules*. One **joule (J)** is about the energy of an apple falling from a table to the floor. This law shows how strongly the energy radiated depends on temperature. If you doubled an object's temperature, for instance, it would radiate not 2 times, but rather 2^4, or 16, times more energy per second from each square meter of its surface. A small change in temperature can make a big difference to the brightness of a star.

*For the sake of completeness, note that the constant σ equals 5.67×10^{-8} J/m^2 s K^4.

different colors. Betelgeuse is cool and looks red, but Rigel is hot and looks blue. The properties of blackbody radiation are described in ■ Reasoning with Numbers 6-1.

Cool objects don't glow at visible wavelengths but still produce blackbody radiation. For example, the human body has a temperature of 310 K and emits blackbody radiation mostly in the infrared part of the spectrum. Infrared security cameras can detect burglars by the radiation they emit, and mosquitoes can track you down in total darkness by homing in on your infrared radiation. Although you emit lots of infrared radiation, you rarely emit higher-energy photons; and you almost never emit an X-ray or gamma-ray photon. Your wavelength of maximum intensity lies in the infrared part of the spectrum.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Blackbody" and "Stefan–Boltzmann Law."

◄ SCIENTIFIC ARGUMENT ►

The infrared radiation coming out of your ear can tell a doctor your temperature. How does that work?

You know two radiation laws, so your argument must use the right one. Doctors and nurses use a handheld device to measure body temperature by observing the infrared radiation emerging from a patient's ear. You might suspect the device depends on the Stefan–Boltzmann law and measures the intensity of the infrared radiation. A person with a fever will emit more energy than a healthy person. However, a healthy person with a large ear canal would emit more than a person with a small ear canal, so measuring intensity would not be accurate. The device actually depends on Wien's law in that it measures the "color" of the infrared radiation. A patient with a fever will emit at a slightly shorter wavelength of maximum intensity, and the infrared radiation emerging from his or her ear will be a tiny bit "bluer" than normal.

Astronomers can measure the temperatures of stars the same way. Adapt your argument for stars. **Use Figure 6-6 to explain how the colors of stars reveal their temperatures.**

◄ ►

6-3 Stellar Spectra

SCIENCE IS A way of understanding nature, and the spectrum of a star tells you a great deal about its temperature, motion, and composition. In later chapters, you will use spectra to study many more astronomical objects such as galaxies and planets, but you can begin with the spectra of stars, including that of the sun.

The Formation of a Spectrum

The spectrum of a star is formed as light passes outward through the gases near its surface. Read ■ **Atomic Spectra** on pages 102–103 and notice that it describes three important properties of spectra and defines 12 new terms that will help you discuss astronomical spectra:

1. There are three kinds of spectra: *continuous spectra; absorption* or *dark-line spectra,* which contain *absorption lines:* and *emission* or *bright-line spectra,* which contain *emission lines.* These spectra are described by *Kirchhoff's laws.* When you see one of these types of spectra, you can recognize the kind of matter that emitted the light.
2. Photons are emitted or absorbed when an electron in an atom makes a *transition* from one energy level to another. The wavelengths of the photons depend on the energy difference between the two levels. Hydrogen atoms can produce many spectral lines in series such as the *Lyman, Balmer,* and *Paschen* series. Only three lines in the Balmer series are visible to human eyes. The emitted photons coming from a hot cloud of hydrogen gas have the same wavelengths as the photons absorbed by hydrogen atoms in the gases of a star.
3. Most modern astronomy books display spectra as graphs of intensity versus wavelength. Be sure you see the connection between dark absorption lines and dips in the graphed spectrum.

Whatever kind of spectrum astronomers look at, the most common spectral lines are the Balmer lines of hydrogen. In the next section, you will see how Balmer lines can tell you the temperature of a star's surface.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Emission and Absorption Spectra."

The Balmer Thermometer

You can use the Balmer absorption lines as a thermometer to find the temperatures of stars. From the discussion of black body radiation, you already know how to estimate temperature from color, but the Balmer lines give you much greater accuracy.

Recall that astronomers use the Kelvin temperature scale when referring to stellar temperatures. These temperatures range from 40,000 K to 2000 K and refer to the temperature of the star's surface. The centers of stars are much hotter—millions of degrees—but the colors and spectra of stars tell you only about the surface because that's where the light comes from.

The Balmer thermometer works because the strength of the Balmer lines depends on the temperature of the star's surface layers. Both hot and cool stars have weak Balmer lines, but medium-temperature stars have strong Balmer lines.

The Balmer absorption lines are produced only by atoms with electrons in the second energy level. If the star is cool, there are few violent collisions between atoms to excite the electrons, so the electrons of most atoms are in the ground state. Electrons in the ground state can't absorb photons in the Balmer series. As a result, you should expect to find weak Balmer absorption lines in the spectra of cool stars.

In the surface layers of hot stars, on the other hand, there are many violent collisions between atoms. These collisions can excite electrons to high energy levels or ionize some atoms by knocking the electron out of the atoms. Consequently, there are few hydrogen atoms with their electrons in the second orbit to form Balmer absorption lines. Hot stars, like cool stars, have weak Balmer absorption lines.

In stars of an intermediate temperature, roughly 10,000 K, the collisions are just right to excite large numbers of electrons into the second energy level. The gas absorbs Balmer wavelength photons very well and produces strong Balmer lines.

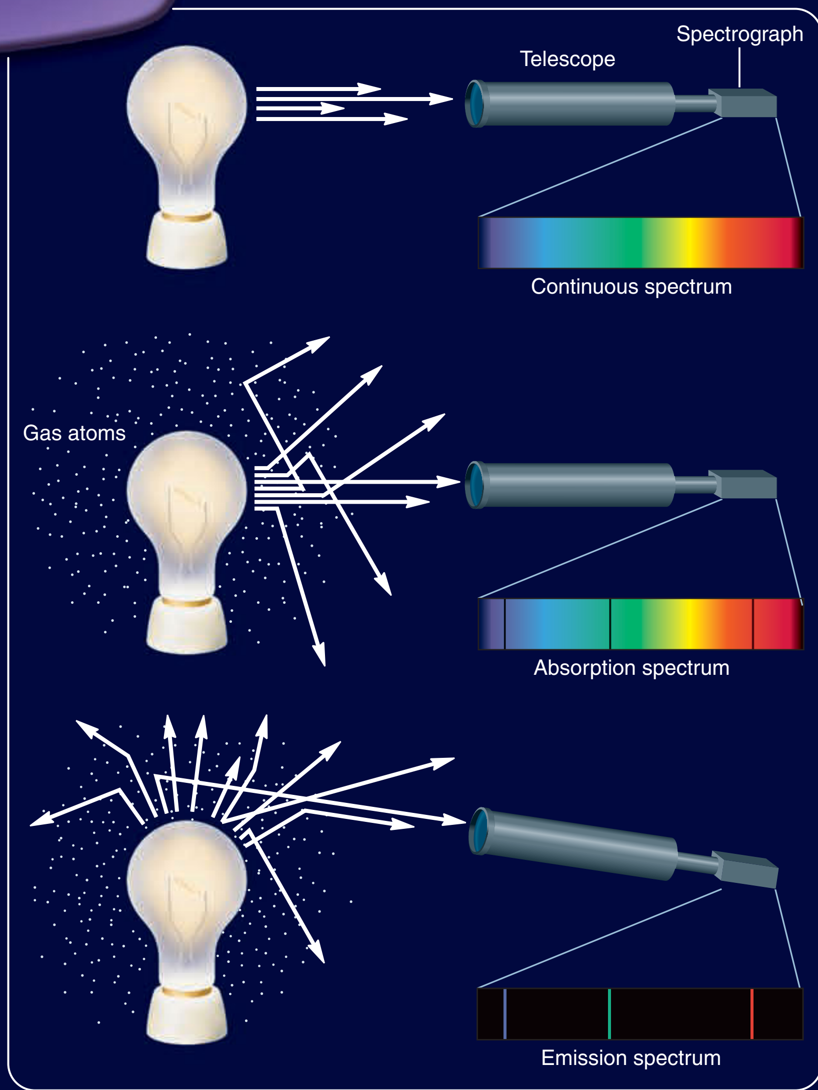
Theoretical calculations can predict just how strong the Balmer lines should be for stars of various temperatures. Such calculations are the key to finding temperatures from stellar spectra. The curve in ■ Figure 6-7a shows the strength of the Balmer lines for various stellar temperatures. But you can see from the graph that a star with Balmer lines of a certain strength might have either of two temperatures, one high and one low. How do you know which is right? You must examine other spectral lines to choose the correct temperature.

You have seen how the strength of the Balmer lines depends on temperature. Temperature has a similar effect on the spectral lines of other elements, but the temperature at which the lines reach their maximum strength differs for each element (Figure 6-7b). If you add a number of chemical elements to your graph, you get a powerful aid for finding the stars' temperatures (Figure 6-7c).

Now you can determine a star's temperature by comparing the strengths of its spectral lines with your graph. For instance, if

Atomic Spectra

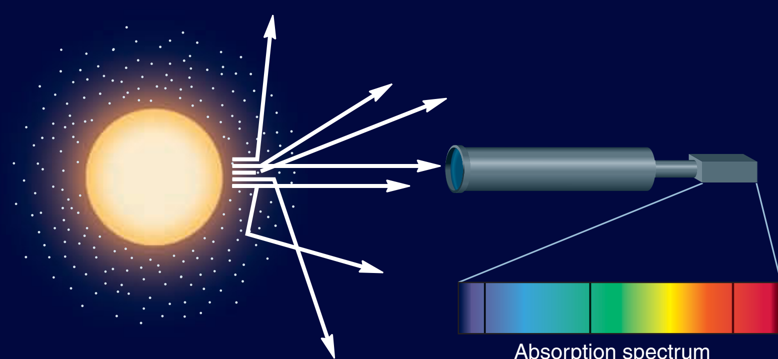
1 To understand how to analyze a spectrum, begin with a simple incandescent lightbulb. The hot filament emits blackbody radiation, which forms a **continuous spectrum.**



An **absorption spectrum** results when radiation passes through a cool gas. In this case you can imagine that the lightbulb is surrounded by a cool cloud of gas. Atoms in the gas absorb photons of certain wavelengths, which are missing from the spectrum, and you see their positions as dark **absorption lines.** Such spectra are sometimes called **dark-line spectra.**

An **emission spectrum** is produced by photons emitted by an excited gas. You could see **emission lines** by turning your telescope aside so that photons from the bright bulb did not enter the telescope. The photons you would see would be those emitted by the excited atoms near the bulb. Such spectra are also called **bright-line spectra.**

1a The spectrum of a star is an absorption spectrum. The denser layers of the photosphere emit blackbody radiation. Gases in the atmosphere of the star absorb their specific wavelengths and form dark absorption lines in the spectrum.



1b In 1859, long before scientists understood atoms and energy levels, the German scientist Gustav Kirchhoff formulated three rules, now known as **Kirchhoff's laws,** that describe the three types of spectra.

KIRCHHOFF'S LAWS

Law I: The Continuous Spectrum

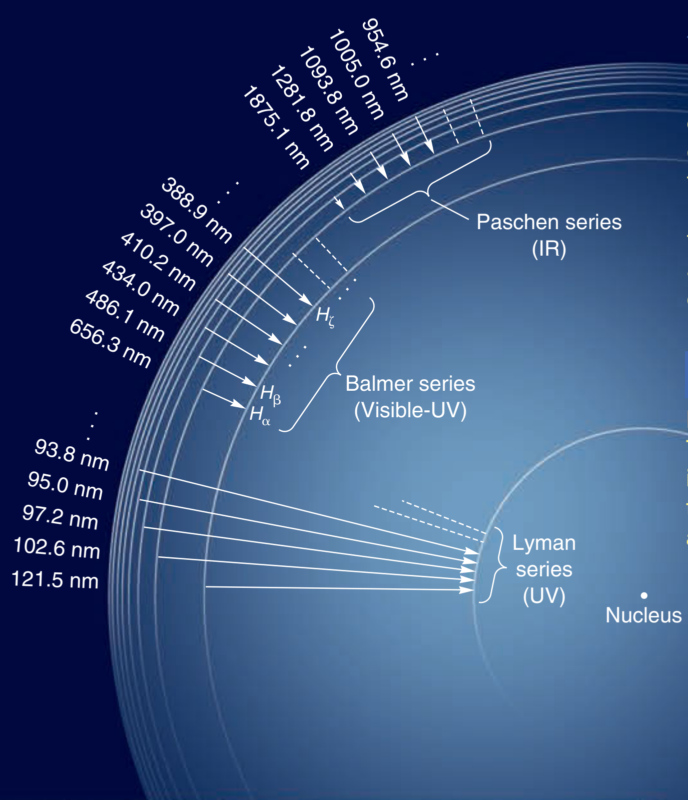
A solid, liquid, or dense gas excited to emit light will radiate at all wavelengths and thus produce a continuous spectrum.

Law II: The Emission Spectrum

A low-density gas excited to emit light will do so at specific wavelengths and thus produce an emission spectrum.

Law III: The Absorption Spectrum

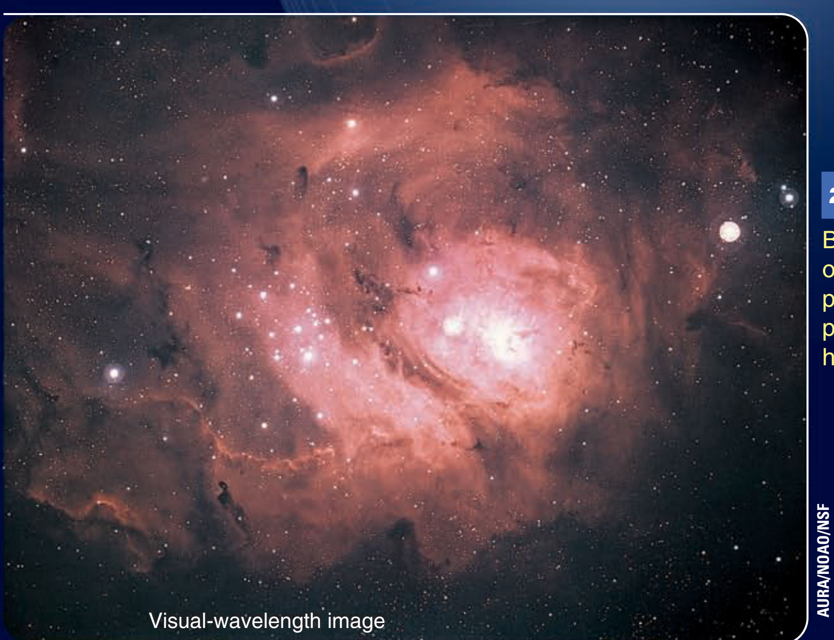
If light comprising a continuous spectrum passes through a cool, low-density gas, the result will be an absorption spectrum.



2 The electron orbits in the hydrogen atom are shown here as energy levels. When an electron makes a **transition** from one orbit to another, it changes the energy stored in the atom. In this diagram, arrows pointed inward represent transitions that result in the emission of a photon. If the arrows pointed outward, they would represent transitions that result from the absorption of a photon. Long arrows represent large amounts of energy and correspondingly short-wavelength photons.

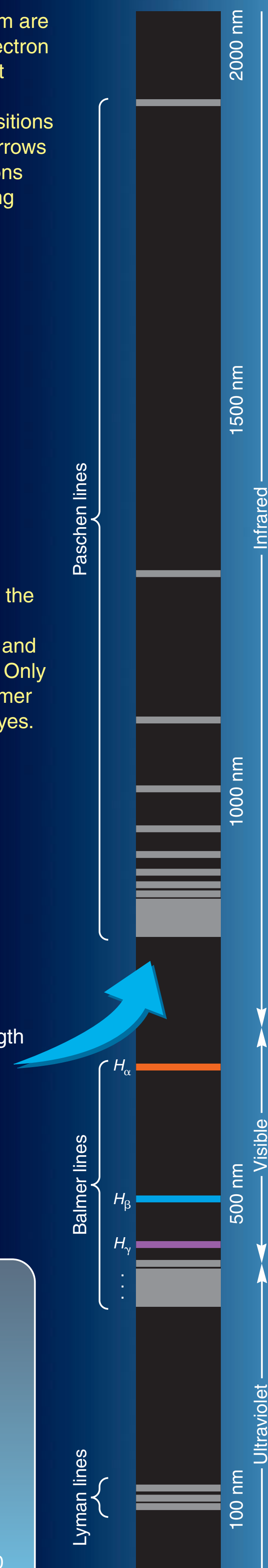
2a Transitions in the hydrogen atom can be grouped into series—the **Lyman series, Balmer series, Paschen series,** and the like. Transitions and the resulting spectral lines are identified by Greek letters. Only the first few transitions in the first three series are shown at left.

2b In this drawing (right) of the hydrogen spectrum, emission lines in the infrared and ultraviolet are shown as gray. Only the first three lines of the Balmer series are visible to human eyes.



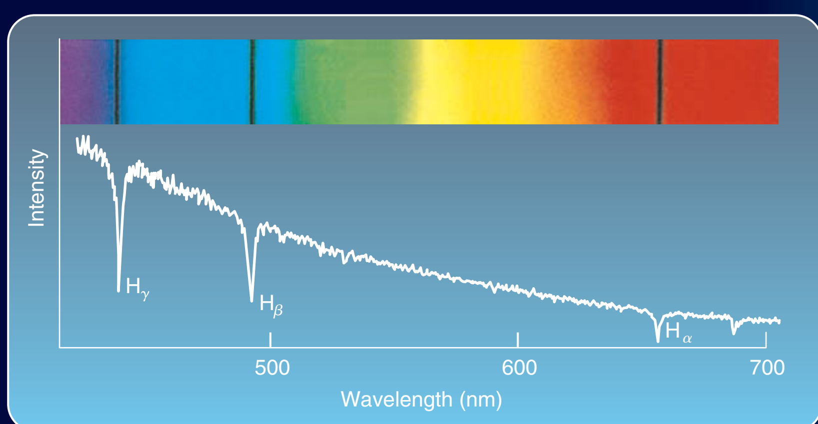
Visual-wavelength image

2c Excited clouds of gas in space emit light at all of the Balmer wavelengths, but you see only the red, blue, and violet photons blending to create the pink color typical of ionized hydrogen.



The shorter-wavelength lines in each series blend together.

3 Modern astronomers rarely work with spectra as bands of light. Spectra are usually recorded digitally, so it is easy to represent them as graphs of intensity versus wavelength. Here the artwork above the graph suggests the appearance of a stellar spectrum. The graph below reveals details not otherwise visible and allows comparison of relative intensities. Notice that dark absorption lines in the spectrum appear as dips in the curve of intensity.



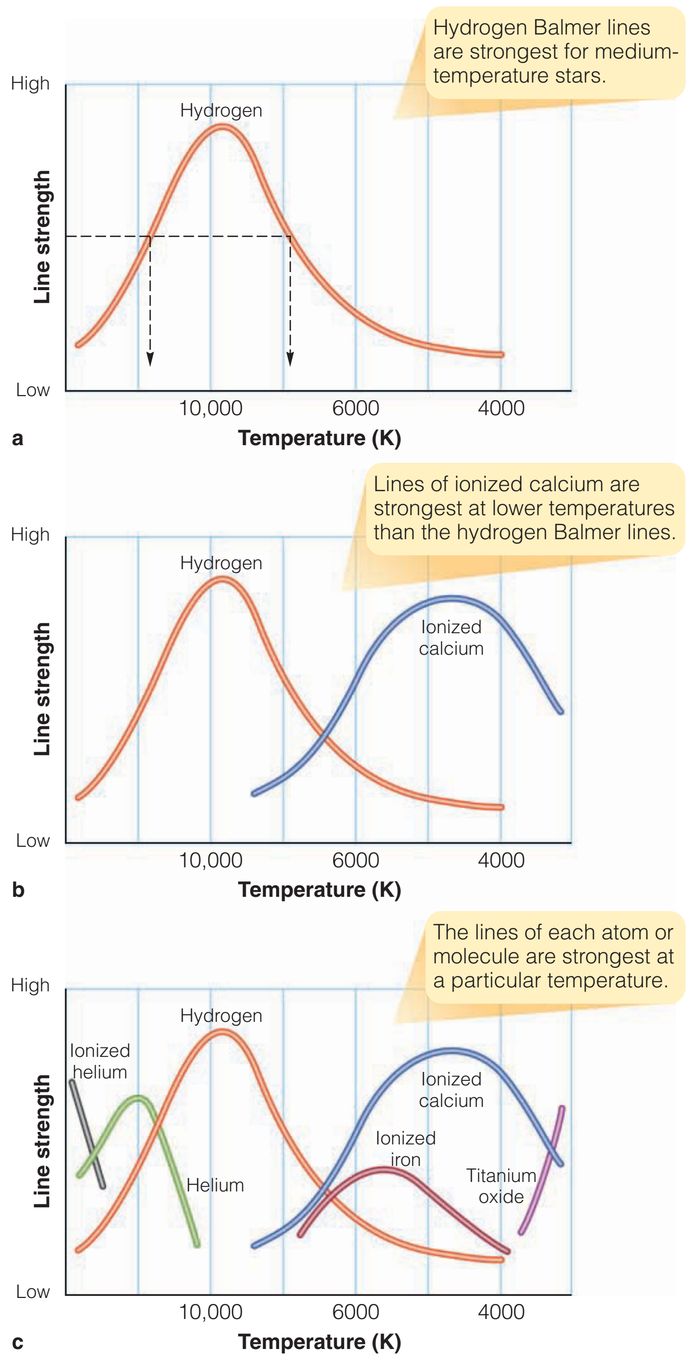


Figure 6-7

The strength of spectral lines can tell you the temperature of a star. (a) Balmer hydrogen lines alone are not enough because they give two answers. Balmer lines of a certain strength could be produced by a hotter star or a cooler star. (b) Adding another atom to the diagram helps, and (c) adding many atoms and molecules to the diagram creates a precise aid to find the temperatures of stars.

you recorded the spectrum of a star and found medium-strength Balmer lines and strong helium lines, you could conclude that it had a temperature of about 20,000 K. But if the star had weak hydrogen lines and strong lines of ionized iron, you would assign it a temperature of about 5800 K, similar to that of the sun.

The spectra of stars cooler than about 3000 K contain dark bands produced by molecules such as titanium oxide (TiO). Because of their structure, molecules can absorb photons at many wavelengths, producing numerous, closely spaced spectral lines that blend together to form bands. These molecular bands appear in the spectra of only the coolest stars because, as mentioned before, molecules in cool stars are not subject to the violent collisions that would break them apart in hotter stars.

From stellar spectra, astronomers have found that the hottest stars have surface temperatures above 40,000 K and the coolest about 2000 K. Compare these with the surface temperature of the sun, about 5800 K.

Spectral Classification

You have seen that the strengths of spectral lines depend on the surface temperature of the star. From this you can conclude that all stars of a given temperature should have similar spectra. If you learn to recognize the pattern of spectral lines produced by a 6000 K star, for instance, you need not use Figure 6-7c every time you see that kind of spectrum. You can save time by classifying stellar spectra rather than analyzing each one individually.

The first widely used classification system was devised by astronomers at Harvard during the 1890s and 1900s. One of the astronomers, Annie J. Cannon, personally inspected and classified the spectra of over 250,000 stars. The spectra were first classified into groups labeled A through Q, but some groups were later dropped, merged with others, or reordered. The final classification includes the seven major **spectral classes,** or **types,** still used today: O, B, A, F, G, K, M.*

This sequence of spectral types, called the **spectral sequence,** is important because it is a temperature sequence. The O stars are the hottest, and the temperature decreases along the sequence to the M stars, the coolest. For maximum precision, astronomers divide each spectral class into ten subclasses. For example, spectral class A consists of the subclasses A0, A1, A2, . . . A8, A9. Next come F0, F1, F2, and so on. This finer division gives a star's temperature to an accuracy within about 5 percent. The sun, for example, is not just a G star, but a G2 star. ■ Table 6-1 breaks down some of the information in Figure 6-7c and presents it in tabular form according to spectral class. For example, if a star has weak Balmer lines and lines of ionized helium, it must be an O star.

Thirteen stellar spectra are arranged in ■ Figure 6-8 from the hottest at the top to the coolest at the bottom. You can easily see how the strength of spectral lines depends on temperature. The Balmer lines are strongest in A stars, where the temperature is moderate but still high enough to excite the electrons in hydro-

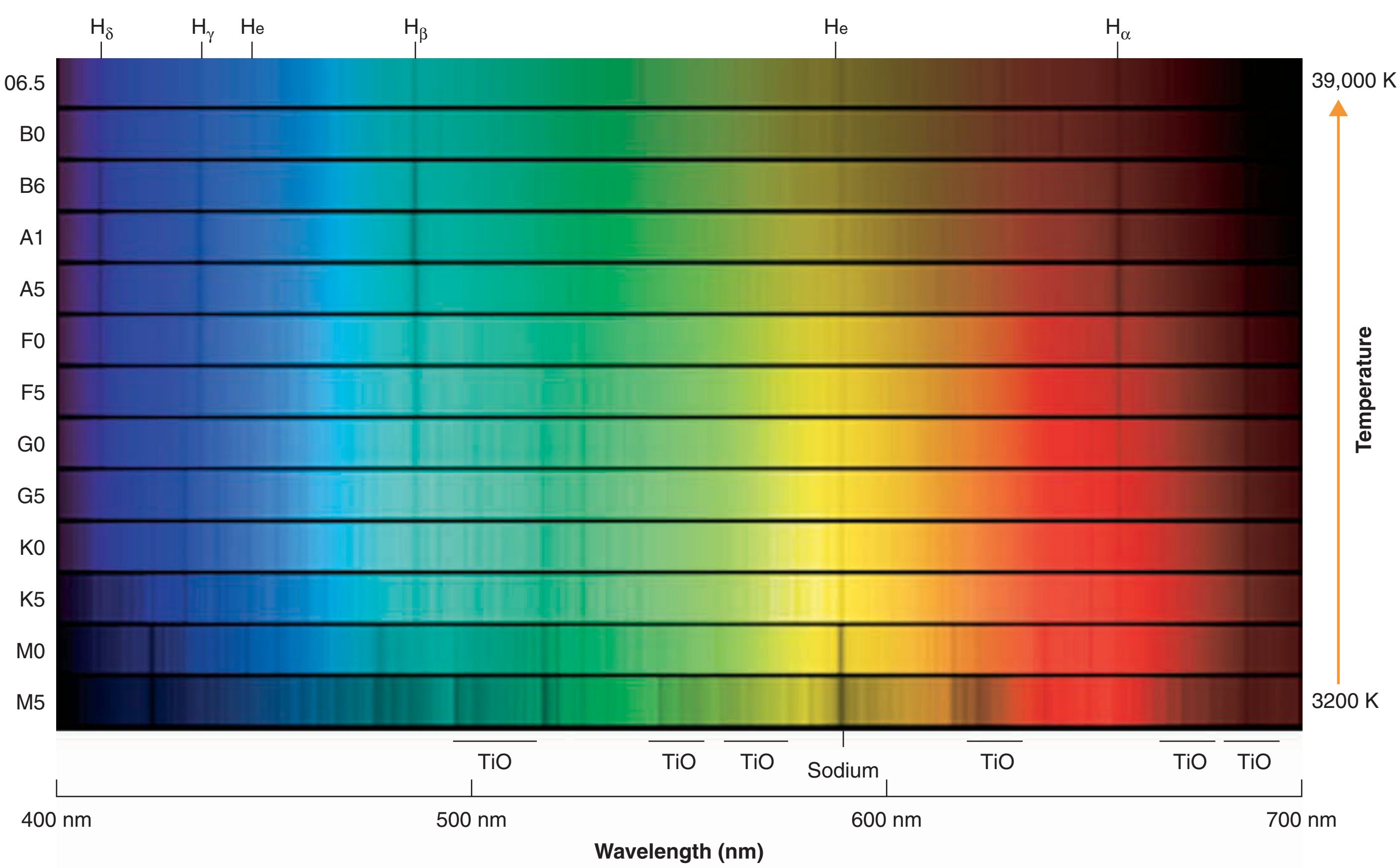
*Generations of astronomy students have remembered the spectral sequence using the mnemonic "Oh, Be A Fine Girl (Guy), Kiss Me." More recent suggestions from students include "Oh Boy, An F Grade Kills Me" and "Only Bad Astronomers Forget Generally Known Mnemonics."

Table 6-1 | Spectral Classes

Spectral Class	Approximate Temperature (K)	Hydrogen Balmer Lines	Other Spectral Features	Naked-Eye Example
O	40,000	Weak	Ionized helium	Meissa (O8)
B	20,000	Medium	Neutral helium	Achernar (B3)
A	10,000	Strong	Ionized calcium weak	Sirius (A1)
F	7500	Medium	Ionized calcium weak	Canopus (F0)
G	5500	Weak	Ionized calcium medium	Sun (G2)
K	4500	Very weak	Ionized calcium strong	Arcturus (K2)
M	3000	Very weak	TiO strong	Betelgeuse (M2)

gen atoms to the second energy level, where they can absorb Balmer wavelength photons. In the hotter stars (O and B), the Balmer lines are weak because the higher temperature excites the electrons to energy levels above the second or ionizes the atoms. The Balmer lines in cooler stars (F through M) are also weak but for a different reason. The lower temperature cannot excite many electrons to the second energy level, so few hydrogen atoms are capable of absorbing Balmer wavelength photons.

Although these spectra are attractive, astronomers rarely work with spectra as color images. Rather, they display spectra as graphs of intensity versus wavelength that show dark absorption lines as dips in the graph (■ Figure 6-9). Such graphs allow more detailed



■ **Figure 6-8**

These spectra show stars from hot O stars at the top to cool M stars at the bottom. The Balmer lines of hydrogen are strongest about A0, but the two closely spaced lines of sodium in the yellow are strongest for very cool stars. Helium lines appear only in the spectra of the hottest stars. Notice that the helium line visible in the top spectrum has nearly but not exactly the same wavelength as the sodium lines visible in cooler stars. Bands produced by the molecule titanium oxide are strong in the spectra of the coolest stars. (AURA/NOAO/NSF)

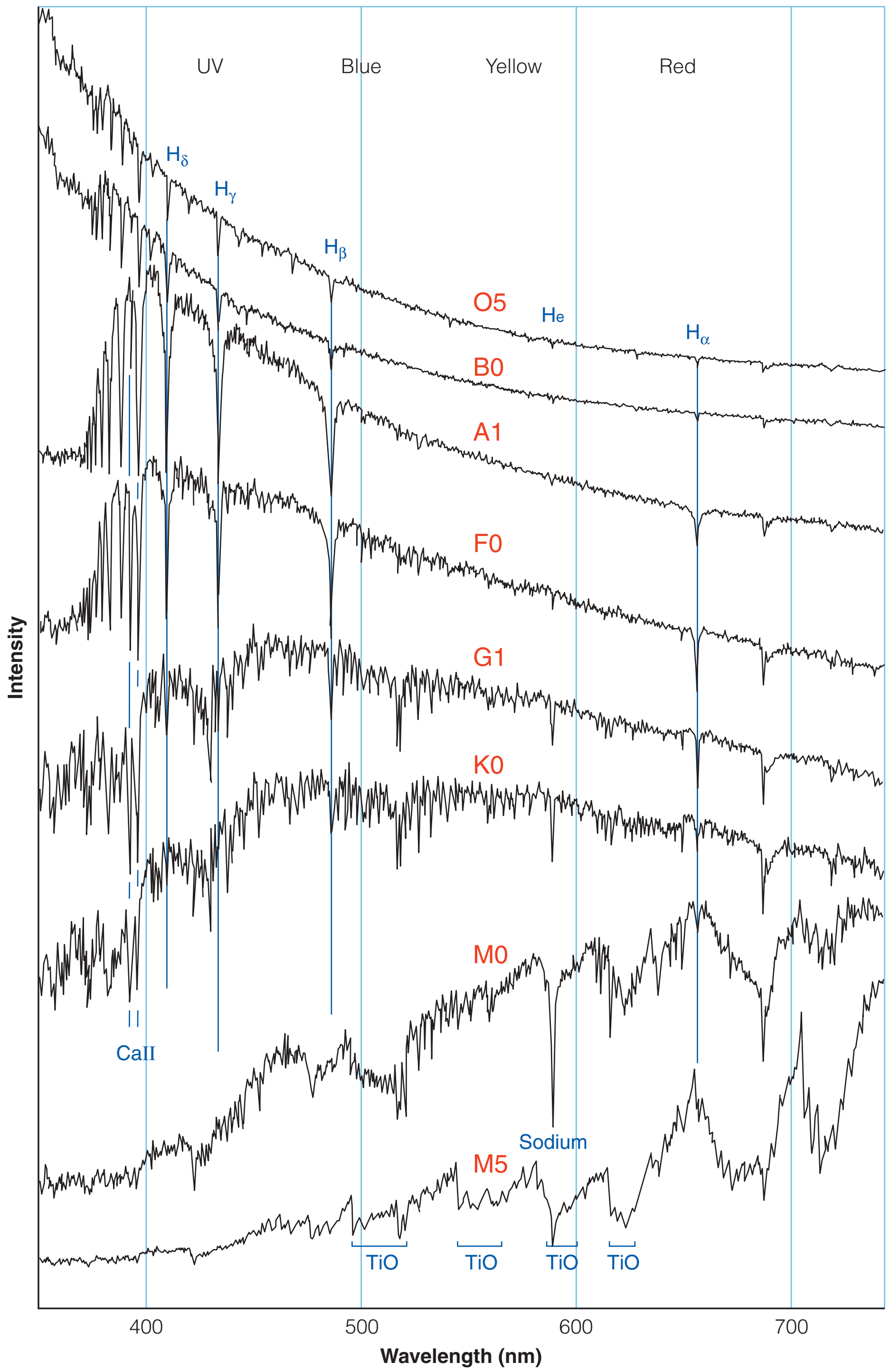


Figure 6-9

Modern digital spectra are often represented as graphs of intensity versus wavelength with dark absorption lines appearing as sharp dips in the curves. The hottest stars are at the top and the coolest at the bottom. Hydrogen Balmer lines are strongest at about A0, while lines of ionized calcium (CaII) are strong in K stars. Titanium oxide (TiO) bands are strongest in the coolest stars. Compare these spectra with Figures 6-7c and 6-8. (Courtesy NOAO, G. Jacoby, D. Hunter, and C. Christian)

analysis than photographs. Notice, for example, that the overall curves are similar to blackbody curves. The wavelength of maximum intensity is in the infrared for the coolest stars and in the ultraviolet for the hottest stars. Look carefully at these graphs, and you can see that helium is visible only in the spectra of the hottest classes and titanium oxide bands only in the coolest. Two lines of ionized calcium increase in strength from A to K and then decrease from K through M. Because the strengths of these spectral lines depend on temperature, it requires only a few moments to study a star's spectrum and determine its temperature.

Now you can learn something new about your Favorite Stars. Sirius, brilliant in the winter sky, is an A1 star; and Vega, bright overhead in the summer sky, is an A0 star. They have nearly the same temperature and color, and both have strong Balmer lines in their spectra. The bright red star in Orion is Betelgeuse, a cool M2 star, but blue-white Rigel is a hot B8 star. Polaris, the North Star, is an F8 star a bit hotter than our sun, and Alpha Centauri, the closest star to the sun, seems to be a G2 star just like the sun.

The study of spectral types is a century old, but astronomers continue to discover new types of stars. The **L dwarfs,** found in 1998, are cooler and fainter than M stars. The spectra of L dwarfs show that they are clearly a different type of star. The spectra of M stars contain bands produced by metal oxides such as titanium oxide (TiO), but L dwarf spectra contain bands produced by molecules such as iron hydride (FeH). The **T dwarfs,** discovered in 2000, are even cooler and fainter than L dwarfs. Their spectra show absorption by methane (CH_4) and water vapor (■ Figure 6-10). The development of giant telescopes and highly sensitive infrared cameras and spectrographs is allowing astronomers to find and study these coolest of stars.

Chemical Composition

Identifying the elements that are present in a star by identifying the lines in the star's spectrum is a relatively straightforward procedure. For example, two dark absorption lines appear in the yellow region of the solar spectrum at the wavelengths 589 nm

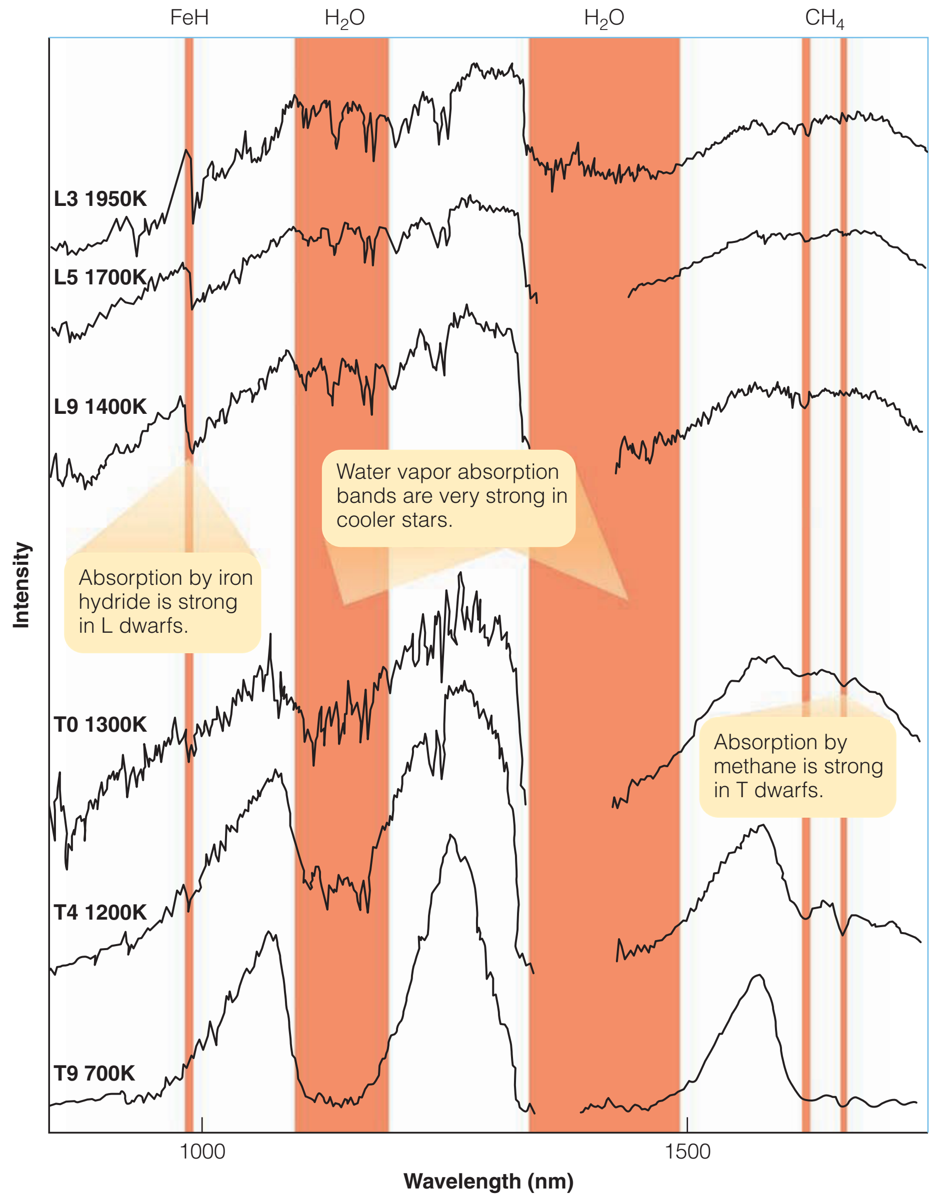


Figure 6-10

These six infrared spectra show the dramatic differences between L dwarfs and T dwarfs. Spectra of M stars show titanium oxide bands (TiO), but L and T dwarfs are so cool that TiO molecule lines are not prominent. Other molecules such as iron hydride (FeH), water (H_2O), and methane (CH_4) can form in these very cool stars. (Adapted from Thomas R. Geballe, Gemini Observatory, from a graph that originally appeared in *Sky and Telescope Magazine,* February 2005, p. 37.)

and 589.6 nm. The only atom that can produce this pair of lines is sodium, so the sun must contain sodium. Over 90 elements in the sun have been identified this way.

However, just because the spectral lines characteristic of an element are missing, you cannot conclude that the element itself is absent. For example, the hydrogen Balmer lines are weak in the sun's spectrum, even though 90 percent of the atoms in the sun are hydrogen. This is because the sun is too cool to produce strong Balmer lines. Astronomers must consider that an element's spectral lines may be absent from a star's spectrum because the star is too cool or too hot to excite those atoms to the energy levels that produce visible spectral lines.

To derive accurate chemical abundances, astronomers must use the physics that describes the interaction of light and matter to analyze a star's spectrum, take into account the star's temperature, and calculate the amounts of the elements present in the star. Such results show that nearly all stars have compositions similar to the sun's—about 91 percent of the atoms are hydrogen, and 8.9 percent are helium, with small traces of heavier elements (■ Table 6-2). You will use these results in later chapters when you study the life stories of the stars, the history of our galaxy, and the origin of the universe.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Stellar Atomic Absorption Lines."

The Doppler Effect

Surprisingly, one of the pieces of information hidden in a spectrum is the velocity of the light source. Astronomers can measure the wavelengths of the lines in a star's spectrum and find the velocity of the star. The **Doppler effect** is the apparent change in the wavelength of radiation caused by the motion of the source.

When astronomers talk about the Doppler effect, they are talking about a shift in the wavelength of electromagnetic radiation. But the Doppler shift can occur in all forms of wave phenomena, including sound waves, so you probably hear the Doppler effect every day without noticing.

The pitch of a sound is determined by its wavelength. Sounds with long wavelengths have low pitches, and sounds with short wavelengths have higher pitches. You hear a Doppler shift every time a car or truck passes you and the pitch of its engine noise drops. Its sound is shifted to shorter wavelengths and higher pitches while it is approaching and is shifted to longer wavelengths and lower pitches after it passes.

To see why the sound waves are shifted in wavelength, consider a fire truck approaching you with a bell clanging once a second. When the bell clangs, the sound travels ahead of the truck to reach your ears. One second later, the bell clangs again, but, during that one second, the fire truck has moved closer to you, so the bell is closer at its second clang. Now the sound has

Table 6-2 | The Most Abundant Elements in the Sun

Element	Percentage by Number of Atoms	Percentage by Mass
Hydrogen	91.0	70.9
Helium	8.9	27.4
Carbon	0.03	0.3
Nitrogen	0.008	0.1
Oxygen	0.07	0.8
Neon	0.01	0.2
Magnesium	0.003	0.06
Silicon	0.003	0.07
Sulfur	0.002	0.04
Iron	0.003	0.1

a shorter distance to travel and reaches your ears a little sooner than it would have if the fire truck were not approaching. If you timed the clangs, you would find that you heard them slightly less than one second apart. After the fire truck passes you and is moving away, you hear the clangs sounding slightly more than one second apart, because now each successive clang of the bell occurs farther from you and the sound travels farther to reach your ears.

Figure 6-11a shows a fire truck moving toward one observer and away from another observer. The position of the bell at each clang is shown by a small black bell. The sound of the clangs spreading outward is represented by black circles. You can see how the clangs are squeezed together ahead of the fire truck and stretched apart behind.

Now you can substitute a source of light for the clanging bell (Figure 6-11b). Imagine the light source emitting waves continuously as it approaches you. Each time the source emits the peak of a wave, it will be slightly closer to you than when it emitted the peak of the previous wave. From your vantage point, the successive peaks of the wave will seem closer together in the same way that the clangs of the bell seemed closer together. The light will appear to have a shorter wavelength, making it slightly bluer. Because the light is shifted slightly toward the blue end of the spectrum, this is called a **blueshift.** After the light source has passed you and is moving away, the peaks of successive waves seem farther apart, so the light has a longer wavelength and is redder. This is a **redshift.** The shifts are much too small to change the color of a star, but they are easily detected in spectra.

The terms *redshift* and *blueshift* are used to refer to any range of wavelengths. The light does not actually have to be red or blue, and the terms apply equally to wavelengths in other parts of the electromagnetic spectrum such as X-rays and radio waves. *Red* and *blue* refer to the direction of the shift, not to actual color.

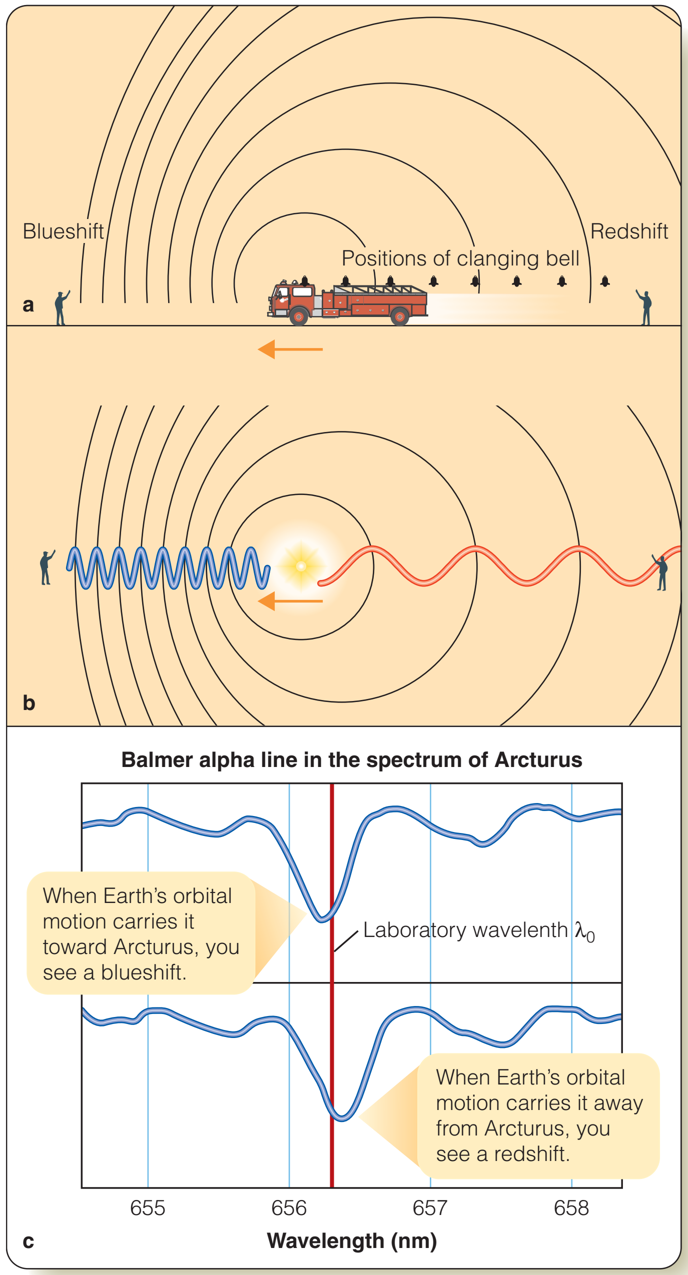


Figure 6-11

The Doppler effect. (a) The clanging bell on a moving fire truck produces sounds that move outward (black circles). An observer ahead of the truck hears the clangs closer together, while an observer behind the truck hears them farther apart. (b) A moving source of light emits waves that move outward (black circles). An observer in front of the light source observes a shorter wavelength (a blueshift), and an observer behind the light source observes a longer wavelength (a redshift). (c) Absorption lines in the spectrum of the bright star Arcturus are shifted to the blue in winter, when Earth's orbital motion carries it toward the star, and to the red in summer when Earth moves away from the star.

The amount of change in wavelength, and thus the magnitude of the Doppler shift, depends on the velocity of the source. A moving car has a smaller Doppler shift than a jet plane, and a slow-moving star has a smaller Doppler shift than one that is moving more quickly. You can measure the velocity of a star by measuring the size of its Doppler shift. Police measure Doppler shifts of passing cars by using radar guns, and astronomers measure the shift of dark lines in a stars' spectrum. ■ Reasoning with Numbers 6-2 shows you how to make a Doppler shift calculation.

When you think about the Doppler effect, it is important to remember two things. Earth itself moves, so a measurement of a Doppler shift really measures the relative motion between Earth and the star. Figure 6-11c shows the Doppler effect in two spectra of the star Arcturus. Lines in the top spectrum are slightly blueshifted because the spectrum was recorded when Earth, in the course of its orbit, was moving toward Arcturus. Lines in the bottom spectrum are redshifted because it was recorded six months later, when Earth was moving away from Arcturus.

The second point to remember is that the Doppler shift is sensitive only to the part of the velocity directed away from you or toward you. This is the **radial velocity (V_r).** You cannot use the Doppler effect to detect any part of the velocity that is perpendicular to your line of sight. A star moving to the left, for example, would have no blueshift or redshift because its distance from Earth would not be decreasing or increasing. This is why police using radar guns park right next to the highway. They want to measure your full velocity as you drive down the highway, not just part of your velocity. This is shown ■ Figure 6-12.

Reasoning with Numbers | 6-2

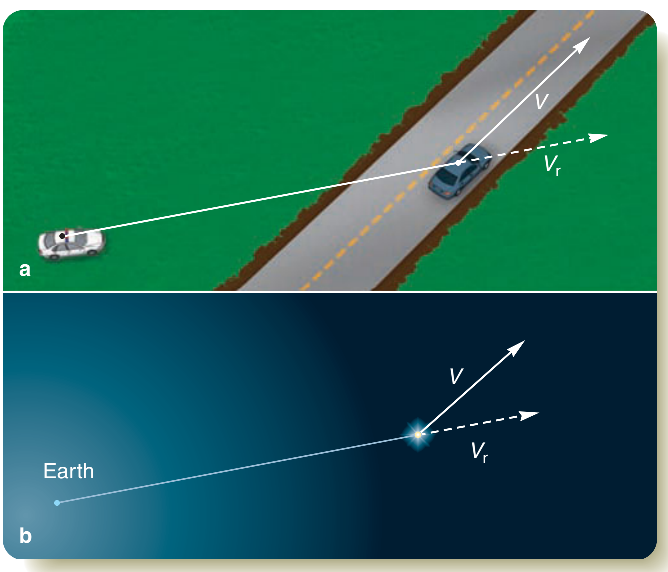
The Doppler Formula

Astronomers can measure radial velocity by using the Doppler effect. The laboratory wavelength λ_0 is the wavelength a certain spectral line would have in a laboratory where the source of the light is not moving. In the spectrum of a star, this spectral line is shifted by some small amount $\Delta\lambda$. If the wavelength is increased (a redshift), $\Delta\lambda$ is positive; if the wavelength is decreased (a blueshift), $\Delta\lambda$ is negative. The radial velocity, V_r, of the star is given by the Doppler formula:

$$\frac{V_r}{c} = \frac{\Delta\lambda}{\lambda_0}$$

That is, the radial velocity divided by the speed of light, c, is equal to $\Delta\lambda$ divided by λ_0. In astronomy, radial velocities are almost always given in kilometers per second, so c is expressed as 300,000 km/s.

For example, suppose the laboratory wavelength of a certain spectral line is 600.00 nm, and the line is observed in a star's spectrum at a wavelength of 600.10 nm. Then $\Delta\lambda$ is +0.10 nm, and the velocity is 0.10/600 multiplied by the speed of light. The radial velocity equals 50 km/s. Because $\Delta\lambda$ is positive, you know the star is receding from you.



■ **Figure 6-12**

(a) Police radar can measure only the radial part of your velocity (V_r) as you drive down the highway, not your true velocity along the pavement (V). That is why police using radar never park far from the highway. (b) From Earth, astronomers can use the Doppler effect to measure the radial velocity (V_r) of a star, but they cannot measure its true velocity, V, through space.

What Are We? Stargazers

Do you suppose chickens ever look at the sky and wonder what the stars are? Probably not. Chickens are very good at the chicken business, but they are not known for big brains and deep thought. Humans, in contrast, have highly evolved, sophisticated brains and are extremely curious. In fact, curiosity may be the most reliable characteristic of intelligence, and curiosity about the stars is a natural extension of our continual attempts to understand the world around us.

For early astronomers like Copernicus and Kepler, the stars were just points of light. There seemed to be no way to learn anything about them. Galileo's telescope revealed surprising details about the planets, but even viewed through a large telescope, the stars are just points of light. Even when later astronomers began to assume that the stars were other suns, the stars seemed forever beyond human knowledge.

As you have seen, the key is understanding how light interacts with matter. In the last 150 years or so, scientists have discovered how atoms and light interact to form spectra, and astronomers have applied those discoveries to the ultimate object of human curiosity—the stars.

Chickens may never wonder what the stars are, or even wonder what chickens are, but humans are curious animals, and we do wonder about the stars and about ourselves. Our yearning to understand the stars is just part of our quest to understand what we are.

Study and Review

Summary

- An atom consists of a **nucleus (p. 95)** surrounded by a cloud of **electrons (p. 95).** The nucleus is made up of positively charged **protons (p. 95)** and uncharged **neutrons (p. 95).**
- The number of protons in an atom determines which element it is. Atoms of the same element (that is, having the same number of protons) with different numbers of neutrons are called **isotopes (p. 96).**
- A neutral atom is surrounded by a number of negatively charged electrons equal to the number of protons in the nucleus. An atom that has lost or gained an electron is said to be **ionized (p. 96)** and is called an **ion (p. 96).**
- Two or more atoms joined together form a **molecule (p. 96).**
- The electrons in an atom are attracted to the nucleus by the **Coulomb force (p. 96).** As described by **quantum mechanics (p. 96)**, the **binding energy (p. 96)** that holds electrons in at atom is limited to certain energies, and that means the electrons may occupy only certain **permitted orbits (p. 96).**
- The size of an electron's orbit depends on its energy, so the orbits can be thought of as **energy levels (p. 97)** with the lowest possible energy level known as the **ground state (p. 97).**
- An **excited atom (p. 98)** is one in which an electron is raised to a higher orbit by a collision between atoms or the absorption of a photon of the proper energy.
- The agitation among the atoms and molecules of an object is called **thermal energy (p. 99)**, and the flow of thermal energy is **heat (p. 99).** In contrast, **temperature (p. 99)** refers to the intensity of the agitation and is expressed on the **Kelvin temperature scale (p. 99)**, which gives temperature above **absolute zero (p. 99).**
- Collisions among the particles in a body accelerate electrons and cause the emission of **blackbody radiation (p. 99).** The hotter an object is, the more it radiates and the shorter is its **wavelength of maximum intensity,** λ_{max} **(p. 99).** This allows astronomers to estimate the temperatures of stars from their colors.
- One **joule (J) (p. 100)** is about the energy of an apple falling from a table to the floor.
- **Kirchhoff's laws (p. 102)** explain that a hot solid, liquid, or dense gas emits electromagnetic radiation at all wavelengths and produces a **continuous spectrum (p. 102).** An excited low-density gas produces an **emission (bright-line) spectrum (p. 102)** containing **emission lines (p. 102).** A light source viewed through a low-density gas produces an **absorption (dark-line) spectrum (p. 102)** containing **absorption lines (p. 102).**
- An atom can emit or absorb a photon when an electron makes a **transition (p. 103)** between orbits.
- Because orbits of only certain energies are permitted in an atom, photons of only certain wavelengths can be absorbed or emitted. Each kind of atom has its own characteristic set of spectral lines. The hydrogen atom has the **Lyman (p. 103) series** of lines in the ultraviolet, the **Balmer series (p. 103)** partially in the visible, and the **Paschen series (p. 103)** (plus others) in the infrared.
- The strength of spectral lines depends on the temperature of the star. For example, in cool stars, the Balmer lines are weak because atoms are not excited out of the ground state. In hot stars, the Balmer lines are weak because atoms are excited to higher orbits or are ionized. Only at medium temperatures are the Balmer lines strong.
- A star's **spectral class** (or **type**) **(p. 104)** is determined by the absorption lines in its spectrum. The resulting **spectral sequence (p. 104)**, OBAFGKM, is important because it is a temperature sequence. By classifying a star, the astronomer learns the temperature of the star's surface.
- Long after the spectral sequence was created, astronomers found the **L dwarfs (p. 106)** and **T dwarfs (p. 106)** at temperatures even cooler than the M stars.
- A spectrum can tell you the chemical composition of the stars. The presence of spectral lines of a certain element shows that that element must be present in the star. But you must proceed with care. Lines of a certain element may be weak or absent if the star is too hot or too cool even if the element is present in the star's atmosphere.
- The **Doppler effect (p. 107)** can provide clues to the motions of the stars. When a star is approaching, you observe slightly shorter wavelengths, a **blueshift (p. 108)**, and when it is receding, you observe slightly longer wavelengths, a **redshift (p. 108).** This Doppler effect reveals a star's **radial velocity (p. 109)**, that part of its velocity directed toward or away from Earth.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why might you say that atoms are mostly empty space?
2. What is the difference between an isotope and an ion?
3. Why is the binding energy of an electron related to the size of its orbit?
4. Explain why ionized calcium can form absorption lines, but ionized hydrogen cannot.
5. Describe two ways an atom can become excited.
6. Why do different atoms have different lines in their spectra?
7. Why does the amount of blackbody radiation emitted depend on the temperature of the object?
8. Why do hot stars look bluer than cool stars?
9. What kind of spectrum does a neon sign produce?
10. Why are Balmer lines strong in the spectra of medium-temperature stars and weak in the spectra of hot and cool stars?
11. Why are titanium oxide features visible in the spectra of only the coolest stars?
12. Explain the similarities among Table 6-1, Figure 6-7c, Figure 6-8, and Figure 6-9.
13. Explain why the presence of spectral lines of a given element in the solar spectrum tells you that element is present in the sun, but the absence of the lines would not mean the element is absent from the sun.
14. Why does the Doppler effect detect only radial velocity?
15. How can the Doppler effect explain shifts in both light and sound?
16. **How Do We Know?** How is the world you see around you determined by a world you cannot see?

Discussion Questions

1. In what ways is the model of an atom a scientific model? In what ways is it incorrect?
2. Can you think of classification systems used to simplify what would otherwise be complex measurements? Consider foods, movies, cars, grades, and clothes.

Problems

1. Human body temperature is about 310 K (98.6°F). At what wavelength do humans radiate the most energy? What kind of radiation do we emit?
2. If a star has a surface temperature of 20,000 K, at what wavelength will it radiate the most energy?
3. Infrared observations of a star show that it is most intense at a wavelength of 2000 nm. What is the temperature of the star's surface?
4. If you double the temperature of a blackbody, by what factor will the total energy radiated per second per square meter increase?
5. If one star has a temperature of 6000 K and another star has a temperature of 7000 K, how much more energy per second will the hotter star radiate from each square meter of its surface?
6. Transition *A* produces light with a wavelength of 500 nm. Transition *B* involves twice as much energy as *A*. What wavelength light does it produce?
7. Determine the temperatures of the following stars based on their spectra. Use Figure 6-7c.
   a. medium-strength Balmer lines, strong helium lines
   b. medium-strength Balmer lines, weak ionized-calcium lines
   c. strong TiO bands
   d. very weak Balmer lines, strong ionized-calcium lines
8. To which spectral classes do the stars in Problem 7 belong?
9. In a laboratory, the Balmer beta line has a wavelength of 486.1 nm. If the line appears in a star's spectrum at 486.3 nm, what is the star's radial velocity? Is it approaching or receding?
10. The highest-velocity stars an astronomer might observe have velocities of about 400 km/s. What change in wavelength would this cause in the Balmer gamma line? (*Hint:* Wavelengths are given on page 103.)

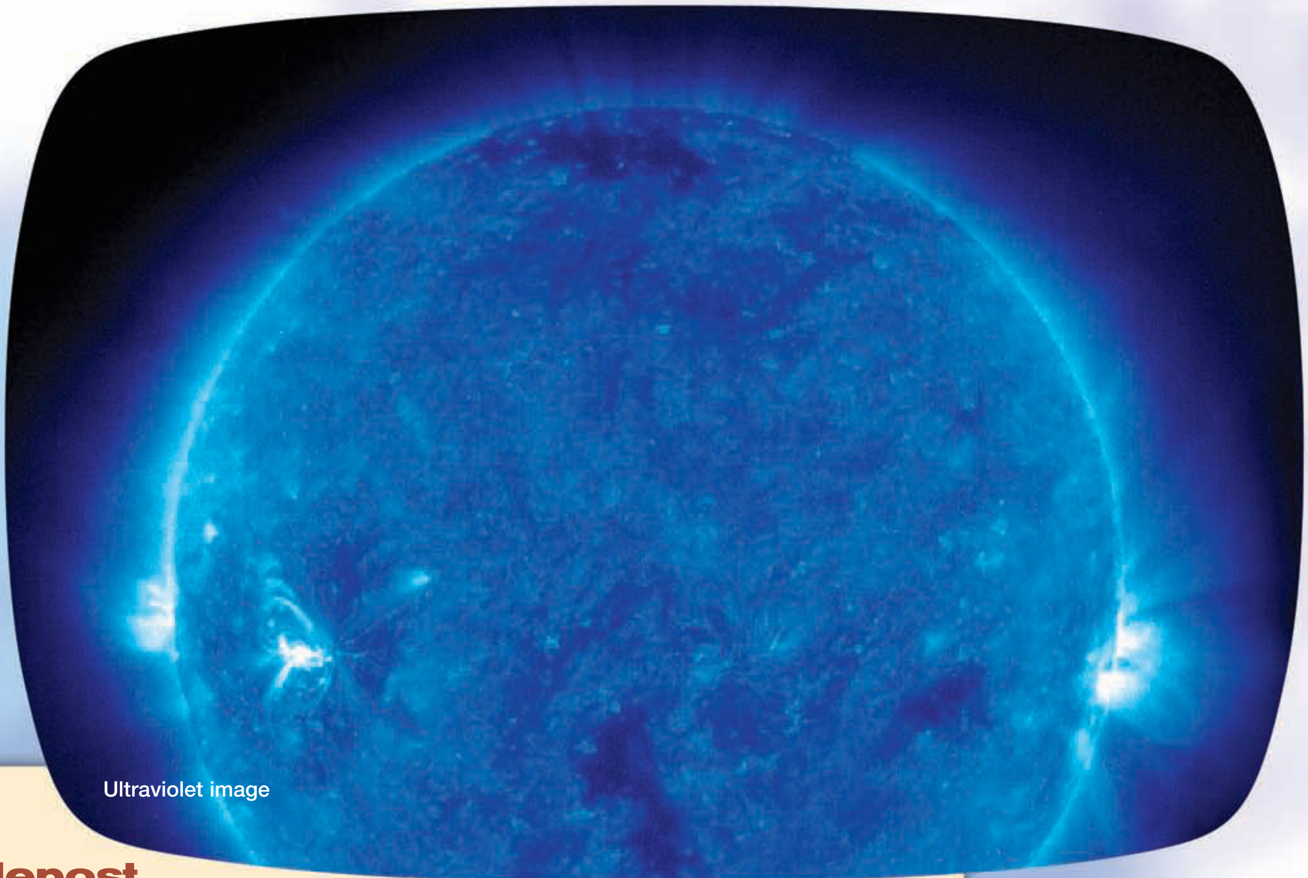
Learning to Look

1. Consider Figure 6-3. When an electron in a hydrogen atom moves from the third orbit to the second orbit, the atom emits a Balmer alpha photon in the red part of the spectrum. In what part of the spectrum would you look to find the photon emitted when an electron in a helium atom makes the same transition?
2. Where should the police car in Figure 6-12 have parked to make a good measurement?
3. The nebula shown at right contains mostly hydrogen excited to emit photons. What kind of spectrum would you expect this nebula to produce?

T. Rector, University of Alaska, and WIYN/NURO/AURA/NSF

4. If the nebula in the image above crosses in front of the star and the nebula and star have different radial velocities, what might the spectrum of the star look like?

7 The Sun



This far-ultraviolet image of the sun made from space reveals complex structure on the surface and clouds of gas being ejected into space. (NASA/SOHO)

Guidepost

The sun is the source of light and warmth in our solar system, so it is a natural object of human curiosity. It is also the one star that is most clearly visible from Earth. The interaction of light and matter, which you studied in Chapter 6, can reveal the secrets of the sun and introduce you to the stars.

In this chapter, you will discover how the analysis of the solar spectrum can paint a detailed picture of the sun's atmosphere and how basic physics has solved the mystery of the sun's core. Here you will answer four essential questions:

What do you see when you look at the sun?

How does the sun make its energy?

What are the dark sunspots?

Why does the sun go through a cycle of activity?

Although this chapter is confined to the center of the solar system, it introduces you to a star and leads your thoughts onward among the stars and galaxies that fill the universe.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

All cannot live on the piazza, but everyone may enjoy the sun.

ITALIAN PROVERB

A WIT ONCE remarked that solar astronomers would know a lot more about the sun if it were farther away. The sun is so close that Earth's astronomers can see swirling currents of gas and arched bridges of magnetic force. The details seem overwhelming. But the sun is just an average star, and in a sense, it is a simple object. It is made up almost entirely of the gases hydrogen and helium confined by its own gravity in a sphere 109 times Earth's diameter (■ Celestial Profile 1). The gases of the sun's surface are hot and radiate the light and heat that make life possible on Earth. That solar atmosphere is where you can begin your exploration.

7-1 The Solar Atmosphere

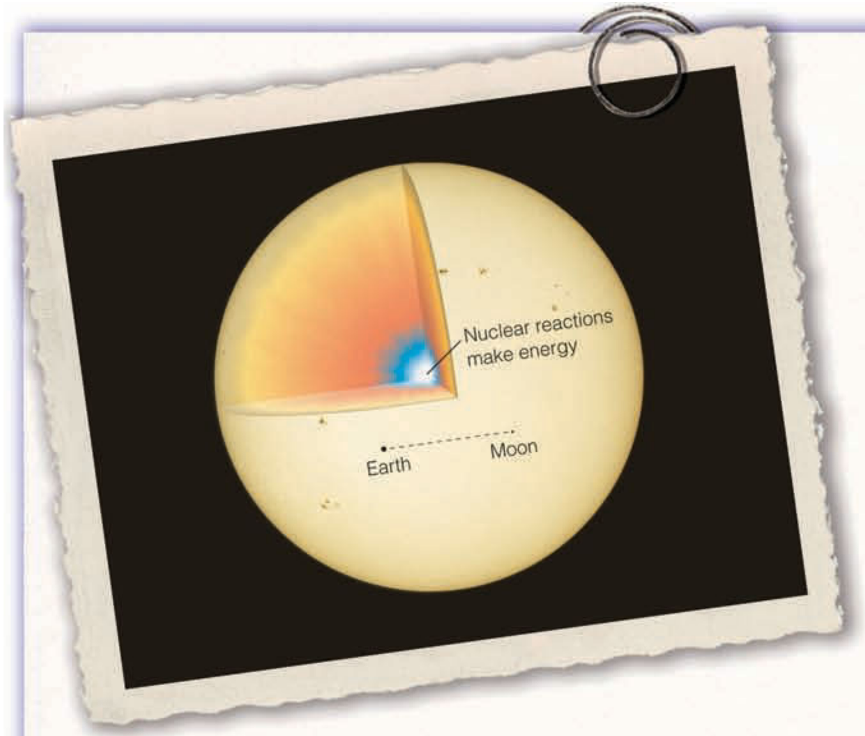
THE SUN'S ATMOSPHERE is made up of three layers. The visible surface is the photosphere, and above that lie the chromosphere and the corona. (You first met these terms in Chapter 3 when you learned about solar eclipses.)

When you look at the sun you see a hot, glowing surface with a temperature of about 5800 K. At that temperature, every square millimeter of the sun's surface must be radiating more energy than a 60-watt lightbulb. With all that energy radiating into space, the sun's surface would cool rapidly if energy did not flow up from the interior to keep the surface hot, so simple logic tells you that energy in the form of heat is flowing outward from the sun's interior. Not until the 1930s did astronomers understand that the sun makes its energy by nuclear reactions at the center. These nuclear reactions are discussed in detail later in this chapter.

For now, you can consider the sun's atmosphere in its quiescent, average state. Later you can add the details of its continuous activity as heat flows outward from its interior and it churns like a pot of boiling soup.

The Photosphere

The visible surface of the sun looks like a smooth layer of gas marked only by a few dark **sunspots** that come and go over a few weeks. Although the photosphere seems to be a distinct surface, it is not solid. In fact, the sun is gaseous from its outer atmosphere right down to its center. The photosphere is the thin layer of gas from which Earth receives most of the sun's light. It is less than 500 km deep and has an average temperature of about 5800 K. If the sun magically shrank to the size of a bowling ball, the photosphere would be no thicker than a layer of tissue paper wrapped around the ball (■ Figure 7-1).



This visible image of the sun shows a few sunspots and is cut away to show the location of energy generation at the sun's center. The Earth–moon system is shown for scale. (Daniel Good)

Celestial Profile 1: The Sun

From Earth:

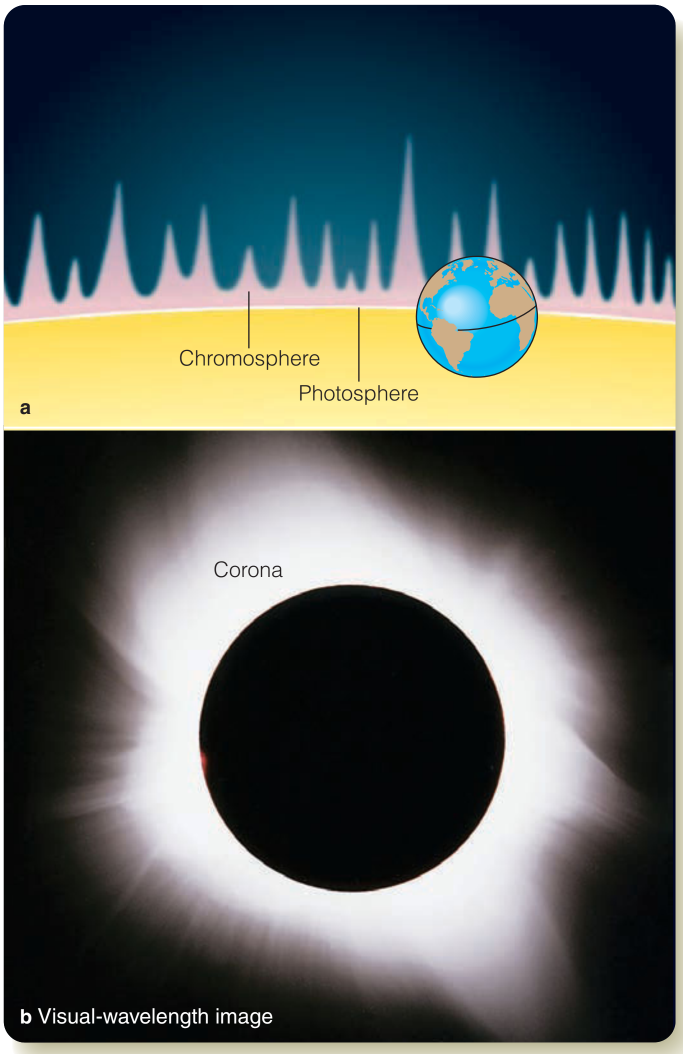
Average distance from Earth	1.00 AU (1.495979 × 10^8 km)
Maximum distance from Earth	1.0167 AU (1.5210 × 10^8 km)
Minimum distance from Earth	0.9833 AU (1.4710 × 10^8 km)
Average angular diameter	0.53° (32 minutes of arc)
Period of rotation	25.38 days at equator
Apparent visual magnitude	−26.74

Characteristics:

Radius	6.9599 × 10^5 km
Mass	1.989 × 10^{30} kg
Average density	1.409 g/cm^3
Escape velocity at surface	617.7 km/s
Luminosity	3.826 × 10^{26} J/s
Surface temperature	5800 K
Central temperature	15 × 10^6 K
Spectral type	G2 V
Absolute visual magnitude	4.83

Personality Point:

In Greek mythology, the sun was carried across the sky in a golden chariot pulled by powerful horses and guided by the sun god Helios. When Phaeton, the son of Helios, drove the chariot one day, he lost control of the horses, and Earth was nearly set ablaze before Zeus smote Phaeton from the sky. Even in classical times, people understood that life on Earth depends critically on the sun.



■ **Figure 7-1**

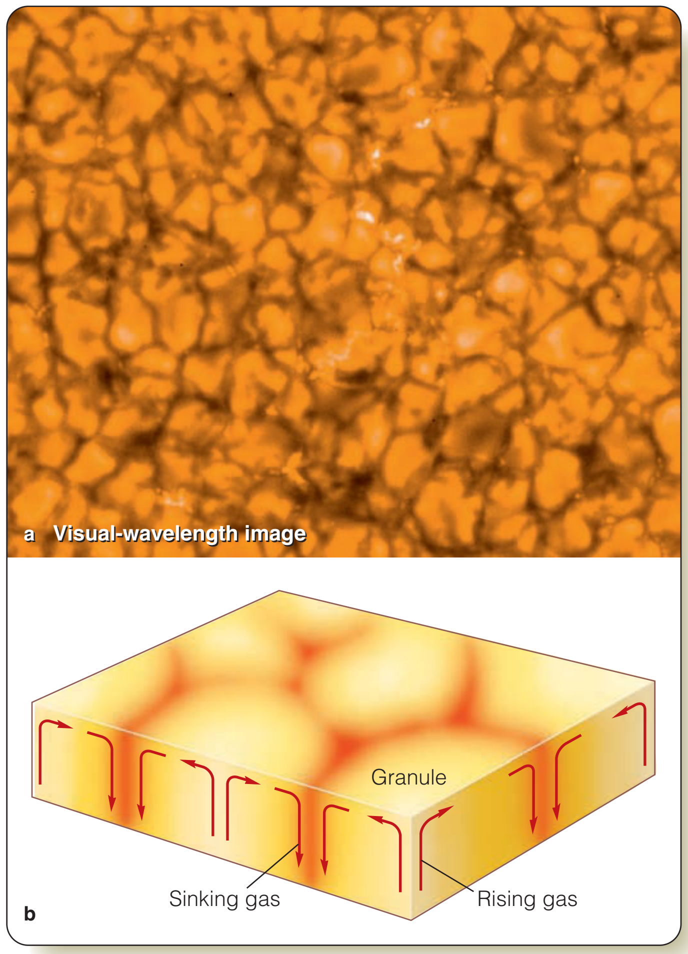
(a) A cross section at the edge of the sun shows the relative thickness of the photosphere and chromosphere. Earth is shown for scale. On this scale, the disk of the sun would be more than 1.5 m (5 ft) in diameter. The corona extends from the top of the chromosphere to great height above the photosphere. (b) This photograph, made during a total solar eclipse, shows only the inner part of the corona. (Daniel Good)

The photosphere is the layer in the sun's atmosphere that is dense enough to emit plenty of light but not so dense that the light can't escape. Below the photosphere, the gas is denser and hotter and therefore radiates plenty of light, but that light cannot escape from the sun because of the outer layers of gas. So you cannot detect light from these deeper layers. Above the photosphere, the gas is less dense and is unable to radiate much light.

Although the photosphere appears to be substantial, it is really a very-low-density gas. Even in its deepest and densest layers, the photosphere is 3400 times less dense than the air you breathe. To find gases as dense as the air at Earth's surface, you would have to descend about 70,000 km below the photosphere, about 10 percent of the way to the sun's center. With fantastically efficient insulation, you could fly a spaceship right through the photosphere.

The spectrum of the sun is an absorption spectrum, and that can tell you a great deal about the photosphere. You know from Kirchhoff's third law that an absorption spectrum is produced when a source of a continuous spectrum is viewed through a gas. In the case of the photosphere, the deeper layers are dense enough to produce a continuous spectrum, but atoms in the photosphere absorb photons of specific wavelengths, producing absorption lines of hydrogen, helium, and other elements.

In good photographs, the photosphere has a mottled appearance because it is made up of dark-edged regions called granules. The overall pattern is called **granulation** (■ Figure 7-2a). Each granule is about the size of Texas and lasts for only 10 to 20 minutes before fading away. Faded granules are continuously replaced by new granules. Spectra of these granules



■ **Figure 7-2**

(a) This ultra-high-resolution image of the photosphere shows granulation. The largest granules here are about the size of Texas. (Hinode JAXA/NASA/PPARC) (b) This model explains granulation as the tops of rising convection currents just below the photosphere. Heat flows upward as rising currents of hot gas and downward as sinking currents of cool gas. The rising currents heat the solar surface in small regions seen from Earth as granules.

show that the centers are a few hundred degrees hotter than the edges, and Doppler shifts reveal that the centers are rising and the edges are sinking at speeds of about 0.4 km/s.

From this evidence, astronomers recognize granulation as the surface effects of convection just below the photosphere. **Convection** occurs when hot fluid rises and cool fluid sinks, as when, for example, a convection current of hot gas rises above a candle flame. You can watch convection in a liquid by adding a bit of cool nondairy creamer to an unstirred cup of hot coffee. The cool creamer sinks, warms, expands, rises, cools, contracts, sinks again, and so on, creating small regions on the surface of the coffee that mark the tops of convection currents. Viewed from above, these regions look much like solar granules.

In the sun, rising currents of hot gas heat small regions of the photosphere, which, being slightly hotter, emit more black body radiation and look brighter. The cool sinking gas of the edges emits less light and thus looks darker (Figure 7-2b). The presence of granulation is clear evidence that energy is flowing upward through the photosphere.

Spectroscopic studies of the solar surface have revealed another less obvious kind of granulation. **Supergranules** are regions a little over twice Earth's diameter that include about 300 granules each. These supergranules are regions of very slowly rising currents that last a day or two. They appear to be produced by larger currents of rising gas deeper under the photosphere.

The Chromosphere

Above the photosphere lies the chromosphere. Solar astronomers define the lower edge of the chromosphere as lying just above the visible surface of the sun, with its upper regions blending gradually with the corona. You can think of the chromosphere as an irregular layer with a depth on average less than Earth's diameter (see Figure 7-1). Because the chromosphere is roughly 1000 times fainter than the photosphere, you can see it with your unaided eyes only during a total solar eclipse when the moon covers the brilliant photosphere. Then, the chromosphere flashes into view as a thin line of pink just above the photosphere. The word *chromosphere* comes from the Greek word *chroma,* meaning "color." The pink color is produced by the combined light of three bright emission lines—the red, blue, and violet Balmer lines of hydrogen.

Astronomers know a great deal about the chromosphere from its spectrum. The chromosphere produces an emission spectrum, and Kirchhoff's second law tells you it must be an excited, low-density gas. The chromosphere is about 10^8 times less dense than the air you breathe.

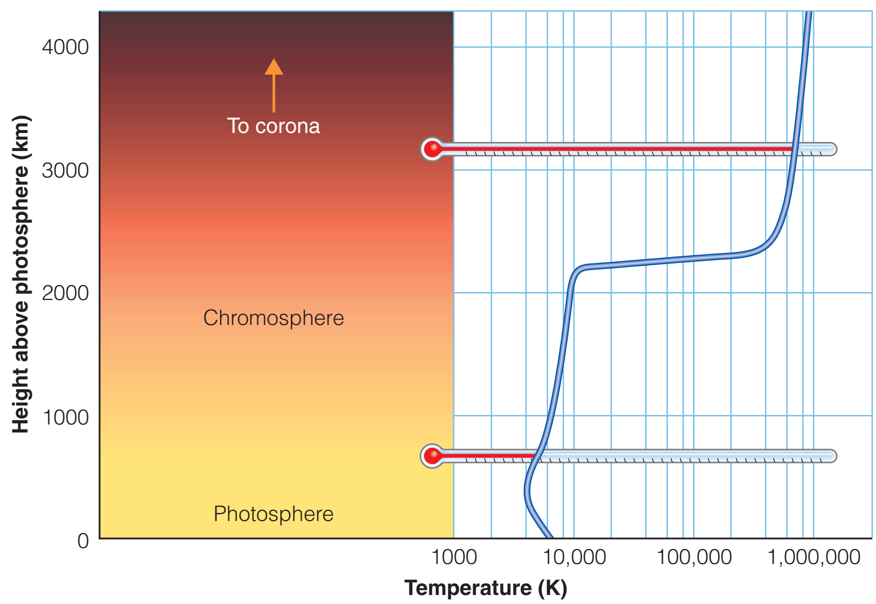
Spectra reveal that atoms in the lower chromosphere are ionized, and atoms in the higher layers of the chromosphere are even more highly ionized. That is, they have lost more electrons. From the ionization state of the gas, astronomers can find the temperature in different parts of the chromosphere. Just above the photosphere the temperature falls to a minimum of about 4500 K and then rises rapidly (■ Figure 7-3) to the extremely high temperatures of the corona.

Solar astronomers can take advantage of some elegant physics to study the chromosphere. The gases of the chromosphere are transparent to nearly all visible light, but atoms in the gas are very good at absorbing photons of specific wavelengths. This produces certain dark absorption lines in the spectrum of the photosphere. A photon at one of those wavelengths is very unlikely to escape from deeper layers. A **filtergram** is an image of the sun made using light in one of those dark absorption lines. Those photons can only have escaped from higher in the atmosphere. In this way, filtergrams reveal detail in the upper layers of the chromosphere. Another way to study these high layers of gas is to record solar images in the far-ultraviolet or in the X-ray part of the spectrum.

■ Figure 7-4 shows a filtergram made at the wavelength of the H_α Balmer line. This image shows complex structure in the chromosphere. **Spicules** are flamelike jets of gas extending upward into the chromosphere and lasting 5 to 15 minutes. Seen at the limb of the sun's disk, these spicules blend together and look like flames covering a burning prairie (Figure 7-1), but they are not flames at all. Spectra show that spicules are cooler gas from the lower chromosphere extending upward into hotter regions. Images at the center of the solar disk show that spicules spring up

■ Figure 7-3

The chromosphere. If you could place thermometers in the sun's atmosphere, you would discover that the temperature increases from 5800 K at the photosphere to 10^6 K at the top of the chromosphere.



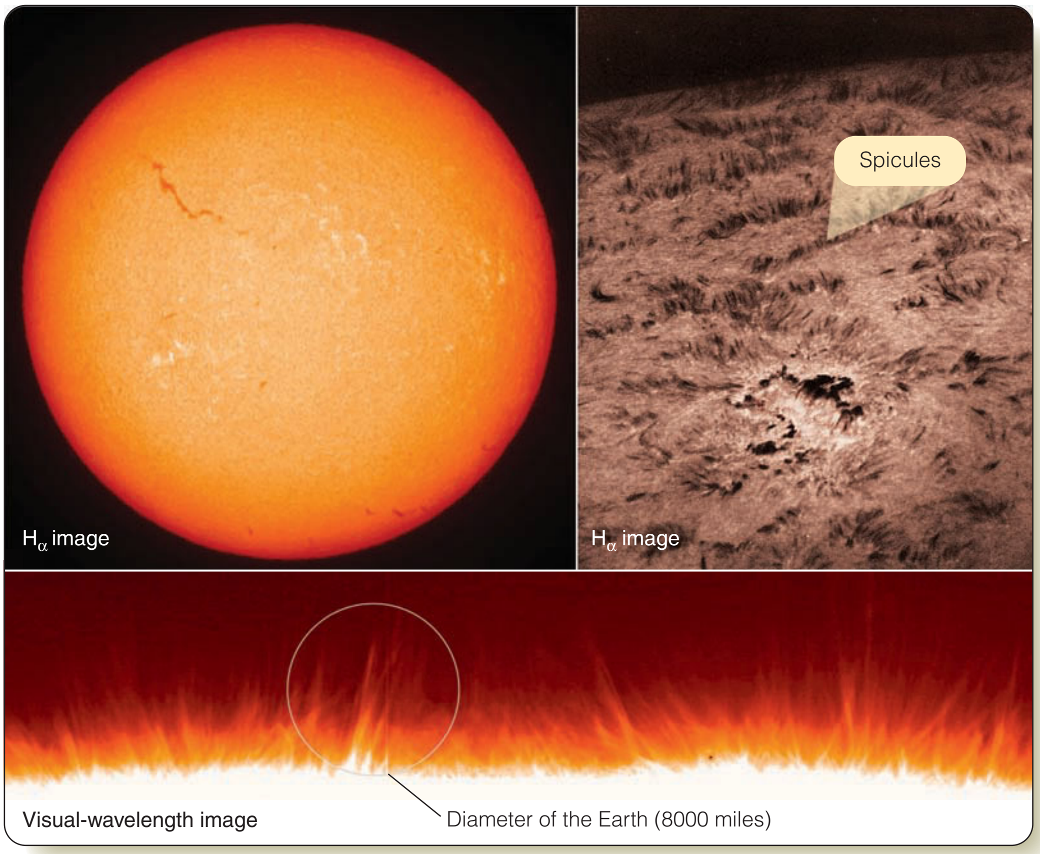


Figure 7-4

H_α filtergrams reveal complex structure in the chromosphere that cannot be seen at visual wavelengths, including spicules springing from the edges of supergranules over twice the diameter of Earth. Seen at the edge of the solar disk, spicules look like a burning prairie, but they are not at all related to burning. Compare with Figure 7-1. (BBSO; © 1971 NOAO/NSO; Hinode)

around the edge of supergranules like weeds around flagstones (Figure 7-4).

Spectroscopic analysis of the chromosphere alerts you that it is a low-density gas in constant motion where the temperature increases rapidly with height. Just above the chromosphere lies even hotter gas.

The Solar Corona

The outermost part of the sun's atmosphere is called the corona, after the Greek word for crown. The corona is so dim that it is not visible in Earth's daytime sky because of the glare of scattered light from the sun's brilliant photosphere. During a total solar eclipse, however, when the moon covers the photosphere, you can see the innermost parts of the corona, as shown in Figure 7-1b. Observations made with specialized telescopes called **coronagraphs** can block the light of the photosphere and record the corona out beyond 20 solar radii, almost 10 percent of the way to Earth. Such images show streamers in the corona that follow magnetic lines of force in the sun's magnetic field (■ Figure 7-5).

The spectrum of the corona can tell you a great deal about the coronal gases and simultaneously illustrate how astronomers analyze a spectrum. Some of the light from the outer corona produces a spectrum with absorption lines that are the same as the photosphere's spectrum. This light is just sunlight reflected from dust particles in the corona. In contrast, some of the light from the corona produces a continuous spectrum that lacks absorption lines, and that happens when sunlight from the photosphere is scattered off free electrons in the ionized coronal gas. Because the coronal gas has a temperature over 1 million K and the electrons travel very fast, the reflected photons suffer large, random Doppler shifts that smear out absorption lines to produce a continuous spectrum.

Superimposed on the corona's continuous spectrum are emission lines of highly ionized gases. In the lower corona, the atoms are not as highly ionized as they are at higher altitudes, and this tells you that the temperature of the corona rises with altitude. Just above the chromosphere, the temperature is about 500,000 K, but in the outer corona the temperature can be 2 million K or more.

The corona is exceedingly hot gas, but it is not very bright. Its density is very low, only 10^6 atoms/cm^3 in its lower regions. That is about a trillion times less dense than the air you breathe. In its outer layers the corona contains only 1 to 10 atoms/cm^3, fewer than in the best vacuum on Earth. Because of this low density, the hot gas does not emit much radiation.

Astronomers have wondered for years how the corona and chromosphere can be so hot. Heat flows from hot regions to cool regions, never from cool to hot. So how can the heat from the photosphere, with a temperature of only 5800 K, flow out into the much hotter chromosphere and corona? Observations made by the SOHO satellite have mapped a **magnetic carpet** of looped magnetic fields extending up through the photosphere. Remember that the gas of the chromosphere and corona has a very low density, so it can't resist movement of the magnetic fields. Turbulence below the photosphere seems to flick the magnetic loops back and forth and whip the gas about, heating the gas. Furthermore, observations with the Hinode spacecraft have revealed magnetic waves generated by turbulence below the photosphere traveling up into the chromosphere and corona and

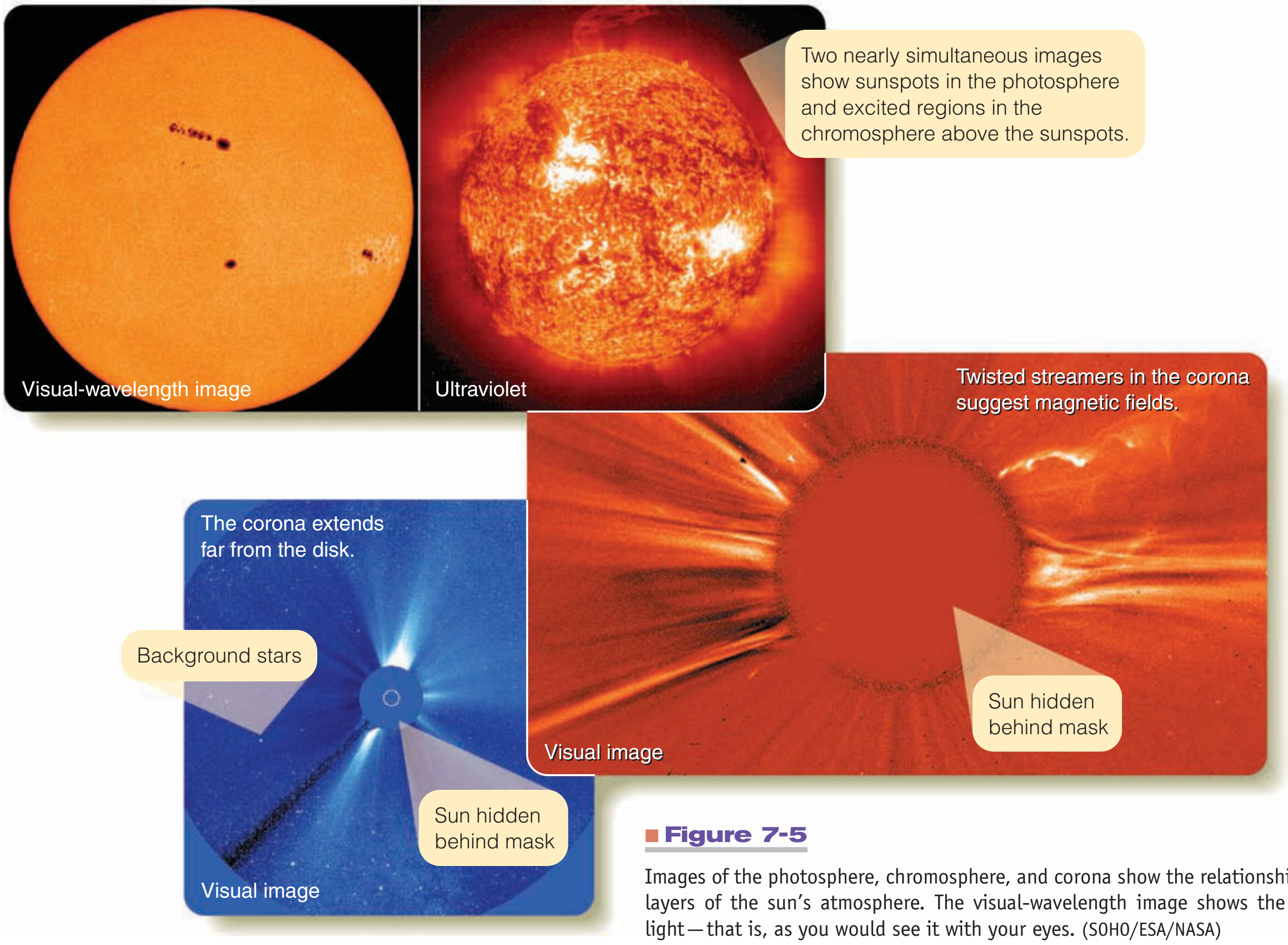


Figure 7-5

Images of the photosphere, chromosphere, and corona show the relationships among the layers of the sun's atmosphere. The visual-wavelength image shows the sun in white light—that is, as you would see it with your eyes. (SOHO/ESA/NASA)

heating the gas. In both cases, energy appears to flow outward as the agitation of the magnetic fields.

Not all of the sun's magnetic field loops back; some of the field leads outward into space. Gas from the solar atmosphere follows along the magnetic fields that point outward and flows away from the sun in a breeze called the **solar wind.** Like an extension of the corona, the low-density gases of the solar wind blow past Earth at 300 to 800 km/s with gusts as high as 1000 km/s. Earth is bathed in the corona's hot breath.

Because of the solar wind, the sun is slowly losing mass, but this is only a minor loss for an object as massive as the sun. The sun loses about 10^7 tons per second, but that is only 10^{-14} of a solar mass per year. Later in life, the sun, like many other stars, will lose mass rapidly in a more powerful wind. You will see in future chapters how this affects stars.

Do other stars have chromospheres, coronae, and stellar winds like the sun? Stars are so far away they never look like more than points of light, but ultraviolet and X-ray observations suggest that the answer is yes. The spectra of many stars contain emission lines in the far-ultraviolet that could have formed only in the low-density, high-temperature gases of a chromosphere and corona. Also, many stars are sources of X-rays, which appear to have been produced by the high-temperature gas in coronae. This observational evidence gives astronomers good reason to believe that the sun, for all its complexity, is a typical star.

The layers of the solar atmosphere are all that astronomers can observe directly, but there are phenomena in those layers that reveal what it's like inside the sun—your next destination.

Below the Photosphere

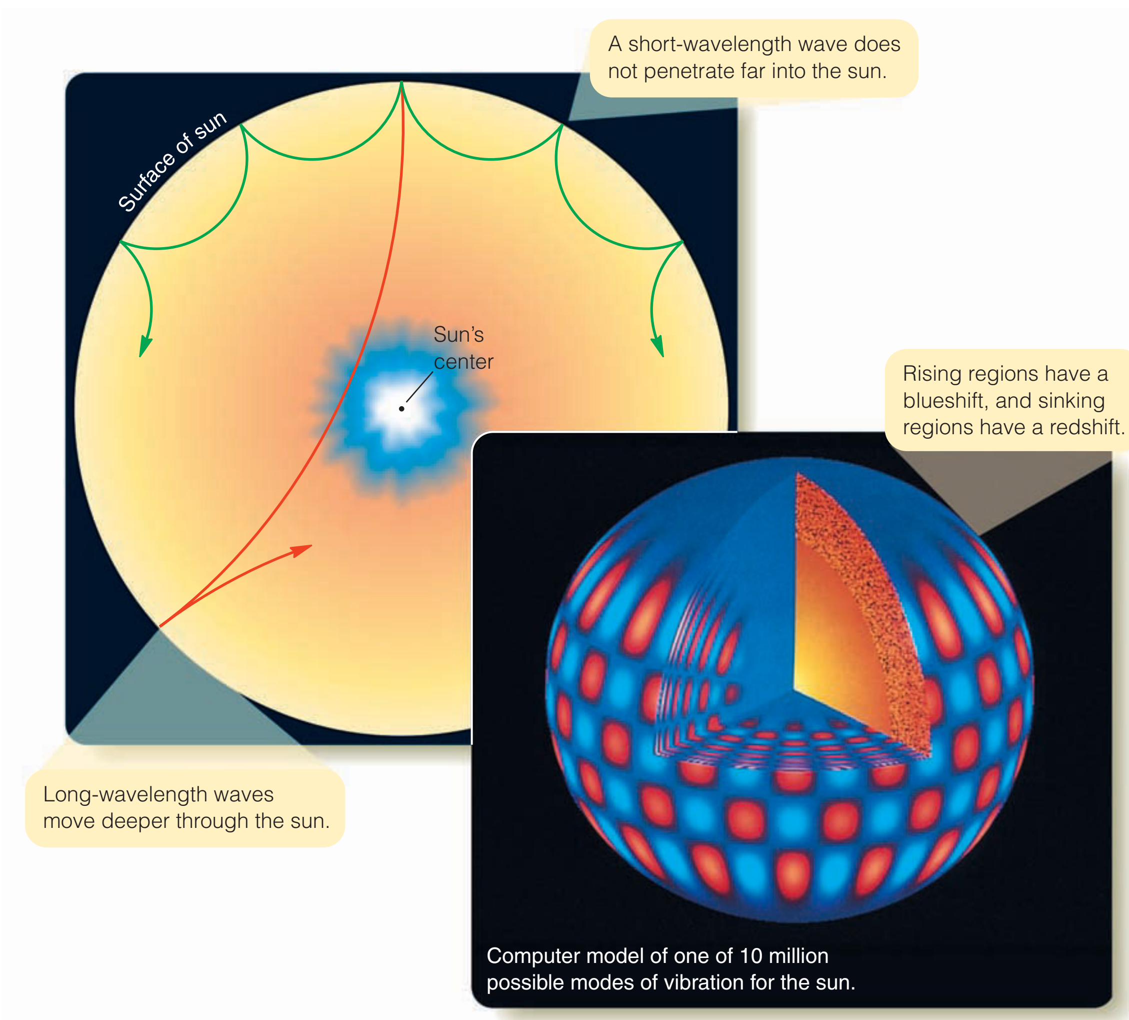
Almost no light emerges from below the photosphere, so you can't see into the solar interior. However, solar astronomers can study naturally occurring vibrations in the sun to explore its depths in a process called **helioseismology.** Random convective movements of gas in the sun constantly produce vibrations—rumbles that would be much too low to hear with human ears even if your ears could survive a visit to the sun's atmosphere. Some of these vibrations resonate in the sun like sound waves in organ pipes. A vibration with a period of 5 minutes is strongest, but the periods range from 3 to 20 minutes. These are very, very low-pitched sounds!

Astronomers can detect these vibrations by observing Doppler shifts in the solar surface. As a vibrational wave travels down into the sun, the increasing density and temperature curve its path, and it returns to the surface, where it makes the photosphere heave up and down by small amounts—roughly plus or minus 15 km. This covers the surface of the sun with a pattern of rising and falling regions that can be mapped using the Doppler effect (■ Figure 7-6). By observing these motions, astronomers can determine which vibrations resonate and become stronger and which become weaker. Short-wavelength waves penetrate less deeply and travel shorter distances than longer-wavelength waves, so the different wavelength vibrations explore different layers in the sun. Just as geologists can study Earth's interior by analyzing vibrations from earthquakes, so solar astronomers can use helioseismology to explore the sun's interior.

You can better understand how helioseismology works if you think of a duck pond. If you stood at the shore of a duck pond and looked down at the water, you would see ripples arriving from all parts of the pond. Because every duck on the pond contributes to the ripples, you could, in principle, study the ripples near the shore and draw a map showing the position and velocity of every duck on the pond. Of course, it would be difficult to untangle the different ripples, so you would need lots of data and a big computer. Nevertheless, all of the information would be there, lapping at the shore.

Helioseismology demands huge amounts of data, so astronomers have used a network of telescopes around the world operated by the Global Oscillation Network Group (GONG). The network can observe the sun continuously for weeks at a time as Earth rotates. The sun never sets on GONG. The SOHO satellite in space can observe solar oscillations continuously and can detect motions as slow as 1 mm/s (0.002 mph). Solar astronomers can then use high-speed computers to separate the different patterns on the solar surface and measure the strength of the waves at many different wavelengths.

Helioseismology has allowed astronomers to map the temperature, density, and rate of rotation inside the sun. They have been able to detect great currents of gas flowing below the photosphere and the emergence of sunspots before they appear in the photosphere. Helioseismology can even locate sunspots on the back side of the sun, sunspots that are not yet visible from Earth.



■ **Figure 7-6**

Helioseismology: The sun can vibrate in millions of different patterns or modes, and each mode corresponds to a different wavelength vibration penetrating to a different level. By measuring Doppler shifts as the surface moves gently up and down, astronomers can map the inside of the sun. (AURA/NOAO/NSF)

◄ **SCIENTIFIC ARGUMENT** ►

What evidence leads astronomers to conclude that temperature increases with height in the chromosphere and corona?

Scientific arguments usually involve evidence, and in astronomy that means observations. Solar astronomers can observe the spectrum of the chromosphere, and they find that atoms there are more highly ionized (have lost more electrons) than atoms in the photosphere. Atoms in the corona are even more highly ionized. That must mean the chromosphere and corona are hotter than the photosphere.

Evidence is the key to understanding how science works. Now it is time to build a new argument. **What evidence leads astronomers to conclude that other stars have chromospheres and coronae like those of the sun?**

◄ ►

7-2 Nuclear Fusion in the Sun

LIKE SOAP BUBBLES, stars are structures balanced between opposing forces that individually would destroy them. The sun is a ball of hot gas held together by its own gravity. If it were not for the sun's gravity, the hot, high-pressure gas in the sun's interior would explode outward. Likewise, if the sun were not so hot, its gravity would compress it into a small dense body. In this section, you will discover how the sun generates its heat.

The sun is powered by nuclear reactions that occur near its center.* The energy keeps the interior hot, and keeps the gas totally ionized. That is, the electrons are not attached to atomic nuclei, so the gas is an atomic soup of rapidly moving particles colliding with each other at high velocities. Nuclear reactions inside stars involve atomic nuclei, not whole atoms.

How exactly can the nucleus of an atom yield energy? The answer lies in the forces that hold the nuclei together.

Nuclear Binding Energy

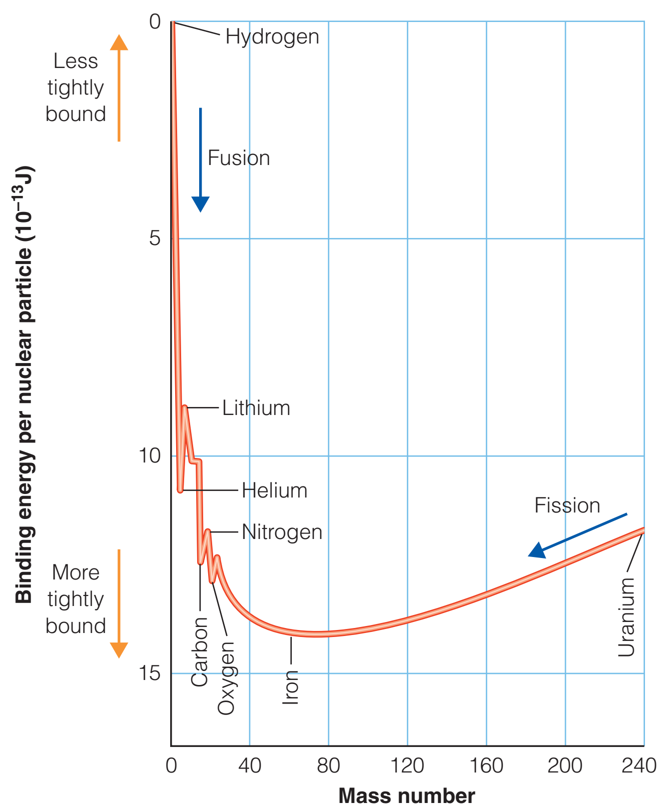
The sun generates its energy by breaking and reconnecting the bonds between the particles *inside* atomic nuclei. This is quite different from the way you would generate energy by burning wood in a fireplace. The process of burning wood extracts energy by breaking and rearranging chemical bonds among atoms in the wood. Chemical bonds are formed by the electrons in atoms, and you saw in Chapter 6 that the electrons are bound to the atoms by the electromagnetic force. So the chemical energy released when these bonds are broken and rearranged originates in the electromagnetic force.

There are only four forces in nature: the force of gravity, the electromagnetic force, the **weak force,** and the **strong force.** The weak force is involved in the radioactive decay of certain kinds of nuclear particles, and the strong force binds together atomic nuclei. Nuclear energy comes from the strong force.

Nuclear power plants on Earth generate energy through **nuclear fission** reactions that split uranium nuclei into less massive fragments. A uranium nucleus contains a total of 235 protons and neutrons, and when it decays, it splits into a range of fragments containing roughly half as many particles. Because the fragments produced are more tightly bound than the uranium nuclei, binding energy is released during uranium fission.

Stars don't use nuclear fission. They make energy in **nuclear fusion** reactions that combine light nuclei into heavier nuclei. The most common reaction, the one that occurs in the sun, fuses hydrogen nuclei (single protons) into helium nuclei, which contain two protons and two neutrons. Because the nuclei produced are more tightly bound than the original nuclei, energy is released.

■ Figure 7-7 shows how tightly different atomic nuclei are bound. The lower in the diagram, the more tightly the particles in a nucleus are held. Notice that both fusion and fission reactions move downward in the diagram toward more tightly bound



■ **Figure 7-7**

The red line in this graph shows the binding energy per particle, the energy that holds particles inside an atomic nucleus. The horizontal axis shows the atomic mass number of each element, the number of protons and neutrons in the nucleus. Both fission and fusion nuclear reactions move downward in the diagram (arrows) toward more tightly bound nuclei. Iron has the most tightly bound nucleus, so no nuclear reactions can begin with iron and release energy.

*Astronomers sometimes use the wrong words when they talk about nuclear reactions inside stars. They may use words like *burn* or *ignite.* What goes on inside stars is not related to simple burning but is comprised of nuclear reactions.

nuclei. They both produce energy by releasing the binding energy of atomic nuclei.

Hydrogen Fusion

The sun fuses together four hydrogen nuclei to make one helium nucleus. Because one helium nucleus has 0.7 percent less mass than four hydrogen nuclei, it seems that some mass vanishes in the process. In fact, that mass is converted to energy, and you could figure out how much by using Einstein's famous equation $E = mc^2$ (■ Reasoning with Numbers 7-1).

You can symbolize the fusion reactions in the sun with a simple nuclear reaction:

$$4\ {}^1\mathrm{H} \rightarrow {}^4\mathrm{He} + \text{energy}$$

In this equation, ^{1}H represents a proton, the nucleus of the hydrogen atom, and ^{4}He represents the nucleus of a helium atom. The superscripts indicate the approximate weight of the nuclei (the number of protons plus the number of neutrons). The actual steps in the process are more complicated than this convenient summary suggests. Instead of waiting for four hydrogen nuclei to collide simultaneously, a highly unlikely event, the process can proceed step-by-step in a chain of reactions—the proton–proton chain.

The **proton–proton chain** is a series of three nuclear reactions that builds a helium nucleus by adding together protons. This process is efficient at temperatures above 10,000,000 K. The sun, for example, manufactures over 90 percent of its energy in this way.

The three steps in the proton–proton chain entail these reactions:

$$\begin{aligned} {}^1\mathrm{H} + {}^1\mathrm{H} &\rightarrow {}^2\mathrm{H} + e^+ + \nu \\ {}^2\mathrm{H} + {}^1\mathrm{H} &\rightarrow {}^3\mathrm{He} + \gamma \\ {}^3\mathrm{He} + {}^3\mathrm{He} &\rightarrow {}^4\mathrm{He} + {}^1\mathrm{H} + {}^1\mathrm{H} \end{aligned}$$

In the first reaction, two hydrogen nuclei (two protons) combine to form a heavy hydrogen nucleus called **deuterium,** emitting a particle called a **positron,** e^+ (a positively charged electron), and a **neutrino,** ν (a subatomic particle having an extremely low mass and a velocity nearly equal to the velocity of light). In the second reaction, the heavy hydrogen nucleus absorbs another proton and, with the emission of a gamma ray, γ, becomes a lightweight helium nucleus. Finally, two lightweight helium nuclei combine to form a common helium nucleus and two hydrogen nuclei. Because the last reaction needs two ^{3}He nuclei, the first and second reactions must occur twice (■ Figure 7-8). The net result of this chain reaction is the transformation of four hydrogen nuclei into one helium nucleus plus energy.

The energy appears in the form of gamma rays, positrons, the energy of motion of the particles, and neutrinos. The gamma rays are photons that are absorbed by the surrounding gas before they can travel more than a fraction of a millimeter. This heats the gas. The positrons produced in the first reaction combine with free electrons, and both particles vanish, converting their mass into gamma rays, which are absorbed and also help keep the gas hot. In addition, when fusion produces new nuclei, they fly apart at high velocity and collide with other particles. This energy of motion helps raise the temperature of the gas. The neutrinos, on the other hand, don't heat the gas. Neutrinos resemble photons except that they almost never interact with other particles. The average neutrino could pass unhindered through a lead wall a light-year thick. Consequently, the neutrinos do not warm the gas but race out of the sun at nearly the speed of light, carrying away roughly 2 percent of the energy produced.

Creating one helium nucleus makes only a small amount of energy, hardly enough to raise a housefly one-thousandth of an inch. Because one reaction produces such a small amount of energy, it is obvious that many reactions are necessary to supply the energy needs of a star. The sun, for example, needs to complete 10^{38} reactions per second, transforming 5 million tons of mass into energy every second. It might sound as if the sun is losing mass at a furious rate, but in its entire 10-billion-year lifetime, the sun will convert less than 0.07 percent of its mass into energy.

It is a **Common Misconception** that nuclear fusion in the sun is tremendously powerful. After all, the fusion of a milligram of hydrogen (roughly the mass of a match head) produces as much energy as burning 30 gallons of gasoline. However, at any one time, only a tiny fraction of the hydrogen atoms are fusing into helium, and the nuclear reactions in the sun are spread through a large volume in its core. Any single gram of matter produces only a little energy. A person of normal mass eating a

Reasoning with Numbers | 7-1

Hydrogen Fusion

When four hydrogen nuclei fuse to make one helium nucleus, a small amount of matter seems to disappear:

$$\begin{aligned} \text{4 hydrogen nuclei} &= 6.693 \times 10^{-27}\ \text{kg} \\ -\ \text{1 helium nucleus} &= 6.645 \times 10^{-27}\ \text{kg} \\ \hline \text{difference in mass} &= 0.048 \times 10^{-27}\ \text{kg} \end{aligned}$$

That mass is converted to energy according to Einstein's equation:

$$\begin{aligned} E &= mc^2 \\ &= (0.048 \times 10^{-27}\ \text{kg}) \times (3 \times 10^8\ \text{m/s})^2 \\ &= 0.43 \times 10^{-11}\ \text{J} \end{aligned}$$

Recall that one joule (J) is roughly equal to the energy of an apple falling from a table to the floor.

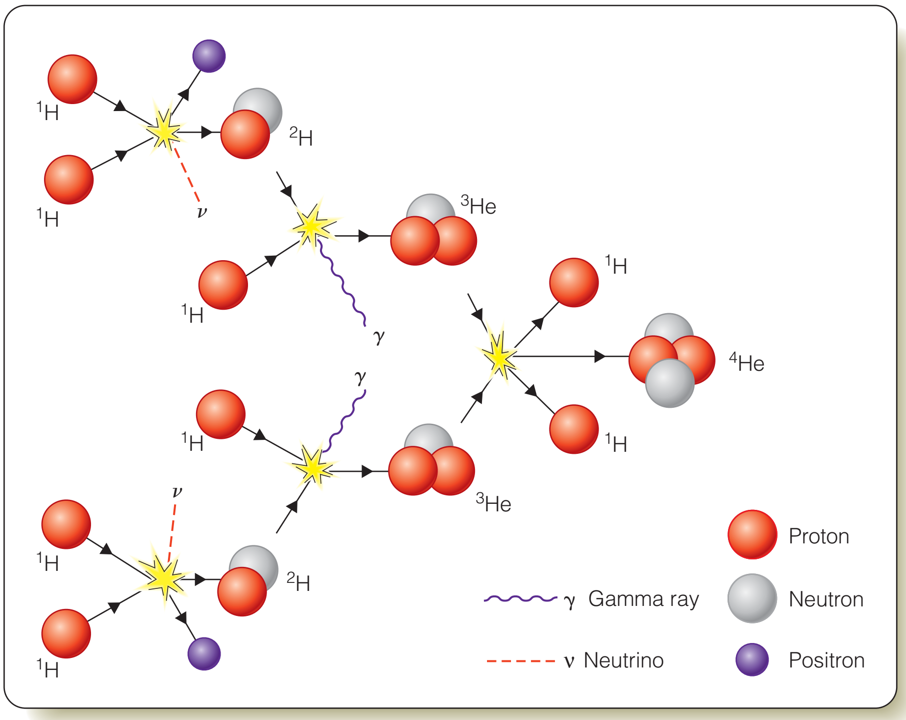


Figure 7-8

The proton–proton chain combines four protons (at far left) to produce one helium nucleus (at right). Energy is produced mostly as gamma rays and as positrons, which combine with electrons and convert their mass into energy. Neutrinos escape, carrying away about 2 percent of the energy produced.

normal diet produces about 4000 times more heat per gram than the matter in the core of the sun. Gram for gram, you are a much better heat producer than the sun. The sun produces a lot of energy because it contains a lot of grams of matter in its core.

Fusion reactions can occur only when the nuclei of two atoms get very close to each other. Because atomic nuclei carry positive charges, they repel each other with an electrostatic force called the Coulomb force. Physicists commonly refer to this repulsion between nuclei as the **Coulomb barrier.** To overcome this barrier and get close together, atomic nuclei must collide violently. Violent collisions are rare unless the gas is very hot, in which case the nuclei move at high speeds and collide violently. (Remember, an object's temperature is related to the speed with which its particles move.)

So nuclear reactions in the sun take place only near the center, where the gas is hot and dense. A high temperature ensures that a few of the collisions between nuclei are violent enough to overcome the Coulomb barrier, and a high density ensures that there are enough collisions, and thus enough reactions, to meet the sun's energy needs.

Energy Transport in the Sun

Now you are ready to follow the energy from the core of the sun to the surface. The surface is cool, only about 5800 K, and the center is over 10 million K, so energy must flow outward from the core.

Because the core is so hot, the photons bouncing around there are gamma rays. Each time a gamma ray encounters an electron, it is deflected or scattered in a random direction; and, as it bounces around, it slowly drifts outward toward the surface. That carries energy outward in the form of radiation, so astronomers refer to the inner parts of the sun as the **radiative zone.**

To examine this process, imagine picking a single gamma ray and following it to the surface. As your gamma ray is scattered over and over by the hot gas, it drifts outward into cooler layers, where the cooler gas tends to emit photons of longer wavelength. Your gamma ray will eventually be absorbed by the gas and reemitted as two X-rays. Now you must follow those two X-rays as they bounce around, and soon you will see them drifting outward into even cooler gas, where they will become a number of longer-wavelength photons. The packet of energy that began as a single gamma ray gets broken down into a large number of lower-energy photons, and it eventually emerges from the sun's surface as about 1800 photons of visible light.

But something else happens along the way. The packet of energy that you began following in the core eventually reaches the outer layers of the sun where the gas is so cool that it is not very transparent to radiation. There, energy backs up like water behind a dam, and the gas begins to churn in convection. Hot blobs of gas rise, and cool blobs sink. In this region, known as the **convective zone,** the energy is carried outward as circulating gas.

The radiative and convective zones are shown in Figure 7-9. The granulation visible on the photosphere is clear evidence of a convective zone just below the photosphere carrying energy upward to the surface.

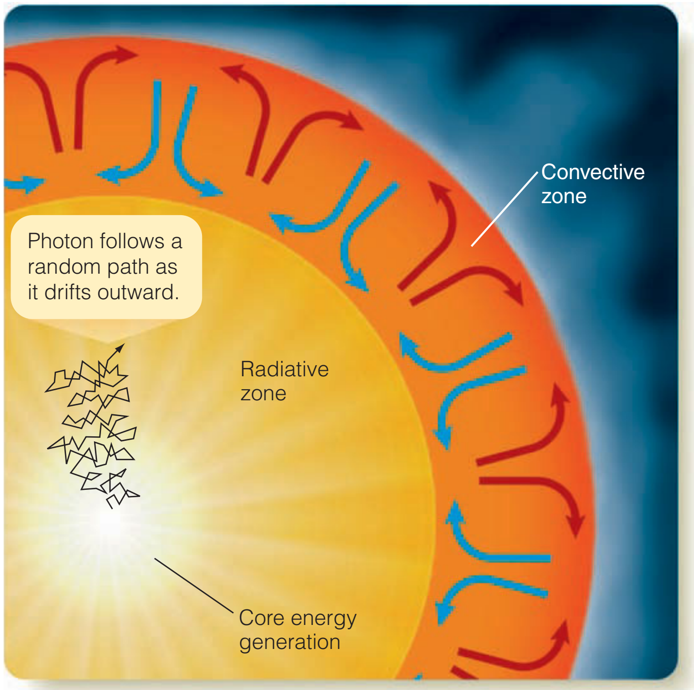
Sunlight is nuclear energy produced in the core of the sun. The energy of a single gamma ray can take a million years to work its way outward, first as radiation and then as convection on its journey to the photosphere.

It is time to ask the critical question that lies at the heart of science. What is the evidence to support this theoretical explanation of how the sun makes its energy?

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Nuclear Fusion."

Counting Solar Neutrinos

Nuclear reactions in the sun's core produce floods of neutrinos that rush out of the sun and off into space. Over 10^{12} solar neutrinos flow through your body every second, but you never feel



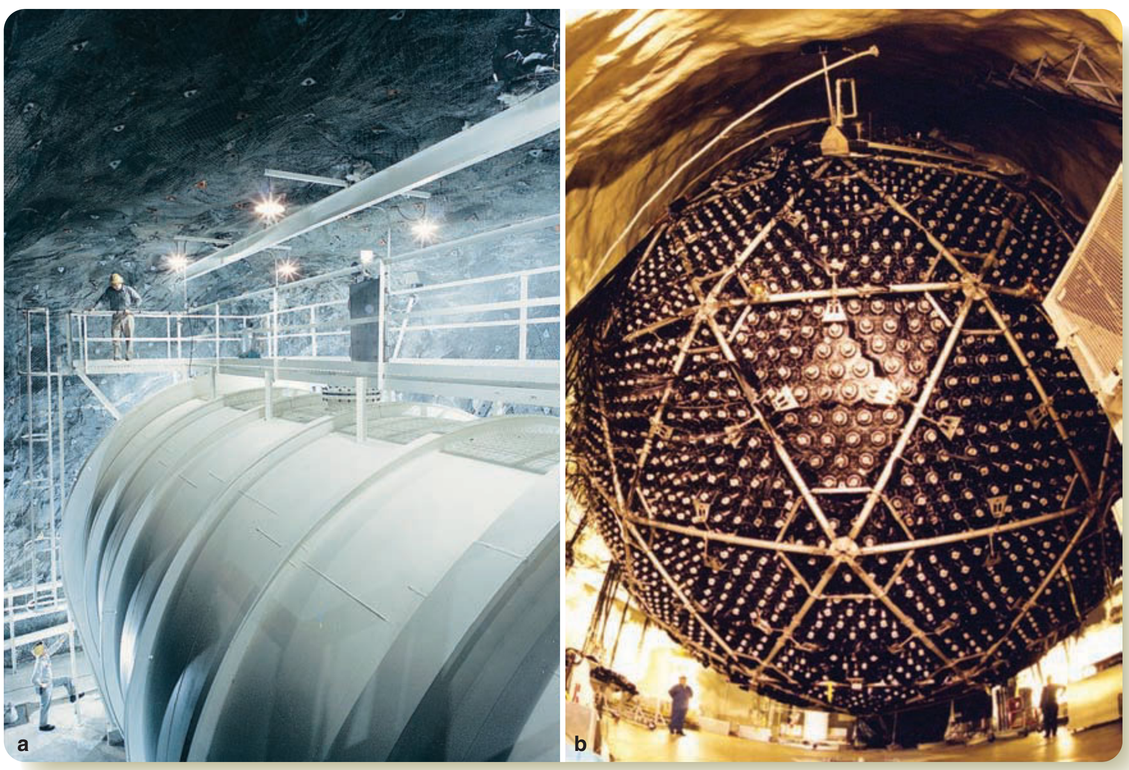
■ **Figure 7-9**

A cross section of the sun. Near the center, nuclear fusion reactions generate high temperatures. Energy flows outward through the radiative zone as photons are randomly defected over and over by electrons. In the cooler, more opaque outer layers, the energy is carried by rising convection currents of hot gas (red arrows) and sinking currents of cooler gas (blue arrows). **Animated!**

them because you are almost perfectly transparent to neutrinos. If you could detect these neutrinos, you could probe the sun's interior. You can't focus neutrinos with a lens or mirror, and they zip right through detectors used to count other atomic particles, but neutrinos of certain energies can trigger the radioactive decay of certain atoms. That gives astronomers a way to count solar neutrinos.

In the 1960s, chemist Raymond Davis Jr. devised a way to count neutrinos produced by hydrogen fusion in the sun. He buried a 100,000-gallon tank of cleaning fluid (perchloroethylene C_2Cl_4) in the bottom of a South Dakota gold mine where cosmic rays could not reach it (■ Figure 7-10a) and counted the number of times a neutrino triggered a chlorine atom into becoming an argon atom. He expected to detect one neutrino a day, but he actually counted one-third as many as expected, only one every three days.

The Davis neutrino experiment created a huge controversy. Were scientists wrong about nuclear fusion in the sun? Did they misunderstand how neutrinos behave? Because astronomers had great confidence in their understanding of the solar interior, they didn't abandon their theories immediately (■ How Do We Know? 7-1). It took over 30 years, but eventually physicists were able to build better detectors, and they discovered that neutrinos oscillate among three different types, which physicists call flavors. Nuclear reactions in the sun produce only one flavor, and the Davis experiment was designed to detect (taste) that flavor.



■ **Figure 7-10**

(a) The Davis solar neutrino experiment used cleaning fluid and could detect only one of the three flavors of neutrinos. (Brookhaven National Laboratory) (b) The Sudbury Neutrino Observatory is a 12-meter-diameter globe containing water rich in deuterium in place of hydrogen. Buried 6800 feet deep in an Ontario mine, it can detect all three flavors of neutrinos and confirms that neutrinos oscillate among the flavors. (Photo courtesy of SNO)

How Do We Know? 7-1

Scientific Confidence

How can scientists be certain of something? Sometimes scientists stick so firmly to their ideas in the face of contradictory claims that it sounds as if they are stubbornly refusing to consider alternatives. One example is the perpetual motion machine, a device that runs continuously with no source of energy. If you could invest in a real perpetual motion machine, you could sell cars that would run without any fuel. That's good mileage.

For centuries people have claimed to have invented a perpetual motion machine, and for just as long scientists have been dismissing these claims as impossible. The problem with a perpetual motion machine is that it violates the law of conservation of energy, and scientists are not willing to accept that the law could be wrong. In fact, the Royal Academy of Sciences in Paris was so sure that a perpetual motion machine was impossible, and so tired of debunking hoaxes, that in 1775 they issued a formal statement refusing to deal with them. The U.S. Patent Office is so skeptical that they won't even consider granting a patent for one without seeing a working model first. Why do scientists seem so stubborn and closed minded on this issue?

Why isn't one person's belief in perpetual motion just as valid as another person's belief in the law of conservation of energy? In fact, the two positions are not equally valid. The confidence physicists have in their law is not a belief or even an opinion; it is an understanding founded on the fact that the law has been tested uncountable times and has never failed. The law is a fundamental truth about nature and can be used to understand what is possible and what is impossible. In contrast, no one has ever successfully demonstrated a perpetual motion machine.

When the first observations of solar neutrinos detected fewer than predicted, some scientists speculated that astronomers misunderstood how the sun makes its energy or that they misunderstood the internal structure of the sun. But many astronomers stubbornly refused to reject their model because the nuclear physics of the proton–proton chain is well understood, and models of the sun's structure have been tested successfully many times. The confidence astronomers felt in their understanding of the sun was an example of scientific certainty, and that confidence in basic natural laws prevented them from abandoning decades of work in the face of a single contradictory observation.

What seems to be stubbornness among scientists is really their confidence in basic principles that have been tested over and over. Those principles are the keel that keeps the ship of science from rocking before every little breeze. Without even looking at that perpetual motion machine, your physicist friends can warn you not to invest.

But in the 8-minute journey from the sun's core to Earth, the neutrinos oscillated so much they were evenly distributed among the three different flavors when they arrived at Earth. That's why the Davis experiment detected only one-third of the number predicted.

In 2007, scientists announced that a supersensitive experiment in a tunnel under the Italian Alps had detected 50 neutrinos a day coming from the sun. The neutrinos have lower energies than those caught by the Davis experiment and are produced by a side reaction that produces beryllium-7. The number of neutrinos detected matches the prediction of models of nuclear fusion in the sun.

The center of the sun seems forever beyond human experience, but counting solar neutrinos provides the evidence to confirm the theories. The sun makes its energy through nuclear fusion.

◄ SCIENTIFIC ARGUMENT ►

Why does nuclear fusion require that the gas be very hot?

This argument has to include the basic physics of atoms and thermal energy. Inside a star, the gas is so hot it is ionized, which means the electrons have been stripped off the atoms leaving bare, positively charged nuclei. In the case of hydrogen, the nuclei are single protons. These atomic nuclei repel each other because of their positive charges, so they must collide with each other at high velocity if they are to overcome that repulsion and get close enough together to fuse. If the atoms in a gas are moving rapidly, then the gas must have a high temperature, so nuclear fusion requires that the gas be very hot. If the gas is cooler than about 10 million K, hydrogen can't fuse because the protons don't collide violently enough to overcome the repulsion of their positive charges.

It is easy to see why nuclear fusion in the sun requires high temperature, but now expand your argument. **Why does it require high density?**

◄ ►

7-3 Solar Activity

THE SUN IS unquiet. It is home to slowly changing spots larger than Earth and vast eruptions that dwarf human imagination. All of these seemingly different forms of solar activity have one thing in common—magnetic fields. The weather on the sun is magnetic.

Observing the Sun

Solar activity is often visible with even a small telescope, but you should be very careful if you try to observe the sun. Sunlight is intense, and when it enters your eye it is absorbed and converted into thermal energy. The infrared radiation in sunlight is especially dangerous because your eyes can't detect it. You don't sense how intense the infrared is, but it is converted to thermal energy in your eyes and can burn and scar your retinas.

It is not safe to look directly at the sun, and it is even more dangerous to look at the sun through any optical instrument such as a telescope, binoculars, or even the viewfinder of a camera. The light-gathering power of such an optical system concentrates the sunlight and can cause severe injury. Never look at the sun with any optical instrument unless you are certain it is safe. ■ Figure 7-11 shows a safe way to observe the sun with a small telescope.

In the early 17th century, Galileo observed the sun and saw spots on its surface; day by day he saw the spots moving across the sun's disk. He rightly concluded that the sun was a sphere and was rotating. If you repeated his observations, you would probably see something that looks like Figure 7-11b. You would see sunspots.

Sunspots

The dark sunspots that you see at visible wavelengths only hint at the complex processes that go on in the sun's atmosphere. To explore those processes, you must analyze images and spectra at a wide range of wavelengths.

Study ■ **Sunspots and the Sunspot Cycle** on pages 126–127 and notice five important points and four new terms:

1. Sunspots are cool spots on the sun's surface caused by strong magnetic fields.
2. Sunspots follow an 11-year cycle, becoming more numerous, reaching a maximum, and then becoming much less numerous. The *Maunder butterfly diagram* shows how the location of sunspots changes during a cycle.
3. The *Zeeman effect* gives astronomers a way to measure the strength of magnetic fields on the sun and provide evidence that sunspots contain strong magnetic fields.
4. The intensity of the sunspot cycle can vary from cycle to cycle and appears to have almost faded away during the *Maunder minimum* in the late 17th century. This seems to have affected Earth's climate.
5. The evidence is clear that sunspots are part of *active regions* dominated by magnetic fields that involve all layers of the sun's atmosphere.

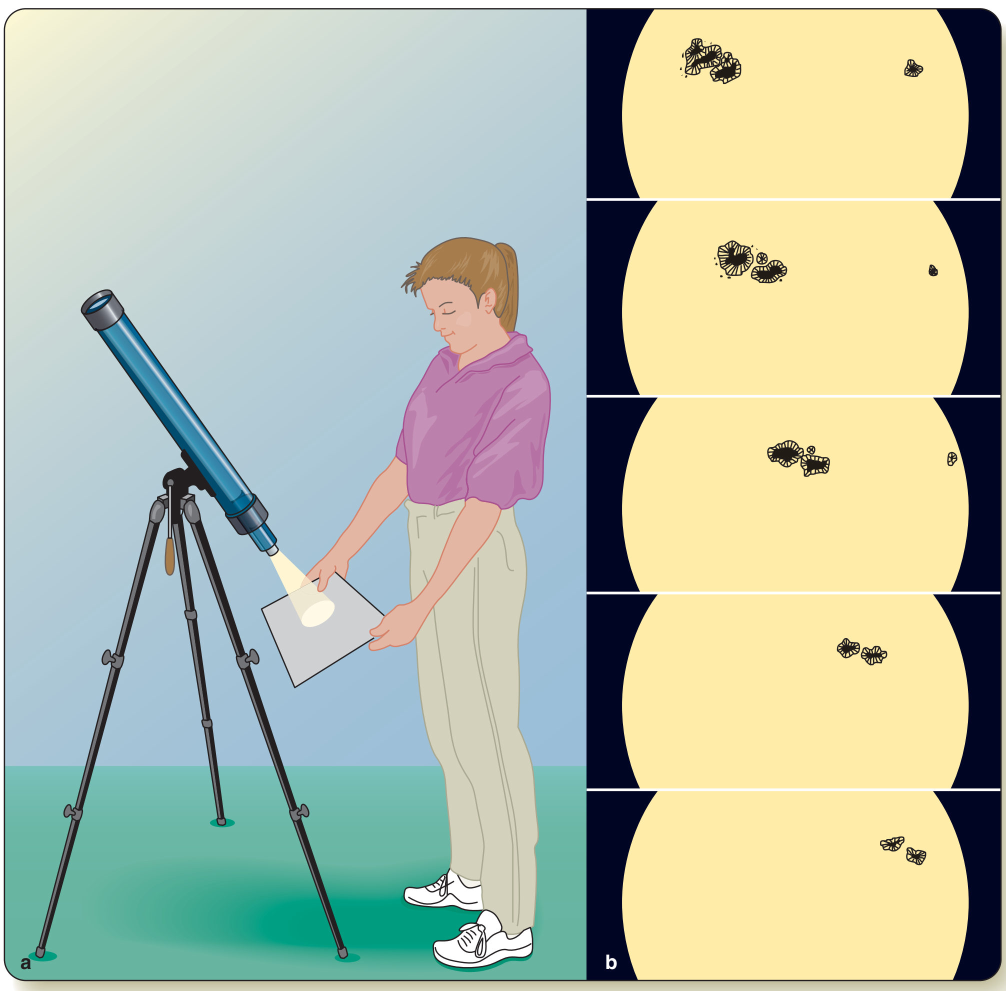
The sunspot groups are merely the visible traces of magnetically active regions. But what causes this magnetic activity? The answer is linked to the waxing and waning of the sun's overall magnetic field.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Zeeman Effect," "Sunspot Cycle I," and "Sunspot Cycle II."

The Sun's Magnetic Cycle

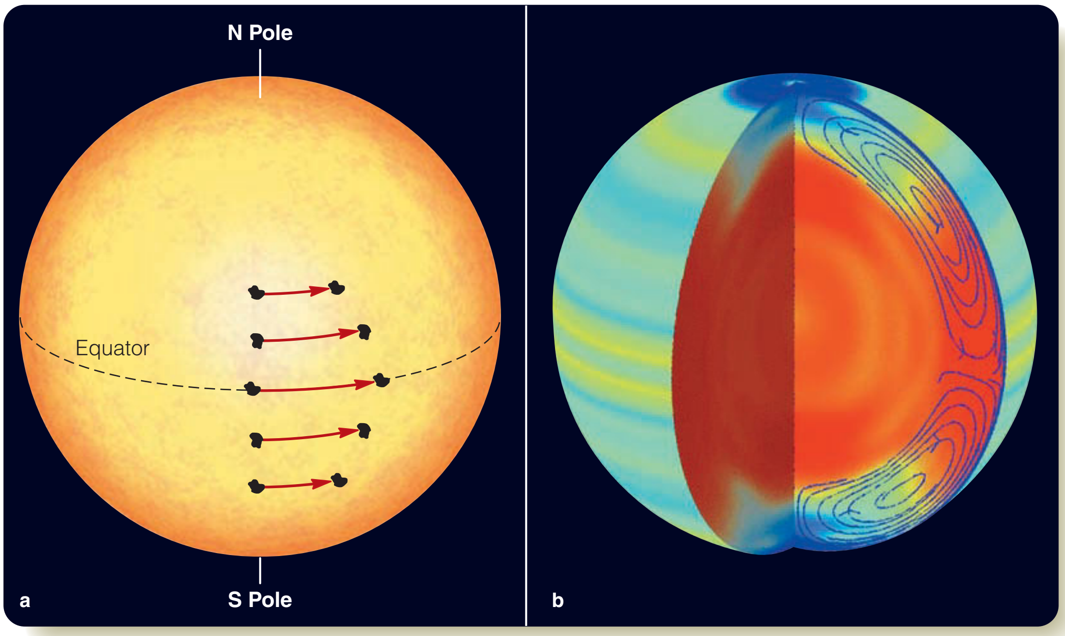
The sun's magnetic field is powered by the energy flowing outward through the moving currents of gas. The gas is highly ionized, so it is a very good conductor of electricity. When an electrical conductor rotates rapidly and is stirred by convection, it can convert some of the energy flowing outward as convection into a magnetic field. This process is called the **dynamo effect,** and it is believed to operate in Earth's core and produce Earth's magnetic field. Helioseismologists have found evidence that the dynamo effect generates the sun's magnetic field at the bottom of the convection zone deep under the photosphere.

The sun's magnetic field cannot be as stable as Earth's. The sun does not rotate as a rigid body. It is a gas from its outermost layers down to its center, so some parts of the sun can rotate faster than other parts. The



■ **Figure 7-11**

(a) Looking through a telescope at the sun is dangerous, but you can always view the sun safely with a small telescope by projecting its image on a white screen. (b) If you sketch the location and structure of sunspots on successive days, you will see the rotation of the sun and gradual changes in the size and structure of sunspots just as Galileo did in 1610.



■ **Figure 7-12**

(a) In general, the photosphere of the sun rotates faster at the equator than at higher latitudes. If you started five sunspots in a row, they would not stay lined up as the sun rotates. (b) Detailed analysis of the sun's rotation from helioseismology reveals regions of slow rotation (blue) and rapid rotation (red). Such studies show that the interior of the sun rotates differentially and that currents similar to the trade winds in Earth's atmosphere flow through the sun. (NASA/SOI)

equatorial region of the photosphere rotates faster than do regions at higher latitudes (■ Figure 7-12a). At the equator, the photosphere rotates once every 25 days, but at latitude 45° one rotation takes 27.8 days. Helioseismology can map the rotation throughout the interior (Figure 7-12b) and even there different levels rotate with different periods. This phenomenon is called **differential rotation,** and it is clearly linked with the magnetic cycle.

Although the magnetic cycle is not fully understood, the **Babcock model** (named for its inventor) explains the magnetic cycle as a progressive tangling of the solar magnetic field. Because the electrons in an ionized gas are free to move, the gas is a very good conductor of electricity, so any magnetic field in the gas is "frozen" into it. If the gas moves, the magnetic field must move with it. The sun's magnetic field is frozen into its gases, and differential rotation wraps this field around the sun like a long string caught on a hubcap. Rising and sinking gas currents twist the field into ropelike tubes, which tend to float upward. The model predicts that sunspot pairs occur where these magnetic tubes burst through the sun's surface (■ Figure 7-13).

Sunspots tend to occur in groups or pairs, and the magnetic field around the pair resembles that around a bar magnet with one end magnetic north and the other end magnetic south, just as you would expect if a magnetic tube emerged through one sunspot in a pair and reentered through the other. At any one time, sunspot pairs south of the sun's equator have reversed polarity compared with those north of the sun's equator. ■ Figure 7-14 illustrates this by showing sunspot pairs south of the sun's equator with magnetic south poles leading and sunspots north of the sun's equator with magnetic north poles leading. At the end of an 11-year sunspot cycle, the new spots appear with reversed magnetic polarity.

The Babcock model explains the reversal of the sun's magnetic field from cycle to cycle. As the magnetic field becomes tangled, adjacent regions of the sun are dominated by magnetic fields that point in different directions. After about 11 years of tangling, the field becomes so complex that adjacent regions of the sun begin changing their magnetic field to agree with neighboring regions. The entire field quickly rearranges itself into a simpler pattern, and differential rotation begins winding it up to start a new cycle. But the newly organized field is reversed, and the next sunspot cycle begins with magnetic north replaced by magnetic south. Consequently, the complete magnetic cycle is 22 years long, and the sunspot cycle is 11 years long.

This magnetic cycle explains the Maunder butterfly diagram. As a sunspot cycle begins, the twisted tubes of magnetic force first begin to float upward and produce sunspot pairs at higher latitude. Consequently the first sunspots in a cycle appear further north and south of the equator. Later in the cycle, when the field is more tightly wound, the tubes of magnetic force arch up through the surface closer to the equator. As a result, the later sunspot pairs in a cycle appear closer to the equator.

Notice the power of a scientific model. The Babcock model may in fact be incorrect in some details, but it provides a framework on which to organize all of the complex solar activity. Even though the models of the sky in Chapter 2 and the atom in Chapter 6 were only partially correct, they served as organizing themes to guide your thinking. Similarly, although the precise details of the solar magnetic cycle are not yet understood, the Babcock model gives you a general picture of the behavior of the sun's magnetic field (■ How Do We Know? 7-2).

If the sun is truly a representative star, you might expect to find similar magnetic cycles on other stars, but stars other than the sun are too distant to be observed as anything but tiny points of light and spots are not directly visible. Some stars, however, vary in brightness over a period of days in a way that reveals they are marked

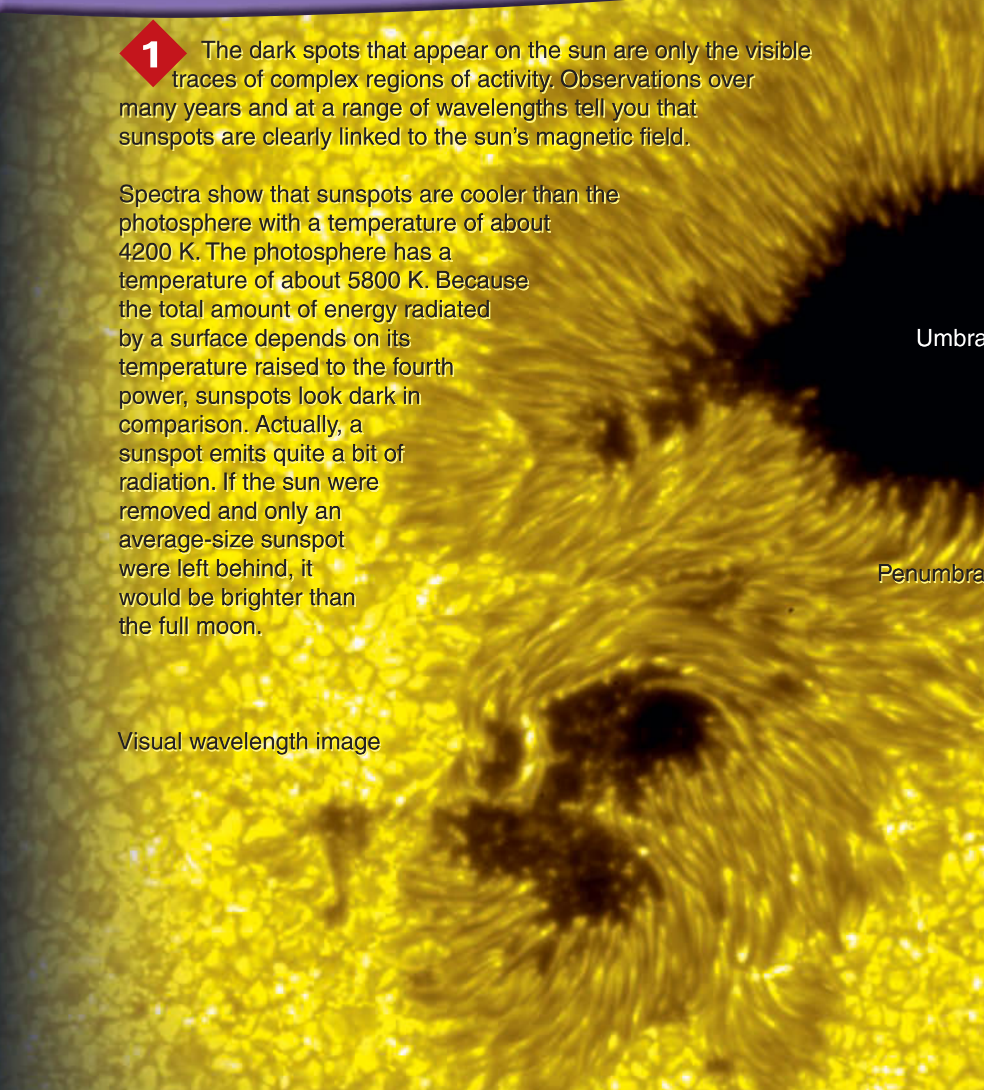
Sunspots and the Sunspot Cycle

1 The dark spots that appear on the sun are only the visible traces of complex regions of activity. Observations over many years and at a range of wavelengths tell you that sunspots are clearly linked to the sun's magnetic field.

Spectra show that sunspots are cooler than the photosphere with a temperature of about 4200 K. The photosphere has a temperature of about 5800 K. Because the total amount of energy radiated by a surface depends on its temperature raised to the fourth power, sunspots look dark in comparison. Actually, a sunspot emits quite a bit of radiation. If the sun were removed and only an average-size sunspot were left behind, it would be brighter than the full moon.

A typical sunspot is about twice the size of Earth, but there is a wide range of sizes. They appear, last a few weeks to as long as 2 months, and then shrink away. Usually, sunspots occur in pairs or complex groups.

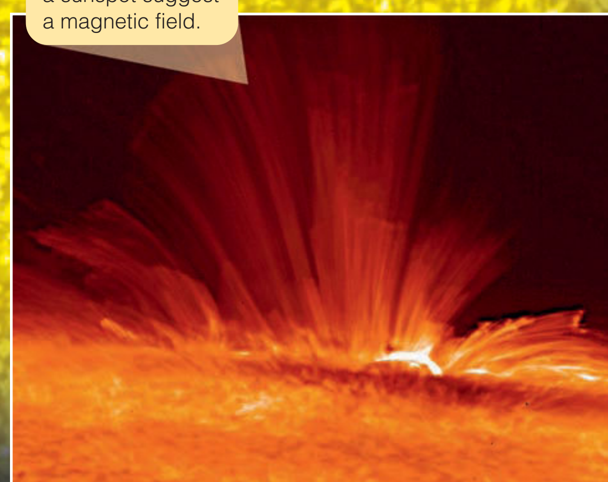
NASA



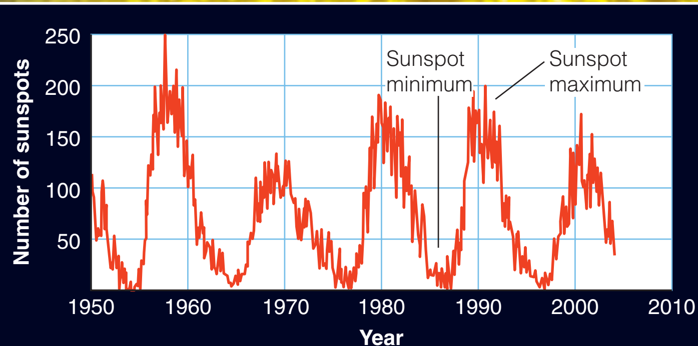
Visual wavelength image

Hinode JAXA/NASA

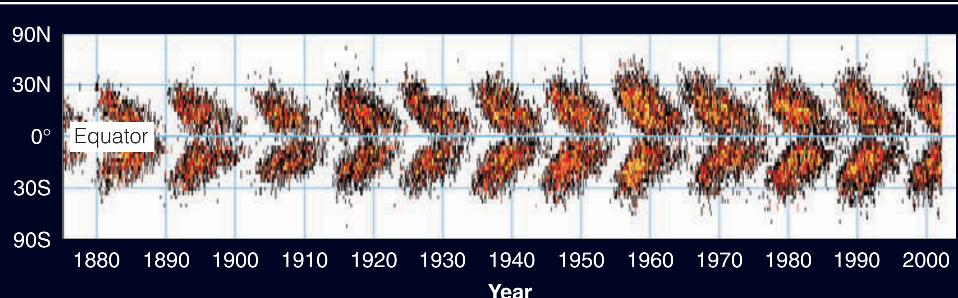
Sunspots are not shadows, but astronomers refer to the dark core of a sunspot as its umbra and the outer, lighter region as the penumbra.


Streamers above a sunspot suggest a magnetic field.

Hinode JAXA/NASA

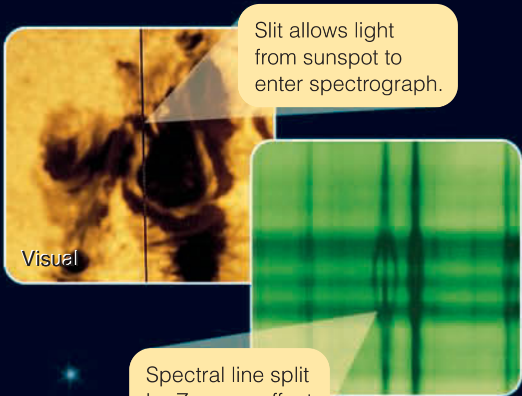


2 The number of spots visible on the sun varies in a cycle with a period of 11 years. At maximum, there are often over 100 spots visible. At minimum, there are very few.

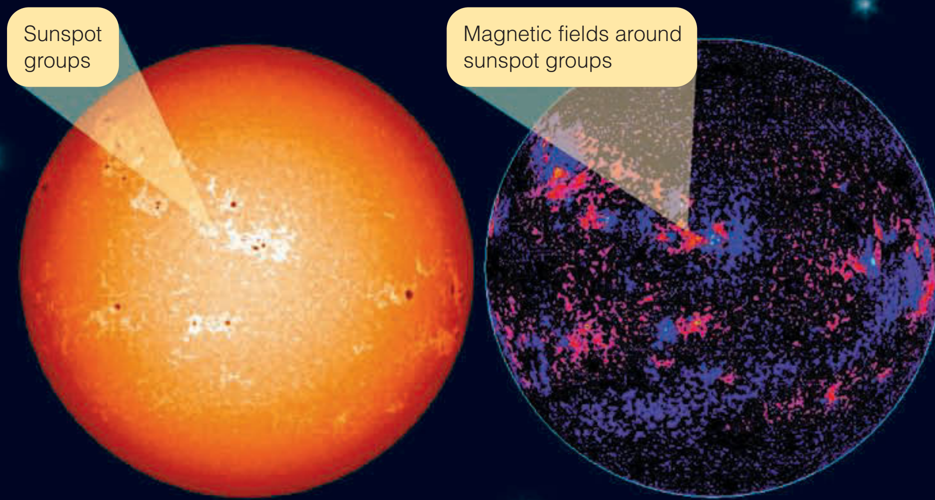


2a Early in the cycle, spots appear at high latitudes north and south of the sun's equator. Later in the cycle, the spots appear closer to the sun's equator. If you plot the latitude of sunspots versus time, the graph looks like butterfly wings, as shown in this **Maunder butterfly diagram,** named after E. Walter Maunder of Greenwich Observatory.

3 Astronomers can measure magnetic fields on the sun using the **Zeeman effect** as shown below. When an atom is in a magnetic field, the electron orbits are altered, and the atom is able to absorb a number of different wavelength photons even though it was originally limited to a single wavelength. In the spectrum, you see single lines split into multiple components, with the separation between the components proportional to the strength of the magnetic field.



AURA/NOAO/NSF

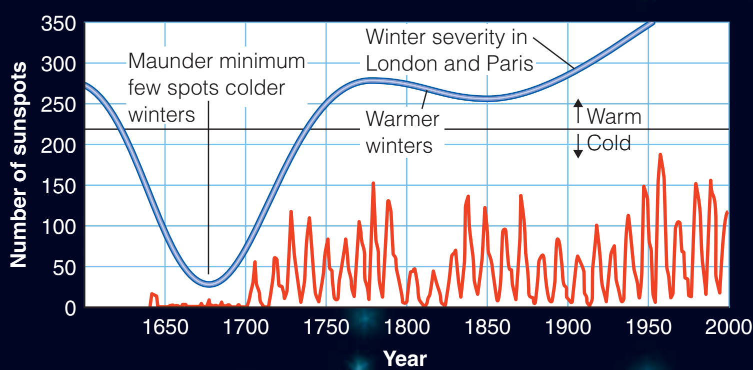


Ultraviolet filtergram | Magnetic image

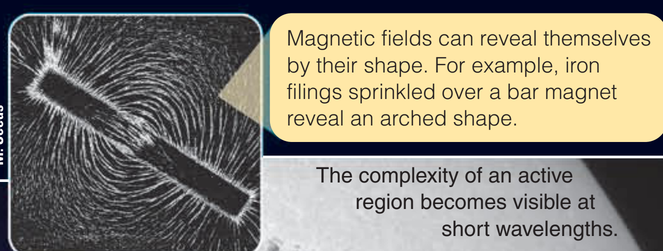
Simultaneous images

J. Harvey/NSO and HAO/NCAR

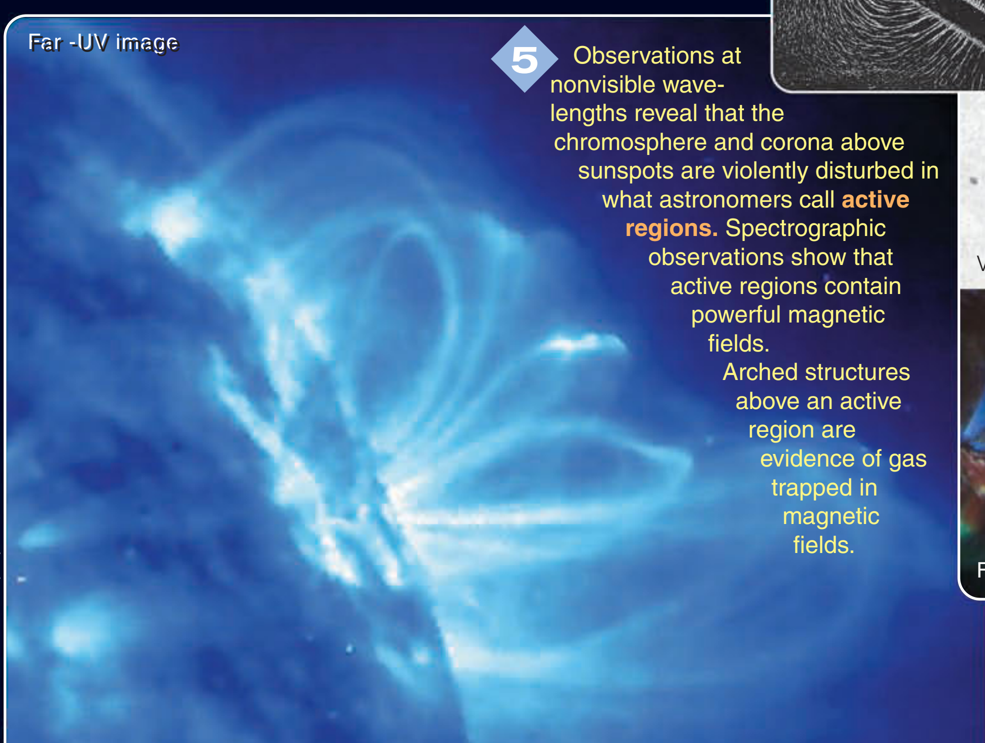
3a Images of the sun above show that sunspots contain magnetic fields a few thousand times stronger than Earth's. The strong fields are believed to inhibit gas motion below the photosphere; consequently, convection is reduced below the sunspot, and the surface there is cooler. Heat prevented from emerging through the sunspot is deflected and emerges around the sunspot, which can be detected in ultraviolet and infrared images.



4 Historical records show that there were very few sunspots from about 1645 to 1715, a phenomenon known as the **Maunder minimum.** This coincides with a period called the "little ice age," a period of unusually cool weather in Europe and North America from about 1500 to about 1850, as shown in the graph at left. Other such periods of cooler climate are known. The evidence suggests that there is a link between solar activity and the amount of solar energy Earth receives. This link has been confirmed by measurements made by spacecraft above Earth's atmosphere.



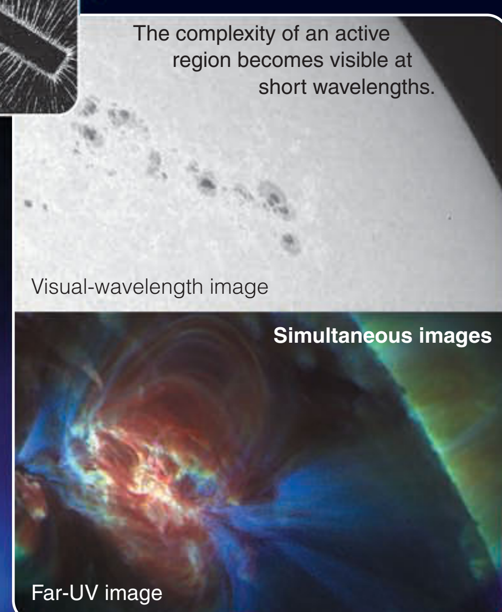
M. Seeds



SOHO/EIT, ESA and NASA

5 Observations at nonvisible wavelengths reveal that the chromosphere and corona above sunspots are violently disturbed in what astronomers call **active regions.** Spectrographic observations show that active regions contain powerful magnetic fields.

Arched structures above an active region are evidence of gas trapped in magnetic fields.



NASA/TRACE

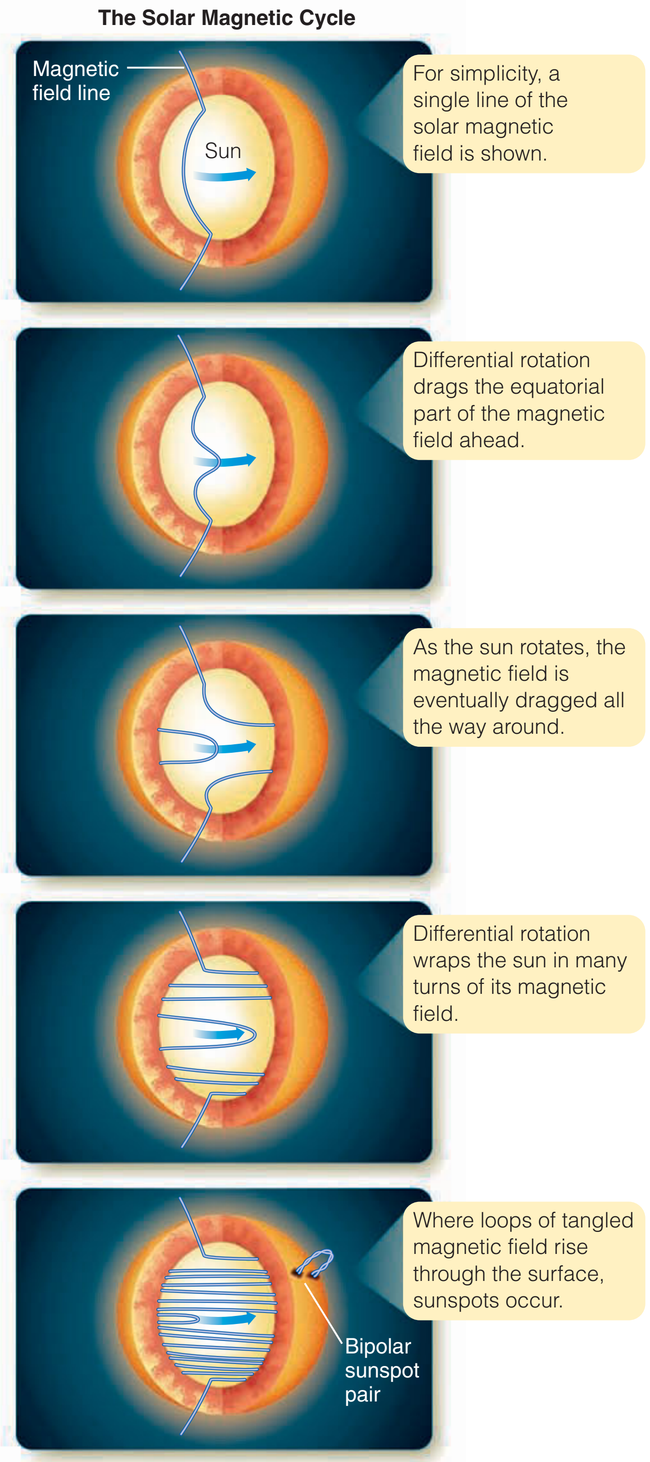


Figure 7-13

The Babcock model of the solar magnetic cycle explains the sunspot cycle as a consequence of the sun's differential rotation gradually winding up the magnetic field near the base of the sun's outer, convective layer.

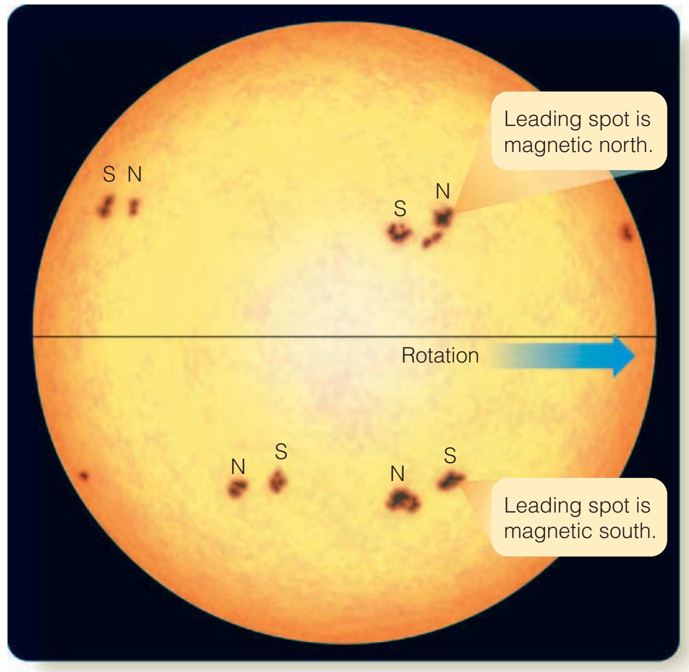


Figure 7-14

In sunspot groups, here simplified into pairs of major spots, the leading spot and the trailing spot have opposite magnetic polarity. Spot pairs in the southern hemisphere have reversed polarity from those in the northern hemisphere.

with dark spots and are rotating. Other stars have features in their spectra that vary cyclically with periods of years, suggesting that they are subject to magnetic cycles much like the sun's. At least one other star, tau Bootis, has been observed to reverse its magnetic field. Once again, the evidence tells you that the sun is a normal star.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Convection and Magnetic Fields."

Chromospheric and Coronal Activity

The solar magnetic fields extend high into the chromosphere and corona, where they produce beautiful and powerful phenomena. Study ■ **Magnetic Solar Phenomena** on pages 130–131 and notice three important points and 7 new terms:

1. All solar activity is magnetic. The arched shapes of *prominences* are produced by magnetic fields, and *filaments* are prominences seen from above.
2. Tremendous energy can be stored in arches of magnetic field, and when two arches encounter each other a *reconnection* can release powerful eruptions called *flares.* Although these eruptions occur far from Earth, they can affect us in dramatic ways, and *coronal mass ejections (CMEs)* can trigger communications blackouts and *auroras.*
3. In some regions of the solar surface, the magnetic field does not loop back. High-energy gas from these *coronal holes* flows outward and produces much of the solar wind.

You may have heard the **Common Misconception** that an auroral display in the night sky is caused by sunlight reflecting off of the ice and snow at Earth's North Pole. It is fun to think of polar bears standing on sunlit slabs off the ice, but that doesn't cause auroras. You know that auroras are produced by gases in

How Do We Know? 7-2

Confirmation and Consolidation

What do scientists do all day?
The scientific method is sometimes portrayed as a kind of assembly line where scientists crank out new hypotheses and then test them through observation. In reality, scientists don't often generate entirely new hypotheses. It is rare that an astronomer makes an observation that disproves a long-held theory and triggers a revolution in science. Then what is the daily grind of science really about?

Many observations and experiments merely confirm already-tested hypotheses. The biologist knows that all worker bees in a hive are sisters. All of the workers are female, and they all had the same mother, the queen bee. A biologist can study the DNA from many workers and confirm that hypothesis. By repeatedly confirming a hypothesis, scientists build confidence in the hypothesis and may be able to extend it. Do all of the workers in a hive have the same father, or did the queen mate with more than one male drone?

A yellow jacket is a wasp from a nest containing a queen wasp. Michael Durham/Getty Images

Another aspect of routine science is consolidation, the linking of a hypothesis to other well-studied phenomena. A biologist can study yellow jacket wasps from a single nest and discover that the wasps, too, are sisters. There must be a queen wasp who lays all of the eggs in a nest. But in a few nests, the scientist may find two sets of unrelated sister workers. Those nests must contain two queens sharing the nest for convenience and protection. From her study of wasps, the biologist consolidates what she knows about bees with what others have learned about wasps and reveals something new: That bees and wasps have evolved in similar ways for similar reasons.

Confirmation and consolidation allow scientists to build confidence in their understanding and extend it to explain more about nature.

Earth's upper atmosphere excited to glowing by energy from the solar wind.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Auroras."

◄ SCIENTIFIC ARGUMENT ►

What kind of activity would the sun have if it didn't rotate differentially?

This is a really difficult question because only one star is visible close up. Nevertheless, you can construct a scientific argument by thinking about the Babcock model. If the sun didn't rotate differentially, its equator traveling faster than its higher latitudes, then the magnetic field might not get wound up, and there might not be a solar cycle. Twisted tubes of magnetic field might not form and rise through the photosphere to produce prominences and flares, although convection might tangle the magnetic field and produce some activity. Is the magnetic activity that heats the chromosphere and corona driven by differential rotation or by convection? It is hard to guess; but, without differential rotation, the sun might not have a strong magnetic field and high-temperature gas above its photosphere.

This is very speculative, but sometimes in the critical analysis of ideas it helps to imagine a change in a single important factor and try to understand what might happen. For example, redo the argument above. **What do you think the sun would be like if it had no convection inside?**

◄ ►

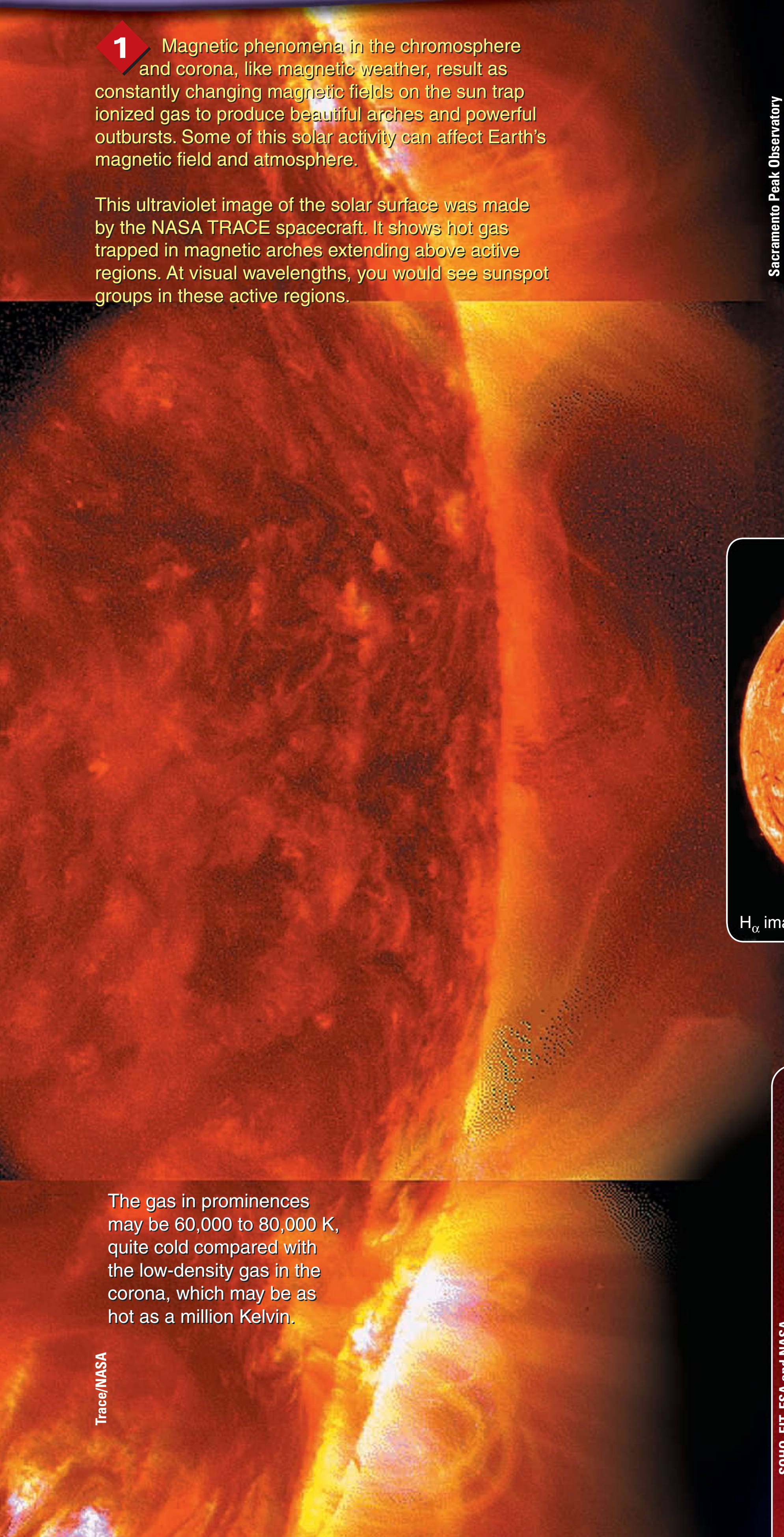
What Are We? Sunlight

We live very close to a star and depend on it for survival. All of our food comes from sunlight that was captured by plants on land or by plankton in the oceans. We either eat those plants directly or eat the animals that feed on those plants. Whether you had salad, seafood, or a cheeseburger for supper last night, you dined on sunlight, thanks to photosynthesis.

Almost all of the energy that powers human civilization comes from the sun through photosynthesis in ancient plants that were buried and converted to coal, oil, and natural gas. New technology is making energy from plant products like corn, soy beans, and sugar. It is all stored sunlight. Windmills generate electrical power, and the wind blows because of heat from the sun. Photocells make electricity directly from sunlight. Even our bodies have adapted to use sunlight to manufacture vitamin D.

Our planet is warmed by the sun, and without that warmth the oceans would be ice and much of the atmosphere would be a coating of frost. Books often refer to the sun as "our sun" or "our star." It is ours in the sense that we live beside it and by its light and warmth, but we can hardly say it belongs to us. It is more correct to say that we belong to the sun.

Magnetic Solar Phenomena

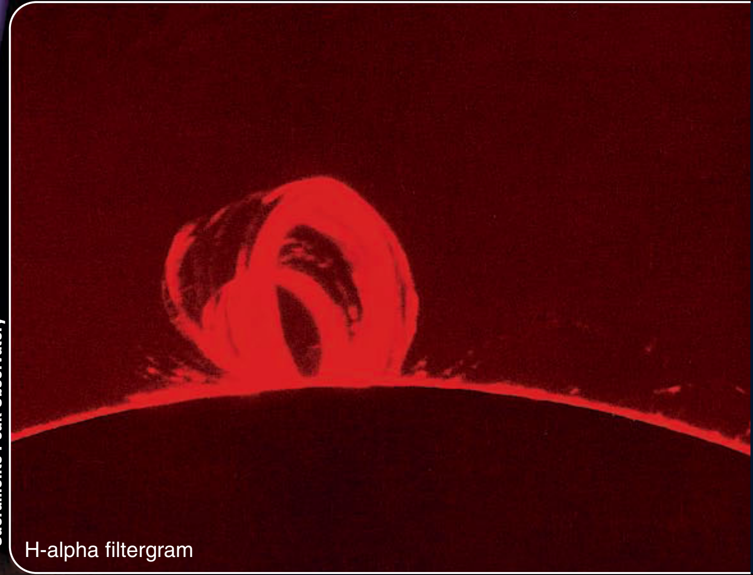


1 Magnetic phenomena in the chromosphere and corona, like magnetic weather, result as constantly changing magnetic fields on the sun trap ionized gas to produce beautiful arches and powerful outbursts. Some of this solar activity can affect Earth's magnetic field and atmosphere.

This ultraviolet image of the solar surface was made by the NASA TRACE spacecraft. It shows hot gas trapped in magnetic arches extending above active regions. At visual wavelengths, you would see sunspot groups in these active regions.

The gas in prominences may be 60,000 to 80,000 K, quite cold compared with the low-density gas in the corona, which may be as hot as a million Kelvin.

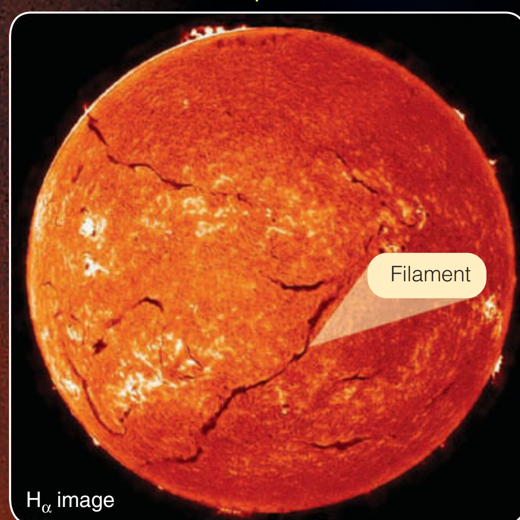
Trace/NASA



H-alpha filtergram

Sacramento Peak Observatory

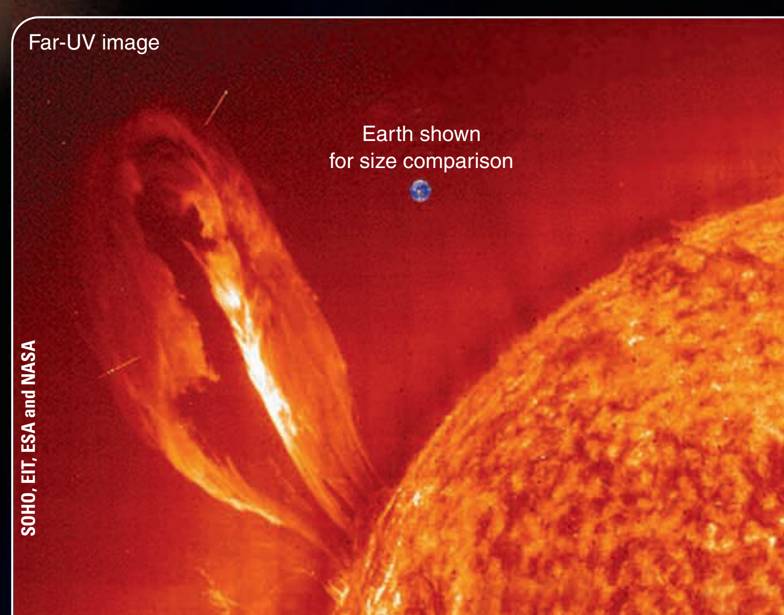
1a A **prominence** is composed of ionized gas trapped in a magnetic arch rising up through the photosphere and chromosphere into the lower corona. Seen during total solar eclipses at the edge of the solar disk, prominences look pink because of the three Balmer emission lines. The image above shows the arch shape suggestive of magnetic fields. Seen from above against the sun's bright surface, prominences form dark **filaments.**



H_α image

NOAA/SEL/USAF

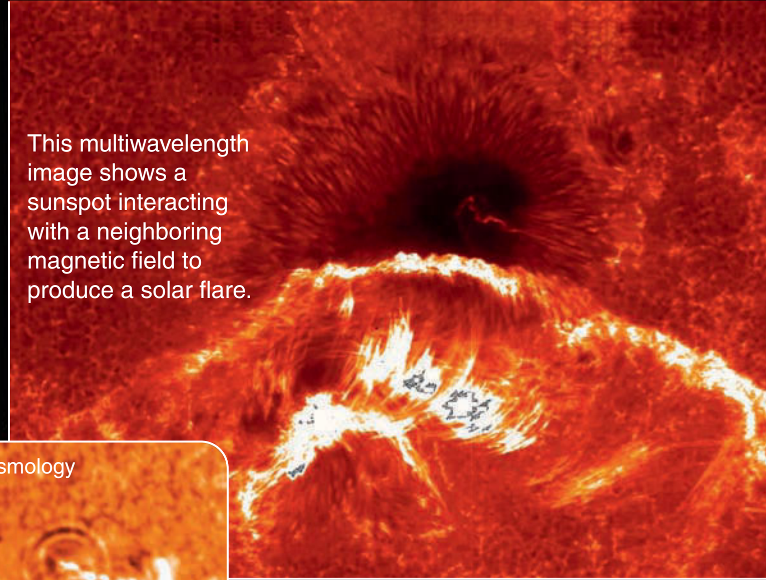
1b Quiescent prominences may hang in the lower corona for many days, whereas eruptive prominences burst upward in hours. The eruptive prominence below is many Earth diameters long.



Far-UV image

SOHO, EIT, ESA and NASA

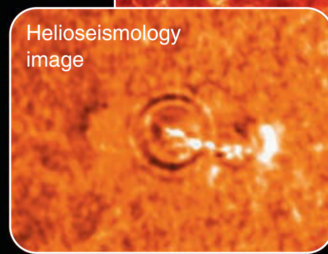
2 Solar **flares** rise to maximum in minutes and decay in an hour. They occur in active regions where oppositely directed magnetic fields meet and cancel each other out in what astronomers call **reconnections.** Energy stored in the magnetic fields is released as short-wavelength photons and as high-energy protons and electrons. X-ray and ultraviolet photons reach Earth in 8 minutes and increase ionization in our atmosphere, which can interfere with radio communications. Particles from flares reach Earth hours or days later as gusts in the solar wind, which can distort Earth's magnetic field and disrupt navigation systems. Solar flares can also cause surges in electrical power lines and damage to Earth satellites.



This multiwavelength image shows a sunspot interacting with a neighboring magnetic field to produce a solar flare.

Hinode JAXA/NASA

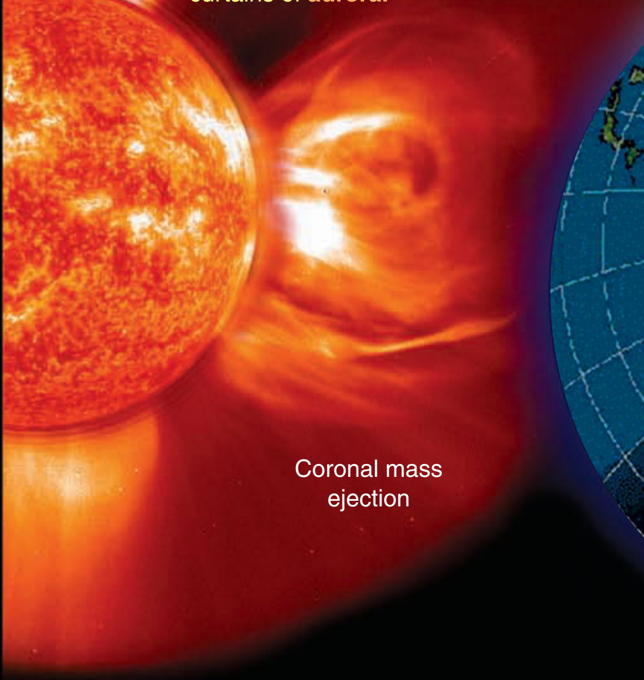
2a At right, waves rush outward at 50 km/sec from the site of a solar flare 40,000 times stronger than the 1906 San Francisco earthquake. The biggest solar flares can be a billion times more powerful than a hydrogen bomb.



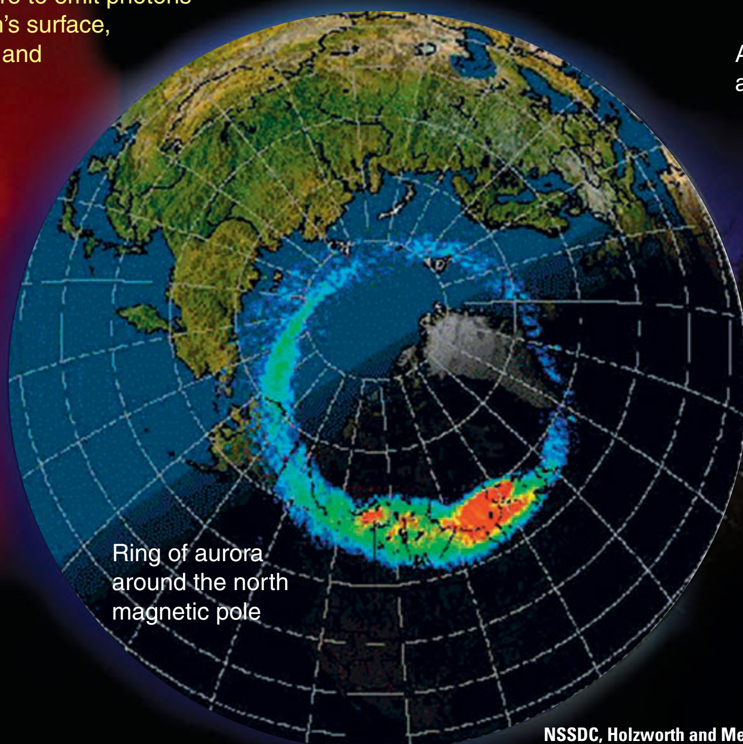
Helioseismology image

SOHO/MDI, ESA, and NASA

2b The solar wind, enhanced by eruptions on the sun, interacts with Earth's magnetic field and can create electrical currents up to a million megawatts. Those currents flowing down into a ring around Earth's magnetic poles excite atoms in Earth's upper atmosphere to emit photons as shown below. Seen from Earth's surface, the gas produces glowing clouds and curtains of **aurora.**



Coronal mass ejection



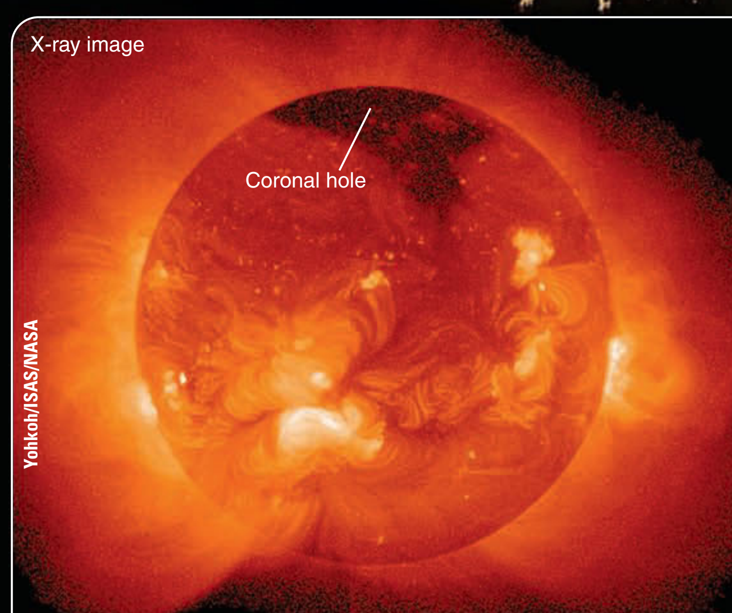
Ring of aurora around the north magnetic pole

NSSDC, Holzworth and Meng

Auroras occur about 130 km above the Earth's surface.

2c Magnetic reconnections can release enough energy to blow large amounts of ionized gas outward from the corona in **coronal mass ejections (CMEs).** If a CME strikes Earth, it can produce especially violent disturbances in Earth's magnetic field.

3 Much of the solar wind comes from **coronal holes,** where the magnetic field does not loop back into the sun. These open magnetic fields allow ionized gas in the corona to flow away as the solar wind. The dark area in this X-ray image at right is a coronal hole.



X-ray image

Yohkoh/ISAS/NASA

Study and Review

Summary

- The sun is very bright, and its light and infrared radiation can burn your eyes, so you must take great care in observing it. At sunset or sunrise when it is safe to look at the sun, you see the sun's photosphere, the level in the sun from which visible photons most easily escape. Dark **sunspots (p. 113)** come and go on the sun, but only rarely are they large enough to be visible to the unaided eye.
- The solar atmosphere consists of three layers of hot, low-density gas: the photosphere, chromosphere, and corona.
- The **granulation (p. 114)** of the photosphere is produced by **convection (p. 114)** currents of hot gas rising from below. Larger **supergranules (p. 115)** appear to be caused by larger convection currents deeper in the sun.
- The chromosphere is most easily visible during total solar eclipses, when it flashes into view for a few seconds. It is a thin, hot layer of gas just above the photosphere, and its pink color is caused by the Balmer emission lines in its spectrum.
- **Filtergrams (p. 115)** of the chromosphere reveal **spicules (p. 115)**, flamelike structures extending upward into the lower corona.
- The corona is the sun's outermost atmospheric layer and can be imaged using a **coronagraph (p. 116).** It is composed of a very-low-density, very hot gas extending many solar radii from the visible sun. Its high temperature—over 2 million K—is believed to be maintained by the magnetic field extending up through the photosphere—the **magnetic carpet (p. 116)**—and by magnetic waves coming from below the photosphere.
- Parts of the corona give rise to the **solar wind (p. 117),** a breeze of low-density ionized gas streaming away from the sun.
- Solar astronomers can study the motion, density, and temperature of gases inside the sun by analyzing the way the solar surface oscillates. Known as **helioseismology (p. 117),** this field of study requires large amounts of data and extensive computer analysis.
- Nuclear reactors on Earth generate energy through **nuclear fission (p. 119),** during which large nuclei such as uranium break into smaller fragments. The sun generates its energy through **nuclear fusion (p. 119),** during which hydrogen nuclei fuse to produce helium nuclei.
- There are only four forces in nature: the electromagnetic force, the gravitational force, the **weak force (p. 119),** and the **strong force (p. 119).** In nuclear fission or nuclear fusion, the energy comes from the strong force.
- Hydrogen fusion in the sun proceeds in three steps known as the **proton–proton chain (p. 120).** The first step in the chain combines two hydrogen nuclei to produce a heavy hydrogen nucleus called **deuterium (p. 120).** The second step forms light helium, and the third step combines the light helium nuclei to form normal helium. Energy is released as **positrons (p. 120), neutrinos (p. 120),** gamma rays, and the rapid motion of particles flying away.
- Fusion can occur only at the center of the sun because charged particles repel each other, and high temperatures are needed to give particles high enough velocities to penetrate this **Coulomb barrier (p. 121).** High densities are needed to provide large numbers of reactions.
- Neutrinos escape from the sun's core at nearly the speed of light, carrying away about 2 percent of the energy. Observations of fewer neutrinos than expected coming from the sun's core are now explained by the oscillation of neutrinos among three different types (flavors). The detection of solar neutrinos confirms the theory that the sun's energy comes from hydrogen fusion.
- Energy flows out of the sun's core as photons traveling through the **radiative zone (p. 121)** and closer to the surface as rising currents of hot gas and sinking currents of cooler gas in the **convective zone (p. 121).**
- Sunspots seem dark because they are slightly cooler than the rest of the photosphere. The average sunspot is about twice the size of Earth. They appear for a month or so and then fade away, and the number of spots on the sun varies with an 11-year cycle.
- Early in a sunspot cycle, spots appear farther from the sun's equator, and later in the cycle they appear closer to the equator. This is shown in the **Maunder butterfly diagram (p. 126).**
- Astronomers can use the **Zeeman effect (p. 127)** to measure magnetic fields on the sun. The average sunspot contains magnetic fields a few thousand times stronger than Earth's. This is part of the evidence that the sunspot cycle is produced by a solar magnetic cycle.
- The sunspot cycle does not repeat exactly each cycle, and the decades from 1645 to 1715, known as the **Maunder minimum (p. 127),** seem to have been a time when solar activity was very low and Earth's climate was slightly colder.
- Sunspots are the visible consequences of **active regions (p. 127)** where the sun's magnetic field is strong. Arches of magnetic field can produce sunspots where the field passes through the photosphere.
- The sun's magnetic field is produced by the **dynamo effect (p. 124)** operating at the base of the convection zone.
- Alternate sunspot cycles have reversed magnetic polarity, which has been explained by the **Babcock model (p. 125),** in which the **differential rotation (p. 125)** of the sun winds up the magnetic field. Tangles in the field arch above the surface and cause active regions visible to your eyes as sunspot pairs. When the field becomes strongly tangled, it reorders itself into a simpler but reversed field, and the cycle starts over.
- Other stars are too far away for starspots to be visible, but spectroscopic observations reveal that many other stars have spots and magnetic fields that follow long-term cycles like the sun's.
- Arches of magnetic field are visible as **prominences (p. 130)** in the chromosphere and corona. Seen from above in filtergrams, prominences are visible as dark **filaments (p. 130)** silhouetted against the bright chromosphere.
- **Reconnections (p. 131)** of magnetic fields can produce powerful **flares (p. 131),** sudden eruptions of X-ray, ultraviolet, and visible radiation plus high-energy atomic particles. Flares are important because they can have dramatic effects on Earth, such as communications blackouts.
- The solar wind originates in regions on the solar surface called **coronoal holes (p. 131),** where the sun's magnetic field leads out into space and does not loop back to the sun.
- **Coronal mass ejections (p. 131)** occur when magnetic fields on the surface of the sun eject bursts of ionized gas that flow outward in the solar wind. Such bursts can produce **auroras (p. 131)** and other phenomena if they strike Earth.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why can't you see deeper into the sun than the photosphere?
2. What evidence can you give that granulation is caused by convection?
3. How are granules and supergranules related? How do they differ?
4. How can astronomers detect structure in the chromosphere?

5. What evidence can you give that the corona has a very high temperature?
6. What heats the chromosphere and corona to a high temperature?
7. How are astronomers able to explore the layers of the sun below the photosphere?
8. Why does nuclear fusion require high temperatures?
9. Why does nuclear fusion in the sun occur only near the center?
10. How can astronomers detect neutrinos from the sun?
11. How can neutrino oscillation explain the solar neutrino problem?
12. What evidence can you give that sunspots are magnetic?
13. How does the Babcock model explain the sunspot cycle?
14. What does the spectrum of a prominence reveal? What does its shape reveal?
15. How can solar flares affect Earth?
16. **How Do We Know?** What does it mean when scientists say they are certain? What does scientific certainty really mean?
17. **How Do We Know?** How does consolidation extend scientific understanding?

Discussion Questions

1. What energy sources on Earth cannot be thought of as stored sunlight?
2. What would the spectrum of an auroral display look like? Why?
3. What observations would you make if you were ordered to set up a system that could warn astronauts in orbit of dangerous solar flares? Such a warning system exists.

Problems

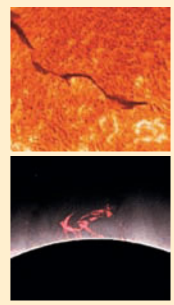
1. The radius of the sun is 0.7 million km. What percentage of the radius is taken up by the chromosphere?
2. The smallest detail visible with ground-based solar telescopes is about 1 second of arc. How large a region does this represent on the sun? (*Hint:* Use the small-angle formula.)
3. What is the angular diameter of a star like the sun located 5 ly from Earth? Is the Hubble Space telescope able to detect detail on the surface of such a star?
4. How much energy is produced when the sun converts 1 kg of mass into energy?
5. How much energy is produced when the sun converts 1 kg of hydrogen into helium? (*Hint:* How does this problem differ from Problem 4?)
6. A 1-megaton nuclear weapon produces about 4×10^{15} J of energy. How much mass must vanish when a 5-megaton weapon explodes?
7. Use the luminosity of the sun, the total amount of energy it emits each second, to calculate how much mass it converts to energy each second.
8. If a sunspot has a temperature of 4200 K and the solar surface has a temperature of 5800 K, how many times brighter is a square meter of the surface compared to a square meter of the sunspot? (*Hint:* Use the Stefan–Boltzmann law, Chapter 6.)
9. A solar flare can release 10^{25} J. How many megatons of TNT would be equivalent? (*Hint:* A 1-megaton bomb produces about 4×10^{15} J.)
10. The United States consumes about 2.5×10^{19} J of energy in all forms in a year. How many years could you run the United States on the energy released by the solar flare in Problem 9?
11. Neglecting energy absorbed or reflected by Earth's atmosphere, the solar energy hitting 1 square meter of Earth's surface is 1370 J/s. How long does it take a baseball diamond (90 ft on a side) to receive 1 megaton of solar energy?

Learning to Look

1. Whenever there is a total solar eclipse, you can see something like the image shown at right. Explain why the shape and extent of the glowing gases is different for each eclipse.

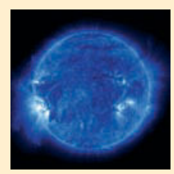
NOAO and Daniel Good

2. The two images here show two solar phenomena. What are they, and how are they related? How do they differ?



Images courtesy Daniel Good and NOAO

3. This image of the sun was recorded in the extreme ultraviolet by the SOHO spacecraft. Explain the features you see.



NASA/SOHO

8 The Family of Stars

The center of the gas cloud around the star V838 Monocerotis is ablaze with the light of brilliant stars, but fainter stars also dot the image. The crosses on the star images are produced by light diffraction in the telescope. (NASA/Hubble Heritage Project)

Guidepost

Science is based on measurement, but measurement in astronomy is very difficult. To discover the properties of stars, astronomers must use their telescopes and spectrographs in ingenious ways to solve the secret code of starlight. The result is a family portrait of the stars.

Here you will find answers to five essential questions about stars:

- **How far away are the stars?**
- **How much energy do stars make?**
- **How big are stars?**
- **How much matter do stars contain?**
- **What is the typical star like?**

With this chapter, you leave our sun behind and begin your study of the billions of stars that dot the sky. In a sense, the star is the basic building block of the universe. If you hope to understand what the universe is, what our sun is, what our Earth is, and what we are, you must understand stars.

Once you know how to find the basic properties of stars, you will be ready to trace the history of the stars from birth to death, a story that begins in the next chapter.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Ice is the silent language of the peak;
and fire the silent language of the star.

CONRAD AIKEN,
AND IN THE HUMAN HEART

DOES YOUR FAMILY include some characters? The family of stars is amazingly diverse. In a photograph, stars differ only slightly in color and brightness, but you are going to discover that some are huge and some are tiny, some are astonishingly hot and some are quite cool, some are ponderously massive and some are weenie little stars hardly massive enough to shine. If your family is as diverse as the family of stars, you must have some peculiar relatives.

Unfortunately, finding out what a star is like is quite difficult. When you look at a star, you look across a vast distance and see only a bright point of light. Just looking tells you almost nothing about a star's energy production, diameter, or mass. Rather than just look at stars, you must analyze starlight with great care. Starlight is the silent language of the sky, and it speaks volumes.

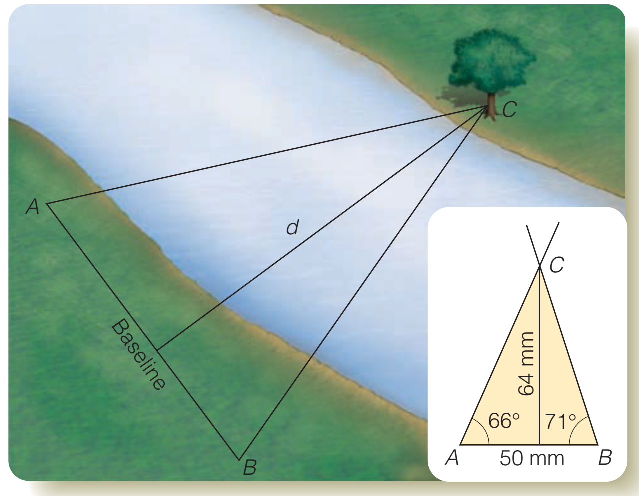
8-1 Measuring the Distances to Stars

ALTHOUGH YOU WANT to learn such things as the size and mass of stars, you immediately meet a detour. To find out almost anything about a star, you must know how far away it is. A quick detour will provide you with a method of measuring the distances to stars.

Distance is the most difficult measurement in astronomy, and astronomers have found a number of ways to estimate the distance to stars. Yet each of those ways depends on a direct geometrical method that is much like the method surveyors use to measure the distance across a river they cannot cross. You can begin by reviewing this method and then apply it to stars.

The Surveyor's Triangulation Method

To measure the distance across a river, a team of surveyors begins by driving two stakes into the ground a known distance apart. The distance between the stakes is the baseline of the measurement. The surveyors then choose a landmark on the opposite side of the river, a tree perhaps, thus establishing a large triangle marked by the two stakes and the tree. Using their surveying instruments, they sight the tree from the two ends of the baseline and measure the two angles on their side of the river (■ Figure 8-1).



■ **Figure 8-1**

You can find the distance *d* across a river by measuring the baseline and the angles *A* and *B* and then constructing a scale drawing of the triangle.

Now that they know two angles of this large triangle and the length of the side between the angles, the surveyors can find the distance across the river by simple trigonometry. Another way to find the distance is to construct a scale drawing. For example, if the baseline is 50 m and the angles are 66° and 71°, you can draw a line 50 mm long to represent the baseline. Each millimeter on your drawing is worth 1 meter. Using a protractor, you can construct angles of 66° and 71° at each end of the baseline, and then, as shown in Figure 8-1, extend the two sides until they meet at *C*. Point *C* on your drawing is the location of the tree. If you measure the height of your triangle, you would find it to be 64 mm and thus conclude that the distance from the baseline to the tree is 64 m.

Modern surveyors use computers to solve these problems, but however you solve the problem, the point is that simple triangulation can reveal the distance across a river.

The Astronomer's Triangulation Method

To find the distance to a star, you must use a very long baseline, the diameter of Earth's orbit. If you take a photograph of a nearby star and then wait 6 months, Earth will have moved halfway around its orbit. You can then take another photograph of the star. This second photograph is taken at a point in space 2 AU (astronomical units) from the point where the first photograph was taken. Thus your baseline equals the diameter of Earth's orbit, or 2 AU, and lines to the star outline a long thin triangle (■ Figure 8-2).

You then have two photographs of the same part of the sky taken from slightly different locations in space. When you examine the photographs, you will discover that the star is not in exactly the same place in the two photographs. This apparent shift in the position of the star is called *parallax,* the apparent change in the position of an object due to a change in the location of the observer. In Chapter 4, you saw an everyday example. Your thumb, held at arm's length, appears to shift position against a

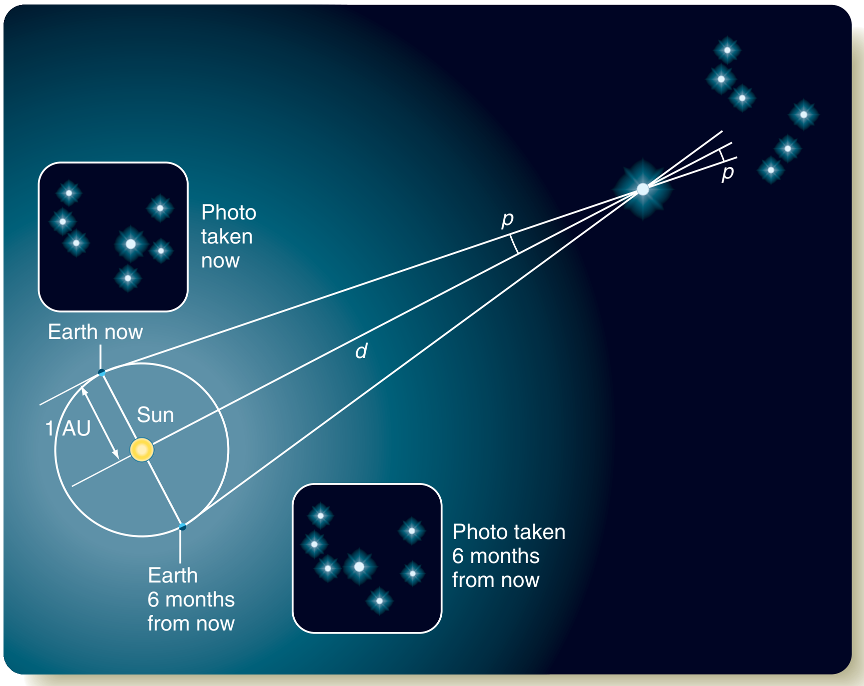


Figure 8-2

You can measure the parallax of a nearby star by photographing it from two points along Earth's orbit. For example, you might photograph it now and again in six months. Half of the star's total change in position from one photograph to the other in this example is its stellar parallax, *p*.

distant background when you look with first one eye and then with the other (see page 44). In this case, the baseline is the distance between your eyes, and the parallax is the angle through which your thumb appears to move when you change eyes. The farther away you hold your thumb, the smaller the parallax.

Because the stars are so distant, their parallaxes are very small angles, usually expressed in seconds of arc. The quantity that astronomers call **stellar parallax *(p)*** is half the total shift of the star, as shown in Figure 8-2. Astronomers measure the parallax, and surveyors measure the angles at the ends of the baseline, but both measurements reveal the same thing—the shape of the triangle and thus the distance to the object in question.

Measuring the parallax *p* is very difficult because it is such a small angle. The star nearest the sun is one of our Favorite Stars, α Centauri. It has a parallax of only 0.76 second of arc, and the more distant stars have even smaller parallaxes. To see how small these angles are, hold a piece of paper edgewise at arm's length. The thickness of the paper covers an angle of about 30 seconds of arc.

You cannot use scale drawings to find the distances to stars because the angles are so small and the distances are so large. Even for the nearest star, the triangle would have to be 300,000 times longer than it was wide. If the baseline in your drawing were 1 cm, the triangle would have to be about 3 km long. ■ Reasoning with Numbers 8-1 describes how you can find the distance from the parallax without drawing scale triangles.

The distances to stars are so large that it is not convenient to use astronomical units. As Reasoning with Numbers 8-1 explains, when you measure distance via parallax, it is convenient to use the unit of distance called a **parsec (pc).** The word *parsec* was created by combining *par*allax and *sec*ond of arc. One parsec equals the distance to an imaginary star that has a parallax of 1 second of arc. A parsec is 206,265 AU, which equals roughly 3.26 ly (light-years).*

The blurring caused by Earth's atmosphere makes star images about 1 second of arc in diameter, and that makes it difficult to measure parallax from Earth. Even when astronomers average together many observations, they cannot measure parallax with an uncertainty smaller than about 0.002 second of arc. If you measure a parallax of 0.02 second of arc, the uncertainty is about 10 percent. Ten percent is about the largest uncertainty in a parallax measurement that astronomers can comfortably tolerate, so ground-based astronomers have not been able to measure the distance to stars further distant than about 50 pc. Since the first stellar parallax was measured in 1838, ground-based astronomers have been able to measure accurate parallaxes for only about 10,000 stars.

In 1989, the European Space Agency launched the satellite Hipparcos to measure stellar parallaxes from orbit above the blurring effects of Earth's atmosphere. The satellite observed for four years, and the data were reduced by highly sophisticated software to produce two parallax catalogs in 1997. One catalog contains 120,000 stars with parallaxes 20 times more accurate than ground-based measurements. The other catalog contains over a million stars with parallaxes as accurate as ground-based parallaxes. The Hipparcos data have given astronomers new insights into the nature of stars.

The European Space Agency plans to launch the GAIA mission in a few years. It will be able to measure the parallax of a billion stars to 10 percent. NASA's planned Space Interferometry Mission will be able to measure the distances of stars out to 25,000 pc.

Go to **academic.cengage.com astronomy/seeds** to see Astronomy Exercises "Parallax I" and "Parallax II."

8-2 Intrinsic Brightness

If you see a light on a dark highway, it is hard to tell how bright it really is. It could be the brilliant headlight on a distant truck or the dim headlight on a nearby bicycle (■ Figure 8-3). How

* The parsec is used throughout astronomy because it simplifies the calculation of distance. However, there are instances in which the light-year is also convenient. Consequently, the chapters that follow use either parsecs or light-years as convenience and custom dictate.

Reasoning with Numbers | 8-1

Parallax and Distance

To find the distance to a star from its measured parallax, imagine that you observe Earth from the star. Figure 8-2 shows that the angular distance you observe between the sun and Earth equals the star's parallax p. To find the distance, recall that the small-angle formula (see Reasoning with Numbers 3-1) relates an object's angular diameter, its linear diameter, and its distance. In this case, the angular diameter is p, and the linear diameter is 1 AU. Then the small-angle formula, rearranged slightly, tells you that the distance to the star in AU is equal to 206,265 divided by the parallax in seconds of arc:

$$d = \frac{206{,}265}{p}$$

Because the parallaxes of even the nearest stars are less than 1 second of arc, the distances in AU are inconveniently large numbers. To keep the numbers manageable, astronomers have defined the parsec as their unit of distance in a way that simplifies the arithmetic. One parsec equals 206,265 AU, so the equation becomes

$$d = \frac{1}{p}$$

Thus, a parsec is the distance to an imaginary star whose parallax is 1 second of arc.

Example: The star Altair has a parallax of 0.20 second of arc. How far away is it?

Solution: The distance in parsecs equals 1 divided by 0.2, or 5 pc:

$$d = \frac{1}{0.2} = 5 \text{ pc}$$

Because 1 pc equals about 3.26 ly, Altair is about 16.3 ly away.

bright an object looks depends not only on how much light it emits but also on its distance.

A sixth-magnitude star just visible to your eye looks faint, but its apparent magnitude doesn't tell you how luminous it really is. Now that you know how to find the distance to stars, you can use those distances to figure out the intrinsic brightness of the stars. *Intrinsic* means "belonging to the thing," so, the intrinsic brightness of a star refers to the total amount of light the star emits.

Brightness and Distance

When you look at a bright light, your eyes respond to the visual wavelength energy falling on your eye's retina. The apparent brightness you perceive is related to the flux of energy entering your eye. **Flux** is the energy in joules (J) per second falling on 1 square meter. Recall that a joule is about as much energy as is released when an apple falls from a table onto the floor. One joule per second is one Watt, a common unit of energy consumption used, for example, to rate lightbulbs.

The apparent brightness of a light source is related in a simple way to its distance. Imagine that you enclosed a lightbulb at the center of a spherical screen. The light that falls on a single square meter of that screen is shown in yellow in Figure 8-3. Now imagine that you doubled the size of the spherical screen. The light that used to cover 1 square meter is now

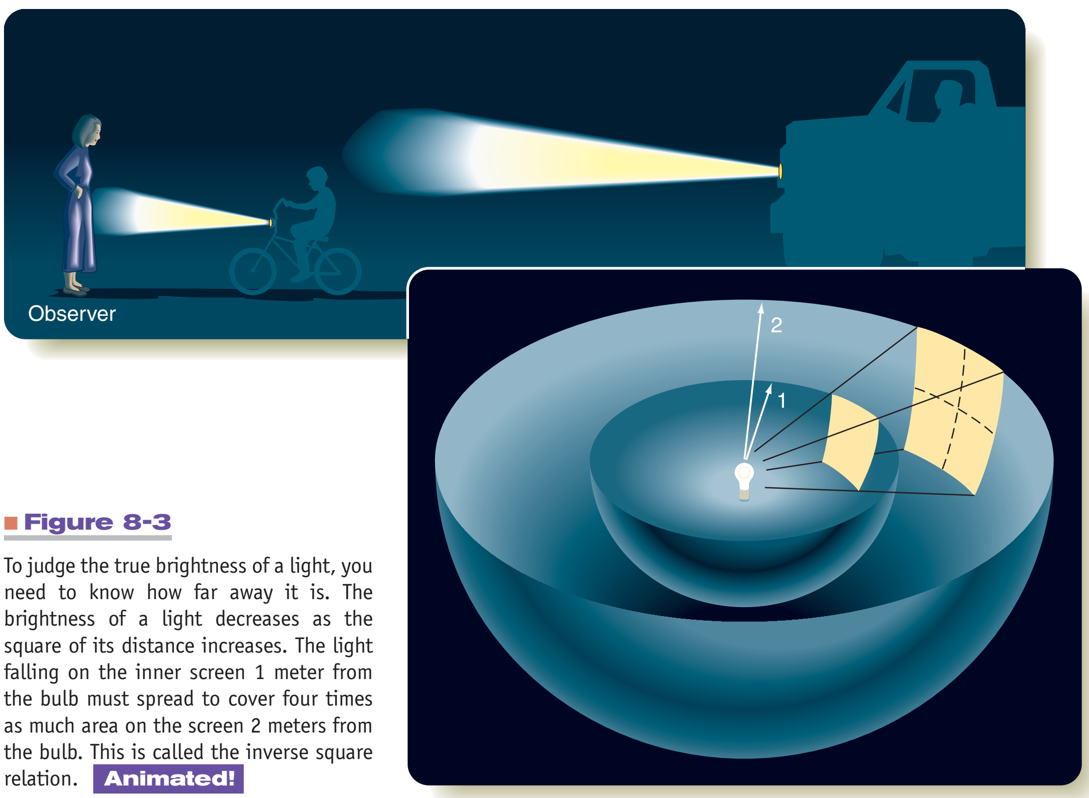


Figure 8-3

To judge the true brightness of a light, you need to know how far away it is. The brightness of a light decreases as the square of its distance increases. The light falling on the inner screen 1 meter from the bulb must spread to cover four times as much area on the screen 2 meters from the bulb. This is called the inverse square relation. **Animated!**

spread out to cover 4 square meters. Now any spot on the larger screen receives only one-fourth as much flux as a spot on the smaller screen. Likewise, if you tripled the size of the spherical screen, a spot on the screen would receive only one-ninth as much flux. The flux is inversely proportional to the square of the distance from the source. This is known as the inverse square relation. (You first encountered the inverse square relation in Chapter 4, where it was applied to the strength of gravity.)

Now you can understand how the brightness of any light source depends on its distance. Its brightness is reduced in proportion to the square of its distance, and that is an important clue to the intrinsic brightness of a star. If you know the apparent magnitude of a star and its distance from Earth, you can use the inverse square law to correct for distance and learn the intrinsic brightness of the star. Astronomers do that using a special kind of magnitude scale described in the next section.

Absolute Visual Magnitude

If all stars were the same distance from Earth, you could compare one with another and decide which was emitting more light and which less. Of course, the stars are scattered at different distances, and you can't shove them around to line them up for comparison. If, however, you know the distance to a star, you can use the inverse square relation to calculate the brightness the star would have at some standard distance. Astronomers have adopted 10 pc as the standard distance and refer to the apparent visual magnitude a star would have if it were 10 pc away as its **absolute visual magnitude (M_V).** This is an expression of the intrinsic brightness of the star.

The symbol for absolute visual magnitude is a capital *M* with a subscript V. The subscript reminds you it is a visual magnitude based only on the wavelengths of light you can see. Other magnitude systems are based on other parts of the electromagnetic spectrum, such as the infrared and ultraviolet.

It is not difficult to find the absolute visual magnitude of a nearby star. You begin by measuring the apparent visual magnitude, which is an easy task in astronomy. Then you find the distance to the star. If the star is nearby, you can measure its parallax and from that find the distance. Once you know the distance, you can use a simple formula to correct the apparent visual magnitude for the distance and find the absolute visual magnitude (■ Reasoning with Numbers 8-2).

How does the sun stack up against other stars? Astronomers can find the sun's absolute visual magnitude because they know the distance to the sun and can measure its apparent visual magnitude. The sun is tremendously bright in the sky, but it is very close. Its absolute visual magnitude is only 4.83. If the sun were only 10 pc from Earth (not a great distance in astronomy), it would look no brighter than the faintest star in the handle of the Little Dipper.

The intrinsically brightest stars known have absolute visual magnitudes of about –8, which means that such a star 10 pc from Earth would be nearly as bright as the moon. Such stars are 13 magnitudes brighter than the sun, so they must be emitting over 100,000 times more light than the sun. Yet the intrinsically faintest stars have absolute visual magnitudes of +15 or fainter. They are ten magnitudes fainter than the sun, meaning they are emitting 10,000 times less light than the sun.

The detour to find the distance to stars had led you to absolute visual magnitude, and some new insights into what stars are like. One last step will tell you how much energy stars generate.

Go to **academic.cengage.com astronomy/seeds** to see Astronomy Exercise "Apparent Brightness and Distance."

Luminosity

The **luminosity (*L*)** of a star is the total energy the star radiates in one second. Hot stars emit a great deal of ultraviolet radiation that you can't see, and cool stars emit infrared. Because absolute visual magnitude includes only visible radiation, astronomers must add a small correction to make up for the invisible energy. Then they can calculate the luminosity of the star from its absolute magnitude.

To make that calculation, you compare the star with the sun. If the corrected absolute magnitude of the star is one magnitude brighter than the sun, then it must be 2.5 times more luminous. If it is five magnitudes brighter, then it must be 100 times more luminous. Astronomers would write the luminosity of the first star as 2.5 $L_\odot$ and the luminosity of the second star as 100 $L_\odot$. To find the luminosity of a star in joules per second, you can just multiply by the luminosity of the sun, 3.8×10^{26} J/s. For example, Favorite Star Aldebaran has a luminosity of about 150 $L_\odot$, which corresponds to 6×10^{28} J/s.

The most luminous stars emit roughly a million times more energy than the sun, and the least luminous stars emit over a thousand times less. Although stars look similar in the sky, they can emit astonishing different amounts of energy. The most luminous emit at least a billion times more energy per second than the least luminous. Clearly, the family of stars contains some interesting characters.

◄ SCIENTIFIC ARGUMENT ►

How can two stars look the same in the sky but have dramatically different luminosities?

You can answer this question by building a scientific argument that relates three factors: the appearance of a star, its true luminosity, and its distance. The further away a star is, the fainter it looks, and that is just the inverse square law. Favorite Stars Vega and Rigel have the same apparent visual magnitude, so your eyes must be receiving the same amount of light from them. But Rigel is much more luminous than Vega, so it must be further away. Parallax observations from the Hipparcos satellite confirm that Rigel is 31 times further away than Vega.

Distance is often the key to understanding the brightness of stars, but temperature can also be important. Build a scientific argument to answer the following: **Why must astronomers make a correction in converting**

Reasoning with Numbers | 8-2

Absolute Magnitude

Apparent visual magnitude tells you how bright a star looks (see Reasoning with Numbers 2-1), but absolute visual magnitude tells you how bright the star really is. The absolute visual magnitude M_V of a star is the apparent visual magnitude of the star if it were 10 pc away. If you know a star's apparent visual magnitude and its distance, you can calculate its absolute visual magnitude. The equation that allows this calculation relates apparent visual magnitude m_V, distance in parsecs d, and absolute visual magnitude M_V:

$$m_V - M_V = -5 + 5 \log_{10}(d)$$

Sometimes it is convenient to rearrange the equation and write it in this form:

$$d = 10^{(m_V - M_V + 5)/5}$$

It is the same equation, so you can use whichever form is most convenient in a given problem. If you know the distance, the first form of the equation is convenient, but if you are trying to find the distance, the second form of the equation is best.

Example: Favorite Star Polaris is 132 pc from Earth and has an apparent magnitude of 2.5. What is its absolute visual magnitude?

Solution: A pocket calculator tells you that $\log_{10}(132)$ equals 2.12, so you substitute into the first equation to get

$$2.5 - M_V = -5 + 5(2.12)$$

Solving for M_V tells you that the absolute visual magnitude of Polaris is −3.1. If it were only 10 pc from Earth, it would dominate the night sky.

the absolute visual magnitude of very hot or very cool stars into luminosities?

◄ ►

8-3 The Diameters of Stars

Now that you know the luminosities of stars, you can find their diameters. You know little about stars until you know their diameters. Are they all the same size as the sun, or are some larger and some smaller?

Recall that astronomers can not see the stars as disks through astronomical telescopes (Chapter 5). All stars look like points of light no matter how big the telescope. Nevertheless, there is a way to find out how big stars really are. If you know their temperatures and luminosities, you can find their diameters. This relationship will introduce you to the most important diagram in astronomy, where you will discover more of the stars' family secrets.

Luminosity, Radius, and Temperature

To use the luminosity and temperature of a star to learn its diameter, you must first understand the two factors that affect a star's luminosity, surface area and temperature. You can eat dinner by candlelight because a candle flame has a small surface area. Although the flame is very hot, it cannot radiate much heat; it has a low luminosity. However, if the candle flame were 12 ft tall, it would have a very large surface area from which to radiate, and, although it might be no hotter than a normal candle flame, its luminosity would drive you from the table (■ Figure 8-4).

In a similar way, a hot star may not be very luminous if it has a small surface area. It could be highly luminous, however, if it were larger and had a larger surface area from which to radiate. Even a cool star could be luminous if it had a large surface area. Because of this dependence on both temperature and surface area, you need to separate the effects of temperature and surface area, and then you can find the diameters of stars. (See ■ Reason-

■ Figure 8-4

Molten lava pouring from a volcano is not as hot as a candle flame, but a lava flow has more surface area and radiates more energy than a candle flame. Approaching a lava flow without protective gear is dangerous. (Karafft/Photo Researchers, Inc.)

Reasoning with Numbers | 8-3

Luminosity, Radius, and Temperature

The luminosity L of a star depends on two things—its size and its temperature. If the star has a large surface area from which to radiate, it can radiate a great deal. Recall from our discussion of black body radiation in Reasoning with Numbers 6-1 that the amount of energy emitted per second from each square meter of the star's surface is σT^4. Thus, the star's luminosity can be written as its surface area in square meters times the amount it radiates from each square meter:

$$L = \text{area} \times \sigma T^4$$

Because a star is a sphere, you can use the formula area = $4\pi R^2$. Then the luminosity is

$$L = 4\pi R^2 \sigma T^4$$

This seems complicated, but if you express luminosity, radius, and temperature in terms of the sun, you get a much simpler form:*

$$\frac{L}{L_\odot} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{T}{T_\odot}\right)^4$$

Example A: Suppose you want to find the luminosity of a star that is 10 times the sun's radius but only half as hot. How luminous is it?

Solution:

$$\frac{L}{L_\odot} = \left(\frac{10}{1}\right)^2 \left(\frac{1}{2}\right)^4 = \frac{100}{1} \times \frac{1}{16} = 6.25$$

The star has 6.25 times the sun's luminosity.

You can also use this formula to find diameters.

Example B: Suppose you found a star whose absolute magnitude is +1 and whose spectrum shows it is twice the sun's temperature. What is the diameter of the star?

Solution: The star's absolute magnitude is four magnitudes brighter than the sun, and you recall from Reasoning with Numbers 2-1 that four magnitudes is a factor of 2.512^4, or about 40. The star's luminosity is therefore about 40 $L_\odot$. With the luminosity and temperature, you can find the radius:

$$\frac{40}{1} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{2}{1}\right)^4$$

Solving for the radius you get:

$$\left(\frac{R}{R_\odot}\right)^2 = \frac{40}{2^4} = \frac{40}{16} = 2.5$$

So the radius is

$$\frac{R}{R_\odot} = \sqrt{2.5} = 1.58$$

The star is 58 percent larger in radius than the sun.

* In astronomy the symbols ⊙ and ⊕ refer respectively to the sun and Earth. Thus $L_\odot$ refers to the luminosity of the sun, $T_\odot$ refers to the temperature of the sun, and so on.

ing with Numbers 8-3.) Astronomers use a special diagram to sort the stars by temperature and size.

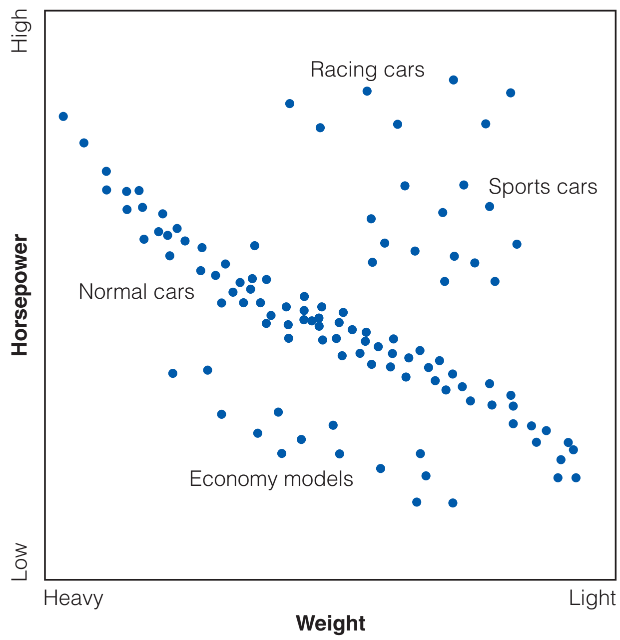
The H–R diagram

The **Hertzsprung–Russell (H–R) diagram,** named after its originators, Ejnar Hertzsprung and Henry Norris Russell, is a graph that separates the effects of temperature and surface area on stellar luminosities and enables astronomers to sort stars according to their diameters. Before you explore the details of the H–R diagram, try looking at a similar diagram you might use to sort automobiles.

You can plot a diagram such as ■ Figure 8-5, showing horsepower versus weight for various makes of cars. In general, the more a car weighs, the more horsepower it has. Most cars fall somewhere along the sequence of cars, running from heavy, high-powered cars at the upper left to light, low-powered models at the lower right. You might call this the main sequence of cars. But some cars have much more horsepower than normal for their weight—the sport or racing models—and lie higher in the diagram. Other cars, the economy models, have less power than normal for cars of the same weight and fall lower in the diagram. Just as this diagram sorts cars into family groups, so the H–R diagram sorts stars into groups according to size.

The H–R diagram is a graph with luminosity on the vertical axis and temperature on the horizontal axis. A star is represented by a point on the graph that marks its luminosity and its temperature. The H–R diagram in ■ Figure 8-6 also contains a scale of spectral type across the top. Because a star's spectral type is determined by its temperature, you could use either spectral type or temperature on the horizontal axis.

In an H–R diagram, the location of a point tells you a great deal about the star it represents. Points near the top of the diagram represent very luminous stars, and points near the bottom represent very-low-luminosity stars. Also, points near the right edge of the diagram represent very cool stars, and points near the



■ **Figure 8-5**

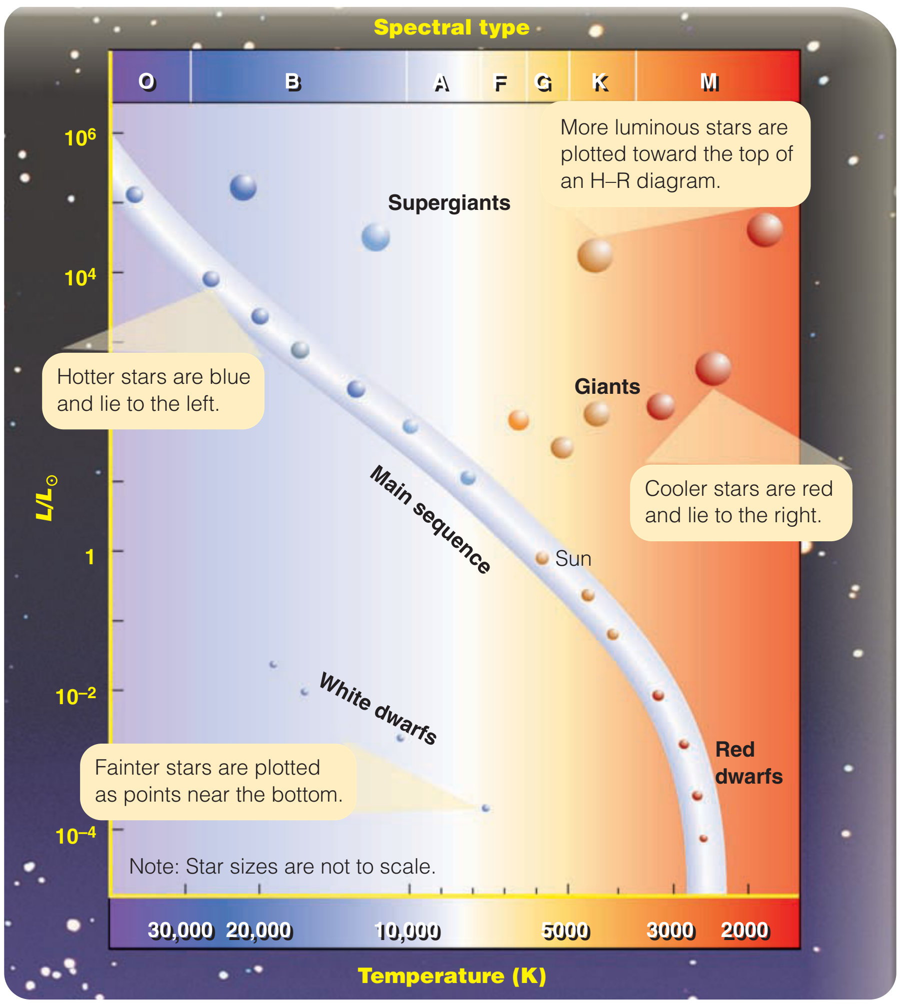
You could analyze automobiles by plotting their horsepower versus their weight and thus reveal relationships between various models. Most would lie somewhere along the main sequence of "normal" cars.

These are called **giant** stars, and they are roughly 10 to 100 times larger than the sun. There are even **supergiant** stars at the top of the H–R diagram that are over a thousand times the sun's diameter.

At the bottom of the H–R diagram lie the economy models, stars that are very low in luminosity because they are very small. At the bottom end of the main sequence, the **red dwarfs** are not only small, but they are also cool, and that gives them low luminosities. In contrast, the **white dwarfs** lie in the lower left of the H–R diagram and are lower in luminosity than you would expect, given their high temperatures. Although some white dwarfs are among the hottest stars known, they are so small they have very little surface area from which to radiate, and that limits them to low luminosities.

■ **Figure 8-6**

In an H–R diagram, a star is represented by a dot that shows the luminosity and temperature of the star. The background color in this diagram indicates the temperature of the stars. The sun is a yellow-white G2 star. Most stars fall along a sequence running from hot luminous stars at upper left to cool low-luminosity stars at lower right. The exceptions — giants, supergiants, and white dwarfs — are discussed in the text.



left edge of the diagram represent very hot stars. Notice in the H–R diagram in Figure 8-6 how the artist has used color to represent temperature. In the diagram, the red stars are cool, and blue stars are hot.

Astronomers use H–R diagrams so often that they usually skip the words "the point that represents the star." Rather, they will say that a star is located in a certain place in the diagram. The location of a star in the H–R diagram has nothing to do with the location of the star in space. Furthermore, a star may move in the H–R diagram as it ages and its luminosity and temperature change, but such motion in the diagram has nothing to do with the star's motion in space.

The **main sequence** is the region of the H–R diagram running from upper left to lower right. It includes roughly 90 percent of all normal stars. In Figure 8-6, the main sequence is represented by a curved line with dots for stars plotted along it. As you might expect, the hot main-sequence stars are more luminous than the cool main-sequence stars.

Notice in the H–R diagram that some cool stars lie above the main sequence. Although they are cool, they are luminous, and that must mean they are larger and have more surface area than main sequence stars of the same temperature.

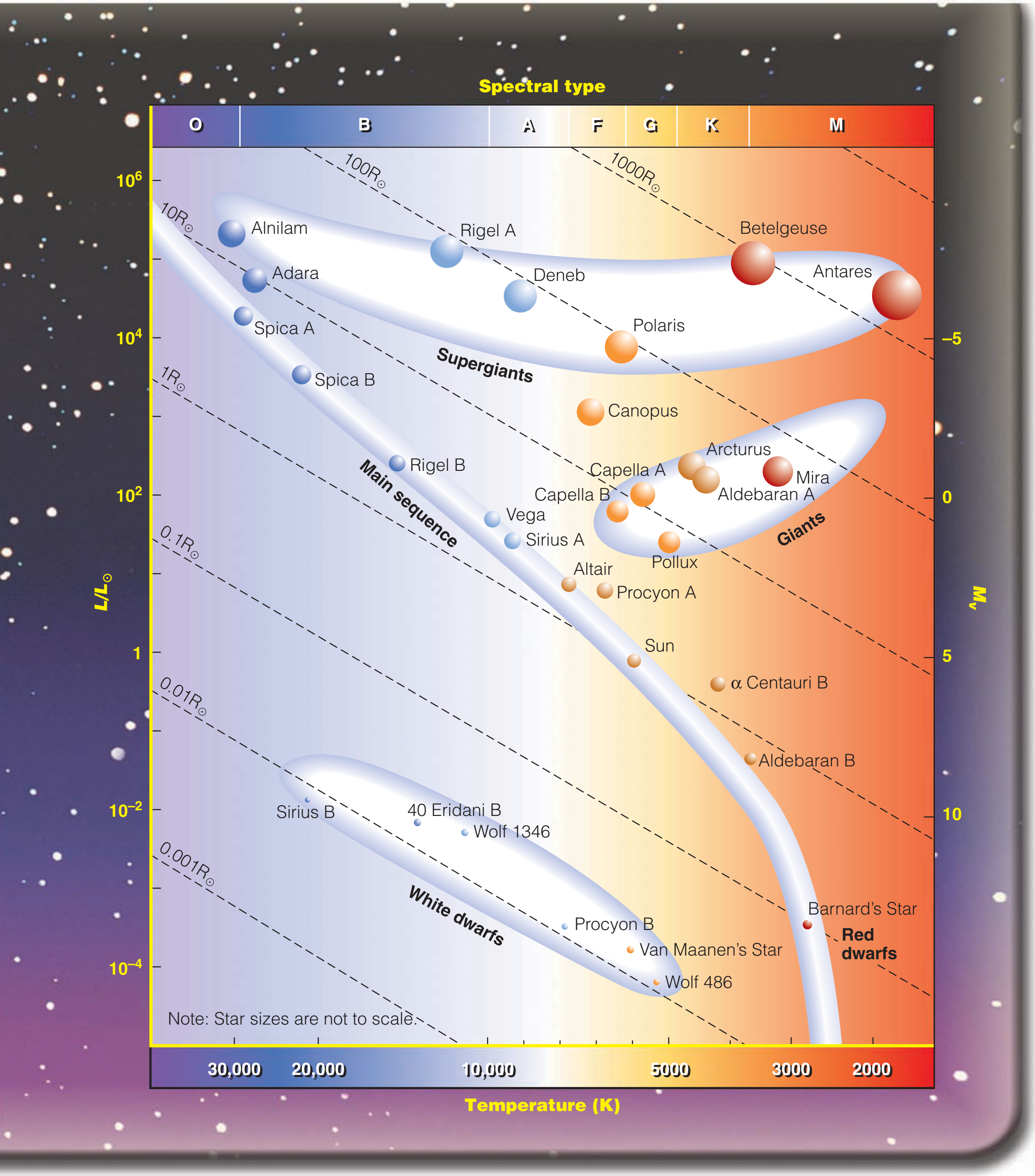
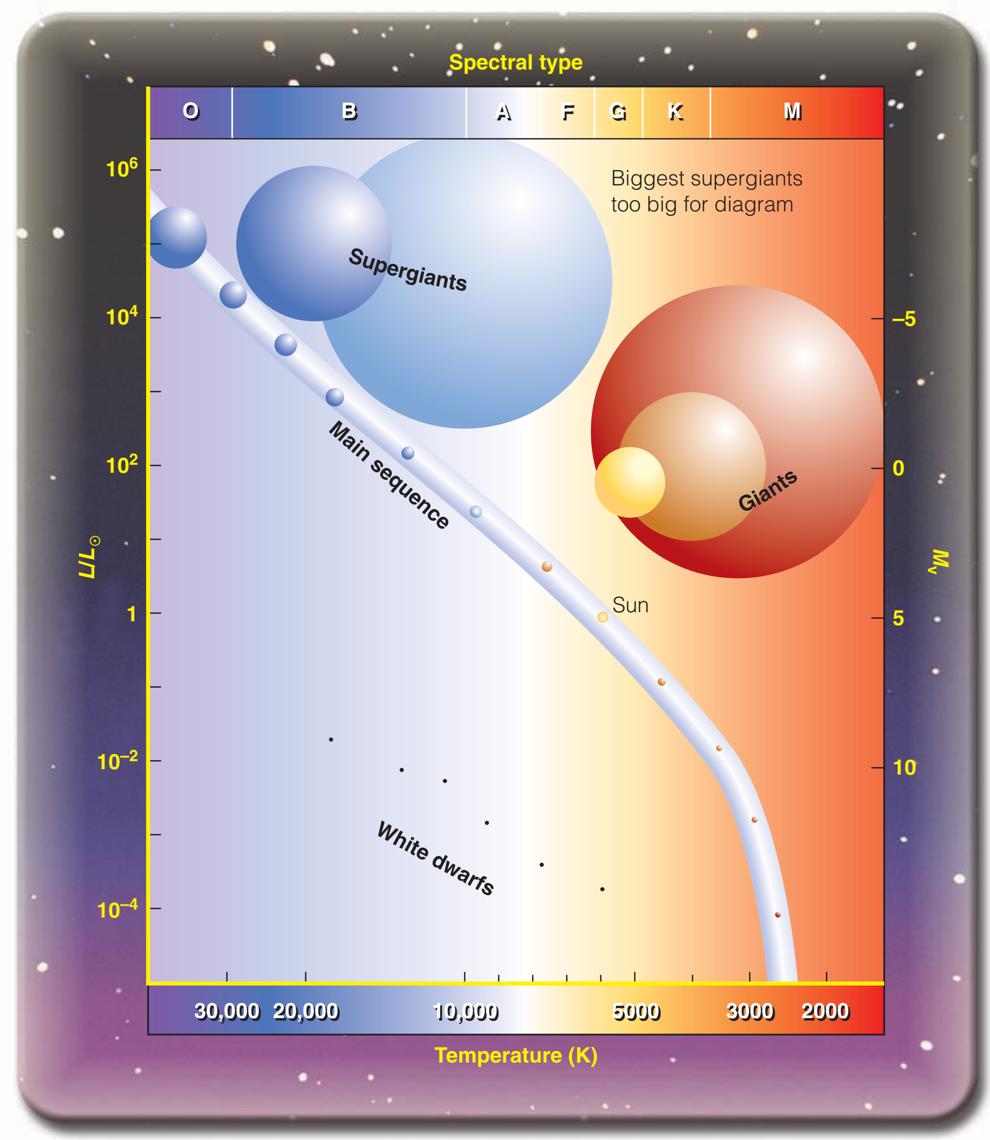


Figure 8-7

An H–R diagram showing the luminosity and temperature of many well-known stars. The dashed lines are lines of constant radius. The star sizes on this diagram are not to scale; try to sketch in the correct sizes for supergiants and white dwarfs using the size of the sun as a guide. (Individual stars that orbit each other are designated A and B, as in Spica A and Spica B.)

The equation in Reasoning with Numbers 8-3 can be used to draw precise lines of constant radius across the H–R diagram, and these lines slope down and to the right across the diagram because cooler stars are fainter than hotter stars of the same size. Figure 8-7 plots the luminosities and temperatures of a number of well-known stars along with lines of constant radius. For example, locate the line labeled $1R_\odot$ (1 solar radius) and notice that it passes through the point representing the sun. Any star whose point is located along this line has a radius equal to the sun's. Next, look at the rest of the stars along the main sequence. They range from a tenth the size of the sun to about ten times as large. Even though the main sequence slopes dramatically down to the right across the diagram, most main-sequence stars are similar in size. In contrast, the white dwarfs at the lower left of the diagram are extremely small—only about the size of Earth—and the giants and supergiants at the upper right are extremely large compared to the stars of the main sequence.

Notice the great range of sizes among stars. The largest stars are 100,000 times larger than the tiny white dwarfs. If the sun were a tennis ball, the white dwarfs would be grains of sand, and



■ **Figure 8-8**

The relative sizes of stars. Giant stars are 10 to 100 times larger than the sun, and white dwarfs are about the size of Earth. (The dots representing white dwarfs here are much too large.) The larger supergiants are 1000 times larger in diameter than the sun and would be about a meter in diameter in this diagram.

the largest supergiants would be as big as football fields (■ Figure 8-8).

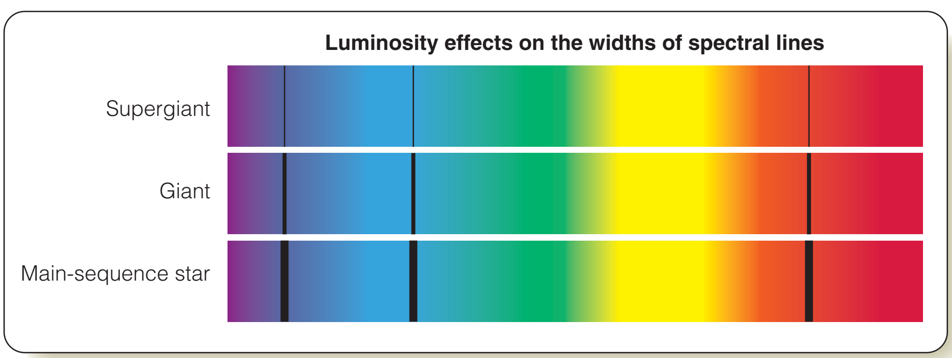
Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Stefan–Boltzmann Law II."

Luminosity Classification

A star's spectrum contains clues as to whether it is a main-sequence star, a giant, or a supergiant. The larger a star is, the less dense its atmosphere is, and that can affect the widths of spectral lines.

If the atoms that produce these lines collide often in a dense gas, their energy levels become distorted and their spectral lines broadened. Hydrogen Balmer lines are an example. In the spectrum of a main-sequence star, the Balmer lines are broad because the star's atmosphere is dense and the hydrogen atoms collide often. In the spectrum of a giant star, the lines are narrower (■ Figure 8-9) because the giant star's atmosphere is less dense, and the hydrogen atoms collide less often. In the spectrum of a supergiant star, the Balmer lines are very narrow.

You can look at a star's spectrum and tell roughly how big it is. Size categories derived from spectra are called **luminosity classes,** because the size of the star is the dominating factor in determining luminosity. Supergiants, for example, are very luminous because they are very large. The luminosity classes are represented by the roman nu-



■ **Figure 8-9**

These schematic spectra show how the widths of spectral lines reveal a star's luminosity classification. Supergiants have very narrow spectral lines, and main-sequence stars have broad lines. In addition, certain spectral lines are more sensitive to this effect than others, so an experienced astronomer can inspect a star's spectrum and determine its luminosity classification.

merals I through V, with supergiants further subdivided into types Ia and Ib, as follows:

Luminosity Classes

Ia Bright supergiant
Ib Supergiant
II Bright giant
III Giant
IV Subgiant
V Main-sequence star

You can distinguish between the bright supergiants (Ia) such as Rigel and the regular supergiants (Ib) such as Polaris, the North Star. The star Adhara is a bright giant (II), Aldebaran is a giant (III), and Altair is a subgiant (IV). Sirius and Vega, like the sun, are main-sequence stars (V). When you describe a star, its luminosity class appears after the spectral type, as in G2 V for the sun. White dwarfs don't enter into this classification, because their spectra are peculiar. Notice that some of our Favorite Stars are unusual; next time you look at Polaris, remind yourself that it is a supergiant.

The positions of the luminosity classes on the H–R diagram are shown in ■ Figure 8-10. Remember that these are rather broad classifications and that the lines on the diagram are only approximate. A star of luminosity class III may lie slightly above or below the line labeled III. A scale of absolute magnitude has been added to the right edge of this H–R diagram for easy reference.

Luminosity classification is subtle and not too accurate, but it is important in modern astronomy. As you will see in the next section, luminosity classification provides a way to estimate the distance to stars that are too far away to have measurable parallaxes.

Spectroscopic Parallax

Astronomers can measure the stellar parallax of nearby stars, but most stars are too distant to have measurable parallaxes. Astronomers can estimate the distances to these stars by the process called **spectroscopic parallax**—the estimation of the distance to a star from its spectral type, luminosity class, and apparent magnitude. Spectroscopic parallax is not an actual measure of parallax, but it does tell you the distance to the star.

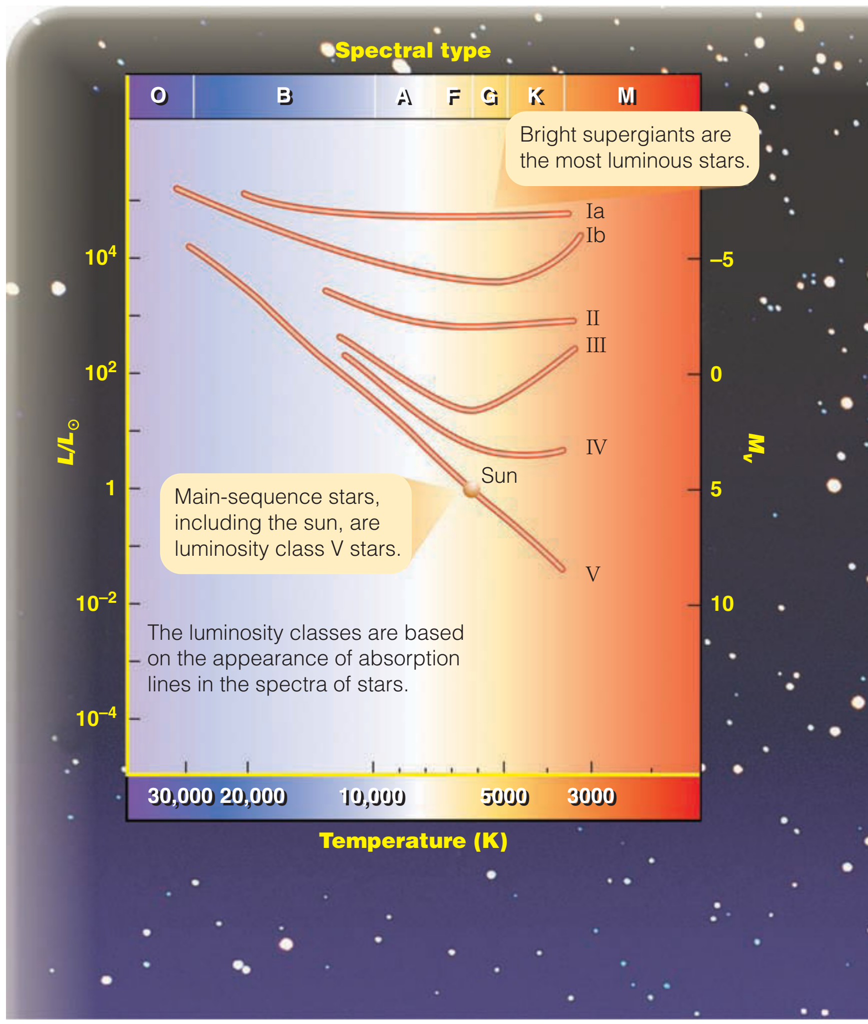
Spectroscopic parallax relies on the location of the star in the H–R diagram. If you record the spectrum of a star, you can determine its spectral class, and that tells you its horizontal location in the H–R diagram. You can also determine its luminosity class by looking at the widths of its spectral lines, and that tells you the star's vertical location in the diagram. Once you plot the point that represents the star in the H–R diagram, you can read off its absolute magnitude. As you have learned earlier in this chapter, you can find the distance to a star by comparing its apparent and absolute magnitudes.

For example, our Favorite Star Betelgeuse is classified M2 Ia, and its apparent magnitude is about 0.05. You can plot this star in an H–R diagram such as Figure 8-10, where you would find that it should have an absolute magnitude of about −7.2. Using the apparent and absolute magnitudes, you can then find the distance using the equation in Reasoning with Numbers 8-2. Spectroscopic parallax places Betelgeuse about 350 pc from Earth. More accurate measurements made by the Hipparcos satellite reveal the distance to be 520 pc, so the result derived from spectroscopic parallax is only approximate. Obviously a direct measurement of the parallax is better, but for more distant stars spectroscopic parallax is often the only way to find their distance.

◄ SCIENTIFIC ARGUMENT ►

What evidence can you give that giant stars really are bigger than the sun?

Scientific arguments are based on evidence, so you need to proceed step-by-step here. Stars exist that have the same spectral type as the sun but are clearly more luminous. Capella, for example, is a G star with an absolute magnitude of 0. Because it is a G star, it must have about the same



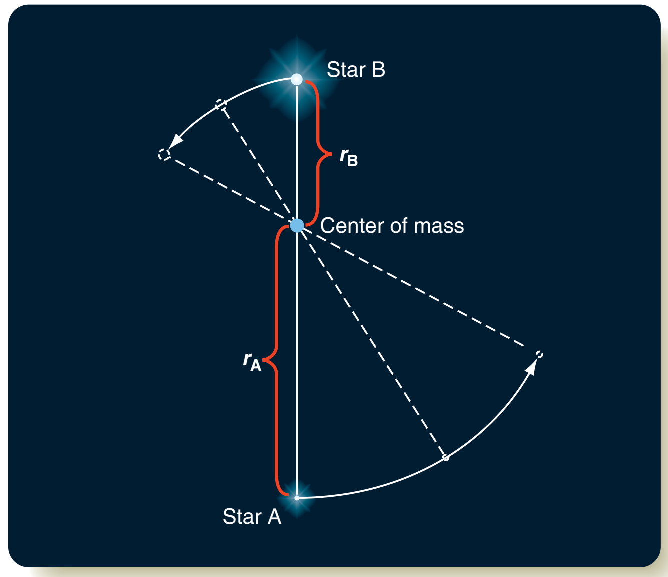
■ **Figure 8-10**
The approximate location of the luminosity classes on the H–R diagram.

temperature as the sun, but its absolute magnitude is almost five magnitudes brighter than the sun's. A magnitude difference of five magnitudes corresponds to an intensity ratio of 100, so Capella must be about 100 times more luminous than the sun. If it has the same surface temperature as the sun but is 100 times more luminous, then it must have a surface area 100 times greater than the sun's. Because the surface area of a sphere is proportional to the square of the radius, Capella must be ten times larger in radius. That is clear observational evidence that Capella is a giant star.

In Figure 8-7, you can see that Procyon B is a white dwarf slightly warmer than the sun but about 10,000 times less luminous. Build a scientific argument based on evidence to resolve this question. **Why do astronomers conclude that white dwarfs must be small stars?**

8-4 The Masses of Stars

To UNDERSTAND STARS well, you must find out how much matter stars contain, that is, their masses. Do they all contain about the same mass as the sun, or are some more massive and others less? Unfortunately, it's difficult to determine the mass of a star. Looking through a telescope, you see only a point of light that reveals nothing about the mass of the star. Gravity is the key. Matter produces a gravitational field, and you can figure out how much matter a star contains if you watch another object move through the star's gravitational field. To find the masses of stars, you must study **binary stars,** two stars that orbit each other.



■ **Figure 8-11**

As stars in a binary star system revolve around each other, the line connecting them always passes through the center of mass, and the more massive star is always closer to the center of mass. **Animated!**

Binary Stars in General

The key to finding the mass of a binary star is understanding orbital motion. Chapter 4 illustrated orbital motion by asking you to imagine a cannonball fired from a high mountain (see page 62). If Earth's gravity didn't act on the cannonball, it would follow a straight-line path and leave Earth forever. Because Earth's gravity pulls it away from its straight-line path, the cannonball follows a curved path around Earth—an orbit. When two stars orbit each other, their mutual gravitation pulls them away from straight-line paths and makes them follow closed orbits around a point between the stars.

Each star in a binary system moves in its own orbit around the system's center of mass, the balance point of the system. If the stars were connected by a massless rod and placed in a uniform gravitational field such as that near Earth's surface, the system would balance at its center of mass like a child's seesaw (see page 63). If one star is more massive than its companion, then the more massive star is closer to the center of mass and travels in a smaller orbit, while the lower-mass star whips around in a larger orbit (■ Figure 8-11). The ratio of the masses of the stars M_A/M_B equals r_B/r_A, the inverse of the ratio of the radii of the orbits. If one star has an orbit twice as large as the other star's orbit, then it must be half as massive. Getting the ratio of the masses is easy, but that doesn't tell you the individual masses of the stars, which is what you really want to know. That takes further analysis.

To find the mass of a binary star system, you must know the size of the orbits and the orbital period—the length of time the stars take to complete one orbit. The smaller the orbits are and the shorter the orbital period is, the stronger the stars' gravity must be to hold each other in orbit. For example, if two stars whirl rapidly around each other in small orbits, then their gravity must be very strong to prevent their flying apart. Such stars would have to be very massive. From the size of the orbits and the orbital period, you can figure out how much mass the stars contain, as explained in ■ Reasoning with Numbers 8-4. Such calculations yield the total mass, which, combined with the ratio of the masses found from the relative sizes of the orbits, can tell you the individual masses of the stars.

Actually, figuring out the mass of a binary star system is not as easy as it might seem from this discussion. The orbits of the two stars may be elliptical; and, although the orbits lie in the same plane, that plane can be tipped at an unknown angle to your line of sight, further distorting the observed shapes of the orbits. Astronomers must find ways to correct for these distortions. In addition, astronomers analyzing binary systems must find the distances to them so they can estimate the true size of the orbits in astronomical units. Finding the masses of binary stars requires a number of steps to get from what you can observe to what you really want to know, the masses. Constructing such sequences of steps is an important part of science (■ How Do We Know? 8-1).

Although there are many different kinds of binary stars, three types are especially useful for determining stellar masses. These are discussed separately in the next sections.

Reasoning with Numbers | 8-4

The Masses of Binary Stars

Johannes Kepler's third law of orbital motion worked only for the planets in our solar system, but when Isaac Newton realized that mass was involved, he made the third law into a general principle. Newton's version of the third law applies to any pair of objects that orbit each other. The total mass of the two objects is related to the average distance a between them and their orbital period P. If the masses are M_A and M_B, then

$$M_A + M_B = \frac{a^3}{P^2}$$

In this formula, a is expressed in AU, P in years, and the mass in solar masses.

Notice that this formula is related to Kepler's third law of planetary motion (see Table 4-1). Almost all of the mass of the solar system is in the sun. If you apply this formula to any planet in our solar system, the total mass is 1 solar mass. Then the formula becomes $P^2 = a^3$, which is Kepler's third law.

In other star systems, the total mass is not necessarily 1 solar mass, and this gives you a way to find the masses of binary stars. If you can find the average distance in AU between the two stars and their orbital period in years, the sum of the masses of the two stars is just a^3/P^2.

Example A: If you observe a binary system with a period of 32 years and an average separation of 16 AU, what is the total mass?

Solution: The total mass equals $16^3/32^2$, which equals 4 solar masses.

Example B: Let's call the two stars in the previous example A and B. Suppose star A is 12 AU away from the center of mass, and star B is 4 AU away. What are the individual masses?

Solution: The ratio of the masses must be 12:4, which is the same as a ratio of 3:1. What two numbers add up to 4 and have the ratio 3:1? Star B must be 3 solar masses, and star A must be 1 solar mass.

How Do We Know? | 8-1

Chains of Inference

How do scientists measure something they can't detect? Sometimes scientists cannot directly observe the things they really want to study, so they must construct chains of inference that connect observable parameters to the unobservable quantities they want to know. You can't observe the mass of stars directly, so you must find a way to use what you can observe, orbital period and angular separation, to figure out step by step the parameters you need to calculate the mass.

Consider another example. Geologists can't measure the temperature and density of Earth's interior directly. There is no way to drill a hole to Earth's center and lower a thermometer or recover a sample. Nevertheless, the speed of vibrations from a distant earthquake depends on the temperature and density of the rock they pass through. Geologists can't measure the speed of the vibrations deep inside Earth; but they can measure the delays in the arrival times at different locations on the surface, and that allows them to work their way back to the speed and, finally, the temperature and density.

Chains of inference can be nonmathematical. Biologists studying the migration of whales can't follow individual whales for years at a time, but they can observe them feeding and mating in different locations; take into consideration food sources, ocean currents, and water temperatures; and construct a chain of inference that leads back to the seasonal migration pattern for the whales.

This chapter contains a number of chains of inference. Almost all sciences use chains of inference. When you can link the observable parameters step by step to the final conclusions, you gain a strong insight into the nature of science.

San Andreas fault: A chain of inference connects earthquakes to conditions inside Earth. (USGS)

Visual Binary Systems

In a **visual binary system,** the two stars are separately visible in the telescope. Only a pair of stars with large orbits can be separated visually; if the orbits are small, the telescope cannot resolve the star images, and you see only a single point of light. In a visual binary system, you can see each star moving around its orbit.

Visual binary systems are common; more than half of all stars are members of binary star systems, and many of those are visual binaries. Favorite Star Polaris has two stellar companions. Few visual binaries, however, can be analyzed completely. Many are so far apart that their periods are much too long for practical mapping of their orbits. Others are so close together that they are hardly visible as separate stars, and it is difficult to map the shape of their orbits. One of the stars that orbits Polaris, for instance, orbits with a period over a thousand years, and the other is so close to Polaris that it is hardly visible even with the Hubble Space Telescope.

Astronomers study visual binary systems by measuring the position of the two stars directly at the telescope or in images. In either case, the astronomers need measurements over many years to map the orbits. The first frame of ■ Figure 8-12 shows a photograph of our Favorite Star Sirius, which is a visual binary system made up of the bright star Sirius A and its white dwarf companion Sirius B. The photo was taken in 1960. Successive frames in Figure 8-12 show the motion of the two stars as observed since 1960 and the orbits the stars follow. The orbital period is 50 years, and astronomers have found accurate masses for both stars.

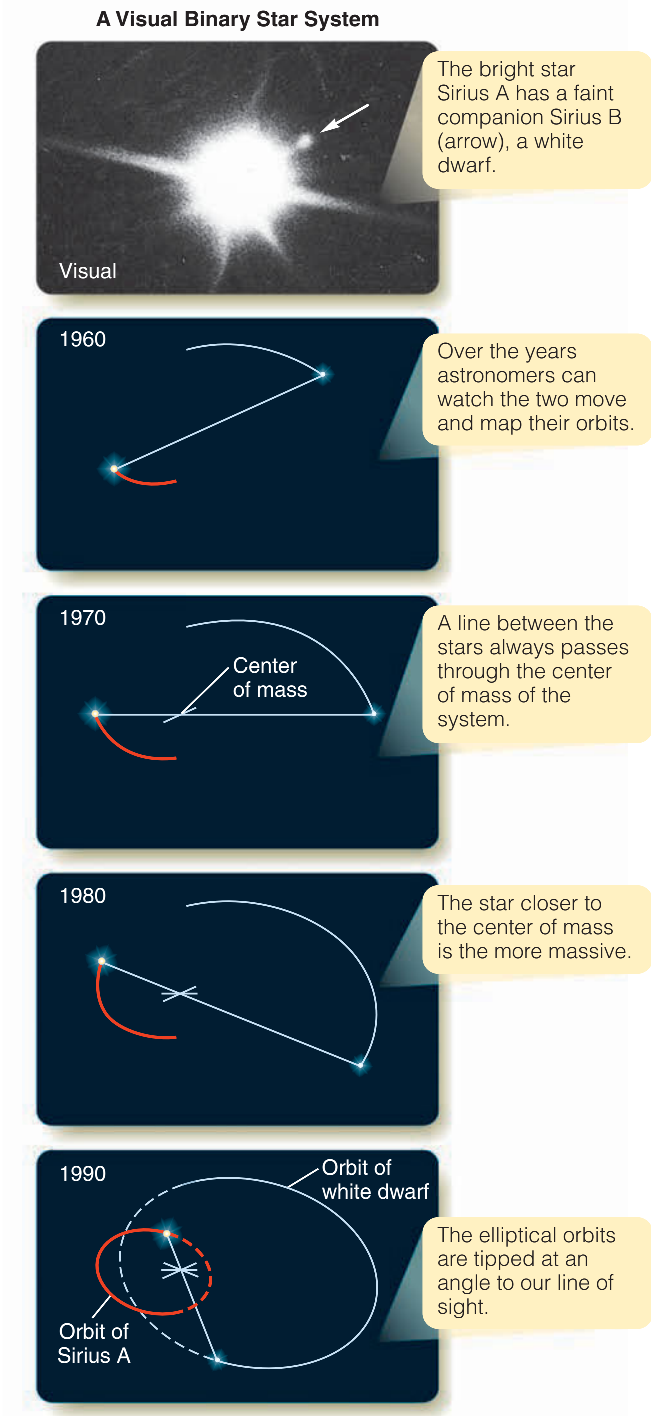
Spectroscopic Binary Systems

If the stars in a binary system are close together, the telescope, limited by diffraction and seeing, reveals only a single point of light. Only by looking at a spectrum, which is formed by light from both stars and contains spectral lines from both, can astronomers tell that there are two stars present and not one. Such a system is a **spectroscopic binary system.**

■ Figure 8-13 shows a sequence of spectra recorded over a few days as the stars in a spectroscopic binary moved around their orbits. You can see how the spectral line in the top spectrum splits into two components that move apart, move together, merge as they cross, and then move apart again. This is the sure sign of a spectroscopic binary.

To understand the spectra in Figure 8-13, look at the diagrams in ■ Figure 8-14, which shows two stars orbiting each other. In the first frame, star A is approaching while star B recedes. In the spectrum, you see a spectral line from star A Doppler shifted toward the blue end of the spectrum while the same spectral line from star B is Doppler shifted toward the red end of the spectrum. As you watch the two stars revolve around their orbits, they alternately approach and recede, and you see small Doppler shifts moving their spectral lines apart and then back together. In a real spectroscopic binary, you can't see the individual stars, but the sight of pairs of spectral lines moving back and forth across each other would alert you that you were observing a spectroscopic binary system.

At first glance, it seems that it should be easy to find the masses of the stars in a spectroscopic binary. You can find the orbital period by waiting to see how long it takes for the spectral



■ **Figure 8-12**

The orbital motion of Sirius A and Sirius B can reveal their individual masses. (Photo: © UC Regents/Lick Observatory)

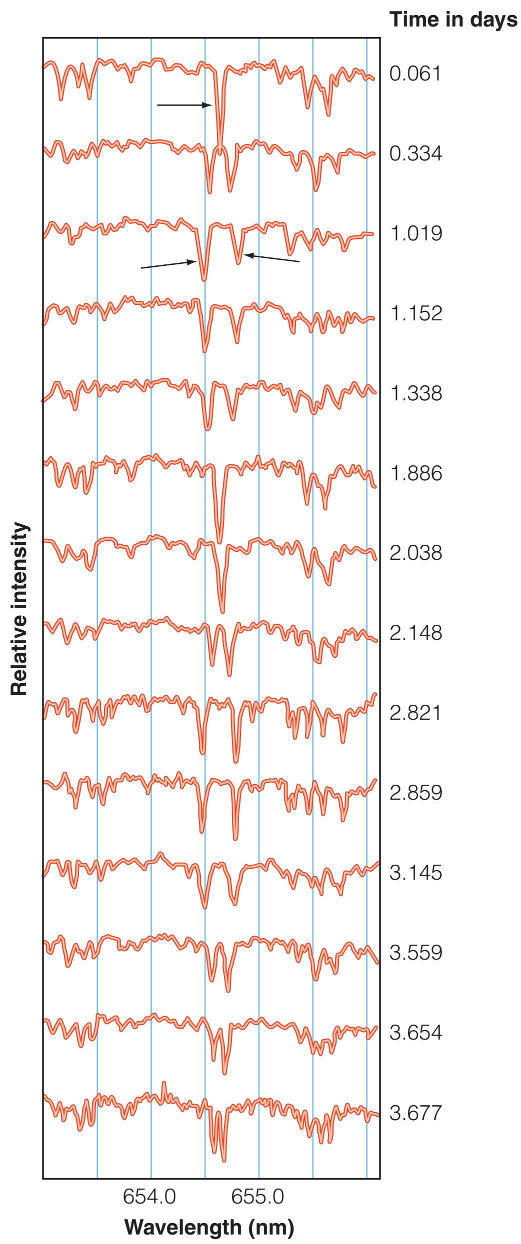


Figure 8-13

Fourteen spectra of the star HD 80715 are shown here as graphs of intensity versus wavelength. A single spectral line (arrow in top spectrum) splits into a pair of spectral lines (arrows in third spectrum), which then merge and split apart again. These changing Doppler shifts reveal that HD 80715 is a spectroscopic binary. (Adapted from data courtesy of Samuel C. Barden and Harold L. Nations)

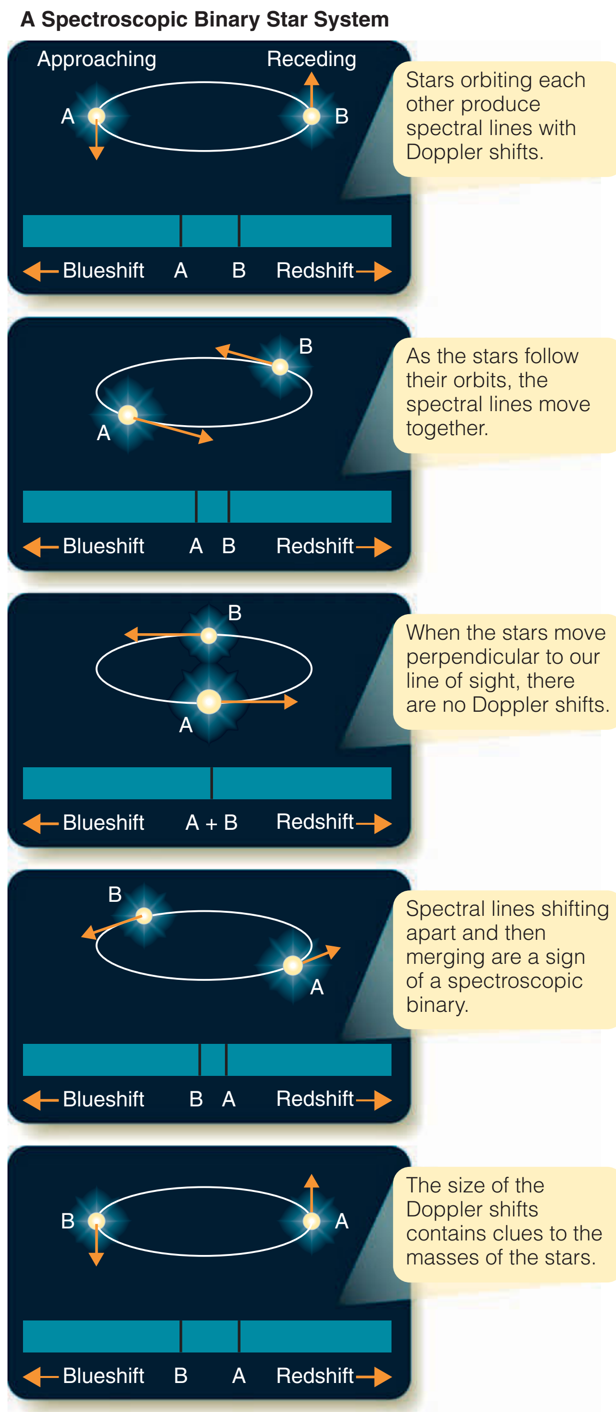


Figure 8-14

From Earth, a spectroscopic binary looks like a single point of light, but the Doppler shifts in its spectrum reveal the orbital motion of the two stars.

lines to return to their starting positions. You can measure the size of the Doppler shifts to find the orbital velocities of the two stars. If you multiply velocity times orbital period, you can find the circumference of the orbit, and from that you can find the radius of the orbit. Now that you know the orbital period and the size of the orbit you should be able to calculate the mass. One important detail is missing, however. You don't know how much the orbits are inclined to your line of sight.

You can find the inclination of a visual binary system because you can see the shape of the orbits. In a spectroscopic binary system, however, you cannot see the individual stars, so you can't find the inclination or untip the orbits. Recall that the Doppler effect only reveals the radial velocity, the part of the velocity directed toward or away from the observer. Because you cannot find the inclination, you cannot correct these radial velocities to find the true orbital velocities. Consequently, you cannot find the true masses. All you can find from a spectroscopic binary system is a lower limit to the masses.

More than half of all stars are in binary systems, and most of those are spectroscopic binary systems. Many of the familiar stars in the sky are actually pairs of stars orbiting each other (■ Figure 8-15).

You might wonder what happens when the orbits of a spectroscopic binary system lie exactly edge-on to Earth. The result is the most informative kind of binary system.

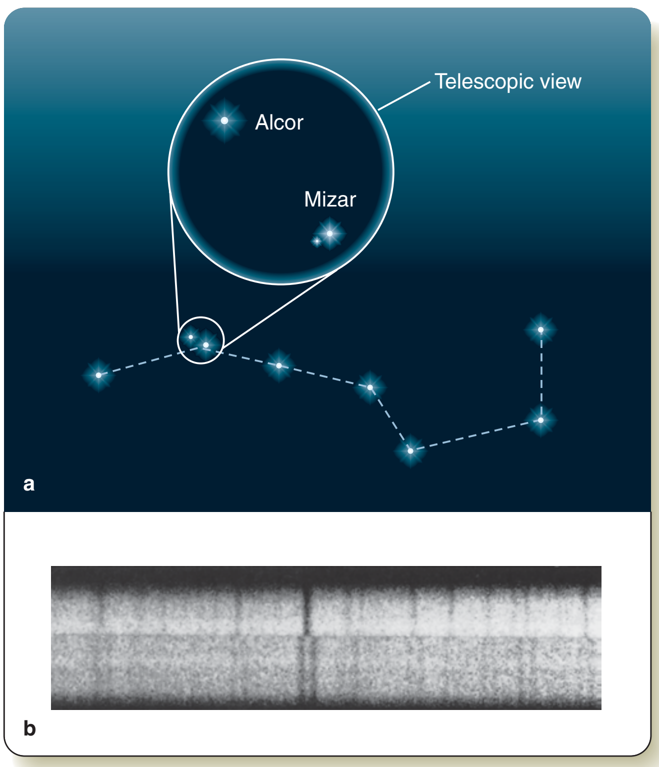
Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Spectroscopic Binaries."

Eclipsing Binary Systems

As mentioned earlier, the orbits of the two stars in a binary system always lie in a single plane. If that plane is nearly edge-on to Earth, then the stars appear to cross in front of each other. Imagine a model of a binary star system in which a cardboard disk represents the orbital plane and balls represent the stars, as in ■ Figure 8-16. If you see the model from the edge, then the balls that represent the stars can move in front of each other as they follow their orbits. The small star crosses in front of the large star, and then, half an orbit later, the large star crosses in front of the small star. When one star moves in front of the other, it blocks some of the light, and the star is eclipsed. Such a system is called an **eclipsing binary system.**

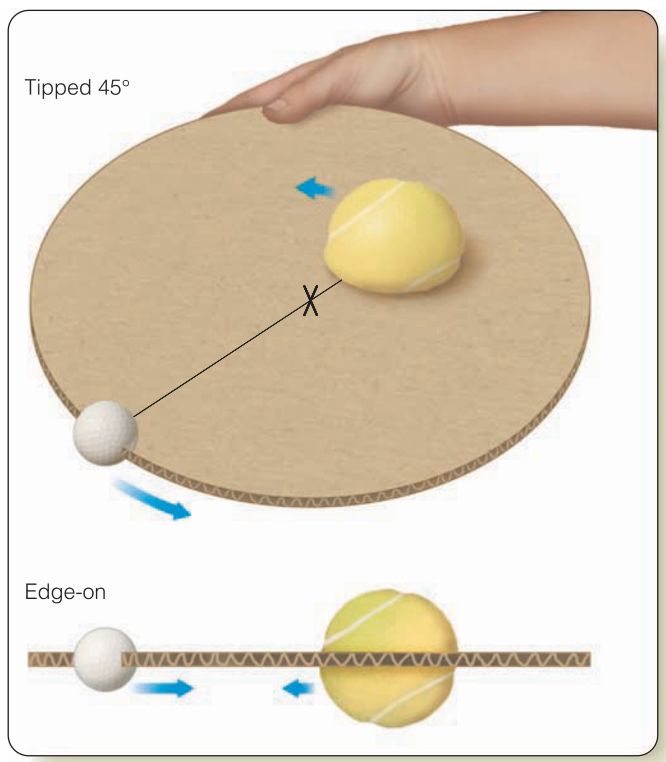
Seen from Earth, the two stars are not visible separately. The system looks like a single point of light. But when one star moves in front of the other star, part of the light is blocked, and the total brightness of the point of light decreases. ■ Figure 8-17 shows a smaller star moving in an orbit around a larger star, first eclipsing the larger star and then being eclipsed itself as it moves behind its companion. The resulting variation in the brightness of the system is shown as a graph of brightness versus time, a **light curve.** Remember, you can't see the individual stars in an eclipsing system. Cover the stars in Figure 8-17 with your fingers and look only at the light curve. If you saw such a light curve, you would immediately recognize the point of light in the sky as an eclipsing binary system.

The light curves of eclipsing binary systems contain tremendous amounts of information, but the curves can be difficult to analyze. Figure 8-17 shows an idealized system. Compare this with ■ Figure 8-18, which shows the light curve of a real system in which the stars have dark spots on their surfaces and are so close to each other that their shapes are distorted.



■ **Figure 8-15**

(a) At the bend of the handle of the Big Dipper lies a pair of stars, Mizar and Alcor. Through a telescope you can discover that Mizar has a fainter companion and so is a member of a visual binary system. (b) Spectra of Mizar recorded at different times show that it is itself a spectroscopic binary system rather than a single star. In fact, both the faint companion to Mizar and the nearby star Alcor are also spectroscopic binary systems. (The Observatories of the Carnegie Institution of Washington)



■ **Figure 8-16**

Imagine a model of a binary system with balls for stars and a disk of cardboard for the plane of the orbits. Only if you view the system edge-on do you see the stars cross in front of each other.

An Eclipsing Binary Star System

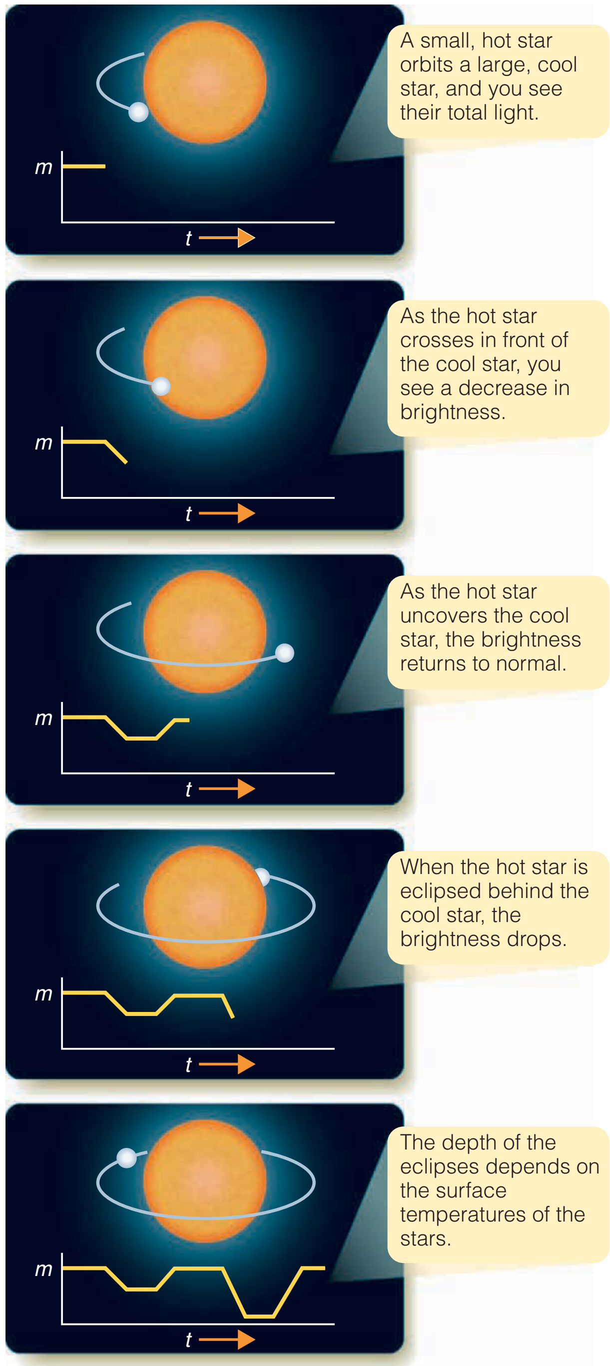


Figure 8-17

From Earth, an eclipsing binary looks like a single point of light, but changes in brightness reveal that two stars are eclipsing each other. Doppler shifts in the spectrum combined with the light curve, shown here as magnitude versus time, can reveal the size and mass of the individual stars.

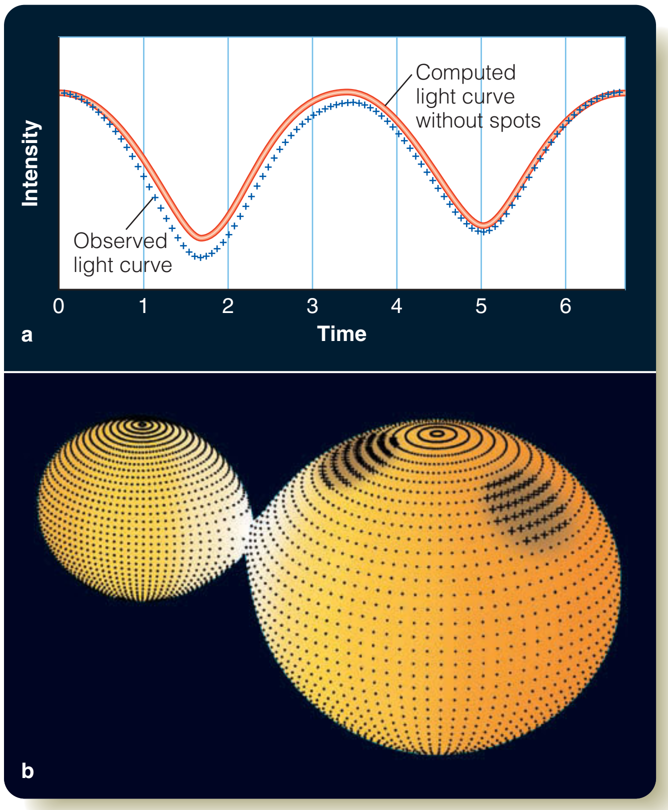


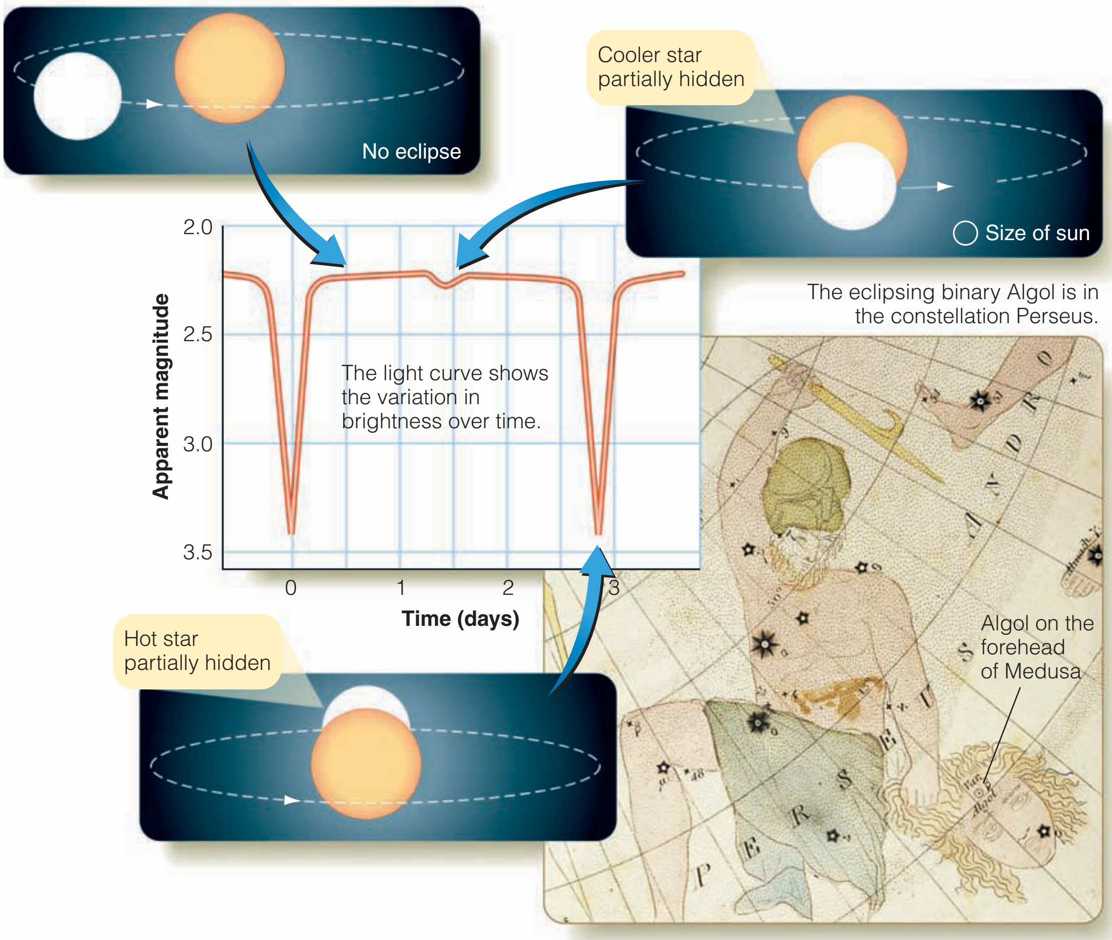
Figure 8-18

The observed light curve of the binary star VW Cephei (lower curve) shows that the two stars are so close together their gravity distorts their shapes. Slight distortions in the light curve reveal the presence of dark spots at specific places on the star's surface. The upper curve shows what the light curve would look like if there were no spots. (Graphics created with Binary Maker 2.0)

Once the light curve of an eclipsing binary system has been accurately observed, you can construct a chain of inference leading to the masses of the two stars. You can find the orbital period easily, and you can get spectra showing the Doppler shifts of the two stars. You can find the orbital velocity because you don't have to untip the orbits; you know they are nearly edge-on, or there would not be eclipses. Then you can find the size of the orbits and the masses of the stars.

Earlier in this chapter you used luminosity and temperature to find the radii of stars, but eclipsing binary systems provide a way to measure the sizes of stars directly. From the light curve you can tell how long it took for the small star to cross the large star. Multiplying this time interval by the orbital velocity of the small star gives the diameter of the larger star. You can also determine the diameter of the small star by noting how long it took to disappear behind the edge of the large star. For example, if it took 300 seconds for the small star to disappear while traveling 500 km/s relative to the large star, then it must be 150,000 km in diameter. Of course, there are complications due to the inclination and eccentricity of orbits, but often these effects can be taken into account.

Algol (β Persei) is one of the best-known eclipsing binaries because its eclipses are visible to the naked eye. Normally, its magnitude is about 2.15, but its brightness drops to 3.4 during eclipses that occur every 68.8 hours. Although the nature of the star was not recognized until 1783, its periodic dimming was



■ **Figure 8-19**

The eclipsing binary Algol consists of a hot B star and a cooler G or K star. The eclipses are partial, meaning that neither star is completely hidden during eclipses. The orbit here is drawn as if the cooler star were stationary.

probably known to the ancients. *Algol* comes from the Arabic for "the demon," and it is associated in constellation mythology with the severed head of Medusa, the sight of whose serpentine locks turned mortals to stone (■ Figure 8-19). Indeed, in some accounts, Algol is the winking eye of the demon.

From the study of binary stars, astronomers have found that the masses of stars range from roughly 0.1 solar mass at the low end to nearly 100 solar masses at the high end. The most massive stars ever found in a binary system have masses of 83 and 82 solar masses. A few other stars are believed to be more massive, 100 solar masses to 150 solar masses, but they do not lie in binary systems, so astronomers must estimate their mass.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Eclipsing Binaries."

◄ SCIENTIFIC ARGUMENT ►

When you look at the light curve for an eclipsing binary system with total eclipses, how can you tell which star is hotter?

Scientists must have good imaginations to visualize objects they cannot see. This scientific argument will exercise your imagination. If you assume that the two stars in an eclipsing binary system are not the same size, then you can refer to them as the larger star and the smaller star. When the smaller star moves behind the larger star, you lose the light coming from the total area of the small star. And when the smaller star moves in front of the larger star, it blocks off light from the same amount of area on the larger star. In both cases, the same amount of area, the same number of square meters, is hidden from your sight. Then the amount of light lost during an eclipse depends only on the temperature of the hidden surface, because temperature is what determines how much a single square meter can radiate per second. When the surface of the hotter star is hidden, the brightness will fall dramatically, but when the surface of the cooler star is hidden, the brightness will not fall as much. So you can look at the light curve and point to the deeper of the two eclipses and say, "That is where the hotter star is behind the cooler star."

Now change the argument to consider the diameters of the stars. **How could you look at the light curve of an eclipsing binary with total eclipses and find the ratio of the diameters?**

◄ ►

8-5 A Survey of the Stars

You have learned how to find the luminosities, diameters, and masses of stars, and now you can put those data (■ How Do We Know? 8-2) together to paint a family portrait of the stars. As in any family portrait, both similarities and differences are important clues to the history of the family. As you begin trying to understand how stars are born and how they die, ask a simple question: What is the average star like? Answering that question is both challenging and illuminating.

Surveying the Stars

If you want to know what the average person thinks about a certain subject, you take a survey. If you want to know what the average star is like, you must survey the stars. Such surveys reveal important relationships among the family of stars.

How Do We Know? 8-2

Basic Scientific Data

Where do large masses of scientific data come from? In a simple sense, science is the process by which scientists look at data and search for relationships, and it sometimes requires large amounts of data. For example, astronomers need to know the masses and luminosities of many stars before they can begin to understand the mass–luminosity relationship.

Compiling basic data is one of the common forms of scientific work—a necessary first step toward scientific analysis and understanding. An archeologist may spend months or even years diving to the floor of the Mediterranean Sea to study an ancient Greek shipwreck. She will carefully measure the position of every wooden timber and bronze fitting. She will photograph and recover everything from broken pottery to tools and weapons. The care with which she records data on the site pays off when she begins her analysis. For every hour the archaeologist spends recovering an object, she may spend days or weeks in her office, a library, or a museum identifying and understanding the object. Why was there a Phoenician hammer on a Greek ship? What does that reveal about the economy of ancient Greece?

Finding, identifying and understanding that ancient hammer contributes only a small bit of information, but the work of many scientists eventually builds a picture of how ancient Greeks saw their world. Solving a single binary star system to find the masses of the stars does not tell an astronomer a great deal about nature. Over the years, however, many astronomers have added their results to the growing data file on stellar masses. Scientific data accumulates and can be analyzed by later generations of scientists.

Collecting mineral samples can be hard work, but it is also fun. Scientists sometimes collect large amounts of data because they enjoy the process. (M. A. Seeds)

Not many decades ago, surveying large numbers of stars was an exhausting task, but modern computers have changed that. Specially designed telescopes controlled by computers can make millions of observations per night, and high-speed computers can compile and analyze these vast surveys and create easy-to-use databases. Those surveys produce mountains of data that astronomers can mine while searching for relationships within the family of stars.

What could you learn about stars from a survey of the stars near the sun? Because the sun is thought to be in a typical place in the universe, such a survey could reveal general characteristics of the stars and might reveal unexpected processes in the formation and evolution of stars. Study ■ **The Family of Stars** on pages 154–155 and notice three important points:

1. Taking a survey is difficult because you must be sure you get an honest sample. If you don't survey enough stars or if you don't notice some kinds of stars, you can get biased results.
2. Most stars are faint, and luminous stars are rare. The most common kinds of stars are the lower-main-sequence red dwarfs and the white dwarfs.
3. A survey reveals that what you see in the sky is deceptive. Stars near the sun are quite faint; but luminous stars, although they are rare, are easily visible even at great distances. Many of the brighter stars in the sky are highly luminous stars that you see even though they lie far away.

The night sky is a beautiful carpet of stars, but they are not all the same. Some are giants and supergiants, and some are dwarfs. The family of the stars is rich in its diversity.

Mass, Luminosity, and Density

If you survey enough stars and plot the data in an H–R diagram, you can see the patterns that hint at how stars are born, how they age, and how they die.

If you label an H–R diagram with the masses of the plotted stars, as in ■ Figure 8-20, you will discover that the main-sequence stars are ordered by mass. The most massive main-sequence stars are the hot stars. As you run your eye down the main sequence, you will find successively lower-mass stars, until you reach the lowest-mass, coolest, faintest main-sequence stars.

Stars that do not lie on the main sequence are not in order according to mass. Giant stars are a jumble of different masses, and supergiants, although they tend to be more massive than giants, are in no particular order in the H–R diagram. In contrast, all white dwarfs have about the same mass, somewhere in the narrow range of from 0.5 to about 1 solar mass.

Because of the systematic ordering of mass along the main sequence, the main-sequence stars follow a **mass–luminosity relation**—the more massive a star is, the more luminous it is (■ Figure 8-21). In fact, the mass–luminosity relation can be expressed as a simple formula (see ■ Reasoning with Numbers 8-5). Giants and supergiants do not follow the mass–luminosity relation very closely, and white dwarfs do not at all. In the next two chapters, the mass–luminosity relation will help you understand how stars are born, live, and die.

The density of stars reveals another pattern in the H–R diagram. The average density of a star is its mass divided by its volume. Stars are not uniform in density but are most dense at their centers and least dense near their surface. The center of the sun, for instance, is about 150 times as dense as water; its density near the visible surface is about 3400 times less dense than Earth's atmosphere at sea level. A star's average density is intermediate between its central and surface densities. The sun's average density is approximately 1 g/cm^3—about the density of water.

Main-sequence stars have average densities similar to the sun's, but giant stars, being large, have low average densities, ranging from 0.1 to 0.01 g/cm^3. The enormous supergiants have still lower densities, ranging from 0.001 to 0.000001 g/cm^3. These densities are thinner than the air you breathe, and if you could insulate yourself from the heat, you could fly an airplane through these stars. Only near the center would you be in any danger, for there the material is very dense—roughly three million times the density of water.

The white dwarfs have masses about equal to the sun's but are very small, only about the size of Earth. That means the matter is compressed to enormous densities. On Earth, a teaspoonful of this material would weigh about 15 tons.

Reasoning with Numbers | 8-5

The Mass–Luminosity Relation

You can calculate the approximate luminosity of a star using a simple equation. A star's luminosity in terms of the sun's luminosity equals its mass in solar masses raised to the 3.5 power:

$$L = M^{3.5}$$

This is the mathematical form of the mass–luminosity relation. It is only an approximation, as shown by the red line in Figure 8-21, but it applies to most stars over a wide range of stellar masses.

You can do simple calculations with this equation if you remember that raising a number to the 3.5 power is the same as cubing it and then multiplying by its square root.

Example: What is the luminosity of a star four times the mass of the sun?

Solution: The star must be 128 times more luminous than the sun because

$$L = M^{3.5} = 4^{3.5} = 4 \times 4 \times 4\sqrt{4} = 64 \times 2 = 128$$

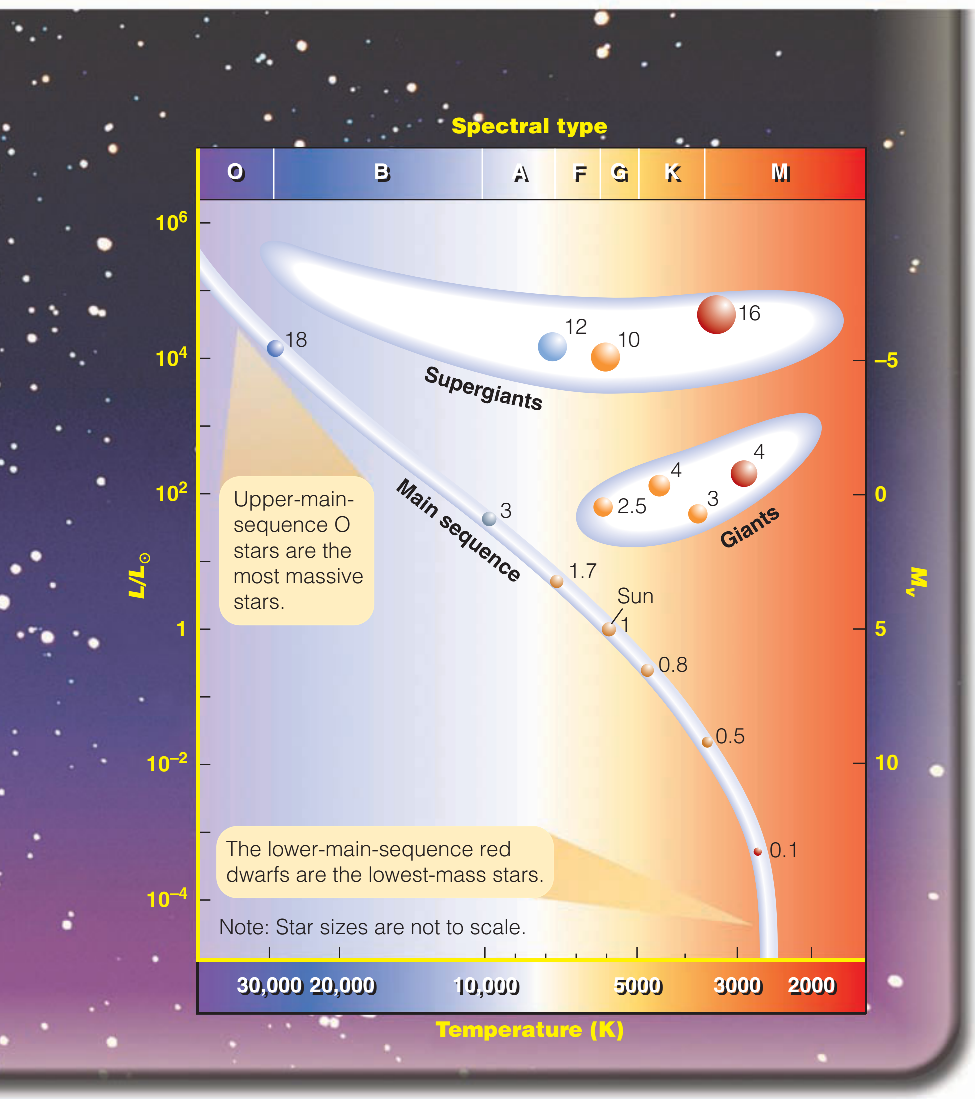


Figure 8-20

The masses of the plotted stars are labeled on this H–R diagram. Notice that the masses of main-sequence stars decrease from top to bottom but that masses of giants and supergiants are not arranged in any ordered pattern.

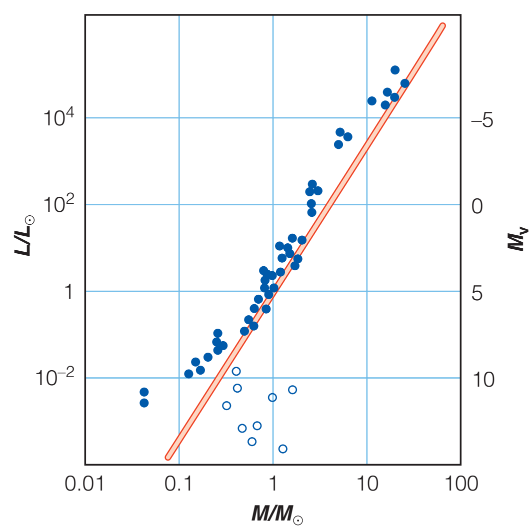


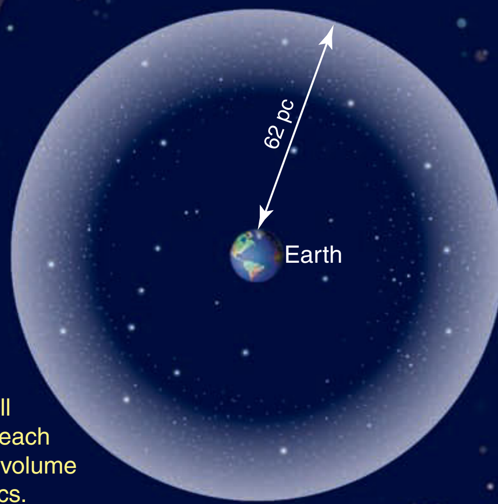
Figure 8-21

The mass–luminosity relation shows that the more massive a main-sequence star is, the more luminous it is. The open circles represent white dwarfs, which do not obey the relation. The red line represents the equation in Reasoning with Numbers 8-5.

The Family of Stars

1 What is the most common kind of star? Are some rare? Are some common? To answer those questions you must survey the stars. To do so you must know their spectral class, their luminosity class, and their distance. Your census of the family of stars produces some surprising demographic results.

1a You could survey the stars by observing every star within 62 pc of Earth. A sphere 62 pc in radius encloses a million cubic parsecs. Such a survey would tell you how many stars of each type are found within a volume of a million cubic parsecs.



2 Your survey faces two problems.

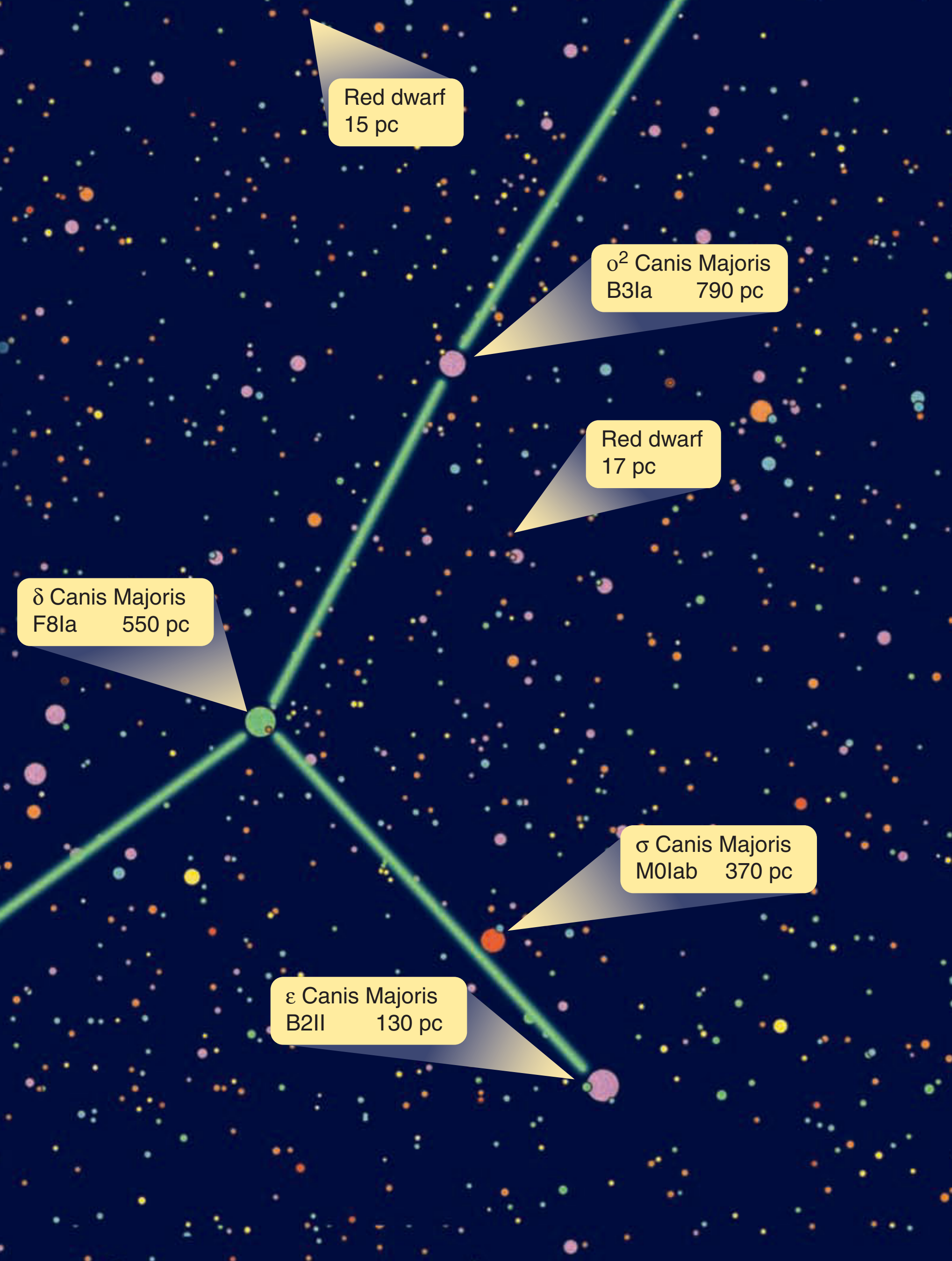
1. The most luminous stars are so rare you find few in your survey region. There are no O stars at all within 62 pc of Earth.

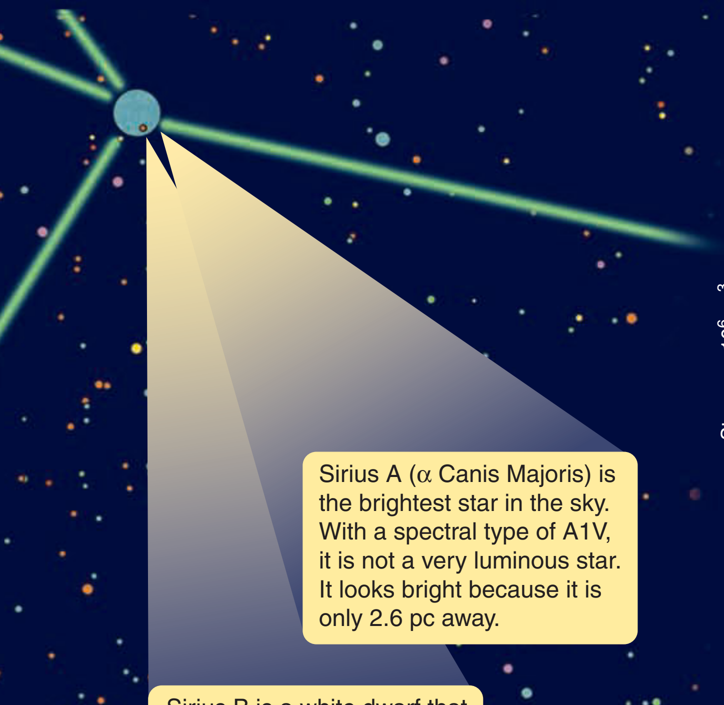
2. Lower-main-sequence M stars, called red dwarfs, and white dwarfs are so faint they are hard to locate even when they are only a few parsecs from Earth. Finding every one of these stars in your survey sphere is a difficult task.

Spectral Class Color Key

- O and B
- A
- F
- G
- K
- M

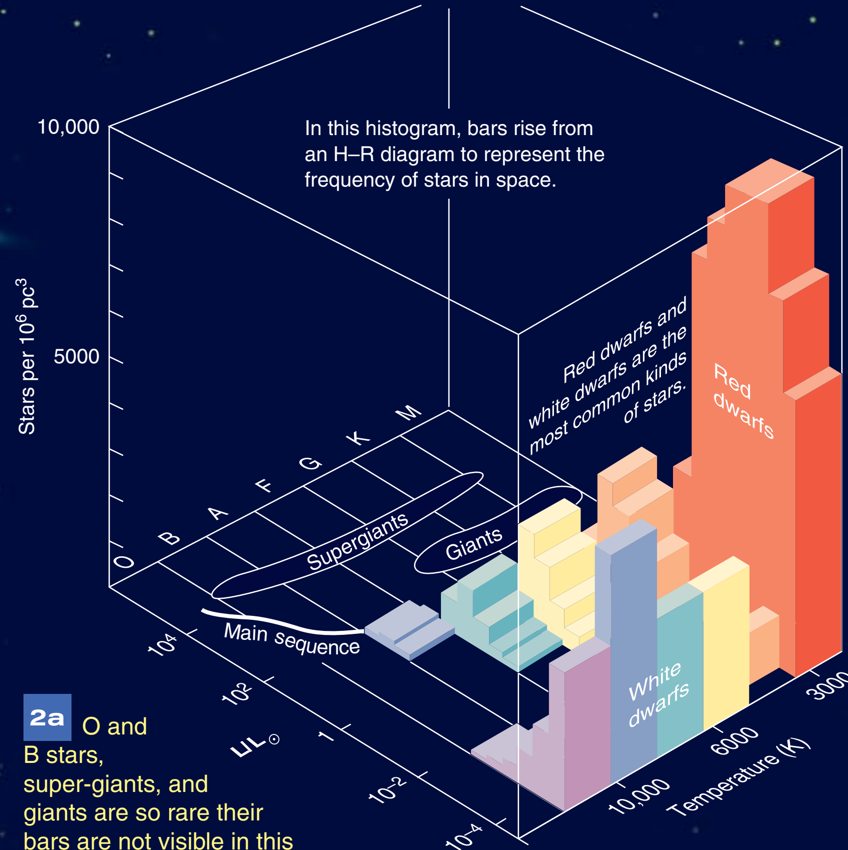
The star chart in the background of these two pages shows most of the constellation Canis Major; stars are represented as dots with colors assigned according to spectral class. The brightest stars in the sky tend to be the rare, highly luminous stars, which look bright even though they are far away. Most stars are of very low luminosity, so nearby stars tend to be very faint red dwarfs.





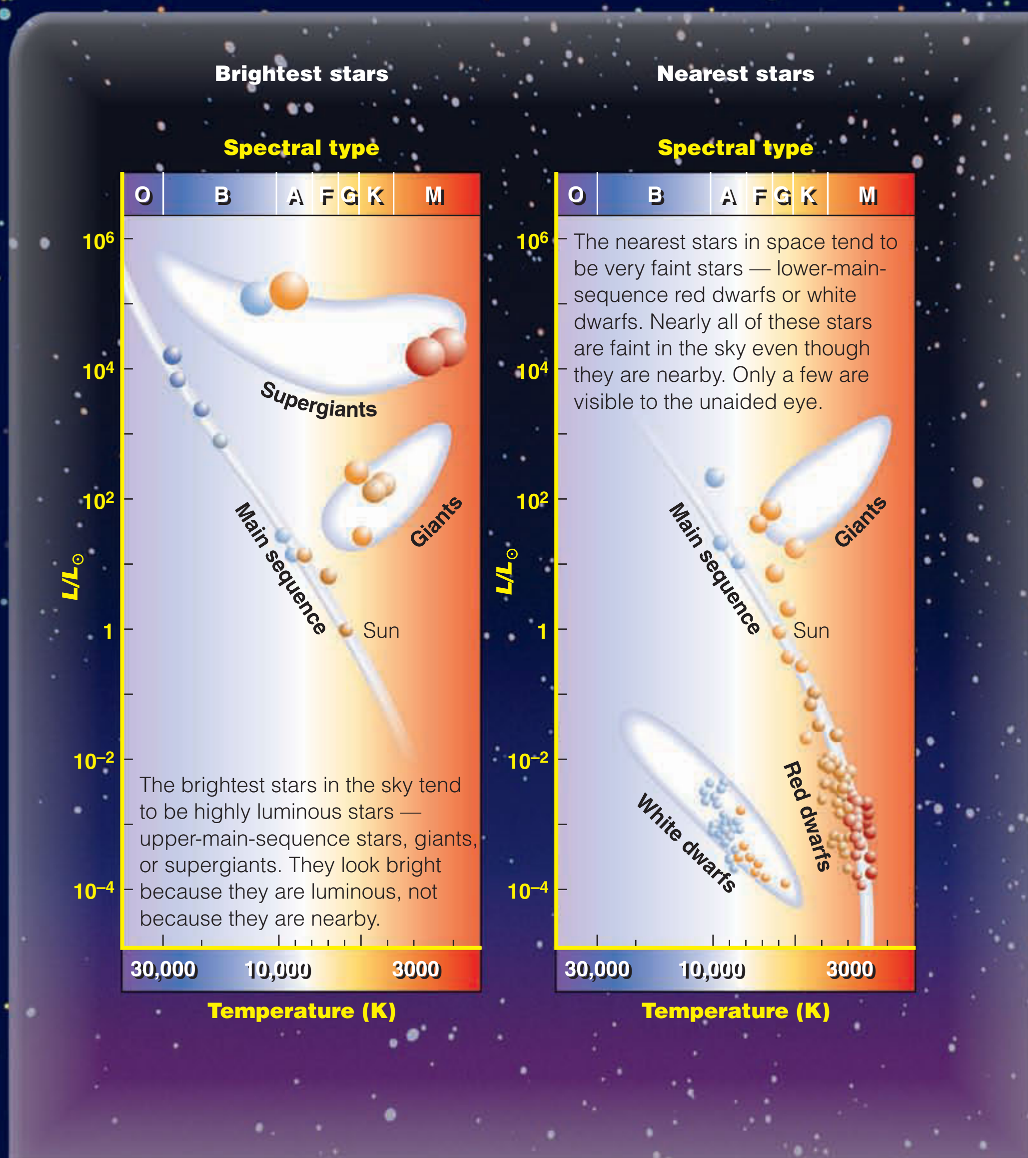
Sirius A (α Canis Majoris) is the brightest star in the sky. With a spectral type of A1V, it is not a very luminous star. It looks bright because it is only 2.6 pc away.

Sirius B is a white dwarf that orbits Sirius A. Although Sirius B is not very far away, it is much too faint to see with the unaided eye.



2a O and B stars, super-giants, and giants are so rare their bars are not visible in this graph.

3 Luminous stars are rare but are easy to see. Most stars are very low luminosity objects. Not a single white dwarf or red dwarf is bright enough to see with the unaided eye. See H–R diagrams at right.



Density divides stars into three groups. Most stars are main-sequence stars with densities like the sun's. Giants and supergiants are very-low-density stars, and white dwarfs are high-density stars. You will see in later chapters that these densities reflect different stages in the evolution of stars.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Mass–Luminosity Relation" and "H–R/Mass–Luminosity 3D Graph."

SCIENTIFIC ARGUMENT

What kind of stars do you see if you look at a few of the brightest stars in the sky?

This argument shows how careful you must be to interpret simple observations. When you look at the night sky, the brightest stars are mostly giants and supergiants. Most of the bright stars in Canis Major, for instance, are supergiants. Sirius, one of our Favorite Stars, is the brightest star in the sky, but it is just a main-sequence star; it looks bright because it is nearby, not because it is very luminous. In general, the supergiants and giants are so luminous that they stand out and look bright, even though they are not nearby. When you look at a bright star in the sky, you are probably looking at a highly luminous star—a supergiant or a giant. You can check the argument above by consulting the tables of the brightest and nearest stars in the Appendix.

Now revise your argument. **What kind of star do you see if you look at a few of the stars nearest to the sun?**

◄ ►

What Are We? Medium

We humans are medium creatures, and we experience medium things. You can see trees and flowers and small insects, but you cannot see the beauty of the microscopic world without ingenious instruments and special methods. Similarly, you can sense the grandeur of a mountain range, but larger objects, such as stars, are too big for our medium senses. You must use your ingenuity and imagination to experience the truth of such large objects. That is what science does for us. We live between the microscopic world and the astronomical world, and science enriches our lives by revealing the parts of the universe beyond our daily experience.

Experience is fun, but it is very limited. You may enjoy experiencing a flower by admiring its color and shape and by smelling its fragrance. But the flower is more wonderful than your experience can reveal. To truly appreciate the flower you need to understand it, to understand how it serves its plant and how the plant came to create such a beautiful blossom.

Humans have a natural drive to understand as well as experience. You have experienced the stars in the night sky, and now you are beginning to understand them as objects ranging from hot O stars to cool red dwarfs. It is natural for you to wonder why these stars are so different. As you explore that story in the following chapters, you will discover that although you have medium senses, you can understand the stars.

Study and Review

Summary

- Distance is critical in astronomy. Your goal in this chapter was to characterize the stars by finding their luminosities, diameters, and masses. Before you could begin, you needed to find their distances. Only by first knowing the distance to a star could you find its other properties.
- Astronomers can measure the distance to nearer stars by observing their **stellar parallaxes (p. 136).** The most distant stars are so far away that their parallaxes are unmeasurably small. Space telescopes above Earth's atmosphere have measured the parallaxes of huge number of stars.
- Stellar distances are commonly expressed in **parsecs (p. 136).** One parsec is 206,265 AU—the distance to an imaginary star whose parallax is 1 second of arc. One parsec equals 3.26 light-years.
- The amount of light received from a star, the light **flux (p. 137),** is related to its distance by the inverse square law. Once you know the distance to a star, you can find its intrinsic brightness expressed as its **absolute visual magnitude (p. 138)**—the apparent magnitude the star would have if it were 10 pc away.
- The **luminosity *(L)* (p. 138)** of a star, found from its absolute magnitude, is a measure of the total energy radiated by the star in one second. Luminosity is often expressed in terms of the luminosity of the sun.
- The **Hertzsprung–Russell (H–R) diagram (p. 140)** is a plot of luminosity versus surface temperature. It is an important graph in astronomy because it sorts the stars into categories by size.
- Roughly 90 percent of normal stars, including the sun, fall on the **main sequence (p. 141),** with the more massive stars being hotter and more luminous.
- The **giants (p. 141)** and **supergiants (p. 141),** however, are much larger and lie above the main sequence. **Red dwarfs (p. 141)** lie at the bottom end of the main sequence. Some of the **white dwarfs (p. 141)** are very hot stars, but they fall below the main sequence because they are so small.
- The large size of the giants and supergiants means their atmospheres have low densities and their spectra have sharper spectral lines than the spectra of main-sequence stars. In fact, it is possible to assign stars to **luminosity classes (p. 143)** by the widths of their spectral lines. Class V stars are main-sequence stars. Giant stars, class III, have sharper lines, and supergiants, class I, have extremely sharp spectral lines.
- Astronomers can use the locations of the luminosity classes in the H–R diagram to estimate the distances to stars in a technique called **spectroscopic parallax (p. 144).**
- The only direct way you can find the mass of a star is by studying **binary stars (p. 145).** When two stars orbit a common center of mass, astronomers find their masses by observing the period and sizes of their orbits. In a **visual binary (p. 147),** both stars are visible and the orbits can be mapped, but in a **spectroscopic binary system (p. 147)** the stars are so close together they look like a single point of light, and the orbits can't be observed directly.

- In an **eclipsing binary system (p. 149),** the orbits are edge on and the stars cross in front of each other. The resulting brightness changes in the **light curve (p. 149)** can reveal the diameters of the stars as well as their masses.
- A survey in the neighborhood of the sun shows that lower-main-sequence stars are the most common type. Giants and supergiants are rare, but white dwarfs are quite common, although they are faint and hard to find.
- The **mass–luminosity relation (p. 152)** says that the more massive a star is, the more luminous it is. Main-sequence stars follow this rule closely, the most massive being the upper-main-sequence stars and the least massive the lower-main-sequence stars. Giants and supergiants do not follow the relation precisely, and white dwarfs not at all.
- Given the mass and diameter of a star, you can find its average density. On the main sequence, the stars are about as dense as the sun, but the giants and supergiants are very-low-density stars. Some are much thinner than air. The white dwarfs, lying below the main sequence, are tremendously dense.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why are Earth-based parallax measurements limited to the nearest stars?
2. Why was the Hipparcos satellite able to make more accurate parallax measurements than ground-based telescopes?
3. What do the words *absolute* and *visual* mean in the definition of absolute visual magnitude?
4. What does luminosity measure that is different from what absolute visual magnitude measures?
5. Why does the luminosity of a star depend on both its radius and its temperature?
6. How can you be sure that giant stars really are larger than main-sequence stars?
7. What evidence shows that white dwarfs must be very small?
8. What observations would you make to classify a star according to its luminosity? Why does that method work?
9. Why does the orbital period of a binary star depend on its mass?
10. What observations would you make to study an eclipsing binary star?
11. Why don't astronomers know the inclination of a spectroscopic binary? How do they know the inclination of an eclipsing binary?
12. How do the masses of stars along the main sequence illustrate the mass–luminosity relation?
13. Why is it difficult to find out how common the most luminous stars are? The least luminous stars?
14. What is the most common kind of star?
15. If you look only at the brightest stars in the night sky, what kind of star are you likely to be observing? Why?
16. **How Do We Know?** What is the missing link in the chain of inference leading from observations of spectroscopic binaries to the masses of the stars?
17. **How Do We Know?** In what way is basic scientific data cumulative, and how do accumulations of such data help later scientists?

Discussion Questions

1. If someone asked you to compile a list of the nearest stars to the sun based on your own observations, what measurements would you make, and how would you analyze them to detect nearby stars?
2. The sun is sometimes described as an average star. Is that true? What is the average star really like?

Problems

1. If a star has a parallax of 0.050 second of arc, what is its distance in pc? In ly? In AU?
2. If you place a screen of area 1 m^2 at a distance of 2.8 m from a 100-watt lightbulb, the light flux falling on the screen will be 1 J/s. To what distance must you move the screen to make the flux striking it equal 0.01 J/s? (This assumes the lightbulb emits all of its energy as light.)
3. If a star has a parallax of 0.016 second of arc and an apparent magnitude of 6, how far away is it, and what is its absolute magnitude?
4. Complete the following table:

m	M_v	*d* (pc)	*p* (seconds of arc)
____	7	10	____
11	____	1000	____
____	−2	____	0.025
4	____	____	0.040

5. The unaided human eye can see stars no fainter than those with an apparent magnitude of 6. If you can see a bright firefly blinking up to 0.5 km away, what is the absolute visual magnitude of the firefly? (*Hint:* Convert the distance to parsecs and use the formula in Reasoning with Numbers 8-2.)
6. If a main-sequence star has a luminosity of 100 $L_\odot$, what is its spectral type? (*Hint:* See Figure 8-7.)
7. If a star is ten times the radius of the sun and half as hot, what will its luminosity be? (*Hint:* See Reasoning with Numbers 8-3.)
8. A B0 V star has an apparent magnitude of 11. Use the method of spectroscopic parallax to estimate the distance to the star. Why might this distance be inaccurate?
9. Find the luminosity and spectral type of a 5-$M_\odot$ main-sequence star.
10. In the following table, which star is brightest in apparent magnitude? Most luminous? Largest? Least dense? Farthest away?

Star	Spectral Type	*m*
a	G2 V	5
b	B1 V	8
c	G2 Ib	10
d	M5 III	19
e	White dwarf	15

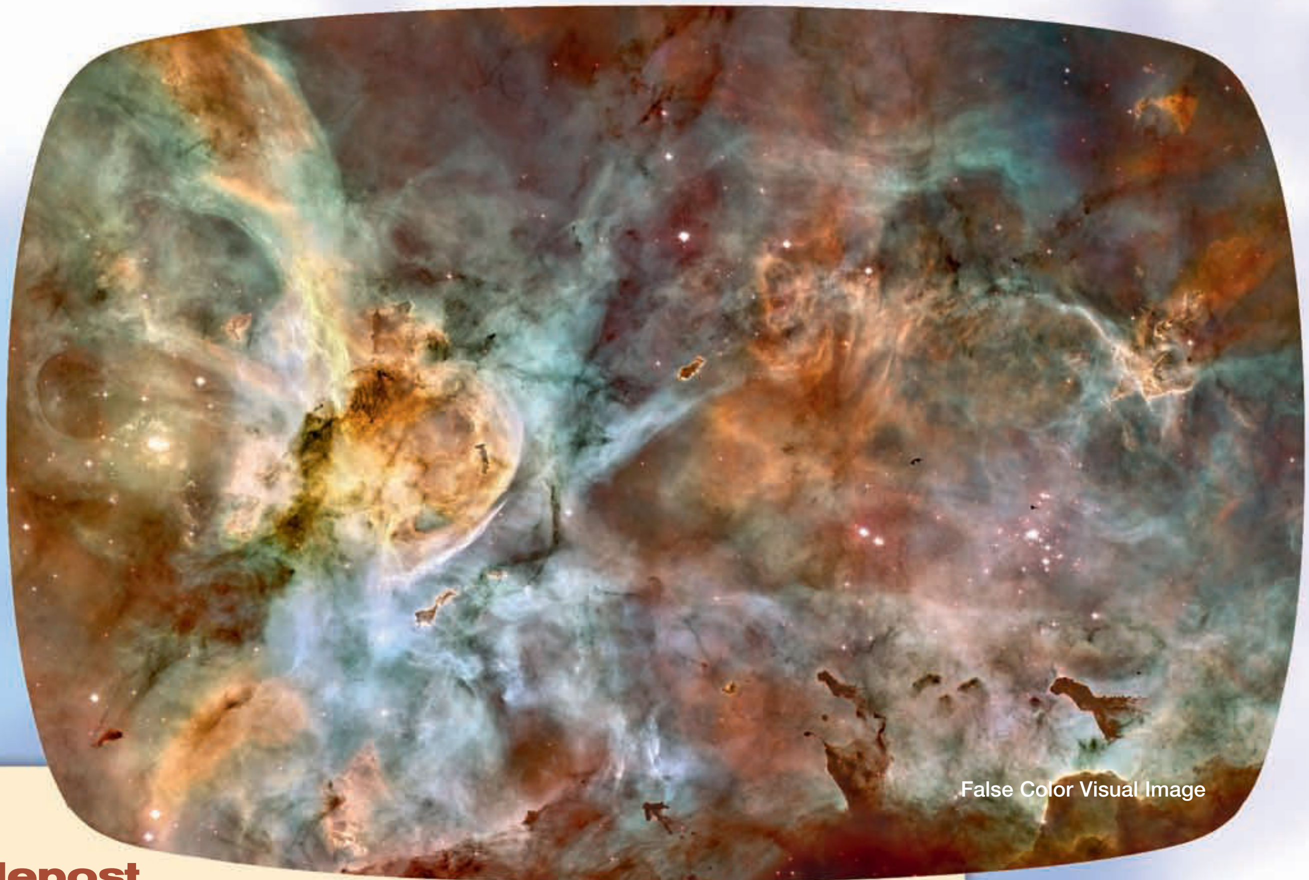
11. If two stars orbit each other with a period of 6 years and a separation of 4 AU, what is their total mass? (*Hint:* See Reasoning with Numbers 8-4.)
12. If the eclipsing binary in Figure 8-17 has a period of 32 days, an orbital velocity of 153 km/s, and an orbit that is nearly edge-on, what is the circumference of the orbit? The radius of the orbit? The mass of the system?
13. If the orbital velocity of the eclipsing binary in Figure 8-17 is 153 km/s and the smaller star becomes completely eclipsed in 2.5 hours, what is its diameter?
14. What is the luminosity of a main-sequence 4-solar-mass star? Of a 9-solar-mass star? Of a 7-solar-mass star?

Learning to Look

1. Look at Figure 8-4. Why is the lava nearest the source brighter and yellower than the lava that is farther away?
2. If all of the stars in the photo here are members of the same star cluster, then they all have about the same distance. Then why are three of the brightest much redder than the rest? What kind of star are they?


NASA

9 The Formation and Structure of Stars



The Carina Nebula is over 50 ly in diameter, and this image of its center, assembled from images recorded at the wavelengths emitted by hydrogen and oxygen, reveals turbulent motions and rapid star formation. (NASA/ESA, N. Smith, Univ. of California, Berkeley, and The Hubble Heritage Team, STScI/AURA)

Guidepost

In the last chapter you discovered some amazing differences among members of the family of stars. This is where you will really begin to see how the universe works. Here you will begin putting together observations and theories to understand how nature makes stars. That will answer four essential questions about stars:

How are stars born?

How do stars make energy?

How do stars maintain their stability?

How long can a star survive?

The most important question you might ask is the key question in science: What's the evidence? Astronomers understand how stars are born and how they make their energy and remain stable because of evidence. Testing theories against evidence is the basic skill required of all scientists, and you will use it over and over in the 11 chapters that follow.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Jim he allowed [the stars] was made, but I allowed they happened. Jim said the moon could'a laid them; well, that looked kind of reasonable, so I didn't say nothing against it, because I've seen a frog lay most as many, so of course it could be done.

MARK TWAIN
THE ADVENTURES OF HUCKLEBERRY FINN

THE STARS ARE not eternal. The stars you see tonight are the same stars your parents, grandparents, and great-grandparents saw. Stars change hardly at all in a human lifetime, but they have a life cycle of their own. Stars are born, and stars die. This chapter begins that story.

In this chapter, you will see how gravity creates stars from the thin gas of space and how nuclear reactions inside stars generate energy. You will see how the flow of that energy outward toward the surface of the star balances gravity and makes the stars stable. In the next chapter, you will complete the story of stars as they exhaust their fuels and die.

How can astronomers know what stars are like when they can't see inside them and don't live long enough to see them evolve? The answer lies in the methods of science. By constructing theories that describe how nature works and then testing those theories against evidence from observations, scientists can unravel some of nature's greatest secrets. They can even understand the origin and structure of the stars.

9-1 The Birth of Stars

THE KEY TO understanding star formation is the correlation between young stars and clouds of gas and dust. Where you find the youngest groups of stars, you also find large clouds of gas and dust. This should lead you to suspect that stars form from such clouds, just as raindrops condense from the water vapor in a thundercloud. But how can these cold clouds contract, heat up, and become stars?

The Interstellar Medium

It is a **Common Misconception** to imagine that space is empty—a vacuum. In fact, the space between the stars is not empty but is filled with low-density gas and dust called the **interstellar medium.**

About 75 percent of the mass of the gas in the interstellar medium is hydrogen, and 25 percent is helium; there are also traces of carbon, nitrogen, oxygen, calcium, sodium, and heavier atoms. Roughly 1 percent of the mass is made up of microscopic dust particles called **interstellar dust.** The dust particles are tiny, about the size of the particles in cigarette smoke, and observations show that they are made mostly of carbon and silicates (rocklike minerals) mixed with or coated with frozen water. The average distance between dust grains is about 10–100 meters.

This interstellar gas and dust is not uniformly distributed through space; it consists of a complex tangle of cool, dense clouds pushed and twisted by currents of hot, low-density gas. Although the cool clouds contain only 10 to 1000 atoms/cm^3 (fewer than any vacuum in a lab on Earth), astronomers refer to them as dense clouds in contrast with the hot, low-density gas that fills the spaces between clouds. That thin gas has a density of only about 0.1 atom/cm^3, which is the same as 1 atom in every 10 cubic centimeters.

The preceding two paragraphs describe the interstellar medium, but you should not accept these facts blindly. Science is based on evidence, so you should demand to know what observational evidence supports these facts. How do astronomers know there is an interstellar medium, and how do they know its properties?

In some cases, the interstellar medium is easily visible as clouds of gas and dust as in the case of the Great Nebula in Orion, an object you can see with your unaided eye (Figure 2-4). Astronomers call such a cloud a **nebula** from the Latin word for cloud. Such nebulae (plural) are clear evidence of an interstellar medium. Study ■ **Three Kinds of Nebulae** on pages 160–161 and notice three important points and four new terms:

1. *Emission nebulae* are produced where very hot stars excite clouds of low-density gas to emit light. The clouds are mostly hydrogen gas ionized by the light from the hot stars, so the nebulae are sometimes called *HII regions.*
2. Where slightly cooler stars illuminate gas clouds containing dust, you see *reflection nebulae,* which provide evidence that the dust in the clouds is made up of very small particles.
3. *Dark nebulae* are produced where dense clouds of gas and dust are silhouetted against background regions filled with stars or bright nebulae.

If a cloud is not too dense, starlight may be able to penetrate it. Stars can be seen through these clouds; but the stars look dimmer because the dust in the clouds scatters some of the light. Because shorter wavelengths are scattered more easily than longer wavelengths, the redder photons are more likely to make it through, so the stars look slightly redder than they should—an effect called **interstellar reddening** (■ Figure 9-1). (This is the same process that makes the setting sun look redder.) Distant stars are dimmed and reddened by intervening gas and dust, clear evidence of an interstellar medium. At near-infrared wavelengths, stars are more easily seen through the dusty interstellar medium because those longer wavelengths are scattered less often.

The thin gas and dust fills the spaces between the denser nebulae. You can see evidence of that in the spectra of distant stars

Three Kinds of Nebulae

1 **Emission nebulae** are produced when a hot star excites the gas near it to produce an emission spectrum. The star must be hotter than about B1 (25,000 K). Cooler stars do not emit enough ultraviolet radiation to ionize the gas. Emission nebulae have a distinctive pink color produced by the blending of the red, blue, and violet Balmer lines. Emission nebulae are also called **HII regions,** following the custom of naming gas with a roman numeral to show its state of ionization. HI is neutral hydrogen, and HII is ionized.

In an HII region, the ionized nuclei and free electrons are mixed. When a nucleus captures an electron, the electron falls down through the atomic energy levels, emitting photons at specific wavelengths. Spectra indicate that the nebulae have compositions much like that of the sun – mostly hydrogen. Emission nebulae have densities of 100 to 1000 atoms per cubic centimeter, better than the best vacuums produced in laboratories on Earth.

Visual-wavelength image

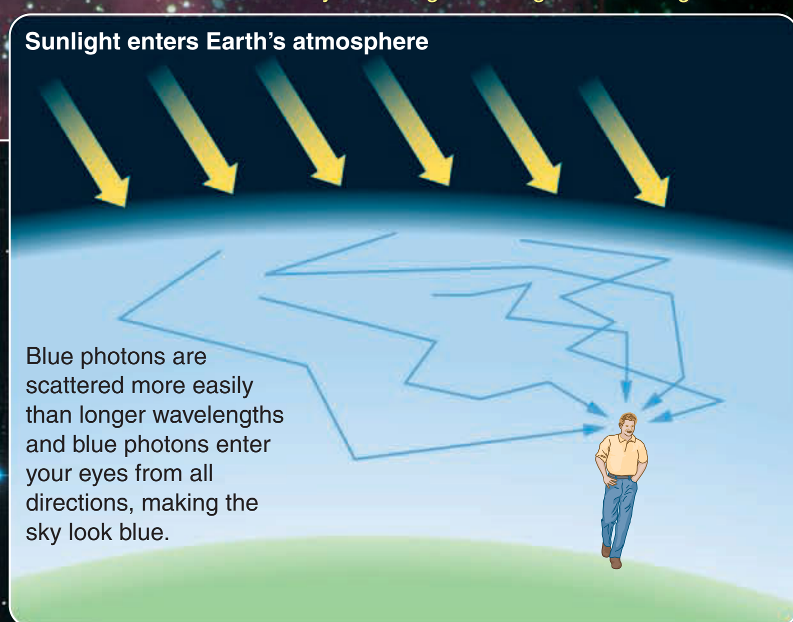
European Southern Observatory

CENGAGENOW™

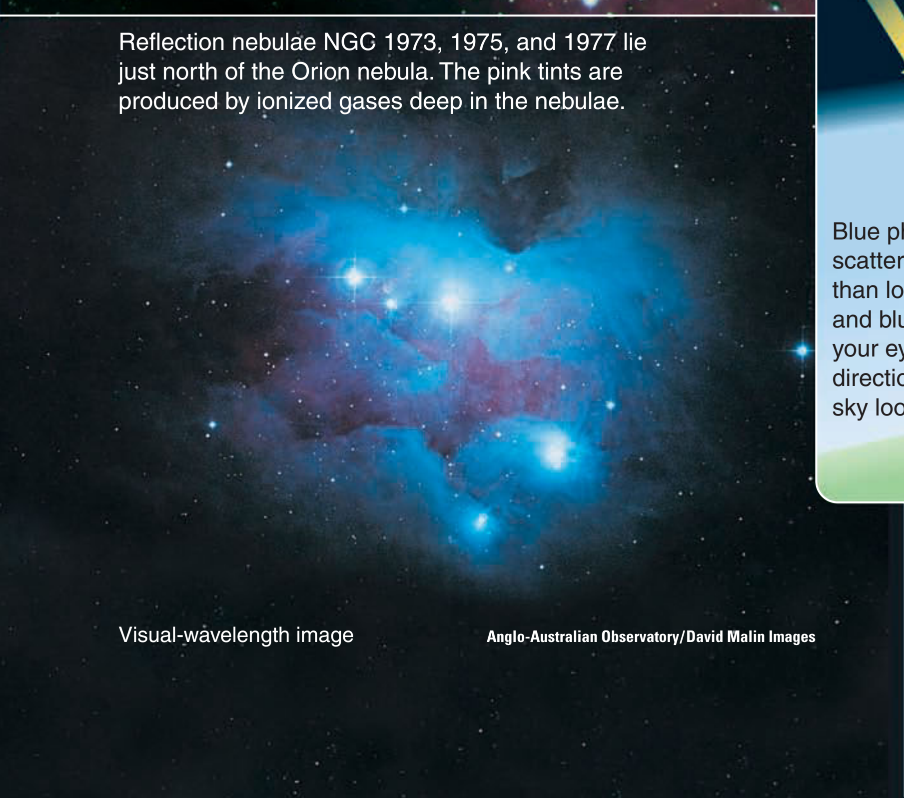
Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Scattering." Watch photons scatter through Earth's atmosphere.

2 A **reflection nebula** is produced when starlight scatters from a dusty nebula. Consequently, the spectrum of a reflection nebula is just the reflected absorption spectrum of starlight. Gas is surely present in a reflection nebula, but it is not excited to emit photons. See image below.

2a Reflection nebulae look blue for the same reason the sky looks blue. Short wavelengths scatter more easily than long wavelengths. See image below.



Reflection nebulae NGC 1973, 1975, and 1977 lie just north of the Orion nebula. The pink tints are produced by ionized gases deep in the nebulae.



Visual-wavelength image

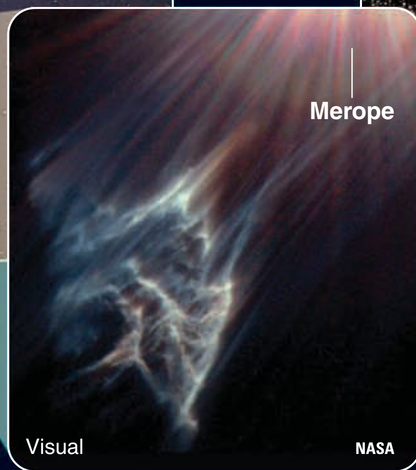
Anglo-Australian Observatory/David Malin Images

2b The blue color of reflection nebulae at left shows that the dust particles must be very small in order to preferentially scatter the blue photons. Interstellar dust grains must have diameters ranging from 0.01 mm down to 100 nm or so.

The hottest star in the Pleiades star cluster is Merope, a B3 star. It is not hot enough to ionize the gas so you see a reflection nebula rather than an emission nebula.

Visual-wavelength image

Caltech



Visual

NASA

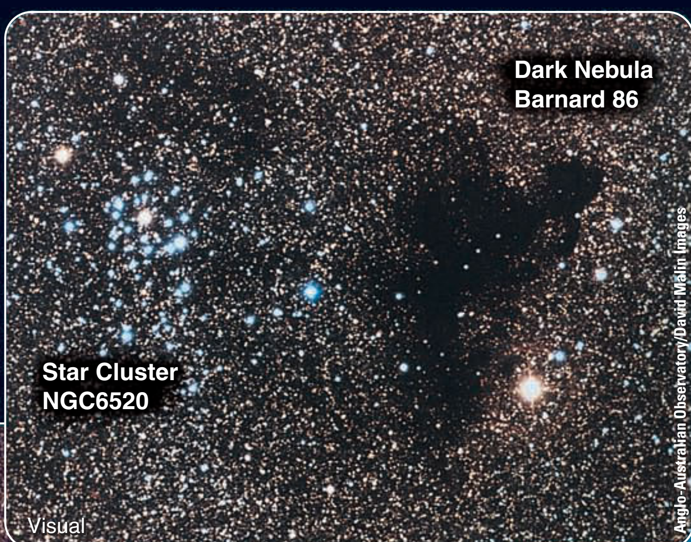
2c A dusty reflection nebula is located very close to the star Merope above. The glare from the star is caused by internal reflections in the telescope, but the wispy nature of the nebula is real. The intense light from the star is pushing the dust particles away and may destroy the little nebula over the next few thousand years.

The Milky Way in Sagittarius contains two nebulae that dramatically demonstrate the difference between emission and reflection nebulae.

Visual

Daniel Good

3 **Dark nebulae** are dense clouds of gas and dust that obstruct the view of more distant stars. Some are generally round, but others are twisted and distorted, as shown at the left, suggesting that even when there are no nearby stars to ionize the gas or produce a reflection nebula, there are breezes and currents pushing through the interstellar medium.



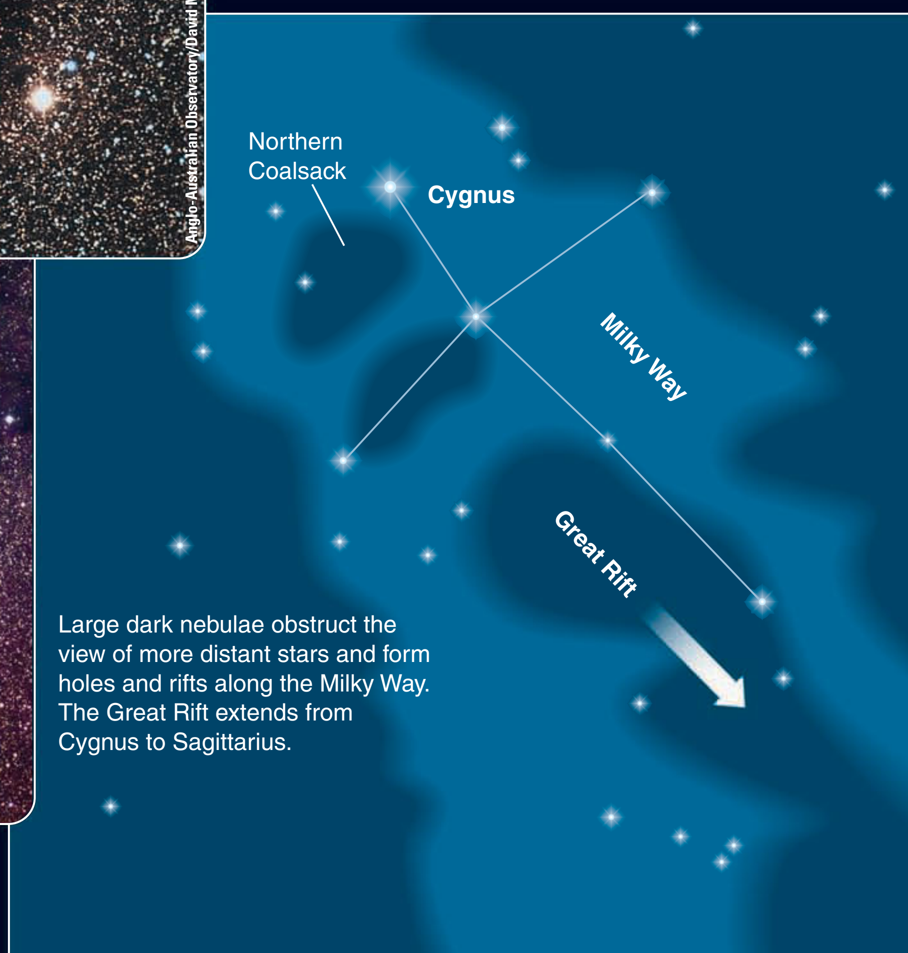
Visual

Anglo-Australian Observatory/David Malin Images

Anglo-Australian Observatory/
David Malin Images

Twisted by intense light from nearby stars, this dark nebula is visible because it obscures more distant stars.

Visual-wavelength image



Large dark nebulae obstruct the view of more distant stars and form holes and rifts along the Milky Way. The Great Rift extends from Cygnus to Sagittarius.

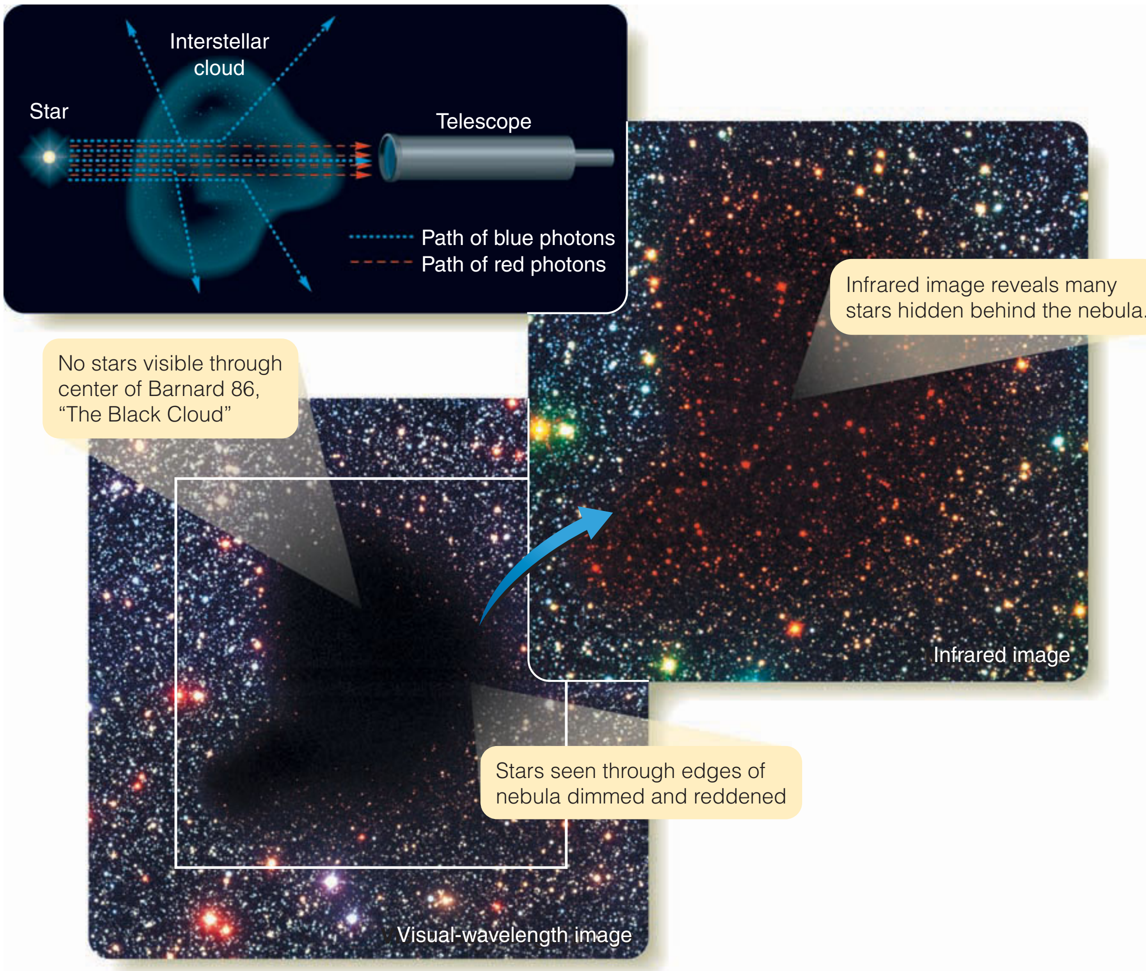


Figure 9-1

Interstellar reddening makes stars seen through a cloud of gas and dust look redder than they should because shorter wavelengths are more easily scattered. If the gas and dust is especially dense, no stars are visible through the cloud at visual wavelengths except near the edges. At the longer wavelengths of the near-infrared, many stars can be detected behind the cloud. (European Southern Observatory)

(Figure 9-2a). As starlight travels through the thin gas of the interstellar medium, gas atoms of elements such as calcium and sodium absorb photons of certain wavelengths, producing narrow **interstellar absorption lines.** You can be sure these lines originate in the interstellar medium because they appear in the spectra of O and B stars—stars that are too hot to form calcium and sodium absorption lines in their own atmospheres. Also, the narrowness of the interstellar lines indicates they could not have been formed in the hot atmospheres of the stars. Recall that in a hot gas, the atoms move rapidly, and the Doppler shifts of the different atoms smear the spectral lines, making them broad. The narrow lines must have formed in the interstellar medium, where gas atoms travel every slowly. The extremely narrow widths of these lines show that the gas of the interstellar medium is very cold, 10 to 50 K. Often multiple interstellar lines appear with slightly different Doppler shifts because the light from the star passed through a number of gas clouds on its way to Earth (Figure 9-2b).

Observations at nonvisible wavelengths provide valuable evidence about the interstellar medium. X-ray observations can detect regions of very hot gas apparently produced by exploding stars, and infrared observations can detect dust in the interstellar medium. Although interstellar dust grains are very small and very cold, there are huge numbers of grains in a cloud, and each

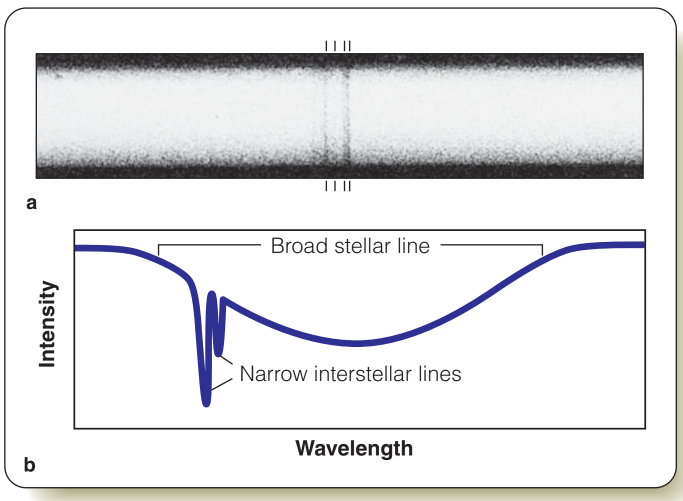


Figure 9-2

Interstellar absorption lines can be recognized in two ways. (a) The B0 supergiant ϵ Orionis is much too hot to show spectral lines of once-ionized calcium (Ca II), yet this short segment of its spectrum reveals narrow, multiple lines of Ca II (tic marks) that must have been produced in the interstellar medium. (The Observatories of the Carnegie Institution of Washington) (b) Spectral lines produced in the atmospheres of stars are much broader than the spectral lines produced in the interstellar medium. (Adapted from a diagram by Binnendijk) In both (a) and (b), the multiple interstellar lines are produced by separate interstellar clouds with slightly different radial velocities.

grain emits infrared radiation at long wavelengths. If the gas is cool enough, molecules can form, and, some molecules in the cold gas emit in the infrared, so infrared observations can detect very cold clouds of gas. Ultraviolet observations have been able to map the distribution of hydrogen in space and reveal that the sun is located just inside a cavity filled with hot gas. Furthermore, radio astronomers can study the radio emissions of specific molecules in the interstellar medium—the equivalent of emission lines in visible light.

■ Figure 9-3 shows regions where hot stars have inflated bubbles of intensely hot gas that push into the interstellar medium and cause the condensation of dust. The hot gas is detectable at X-ray wavelengths, and the dust is easily observed in the infrared. New stars can be born where the gas and dust is compressed.

There is no shortage of evidence that there is an interstellar medium. Now you can look for evidence linking clouds of gas and dust with the birth of stars.

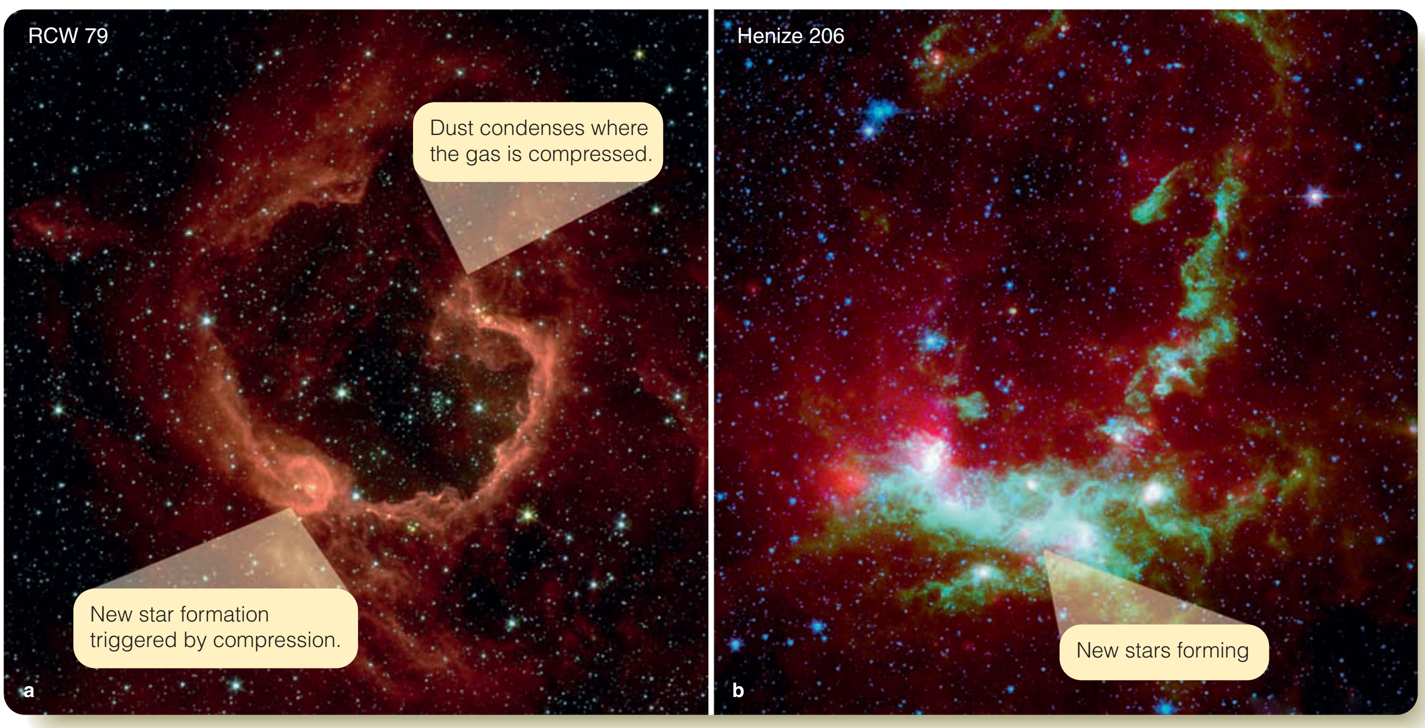
Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Why Is the Sky Blue?"

The Formation of Stars from the Interstellar Medium

To study the formation of stars, you must continue to compare theory with evidence. Theory predicts that over time the combined gravitational attraction of the atoms in a cloud of gas will draw the gas inward, pulling every atom toward the center. But not every cloud will collapse and form stars; the thermal energy in a cloud resists collapse. Temperature is a measure of the motion of the atoms or molecules in a material—in a hot gas, the atoms move more rapidly than do those in a cool gas. The interstellar clouds are very cold, but even at a temperature of only 10 K, the average hydrogen atom moves about 0.5 km/s (1100 mph). This thermal motion would make the cloud drift apart if gravity were too weak to hold it together.

Other factors can help a cloud resist its own gravity. Observations show that clouds are turbulent places with currents of gas pushing through and colliding with each other. Also, magnetic fields in clouds may resist being squeezed. These three factors—thermal motion, turbulence, and magnetic fields—resist gravity, so only the densest clouds are likely to contract.

The densest interstellar clouds contain from 10^3 to 10^5 atoms/cm^3, include from a few hundred thousand to a few million solar masses, and have temperatures as low as 10 K. In such clouds hydrogen can exist as molecules (H_2) rather than as atoms. These dense clouds are called **molecular clouds,** and the largest are called giant molecular clouds. Although hydrogen molecules cannot be detected by radio telescopes, the clouds can be mapped by the radio emission of carbon monoxide molecules (CO) present in small amounts in the gas. Stars form in these clouds when the densest parts become unstable and contract under the influence of their own gravity.



■ **Figure 9-3**

(a) Nebula RCW 79 is a bubble of intensely hot gas about 70 ly in diameter. It was produced over the last million years as hot gas and radiation flowed away from massive, hot stars near the center of the bubble. (NASA/JPL-Caltech/E. Churchwell, Univ. Wisconsin-Madison) (b) Millions of years ago a massive star exploded as a supernova in the upper left quarter of this image, and hot gas from that explosion is now pushing into cooler gas and causing the condensation of dust, which radiates strongly at infrared wavelengths. (NASA/JPL-Caltech/V. Gorjian)

Most clouds do not appear to be gravitationally unstable and will not contract to form stars on their own. However, a stable cloud colliding with a **shock wave** (the astronomical equivalent of a sonic boom) can be compressed and disrupted into fragments. Theoretical calculations show that some of these fragments can become dense enough to collapse under the influence of their own gravity and form stars (■ Figure 9-4).

Shock waves are necessary to trigger star formation, and space is filled with shock waves. A shock wave is a sudden change is gas pressure, and a number of processes can produce them. The sudden blast of light, especially ultraviolet radiation, from a newborn massive star can ionize and drive away nearby gas, forming an expanding shock wave. The collision of two interstellar clouds can produce a shock wave. Supernova explosions (exploding stars described in the next chapter) produce powerful shock waves. Examples of some of these processes are shown in Figure 9-3 and ■ Figure 9-5.

Although these are important sources of shock waves, the dominant trigger of star formation in our galaxy may be the spiral arms themselves. In Chapter 1, you learned that our galaxy contains spiral arms; as interstellar clouds encounter these spiral arms, the clouds are compressed, and star formation can be triggered.

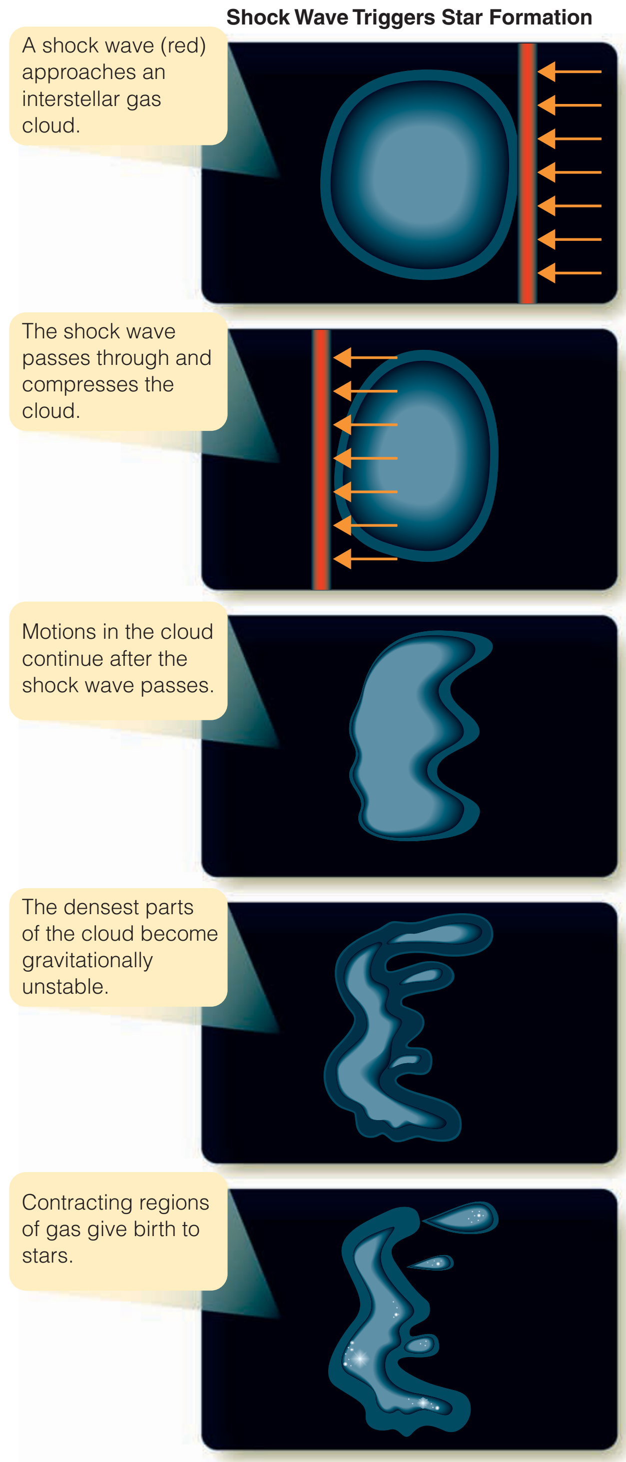
Once begun, star formation can spread like a grass fire. Astronomers have found a number of giant molecular clouds in which stars are forming in a repeating cycle. Both high-mass and low-mass stars form in such a cloud, but low-mass stars are not powerful enough to keep the star formation going. When massive stars form, however, their intense radiation and eventual supernovae explosions push back the surrounding gas and compress it. This compression in turn can trigger the formation of more stars, some of which will be massive. Thus, a few massive stars can drive a continuing cycle of star formation in a giant molecular cloud.

A collapsing cloud of gas breaks up because of instabilities in the contracting cloud and produces 10 to 1000 stars or more. Stars held together in a stable group by their combined gravity are called a **star cluster.** An **association** is a group of stars that are not gravitationally bound to one another. The stars in an association drift away from each other in a few million years. The youngest associations are rich in young stars, including O and B stars.

The Formation of Protostars

To follow the story of star formation further, you need to concentrate on a single fragment of a collapsing cloud as it forms a star. You might be wondering how the unimaginably cold gas of an interstellar cloud can heat up to form a star. The answer is gravity.

Once part of a cloud is triggered to collapse, gravity draws each atom toward the center. At first the atoms fall unopposed; they hardly ever collide with each other. In this free-fall contraction, the atoms pick up speed as they fall until, by the time the gas becomes dense enough for the atoms to collide often, they are traveling very fast. Now collisions convert the inward velocities of the atoms into random motions. Recall that temperature is a measure of the random velocities of the atoms in a gas. The in-



■ **Figure 9-4**

In this summary of a computer model, an interstellar gas cloud is triggered into star formation by a passing shock wave. The events summarized here might span about 6 million years.

Figure 9-5

(a) The blast of light and ultraviolet radiation from an earlier generation of massive stars has compressed neighboring gas and triggered the formation of more stars. Those stars are now triggering the birth of a third generation at right. (NASA, ESA, The Hubble Heritage Team, AURA/STScI) (b) Nebula DR 6, dubbed the Galactic Ghoul, is about 3.5 ly in diameter. It was formed by the hot winds and radiation from 10 massive stars at its center that are inflating the nebula, compressing nearby gas, and triggering more star formation. (NASA/JPL-Caltech/S. Carey)

ward collapse of the cold gas converts gravitational energy into high random velocities, causing the temperature to rise.

The initial collapse forms a dense core of gas, and, as more gas falls in, a warm protostar develops buried deep in the dusty gas. A **protostar** is an object that will eventually become a star.

When a star or protostar changes, it is said to evolve, and you can follow that evolution in the H–R diagram. As the object's temperature and luminosity change, its location in the H–R diagram shifts, and you can draw an arrow called an **evolutionary track** to represent these changes. Because protostars start out very cool and very faint, you must extend the H–R diagram to the right to include very low temperatures. ■ Figure 9-6, which includes temperatures below 100 K, shows the initial collapse of a 1-solar-mass cloud fragment as its temperature rises to more than 1000 K. The exact process is poorly understood, in part because the dusty cloud hides the protostar from sight during its contraction. If you could see a developing protostar, it would be a luminous red object a few thousand times larger than the sun, and you would plot it in the red-giant region of the H–R diagram. It is not, however, a real red giant, and it is invisible inside its dusty cloud.

The hot gas inside the protostar resists gravity, and the star can continue to contract only as fast as it can radiate energy into space. Although this contraction is much slower than the free-fall contraction, the star must continue to contract because its interior is not hot enough to generate nuclear energy.

Throughout its contraction, the protostar converts its gravitational energy into thermal energy. Half of this thermal energy radiates into space, but the remaining half raises the protostar's internal temperature. As the internal temperature climbs, the gas becomes ionized, becoming a mixture of positively charged atomic nuclei and free electrons. When the center gets hot enough, nuclear reactions begin generating energy, the protostar halts its contraction, and, having absorbed part of its cocoon of gas and dust and blown away the rest, it becomes a stable, main-sequence star.

The time it takes for a cool interstellar gas cloud to contract to the main sequence depends on its mass. The more massive the protostar, the stronger its gravity and the faster it contracts (■ Figure 9-7). The sun took about 30 million years to reach the main sequence, but a 15-solar-mass star can contract in only 160,000 years. Conversely, a star of 0.2 solar mass takes 1 billion years to reach the main sequence.

So far the story of the formation of a star from the interstellar medium has been based on theory. By understanding what

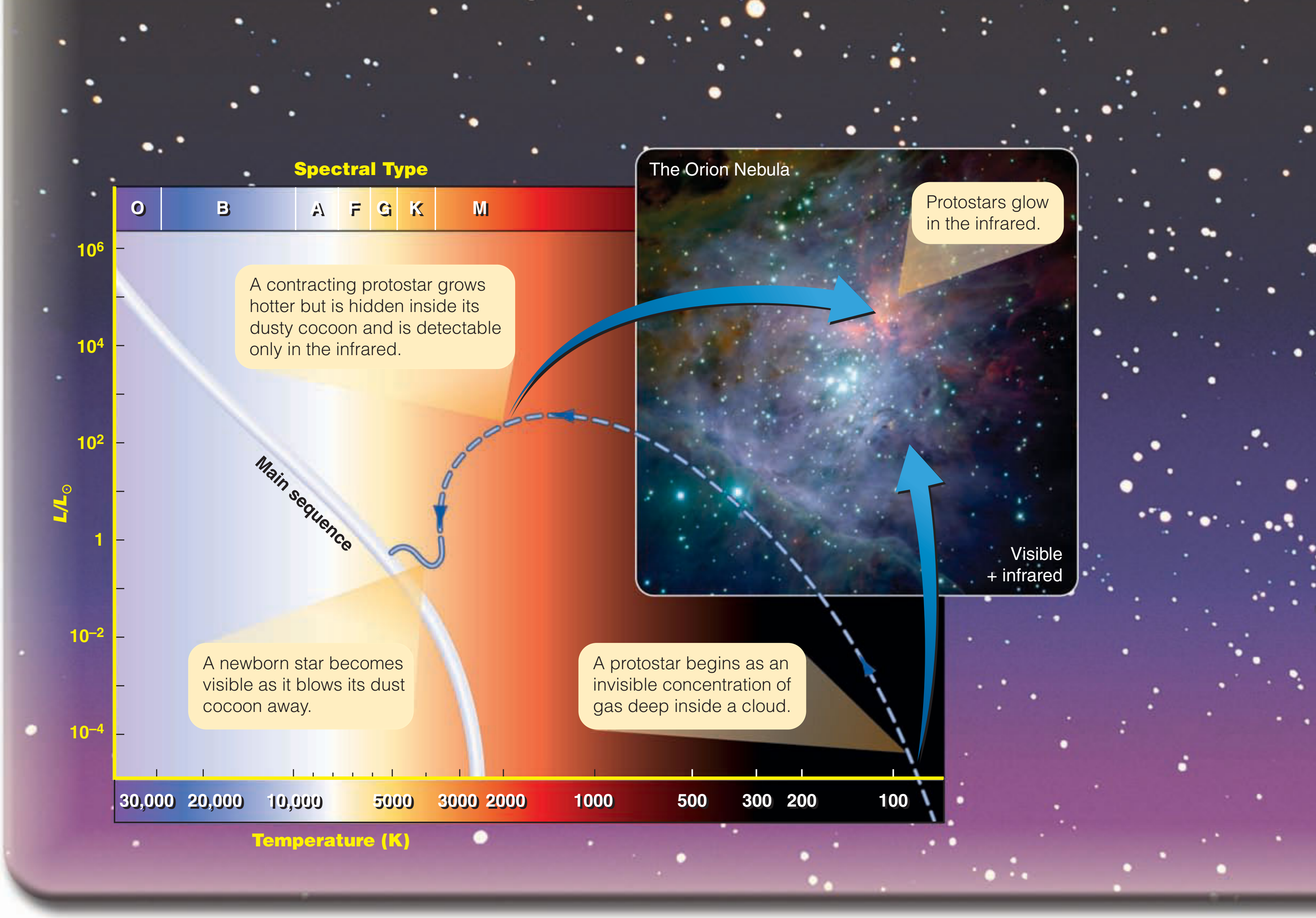


Figure 9-6

This H–R diagram has been extended to very low temperatures to show schematically the contraction of a dim, cool protostar. At visual wavelengths, protostars are invisible because they are deep inside dusty clouds of gas, but they are detectable at infrared wavelengths. The Orion Nebula contains both protostars and newborn stars that are just blowing their dust cocoons away. (ESO)

the interstellar medium is like and by knowing how the laws of physics work, astronomers have been able to tell the story of how stars form. But you can't accept a scientific theory without testing it, and that means you must compare the theory with the evidence. That constant checking of theories against evidence is the distinguishing characteristic of science, so it is time to carefully separate theory from evidence (■ How Do We Know? 9-1) and ask how much of the story of star formation can be observed.

Observations of Star Formation

In astronomy, evidence means observations, so you should expect astronomers to have observations that confirm their theories of star formation. Unfortunately, a protostar is not easy to observe. The protostar stage lasts for less than 0.1 percent of a star's total lifetime, so, although that is a long time in human terms, you cannot expect to find many stars in the protostar stage. Furthermore, protostars form deep inside clouds of dusty gas that absorb any light the protostar might emit. Only when the protostar is hot enough to drive away its enveloping cloud of gas and dust can it finally be seen at visual wavelengths. The **birth line** in the H–R diagram in Figure 9-7 shows where contracting protostars first become visible. Protostars cross the birth line shortly before they reach the main sequence. That means the early evolution of a protostar is hidden from sight.

Although astronomers cannot see protostars at visible wavelengths before the protostars cross the birth line, the forming stars can be detected in the infrared. The surrounding dust absorbs light from the protostar and grows warm, and warm dust radiates copious infrared. Infrared observations made by orbiting infrared telescopes reveal many bright sources of infrared radiation that are protostars buried in dust clouds.

Study ■ **Observational Evidence of Star Formation** on pages 168–169 and notice four important points and four new terms:

1. You can be sure that star formation is going on right now because you can find regions containing stars so young they must have formed recently. *T Tauri stars,* for example, are still in the process of contracting.

How Do We Know? 9-1

Separating Facts from Theories

When scientists disagree, what do they debate? The fundamental work of science is testing theories by comparing them with facts. The facts are evidence of how nature works and represent reality. Theories are attempts to explain how nature works. Scientists are very careful to distinguish between the two.

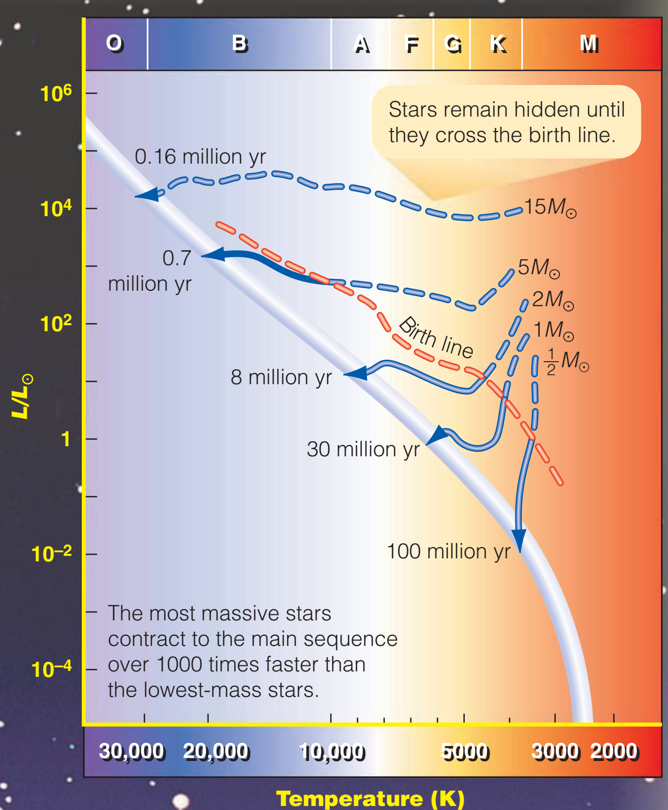
Scientific facts are those observations or experimental results of which scientists are confident. Ornithologists might note that fewer mountain thrushes are returning each spring to a certain mountain valley. Counting bird populations reliably is difficult and requires special techniques, but if the scientists made the observations correctly, they can be confident of their result and treat it as a fact.

To explain the declining population of thrushes, the scientists might consider a number of theories such as global warming or chemical pollution in the food chain. The ornithologists are free to combine or adjust their theories to better explain the bird migration, but they are not free to adjust their facts. Scientific facts are the hard pebbles of reality that can't be changed.

New facts can aggravate debates that are already politically charged. The declining number of mountain thrushes, for example, could be unwelcome news because addressing the root problem might cost taxpayer dollars or hurt local business interests. Nonscientists sometimes debate an issue by trying to adjust or even deny the facts, but scientists are not free to ignore a fact because it is unpopular or inconvenient. Scientists debate an issue by arguing about which theory applies or how a theory could be adjusted to fit the observed facts, but, once established, the facts themselves are not in question.

Whether scientists are measuring the density of an emission nebula or the size of a bird population, the final data become the reality against which theories are tested. When Galileo said we should "read the book of nature," he meant we should consult reality as the final check on our understanding.

In science, evidence is made up of facts, which could range from precise numerical measurement to the observation of the shape of a flower. (M. Seeds)



2 Visual and infrared observations can reveal small dusty clouds of gas called *Bok globules* that seem to be in the process of forming stars.

3 In the H-R diagram, newborn stars lie between the main sequence and the birth line—just where you would expect to find stars that have recently blown away their dust cocoons.

4 Finally, observations provide clues to the process by which stars form. Disks around protostars eject gas in *bipolar flows* that push into the surrounding interstellar medium and produce *Herbig–Haro objects.* Also, observations of small, dense cores of gas and dust within larger nebulae show how star formation begins.

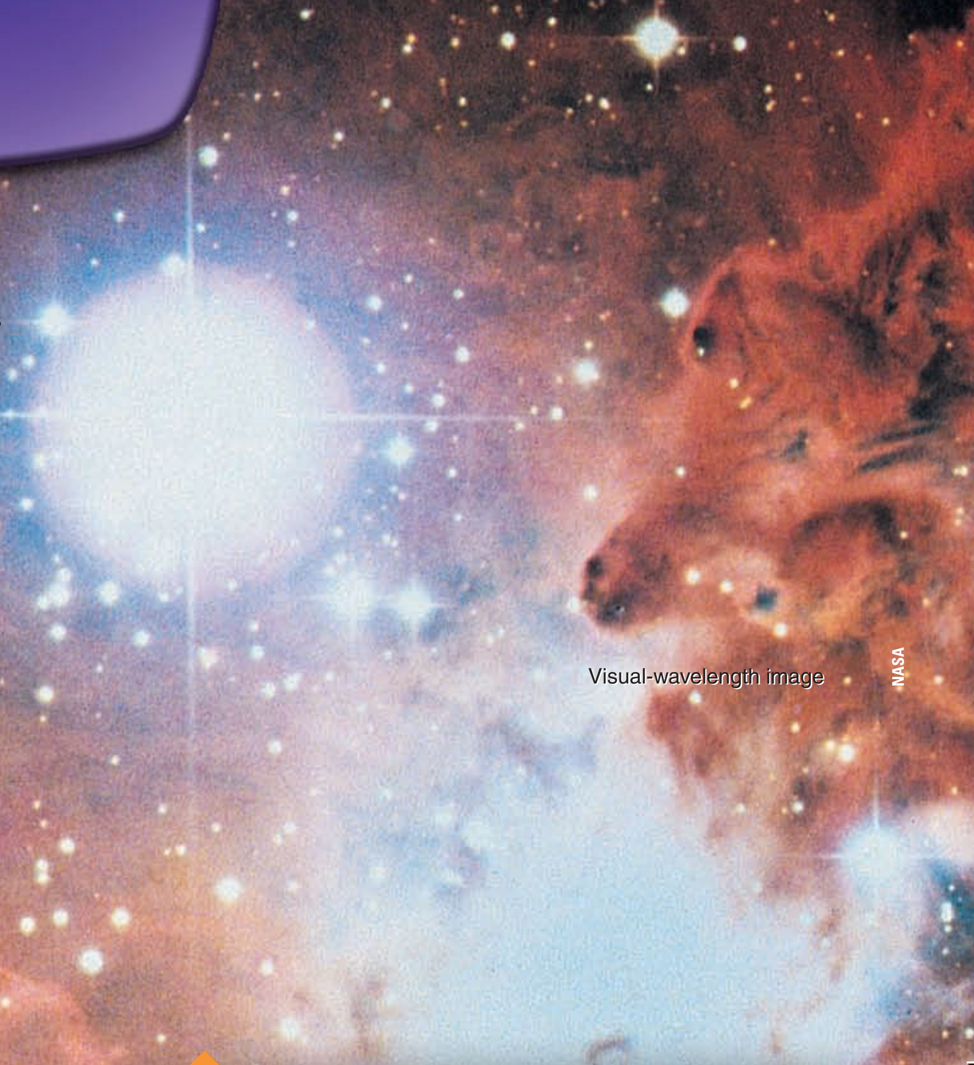
Figure 9-7

The more massive a protostar is, the faster it contracts. A 1-$M_\odot$ star requires 30 million years to reach the main sequence. (Recall that $M_\odot$ means "solar mass.") The dashed line is the birth line, where contracting protostars first become visible as they dissipate their surrounding clouds of gas and dust. Compare with Figure 9-6, which shows the formation of a star of about 1 $M_\odot$ as a dashed line up to the birth line and as a solid line from the birth line to the main sequence. (Illustration design by M.A. Seeds)

Observational Evidence of Star Formation

1 The nebula around the star S Monocerotis is bright with hot stars. Such stars live short lives of only a few million years, so they must have formed recently. Such regions of young stars are common. The entire constellation of Orion is filled with young stars and clouds of gas and dust.

Nebulae containing young stars usually contain **T Tauri** stars. These stars fluctuate irregularly in brightness, and many are bright in the infrared, suggesting they are surrounded by dust clouds and in some cases by dust disks. Spectra show that matter is falling into T Tauri stars while winds blow outward. The T Tauri stars are newborn stars just blowing away their dust cocoons. T Tauri stars appear to have ages ranging from 100,000 years to 10,000,000 years. Spectra of T Tauri stars show signs of an active chromosphere as we might expect from young, rapidly rotating stars with powerful dynamos and strong magnetic fields.


Visual-wavelength image

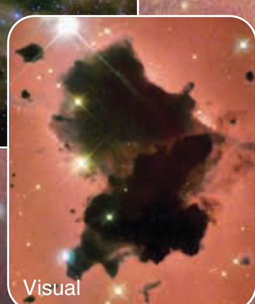
NASA


Visual

Skyfactory.org


Infrared

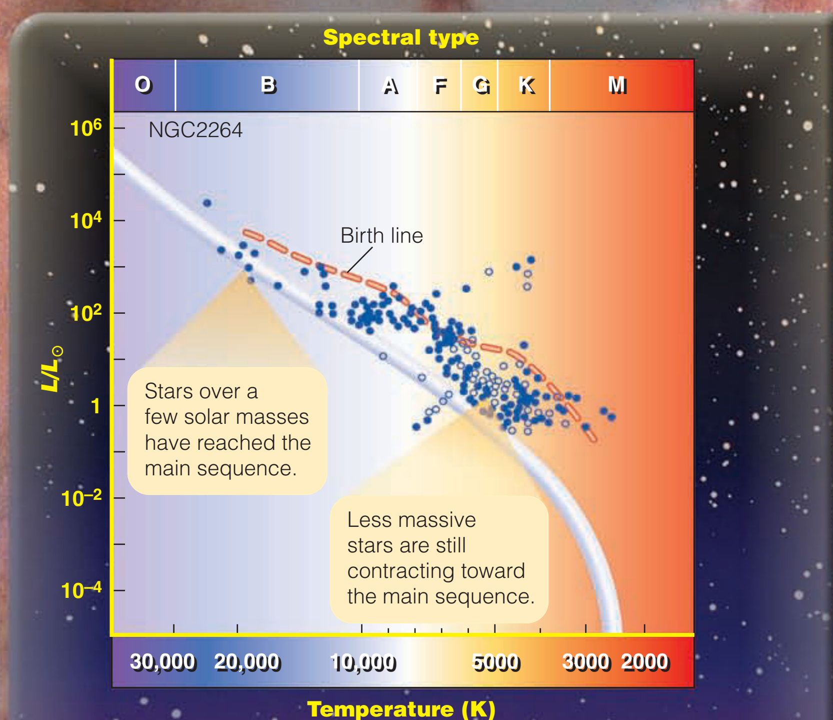
NASA/JPL-Caltech/W. Reach


Visual

NASA

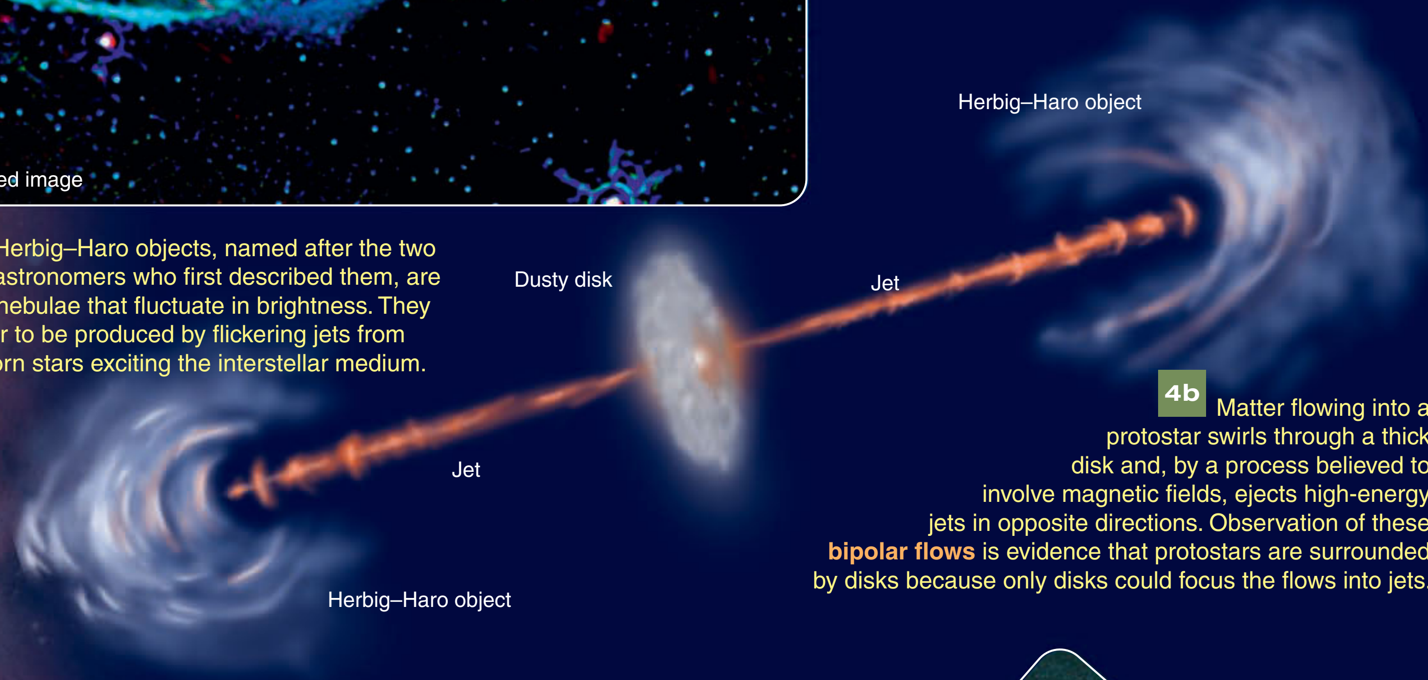
2 The Elephant Trunk (above) is a globule of dark nebula compressed and twisted by radiation and winds from a luminous star to the left of this image. Infrared observations reveal that it contains six protostars (pink images at lower edge) not detectable in visual images. The smallest dark globules are called **Bok globules** (above), named after astronomer Bart Bok. Only a light-year or so in diameter, they contain from 10 to 1000 solar masses.

3 The star cluster NGC 2264, embedded in the nebula on this page, is only a few million years old. Lower-mass stars have not yet reached the main sequence, and the cluster contains many T Tauri stars (open circles), which are found above and to the right of the main sequence, near the birth line. The faintest stars in the cluster were too faint to be observed in this study.

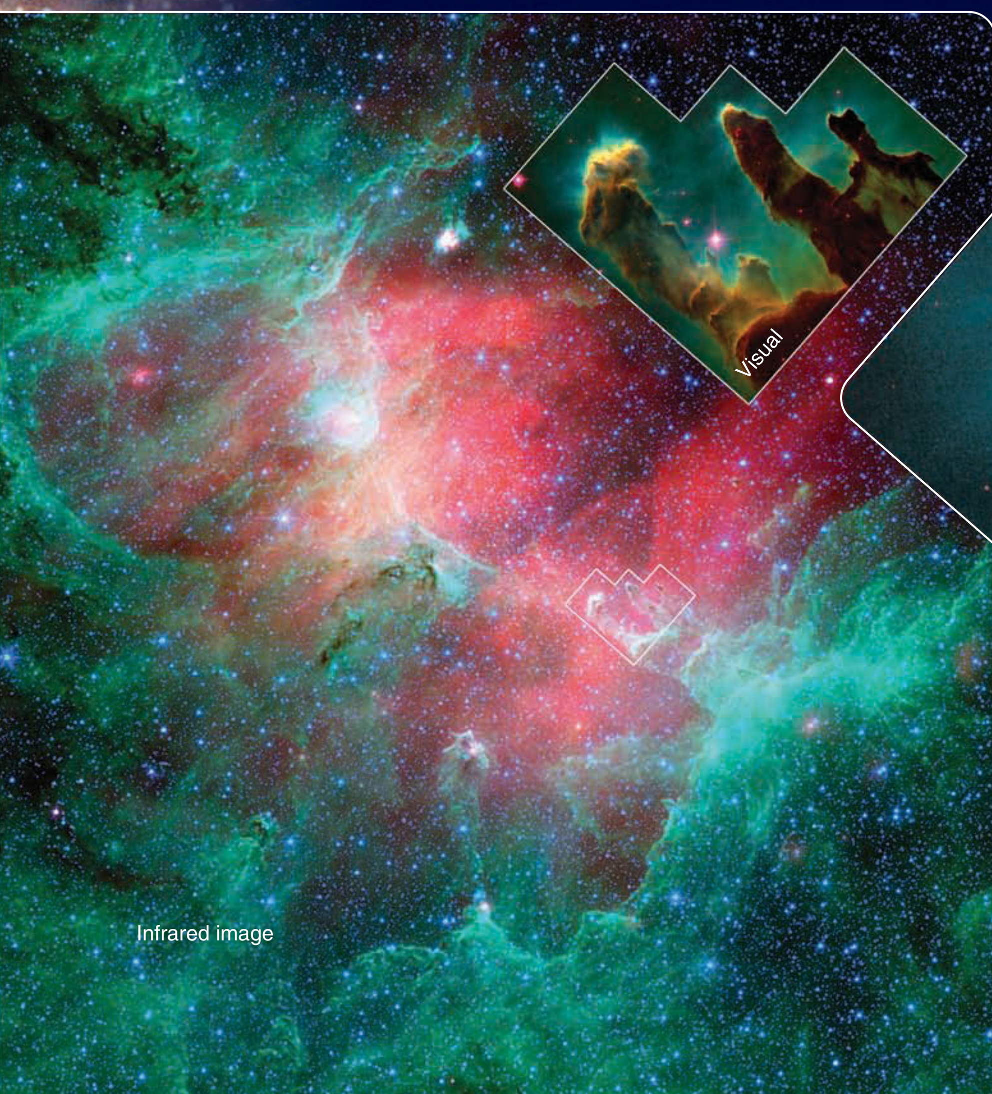


4 At the center of this image, a newborn star is emitting powerful jets to left and right. Where the jets strike the interstellar medium, they produce **Herbig–Haro objects.** Such jets can be over a light-year long and contain gas traveling at over 100 km/s.

4a Herbig–Haro objects, named after the two astronomers who first described them, are small nebulae that fluctuate in brightness. They appear to be produced by flickering jets from newborn stars exciting the interstellar medium.



4b Matter flowing into a protostar swirls through a thick disk and, by a process believed to involve magnetic fields, ejects high-energy jets in opposite directions. Observation of these **bipolar flows** is evidence that protostars are surrounded by disks because only disks could focus the flows into jets.



4c Radiation and winds from massive stars have shaped this nebulosity, and a recent supernova has heated some of the dust (red). Shock waves from the explosion will destroy the Eagle Nebula (inset) within about 1000 years. Erosion of part of the Eagle Nebula has exposed small globules of denser gas and dust (above). About 15 percent of these have formed protostars. Because these objects were first found in the Eagle Nebula, astronomers have enjoyed calling them EGGS—evaporating gaseous globules.

Star formation in a cloud of gas can produce lots of stars, but as soon as a massive star begins to shine, the cloud begins to suffer. The sudden blast of light and gas flowing away from the hot star blows the gas and dust away, and only the densest blobs of gas and dust near the star can resist. Like sunshades, the dense blobs protect the gas and dust behind them, and the result is **star formation pillars** that point like accusing fingers back at the massive star (■ Figure 9-8). You can now recognize the pillars in the Eagle Nebula (p. 169) as star formation pillars, dramatic evidence of the formation of at least one massive star.

All of these observations confirm the theories that describe contracting protostars. Although star formation still holds many mysteries, the general process seems clear. In at least some cases, interstellar gas clouds are compressed by passing shock waves, and the clouds' gravity, acting unopposed, draws the matter inward to form protostars.

Now you are ready to visit one of the most active regions of star formation visible from Earth, a nebula you can observe yourself.

The Orion Nebula

On a clear winter night, you can see with your naked eye the Great Nebula of Orion as a fuzzy wisp in Orion's sword. With binoculars or a small telescope it is striking, and through a large telescope it is breathtaking. At the nebula's center lie four brilliant blue-white stars known as the Trapezium, the brightest of a cluster of a few hundred stars. Surrounding the stars are the glowing filaments of a nebula more than 8 pc across. Like a great thundercloud illuminated from within, the churning currents of gas and dust suggest immense power. The significance of the Orion Nebula lies hidden, figuratively and literally, beyond the visible nebula. The entire region is ripe with star formation.

You should not be surprised to find star formation in Orion. The constellation is a brilliant landmark in the winter sky because it is marked by hot, blue stars. These stars are bright in your sky not because they are nearby but because they are tremendously luminous. These O and B stars cannot live more than a few million years, so you can conclude that they must have been born recently. Furthermore, the constellation contains large numbers of T Tauri stars, which are known to be young. Orion is rich with young stars.

The history of star formation in the constellation of Orion is written in its stars. The stars at Orion's west shoulder are about 12 million years old, whereas the stars of Orion's belt are about 8 million years old. The stars of the Trapezium at the center of the Great Nebula are no older than 2 million years. Apparently, star formation began near the west shoulder, and the massive stars that formed there triggered the formation of the stars you see in Orion's belt. That star formation probably triggered the formation of the stars you see in the Great Nebula. Like a grass fire, star formation has swept across Orion from northwest to southeast.

Study ■ **Star Formation in the Orion Nebula** on pages 172–173 and notice four points:

1. The nebula you see is only a small part of a vast, dusty molecular cloud. You see the nebula because the stars born within it have ionized the nearby gas and driven it outward, breaking out of the much larger molecular cloud.
2. A single very hot star is almost entirely responsible for producing the ultraviolet photons that make the nebula glow, and hot winds blowing from the most massive stars are inflating the nebula like a bubble.
3. Infrared observations reveal clear evidence of active star formation deeper in the molecular cloud just to the northwest of the Trapezium.
4. Finally, many stars visible in the Orion Nebula are surrounded by disks of gas and dust. Such disks do not last long and are clear evidence that the stars are very young. Note that these dusty disks are where planets form.

■ **Figure 9-8**

Light and gas flowing away from the massive star Eta Carinae out of the picture at the top are eroding this nebula and blowing parts of it away. Dense parts of the nebula are slower to erode and form star-formation pillars that point back at the massive star. (NASA/JPL-Caltech/N. Smith, Univ. of Colorado at Boulder)

In the next million years, the familiar outline of the Great Nebula will change, and a new nebula may begin to form as the protostars in the molecular cloud ionize the gas, drive it away, and become visible. Other centers of star formation may develop and then dissipate as massive stars are born and force the gas to expand. If enough massive stars are born, they can blow the entire molecular cloud apart and bring the successive generations of star formation to a final conclusion. The Great Nebula in Orion and its invisible molecular cloud are a beautiful and dramatic example of the continuing cycle of star formation.

◄ SCIENTIFIC ARGUMENT ►

What did Orion look like to the ancient Egyptians, to the first humans, and to the dinosaurs?

Scientific arguments can do more that support a theory; they can change the way you think of the world around you. The Egyptian civilization began only a few thousand years ago, and that is not very long in terms of the history of Orion. Today's hot, young stars are a few million years old, so the Egyptians saw the same constellation you see. (They called it Osiris.) Even the Orion Nebula hasn't changed very much in a few thousand years, and Egyptians may have admired it in the dark skies along the Nile.

Our oldest human ancestors lived about 4 million years ago, and that was about the time when the youngest stars in Orion were forming. Your earliest ancestors may have looked up and seen some of the stars you see, but some stars have formed since that time. Also, the Great Nebula is excited by the Trapezium stars, and they are not more than a few million years old, so your early ancestors probably didn't see the Great Nebula.

The last dinosaurs died about 65 million years ago, long before the birth of the brightest stars in Orion. The dinosaurs, had they the brains to appreciate the view, might have seen bright stars along the Milky Way, but they didn't see Orion. All of the stars in the sky are moving through space, and the sun is orbiting the center of our galaxy. Over many millions of years, the stars move appreciable distances across the sky. The night sky above the dinosaurs contained totally different star patterns.

The Orion Nebula is the product of a giant molecular cloud, but such a cloud can't continue spawning new stars forever. Focus your argument to answer the following: **What processes limit star formation in a molecular cloud?**

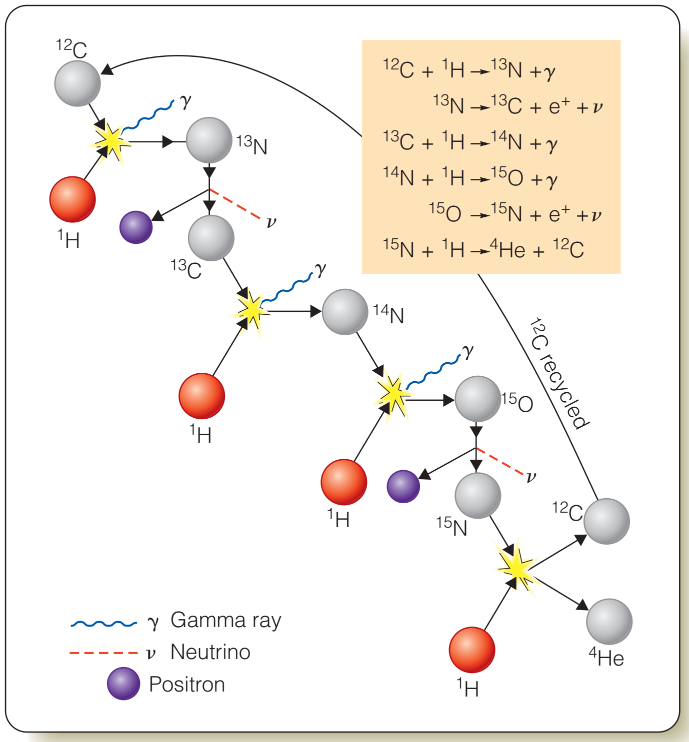
◄ ►

9-2 Fusion in Stars

Gravity makes protostars contract, and the contraction stops when the internal temperature rises high enough to start nuclear fusion. The outward flow of energy stops the contraction. When you studied the sun, you saw how it fuses hydrogen into helium in a chain of reactions called the proton–proton chain. Some stars generate energy in different ways, so it is time to explore those other fusion reactions.

The CNO Cycle

Main-sequence stars more massive than the sun fuse hydrogen into helium using the **CNO (carbon–nitrogen–oxygen) cycle,** a process that uses carbon, nitrogen, and oxygen as steppingstones (■ Figure 9-9). The CNO cycle begins with a carbon nucleus and transforms it first into a nitrogen nucleus, then into an oxygen nucleus, and then back to a carbon nucleus. The carbon is unchanged in the end, but along the way four hydrogen nuclei are fused to make a helium nucleus plus energy, just as in the proton–proton chain.



■ **Figure 9-9**

The CNO cycle uses ^{12}C as a catalyst to combine four hydrogen atoms (^1H) to make one helium atom (^4He) plus energy. The carbon atom reappears at the end of the process, ready to start the cycle over.

The CNO cycle requires a higher temperature because it begins with a carbon nucleus combining with a hydrogen nucleus. A carbon nucleus has a charge six times higher than hydrogen, so the Coulomb barrier is high, and the particles must collide at high velocities to force the particles close enough together. The CNO cycle requires temperatures higher than 16,000,000 K. The center of the sun is not quite hot enough, but stars more massive than about 1.1 solar masses have hotter cores and use the CNO cycle instead of the less efficient proton–proton chain.

Heavy-Element Fusion

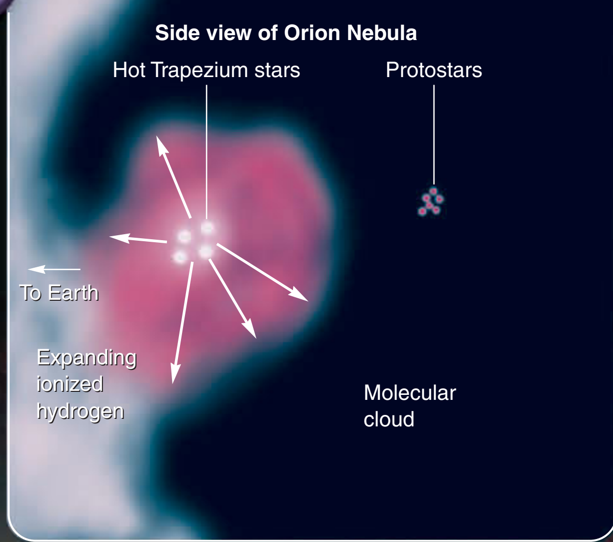
In the later stages of its life, when it has exhausted its hydrogen fuel, a star may fuse other nuclear fuels such as helium and carbon. Because these nuclei have higher positive charges, their Coulomb barriers are higher, and the nuclear reactions require higher temperatures. Helium fusion requires a temperature of at least 100 million K.

You can summarize the helium-fusion process in two steps:

$$^4He + {}^4He \rightarrow {}^8Be + \gamma$$
$$^8Be + {}^4He \rightarrow {}^{12}C + \gamma$$

Star Formation in the Orion Nebula

1 The visible Orion Nebula shown below is a pocket of ionized gas on the near side of a vast, dusty molecular cloud that fills much of the southern part of the constellation Orion. The molecular cloud can be mapped by radio telescopes. To scale, the cloud would be many times larger than this page. As the stars of the Trapezium were born in the cloud, their radiation has ionized the gas and pushed it away. Where the expanding nebula pushes into the larger molecular cloud, it is compressing the gas (see diagram at right) and may be triggering the formation of the protostars that can be detected at infrared wavelengths within the molecular cloud.

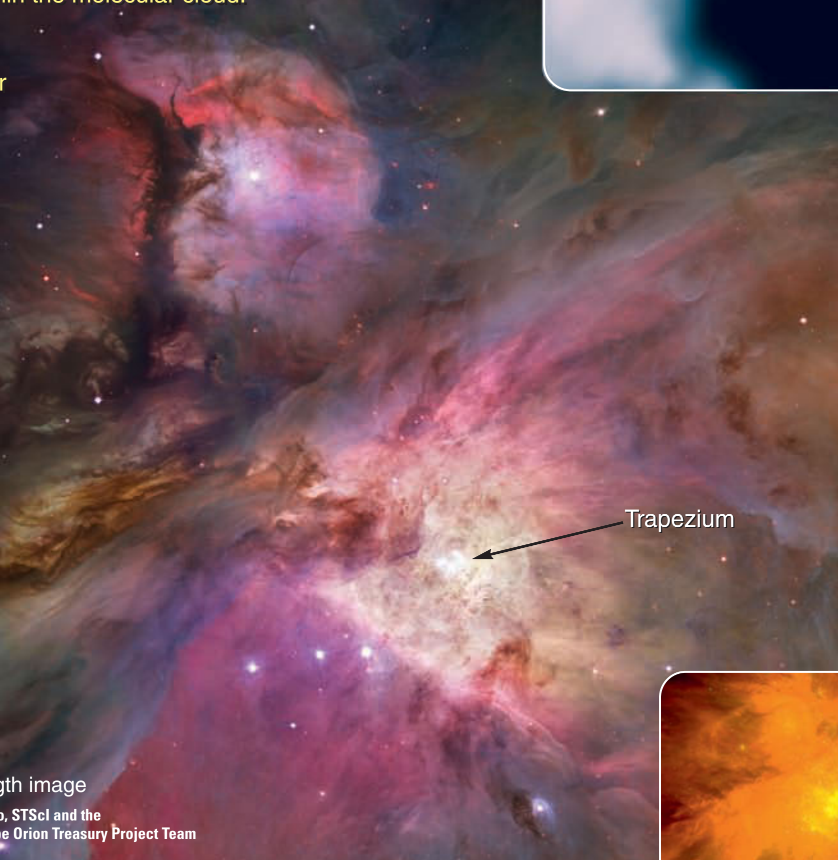


Hundreds of stars lie within the nebula, but only the four brightest, those in the Trapezium, are easy to see with a small telescope. A fifth star, at the narrow end of the Trapezium, may be visible on nights of good seeing.

The cluster of stars in the nebula is less than 2 million years old. This must mean the nebula is similarly young.

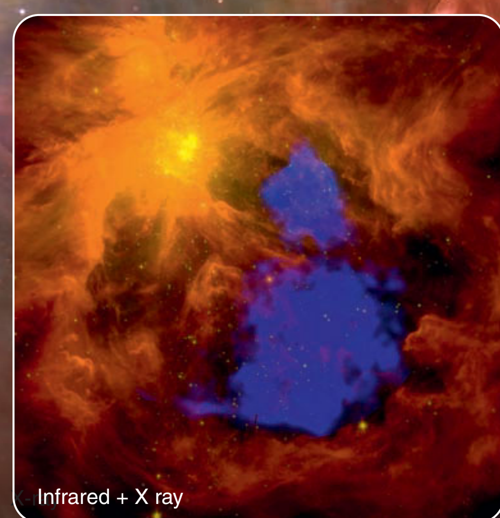
NASA

The near-infrared image above reveals about 50 low-mass, very cool stars that must have formed recently.



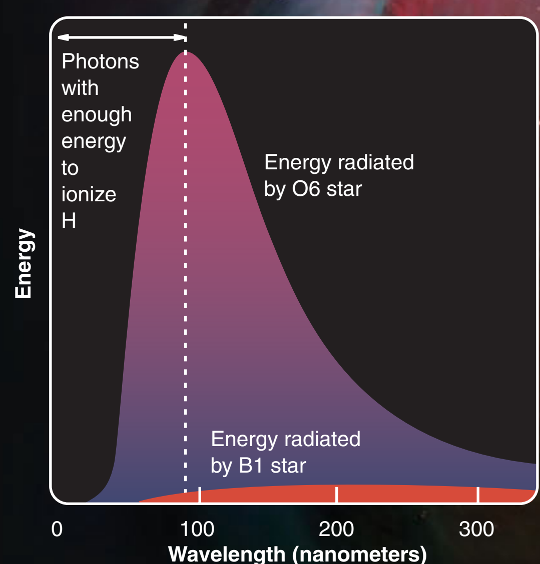
Visual-wavelength image

NASA,ESA, M. Robberto, STScI and the Hubble Space Telescope Orion Treasury Project Team



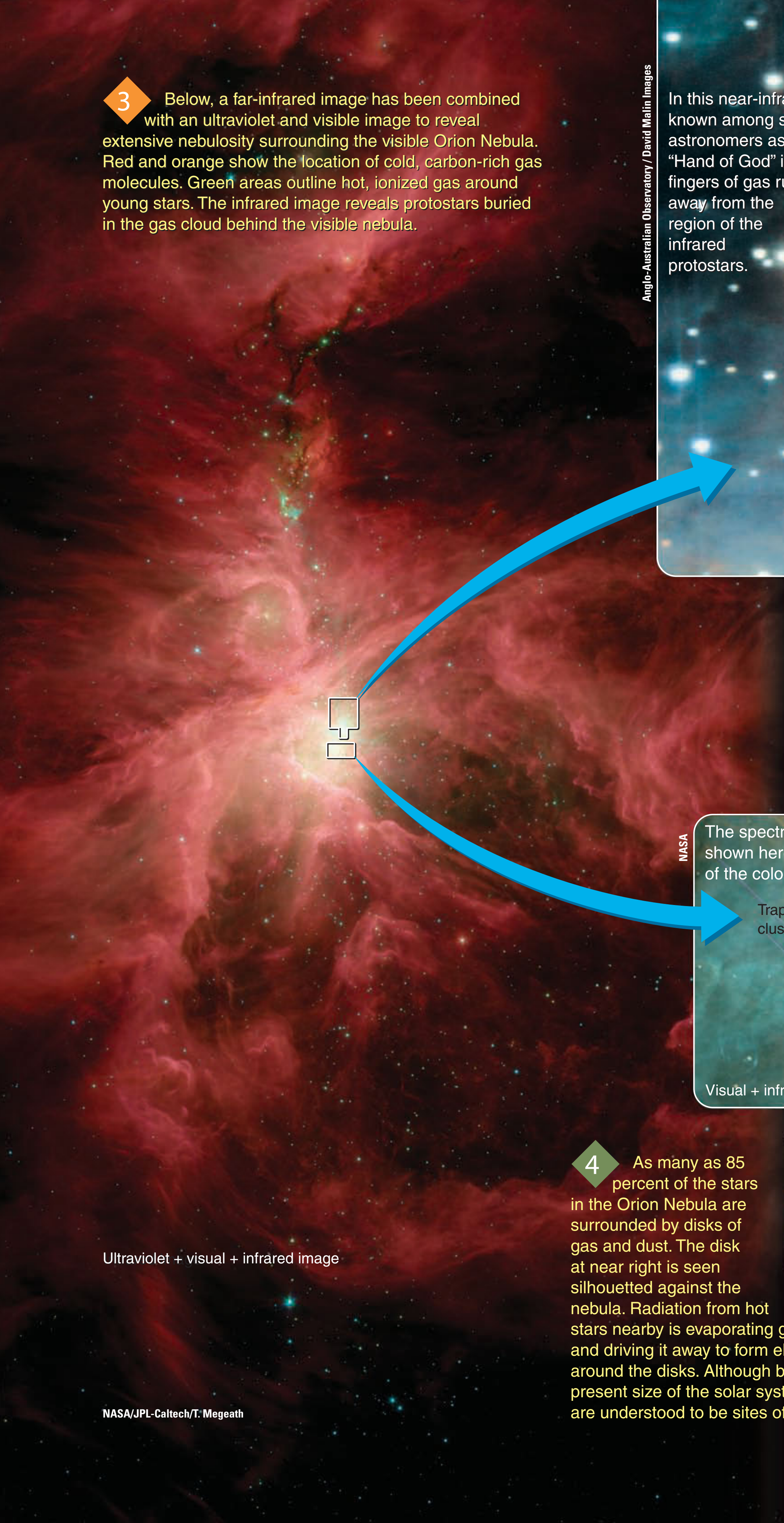
ESA XMM-Newton and NASA Spitzer data

Million-degree gas (blue) is produced by massive stars and inflates the Orion Nebula like a bubble.



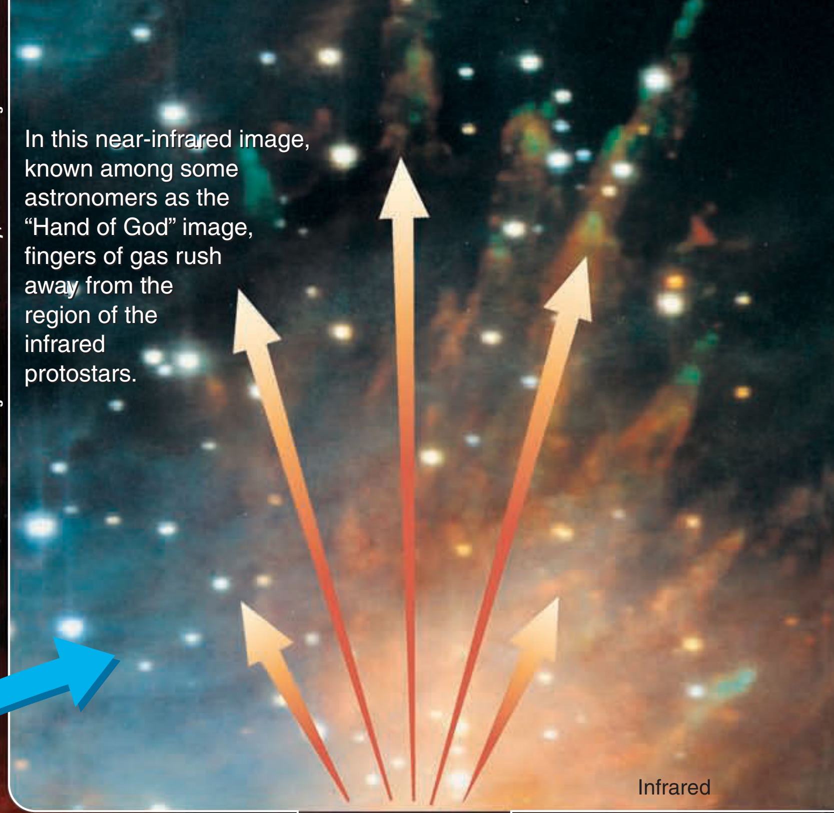
2 Of all the stars in the Orion Nebula, only one is hot enough to ionize the gas. Only photons with wavelengths shorter than 91.2 nm can ionize hydrogen. The second-hottest stars in the nebula are B1 stars, and they emit little of this ionizing radiation. The hottest star, however, is an O6 star 30 times the mass of the sun. At a temperature of 40,000 K, it emits plenty of photons with wavelengths short enough to ionize hydrogen. Remove that one star, and the nebula would turn off its emission.

3 Below, a far-infrared image has been combined with an ultraviolet and visible image to reveal extensive nebulosity surrounding the visible Orion Nebula. Red and orange show the location of cold, carbon-rich gas molecules. Green areas outline hot, ionized gas around young stars. The infrared image reveals protostars buried in the gas cloud behind the visible nebula.



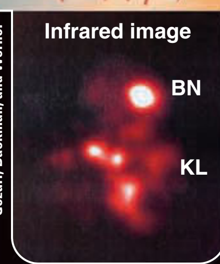
Ultraviolet + visual + infrared image

NASA/JPL-Caltech/T. Megeath



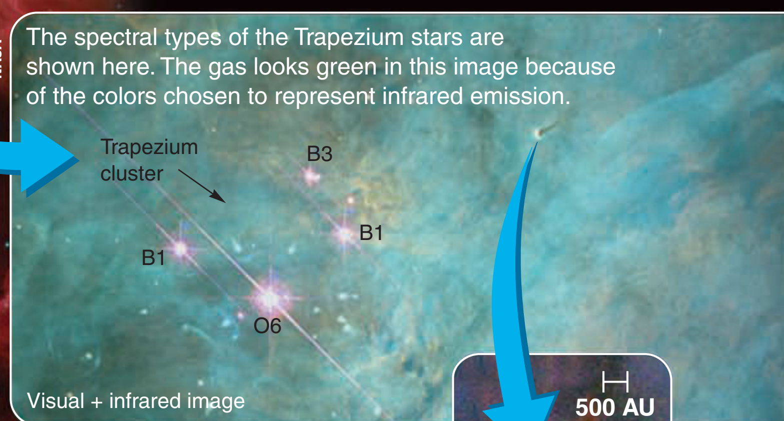
In this near-infrared image, known among some astronomers as the "Hand of God" image, fingers of gas rush away from the region of the infrared protostars.

Anglo-Australian Observatory/ David Malin Images



Gezari, Backman, and Werner

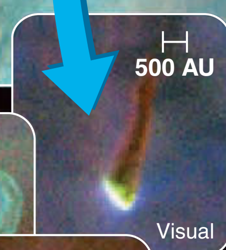
The Becklin-Neugebauer object (BN) is a hot B star just reaching the main sequence. It is not detectable at visual wave-lengths. The Kleinmann-Low nebula (KL) is a cluster of cool young protostars detectable only in the infrared.



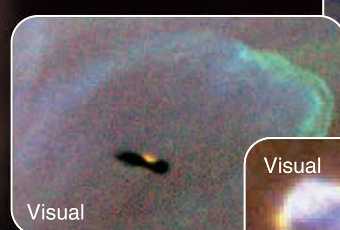
The spectral types of the Trapezium stars are shown here. The gas looks green in this image because of the colors chosen to represent infrared emission.

NASA

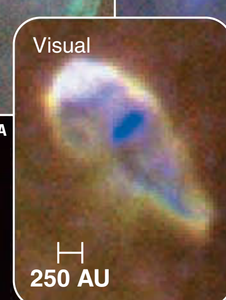
4 As many as 85 percent of the stars in the Orion Nebula are surrounded by disks of gas and dust. The disk at near right is seen silhouetted against the nebula. Radiation from hot stars nearby is evaporating gas from the disks and driving it away to form elongated nebulae around the disks. Although bigger than the present size of the solar system, such disks are understood to be sites of planet formation.



NASA



NASA



NASA

Because a helium nucleus is called an alpha particle, these reactions are commonly known as the **triple-alpha process.** Helium fusion is complicated by the fact that beryllium-8, produced in the first reaction of the process, is very unstable and may break up into two helium nuclei before it can absorb another helium nucleus. Three helium nuclei can also form carbon directly, but such a triple collision is unlikely.

At temperatures above 600,000,000 K, carbon fuses rapidly in a complex network of reactions illustrated in ■ Figure 9-10, where each arrow represents a different nuclear reaction. The process is complicated because nuclei can react by adding a proton, a neutron, or a helium nucleus or by combining directly with other nuclei. Unstable nuclei can decay by ejecting an electron, a positron, or a helium nucleus or by splitting into fragments.

Reactions at still higher temperatures can convert magnesium, aluminum, and silicon into yet heavier atoms. These reactions involving heavy elements will be important in the study of the deaths of massive stars in the next chapter.

The Pressure–Temperature Thermostat

Nuclear reactions in stars manufacture energy and heavy atoms under the supervision of a built-in thermostat that keeps the reactions from erupting out of control. That thermostat is the relation between gas pressure and temperature.

In a star, the nuclear reactions generate just enough energy to balance the inward pull of gravity. Consider what would happen if the reactions began to produce too much energy. The extra energy flowing out of the star would force it to expand. The expansion would lower the central temperature and density and slow the nuclear reactions until the star regained stability.

The same thermostat that keeps the reactions from running too fast also keeps the reactions from slowing down. Suppose the nuclear reactions began making too little energy. Then the star would contract slightly, increasing the central temperature and density and increasing the nuclear energy generation until equilibrium was regained.

The stability of a star depends on the relation between gas pressure and temperature. If an increase or decrease in temperature produces a corresponding change in pressure, then the thermostat is working correctly, and the star is stable. You will see in this chapter how the thermostat accounts for the mass–luminosity relation. In the next chapter, you will see what happens to a star when the thermostat breaks down completely and the nuclear fires rage unregulated.

Now that you know how stars form and how they make their energy, you are ready to descend inside the stars and see how they work.

9-3 Stellar Structure

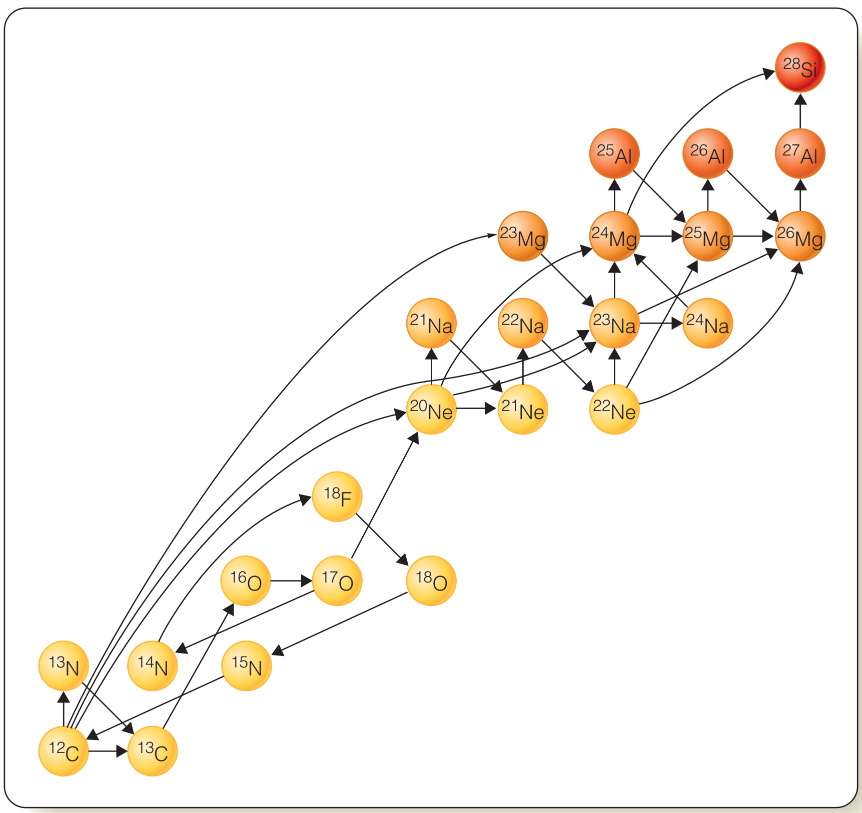
THE NUCLEAR FUSION at the centers of stars heats their interiors, creates high gas pressures, and balances the inward force of gravity. If there is a single idea in modern astronomy that can be called critical, it is this concept of balance. Stars are simple, elegant power sources held together by their own gravity and supported by their nuclear fusion.

Having explored the births of stars and the way they generate energy, you can now consider the structure of a star—the variation in temperature, density, pressure, and so on from the surface of the star to its center. A star's structure depends on how it generates its energy, on four simple laws of structure, and on what it is made of.

It will be easier to think about stellar structure if you imagine that the star is divided into concentric shells like those in an onion. You can then discuss the temperature, density, pressure, and so on in each shell. Of course, these helpful shells do not really exist; stars have no separable layers. The shells are just a mathematical convenience.

The Laws of Mass and Energy

The first two laws of stellar structure have something in common—they are both conservation laws. They say that certain things cannot be cre-



■ **Figure 9-10**

Carbon fusion involves many possible reactions that build numerous heavy atoms.

ated out of nothing or vanish into nothing. Such conservation laws apply to everything in the universe, but you can use them to understand the stars.

The law of **conservation of mass** says that the total mass of a star must equal the sum of the masses of its shells. This is like saying the weight of a cake must equal the sum of the weights of its layers.

The law of **conservation of energy** says that the amount of energy flowing out of the top of a layer in the star must be equal to the amount of energy coming in at the bottom plus whatever energy is generated within the layer. That means that the energy leaving the surface of the star, its luminosity, must equal the sum of the energies generated in all the layers inside the star. This is like saying that all the new cars driving out of a factory must equal the sum of all the cars made on each of the production lines.

These two laws may seem so familiar and so obvious that you hardly need to state them, but they are important clues to the structure of stars. The third law of stellar structure is familiar because you have been using a closely related law in the preceding sections.

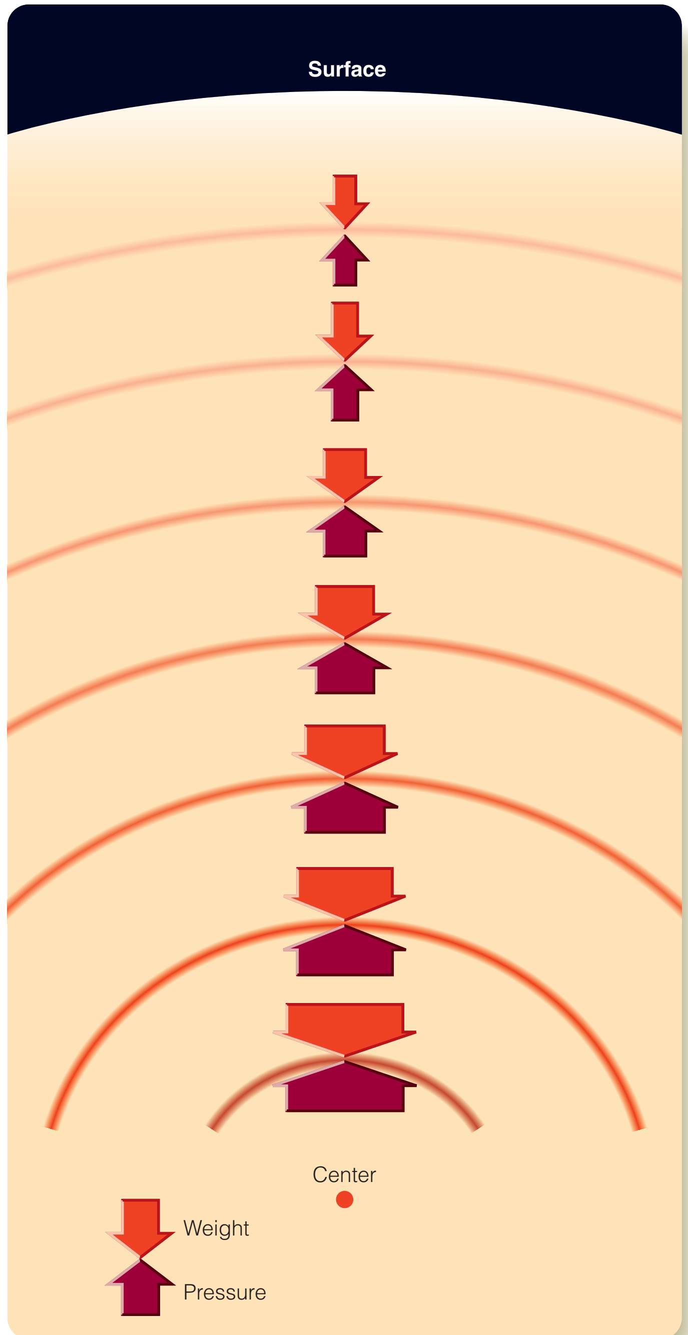
Hydrostatic Equilibrium

When you think about a star, it is helpful to think of it as if it were made up of layers. The weight of each layer must be supported by the layer below. The deeper layers must support the weight of all of the layers above. Because the inside of a star is made up of gas, the weight pressing down on a layer must be balanced by the gas pressure in the layer. If the pressure is too low, the weight from above will compress the layer, and if the pressure is too high, the layer will expand and lift the layers above.

This balance between weight and pressure is called **hydrostatic equilibrium.** The prefix *hydro* (from the Greek word for water) tells you the material is a fluid, the gases of a star, and the suffix *static* tells you the fluid is stable, neither expanding nor contracting.

■ Figure 9-11 illustrates this hydrostatic balance. The weight pressing down on each layer is shown by lighter red arrows, which grow larger with increasing depth as the weight grows larger. The pressure in each layer is shown by darker red arrows, which must grow larger with increasing depth to support the weight.

The law of hydrostatic equilibrium is the third law of stellar structure, and it can tell you something important about the inside of a star. The pressure in a gas depends on the temperature and density of the gas. Deep in the star, the pressure must be high, and that means that the temperature and density of the gas must also be high. Hydrostatic equilibrium tells you that temperature must increase with depth inside a star as each layer maintains the pressure needed to support the weight pressing downward. The layers are kept hot, as you have seen, by the energy flowing outward from the core of the star.



■ **Figure 9-11**

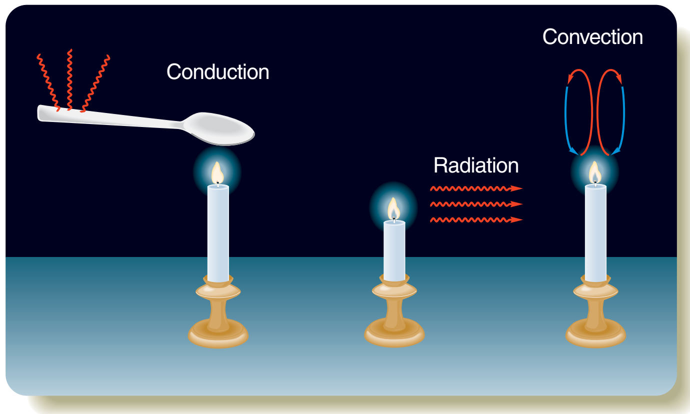
The law of hydrostatic equilibrium says the pressure in each layer of a stable star balances the weight on that layer. Thus, as the weight increases from the surface of a star to its center, the pressure also increases.

Now you should recognize hydrostatic equilibrium. It is closely related to the pressure–temperature thermostat discussed earlier.

Of course, exactly how hydrostatic equilibrium works depends on what an object is made of. Nearly all stars are made of hydrogen and helium gas, with some trace of heavier elements. To fully understand how a star works, astronomers must describe exactly how that gas responds to changes in temperature and pressure. Hydrostatic equilibrium also applies to planets, including Earth, but Earth is made of rock and metal, so understanding how hydrostatic equilibrium supports Earth requires that Earth scientists know how rock and metal respond to changes in temperature and pressure.

Although the law of hydrostatic equilibrium can tell you some things about the inner structure of stars, you need one more law to completely describe the interior of a star. You need a law that describes the flow of energy from the center to the surface.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Hydrostatic Equilibrium."



■ **Figure 9-12**

The three modes by which energy may be transported from the flame of a candle, as shown here, are the three modes of energy transport within a star. **Animated!**

Energy Transport

The surface of a star radiates light and heat into space and would quickly cool if that energy were not replaced. Because the inside of the star is hotter than the surface, energy must flow outward from the core, where it is generated, to the surface, where it radiates away. The flow of energy through the shells determines their temperature, which, as you saw previously, determines how much weight each shell can balance. To understand the structure of a star, you must understand how energy moves from the center through the shells to the surface.

In the sun, energy flows outward from the core as radiation and then, in the sun's outer layers, as convection. Other stars are similar to the sun, but there can be differences. Here you can apply what you know about the sun to stars in general.

The law of **energy transport** says that energy must flow from hot regions to cooler regions by conduction, convection, or radiation. Conduction is the most familiar form of heat flow. If you hold the bowl of a spoon in a candle flame, the handle of the spoon grows warmer. Heat, in the form of motion among the molecules of the spoon, is conducted from molecule to molecule up the handle, until the molecules of metal under your fingers begin to move faster and you sense heat (■ Figure 9-12). Conduction requires close contact between the molecules. Because the particles (atoms, ions, and electrons) in most stars are not in close contact, conduction is unimportant. Conduction is significant in white dwarfs, which have tremendous internal densities.

The transport of energy by radiation is another familiar experience. Put your hand beside a candle flame, and you can feel the heat. What you actually feel are infrared photons radiated by the flame (Figure 9-12). Because photons are packets of energy, your hand grows warm as it absorbs them. Recall that radiation is the principal means of energy transport in the sun's interior, where photons are absorbed and reemitted in random directions over and over as they work their way outward through the radiative zone.

The flow of energy by radiation depends on how difficult it is for the photons to move through the gas. If the gas is cool and dense, the photons are more likely to be absorbed or scattered, preventing the radiation from getting through easily. Such a gas is opaque. In a hot, thin gas, the photons can get through more easily; such a gas is less opaque. The **opacity** of the gas, its resistance to the flow of radiation, depends strongly on its temperature.

If the opacity is high, radiation cannot flow through the gas easily, and it backs up like water behind a dam. When enough heat builds up, the gas begins to churn as hot gas rises and cool gas sinks. This heat-driven circulation of a fluid is convection, the third way energy can move in a star. You are familiar with convection; the rising wisp of smoke above a candle flame is carried by convection. Energy is carried upward in these convection currents as rising hot gas (red in Figure 9-12) and also as sinking cool gas (blue in Figure 9-12).

Convection is important in stars both because it carries energy and because it mixes the gas. Convection currents flowing through the layers of a star tend to homogenize the gas, giving it a uniform composition throughout the convective zone. As you might expect, this mixing affects the fuel supply of the nuclear reactions, just as the stirring of a campfire makes it burn more efficiently.

The four laws of stellar structure are summarized in ■ Table 9-1. These laws, properly understood, can tell you how stars are born, how they live, and how they die.

Stellar Models

The laws of stellar structure, described in general terms in the previous sections, can be written as mathematical equations. By solving those equations in a special way, astronomers can build a mathematical model of the inside of a star.

If you wanted to build a model of a star, you would have to divide the star into about 100 concentric shells and then write

How Do We Know? 9-2

Mathematical Models

How can scientists study aspects of nature that cannot be observed directly? One of the most powerful tools in science is the mathematical model, a group of equations carefully designed to mimic the behavior of objects and processes that scientists want to study. Astronomers build mathematical models of stars to study the structure hidden deep inside them. Models can speed up the slow evolution of stars and slow down the rapid processes that generate energy. Stellar models are based on only four equations, but other models are much more complicated and may require many more equations.

For example, scientists and engineers designing a new airplane don't just cross their fingers, build it, and ask a test pilot to try it out. Long before any metal parts are made, mathematical models are created to test whether the wing design will generate enough lift, whether the fuselage can support the strain, and whether the rudder and ailerons can safely control the plane during takeoff, flight, and landing. Those mathematical models are put through all kinds of tests: Can a pilot fly with one engine shut down? Can the pilot recover from sudden turbulence? Can the pilot land in a crosswind? By the time the test pilot rolls the plane down the runway for the first time, the mathematical models have flown many thousands of miles.

Scientific models are only as good as the assumptions that go into them and must be compared with the real world at every opportunity. If you are an engineer designing a new airplane, you can test your mathematical models by making measurements in a wind tunnel. Models of stars are much harder to test against reality, but they do predict some observable things. Stellar models predict the existence of a main sequence, the mass–luminosity relation, the observed numbers of giant and supergiant stars, and other characteristics of H–R diagrams. Without mathematical models, astronomers would know little about the lives of the stars, and designing new airplanes would be a very dangerous business.

Before any new airplane flies, engineers build mathematical models to test its stability. (The Boeing Company)

Table 9-1 | The Four Laws of Stellar Structure

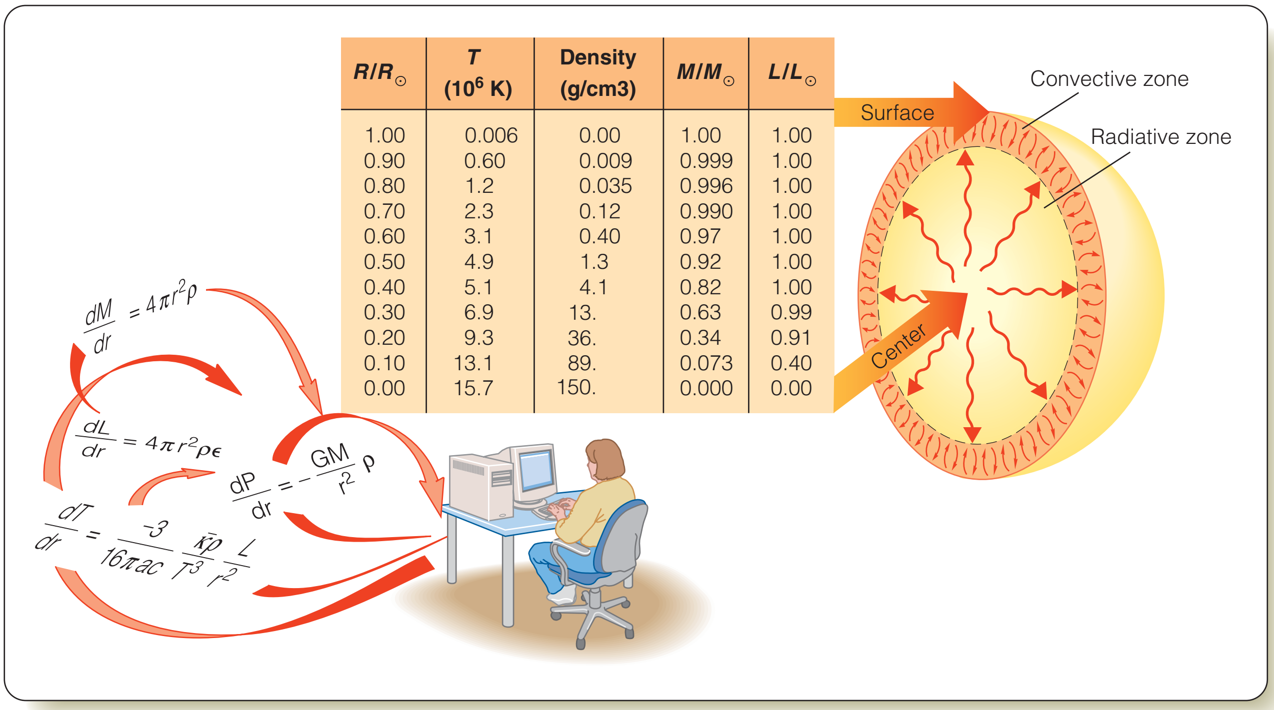
I. Conservation of mass	Total mass equals the sum of shell masses.
II. Conservation of energy	Total luminosity equals the sum of energy generated in each shell.
III. Hydrostatic equilibrium	The weight on each layer is balanced by the pressure in that layer.
IV. Energy transport	Energy moves from hot to cool regions by conduction, radiation, or convection.

down the four equations of stellar structure for each shell. You would then have 400 equations that would have 400 unknowns, namely, the temperature, density, mass, and energy flow in each shell. Solving 400 equations simultaneously is not easy, and the first such solutions, done by hand before the invention of the electronic computer, took months of work. Now a properly programmed computer can solve the equations in a few seconds and print a table of numbers that represents the conditions in each shell of the star. Such a table is a **stellar model**—a mathematical description of the inside of a star (■ How Do We Know? 9-2).

The table shown in ■ Figure 9-13 is a model of the sun. The bottom line, for radius equal to 0.00, represents the center of the sun, and the top line, for radius equal to 1.00, represents the surface. The other lines in the table tell you the temperature and density in each shell, the mass inside each shell, and the fraction of the sun's luminosity flowing outward through the shell. You can use the table to study conditions in the sun. For example, the bottom line tells you the temperature at the center of the sun is over 15 million Kelvin. At such a high temperature the gas is highly transparent, and energy flows as radiation. Nearer the surface, the temperature is lower, the gas is more opaque, and energy is carried by convection.

Stellar models also let astronomers look into a star's past and future. In fact, astronomers can use models as time machines to follow the evolution of stars over billions of years. To look into a star's future, astronomers can use a stellar model to determine how fast the star uses its fuel in each shell. As the fuel is consumed, the chemical composition of the gas changes, the opacity changes, and the amount of energy generated declines. By calculating the rates of these changes, astronomers can predict what the star will look like a few million years in the future. They can then repeat the process over and over and step-by-step follow the evolution of the star as it ages.

Although this sounds simple, it is actually a highly challenging problem involving nuclear and atomic physics, thermodynamics, and sophisticated computational methods. Only since the 1950s have electronic computers made the rapid calculation of stellar models possible, and the advance of astronomy since then has been heavily influenced by the use of such models to study the structure and evolution of stars. The summary of star formation in this chapter is based on thousands of stellar models.



$R/R_\odot$	T (10^6 K)	Density (g/cm3)	$M/M_\odot$	$L/L_\odot$
1.00	0.006	0.00	1.00	1.00
0.90	0.60	0.009	0.999	1.00
0.80	1.2	0.035	0.996	1.00
0.70	2.3	0.12	0.990	1.00
0.60	3.1	0.40	0.97	1.00
0.50	4.9	1.3	0.92	1.00
0.40	5.1	4.1	0.82	1.00
0.30	6.9	13.	0.63	0.99
0.20	9.3	36.	0.34	0.91
0.10	13.1	89.	0.073	0.40
0.00	15.7	150.	0.000	0.00

Figure 9-13

A stellar model is a table of numbers that represent conditions inside a star. Such tables can be computed using the four laws of stellar structure, shown here in mathematical form. The table in this figure describes the sun. (Illustration design by author)

You will continue to rely on theoretical models as you study main-sequence stars in the next section and the deaths of stars in the next chapter.

◄ SCIENTIFIC ARGUMENT ►

What would happen if the sun stopped generating energy?

Sometimes one of the best ways to test your understanding is to build an argument based on an altered situation. Stars are supported by the outward flow of energy generated by nuclear fusion in their interiors. That energy keeps each layer of the star just hot enough for the gas pressure to support the weight of the layers above. Each layer in the star must be in hydrostatic equilibrium; that is, the inward weight must be balanced by outward pressure. If the sun stopped making energy in its interior, nothing would happen at first, but over many thousands of years the loss of energy from its surface would reduce the sun's ability to withstand its own gravity, and it would begin to contract. You wouldn't notice much for 100,000 years or so, but eventually the sun would lose its battle with gravity.

Stars are elegant in their simplicity—nothing more than a cloud of gas held together by gravity and warmed by nuclear fusion. Now build a different argument. **How does the law of hydrostatic equilibrium assure you that stars are hot inside?**

◄ ►

9-4 Main-Sequence Stars

When a contracting protostar begins to fuse hydrogen, it stops contracting and becomes a stable main-sequence star. The most massive stars are so hot that they light up the remaining nearby gas in a beautiful nebula, as if announcing their birth (■ Figure 9-14). That gas is quickly blown away, and the stars begin their long, uneventful lives as main-sequence stars. If you discount the peculiar white dwarfs, then 90 percent of all true stars are main-sequence stars.

The Mass–Luminosity Relation

Observations of the temperatures and luminosities of stars show that main-sequence stars obey a simple rule—the more massive a star is, the more luminous it is. That rule, the mass–luminosity relation discussed in the previous chapter, is the key to understanding the stability of main-sequence stars. In fact, the mass–luminosity relation is predicted by the theories of stellar structure, giving astronomers direct observational confirmation of those theories.

To understand the mass–luminosity relation, you must recall the law of hydrostatic equilibrium, which says that pressure balances weight, and the pressure–temperature thermostat, which regulates energy production. A star that is more massive than the sun has more weight pressing down on its interior, so the interior must have a higher pressure to balance that weight. That means the massive star's automatic pressure–temperature thermostat must keep the gas in its interior hot and the pressure high. A star less massive than the sun has less weight on its interior and thus needs less internal pressure; therefore, its pressure–temperature thermostat is set lower. To sum up, massive stars are more lumi-

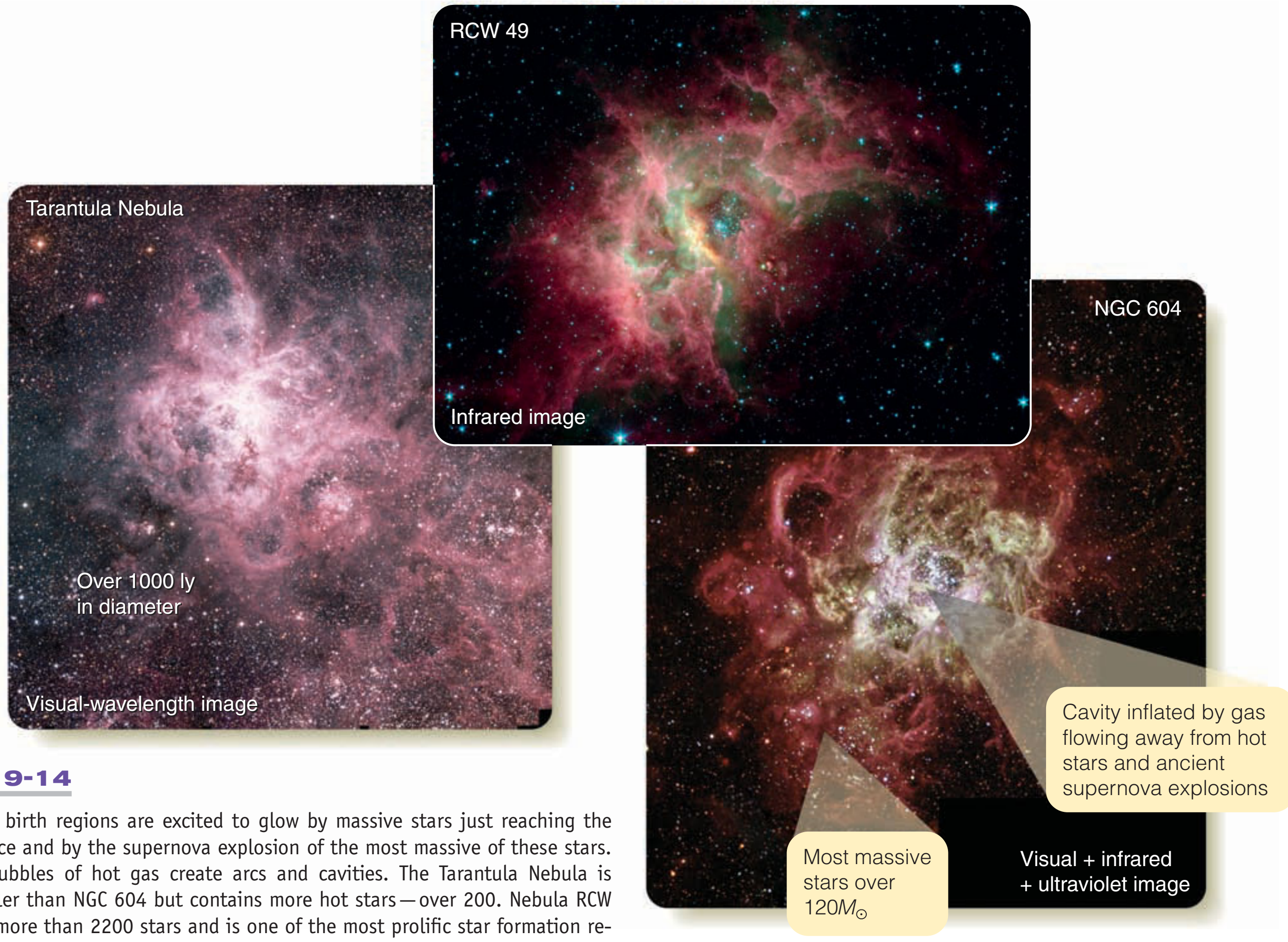


Figure 9-14

Monster star birth regions are excited to glow by massive stars just reaching the main sequence and by the supernova explosion of the most massive of these stars. Expanding bubbles of hot gas create arcs and cavities. The Tarantula Nebula is slightly smaller than NGC 604 but contains more hot stars—over 200. Nebula RCW 49 contains more than 2200 stars and is one of the most prolific star formation regions in our galaxy. (Tarantula: ESO; RCW 49: NASA/JPL-Caltech/E. Churchwell, Univ. of Wisconsin; NGC 604: NASA/Hubble Heritage Team/AURA/STScI)

nous because they must support more weight by making more energy.

The mass–luminosity relation tells you why the main sequence must have a lower end. Masses below about 0.08 solar mass do not have high pressures in their cores. Their central temperatures are too cool to allow hydrogen fusion. Called **brown dwarfs,** such objects are only about ten times larger than Earth, and although they are still warm from contraction, they do not generate energy. They have contracted as far as they can and are slowly cooling off.

Brown dwarfs fall in the gap between low-mass M stars and massive planets like Jupiter. They would look red to your eyes, but they emit most of their energy in the infrared. The warmer brown dwarfs fall in spectral class L and the cooler in spectral class T. Brown dwarfs are so cool that liquid and solid particles of silicates, metals, and other minerals can condense to form cloud layers in their atmospheres. Unlike stars, brown dwarfs appear to have rocky weather.

Because they are so small and cool, brown dwarfs are very-low-luminosity objects and are difficult to find. Nevertheless, hundreds are known. They may be as common as M stars. The evidence shows that nature does indeed make brown dwarfs when a protostar does not have enough mass to begin hydrogen fusion. This observational detection of the lower end of the main sequence further confirms the theories of stellar structure.

Now that you know how stars form and how they maintain their stability through the mass–luminosity relation, you can predict the evolution of main-sequence stars.

The Life of a Main-Sequence Star

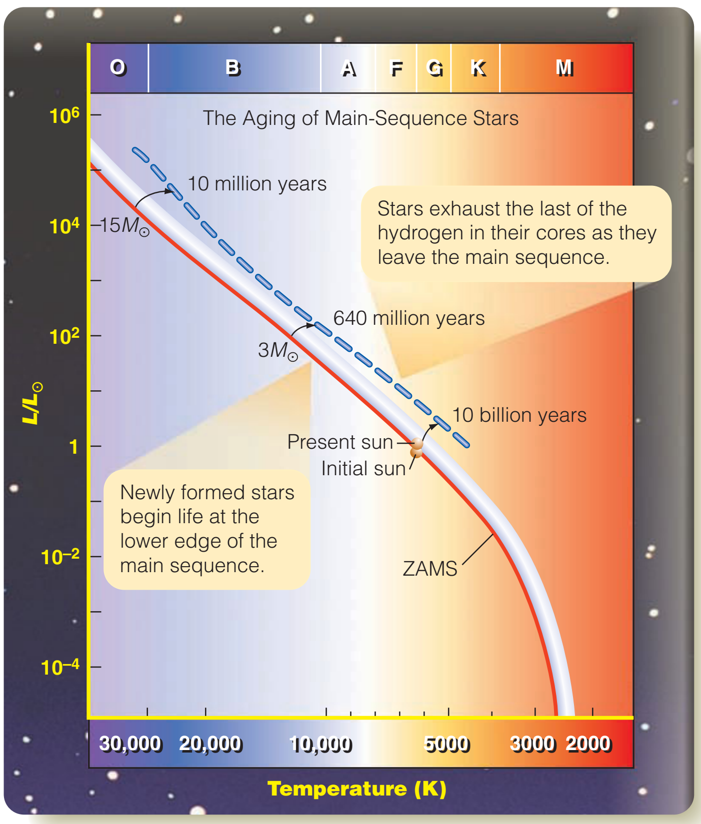
While a star is on the main sequence, it is stable, so you might think its life would be totally uneventful. But a main-sequence star balances its gravity by fusing hydrogen, and as the star gradually uses up its fuel, that balance must change. Thus, even the stable main-sequence stars are slowly changing as they consume their hydrogen fuel.

Recall that hydrogen fusion combines four nuclei into one. As a main-sequence star fuses its hydrogen, the total number of particles in its interior decreases. Each newly made helium nucleus exerts the same pressure as one hydrogen nucleus, but because the gas has fewer nuclei, its total pressure is less. This unbalances the gravity–pressure stability, and gravity squeezes the core of the star more tightly. As the core slowly contracts, its temperature increases, and the nuclear reactions run faster, releasing more energy. This additional energy flowing outward forces the outer layers to expand. As the star becomes gradually larger, it becomes more luminous, and eventually the expansion begins to cool the surface.

As a result of these gradual changes in main-sequence stars, the main sequence is not a sharp line across the H–R diagram but rather a band. Stars begin their stable lives fusing hydrogen on the lower edge of this band, which is known as the **zero-age main sequence (ZAMS),** but gradual changes in luminosity and surface temperature move the stars upward and slightly to the right, as shown in ■ Figure 9-15. By the time they reach the upper edge of the main sequence, they have exhausted nearly all the hydrogen in their centers. Thus you find main-sequence stars scattered throughout the band at various stages of their main-sequence lives.

■ Figure 9-15

Contracting protostars reach stability at the lower edge of the main sequence, the zero-age main sequence (ZAMS). As a star converts hydrogen in its core into helium, it moves slowly across the main sequence, becoming slightly more luminous and slightly cooler. Once a star consumes all of the hydrogen in its core, it can no longer remain a stable main-sequence star. More massive stars age rapidly, but less massive stars use up the hydrogen in their cores more slowly and live longer main-sequence lives.



The sun is a typical main-sequence star; as it undergoes these gradual changes, Earth will suffer. When the sun began its main-sequence life about 5 billion years ago, it was only about 70 percent as luminous as it is now, and by the time it leaves the main sequence in another 5 billion years, the sun will have twice its present luminosity. Long before that, the rising luminosity of the sun will drastically modify Earth's climate, and ultimately drive away our oceans and atmosphere. Life on Earth will probably not survive these changes in the sun, but we have a billion years or more to prepare.

The average star spends 90 percent of its life on the main sequence. This explains why 90 percent of all true stars are main-sequence stars—you are most likely to see a star during that long, stable period while it is on the main sequence. To illustrate, imagine that you photograph a crowd of 20,000 people. Everyone sneezes now and then, but the act of sneezing is very short compared with a human lifetime, so you would expect to find that very few people in your photograph are caught in the act of sneezing. Rather, most people in the photo would be in the much more common nonsneezing state. Your short human lifetime is like a snapshot of the universe; you see most stars on the main sequence, where they spend most of their time.

The number of years a star spends on the main sequence depends on its mass (■ Table 9-2). Massive stars consume fuel rapidly and live short lives, but low-mass stars conserve their fuel and shine for billions of years. For example, a 25-solar-mass star will exhaust its hydrogen and die in only about 7 million years. The sun has enough fuel to last about 10 billion years. The red dwarfs, although they have little fuel, use it up very slowly and may be able to survive for 100 billion years or more. (■ Reasoning with Numbers 9-1 explains how you can quickly estimate the life expectancies of stars from their masses.)

Nature makes more low-mass stars than high-mass stars, but this fact is not sufficient to explain the vast numbers of low-mass stars that fill the sky. An additional factor is stellar lifetime. Because low-mass stars live long lives, there are more of them in the sky than massive stars. Look at page 155 and notice how much

■ Table 9-2 | Main-Sequence Stars

Spectral Type	Mass (Sun = 1)	Luminosity (Sun = 1)	Years on Main Sequence
O5	40	405,000	1×10^6
B0	15	13,000	11×10^6
A0	3.5	80	440×10^6
F0	1.7	6.4	3×10^9
G0	1.1	1.4	8×10^9
K0	0.8	0.46	17×10^9
M0	0.5	0.08	56×10^9

more common the lower-main-sequence stars are than the massive O and B stars. The main-sequence K and M stars are so faint they are difficult to locate, but they are very common. The main-sequence O and B stars are luminous and easy to locate; but, because of their fleeting lives, there are never more than a few visible in the sky.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Mass–Lifetime Relation."

What Are We? Explainers

On cold winter nights when the sky is clear and the stars are bright, Jack Frost paints icy lacework across your windowpane. That's a fairy tale, of course, but it is a graceful evocation of the origin of frost. We humans are explainers, and one way to explain the world around us is to create myths.

Where did the stars come from? An ancient Aztec myth tells the story of the origin of the moon and stars. The stars, known as the Four Hundred Southerners, and the moon, the goddess Coyolxauhqui, plotted to murder their unborn brother, the great war god Huitzilopochtli. Hearing their plotting, he leaped from the womb fully armed, hacked Coyolxauhqui into pieces, and chased the stars away. At night you can see the Four Hundred Southerners scattered across the sky, and each month you can see the moon chopped into pieces as it passes through its phases.

Stories like these explain the origins of things and can make our universe more understandable and more comfortable. Science is a natural extension of our need to explain the world. The stories have become sophisticated scientific theories and are tested over and over against reality, but we humans build those theories for the same reason people used to tell mythical tales.

Reasoning with Numbers | 9-1

The Life Expectancies of Stars

You can estimate the amount of time a star spends on the main sequence—its life expectancy, τ—by estimating the amount of fuel it has and the rate at which it consumes that fuel:

$$\tau = \frac{\text{fuel}}{\text{rate of consumption}}$$

The amount of fuel a star has is proportional to its mass M, and the rate at which it uses up its fuel is proportional to its luminosity L. Thus its life expectancy must be proportional to M/L. You can simplify this equation further because, as you saw in the last chapter, the luminosity of a star depends on its mass raised to the 3.5 power ($L = M^{3.5}$). So the life expectancy is

$$\tau = \frac{M}{M^{3.5}}$$

which is the same as

$$\tau = \frac{1}{M^{2.5}}$$

If you express the mass in solar masses, the lifetime will be in solar lifetimes.

Example: How long can a 4-solar-mass star live?

Solution:

$$\tau = \frac{1}{4^{2.5}} = \frac{1}{4 \cdot 4\sqrt{4}}$$
$$= \frac{1}{32} \text{ solar lifetimes}$$

Studies of solar models show that the sun, presently 5 billion years old, can last another 5 billion years. So a solar lifetime is approximately 10 billion years, and a 4-solar-mass star will last for about

$$\tau = \frac{1}{32} \times (10 \times 10^9 \text{ yr})$$
$$= 310 \times 10^6 \text{ years}$$

Study and Review

Summary

- Stars are born from the gas and dust of the **interstellar medium (p. 159).**
- Astronomers know there is an interstellar medium because they can see **nebulae (p. 159):** glowing **emission nebulae (p. 160),** also called **HII regions (p. 160);** blue **reflection nebulae (p. 160);** and **dark nebulae (p. 161).** Also, they can detect **interstellar absorption lines (p. 162)** in the spectra of distant stars. The interstellar medium makes distant stars look fainter, and the **interstellar dust (p. 159)** makes distant stars look redder than expected, an effect called **interstellar reddening (p. 159).**
- In large, dense, cool clouds of gas, hydrogen can exist as molecules rather than as atoms. These **molecular clouds (p. 163)** are sites of star formation. Such clouds can be triggered to collapse by collision with a **shock wave (p. 164),** which compresses and fragments the gas cloud. A single fragment of a cloud can produce a **protostar (p. 165),** and the entire cloud can produce a **star cluster (p. 164)** containing hundreds of stars. An **association (p. 164)** is a group of stars that have formed together but are not close enough to each other to be held in a star cluster.
- A contracting protostar is large and cool and follows an **evolutionary track (p. 165)** in the H–R diagram that leads it through the red-giant region. It is not visible during this stage because it is surrounded by a cocoon of dust and gas. The dust in the cocoon absorbs the protostar's light and reradiates it as infrared radiation. Many infrared sources are probably protostars. A protostar becomes visible as it crosses the **birth line (p. 166)** in the H-R diagram and blows away its cocoon of gas and dust.
- **T Tauri stars (p. 168)** are young stars between the main sequence and the birth line. They are absorbing and blowing away their surrounding dust and gas. Very young star clusters contain large numbers of T Tauri stars.
- The smallest, darkest clouds are called **Bok globules (p. 168).** Many are in the process of forming stars.
- Some protostars have been found emitting **bipolar flows (p. 169)** of gas as they expel material from their surrounding disks. Where these flows strike existing clouds of gas, astronomers can see nebulae called **Herbig–Haro objects (p. 169).** The bipolar flows are evidently focused by disks of gas and dust around the protostars.
- Gas and radiation flowing away from a newly formed massive star can blow away nearby gas and dust forming **star formation pillars (p. 172).** Where nearby gas and dust clouds are compressed, new star formation can be triggered.
- The Great Nebula in Orion is an active region of star formation. The bright stars in the center of the nebula formed within the last few million years, and infrared telescopes detect protostars buried inside the molecular cloud that lies behind the visible nebula.
- As a protostar grows hot enough to begin hydrogen fusion at its core, it settles onto the main sequence. Stars of the sun's mass or less make their energy by fusing hydrogen into helium in a process called the proton–proton chain. More massive main-sequence stars fuse hydrogen into helium in the **CNO cycle (p. 173).**
- Later in their evolution, stars can fuse helium into carbon in the **triple-alpha process (p. 174).** Some stars can fuse carbon into even heavier elements.
- The relationship between pressure and temperature, the pressure–temperature thermostat, ensures that the star generates just enough energy to support itself.
- The four basic laws that describe the inside of a star include the law of **conservation of mass (p. 175)** and the law of **conservation energy (p. 175).** The law of **hydrostatic equilibrium (p. 175)** shows how the pressure in a layer supports the weight pressing down, and the law of **energy transport (p. 176)** says that energy must flow outward by conduction, convection, or radiation. Conduction is not usually important inside stars. The **opacity (p. 176)** of a gas is its resistance to the flow of radiation.
- Astronomers can study the interiors of stars and the way they change over time by calculating detailed **stellar models (p. 177)** based on the four laws above.
- The mass-luminosity relation among main-sequence stars can be understood from stellar models. More massive stars have more weight to support, and their pressure–temperature thermostats must make more energy. That makes them more luminous.
- The main sequence has a lower end because stars less massive than 0.08 solar mass cannot get hot enough to begin hydrogen fusion. Such objects become **brown dwarfs (p. 179).**
- A contracting protostar begins its life on the **zero-age main sequence (p. 180),** but as it combines hydrogen nuclei to make helium nuclei, the total number of nuclei in its core declines. The core slowly contracts, and the outer layers gradually expand, making the star move upward and to the right across the band of the main sequence.
- How long a star can remain on the main sequence depends on its mass. The more massive a star is, the faster it uses up its hydrogen fuel. A 25-solar-mass star will exhaust its hydrogen and die in only about 7 million years, but the sun is expected to last for 10 billion years.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What evidence can you cite that the spaces between the stars are not empty?
2. What evidence can you cite that the interstellar medium contains both gas and dust?
3. Why would an emission nebula near a hot star look red, while a reflection nebula near its star looks blue?
4. Why do astronomers rely heavily on infrared observations to study star formation?
5. What observational evidence can you cite that star formation is a continuous process?
6. How are Herbig–Haro objects related to star formation?
7. What evidence can you cite that star formation is happening right now in the Orion Nebula?
8. How do the proton–proton chain and the CNO cycle resemble each other? How do they differ?
9. Why does the CNO cycle require a higher temperature than the proton–proton chain?
10. How does the pressure–temperature thermostat control the nuclear reactions inside stars?
11. Step-by-step, explain how energy flows from the sun's core to Earth.
12. Why is there a mass–luminosity relation?
13. Why is there a lower limit to the mass of a main-sequence star?
14. Why does a star's life expectancy depend on its mass?
15. **How Do We Know?** Why might you say that scientists who ignore inconvenient evidence are fooling themselves?
16. **How Do We Know?** How can mathematical models help you understand natural processes that occur too fast to observe?

Discussion Questions

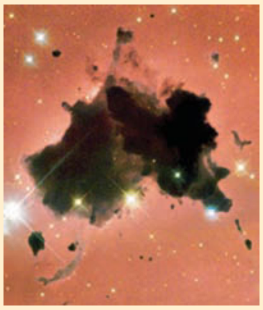
1. When you see distant streetlights through smog, they look dimmer and redder than they do normally. But when you see the same streetlights through fog or falling snow, they look dimmer but not redder. Use your knowledge of the interstellar medium to discuss the relative sizes of the particles in smog, fog, and snow compared with the wavelength of light.
2. If planets form in disks around protostars as a natural by-product of star formation, which do you think are more common — stars or planets?

Problems

1. The interstellar medium dims starlight by about 1.9 magnitudes/1000 pc. What fraction of photons survives a trip of 1000 pc? (*Hint:* See Reasoning with Numbers 2-1.)
2. A small Bok globule has a diameter of 20 arc seconds. If the nebula is 1000 pc from Earth, what is the diameter of the globule?
3. If a giant molecular cloud has a diameter of 30 pc and drifts relative to neighboring clouds at 20 km/s, how long will it take to travel its own diameter?
4. If the dust cocoon around a protostar emits radiation most strongly at a wavelength of 30 microns, what is the temperature of the dust? (*Hint:* See Reasoning with Numbers 6-1.)
5. The gas in a bipolar flow can travel as fast as 300 km/s. How long would it take to travel 1 light-year?
6. Circle all ^{1}H and ^{4}He nuclei in Figure 9-9. Explain how both the proton–proton chain and the CNO cycle can be summarized as $4\ ^1\text{H} \rightarrow\ ^4\text{He} +$ energy.
7. In the model shown in Figure 9-13, how much of the sun's mass is hotter than 13,000,000 K?
8. If a brown dwarf has a surface temperature of 1500 K, at what wavelength will it emit the most radiation? (*Hint:* See Reasoning with Numbers 6-1.)
9. What is the life expectancy of a 16-solar-mass star?
10. If the O6 V star in the Orion Nebula is magnitude 5.4, how far away is the nebula? (*Hint:* Use spectroscopic parallax.)
11. The hottest star in the Orion Nebula has a surface temperature of 40,000 K. At what wavelength does it radiate the most energy? (*Hint:* See Reasoning with Numbers 6-1.)

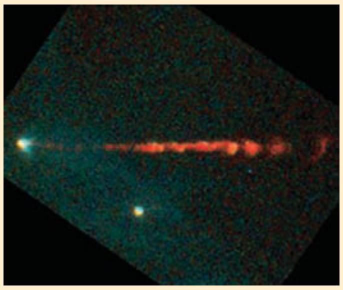
Learning to Look

1. The image at right shows two nebulae, one pink in the background and one black in the foreground. What kind of nebulae are these, and how are they related to star formation?



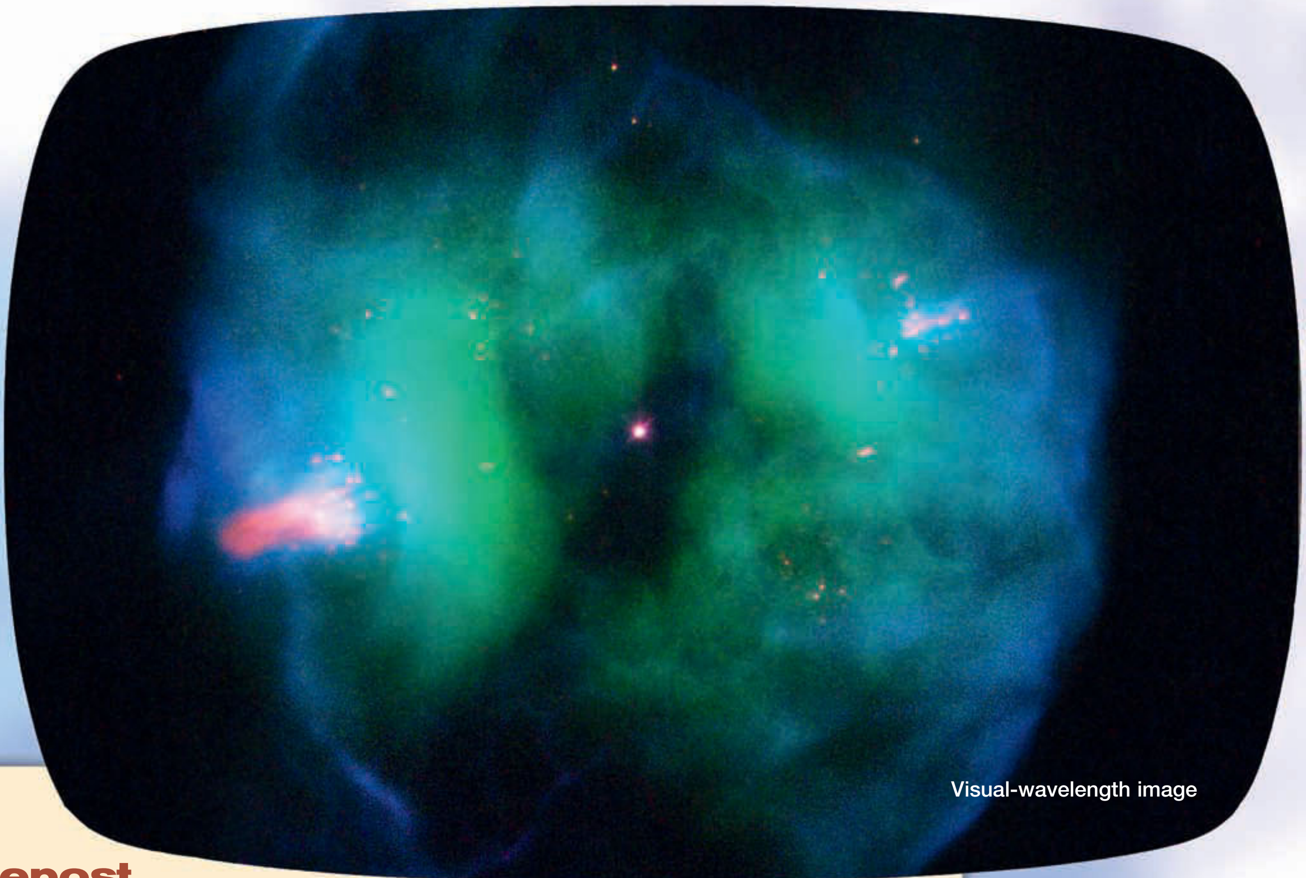
NASA and the Hubble Heritage Team, STScI/AURA

2. In Figure 9-5, a dark globule of dusty gas is located at top right. What do you think that globule would look like if you could see it from the other side?
3. In what ways are the nebulae in Figure 9-3 similar to the Orion Nebula?
4. The star at right appears to be ejecting a jet of gas. What is happening to this star?



STScI and NASA

10 The Deaths of Stars



A dying star has ejected gas to form the nebula NGC 2371 and then collapsed into a small object with a surface temperature over 20 times hotter than the sun. It will eventually become a white dwarf. *(NASA, ESA, and the Hubble Heritage Team)*

Guidepost

The preceding chapter described how stars are born and how they resist their own gravity by fusing nuclear fuels in the center and keeping their central pressure high. But they can last only as long as their fuel.

Now you are ready to consider the fate of the stars. Here you will find the answers to five essential questions:

What happens to a star when it uses up the last of the hydrogen in its core?

What evidence shows that stars really evolve?

How will the sun die?

What happens if an evolving star is in a binary star system?

How do massive stars die?

The deaths of stars are important because life on Earth depends on the sun but also because the deaths of massive stars create the atomic elements of which you are made. If stars didn't die, you would not exist.

In the chapters that follow, you will discover that some of the matter that was once stars becomes trapped in dead ends—white dwarfs, neutron stars, and black holes. But some matter from dying stars escapes back into the interstellar medium and is incorporated into new stars and the planets that circle them. The deaths of stars are part of a great cycle of stellar birth and death that includes your sun, your planet, and you.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Natural laws have no pity.

ROBERT HEINLEIN
THE NOTEBOOKS OF LAZARUS LONG

GRAVITY IS PATIENT—so patient it can kill stars. Stars generate tremendous energy resisting their own gravity, but no star has an infinite supply of fuel for its nuclear reactions. When the fuel runs out, the star dies.

All over the sky, astronomers find beautiful nebulae that were puffed gently into space by dying stars. In addition, astronomers occasionally see a new star appear in the sky, grow brighter, then fade away after a few weeks or a year. You will discover that what looks like a new star in the sky, is either a **nova,** the eruption of a very old dying star, or a **supernova,** the violent explosive death of an aging star. Modern astronomers find a few novae (plural of *nova*) each year, but supernovae (plural) are so rare that there are only one or two each century in our galaxy. Astronomers know that stars die because they occasionally see supernovae flare in other galaxies and because telescopes reveal the remains of stars that have already died (■ Figure 10-1).

The mass of a star is critical in determining its fate. Massive stars can die in violent supernova explosions, but lower-mass stars die quiet deaths. To follow the evolution of stars to their graves, you can start by following the life story of a sunlike, medium-mass star as it becomes a giant star. Then you can see how stars of different masses end their lives.

10-1 Giant Stars

A MAIN-SEQUENCE STAR generates its energy by nuclear fusion reactions that combine hydrogen to make helium. The period during which the star fuses hydrogen lasts a long time, and the star remains on the main sequence for 90 percent of its total existence as an energy-generating star. When the hydrogen is exhausted, however, the star begins to evolve rapidly.

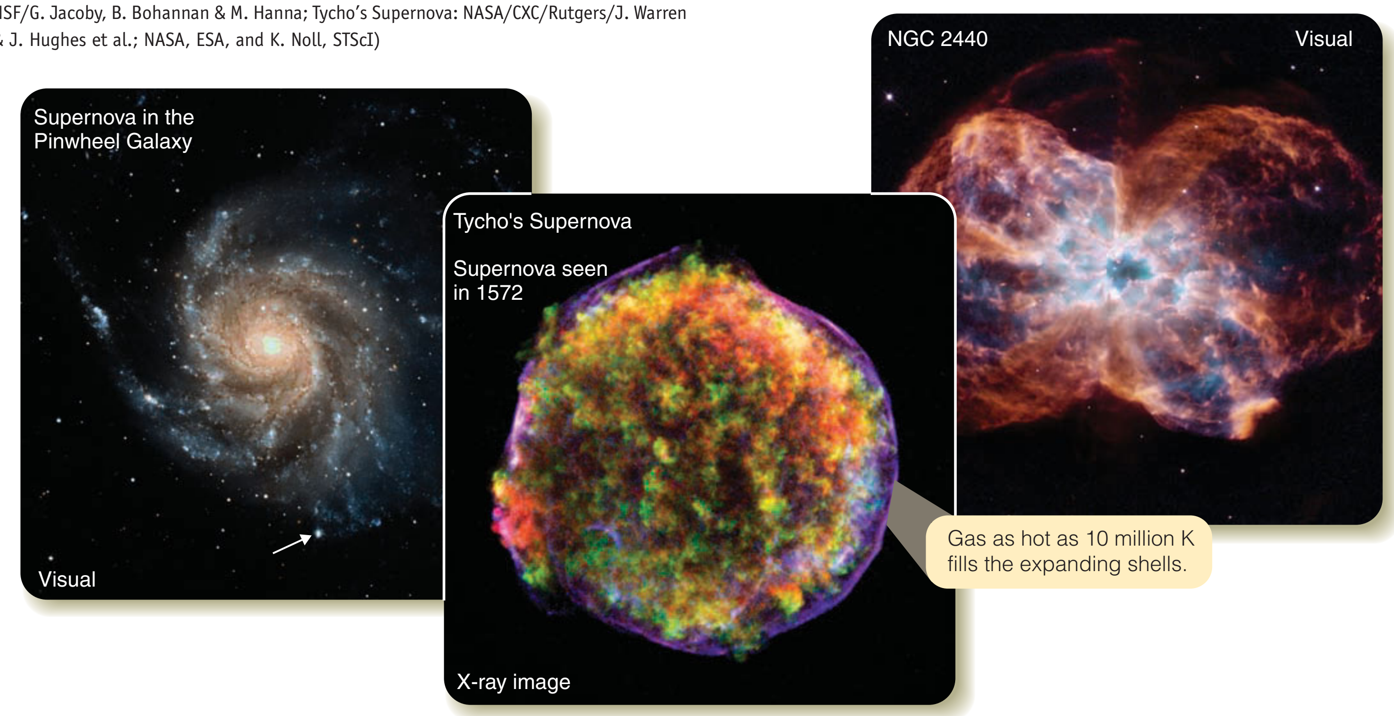
Expansion into a Giant

The nuclear reactions in a main-sequence star's core fuse hydrogen to produce helium. Because the core is cooler than 100,000,000 K, the helium cannot fuse in nuclear reactions, so it accumulates at the star's center like ashes in a fireplace. Initially, this helium ash has little effect on the star, but as hydrogen is exhausted and the stellar core becomes almost pure helium, the star loses the ability to generate the nuclear energy that opposes gravity. As soon as the energy generation starts to die down, gravity begins making the core contract.

Although the core of helium ash cannot generate nuclear energy, it does grow hotter as it contracts because it is converting gravitational energy into thermal energy (see previous chapter). The rising temperature heats the unprocessed hydrogen just outside the core, hydrogen that was never before hot enough to fuse. Soon, hydrogen fusion begins in a spherical layer or shell around

■ **Figure 10-1**

Evidence that stars die: Supernova explosions are rare in any one galaxy, but each year astronomers see a few erupt in other galaxies. In our own galaxy, astronomers find expanding shells filled with hot, low-density gas produced by past supernova explosions of massive stars. In contrast, NGC 2440 is the remains of a lower-mass star that was much like the sun. (Pinwheel: NOAO/AURA/NSF/G. Jacoby, B. Bohannan & M. Hanna; Tycho's Supernova: NASA/CXC/Rutgers/J. Warren & J. Hughes et al.; NASA, ESA, and K. Noll, STScI)



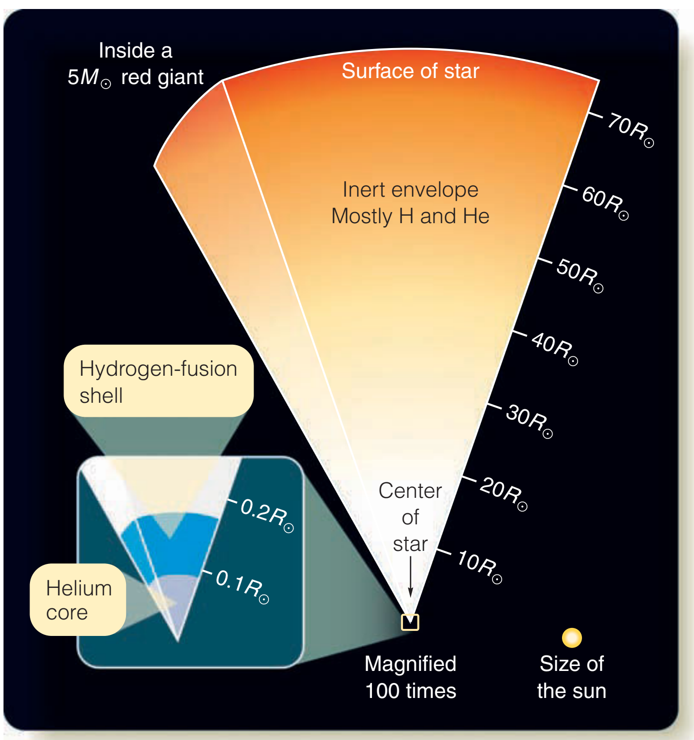
the exhausted core of the star. Like a grass fire burning outward from an exhausted campfire, the hydrogen-fusion shell creeps outward, leaving helium ash behind and increasing the mass of the helium core.

The flood of energy produced by the hydrogen-fusion shell pushes toward the surface, heating the outer layers of the star and forcing them to expand dramatically (■ Figure 10-2). Stars like the sun become giant stars of 10 to 100 solar radii, and the most massive stars become supergiants some 1000 times larger than the sun. This explains the large diameters and low densities of the giant and supergiant stars. In Chapter 8, you learned about the large sizes and low densities of giant and supergiant stars. Now you understand that these stars were once normal main-sequence stars that expanded when hydrogen shell fusion began.

The expansion of its envelope dramatically changes a star's location in the H–R diagram. Just as contraction heats a star, expansion cools it. As the outer layers of gas expand, energy is absorbed in lifting and expanding the gas. The loss of that energy lowers the temperature of the gas. Consequently the point that represents the star in the H–R diagram moves to the right relatively quickly (in less than a million years for a star of 5 solar masses). A massive star moves to the right across the top of the H–R diagram and becomes a supergiant, while a medium-mass star like the sun becomes a red giant (■ Figure 10-3). As the radius of a giant star continues to increase, its enlarging surface area makes the star more luminous, moving its point upward in the H–R diagram. Favorite Star Aldebaran, the glowing red eye of Taurus the Bull, is such a red giant, with a diameter 25 times that of the sun but a much cooler surface temperature.

Degenerate Matter

Although the hydrogen-fusion shell can force the envelope of the star to expand, it cannot stop the contraction of the helium core. Because the core is not hot enough to fuse helium, gravity squeezes it tighter, and it becomes very small. If you were to represent the helium core of a giant star with a baseball, the outer envelope of the star would be about the size of a baseball stadium. Yet the core would contain about 12 percent of the star's mass. When gas is compressed to such extreme densities, it be-

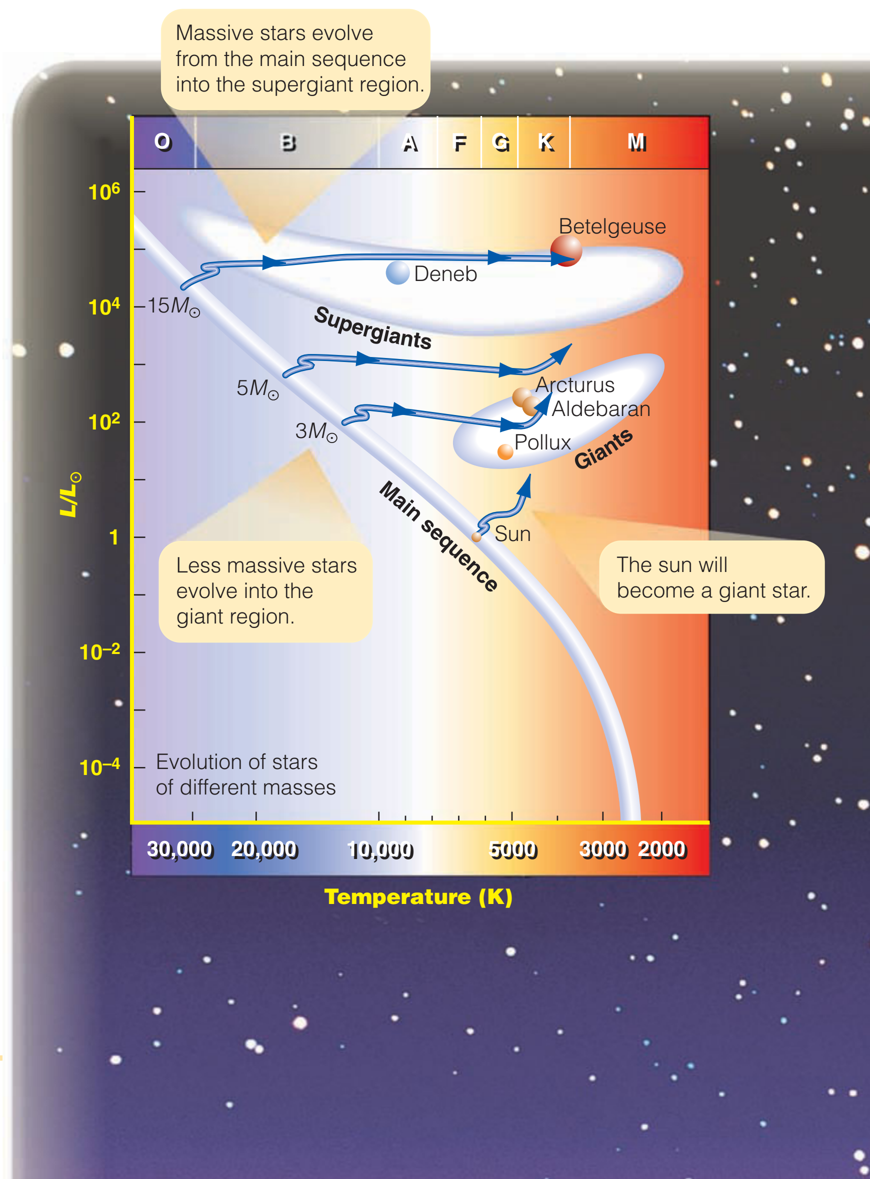


■ **Figure 10-2**

When a star runs out of hydrogen at its center, the core contracts to a small size, becomes very hot, and begins nuclear fusion in a shell (blue). The outer layers of the star expand and cool. The red giant star shown here has an average density much lower than the air at Earth's surface. Here $M_\odot$ stands for the mass of the sun, and $R_\odot$ stands for the radius of the sun. (Illustration design by author)

■ **Figure 10-3**

The evolution of a massive star moves the point that represents it in the H–R diagram to the right of the main sequence into the region of the supergiants such as Deneb and Betelgeuse. The evolution of medium-mass stars moves their points in the H–R diagram into the region occupied by giants such as those shown here.



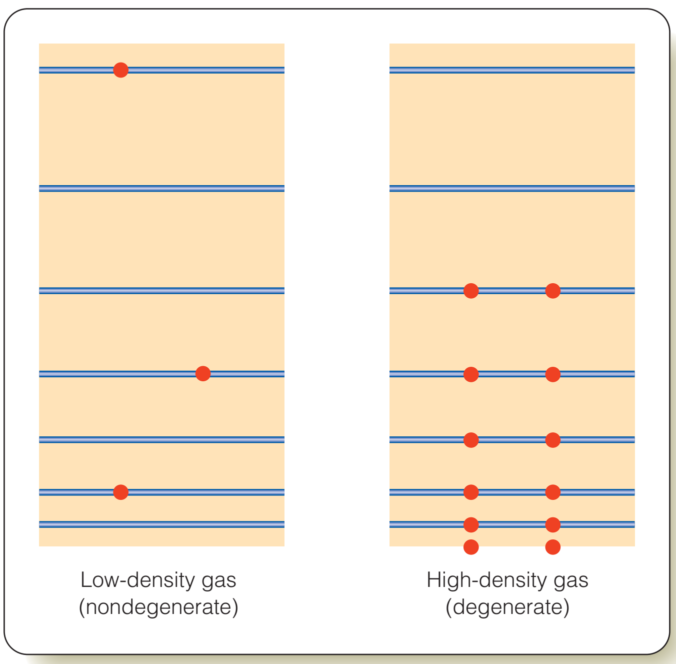
gins to behave in astonishing ways that can affect the evolution of a star. To continue the story of stellar evolution, you must consider the behavior of gas at extremely high densities.

Normally, the pressure in a gas depends on its temperature. The hotter a gas is, the faster its particles move, and the more pressure it exerts. The gas inside a star is ionized, so there are two kinds of particles, atomic nuclei and free electrons. Under normal conditions the gas in a star follows the same pressure/temperature laws as other gases, but if the gas is compressed to very high densities, as in the core of a giant star, two laws of quantum mechanics come into play, and the difference between electrons and nuclei becomes important.

First, quantum mechanics says that the moving electrons confined in the star's core can have only certain amounts of energy, just as the electron in an atom can occupy only certain energy levels (see Chapter 6). You can think of these permitted energies as the rungs of a ladder. An electron can occupy any rung but not the spaces between.

The second quantum mechanical law (called the Pauli exclusion principle) says that two identical electrons cannot occupy the same energy level. Because electrons spin in one direction or the other, two electrons can occupy a single energy level if they spin in opposite directions. That level is then completely filled, and a third electron cannot enter because, whichever way it spins, it will be identical to one or the other of the two electrons already in the level.

A low-density gas has few electrons per cubic centimeter, so there are plenty of energy levels available (■ Figure 10-4). If a gas becomes very dense, however, nearly all of the lower energy levels are occupied. In such a gas, a moving electron cannot slow down; slowing down would decrease its energy, and there are no open energy levels for it to drop down to. It can speed up only if it can absorb enough energy to leap to the top of the energy ladder, where there are empty energy levels.



■ **Figure 10-4**

Electron energy levels are arranged like rungs on a ladder. In a low-density gas many levels are open, but in a degenerate gas all lower-energy levels are filled.

When a gas is so dense that the electrons are not free to change their energy, astronomers call it **degenerate matter.** Although it is a gas, it has two peculiar properties that can affect the star. First, the degenerate gas resists compression. To compress the gas, you must push against the moving electrons, and changing their motion means changing their energy. That requires tremendous effort, because you must boost them to the top of the energy ladder. That is why degenerate matter, though still a gas, is harder to compress than the toughest hardened steel.

Second, the pressure of degenerate gas does not depend on temperature. To see why, note that the pressure depends on the speed of the electrons, which cannot be changed without tremendous effort. The temperature, however, depends on the motion of all the particles in the gas, both electrons and nuclei. If you add heat to the gas, most of that energy goes to speed up the motions of the nuclei, which move slowly and don't contribute much to the pressure. Only a few electrons can absorb enough energy to reach the empty energy levels at the top of the energy ladder. That means that changing the temperature of the gas has almost no effect on the pressure.

These two properties of degenerate matter become important when stars end their main-sequence lives (■ How Do We Know? 10-1). Eventually, many stars collapse into white dwarfs, and you will discover that these tiny stars are made of degenerate matter. But long before that, the cores of many giant stars become so dense that they are degenerate, a situation that can produce a cosmic bomb.

Helium Fusion

Hydrogen fusion in main-sequence stars leaves behind helium ash, which cannot fuse because the temperature is too low. Helium nuclei have a positive charge twice that of a proton, so, to overcome the repulsion between nuclei, they must collide at a high velocity; but the temperature in the core isn't high enough to produce those collisions. As a giant star fuses hydrogen in an expanding shell, its inert core of helium contracts and grows hotter. When the temperature of the core finally reaches 100,000,000 K, it begins to fuse helium nuclei to make carbon (see previous chapter).

How a star begins helium fusion depends on its mass. Stars more massive than about 3 solar masses contract rapidly, their helium-rich cores heat up, and helium fusion begins gradually. But less-massive stars evolve more slowly, and their cores contract so much that the gas becomes degenerate. On Earth, a teaspoon

How Do We Know? 10-1

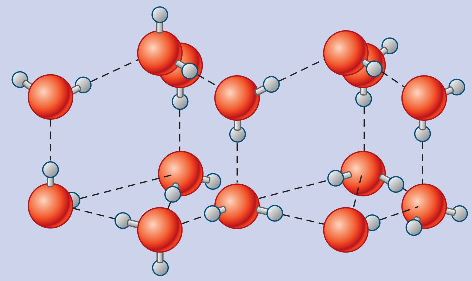
Toward Ultimate Causes

How does a scientist search for natural causes lead into the world of subatomic particles? Scientists search for causes. They are not satisfied to know that a certain kind of star dies by exploding. They want to know why it explodes. They want to find the causes for the natural events they see, and that search for ultimate causes often leads into the atomic world.

For example, why do icebergs float? When water freezes, it becomes less dense than liquid water, so it floats. That answers the question, but you can search for a deeper cause.

Why is ice less dense than water? Water molecules are made up of two hydrogen atoms bonded to an oxygen atom, and the oxygen is so good at attracting electrons, the hydrogen atoms are left needing a bit more negative charge. They are attracted to atoms in nearby molecules. That means the hydrogen atoms in water are constantly trying to stick to other molecules. When water is warm, the thermal motion prevents these hydrogen bonds from forming; but when water freezes, the hydrogen atoms link the water molecules together. Because of the angles at which the bonds form, the molecules leave open spaces between molecules, and that makes ice less dense than water.

But scientists can continue their search for causes. Why do electrons have negative charge? What is charge? Nuclear particle physicists are tying to understand those properties of matter. Sometimes the properties of very large things such as supernovae are determined by the properties of the tiniest particles. Science is exciting because the simple observation that ice floats in your lemonade can lead you toward ultimate causes and some of the deepest questions about how nature works.



Ice has a low density and floats because of the way electrons (blue) link to oxygen (red) when water freezes.

of the gas would weigh more than an automobile. In this degenerate matter, the pressure does not depend on temperature, and that means the pressure–temperature thermostat does not regulate energy production. When the temperature becomes hot enough, helium fusion begins to make energy, and the temperature rises, but pressure does not increase because the gas is degenerate. The higher temperature increases the helium fusion even further, and the result is a runaway explosion called the **helium flash** in which, for a few minutes, the core of a star can generate more energy per second than does an entire galaxy.

Although the helium flash is sudden and powerful, it does not destroy the star. In fact, if you were observing a giant star as it experienced the helium flash, you would probably see no outward evidence of the eruption. The helium core is quite small (Figure 10-2), and all of the energy of the explosion is absorbed by the distended envelope. In addition, the helium flash is a very short-lived event in the life of a star. In a matter of minutes to hours, the core of the star becomes so hot that a significant number of electrons get boosted to empty energy levels at the top of the energy ladder. That increases the pressure and ends the degenerate conditions, and the pressure–temperature thermostat brings the helium fusion under control. From that point on, the star proceeds to fuse helium steadily in its core.

There are two reasons why you should know about the helium flash. First, it is so violent and so sudden, it makes it difficult to compute models of stars and be sure how they evolve. Astronomers have to exercise ingenuity to get past the helium flash and follow the further evolution of stars. Second, the helium flash is a good illustration of how science reveals a hidden universe. Astronomers would never have known about the helium flash were it not for the theoretical calculation of stellar models.

The sun will experience a helium flash in a few billion years, but stars less massive than about 0.4 solar mass never get hot enough to ignite helium. Stars more massive than 3 solar masses ignite helium before their contracting cores become degenerate.

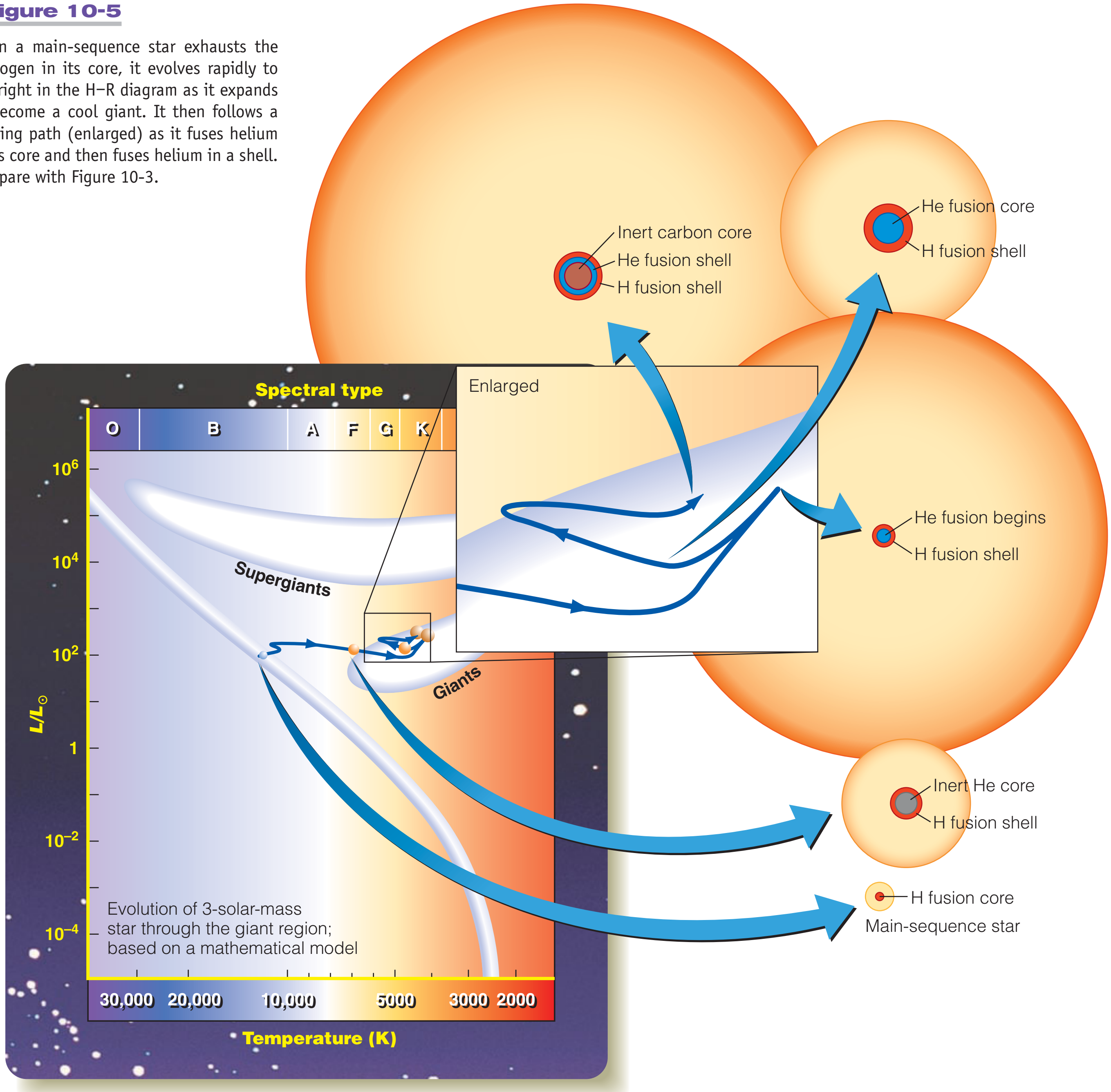
Whether a star experiences a helium flash or not, the ignition of helium in the core changes the structure of the star. The star now makes energy both in its helium-fusion core and in its hydrogen fusion shell. The energy flowing outward from the core can halt the contraction of the core, and the distended envelope of the star contracts and grows hotter. Consequently, the point that represents the star in the H–R diagram moves downward and to the left toward the hot side of the H–R diagram (■ Figure 10-5).

Helium fusion produces carbon, and some of the carbon nuclei absorb helium nuclei to form oxygen. A few of the oxygen nuclei can absorb helium nuclei and form neon and then magnesium. Some of these reactions release neutrons, which, having no charge, are more easily absorbed by nuclei to gradually build even heavier nuclei. These reactions are not important as energy producers, but they are slow-cooker processes that form small traces of heavier elements right up to bismuth, nearly four times heavier than iron. Many of the atoms in your body were produced this way.

As the helium fuel is used up, the accumulation of carbon and oxygen atoms creates an inert core too cool to fuse. Once again,

Figure 10-5

When a main-sequence star exhausts the hydrogen in its core, it evolves rapidly to the right in the H–R diagram as it expands to become a cool giant. It then follows a looping path (enlarged) as it fuses helium in its core and then fuses helium in a shell. Compare with Figure 10-3.



the core contracts and heats up, and soon a helium-fusion shell ignites below the hydrogen-fusion shell. Now that the star makes energy in two fusion shells, it quickly expands, and its surface cools once again. The point that represents the star in the H–R diagram moves back to the right, completing a loop (Figure 10-5).

What happens to a star after helium fusion depends on its mass, but no matter what tricks the star plays to delay its end, it cannot survive long. It must eventually collapse and end its career as a star. The remainder of this chapter will trace the details of this process of stellar death, but before you begin that story, you must ask the most important question in science: What is the evidence for what you have learned so far? What evidence shows that stars actually evolve as theories predict? You will find the answers in clusters of stars.

Star Clusters: Evidence of Evolution

Just as Sherlock Holmes studies peculiar dust on a lamp shade as evidence that will solve a mystery, astronomers look at star clusters and say, “Aha!” A photo of a star cluster freezes a moment in the evolution of the cluster and makes the evolution of the stars visible to human observers.

Because they formed nearly simultaneously from the same gas cloud, the stars in a cluster have about the same age and composition; so any differences you see among them are due to their difference in mass. That means that when you look at a cluster, you can see the effects of stellar evolution as it acts on otherwise similar stars of different mass. Study ■ **Star Cluster H–R Diagrams** on pages 192–193 and notice three important points and four new terms:

1. There are two kinds of star clusters, *open clusters* and *globular clusters.* They look different, but they are similar in the way their stars evolve. You will learn even more about these clusters in a later chapter.
2. You can estimate the age of a star cluster by observing the *turnoff point* in the distribution of the points that represent its stars in the H–R diagram.
3. Finally, the shape of a star cluster's H–R diagram is governed by the evolutionary path the stars take. The H-R diagrams of older clusters are especially clear in outlining how stars evolve away from the main sequence to the giant region, then move left along the *horizontal branch* before evolving back into the giant region. By comparing clusters of different ages, you can visualize how stars evolve almost as if you were watching a film of a star cluster evolving over billions of years.

If it were not for star clusters, astronomers would have little confidence in the theories of stellar evolution. Star clusters make evolution visible and assure astronomers that they really do understand how stars are born, live, and die.

◄ SCIENTIFIC ARGUMENT ►

Why is it only lower-mass stars that outline the horizontal branch?

This argument depends on timing. If a star cluster is young, it may contain a few massive stars, but, because massive stars are so rare and evolve so rapidly, you are unlikely to see more than a few of these stars evolving through the giant or supergiant regions of the H–R diagram. Lower-mass stars are very common and evolve slowly, so in an older star cluster, you can see lots of stars in various stages of the post-main-sequence evolution. That outlines the horizontal branch.

Now construct a different argument. **What evidence can you cite that giant stars are main-sequence stars that have expanded to large diameters?**

◄ ►

10-2 The Deaths of Lower-Main-Sequence Stars

Contracting stars heat up by converting gravitational energy into thermal energy. Low-mass stars have little gravitational energy, so when they contract, they don't get very hot. This limits the fuels they can ignite. In the previous chapter, you saw that protostars less massive than 0.08 solar mass cannot get hot enough to ignite hydrogen. This section will concentrate on stars more massive than 0.08 solar mass but no more than a few times the mass of the sun.

Structural differences divide the lower-main-sequence stars into two subgroups—very-low-mass red dwarfs and medium-mass stars such as the sun. The critical difference between the two groups is the extent of interior convection. If the star is convective, fuel is constantly mixed, and its resulting evolution is drastically altered.

Red Dwarfs

Stars between 0.08 and about 0.4 solar mass—the red dwarfs—have two advantages over more massive stars. First, they have very small masses, and that means they have very little weight to support. Their pressure–temperature thermostats are set low, and they consume their hydrogen fuel very slowly. The discussion of the life expectancies of stars in the previous chapter concluded that the red dwarfs should live very long lives.

The red dwarfs have a second advantage in that they are totally convective. That is, they are stirred by circulating currents of hot gas rising from the interior and cool gas sinking inward. This means the stars are mixed like a pot of soup that is constantly stirred as it cooks. Hydrogen is consumed uniformly throughout the star, which means the star is not limited to the fuel in its core. It can use all of its hydrogen to prolong its life on the main sequence.

Because a red dwarf is mixed by convection, it cannot develop an inert helium core surrounded by unprocessed hydrogen. This is why it can never ignite a hydrogen shell and cannot become a giant star. What astronomers know about stellar evolution indicates that these red dwarfs should use up nearly all of their hydrogen and live very long lives on the lower main sequence, surviving for a hundred billion years or more. Of course, astronomers can't test this part of their theories because the universe is only 13.7 billion years old, so not a single red dwarf has died of old age anywhere in the universe. Every red dwarf that has ever been born is still shining today.

Medium-Mass Stars

Stars like the sun eventually become hot enough to ignite helium as they pass through their giant phase, but, if they contain less than 4 solar masses,* they do not get hot enough to ignite carbon, the next fuel after helium. When they reach that impasse, they collapse and become white dwarfs. There are two keys to the evolution of these sunlike stars, the lack of complete mixing and mass loss.

The interiors of medium-mass stars are not completely mixed (■ Figure 10-6). Stars of 1.1 solar masses or less have no convection near their centers, so they are not mixed at all. Stars

*This mass limit is uncertain, as are many of the masses quoted here. The evolution of stars is highly complex, and such parameters are difficult to specify.

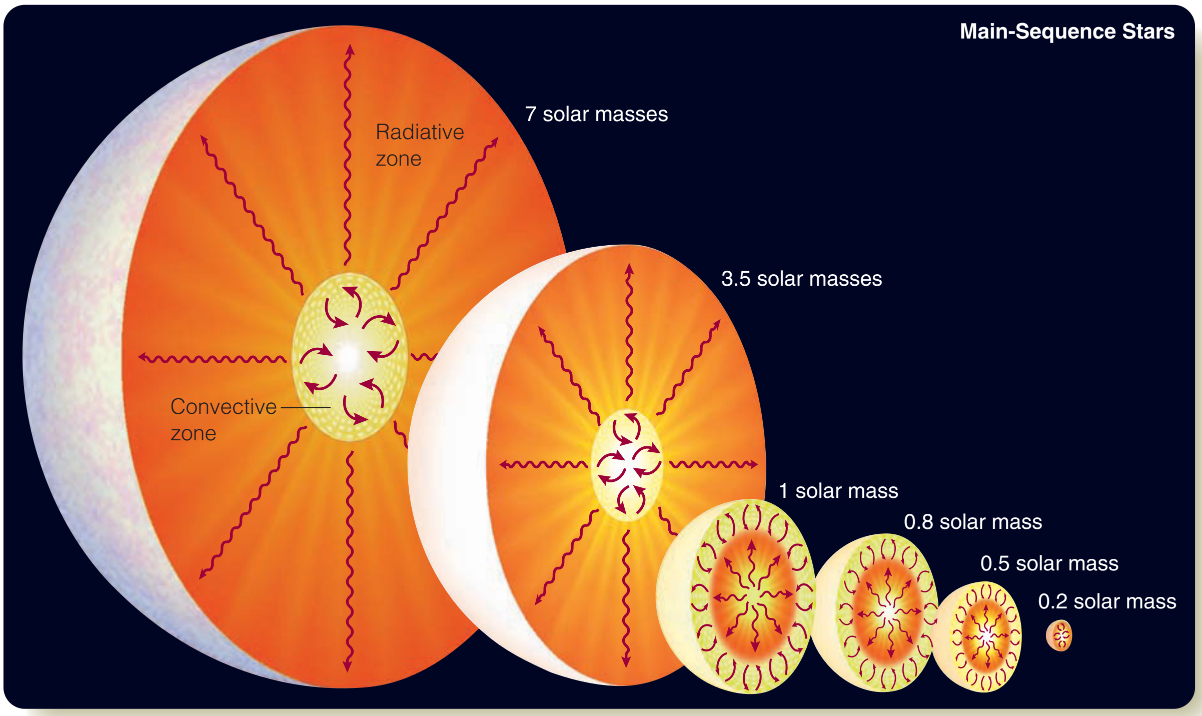


Figure 10-6

Inside main-sequence stars. The more-massive stars have small convective interiors and radiative envelopes. Stars like the sun have radiative interiors and convective envelopes. The lowest-mass stars are convective throughout. The "cores" of the stars where nuclear fusion occurs (not shown) are smaller than the interiors. (Illustration design by author)

with a mass greater than 1.1 solar masses have small zones of convection at their centers, but this mixes no more than about 12 percent of the star's mass. Medium-mass stars, whether they have convective cores or not, are not thoroughly mixed, and the helium ash accumulates in an inert helium core surrounded by unprocessed hydrogen. Recall from earlier in this chapter that when this core contracts, the unprocessed hydrogen ignites in a shell and swells the star into a giant.

In the giant stage, the core of the star contracts, and the envelope expands. The star fuses helium first in its core and then in an expanding shell surrounding a core of carbon and oxygen. This core contracts and grows hotter; but, because the star has too low a mass, the core cannot get hot enough to ignite carbon fusion. The carbon–oxygen core is a dead end for these medium-mass stars.

All of this discussion is based on theoretical models of stars and a general understanding of how stars evolve. Does it really happen? Astronomers need observational evidence to confirm their theories, and the gas that is expelled from these giant stars gives visible evidence that sunlike stars do indeed die in this way.

Planetary Nebulae

When a medium-mass star like the sun becomes a distended giant, its atmosphere cools. As it cools, it becomes more opaque, and light has to push against it to escape. At the same time, the fusion shells become so thin they are unstable and begin to flare, which also pushes the atmosphere outward. Because of this outward pressure, an aging giant can expel its outer atmosphere in repeated surges to form one of the most interesting objects in astronomy, a **planetary nebula,** so called because through a small telescope it looks like the greenish-blue disk of a planet like Uranus or Neptune. In fact, a planetary nebula has nothing to do with a planet. It is composed of ionized gases expelled by a dying star.

Study **The Formation of Planetary Nebulae** on pages 194–195 and notice four things:

1. You can understand what planetary nebulae are like by using simple observational principles such as Kirchhoff's laws and the Doppler effect.
2. Notice the model that astronomers have developed to explain planetary nebulae. The real nebulae are more complex than the simple model of a slow wind and a fast wind, but the model provides a way to organize the observed phenomena.
3. Oppositely directed jets (much like bipolar flows from protostars) produce many of the asymmetries seen in planetary nebulae.
4. The star itself must finally contract into a white dwarf.

Most astronomy books say that the sun will form a planetary nebula, but that may not happen. To ionize the gas and light up

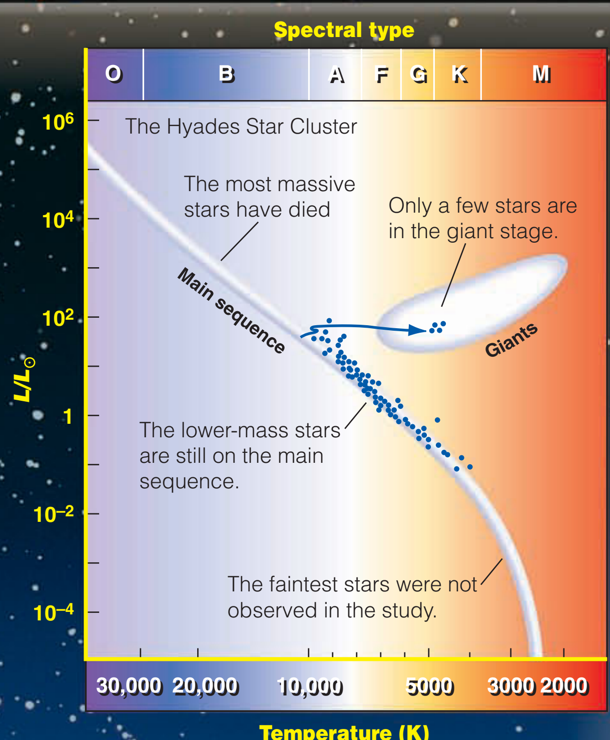
Star Cluster H–R Diagrams

1 An **open cluster** is a collection of 10 to 1000 stars in a region about 25 pc in diameter. Some open clusters are quite small and some are large, but they all have an open, transparent appearance because the stars are not crowded together.

Open Cluster
The Jewel Box

1a A **globular cluster** can contain 10^5 to 10^6 stars in a region only 10 to 30 pc in diameter. The term “globular cluster” comes from the word “globe,” although globular cluster is pronounced like “glob of butter.” These clusters are nearly spherical, and the stars are much closer together than the stars in an open cluster.

Astronomers can construct an H–R diagram for a star cluster by plotting a point to represent the luminosity and temperature of each star.

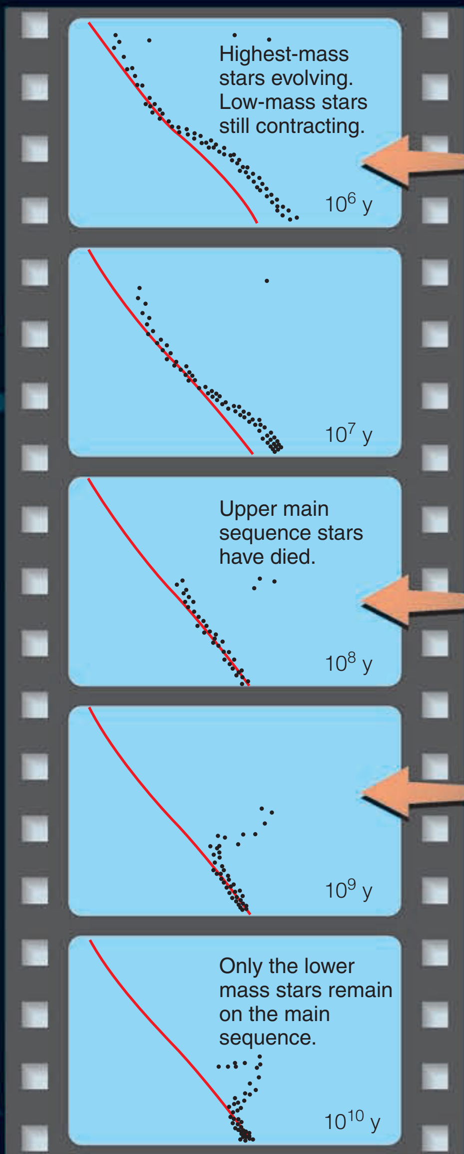


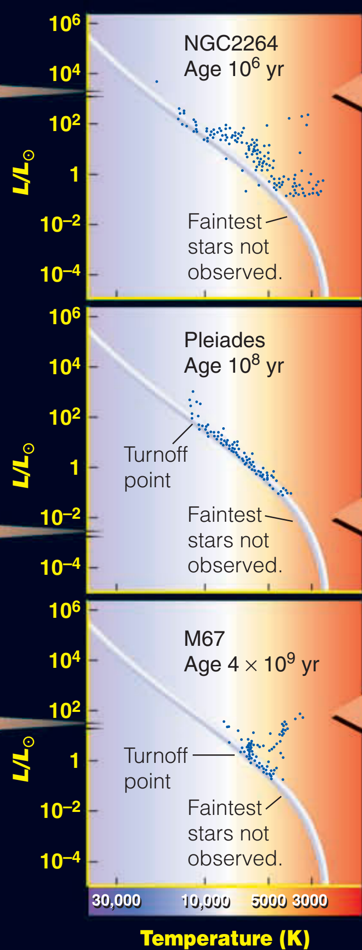
2 The H–R diagram of a star cluster can make the evolution of stars visible. The key is to remember that all of the stars in the star cluster have the same age but differ in mass. The H–R diagram of a star cluster provides a snapshot of the evolutionary state of the stars at the time you happen to be alive. The diagram here shows the 650-million-year-old star cluster called the Hyades. The upper main sequence is missing because the more massive stars have died, and our snapshot catches a few medium-mass stars leaving the main sequence to become giants.

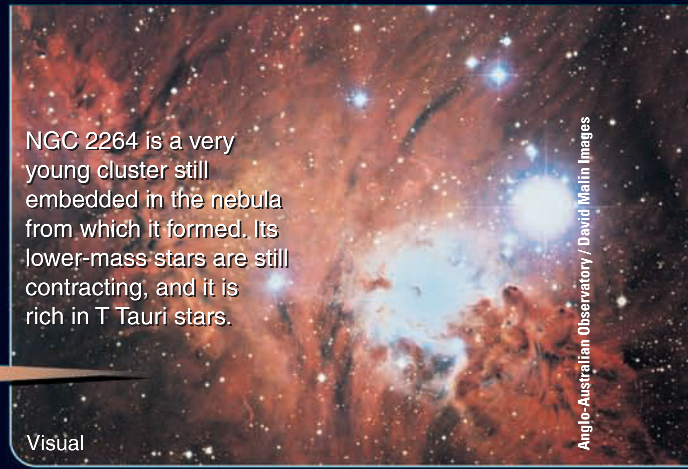
As a star cluster ages, its main sequence grows shorter like a candle burning down. You can judge the age of a star cluster by looking at the **turnoff point**, the point on the main sequence where stars evolve to the right to become giants. Stars at the turnoff point have lived out their lives and are about to die. Consequently, the life expectancy of the stars at the turnoff point equals the age of the cluster.

CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure “Cluster Turnoff” and notice how the shape of a cluster’s H–R diagram changes with time.

3 From theoretical models of stars, you could construct a film to show how the H–R diagram of a star cluster changes as it ages. You can then compare theory (left) with observation (right) to understand how stars evolve. Note that the time step for each frame in this film increases by a factor of 10.

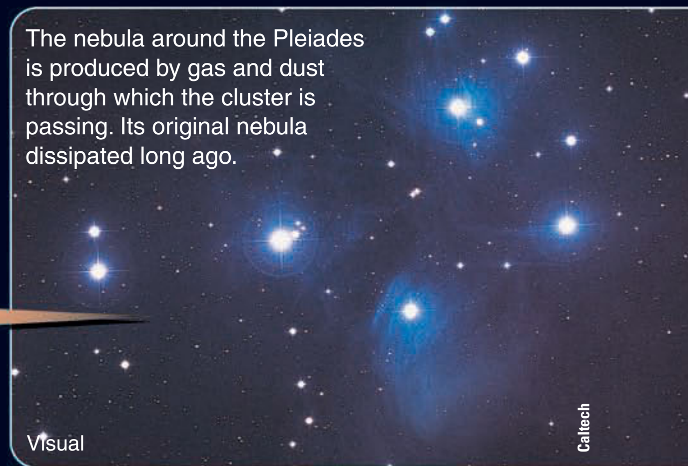






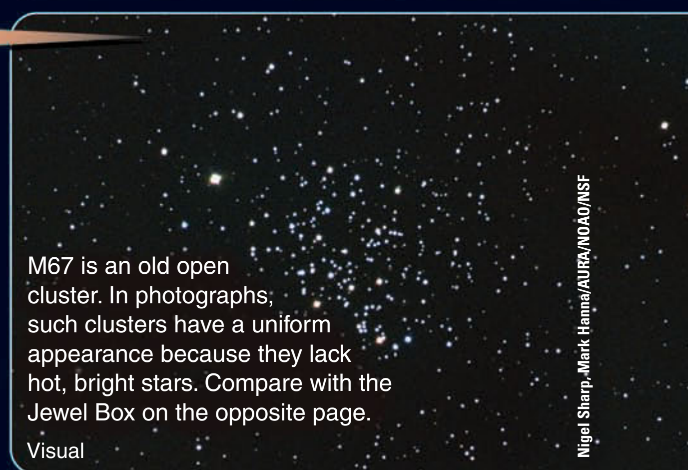
NGC 2264 is a very young cluster still embedded in the nebula from which it formed. Its lower-mass stars are still contracting, and it is rich in T Tauri stars.

Visual



The nebula around the Pleiades is produced by gas and dust through which the cluster is passing. Its original nebula dissipated long ago.

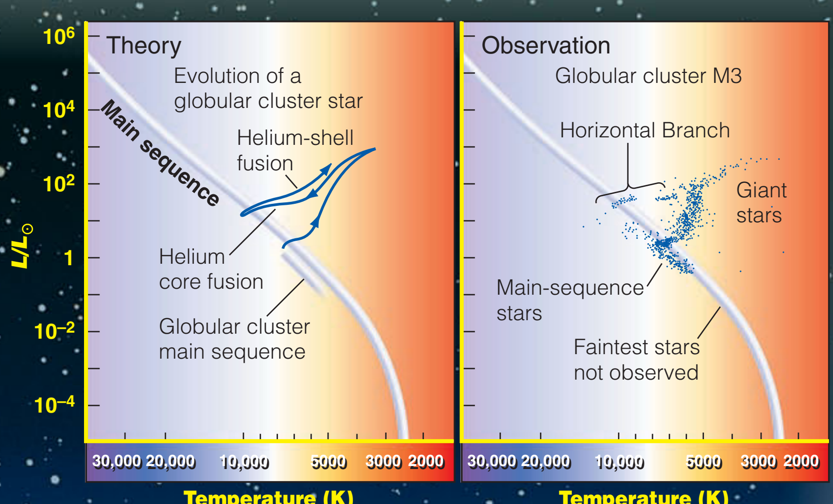
Visual



M67 is an old open cluster. In photographs, such clusters have a uniform appearance because they lack hot, bright stars. Compare with the Jewel Box on the opposite page.

Visual

Globular cluster H–R diagrams resemble the last frame in the film, which tells you that globular clusters are very old.



3a The H–R diagrams of globular clusters have very faint turnoff points showing that they are very old clusters. The best analysis suggests these clusters are about 11 billion years old.

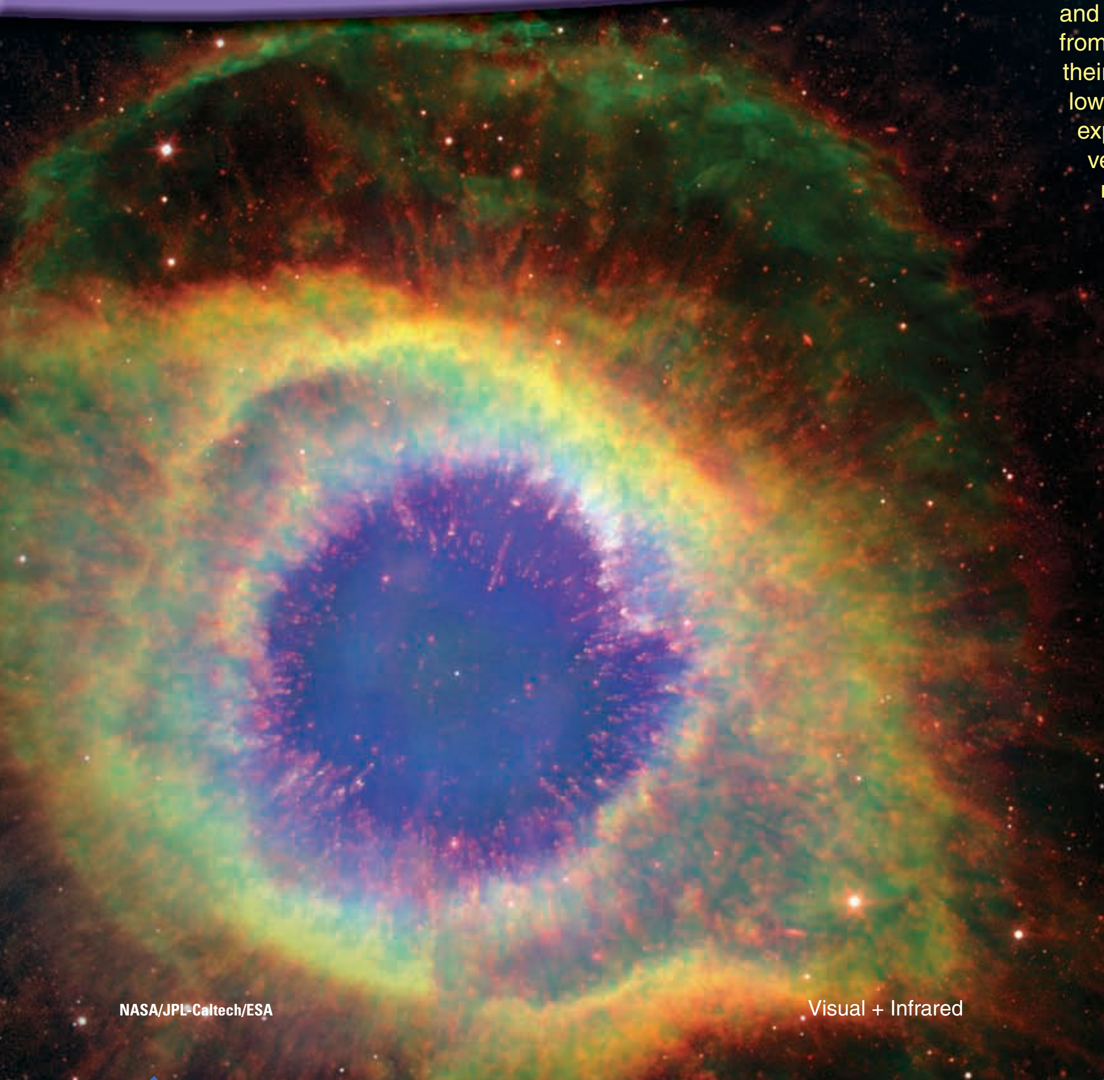
The **horizontal branch** stars are giants fusing helium in their cores and then in shells. The shape of the horizontal branch outlines the evolution of these stars.

The main-sequence stars in globular clusters are fainter and bluer than the zero-age main sequence. Spectra reveal that globular cluster stars are poor in elements heavier than helium, and that means their gases are less opaque. That means energy can flow outward more easily, which makes the stars slightly smaller and hotter. Again the shape of star cluster H–R diagrams illustrates principles of stellar evolution.

The Formation of Planetary Nebulae

1 Simple observations tell astronomers what planetary nebulae are like. Their angular size and their distances indicate that their radii range from 0.2 to 3 ly. The presence of emission lines in their spectra assures that they are excited, low-density gas. Doppler shifts show they are expanding at 10 to 20 km/s. If you divide radius by velocity, you find that planetary nebulae are no more than about 10,000 years old. Older nebulae evidently become mixed into the interstellar medium.

Astronomers find about 1500 planetary nebulae in the sky. Because planetary nebulae are short-lived formations, you can conclude that they must be a common part of stellar evolution. Medium-mass stars up to a mass of about 8 solar masses are destined to die by forming planetary nebulae.

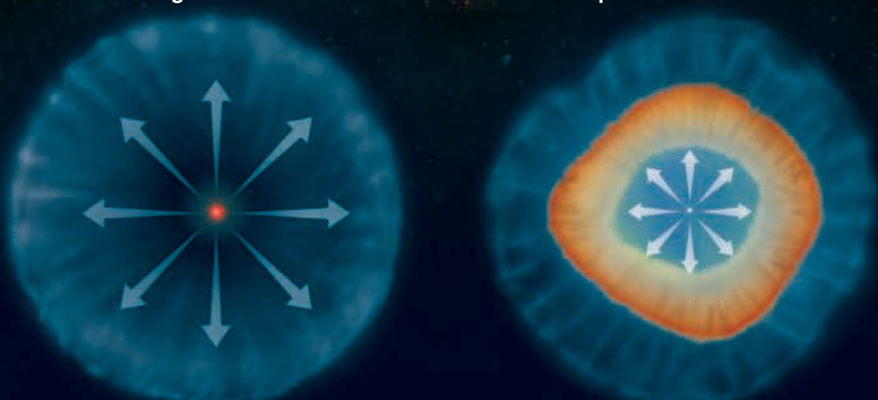


The Helix Nebula is 2.5 ly in diameter, and the radial texture shows how light and winds from the central star are pushing outward.

2 The process that produces planetary nebulae involves two stellar winds. First, as an aging giant, the star gradually blows away its outer layers in a slow breeze of low-excitation gas that is not easily visible. Once the hot interior of the star is exposed, it ejects a high-speed wind that overtakes and compresses the gas of the slow wind like a snowplow, while ultraviolet radiation from the hot remains of the central star excites the gases to glow like a giant neon sign.

Slow stellar wind from a red giant

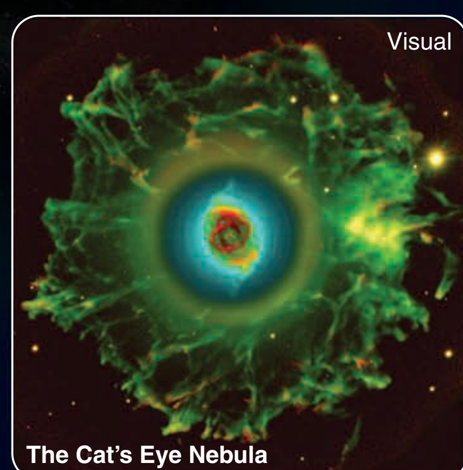
Fast wind from exposed interior



The gases of the slow wind are not easily detectable.

You see a planetary nebula where the fast wind compresses the slow wind.

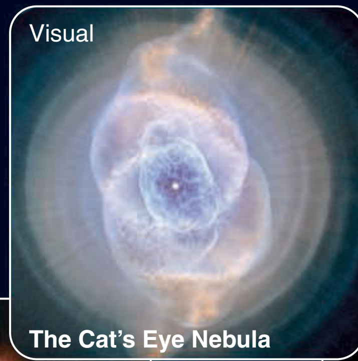
2a The Cat's Eye, below, lies at the center of an extended nebula that must have been exhaled from the star long before the fast wind began forming the visible planetary nebula. See other images of the nebula on opposite page.



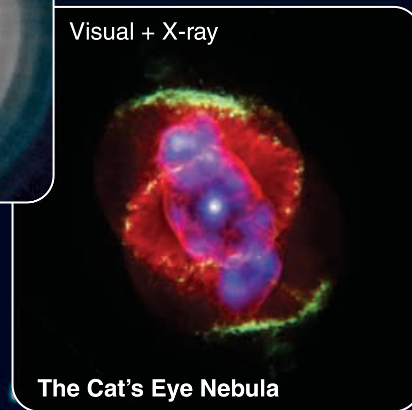
The Cat's Eye Nebula

Roman Corradi/Nordic Optical Telescope

3 Images from the Hubble Space Telescope reveal that asymmetry is the rule in planetary nebulae rather than the exception. A number of causes have been suggested. A disk of gas around a star's equator might form during the slow-wind stage and then deflect the fast wind into oppositely directed flows. Another star or planets orbiting the dying star, rapid rotation, or magnetic fields might cause these peculiar shapes. The Hour Glass Nebula seems to have formed when a fast wind overtook an equatorial disk (white in the image). The nebula Menzel 3, as do many planetary nebulae, shows evidence of multiple ejections.



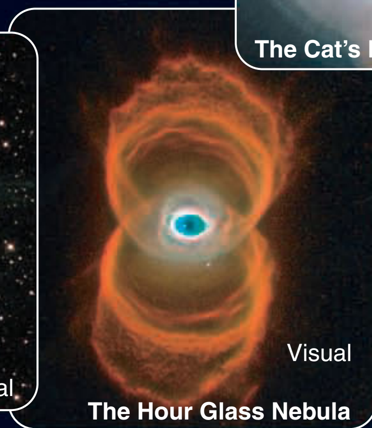
Some shapes suggest bubbles being inflated in the interstellar medium. The Cat's Eye is shown at left, below, and on the facing page.



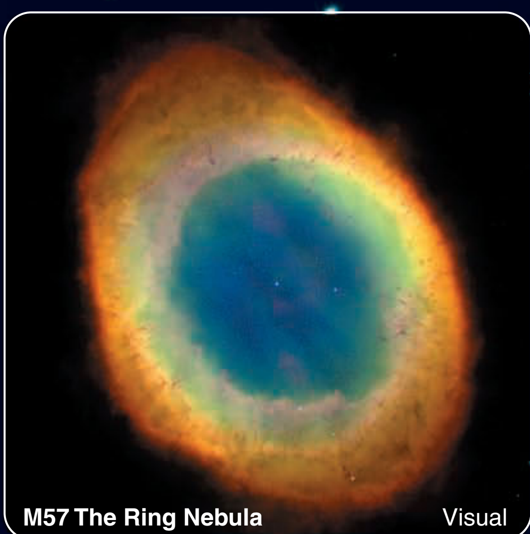
3.3 Visual NASA

The purple glow in the image above is a region of X-ray bright gas with a temperature measured in millions of degrees. It is apparently driving the expansion of the nebula.

NASA



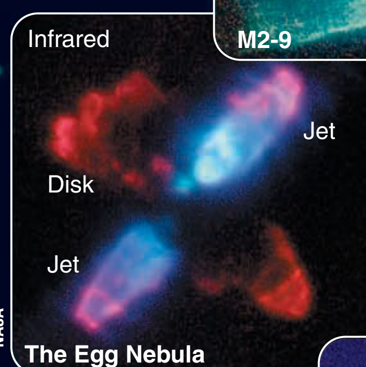
NASA



Hubble Heritage Team, AURA/STScI/NASA

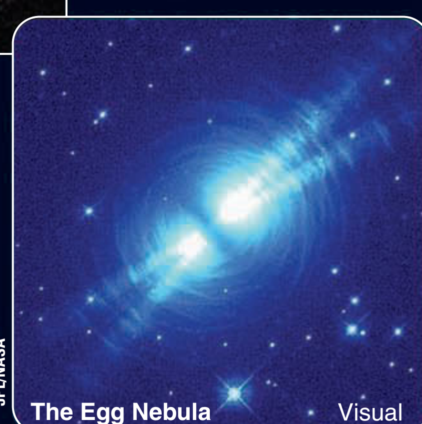
Some planetary nebulae, such as M2-9, are highly elongated, and it has been suggested that the Ring Nebula, at left, is a tubular shape that happens to be pointed roughly at Earth.

NASA

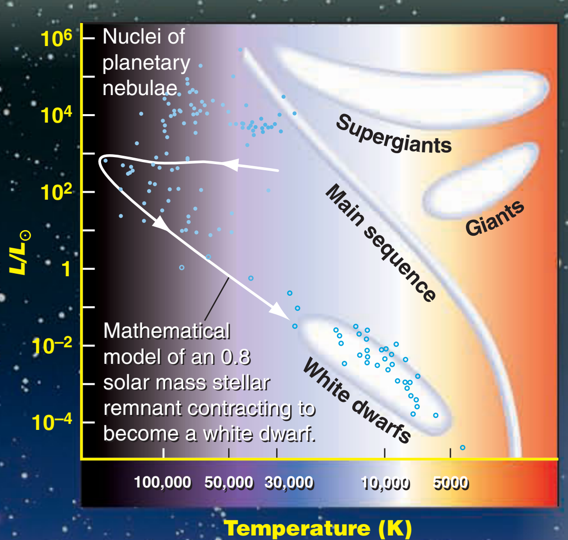


NASA

At visual wavelengths, the Egg Nebula is highly elongated, as shown below. The infrared image at left reveals an irregular, thick disk from which jets of gas and dust emerge. Such beams may create many of the asymmetries in planetary nebulae.



JPL/NASA



4 Once an aging giant star blows its surface into space to form a planetary nebula, the remaining hot interior collapses into a small, intensely hot object containing a carbon and oxygen interior surrounded by hydrogen and helium fusion shells and a thin atmosphere of hydrogen. The fusion gradually dies out, and the core of the star evolves to the left of the conventional H–R diagram to become the intensely hot nucleus of a planetary nebulae. Mathematical models show that these nuclei cool slowly to become white dwarfs.

a planetary nebula, a star must become a white dwarf with a temperature of at least 25,000 K. Mathematical models show that a collapsing star of less than 0.55 solar mass can take as long as a million years to heat up enough to ionize its nebulae, and by that time, the expelled gases are long gone. Models of the sun are not precise enough to indicate how much mass will be left once it ejects its outer layers. If it is left with too little mass, it may heat too slowly. Also, some research suggests that a star needs a binary companion to speed up its spin and make it eject a planetary nebula. The sun, of course, has no binary companion. This is an area of active research, and there are no firm conclusions. Are you disappointed that the sun may not light up its own planetary nebula? At least that potential embarrassment lies a few billion years in the future.

Medium-mass stars die by ejecting gas into space and contracting into white dwarfs. You have found evidence regarding the deaths of medium-mass stars in observations of planetary nebulae. Now you can turn your attention to the evidence revealed by white dwarfs themselves.

White Dwarfs

When you surveyed the stars you discovered that white dwarfs are the second most common kind of star (see page 155). Only red dwarfs are more abundant. Now you can recognize the billions of white dwarfs in our galaxy as the remains of medium-mass stars.

The first white dwarf discovered was the faint companion to Favorite Star Sirius. In that visual binary system, the bright star is Sirius A. The white dwarf, Sirius B, is 10,000 times fainter than Sirius A. The orbital motions of the stars (shown in Figure 8-12) reveal that the white dwarf contains 0.98 solar mass, and its blue-white color tells you that its surface is hot, about 45,000 K. Because it is both very hot and very low luminosity, it must have a small surface area (see Reasoning with Numbers 8-3)—in fact, it is about the size of Earth. Dividing its mass by its volume reveals that it is very dense—about 2×10^6 g/cm^3. On Earth, a teaspoonful of Sirius B material would weigh more than 11 tons.

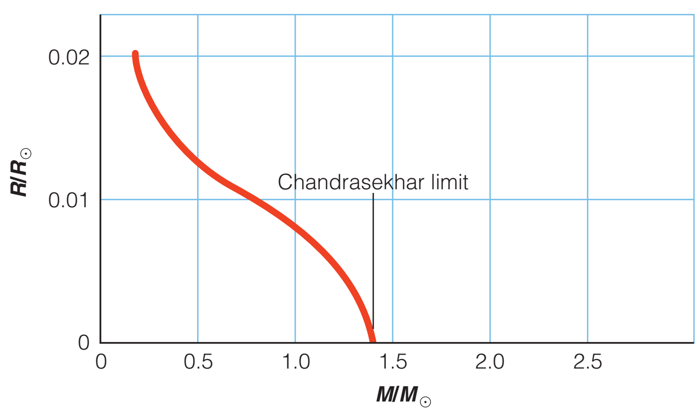
A normal star is supported by energy flowing outward from its core, but a white dwarf cannot generate energy by nuclear fusion. It has exhausted its hydrogen and helium fuels and converted them into carbon and oxygen. When a star collapses into a white dwarf, it converts gravitational energy into thermal energy. Its interior becomes very hot, but it cannot get hot enough to fuse its carbon-oxygen interior. Instead the star contracts until it becomes degenerate. Although a tremendous amount of energy flows out of the hot interior, it is not the energy flow that supports the star. The white dwarf is supported against its own gravity by the pressure of its degenerate electrons.

The interior of a white dwarf is mostly carbon and oxygen nuclei immersed in a whirling storm of degenerate electrons. Theory predicts that as the star cools these particles will lock together to form a crystal lattice, so there may be some truth in thinking of the interiors of aging white dwarfs as great crystals of carbon and oxygen. Near the surface, where the pressure is lower, a layer of ionized gases makes up a hot atmosphere. A 150-lb human would weigh 50 million pounds on the surface of a white dwarf. That strong gravity pulls the atmosphere down into a shallow layer. If Earth's atmosphere were equally shallow, people on the top floors of skyscrapers would have to wear spacesuits.

Clearly, a white dwarf is not a true star. It generates no nuclear energy, is almost totally degenerate, and, except for a thin layer at its surface, contains no gas. Instead of calling a white dwarf a "star," you can call it a "compact object." The next chapter discusses two other compact objects, neutron stars and black holes.

A white dwarf 's future is bleak. As it radiates energy into space, its temperature gradually falls, but it cannot shrink any smaller because its degenerate electrons resist getting closer together. Degenerate matter is a very good thermal conductor, so as heat flows to the surface and escapes into space, the white dwarf gets fainter and cooler, moving downward and to the right in the H–R diagram. Because a white dwarf contains a tremendous amount of heat, it needs billions of years to radiate that heat through its small surface area. Eventually, such objects may become cold and dark, so-called **black dwarfs.** Our galaxy is not old enough to contain black dwarfs. The coolest white dwarfs in our galaxy are about the temperature of the sun.

Perhaps the most interesting thing astronomers have learned about white dwarfs has come from mathematical models. The equations predict that if you added mass to a white dwarf, its radius would *shrink* because added mass would increase its gravity and squeeze it tighter. If you added enough to raise its total mass to about 1.4 solar masses, the equations predict that its radius would shrink to zero (■ Figure 10-7). This is called the **Chandrasekhar limit** after Subrahmanyan Chandrasekhar, the astronomer who discovered it. It seems to imply that a star more



■ **Figure 10-7**

The more massive a white dwarf is, the smaller its radius. Stars more massive than the Chandrasekhar limit of 1.4 $M_\odot$ cannot be white dwarfs.

massive than 1.4 solar masses could not become a white dwarf unless it shed mass in some way.

Stars do lose mass. Observations provide clear evidence that young stars have strong stellar winds, and aging giants and supergiants also lose mass rapidly (■ Figure 10-8). This suggests that stars more massive than the Chandrasekhar limit can eventually die as white dwarfs if they reduce their mass. Theoretical models show that stars that begin life with as much as 8 solar masses should lose mass fast enough to reduce their mass below 1.4 solar masses and eventually collapse to form white dwarfs. With mass loss, a wide range of medium-mass stars can eventually die as white dwarfs.

◄ SCIENTIFIC ARGUMENT ►

What evidence can you site to show that large numbers of stars die by producing planetary nebula?

You can begin your argument by noting that planetary nebulae are only a light-year or so in radius and that Doppler shifts show that they are expanding at 10 to 20 km/s. Dividing the radius by the velocity, tells you that a typical planetary nebula is only about 10,000 years old. That means that the nebulae don't last very long. Nevertheless, astronomers find 1500 of them visible in the sky. To be so common but so short lived, planetary nebulae must be produced in large numbers as medium-mass stars blow their outer layers into space.

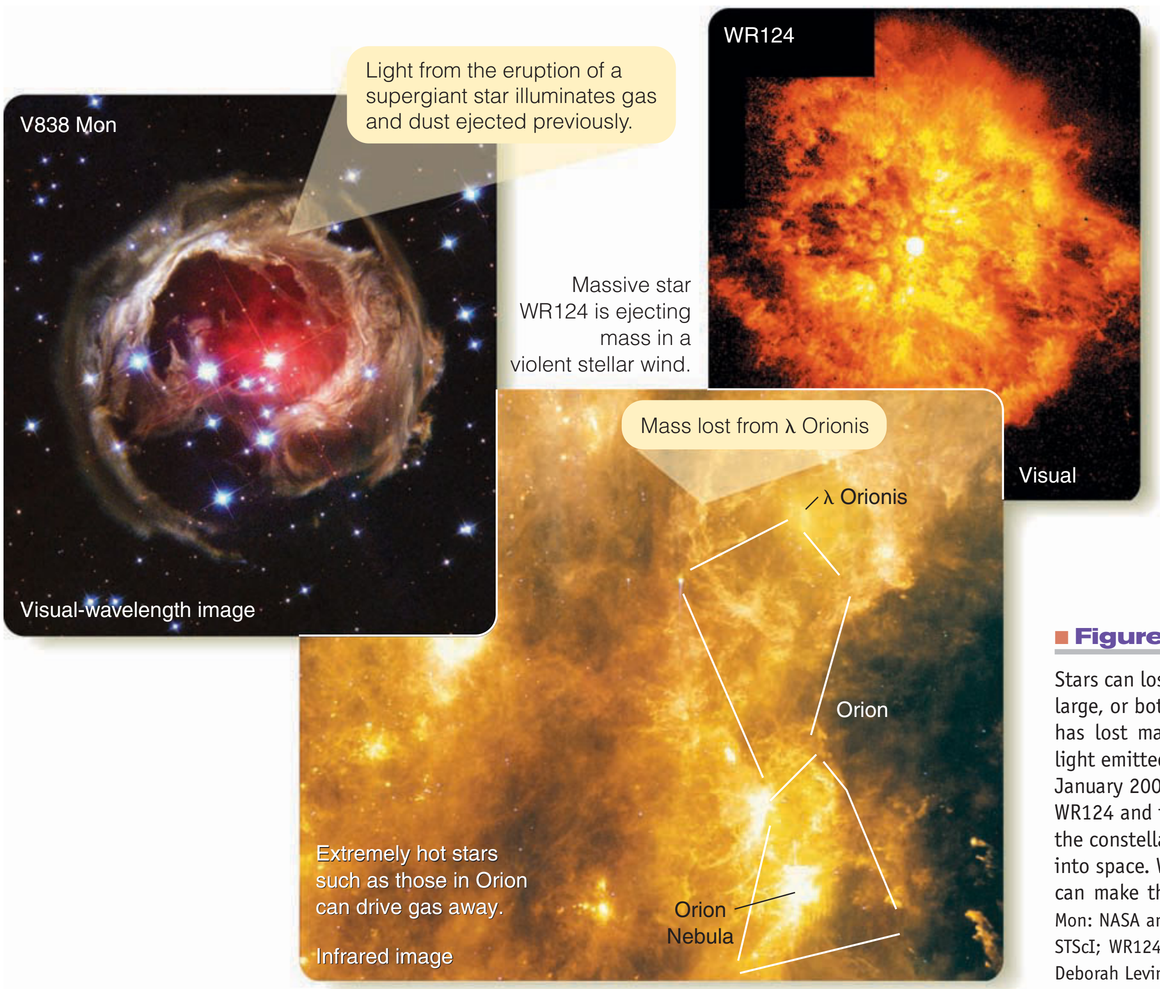
Now review more evidence. **Use Favorite Star Sirius to explain how you know that white dwarfs are very dense.**

◄ ►

10-3 The Evolution of Binary Systems

Stars in binary systems can evolve independently of each other if their orbits are large. In this situation, one of the stars can swell into a giant and collapse without disturbing its companion. But some binary stars orbit as close to each other as 0.1 AU, and when one of those stars begins to swell into a giant, its companion can suffer in peculiar ways.

These interacting binary stars are interesting in their own right. The stars share a complicated history and can experience strange and violent phenomena as they evolve. But such systems are also important because they can help astronomers understand the ultimate fate of stars. In the next chapter you will see how astronomers use interacting binary stars to search for black holes.



■ **Figure 10-8**

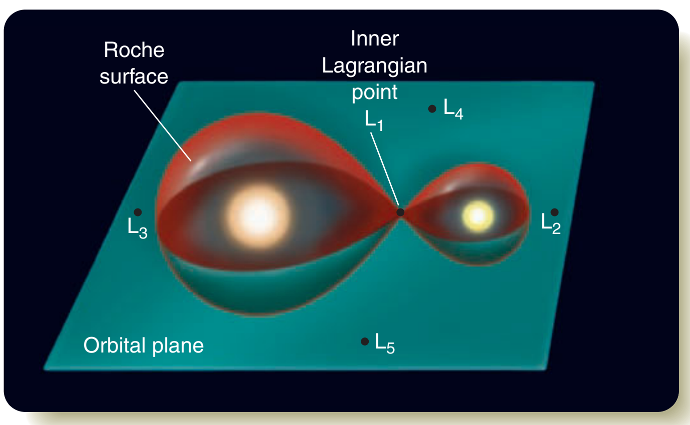
Stars can lose mass if they are very hot, very large, or both. The red supergiant V838 Mon has lost mass in the past, as revealed by light emitted when it erupted temporarily in January 2002. A young massive star such as WR124 and the hot, blue stars that make up the constellation Orion constantly lose mass into space. Warmed dust in these gas clouds can make them glow in the infrared. (V838 Mon: NASA and The Hubble Heritage Team AURA/STScI; WR124: NASA; Orion: NASA/IPAC courtesy Deborah Levine)

Mass Transfer

Binary stars can sometimes interact by transferring mass from one star to the other. The gravitational fields of the two stars, combined with the rotation of the binary system, define a dumbbell-shaped volume around the pair of stars called the **Roche lobes.** The surface of this volume is called the **Roche surface.** The size of the Roche lobes depends on the mass of the stars and on the distance between the stars. If the stars are far apart, the lobes are very large, and the stars easily control their own mass. If the stars are close together, however, the lobes are small and can interfere with the evolution of the stars. Matter inside each star's Roche lobe is gravitationally bound to the star, but matter that leaves a star's Roche lobe can fall into the other star or leave the binary system completely.

The **Lagrangian points** are places in the orbital plane of a binary star system where a bit of matter can reach stability. For astronomers, the most important of these points is the **inner Lagrangian point,** where the two Roche lobes meet (■ Figure 10-9). If matter can leave a star and reach the inner Lagrangian point, it can flow onto the other star. Thus, the inner Lagrangian point is the connection through which the stars can transfer matter.

In general, there are only two ways matter can escape from a star and reach the inner Lagrangian point. First, if a star has a strong stellar wind, some of the gas blowing away from it can pass through the inner Lagrangian point and be captured by the other star. Second, if an evolving star expands so far that it fills its Roche lobe, which can occur if the stars are close together and the lobes are small, then matter can overflow through the inner Lagrangian point onto the other star. Mass transfer driven by a stellar wind tends to be slow, but mass can be transferred rapidly by an expanding star.



■ **Figure 10-9**

A pair of binary stars control the region of space located inside the Roche surface. The Lagrangian points are locations of stability, with the inner Lagrangian point making a connection through which the two stars can transfer matter.

Evolution with Mass Transfer

Mass transfer between stars can affect the evolution of the stars in surprising ways. In fact, it is the solution to a problem that puzzled astronomers for many years.

In some binary systems, the less-massive star has become a giant, while the more-massive star is still on the main sequence. If higher-mass stars evolve faster than lower-mass stars, how do the lower-mass stars in such binaries manage to leave the main sequence first? This is called the Algol paradox, after the binary system Algol (Figure 8-19).

Mass transfer explains how this could happen. Imagine a binary system that contains a 5-solar-mass star and a 1-solar-mass companion. The two stars formed at the same time, so the higher-mass star, evolving faster, leaves the main sequence first. When it expands into a giant, however, it fills its Roche lobe and transfers matter to the low-mass companion. The higher-mass star loses mass and evolves into a lower-mass star, and the companion gains mass and becomes a higher-mass star that is still on the main sequence. This explains how there could be a system such as Algol that contains a 5-solar-mass main-sequence star and a 1-solar-mass giant.

The first four frames of ■ Figure 10-10 show mass transfer producing a system like Algol. The last frame shows an additional stage in which the giant star has collapsed to form a white dwarf and the more massive companion has expanded and is transferring matter back to the white dwarf. Such systems can become the site of tremendous explosions. To see how this can happen, you need to think about how mass falls into a star.

Accretion Disks

Matter flowing from one star to another cannot fall directly into the star. Rather, because of conservation of angular momentum, it must flow into a whirling disk around the star.

Angular momentum refers to the tendency of a rotating object to continue rotating. All rotating objects possess some angular momentum, and in the absence of external forces, an object maintains (conserves) its total angular momentum. An ice skater takes advantage of conservation of angular momentum by starting a spin slowly with her arms extended and then drawing them in. As her mass becomes concentrated closer to her axis of rotation, she spins faster (■ Figure 10-11). The same effect causes the slowly circulating water in a bathtub to spin in a whirlpool as it approaches the drain.

Mass transferred through the inner Lagrangian point in a binary system toward a star must conserve its angular momentum. If the star is small enough, as in the case of a white dwarf, the mass will form a rapidly rotating whirlpool called an **accretion disk** (■ Figure 10-12).

Two important things happen in an accretion disk. First, the gas in the disk grows very hot due to friction and tidal forces. The disk also acts as a brake, shifting angular momentum outward in the disk and allowing the innermost matter to fall into

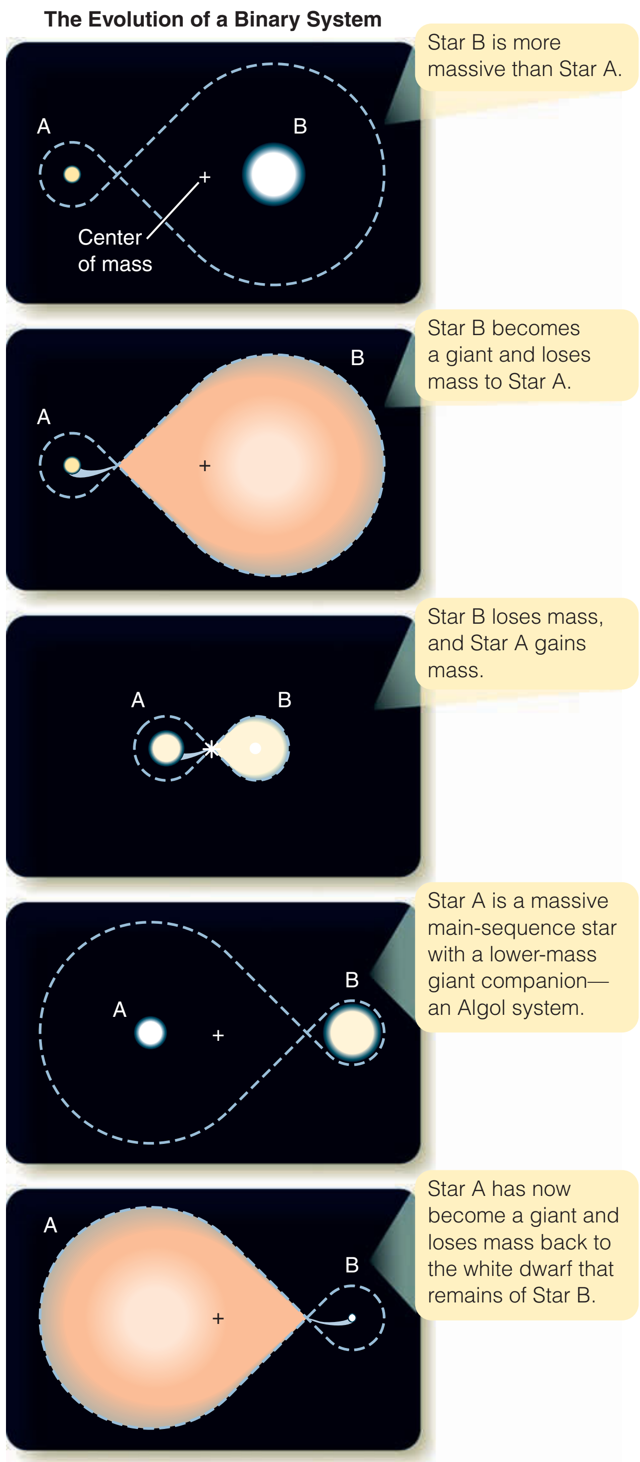


Figure 10-10

A pair of stars orbiting close to each other can exchange mass and modify their evolution.

the white dwarf. The interior parts of an accretion disk around a white dwarf are violent places. The temperature of the gas can exceed a million Kelvin, causing the gas to emit X-rays, and the matter falling inward can produce a violent explosion when enough accumulates on the white dwarf.

Figure 10-11

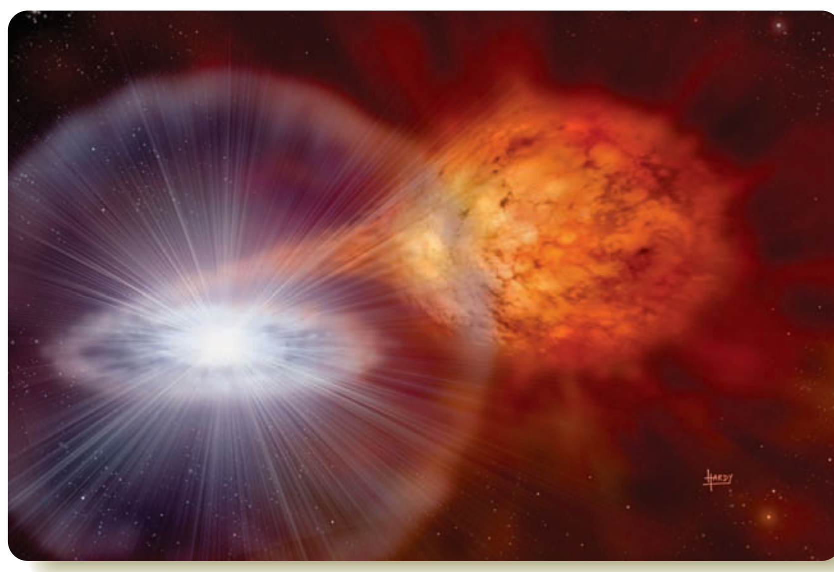
A skater demonstrates conservation of angular momentum when she spins faster by drawing her arms and legs closer to her axis of rotation.

Novae

At the beginning of this chapter you read that the word *nova* refers to what seems to be a new star appearing in the sky for a while and then fading away. Modern astronomers know that a nova is not a new star but an old star flaring up. After a nova fades, astronomers can photograph the spectrum of the remaining faint point of light. Invariably, they find a short-period spectroscopic binary containing a normal star and a white dwarf. A nova is evidently an explosion involving a white dwarf.

Figure 10-12

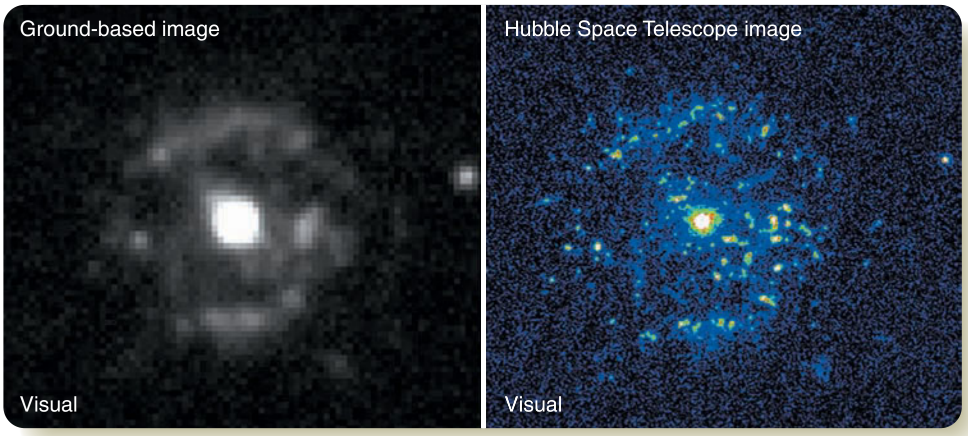
Matter from an evolving red giant falls into a white dwarf and forms a whirling accretion disk. Friction and tidal forces can make the disk very hot. Such systems can lead to nova explosions on the surface of the white dwarf, as shown in this artist's impression. (David A. Hardy, *www.astroart.org*, and PPARC)



Observational evidence can tell you how nova explosions occur. As the explosion begins, spectra show blueshifted absorption lines, which tells you the gas is dense and coming toward you at a few thousand kilometers per second. After a few days, the spectral lines change to emission lines, telling you the gas has thinned. The blueshifts remain, so you can conclude that an expanding cloud of debris has been ejected into space.

Nova explosions occur when mass transfers from a normal star through the inner Lagrangian point into an accretion disk around the white dwarf. As the matter loses its angular momentum in the inner accretion disk, it settles onto the surface of the white dwarf and forms a layer of unused nuclear fuel—mostly hydrogen. As the layer deepens, it becomes denser and hotter until the hydrogen fuses in a sudden explosion that blows the surface off of the white dwarf. Although the expanding cloud of debris contains less than 0.0001 solar mass, it is hot, and its expanding surface area makes it very luminous. Nova explosions can become 100,000 times more luminous than the sun. As the debris cloud expands, cools, and thins over a period of weeks and months, the nova fades from view.

The explosion of this material hardly disturbs the white dwarf and its companion star. Mass transfer quickly resumes, and a new layer of fuel begins to accumulate. How fast the fuel builds up depends on the rate of mass transfer. You can expect novae to repeat each time an explosive layer accumulates. Many novae take thousands of years to build an explosive layer, but some take only decades (■ Figure 10-13).



■ **Figure 10-13**

Nova T Pyxidis erupts about every two decades, expelling shells of gas into space. The shells of gas are visible from ground-based telescopes, but the Hubble Space Telescope reveals much more detail. The shell consists of knots of excited gas that presumably form when a new shell overtakes and collides with a previous shell. (M. Shara and R. Williams, STScI; R. Gilmozzi, ESO; and NASA)

The End of Earth

Astronomy is about us. Although this chapter has discussed the deaths of stars, it has also been discussing the future of our planet. The sun is a medium-mass star and must eventually die by becoming a giant, possibly producing a planetary nebula, and collapsing into a white dwarf. That will spell the end of Earth.

Mathematical models of the sun suggest that it may survive for an additional 5 billion years or so, but it is already growing more luminous as it fuses hydrogen into helium. In a few billion years, it will exhaust hydrogen in its core and swell into a giant star about 100 times its present radius. That giant sun will be about as large as the orbit of Earth, so that will mark the end of our world. The sun's growing luminosity will certainly evaporate Earth's oceans, drive away the atmosphere, and even vaporize much of Earth's crust. Models predict that the expanding sun will eventually become large enough to totally engulf Earth.

While it is a giant star, the sun will lose mass into space, and most of the atoms that were in Earth will be part of the expanding nebula around the sun. If it becomes hot enough, it will ionize the expelled gas and light it up as a planetary nebula. Your atoms will be part of that nebula.

There is no danger that the sun will explode as a nova; it has no binary companion. And, as you will see, the sun is not massive enough to die the violent death of the massive stars.

The most important lesson of astronomy is that we are part of the universe and not just observers. The atoms we are made of are destined to return to the interstellar medium in just a few billion years. That's a long time, and it is possible that the human race will migrate to other planetary systems before then. That might save the human race, but our planet is stardust.

◄ SCIENTIFIC ARGUMENT ►

How does spectroscopic evidence tell you what a nova explosion is like? For this argument you need to use your knowledge of basic spectroscopy. As soon as a nova is seen, astronomers rush to telescopes to record spectra, and they see blueshifted absorption lines. The blueshifts are Doppler shifts showing that the near side of the object is coming toward Earth. The absorption lines must be formed by fairly dense gas seen through thinner gas, much like the atmosphere of a star, so the surface of the star must be expanding rapidly outward. Later the spectrum becomes an emission spectrum, and Kirchhoff's laws tell you that the gas must have thinned enough to become transparent. The continued blueshift shows that the expansion is continuing.

Now review the postexplosion evidence. **How do observations of novae long after they have faded provide evidence that white dwarfs are involved?**

◄ ►

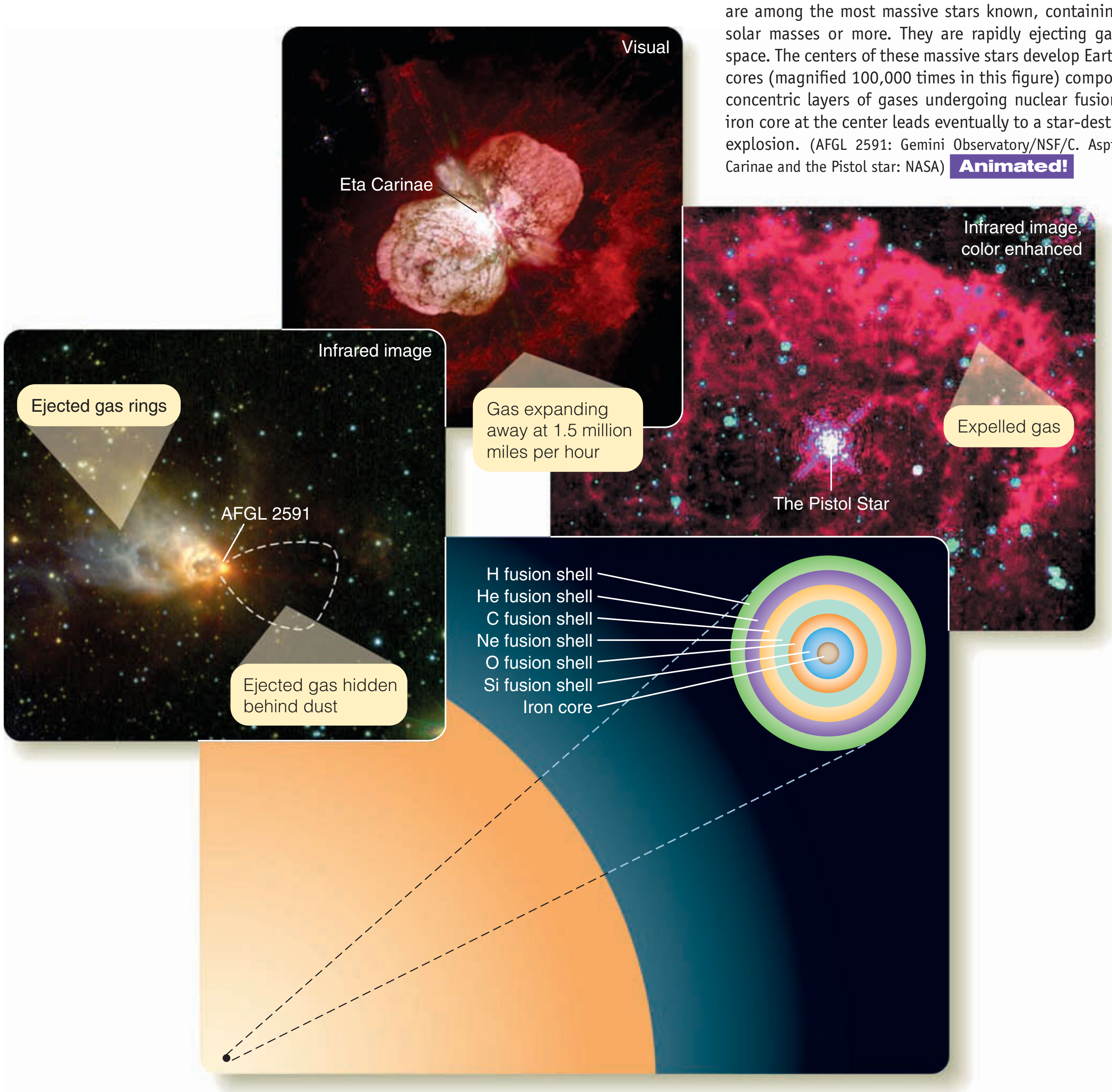
10-4 The Deaths of Massive Stars

You have seen that low- and medium-mass stars die relatively quietly as they exhaust their hydrogen and helium and then drive away their surface layers to form planetary nebulae. In contrast, massive stars live spectacular lives (■ Figure 10-14) and then destroy themselves in violent explosions.

Nuclear Fusion in Massive Stars

Stars on the upper main sequence have too much mass to die as white dwarfs, but their evolution begins much like that of their lower-mass cousins. They consume the hydrogen in their cores and ignite hydrogen shells; as a result, they expand into giants, or, for the most massive stars, supergiants. Next, their cores contract and fuse helium—first in the core and then in a shell, producing a carbon–oxygen core.

Unlike medium-mass stars, the massive stars do become hot enough to ignite carbon fusion at a temperature of about 1 billion Kelvin. Carbon fusion produces more oxygen and neon. As soon as the carbon is exhausted in the core, the core contracts, and carbon ignites in a shell. This pattern of core ignition and shell ignition continues with fuel after fuel, and the star develops the layered structure as shown in Figure 10-14, with a hydrogen-fusion shell



■ **Figure 10-14**

Massive stars live fast and die young. The three shown here are among the most massive stars known, containing 100 solar masses or more. They are rapidly ejecting gas into space. The centers of these massive stars develop Earth-size cores (magnified 100,000 times in this figure) composed of concentric layers of gases undergoing nuclear fusion. The iron core at the center leads eventually to a star-destroying explosion. (AFGL 2591: Gemini Observatory/NSF/C. Aspin; Eta Carinae and the Pistol star: NASA) **Animated!**

above a helium-fusion shell above a carbon-fusion shell above . . . After carbon fuses, oxygen, neon, and magnesium fuse to make silicon and sulfur, and then the silicon fuses to make iron.

The fusion of these nuclear fuels goes faster and faster as the massive star evolves. Recall that massive stars must consume their fuels rapidly to support their great weight, but other factors also cause the heavier fuels like carbon, oxygen, and silicon to fuse at increasing speeds. For one thing, the amount of energy released per fusion reaction decreases as the mass of the fusing atom increases. To support its weight, a star must fuse oxygen much faster than it fused hydrogen. Also, there are fewer nuclei in the core of the star by the time heavy nuclei begin to fuse. Four hydrogens make a helium nucleus, and three heliums make a carbon, so there are 12 times fewer nuclei of carbon available for fusion than there were hydrogen. This means the fusion of heavy elements goes very quickly in massive stars (■ Table 10-1). Hydrogen core fusion can last 7 million years in a 25-solar-mass star, but that same star will fuse the oxygen in its core in 6 months and its silicon in a day.

Supernova Explosions of Massive Stars

Theoretical models of evolving stars combined with nuclear physics allow astronomers to describe what happens inside a massive star when the last of its nuclear fuels are exhausted. The death of a massive star begins with iron nuclei and ends in cosmic violence.

Silicon fusion produces iron, the most tightly bound of all atomic nuclei (see Figure 7-7). Nuclear fusion is able to release energy by combining less tightly bound nuclei into a more tightly bound nucleus, but once the gas in the core of the star has been converted to iron, there are no nuclear reactions that can combine iron nuclei and release energy. The iron core is a dead end in the evolution of a massive star.

As a star develops an iron core, energy production begins to decline, and the core contracts. For nuclei less massive than iron, such contraction heats the gas and ignites new fusion fuels, but nuclear reactions involving iron remove energy from the core in two ways. First, the iron nuclei begin capturing electrons and breaking into smaller nuclei. The gas is so dense it is degenerate,

■ Table 10-1 | Heavy-Element Fusion in a 25-$M_\odot$ Star

Fuel	Time	Percentage of Lifetime
H	7,000,000 years	93.3
He	500,000 years	6.7
C	600 years	0.008
O	0.5 years	0.000007
Si	1 day	0.00000004

and the degenerate electrons helped support the core. The loss of some of the electrons allows the core to contract even faster. Second, temperatures are so high that the average photon is a high-energy gamma ray, and these gamma rays are absorbed by atomic nuclei, causing them to break into smaller fragments. The removal of the gamma rays also allows the core to contract even faster. Although the core of the star cannot generate energy by nuclear fusion, it can draw on the tremendous energy stored in its gravitational field. As the core contracts, the temperature shoots up, but it is not enough to stop the contraction, and the core of the star collapses inward in less than a tenth of a second.

This collapse happens so rapidly that the most powerful computers are unable to predict the details. Consequently, models of supernova explosions contain approximations. Nevertheless, the models predict that the collapse of the core produces an immense supernova explosion in which the star's outer layers are blasted outward. The core of the star must quickly become a neutron star or a black hole, the subjects of the next chapter.

To understand how the inward collapse of the core can produce an outward explosion, it helps to think about a traffic jam. The collapse of the innermost part of the degenerate core allows the rest of the core to fall inward creating a tremendous traffic jam as all of the nuclei fall toward the center. It is as if every car owner in Indiana suddenly tried to drive into downtown Indianapolis. There would be a traffic jam not only downtown but also in the suburbs; and, as more cars arrived, the traffic jam would spread outward. Similarly, although the innermost core collapses inward, a shock wave (a "traffic jam") develops and moves outward through the rest of the star.

The shock wave moves outward through the star aided by two additional sources of energy. First, when the iron nuclei in the core are disrupted, they produce a flood of neutrinos. In fact, for a short time the collapsing core produces more energy per second than all of the stars in all of the galaxies visible in the universe, and 99 percent of that energy is in the form of neutrinos. This flood of neutrinos carries large amounts of energy out of the core, allowing it to collapse further, and helps heat the gas outside the core and accelerate the outward-bound shock wave. The torrent of energy flowing out of the core also triggers tremendous turbulence, and intensely hot gas rushes outward from the interior (■ Figure 10-15). Again, this rising hot gas carries energy out into the envelope and helps drive the shock wave outward. Within a few hours, the shock wave bursts outward through the surface of the star and blasts it apart.

The supernova seen from Earth is the brightening of the star as its distended envelope is blasted outward by the shock wave. As months pass, the cloud of gas expands, thins, and fades, and the rate at which it fades matches the decay rate of certain radioactive nuclei produced in the explosion. The violence in the outer layers can create densities and temperatures high enough to trigger nuclear fusion reactions that produce as much as half a solar mass of radioactive nickel-56. The nickel gradually decays

The Exploding Core of a Supernova

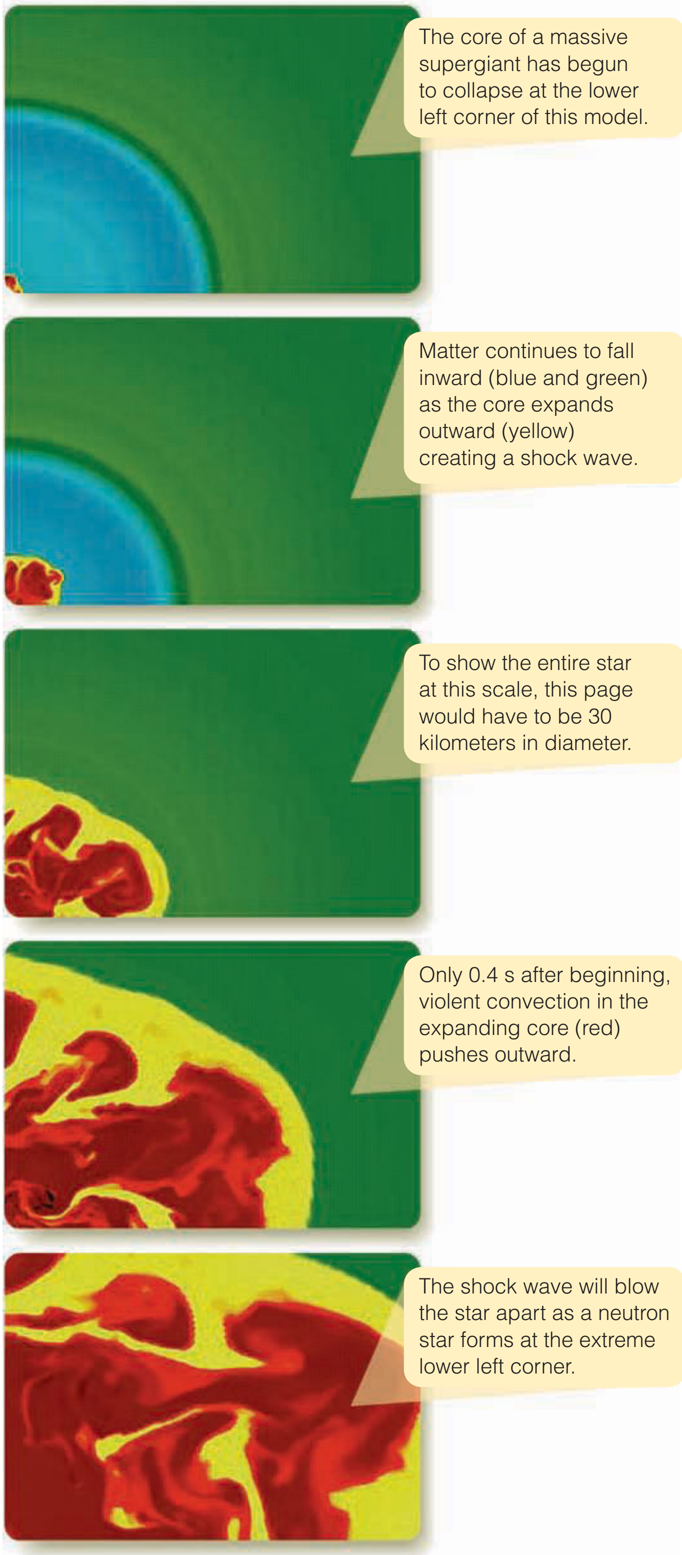


Figure 10-15

As the iron core of a massive star begins to collapse, intensely hot gas triggers violent convection. Even as the outer parts of the core continue to fall inward, the turbulence blasts outward and reaches the surface of the star within hours, creating a supernova eruption. This diagram is based on mathematical models and shows only the exploding core of the star. On this scale a diagram showing the entire star would be over 30 km in diameter. (Courtesy Adam Burrows, John Hayes, and Bruce Fryxell)

to form radioactive cobalt, which decays to form normal iron. Essentially all of the iron in the core of the star is destroyed when the core collapses, but more iron is produced in the outer layers, and that releases energy that keeps the supernova glowing.

The presence of nuclear fusion in the outer layers of the supernova testifies to the violence of the explosion. A typical supernova is equivalent to the explosion of 10^{28} megatons of TNT—about 5 million solar masses of high explosive. But, of course, the explosion is entirely silent. It is a **Common Misconception** promoted by science fiction movies and television that explosions in space are accompanied by sound. You know that's not true. Space is nearly a perfect vacuum, and sound can't travel through a vacuum. Supernova explosions are among the most violent events in nature, but they are silent.

Collapsing massive stars can trigger violent supernova explosions. There is, however, more than one kind of supernova.

Types of Supernovae

In studying supernovae in other galaxies, astronomers have noticed that there are a number of different types. **Type II supernovae** have spectra containing hydrogen lines and appear to be produced by the collapse and explosion of a massive star, the process discussed in the previous section. The hydrogen lines are produced by the outer layers of the star, which are rich in hydrogen. **Type I supernovae** have no hydrogen lines in their spectra, and astronomers have found at least two ways a supernova could occur and lack hydrogen—type Ia and type Ib supernovae.

A **type Ia supernova** occurs when a white dwarf gaining mass in a binary star system exceeds the Chandrasekhar limit and collapses. The collapse of a white dwarf is different from the collapse of a massive star because the core of the white dwarf contains usable fuel. As the collapse begins, the temperature and density shoot up, and the carbon–oxygen core begins to fuse in violent nuclear reactions. In a flicker of a stellar lifetime, the carbon–oxygen interior is entirely consumed, and the outermost layers are blasted away in a violent explosion that at its brightest is about six times more luminous than a type II supernova. The white dwarf is entirely destroyed, and no neutron star or black hole is left behind. No hydrogen lines are seen in the spectrum of a type Ia supernova explosion because white dwarfs contain very little hydrogen.

The less common **type Ib supernova** is thought to occur when a massive star in a binary system loses its hydrogen-rich outer layers to its companion star. The remains of the massive star could develop an iron core and collapse, as described in the previous section, producing a supernova explosion that lacked hydrogen lines in its spectrum. A type Ib supernova is just a type II supernova in which the massive star has lost its atmosphere.

Astronomers working with the largest and fastest computers are using modern theory to try to understand supernova explosions. But the companion to theory is observation, so you should ask what observational evidence supports this story of supernova explosions.

Observations of Supernovae

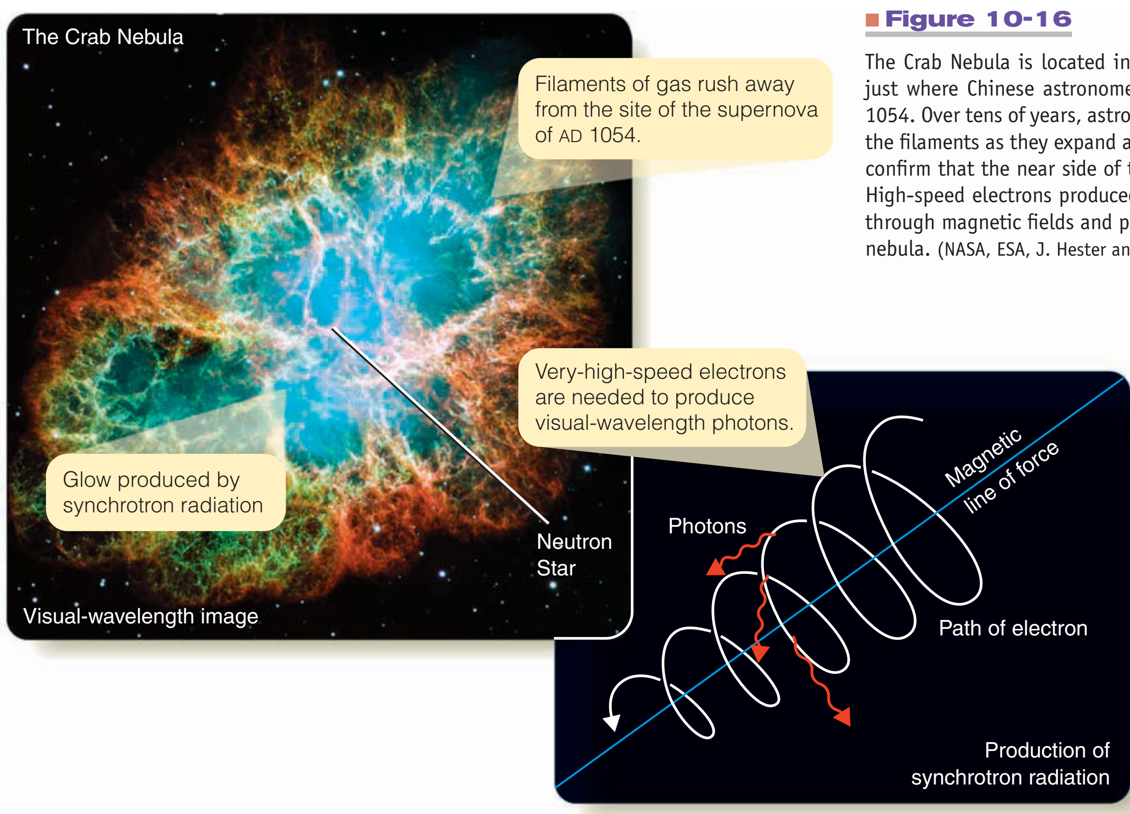
In AD 1054, Chinese astronomers saw a "guest star" appear in the constellation we know as Taurus the Bull. The star quickly became so bright it was visible in the daytime, and then, after a month, it slowly faded, taking almost two years to vanish from sight. When modern astronomers turned their telescopes to the location of the "guest star," they found a peculiar nebula now known as the Crab Nebula for its many-legged shape. In fact, the legs of the Crab Nebula are filaments of gas that are moving away from the site of the explosion at about 1400 km/s. Comparing the radius of the nebula, 1.35 pc, with its velocity of expansion reveals that the nebula began expanding nine or ten centuries ago, just when the "guest star" made its visit. The Crab Nebula is clearly the remains of the supernova seen in AD 1054 (■ Figure 10-16). In the next chapter, you will meet the neutron star found at the center of the Crab Nebula.

The blue glow of the Crab Nebula is produced by **synchrotron radiation.** This form of electromagnetic radiation is produced by rapidly moving electrons spiraling through magnetic fields and is common in the nebulae produced by supernovae. In the case of the Crab Nebula, the electrons travel so fast that they emit visual wavelengths; but, in most such nebulae, the electrons move more slowly, and the synchrotron radiation is at radio wavelengths. By now, the high-speed electrons in the Crab Nebula should have slowed, but the synchrotron radiation at visual wavelengths is still strong. That is evidence that the electrons are being energized by the neutron star at the center of the nebula.

Supernovae are rare. Only a few have been seen with the naked eye in recorded history. Arab astronomers saw one in AD 1006, and the Chinese saw one in AD 1054. European astronomers observed two—one in AD 1572 (Tycho's supernova) and one in AD 1604 (Kepler's supernova). In addition, the guest stars of AD 185, 386, 393, and 1181 may have been supernovae.

In the centuries following the invention of the astronomical telescope in 1609, no supernova was bright enough to be visible to the naked eye. Supernovae are sometimes seen in distant galaxies, but they are faint and hard to study. Then, in the early morning hours of February 24, 1987, astronomers around the world were startled by the discovery of a naked-eye supernova still growing brighter in the southern sky (■ Figure 10-17). The supernova, known officially as SN1987A, occurred only 53,000 pc away in the Large Magellanic Cloud, a small galaxy near our own Milky Way Galaxy. This first naked-eye supernova in 383 years gave astronomers a ringside seat for the most spectacular event in stellar evolution.

One observation of SN1987A is critical in that it confirms the theory of core collapse. At 2:35 AM EST on February 23, 1987, nearly 4 hours before the supernova was first seen, a blast of neutrinos swept through Earth, including perhaps 20 trillion that passed harmlessly through each human body on Earth. Instruments buried in a salt mine under Lake Erie and in a lead mine in Japan, though designed for another purpose, recorded 19 neutrinos in less than 15 seconds. Neutrinos are so difficult to detect that the 19 neutrinos actually detected meant that some



■ **Figure 10-16**

The Crab Nebula is located in the constellation Taurus the Bull, just where Chinese astronomers saw a brilliant guest star in AD 1054. Over tens of years, astronomers can measure the motions of the filaments as they expand away from the center. Doppler shifts confirm that the near side of the nebula is moving toward Earth. High-speed electrons produced by the central neutron star spiral through magnetic fields and produce the foggy glow that fills the nebula. (NASA, ESA, J. Hester and A. Loll, Arizona State University)

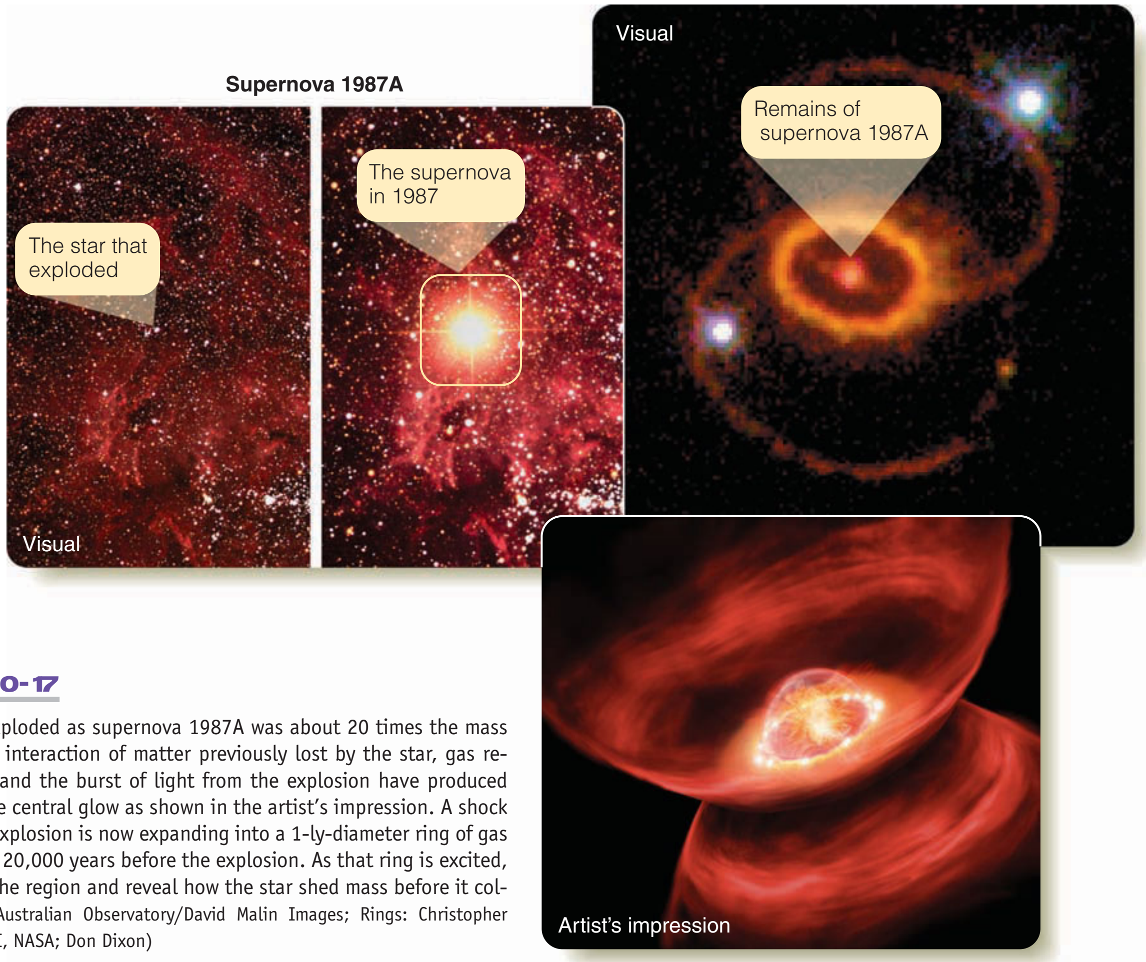


Figure 10-17

The star that exploded as supernova 1987A was about 20 times the mass of the sun. The interaction of matter previously lost by the star, gas recently ejected, and the burst of light from the explosion have produced rings around the central glow as shown in the artist's impression. A shock wave from the explosion is now expanding into a 1-ly-diameter ring of gas ejected roughly 20,000 years before the explosion. As that ring is excited, it will light up the region and reveal how the star shed mass before it collapsed. (Anglo-Australian Observatory/David Malin Images; Rings: Christopher Burrows, ESA/STcI, NASA; Don Dixon)

10^{17} neutrinos must have passed through the detectors in those 15 seconds. Furthermore, the neutrinos arrived from the direction of the supernova. Thus, astronomers concluded that the burst of neutrinos was released when the iron core collapsed into a neutron star, and the supernova itself was seen hours later when the shock wave blasted the star's surface into space.

Most supernovae are seen in distant galaxies (■ Figure 10-18), and careful observations allow astronomers to compare the behavior of different types. Type Ia supernovae, caused by the collapse of white dwarfs, are more luminous at maximum brightness and decline rapidly at first and then more slowly. Type II supernovae, produced by the collapse of massive stars, are not as bright at maximum. They decline in a more irregular way. Recall that type Ia supernovae have no hydrogen lines in their spectra, but type II supernovae do.

SN1987A was a type II supernova, although its light curve is not typical (■ Figure 10-19). It was produced by the explosion of a hot, blue supergiant rather than the usual cool, red supergiant. Evidently, the star was a red supergiant a few thousand years ago but had contracted and heated up slightly, becoming smaller, hotter, and bluer before it exploded. Astronomers believe that most type II supernovae are caused by the collapse of red supergiants.

Although supernova explosions fade to obscurity in a year or two, expanding shells of gas mark the sites of the explosions. The gas, originally expelled at 10,000 to 20,000 km/s, may carry away a fifth of the mass of the exploding star. The collision of that expanding gas with the surrounding interstellar medium can sweep up even more gas and excite it to produce a **supernova remnant,** the nebulous remains of a supernova explosion (■ Figure 10-20).

Some supernova remnants, such as Cassiopeia A, show evidence of jets of matter rushing outward in opposite directions. These may have been ejected as the rotating star collapsed, and, conserving angular momentum, spun up to very high speeds. The first matter blown outward from such a rapidly rotating star could have emerged as jets from its poles. Astronomers are just beginning to understand the details of such violent explosions.

Supernova remnants look quite delicate and do not survive very long—a few tens of thousands of years—before they gradually mix with the interstellar medium and vanish. The Crab Nebula is a young remnant, only about 950 years old and about 8.8 ly in diameter. Older remnants can be larger. Some supernova remnants are visible only at radio and X-ray wavelengths. They have become too tenuous to emit detectable light, but the collision of the expanding hot gas with the interstellar medium

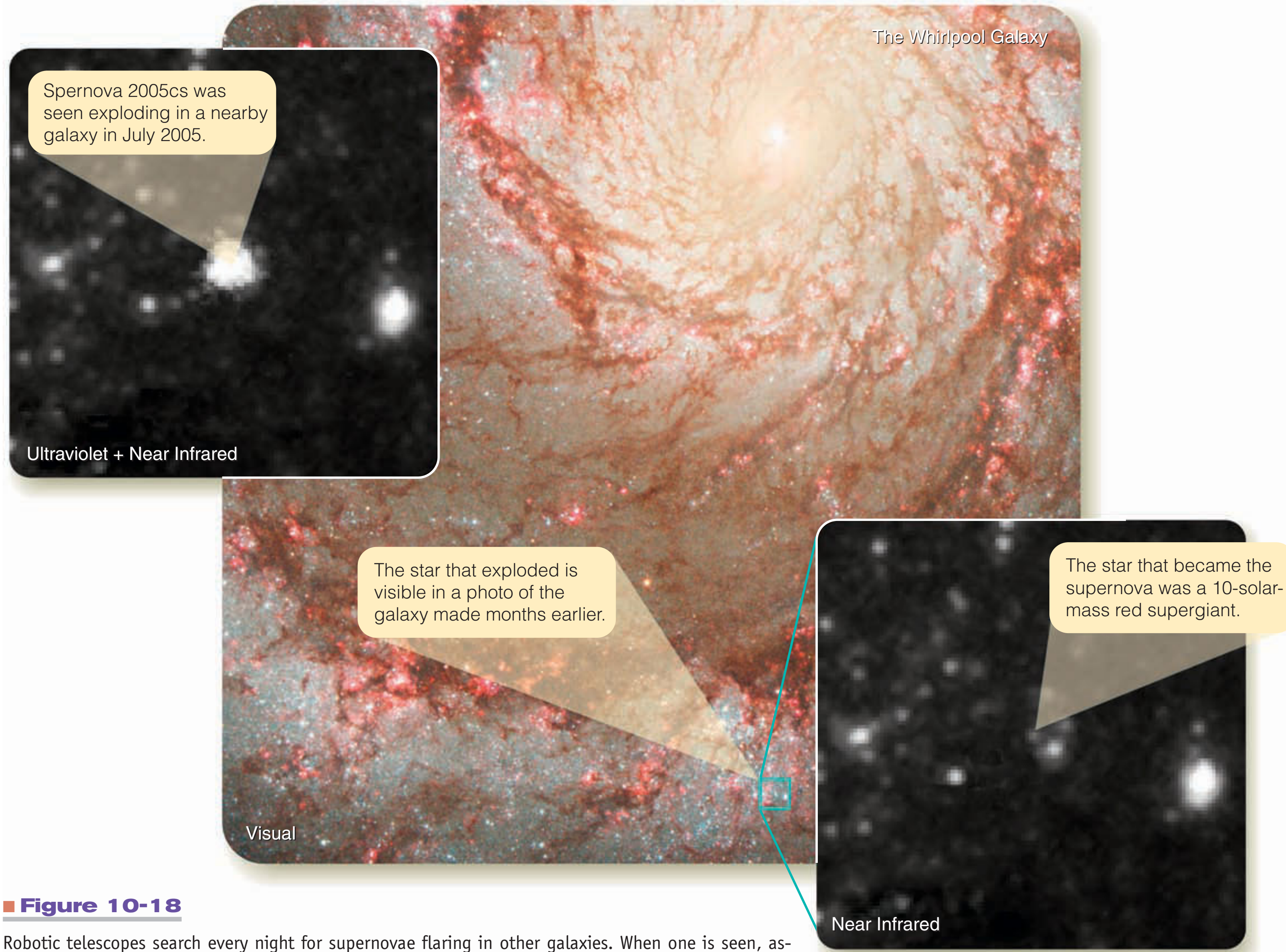


Figure 10-18

Robotic telescopes search every night for supernovae flaring in other galaxies. When one is seen, astronomers can obtain spectra and record the supernova's rise in brightness and its decline to study the physics of exploding stars. If a supernova is seen in a nearby galaxy, it is sometimes possible to identify the star in earlier photos. (NASA, ESA, W. Li and A. Filippenko, Berkeley, S. Beckwith, STScI, and The Hubble Heritage Team, STScI/AURA)

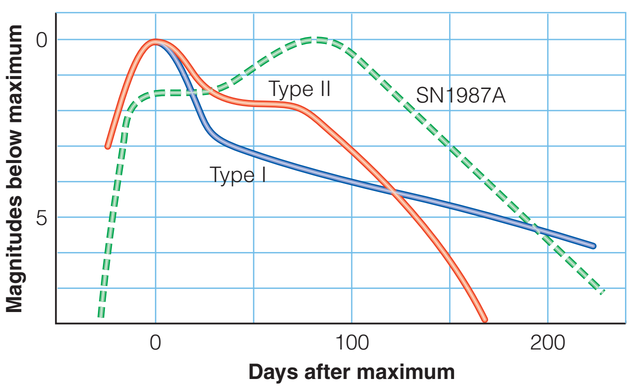


Figure 10-19

Type I supernovae decline rapidly at first and then more slowly, but type II supernovae pause for about 100 days before beginning a steep decline. Supernova 1987A was odd in that it did not rise directly to maximum brightness. These light curves have been adjusted to the same maximum brightness.

can generate radio and X-ray radiation. You saw in the previous chapter that the compression of the interstellar medium by expanding supernova remnants can also trigger star formation.

◄ SCIENTIFIC ARGUMENT ►

What evidence do astronomers have that Supernova 1987A formed a neutron star?

The critical observation consists of only 19 neutrinos detected coming from the direction of the supernova. Because neutrinos are so difficult to detect, those 19 indicate that a huge flood of neutrinos passed through Earth just before the supernova was seen. Theory says the collapse of the core into a ball of neutrons should release a tremendous burst of neutrinos, and astronomers link the neutrinos that were detected with the formation of that ball of neutrons. Notice that this evidence depends on a theory. That's not unusual, but scientists are very careful in analyzing such evidence to be sure the background theory is right. Only then is the evidence meaningful.

Now create a new argument. **What evidence can you cite that type Ia supernovae are not produced by massive stars?**

◄ ►

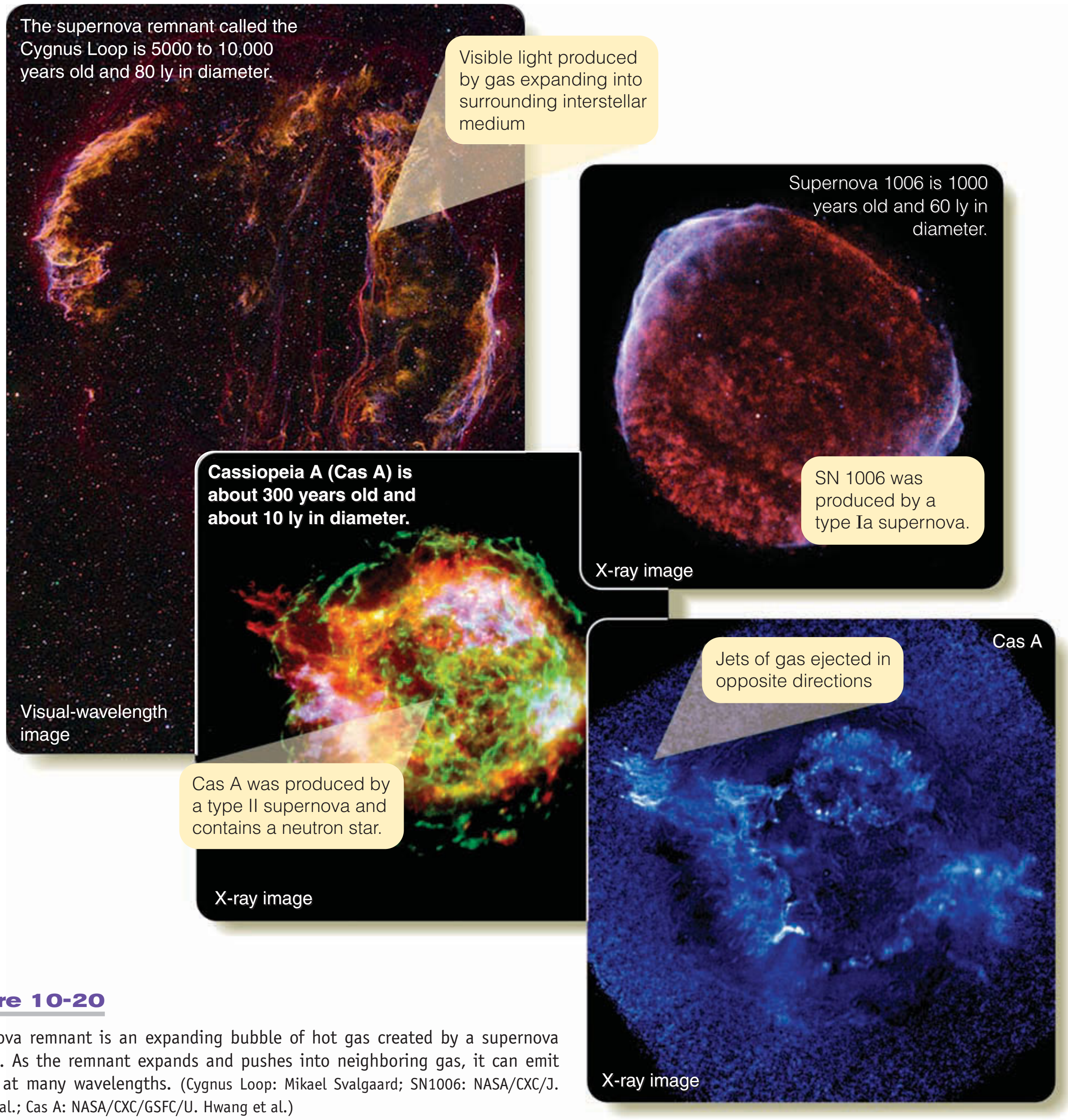


Figure 10-20

A supernova remnant is an expanding bubble of hot gas created by a supernova explosion. As the remnant expands and pushes into neighboring gas, it can emit radiation at many wavelengths. (Cygnus Loop: Mikael Svalgaard; SN1006: NASA/CXC/J. Hughes et al.; Cas A: NASA/CXC/GSFC/U. Hwang et al.)

What Are We? Stardust

You are made of atoms that were cooked up inside stars. Gravity draws matter together to make stars, and although nuclear fusion delays gravity's final victory, stars must eventually die. That process of star life and star death manufactures atoms heavier than helium and spreads them back into the interstellar medium where they can become part of the gas clouds that form new stars. All of the atoms in your body except for the hydrogen were made inside stars.

Some of your atoms, such as the carbon, were cooked up in the cores of medium-mass stars like the sun and were puffed out into space when those stars died and produced planetary nebulae. Some of your atoms, such as the calcium in your bones, were made inside massive stars and were blown out into space during a type II supernova explosion. Many of the iron atoms in your blood were made by the sudden fusing of carbon when white dwarfs collapsed in type Ia supernova explosions. In fact, a few of your heavier atoms, such as iodine in your thyroid gland and selenium in your nerve cells, were produced in the raging violence of supernova explosions.

You are made of star stuff scattered into space long ago by the violent deaths of stars. What are we? We are stardust.

Study and Review

Summary

- A **nova (p. 185)** appears to be a new star that becomes visible and then fades after a few weeks. **Supernovae (p. 185)** are more luminous and last longer. Both are associated with the deaths of stars.
- Main-sequence stars like the sun can expand and become giant stars when they use up the hydrogen fuel in their cores. As that happens, the core contracts and heats up. Hydrogen fusion begins in a spherical layer around the core—a hydrogen-fusion shell.
- Energy from the hydrogen-fusion shell swells the star into a cool giant 10 to 100 times larger in diameter than the sun. Massive stars swell to become supergiants up to 1000 times larger in diameter than the sun.
- The contraction of the star's core eventually ignites helium, first in the core and later in a shell. If the star is massive enough, it can eventually fuse carbon and other elements.
- In **degenerate matter (p. 187),** the density is so high that quantum mechanical effects prevent electrons from changing their energies. Such matter is very difficult to compress, and its pressure does not depend on its temperature.
- If a star's mass lies between about 0.4 and 3 solar masses, its helium core becomes degenerate before the helium ignites. Because pressure does not depend on temperature, there is no pressure–temperature thermostat to control the reactions, and when helium fusion ignites, the core explodes in a **helium flash (p. 188).** All of the energy produced is absorbed by the star.
- As the giant star fuses helium in its core and hydrogen in a shell, it moves toward the hot side of the H–R diagram. As soon as it exhausts the helium in its core and begins fusing helium in a shell, it moves back toward the cool side, producing a loop in its evolutionary path.
- You can see evidence of stellar evolution in the H–R diagrams of star clusters. Stars in both **open clusters (p. 192)** and in **globular clusters (p. 192)** evolve in similar ways. The stars begin their evolution at the same time but evolve at different rates, depending on their masses. The most-massive stars leave the main sequence first and are followed later by progressively less-massive stars. This makes the evolution of stars visible in the H–R diagram.
- You can estimate the age of a star cluster from the **turnoff point (p. 192)** in its H–R diagram.
- In old clusters, stars fusing helium follow a loop in the H–R diagram that is visible in the diagrams of star clusters as the **horizontal branch (p. 193).**
- Red dwarfs less massive than about 0.4 solar mass are completely mixed and will have very little hydrogen left when they die. They cannot ignite a hydrogen-fusion shell, so they cannot become giant stars. They will remain on the main sequence for many times the present age of the universe.
- Medium-mass stars like the sun become cool giants and fuse helium but cannot fuse carbon. They eventually blow away their outer layers and collapse into hot white dwarfs. Ultraviolet radiation from the white dwarfs ionizes the gas to produce **planetary nebulae (p. 191).**
- White dwarfs cannot contract as they cool and will eventually become **black dwarfs (p. 196).** The **Chandrasekhar limit (p. 196)** shows that no white dwarf more massive than 1.4 solar masses can be stable. Presumably more-massive stars can become white dwarfs only if they shed mass.
- Close binary stars evolve in complex ways because they can transfer mass from one star to the other. The two stars are enclosed by a dumbbell region known as the **Roche lobes (p. 198).** If mass from one star crosses the surface of these lobes, the **Roche surface (p. 198),** it can fall into the other star. The **Lagrangian points (p. 198)** in the rotating system are points where mass can remain stable. Matter can flow between the stars through the **inner Lagrangian point (p. 198),** which connects the two Roche lobes.
- As a star becomes a giant, it can expand and fill its Roche lobe, spilling mass to the other star. Mass transfer explains why some binary systems contain a main-sequence star more massive than its giant companion—the Algol paradox.
- Mass that is transferred from one star to the other must conserve **angular momentum** and can form a whirling **accretion disk (p. 198)** around the receiving star. Accretion disks can become hot enough to emit light and even X-rays.
- Mass transferred onto the surface of a white dwarf can build up a layer of fuel that erupts in a nova explosion. A white dwarf can erupt repeatedly so long as mass transfer continues to form new layers of fuel.
- The evolution of the sun into a giant and then its collapse into a white dwarf will end life on Earth.
- Stars more massive than about 8 solar masses cannot lose mass fast enough to reduce their mass low enough to die by ejecting a planetary nebula and collapsing into a white dwarf. Such massive stars must die more violent deaths.
- The massive stars on the upper main sequence fuse nuclear fuels one after the other, producing a layering of fusion shells, but such stars cannot fuse iron because iron is the most tightly bound of all atomic nuclei. When an aging massive star forms an iron core, the core collapses and triggers a supernova explosion known as a **type II supernova (p. 203).**
- The spectra of type II supernovae contain hydrogen lines, but the spectra of **type I supernovae (p. 203)** do not. At least two causes of type I supernovae are known.
- A **type Ia supernova (p. 203)** can occur when mass transferred onto a white dwarf pushes it over the Chandrasekhar limit and it collapses suddenly, fusing all of its carbon at once. A **type Ib supernova (p. 203)** occurs when a massive star in a binary system loses its outer layers of hydrogen before it explodes.
- A supernova expels an expanding shell of gas called a **supernova remnant (p. 205).** The supernova of AD 1054 produced a supernova remnant known as the Crab Nebula, which emits **synchrotron radiation (p. 204),** evidence of a powerful energy source remaining inside the remnant.
- The supernova 1987A is only a few years old, but its expanding gases will eventually form a supernova remnant. Neutrinos observed coming from the direction of the supernova are evidence that the core collapsed and formed a neutron star.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why does helium fusion require a higher temperature than hydrogen fusion?
2. How can the contraction of an inert helium core trigger the ignition of a hydrogen-fusion shell?
3. Why does the expansion of a star's envelope make it cooler and more luminous?
4. Why is degenerate matter so difficult to compress?
5. How does the presence of degenerate matter in a star trigger the helium flash?
6. How can star clusters confirm astronomers' theories of stellar evolution?
7. Why don't red dwarfs become giant stars?

8. What causes an aging giant star to produce a planetary nebula?
9. Why can't a white dwarf contract as it cools? What is its fate?
10. Why can't a white dwarf have a mass greater than 1.4 solar masses?
11. How can a star of as much as 8 solar masses form a white dwarf when it dies?
12. How can you explain the Algol paradox?
13. How can the inward collapse of the core of a massive star produce an outward explosion?
14. What is the difference between type I and type II supernovae?
15. What is the difference between a supernova explosion and a nova explosion?
16. **How Do We Know?** In what ways do the appearance of supernova explosions depend on the properties of subatomic particles?

Discussion Questions

1. How do you know the helium flash occurs if it cannot be observed? Can you accept something as real if you can never observe it?
2. False-color radio images and time-exposure photographs of astronomical images show aspects of nature you can never see with unaided eyes. Can you think of common images in newspapers or on television that reveal phenomena you cannot see?

Problems

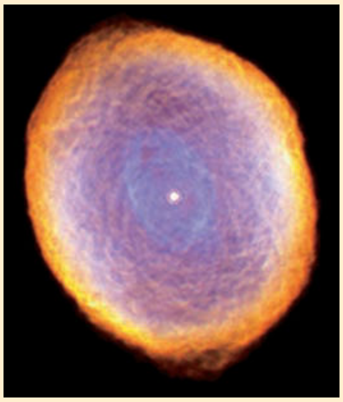
1. About how long will a 0.4-solar-mass star spend on the main sequence? (*Hint:* See Reasoning with Numbers 9-1.)
2. If the stars at the turnoff point in a star cluster have masses of about 4 solar masses, how old is the cluster?
3. About how far apart are the stars in an open cluster? In a globular cluster? (*Hint:* What share of the cluster's volume belongs to a single star?)
4. The Ring Nebula in Lyrae is a planetary nebula with an angular diameter of 76 arc seconds and is 5000 ly from Earth. What is its linear diameter? (*Hint:* See Reasoning with Numbers 3-1.)
5. If the Ring Nebula is expanding at a velocity of 15 km/s, typical of planetary nebulae, how old is it?
6. Suppose a planetary nebula is 1 pc in radius. If the Doppler shifts in its spectrum show it is expanding at 30 km/s, how old is it? (*Hints:* 1 pc equals 3×10^{13} km, and 1 year equals 3.16×10^7 seconds.)
7. If a star the size of the sun expands to form a giant 20 times larger in radius, by what factor will its average density decrease? (*Hint:* The volume of a sphere is $\frac{4}{3}\pi r^3$.
8. If a star the size of the sun collapses to form a white dwarf the size of Earth, by what factor will its density increase? (*Hints:* The volume of a sphere is $\frac{4}{3}\pi r^3$. See Appendix A for the radii of the sun and Earth.)
9. The Crab Nebula is now 1.35 pc in radius and is expanding at 1400 km/s. About when did the supernova occur? (*Hint:* 1 pc equals 3.1×10^{13} km.)
10. If the Cygnus Loop is 25 pc in diameter and is 10,000 years old, with what average velocity has it been expanding? (*Hints:* 1 pc equals 3.1×10^{13} km, and 1 year equals 3.16×10^7 seconds.)
11. Observations show that the gas ejected from SN1987A is moving at about 10,000 km/s. How long will it take to travel one astronomical unit? One parsec? (*Hints:* 1 AU equals 1.5×10^8 km, and 1 pc equals 3.1×10^{13} km.)

Learning to Look

1. The star cluster in the photo at the right contains many hot, blue, luminous stars. Sketch its H–R diagram and discuss its probable age.

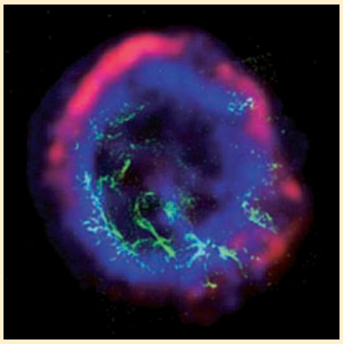
NASA/Walborn, Maíz-Apellániz, and Barbá

2. What processes caused a medium-mass star to produce the nebula at the right? The nebula is now about 0.1 ly in diameter and still expanding. What will happen to it?



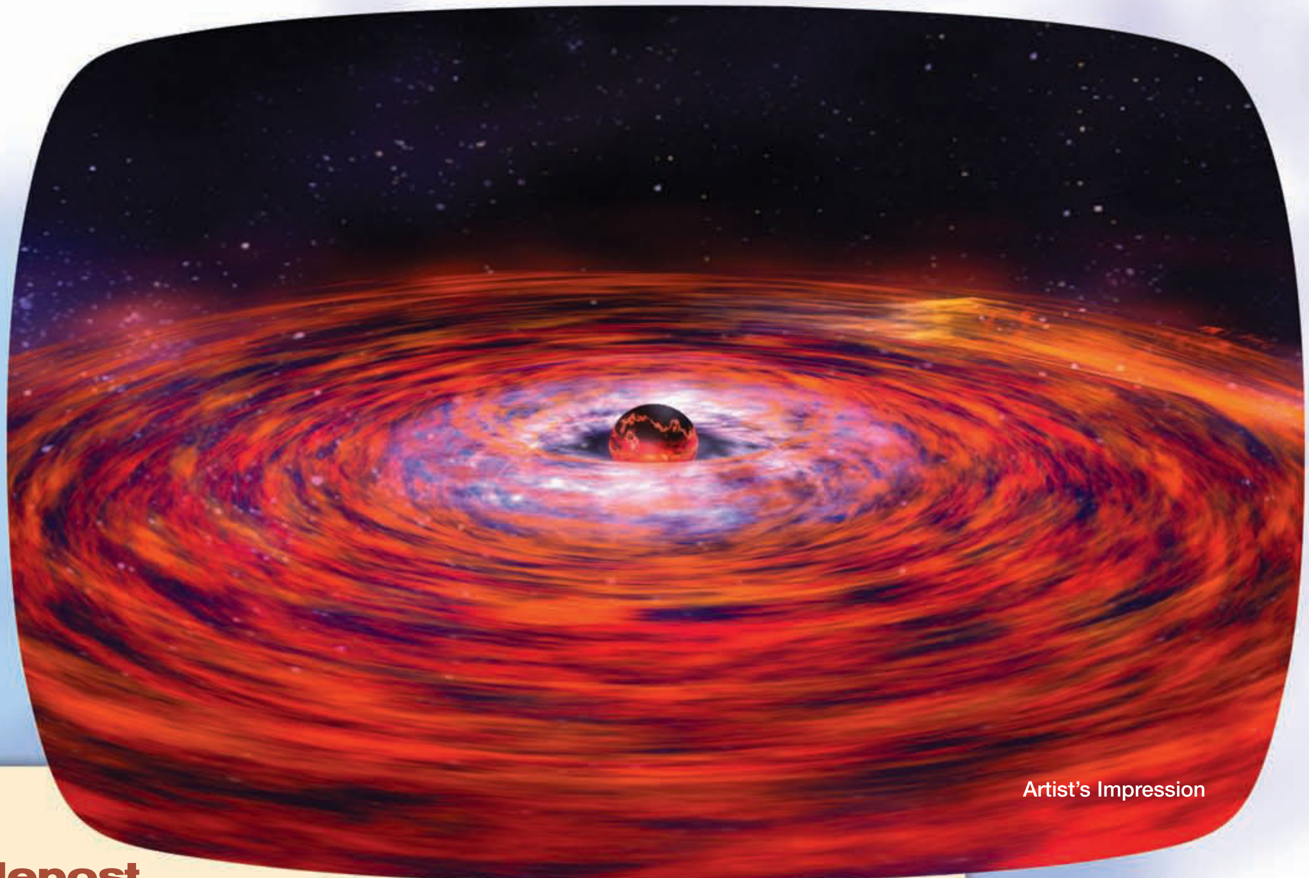
NASA/Hubble Heritage Team/STScI/AURA

3. The image at right combines X-ray (blue), visible (green), and radio (red) images. Observations show the sphere is expanding at a high speed and is filled with very hot gas. What kind of object produced this nebula? Roughly how old do you think it must be?



NASA/CAC/SAO/CSIRO/ATNF/ATCA

11 Neutron Stars and Black Holes



A neutron star containing roughly the mass of the sun and only about 10 km in radius draws matter inward through a whirling disk of gas hot enough to emit X-rays. (NASA/Dana Berry)

Guidepost

In the last two chapters you have traced the story of stars from birth to death. By now you are asking a simple question, "What's left?" The answer depends on the mass of the star. You already know that stars like the sun leave behind white dwarfs, but more massive stars leave behind the strangest beasts in the cosmic zoo.

Now you are ready to meet neutron stars and black holes, and your exploration will answer four essential questions:

How does theory predict the existence of neutron stars?

How do astronomers know neutron stars really exist?

How does theory predict the existence of black holes?

How can astronomers be sure that black holes really exist?

This chapter will show you clear examples of how astronomers combine observations and theory to understand nature.

This chapter ends the story of individual stars, but it does not end the story of stars. In the next chapter, you will begin exploring the giant communities in which stars live—the galaxies.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Almost anything is easier to get into than out of.

AGNES ALLEN

GRAVITY ALWAYS WINS. However a star lives, it must eventually die by collapsing into one of three final states—a white dwarf, neutron star, or black hole. These objects, often called compact objects, are small monuments to the power of gravity. Almost all of the energy available has been squeezed out of compact objects, and you find them in their final, high-density states.

In this chapter, you must compare evidence and theory with great care. Theory predicts the existence of these objects; but, by their nature, they are difficult to detect. To confirm the theories, astronomers have searched for real objects that can be identified as having the properties predicted by theory. That is, they have searched out real neutron stars and real black holes. No medieval knight ever rode off on a more difficult quest.

11-1 Neutron Stars

A **NEUTRON STAR** contains a little over 1 solar mass compressed to a radius of about 10 km. Its density is so high that the matter is stable only as a fluid of neutrons. Theory predicts that such an object would spin a number of times a second, be nearly as hot at its surface as the inside of the sun, and have a magnetic field a trillion times stronger than Earth's. Two questions should occur to you immediately. First, how could any theory predict such a wondrously unbelievable star? And second, do such neutron stars really exist?

Theoretical Prediction of Neutron Stars

The neutron was discovered in the laboratory in February 1932, and physicists were immediately fascinated by the new particle. Only two years later, in January 1934, two Caltech astronomers published a seminal paper. Walter Baade and Fritz Zwicky showed that some novae in historical records were much more luminous than most, and suggested that they were caused by the explosive collapse of a massive star in an explosion they called a supernova. The core of the star, they proposed, would form a small and tremendously dense sphere of neutrons, and Zwicky coined the term "neutron star."

Over the following years, scientists applied the principles of quantum mechanics to see if such an object was indeed possible. Neutrons spin in much the way that electrons do, which means that neutrons must obey the Pauli exclusion principle. That means that if neutrons are packed together tightly enough, they can become degenerate just as electrons do. White dwarfs are supported by degenerate electrons, and quantum mechanics predicts that an even denser mass of neutrons might support itself by the pressure of degenerate neutrons.

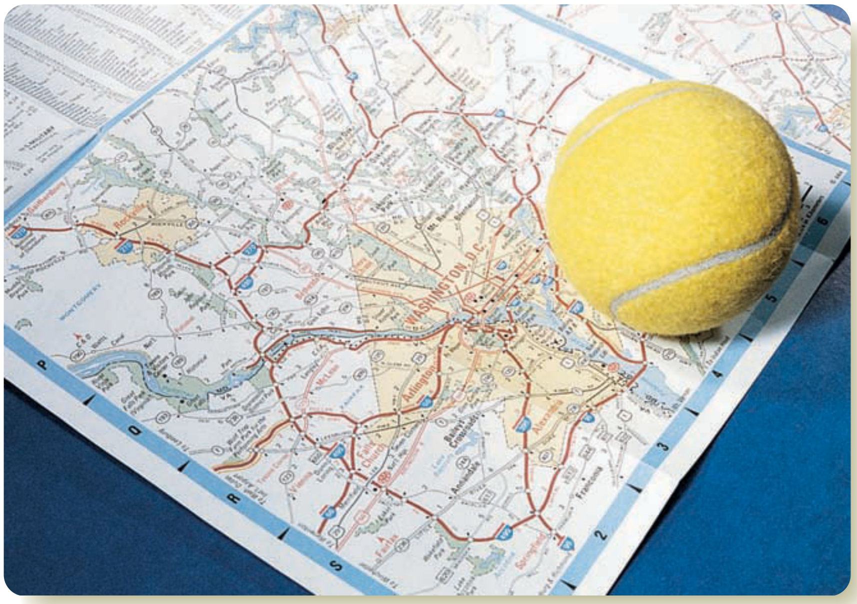
How does the core of a collapsing star become a mass of neutrons? Atomic physics provides an explanation. Although the outer parts of a star exploding as a supernova are blown outward, the core collapses inward. If the collapsing core is more massive than the Chandrasekhar limit of 1.4 solar masses, then it cannot reach stability as a white dwarf. The weight is too great to be supported by degenerate electrons. The collapse of the core continues, and the atomic nuclei are broken apart by gamma rays. Almost instantly, the increasing density forces the freed protons to combine with electrons and become neutrons. In a fraction of a second, the collapsing core becomes a contracting ball of neutrons. After the envelope of the star is blasted away, the core is left behind as a neutron star.

Which stars produce neutron stars? As you saw in the previous chapter, a star of 8 solar masses or less can lose enough mass to die by forming a planetary nebula and leaving behind a white dwarf. More massive stars will lose mass rapidly, but they cannot shed mass fast enough to reduce their mass below the Chandrasekhar limit, so it seems likely that they must die in supernova explosions. Theoretical calculations suggest that stars that begin life on the main sequence with 8 to roughly 20 solar masses will leave behind neutron stars. Stars more massive are thought to form black holes.

How massive can a neutron star be? That is a critical question and a difficult one to answer because scientists don't know the strength of pure neutron material. They can't make such matter in the laboratory, so its properties must be predicted theoretically. The most widely accepted calculations suggest that a neutron star cannot be more massive than 2 to 3 solar masses. If a neutron star were more massive than that, the degenerate neutrons would not be able to support the weight, and the object would collapse (presumably into a black hole). Two of the most massive neutron stars observed have masses of 1.94 and 2.74 solar masses, which confirms the theory.

How big are neutron stars? Mathematical models predict that a neutron star should be only 10 or so kilometers in radius (■ Figure 11-1), which, combined with a typical mass, means it must have a density of almost 10^{15} g/cm^3. On Earth, a sugar-cube-sized lump of this material would weigh 100 million tons. This is roughly the density of an atomic nucleus, so you can think of a neutron star as matter with all of the empty space squeezed out of it.

Simple physics, the same physics you have used in previous chapters to understand normal stars, predicts that neutron stars should be hot, spin rapidly, and have strong magnetic fields. You have seen that contraction heats the gas in a star. As gas particles fall inward, they pick up speed; and, when they collide, their high speeds become thermal energy. The sudden collapse of the core of a massive star to a radius of 10 km should heat it to millions of degrees. Furthermore, neutron stars should cool slowly because the heat can escape only from the surface, and neutron stars are so small they have little surface from which to radiate.



■ **Figure 11-1**

A tennis ball and a road map illustrate the relative size of a neutron star. Such an object, containing slightly more than the mass of the sun, would fit with room to spare inside the beltway around Washington, D.C. (M. Seeds)

The conservation of angular momentum predicts that neutron stars should spin rapidly. All stars rotate because they form from swirling clouds of interstellar matter. As a star collapses, it must rotate faster because it conserves angular momentum. Recall the example of an ice skater spinning slowly with her arms extended and then speeding up as she pulls her arms closer to her body (see Figure 10-11). In the same way, a collapsing star must spin faster as it pulls its matter closer to its axis of rotation. If the sun collapsed to a radius of 10 km, its period of rotation would increase from once every 25 days to over 20 times a second. You might expect the collapsed core of a massive star to rotate 10 or 100 times a second.

It isn't hard to understand why a neutron star should have a powerful magnetic field. The gas of a star is ionized, and that means the magnetic field cannot move easily through the gas. When the star collapses, the magnetic field is squeezed into a smaller area, which can make the field as much as a billion times stronger. Because some stars start with magnetic fields over 1000 times stronger than the sun's, a neutron star could have a magnetic field as much as a trillion times stronger than the sun's. For comparison, that is about 10 million times stronger than any magnetic field ever produced in the laboratory.

Theory predicts the properties of neutron stars, but it also predicts that they should be difficult to observe. Neutron stars are very hot, so from your understanding of black body radiation you can predict they will radiate most of their energy in the X-ray part of the spectrum, radiation that could not be observed in the 1940s and 1950s because astronomers could not put their telescopes above Earth's atmosphere. Also, the small surface areas of neutron stars mean that they will be faint objects. Consequently, astronomers of the mid-20th century were not surprised that none of the newly predicted neutron stars had been found. Neutron stars were, at that point, entirely theoretical objects.

The Discovery of Pulsars

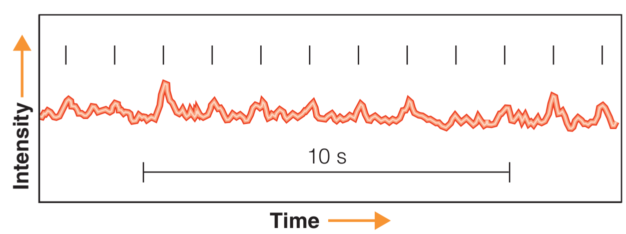
In November 1967, Jocelyn Bell, a graduate student at Cambridge University in England, found a peculiar pattern in the data from a radio telescope. Unlike other radio signals from celestial bodies, this was a series of regular pulses (■ Figure 11-2). At first she and the leader of the project, Anthony Hewish, thought the signal was interference, but they found it day after day in the same place in the sky. Clearly, it was celestial in origin.

Another possibility, that it came from a distant civilization, led them to consider naming it LGM, for Little Green Men. But within a few weeks, the team found three more objects in other parts of the sky pulsing with different periods. The objects were clearly natural, and the team dropped the name LGM in favor of **pulsar**—a contraction of *pulsing star.* The pulsing radio source Bell had observed with her radio telescope was the first known pulsar.

As more pulsars were found, astronomers argued over their nature. Periods ranged from 0.033 to 3.75 seconds and were nearly as exact as an atomic clock. Months of observation showed that many of the periods were slowly growing longer by a few billionths of a second per day. Whatever produced the regular pulses had to be highly precise, nearly as exact as an atomic clock, and gradually slowing down.

It was easy to eliminate possibilities. Pulsars could not be stars. A normal star, even a small white dwarf, is too big to pulse that fast. Nor could a star with a hot spot on its surface spin fast enough to produce the pulses. Even a small white dwarf would fly apart if it spun 30 times a second.

The pulses themselves gave the astronomers a clue. The pulses lasted only about 0.001 second, placing an upper limit on the size of the object producing the pulse. If a white dwarf

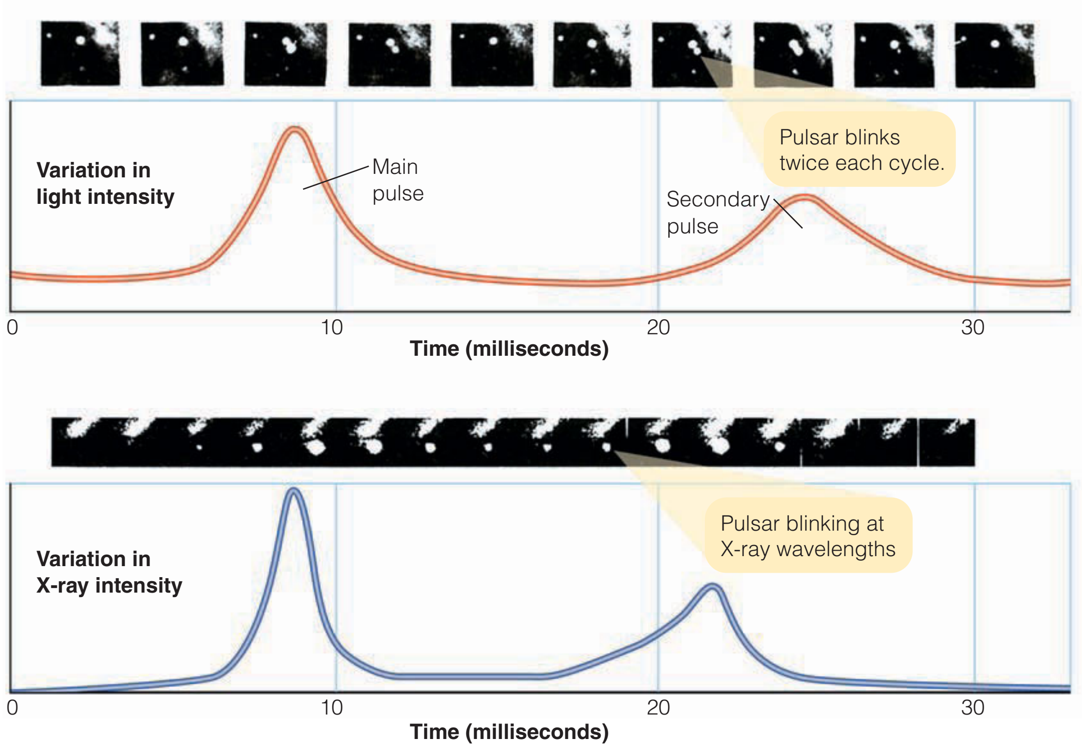


■ **Figure 11-2**

The 1967 detection of regularly spaced pulses in the output of a radio telescope led to the discovery of pulsars. This record of the radio signal from the first pulsar, CP1919, contains regularly spaced pulses (marked by ticks). The period is 1.33730119 seconds.

blinked on and then off in that interval, an observer would not see a 0.001-second pulse. That's because the near side of the white dwarf would be about 6000 km closer to Earth, and light from the near side would arrive 0.022 second before the light from the bulk of the white dwarf. Thus its short blink would be smeared out into a longer pulse. This is an important principle in astronomy—an object cannot change its brightness appreciably in an interval shorter than the time light takes to cross its diameter. If pulses from pulsars are no longer than 0.001 second, then the objects cannot be larger than 300 km (190 miles) in diameter.

Only a neutron star is small enough to be a pulsar. In fact, a neutron star is so small that it couldn't pulsate slowly enough, but it can spin as fast as 1000 times a second without flying apart. The missing link between pulsars and neutron stars was found in late 1968, when astronomers discovered a pulsar at the heart of the Crab Nebula (Figure 10-16). The Crab Nebula is a supernova remnant, and theory predicts that some supernovae leave behind a neutron star. The short pulses and the discovery of the pulsar in the Crab Nebula were strong evidence that pulsars are neutron stars.



■ **Figure 11-3**

High-speed images of the Crab Nebula pulsar show it pulsing at visual wavelengths and at X-ray wavelengths. (© AURA, Inc., NOAO, KPNO) The period of pulsation is 33 milliseconds, and each cycle includes two pulses as its two beams of unequal intensity sweep over Earth. (Courtesy F. R. Harnden, Jr., from *The Astrophysical Journal,* published by the University of Chicago Press; © 1984 The American Astronomical Society)

A Model Pulsar

Scientists often work by building a model of a natural phenomenon—not a physical model made of plastic and glue, but an intellectual conception of how nature works in a specific instance. The astronomer's model may be limited and incomplete, but it helps them organize their understudying.

The modern model of a pulsar has been called the **lighthouse model** and is shown in ■ **The Lighthouse Model of a Pulsar** on pages 214–215. Notice three important points:

1. A pulsar does not pulse but rather emits beams of radiation that sweep around the sky as the neutron star rotates. If the beams do not sweep over Earth, the pulses will not be detectable by Earth's radio telescopes.
2. The mechanism that produces the beams involves extremely high energies and is not fully understood.
3. Modern space telescopes observing from above Earth's atmosphere can image details around young neutron stars and even locate isolated neutron stars whose beams of electromagnetic radiation do not sweep over Earth.

Neutron stars are not simple objects, and modern astronomers need both general relativity and quantum mechanics to try to understand them. Nevertheless, astronomers know enough to tell the life story of pulsars.

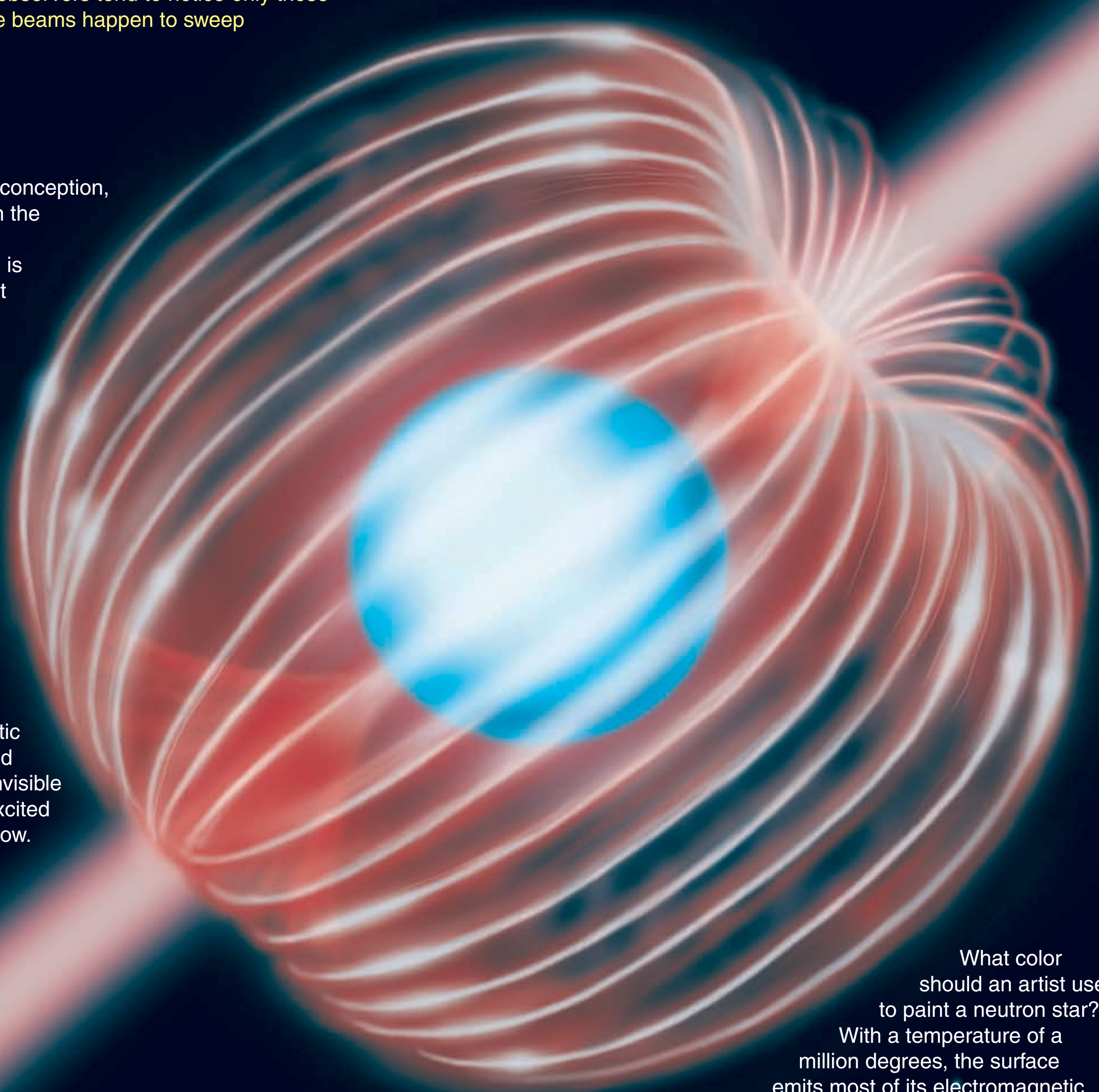
The Evolution of Pulsars

When a pulsar first forms, it is spinning fast, perhaps a hundred times a second. The energy it radiates into space comes from its energy of rotation, so as it blasts beams of radiation outward, its rotation slows. The average pulsar is apparently only a few million years old, and the oldest are about 10 million years old. Presumably, older neutron stars rotate too slowly to generate detectable radio beams.

You can expect that a young neutron star should emit powerful beams of radiation. The Crab Nebula is an example. Only about 950 years old, the Crab pulsar is so powerful it emits photons of radio, infrared, visible, X-ray, and gamma-ray wavelengths (■ Figure 11-3). Careful measurements of its brightness with high-speed instruments show that it blinks twice for every rotation. When one beam sweeps almost directly over Earth, astronomers detect a strong pulse. Half a rotation later, the edge of the other beam brushes over Earth, and astronomers detect a weaker pulse.

The Lighthouse Model of a Pulsar

1 Astronomers think of pulsars not as pulsing objects, but rather as objects emitting beams. As they spin, the beams sweep around the sky; when a beam sweeps over Earth, observers detect a pulse of radiation. Understanding the details of this lighthouse model is a challenge, but the implications are clear. Although a neutron star is only a few kilometers in radius, it can produce powerful beams. Also, observers tend to notice only those pulsars whose beams happen to sweep over Earth.



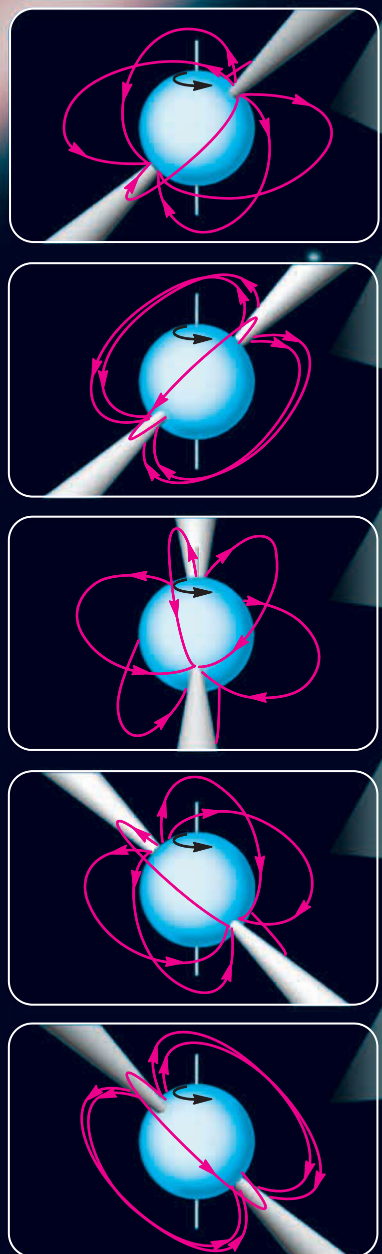
In this artist's conception, gas trapped in the neutron star's magnetic field is excited to emit light and outline the otherwise invisible magnetic field.

Beams of electromagnetic radiation would probably be invisible unless they excited local gas to glow.

What color should an artist use to paint a neutron star? With a temperature of a million degrees, the surface emits most of its electromagnetic radiation at X-ray wavelengths. Nevertheless, it would probably look blue-white to your eyes.

2 How a neutron star can emit beams is one of the challenging problems of modern astronomy, but astronomers have a general idea. A neutron star contains a powerful magnetic field and spins very rapidly. The spinning magnetic field generates a tremendously powerful electric field, and the field causes the production of electron–positron pairs. As these charged particles are accelerated through the magnetic field, they emit photons in the direction of their motion, which produce powerful beams of electromagnetic radiation emerging from the magnetic poles.

Neutron Star Rotation with Beams



As in the case of Earth, the magnetic axis of a neutron star could be inclined to its rotational axis.

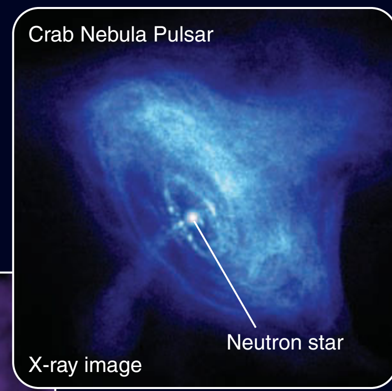
The rotation of the neutron star will sweep its beams around like beams from a lighthouse.

While a beam points roughly toward Earth, observers detect a pulse.

While neither beam is pointed toward Earth, observers detect no energy.

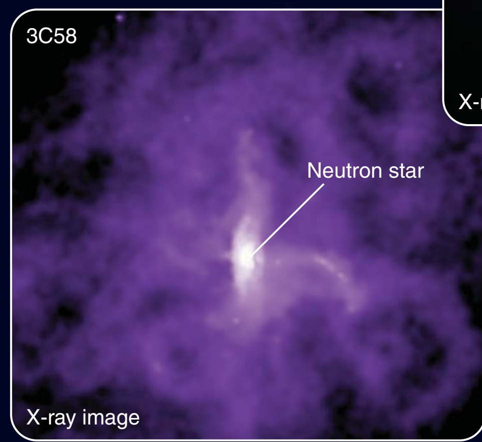
Beams may not be as exactly symmetric as in this model.

3 X-ray observations of young pulsars show that they are surrounded by disks of excited matter and emit powerful jets of excited gas. The disks and jets are shaped by electromagnetic fields and the jets may curve if they encounter magnetic fields.



NASA/CXC/ASU/J. Hester et al.

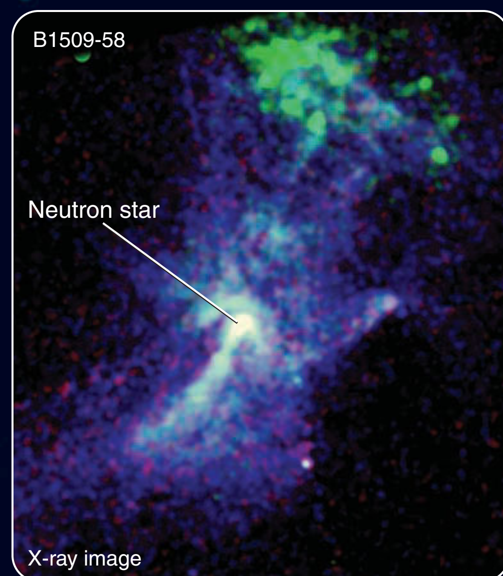
The Crab Nebula Pulsar above was produced by the supernova of AD 1054. It pulses 30 times a second, and x-ray images can detect rapid flickering changes in the disk.



NASA/CXC/SAO/P. Slane et al.

The pulsar 3C58 above was produced by the supernova seen in AD 1181. It pulses 15 times per second and is ejecting jets in both directions.

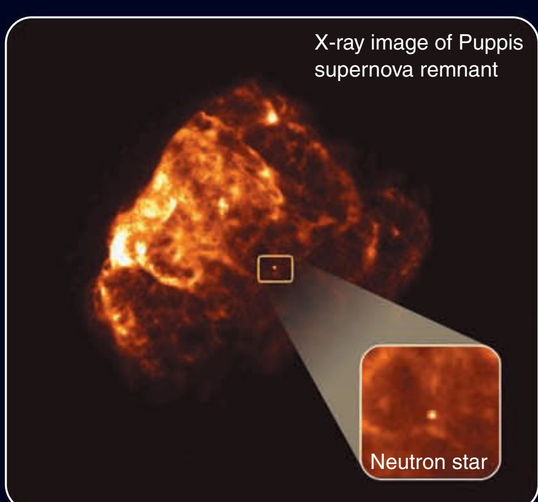
The pulsar B1509-58 at right is only 1700 years old, young for a pulsar. It is ejecting a thin jet almost 20 ly long. The nebulosity at the top of the image is part of the enclosing supernova remnant excited by energy from the pulsar.



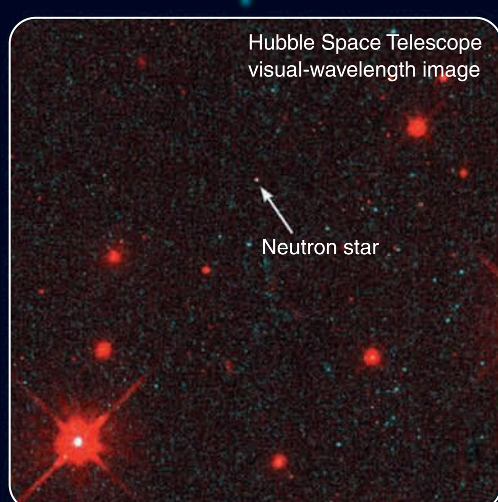
NASA/CXC/SAO/B. Gaensler et al.

CENGAGENOW™

Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see the Active Figure called "Neutron Star." Adjust the inclination of the neutron star's magnetic field to produce pulses.



NASA



NASA

3a If a pulsar's beams do not sweep over Earth, observers detect no pulses, and the neutron star is difficult to find. A few such objects are known, however. The Puppis A supernova remnant is about 4000 years old and contains a point source of X-rays believed to be a neutron star. The isolated neutron star in the right-hand image has a temperature of 700,000 K.

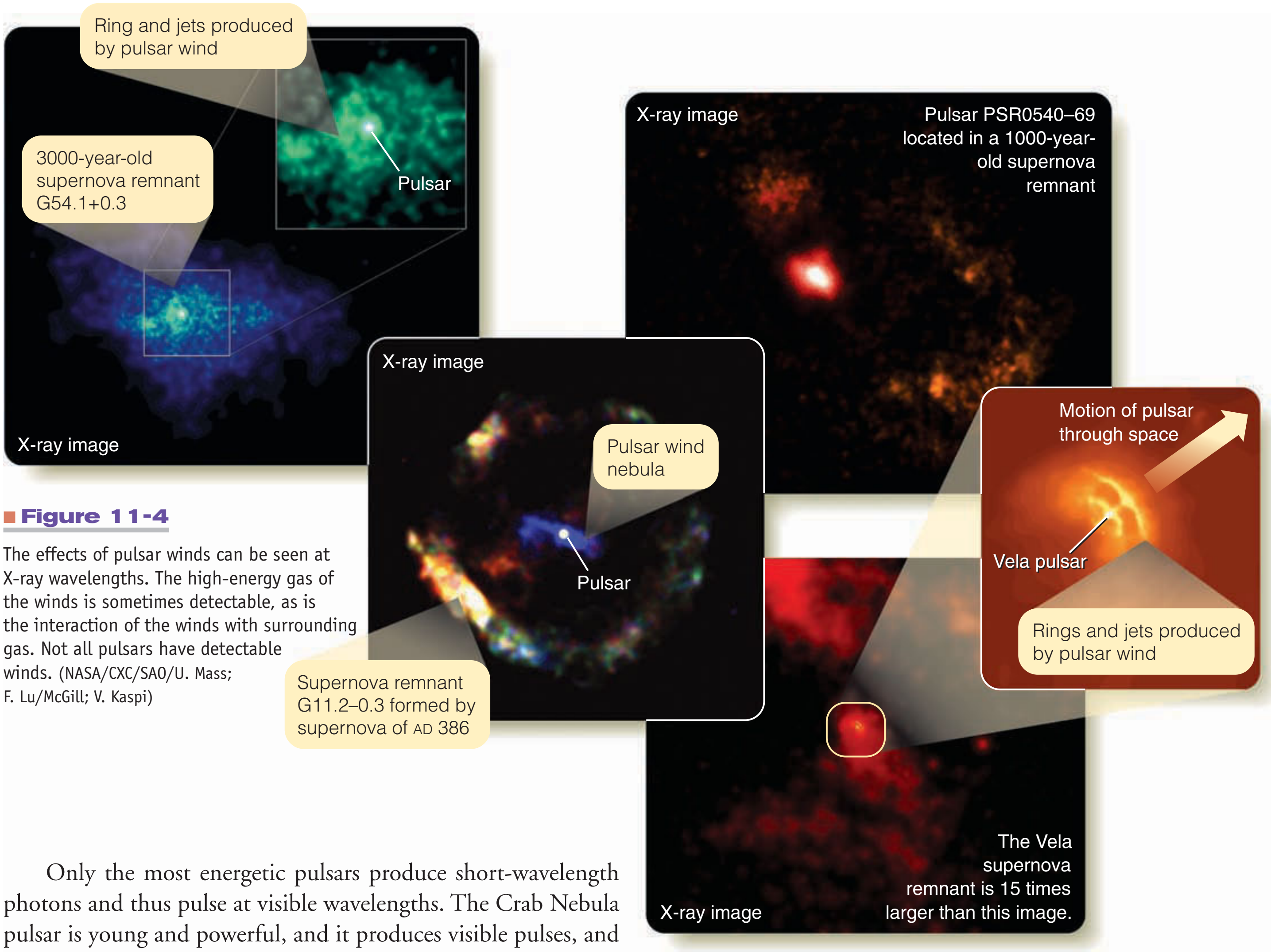


Figure 11-4

The effects of pulsar winds can be seen at X-ray wavelengths. The high-energy gas of the winds is sometimes detectable, as is the interaction of the winds with surrounding gas. Not all pulsars have detectable winds. (NASA/CXC/SAO/U. Mass; F. Lu/McGill; V. Kaspi)

Only the most energetic pulsars produce short-wavelength photons and thus pulse at visible wavelengths. The Crab Nebula pulsar is young and powerful, and it produces visible pulses, and so does another young pulsar called the Vela pulsar (located in the Southern Hemisphere constellation Vela). Compared with most pulsars, the Vela pulsar is fast, pulsing about 11 times a second and, like the Crab Nebula pulsar, is located inside a supernova remnant. Its age is estimated at a relatively young 20,000 to 30,000 years.

The energy in the beams is only a small part of the energy emitted by a pulsar. Roughly 99.9 percent of the energy flowing away from a pulsar is carried as a **pulsar wind** of high-speed atomic particles. This can produce small, high-energy nebulae near a young pulsar (■ Figure 11-4).

You might expect to find all pulsars inside supernova remnants, but the statistics must be examined with care. Not every supernova remnant contains a pulsar, and not every pulsar is located inside a supernova remnant. Many supernova remnants probably do contain pulsars, but their beams never sweep over Earth. Also, some pulsars move through space at high velocity (■ Figure 11-5), quickly leaving their supernova remnants behind. Evidently supernova explosions can occur asymmetrically, perhaps because of the violent turbulence in the exploding core, and that can give a neutron star a high velocity through space.

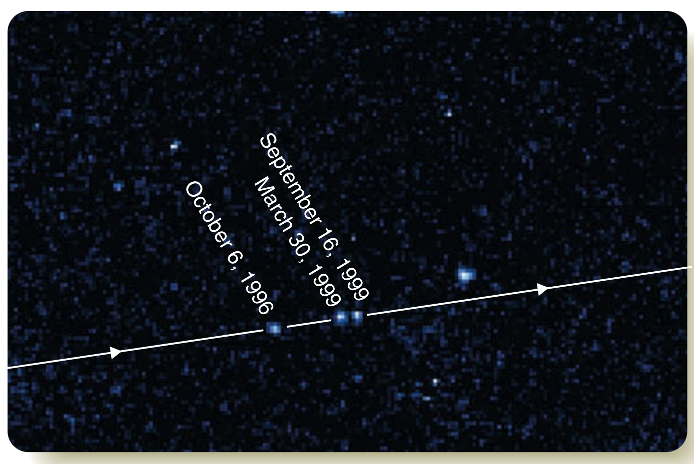


Figure 11-5

Many neutron stars have high velocities through space. Here the neutron star known as RXJ185635-3754 was photographed on three different dates as it rushed past background stars. (V. Kaspi/NASA)

Some supernovae probably occur in binary systems and fling the two stars apart at high velocity. In any case, pulsars are known to have such high velocities that many probably escape the disk of our galaxy. Finally, you must remember that a pulsar can remain detectable for 10 million years or so, but a supernova remnant cannot survive more than about 50,000 years before it is mixed into the interstellar medium. For all these reasons, you should not be surprised that most pulsars are not in supernova remnants and that most supernova remnants do not contain pulsars.

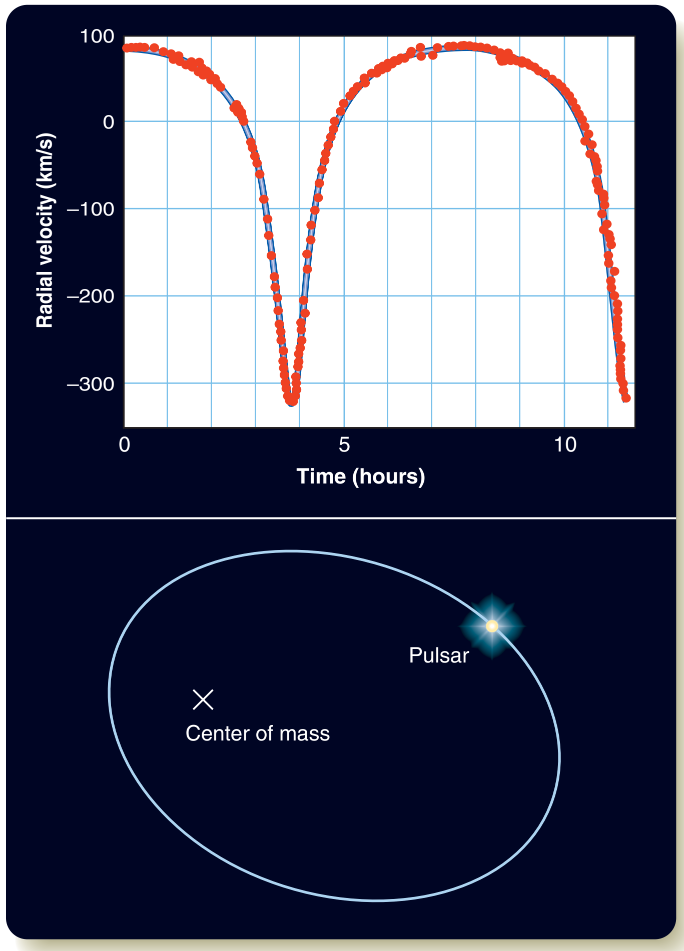
Astronomers conclude that the February 1987 explosion of Supernova 1987A formed a neutron star. As you read in the previous chapter, a burst of neutrinos was detected passing through Earth; theory predicts that the collapse of a massive star's core into a neutron star produces such a burst of neutrinos, so the detection of the neutrinos is evidence that the supernova produced a neutron star. At first the neutron star will be hidden at the center of the expanding shells of gas ejected into space; but, as the gas expands and thins, astronomers might be able to see the neutron star. If its beams don't sweep over Earth, astronomers should eventually be able to detect its X-ray and gamma-ray emission. Although no neutron star has yet been detected, astronomers continue to watch the site, hoping to see a newborn pulsar.

One reason pulsars are so fascinating is the extreme conditions found in spinning neutron stars. To see natural processes of even greater violence, you have only to look at pulsars in binary systems.

Binary Pulsars

Over 1000 pulsars are now known, and some are located in binary systems. These pulsars are of special interest because astronomers can learn more about the neutron star by studying the orbital motions of the binary. Also, in some cases, mass can flow from the companion star onto the neutron star, and that produces high-energy violence.

The first binary pulsar was discovered in 1974 when astronomers Joseph Taylor and Russell Hulse noticed that the pulse period of the pulsar PSR1913+16 was changing. The period first grew longer and then grew shorter in a cycle that took 7.75 hours. Thinking of the Doppler shifts seen in spectroscopic binaries, the radio astronomers realized that the pulsar had to be in a binary system with an orbital period of 7.75 hours. When the orbital motion of the pulsar carries it away from Earth, astronomers see the pulse period lengthen slightly—a redshift. Then, when the pulsar rounds its orbit and approaches Earth, they see the pulse period shorten slightly—a blueshift. From these changing Doppler shifts, Taylor and Hulse could calculate the radial velocity of the pulsar around its orbit just as if it were a spectroscopic binary star. The resulting graph of radial velocity versus time could be analyzed to find the shape of the pulsar's orbit (■ Figure 11-6). The analysis of PSR1913+16 showed that the binary system consists of two neutron stars separated by a distance roughly equal to the radius of our sun.



■ Figure 11-6

The radial velocity of pulsar PSR1913+16 can be found from the Doppler shifts in its pulsation. Analysis of the radial velocity curve allows astronomers to determine the pulsar's orbit. Here the center of mass does not appear to be at a focus of the elliptical orbit because the orbit is inclined. (Adapted from data by Joseph Taylor and Russell Hulse)

Yet another surprise was hidden in the motion of PSR1913+16. In 1916, Einstein's general theory of relativity described gravity as a curvature of space-time. Einstein realized that any rapid change in a gravitational field should spread outward at the speed of light as **gravitational radiation.** Gravity waves have not been detected yet, but Taylor and Hulse were able to show that the orbital period of the binary pulsar is slowly growing shorter because the stars are radiating orbital energy away as gravitational radiation and gradually spiraling toward each other. (Normal binary stars are too far apart and orbit too slowly to emit significant gravitational radiation.) Taylor and Hulse won the Nobel Prize in 1993 for their work with binary pulsars.

Dozens of pulsars have been found orbiting stars of various kinds; by analyzing the Doppler shifts in their pulse periods, astronomers can estimate the masses of the neutron stars. Typical masses are about 1.35 solar masses, in good agreement with models of neutron stars.

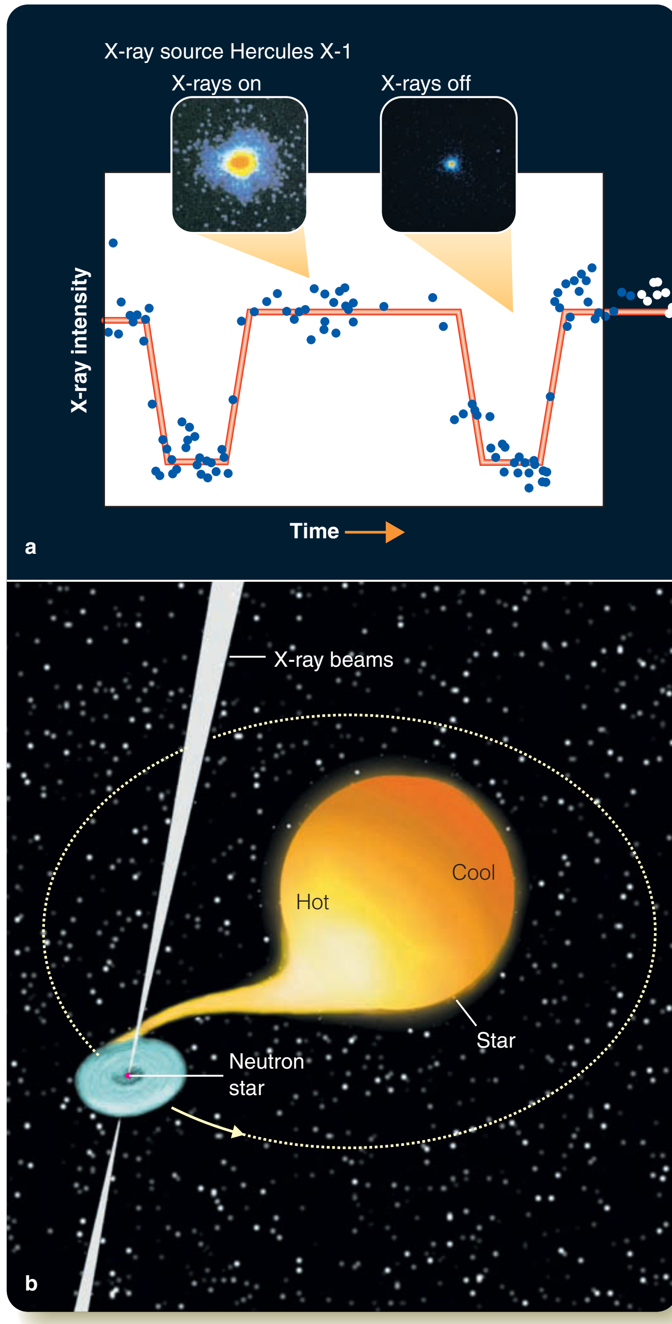
In 2004, radio astronomers announced the discovery of a double pulsar: two pulsars that orbit each other in only 2.4 hours. The spinning beams from both pulsars sweep over Earth (■ Figure 11-7). One spins 44 times a second, and the other spins once every 2.8 seconds. This system is a pulsar jackpot because the orbits are nearly edge-on to Earth and the powerful magnetic fields and the gas trapped in the fields eclipse each other, giving astronomers a chance to study their size and structure. Furthermore, the theory of general relativity predicts that these pulsars are emitting gravitational radiation and that their separation is decreasing by 7 mm per year. The two neutron stars will merge in 85 million years, presumably to trigger a violent explosion. In the meantime, the steady decrease in orbital period can be measured and gives astronomers a further test of general relativity and gravitational radiation.

In addition to producing gravitational radiation, a neutron star's intense gravitational field means that binary pulsars can be sites of tremendous violence if matter is transferred from a star to a neutron star. The gravitational field is so strong that an astronaut stepping onto the surface of a neutron star would be instantly smushed into a layer of matter only 1 atom thick. Matter falling into this gravitational field releases titanic amounts of energy. If you dropped a single marshmallow onto the surface of a neutron star from a distance of 1 AU, it would hit with an impact equivalent to a 3-megaton nuclear warhead. In general, a particle falling from a large distance to the surface of a neutron star will release energy equivalent to 0.2 mc^2, where m is the particle's mass at rest. Even a small amount of matter flowing from a companion star to a neutron star can generate high temperatures and release X-rays and gamma rays.

■ **Figure 11-7**

Artist's impression of the double pulsar. One star must have exploded to form a pulsar, and later the other star did the same. Gravitational radiation causes the neutron stars to spiral toward each other, and they will merge in 85 million years, presumably to trigger a violent explosion. (John Rowe Animations)

Hercules X-1 is an example of such an active system; it contains a 2-solar-mass star and a neutron star that orbit each other with a period of 1.7 days (■ Figure 11-8). Matter flowing from the normal star into an accretion disk around the neutron star reaches temperatures of millions of degrees and emits a powerful X-ray glow. Interactions with the neutron star's magnetic field



■ **Figure 11-8**

(a) Sometimes the X-ray pulses from Hercules X-1 are on, and sometimes they are off. A graph of X-ray intensity versus time looks like the light curve of an eclipsing binary. (Insets: J. Trümper, Max-Planck Institute) (b) In Hercules X-1, matter flows from a star into an accretion disk around a neutron star producing X-rays, which heat the near side of the star to 20,000 K compared with only 7000 K on the far side. X-rays turn off when the neutron star is eclipsed behind the star.

produce beams of X-rays that sweep around with the rotating neutron star (Figure 11-8b). Earth receives a pulse of X-rays every time a beam points this way. The X-rays shut off completely every 1.7 days when the neutron star is eclipsed behind the normal star. Hercules X-1 is a complex system with many different high-energy processes going on simultaneously, but this quick analysis serves to illustrate how complex and powerful such binary systems are during mass transfer.

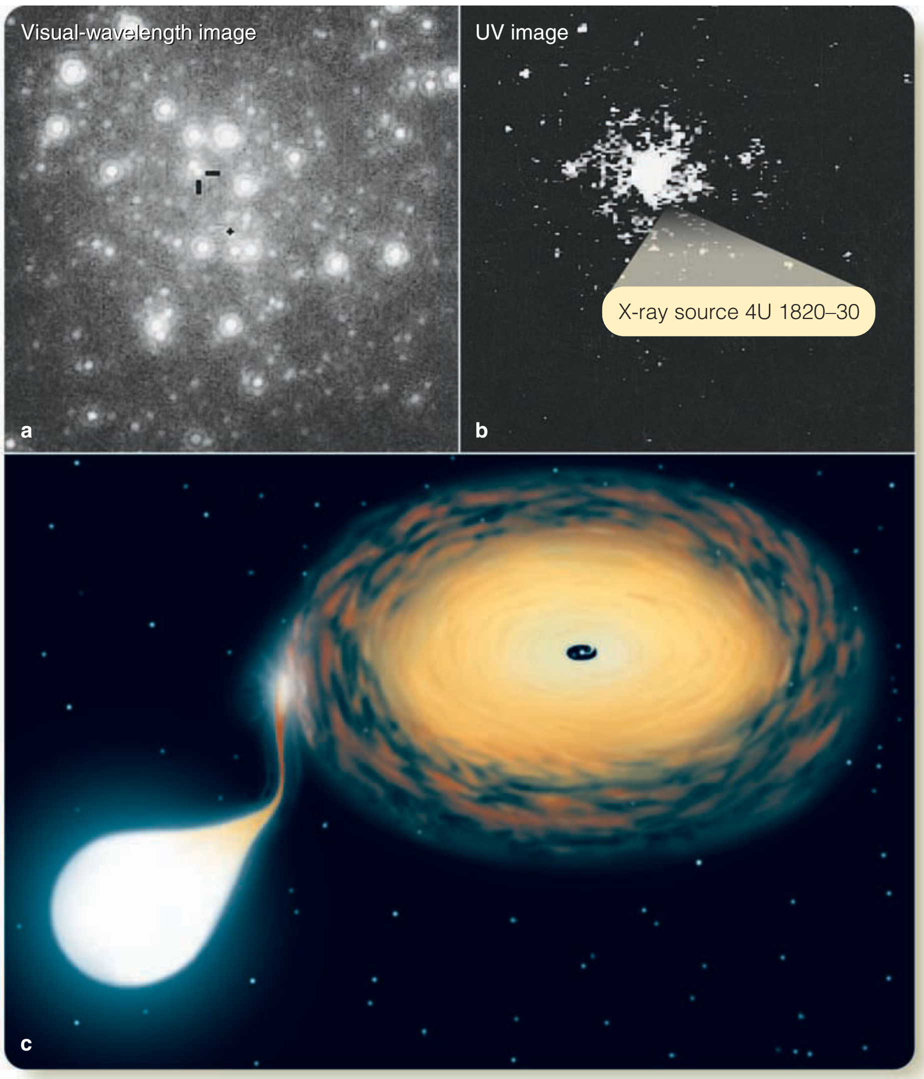
The X-ray source 4U 1820-30 illustrates another way neutron stars can interact with normal stars. In this system, a neutron star and a white dwarf orbit their center of mass with a period of only 11 minutes (■ Figure 11-9a and b). The separation between the two objects is only about one-third the distance between Earth and the moon. To explain how such a very close pairing could originate, theorists suggest that a neutron star collided with a giant star and went into an orbit inside the star. (Recall the low density of the outer envelope of giant stars.) The neutron star would have gradually eaten away the giant star's envelope from the inside, leaving the white dwarf behind. Matter still flows from the white dwarf into an accretion disk and then down to the surface of the neutron star (Figure 11-9c), where it accumulates until it ignites to produce periodic bursts of X-rays. Objects called **X-ray bursters** are thought to be such binary systems involving mass transferred to a neutron star. Notice the similarity between this mechanism and that responsible for novae (Chapter 10).

The Fastest Pulsars

Your knowledge of pulsars suggests that newborn pulsars should blink rapidly, and old pulsars should blink slowly. In fact, the handful that blink the fastest may be quite old. One of the fastest known pulsars is cataloged as Terzan 5ad in the constellation Sagittarius. It pulses 716 times a second and is slowing down only slightly. The energy stored in the rotation of a neutron star at this rate is equal to the total energy of a supernova explosion, so it seemed difficult at first to explain this pulsar. It now appears that Terzan 5ad is an old neutron star that has gained mass and rotational energy from a companion in a binary system. Like water hitting a mill wheel, the matter falling on the neutron star has spun it up to 716 rotations per second. With its weak magnetic field, it slows down very gradually and will continue to spin for a very long time.

A number of other very fast pulsars have been found. They are known generally as **millisecond pulsars** because their pulse periods and therefore their periods of rotation are almost as short as a millisecond (0.001 s). This rapid rotation produces some fascinating physics. If a neutron star 10 km in radius spins 716 times a second, as does Terzan 5ad, then its period is 0.0014 second, and the equator of the neutron star must be traveling about 45,000 km/s. That is fast enough to flatten the neutron star into an ellipsoidal shape and is nearly fast enough to break it up. The fastest known pulsar, XTEJ1739-285, spins 1122 times a second. It appears to have been spun up by matter flowing from its companion star and is very near the breakup speed for an object composed of pure neutrons.

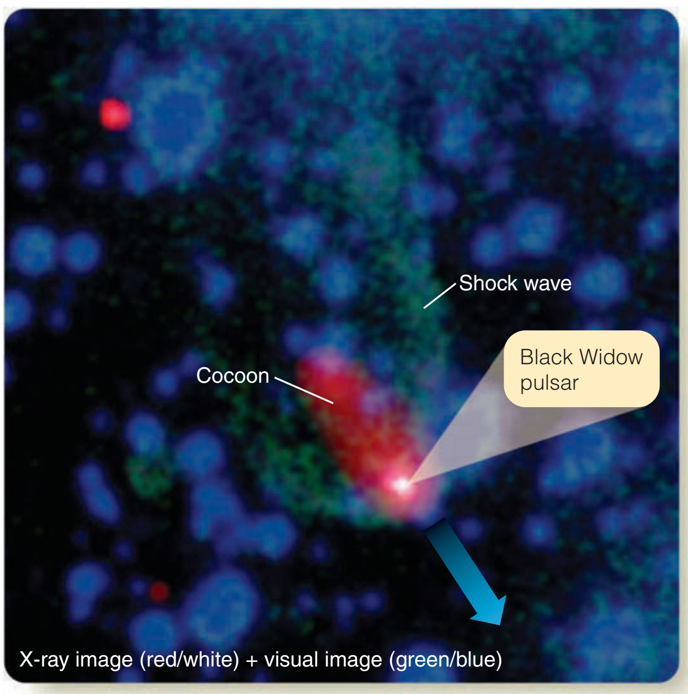
All scientists should be made honorary citizens of Missouri, the "Show Me" state, because scientists demand evidence. The hypothesis that the millisecond pulsars are spun



■ **Figure 11-9**

(a) At visible wavelengths, the center of star cluster NGC 6624 is crowded with stars. (b) In the ultraviolet, one object stands out, an X-ray source consisting of a neutron star orbiting a white dwarf. (c) An artist's conception shows matter flowing from the white dwarf into an accretion disk around the neutron star. (a and b, Ivan King and NASA/ESA; c, Dana Berry, STScI) up by mass transfer from a companion star is quite reasonable, but astronomers demand evidence, and evidence has been found. For example, the pulsar PSRJ1740-5340 has a period of 42 milliseconds and is orbiting with a bloated red star from which it is gaining mass. This appears to be a pulsar caught in the act of being spun up to high speed. For another example, consider the X-ray source XTEJ1751-305, a pulsar with a period of only 2.3 milliseconds. X-ray observations show that it is gaining mass from a companion star. The orbital period is only 42 minutes, and the mass of the companion star is only 0.014 solar mass. The evidence suggests this neutron star has devoured all but the last morsel of its binary partner.

Although some millisecond pulsars have binary companions, some are solitary neutron stars. A pulsar known as the Black Widow may reveal how a fast pulsar can lack a companion. The Black Widow has a period of 1.6 milliseconds, meaning it is spinning 622 times per second, and it orbits with a low-mass companion. Presumably the neutron star was spun up by mass flowing from the companion, but spectra show that the blast of radiation and high-energy particles from the neutron star is now boiling away the surface of the companion. The Black Widow has eaten its fill and is now evaporating the remains of its companion. It will soon be a solitary millisecond pulsar (■ Figure 11-10).



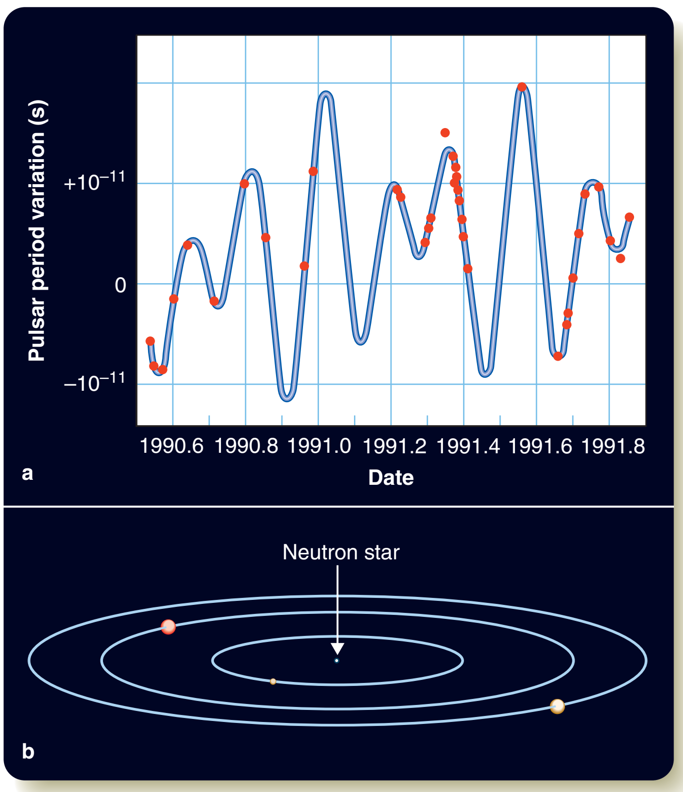
■ Figure 11-10

The Black Widow pulsar and its companion star are moving rapidly through space, creating a shock wave like the bow wave of a speedboat. The shock wave confines high-energy particles shed by the pulsar into an elongated cocoon (red). (X-ray: NASA/CXC/ASTRON/B. Stappers et al.; Optical: AAO/J. Bland- Hawthorn & H. Jones)

"Show me," say scientists; and, in the case of neutron stars, the evidence seems very strong. Of course, you can never prove that a theory is absolutely true (■ How Do We Know? 11-1), but the evidence for neutron stars is so strong that astronomers have great confidence that they really do exist. Other theories that describe how they emit beams of radiation and how they form and evolve are less certain, but continuing observations at many wavelengths are expanding astronomers' understanding of these last embers of massive stars. In fact, precise observations have turned up objects no one expected.

Pulsar Planets

Because a pulsar's period is so precise, astronomers can detect tiny variations by comparing their observations with atomic clocks. When astronomers checked pulsar PSR1257+12, they found variations in the period of pulsation much like those caused by the orbital motion of a binary pulsar (■ Figure 11-11a). However, in the case of PSR1257+12, the variations were much smaller; and, when they were interpreted as Doppler shifts, it became evident that the pulsar was being orbited by at least two objects



■ Figure 11-11

(a) The dots in this graph are observations showing that the period of pulsar PSR1257+12 varies from its average value by a fraction of a billionth of a second. The blue line shows the variation that would be produced by planets orbiting the pulsar. (b) As the planets orbit the pulsar, they cause it to wobble by less than 800 km, a distance that is invisibly small in this diagram. (Adapted from data by Alexander Wolszczan)

How Do We Know? 11-1

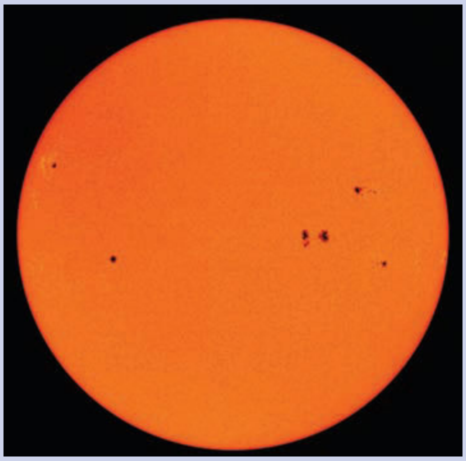
Theories and Proof

Why do astronomers say a theory is confirmed but never say that it is proven? People say dismissively of a theory they dislike, "That's only a theory," as if a theory were just a random guess. In fact, a theory can be a well-tested truth in which all scientists have great confidence. Yet no matter how many tests and experiments you conduct, you can never prove that any scientific theory is absolutely true. It is always possible that the next observation you make will disprove the theory.

There have always been theories about why the sun is hot. Some astronomers once thought the sun was a ball of burning coal, and over a century ago most astronomers accepted the theory that the sun was hot because gravity was making it contract. In the late 19th century, geologists showed that Earth was much older than the sun could be if it was powered by gravity, so the gravity theory had to be wrong. It wasn't until 1920 that another promising theory was proposed by Sir Arthur Eddington, who suggested the sun was powered somehow by the energy in atomic nuclei. In 1938 the German-American astrophysicist Hans Bethe showed how nuclear fusion could power the sun. He won the Nobel Prize in 1967.

No one will ever go to the center of the sun, so you can't *prove* the fusion theory is right. Many observations and model calculations support this theory, and in the chapter on the sun you saw further evidence in the neutrinos that have been detected coming from the sun's core. Nevertheless, there remains some tiny possibility that all the observations and models are misunderstood and that the theory will be overturned by some future discovery. Astronomers have tremendous confidence that the sun is powered by fusion and not gravity or coal, but a scientific theory can never be proven conclusively correct.

There is a great difference between a theory that is a far-fetched guess and a scientific theory that has undergone decades of testing and confirmation with observations, experiments, and models. But no theory can ever be proven absolutely true. It is up to you as a consumer of knowledge and a responsible citizen to distinguish between a flimsy guess and a well-tested theory that deserves to be treated like truth—at least pending further information.



Technically it is still a theory, but astronomers have tremendous confidence that the sun gets its power from nuclear fusion and not from burning coal. (SOHO/MDI)

with planetlike masses of 4.1 and 3.8 Earth masses. The gravitational tugs of these planets make the pulsar wobble about the center of mass of the system by no more than 800 km, producing the tiny changes in period (Figure 11-11b).

Astronomers greeted this discovery with both enthusiasm and skepticism. As usual, they looked for ways to test the hypothesis. Simple gravitational theory predicts that planets in the same system should interact and slightly modify each other's orbits. When the data were analyzed, that interaction was found, further confirming the existence of the planets. In fact, later data revealed the presence of a third planet of about twice the mass of Earth's moon. This illustrates the astonishing precision of studies based on pulsar timing.

Astronomers wonder how a neutron star can have planets. The three planets that orbit PSR1257+12 are closer to the pulsar than Venus is to the sun. Any planets that orbited a star that closely would have been absorbed or vaporized when the star expanded to become a supergiant. Furthermore, the supernova explosion would have suddenly reduced the mass of the star and allowed any orbiting planets to escape from their orbits. It seems more likely that these planets are the remains of a stellar companion that was devoured by the neutron star. In fact, PSR1257+12 spins very fast (161 pulses per second), suggesting that it was spun up in a binary system.

PSR1257+12 is not unique. Another planet has been found orbiting a neutron star that is part of a binary system with a white dwarf. Because this system is located in a very old star cluster and contains a white dwarf, astronomers suspect that the planet may be very old. Planets probably orbit other neutron stars, and small shifts in the timing of the pulses may eventually reveal their presence.

You can imagine what these worlds might be like. Formed from the remains of elderly stars, they might have chemical compositions richer in heavy elements than Earth. You can imagine visiting these worlds, landing on their surfaces, and hiking across their valleys and mountains. Above you, the neutron star would glitter in the sky, a tiny point of light.

◄ SCIENTIFIC ARGUMENT ►

Why are neutron stars easier to detect at X-ray wavelengths?
This argument draws together a number of ideas you know from previous chapters. First, recall that a neutron star is very hot because of the heat released when it contracts to a radius of 10 km. It could easily have a surface temperature of 1,000,000 K, and Wien's law (Chapter 6) tells you that such an object will radiate most intensely at a very short wavelength—X-rays. Normal stars are much cooler and emit only weak X-rays unless they have hot accretion disks. At visual wavelengths, stars are bright, and neutron stars are faint, but at X-ray wavelengths, the neutron stars stand out from the crowd.

Now build a new argument as if you were seeking funds for a research project. **What observations would you make to determine whether a newly discovered pulsar was young or old, single or a member of a binary system, alone or accompanied by planets?**

◄ ►

11-2 Black Holes

You have now studied white dwarfs and neutron stars, two of the three end states of dying stars. Now it's time to think about the third end state—black holes.

Although the physics of black holes is difficult to discuss without using sophisticated mathematics, simple logic is sufficient to predict that they should exist. The problem is to confirm that they are real. What objects observed in the heavens could be real black holes? More difficult than the search for neutron stars, the quest for black holes has nevertheless met with success.

To begin, you must consider a simple question. How fast must an object travel to escape from the surface of a celestial body?

Escape Velocity

Suppose you threw a baseball straight up. How fast must you throw it if it is not to come down? Of course, gravity will always pull back on the ball, slowing it, but if the ball is traveling fast enough to start with, it will never come to a stop and fall back.

Such a ball will escape from Earth. In Chapter 4 you learned that the escape velocity is the initial velocity an object needs to escape from a celestial body (■ Figure 11-12). Whether you are discussing a baseball leaving Earth or a photon leaving a collapsing star, escape velocity depends on two things, the mass of the celestial body and the distance from the center of mass to the escaping object. If the celestial body has a large mass, its gravity is strong, and you need a high velocity to escape; but, if you begin your journey farther from the center of mass, the velocity needed is less. For example, to escape from Earth, a spaceship would have to leave Earth's surface at 11 km/s (25,000 mph), but if you could launch spaceships from the top of a tower 1000 miles high, the escape velocity would be only 10 km/s (22,000 mph).

If you could make an object massive enough or small enough, its escape velocity could be greater than the speed of light. Relativity says that nothing can travel faster than the speed of light, so even photons, which have no mass, would be unable to escape. Such a small, massive object could never be seen because light could not leave it.

Long before Einstein and relativity, the Reverend John Mitchell, a British gentleman astronomer, realized the peculiar consequences of Newton's laws of gravity and motion. In 1783, he pointed out that an object 500 times the radius of the sun but of the same density would have an escape velocity greater than the speed of light. Then, "all light emitted from such a body would be made to return towards it." Mitchell didn't know it, but he was talking about a black hole.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Black Hole."

■ **Figure 11-12**

Escape velocity, the velocity needed to escape from a celestial body, depends on mass. The escape velocity at the surface of a very small body would be so low you could jump into space. Earth's escape velocity is much larger, about 11 km/s (25,000 mph).

Schwarzschild Black Holes

If the core of a star contains more than 3 solar masses when it collapses, no force can stop it. It cannot stop collapsing when it reaches the density of a white dwarf because degenerate electrons cannot support the weight, and it cannot stop when it reaches the density of a neutron star because not even degenerate neutrons can support the weight. No force remains to stop the object from collapsing to zero radius.

As an object collapses, its density and the strength of its surface gravity increase; and if an object collapses to zero radius, its density and gravity become infinite. Mathematicians call such a point a **singularity,** but in physical terms it is difficult to imagine an object of zero radius. Some theorists believe that a singularity is impossible and that the laws of quantum physics must somehow halt the collapse at some subatomic radius roughly 10^{20} times smaller than a proton. Astronomically, it seems to make little difference.

If the contracting core of a star becomes small enough, the escape velocity in the region of space around it is so large that no light can escape. This means you can receive no information about the object or about the region of space near it. Because it emits no light, such a region is called a **black hole.** If the core of an exploding star collapsed into a black hole, the expanding outer layers of the star could produce a supernova remnant, but the core would vanish without a trace.

To understand black holes, you must consider relativity. In 1916, Albert Einstein published a mathematical theory of space and time that became known as the general theory of relativity. Einstein treated space and time as a single entity called space-time. His equations showed that gravity could be described as a curvature of space-time, and almost immediately the astronomer Karl Schwarzschild found a way to solve Einstein's equations to describe the gravitational field around a single, nonrotating, electrically neutral lump of matter. That solution contained the first general relativistic description of a black hole, and nonrotating, electrically neutral black holes are now known as Schwarzschild black holes. In recent decades, theorists such as Roy Kerr and Stephen Hawking have found ways to apply the sophisticated mathematical equations of the general theory of relativity and quantum mechanics to charged, rotating black holes. For this discussion, the differences are minor, and you may proceed as if all black holes were Schwarzschild black holes.

Schwarzschild's solution shows that if matter is packed into a small enough volume, then space-time curves back on itself. Objects can still follow paths that lead into the black hole, but no path leads out, so nothing can escape. Because not even light can escape, the inside of the black hole is totally beyond the view of an outside observer. The **event horizon** is the boundary between the isolated volume of space-time and the rest of the universe, and the radius of the event horizon is called the **Schwarzschild radius,** R_S. A collapsing stellar core must shrink inside its Schwarzschild radius to become a black hole (■ Figure 11-13).

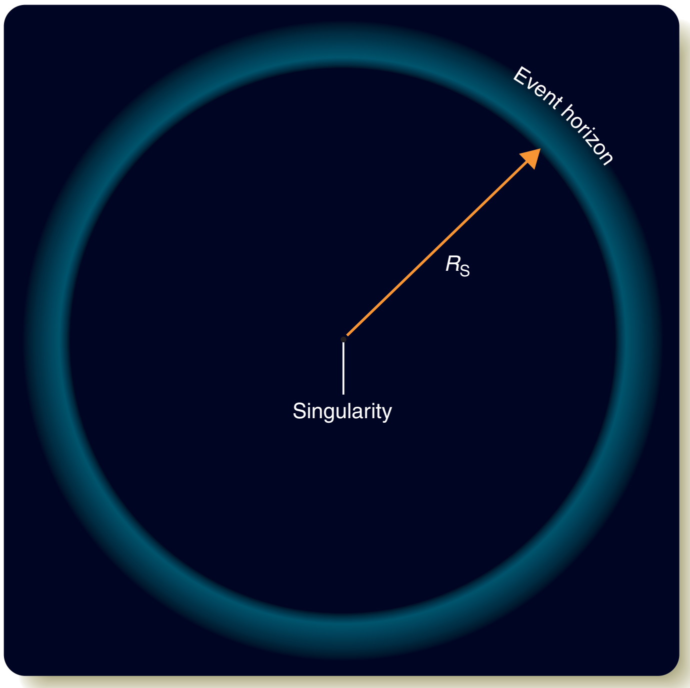
Although Schwarzschild's work was highly mathematical, his conclusion is quite simple. The Schwarzschild radius (in meters) depends only on the mass of the object (in kilograms):

$$R_S = \frac{2GM}{c^2}$$

In this simple formula, G is the gravitational constant, M is the mass, and c is the speed of light. A bit of arithmetic shows that a 1-solar-mass black hole has a Schwarzschild radius of 3 km, a 10-solar-mass black hole has a Schwarzschild radius of 30 km, and so on (■ Table 11-1). Even a very massive black hole would not have a very large event horizon.

Every object has a Schwarzschild radius determined by its mass, but not every object is a black hole. For example, Earth has a Schwarzschild radius of about 1 cm, but it could become a black hole only if you squeezed it inside that radius. Fortunately, Earth will not collapse spontaneously to become a black hole because the strength of the rock and metal in its interior is sufficient to support its weight. Only exhausted stellar cores more massive than about 3 solar masses can form black holes under the sole influence of their own gravity.

It is a **Common Misconception** to think of black holes as giant vacuum cleaners that will eventually suck in everything in the universe. A black hole is just a gravitational field, and at a reasonably large distance its gravity is no greater than that of a normal object of similar mass. If the sun were replaced by a 1-solar-mass black hole, the orbits of the planets would not change at all. ■ Figure 11-14 illustrates this by representing gravitational fields as curvature of the fabric of space-time. Normal, uncurved space-time is represented by a flat plane, and the



■ **Figure 11-13**

A black hole forms when an object collapses to a small size (perhaps to a singularity) and the escape velocity becomes so great light cannot escape. The boundary of the black hole is called the event horizon because any event that occurs inside is invisible to outside observers. The radius of the black hole R_S is the Schwarzschild radius. **Animated!**

■ **Table 11-1 | The Schwarzschild Radius**

	Mass $M_\odot$	R_S
Star	10	30 km
Star	3	9 km
Star	2	6 km
Sun	1	3 km
Earth	0.000003	0.9 cm

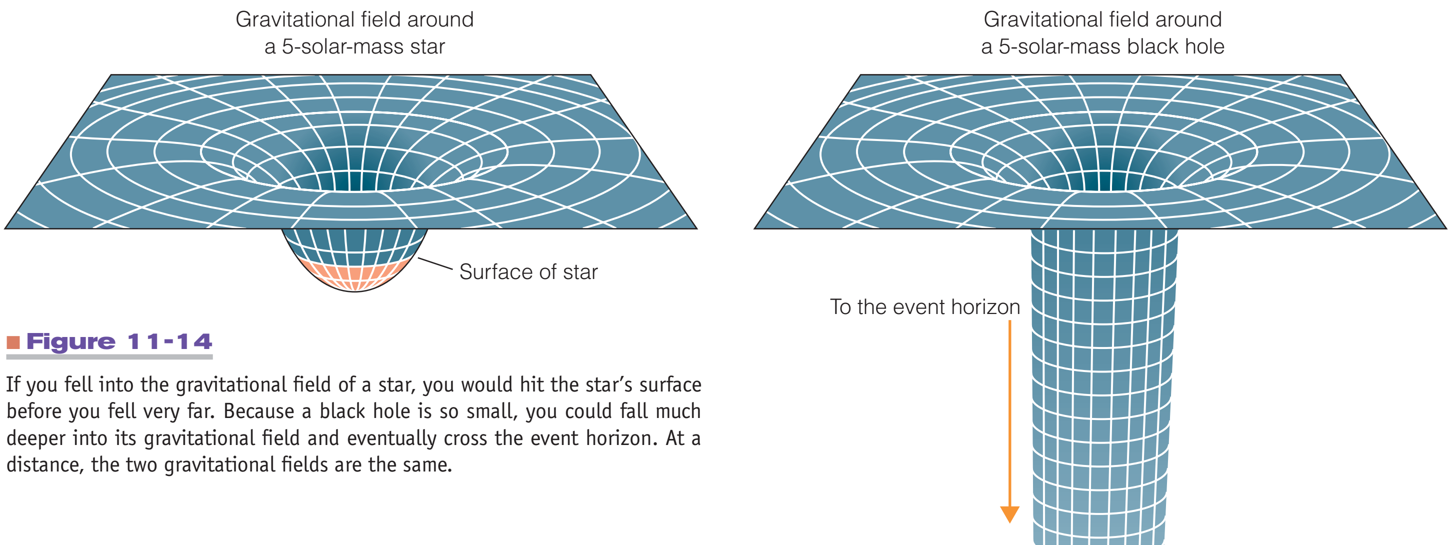


Figure 11-14

If you fell into the gravitational field of a star, you would hit the star's surface before you fell very far. Because a black hole is so small, you could fall much deeper into its gravitational field and eventually cross the event horizon. At a distance, the two gravitational fields are the same.

presence of a mass such as a star curves the plane to produce a depression. The extreme curvature around a black hole produces a deep funnel-shaped surface in this graphic representation. You can see from the graphs that the gravity of a black hole becomes extreme only when you approach close to it.

Now you can check off another **Common Misconception** that may strike you as silly. Because of special effects in movies and TV, some people think black holes should actually look like funnels. Of course, the graphs of the strength of gravity around black holes look like funnels, but black holes themselves are not shaped like funnels. If you could approach a black hole, you might be able to see hot gas swirling inward, but you wouldn't be able to see the black hole itself.

Leaping into a Black Hole

Before you can search for real black holes, you must understand what theory predicts about a black hole. To explore that idea, you can imagine leaping, feet first, into a Schwarzschild black hole.

If you were to leap into a black hole of a few solar masses from a distance of an astronomical unit, the gravitational pull would not be very large, and you would fall slowly at first. Of course, the longer you fell and the closer you came to the center, the faster you would travel. Your wristwatch would tell you that you fell for about two months by the time you reached the event horizon.

Your friends who stayed behind would see something different. They would see you falling more and more slowly as you came closer to the event horizon because, as explained by general relativity, time slows down in curved space-time. This is known as **time dilation.** In fact, your friends would never actually see you cross the event horizon. To them you would fall more and more slowly until you seemed hardly to move. Generations later, your descendants could focus their telescopes on you and see you still inching closer to the event horizon. You, however, would have sensed no slowdown and would conclude that you had reached the event horizon after about two months.

Another relativistic effect would make it difficult to see you with normal telescopes. As light travels out of a gravitational field, it loses energy, and its wavelength grows longer. This is known as the **gravitational redshift.** Although you would notice no effect as you fell toward the black hole, your friends would need to observe at longer and longer wavelengths in order to detect you.

While these relativistic effects seem merely peculiar, other effects would be quite unpleasant. If you were falling feet first, you would feel your feet, which would be closer to the black hole, being pulled in more strongly than your head. This is a tidal force, and at first it would be minor. But as you fell closer, the tidal force would become very large. Another tidal force would compress you as both your left and your right side fell toward the center of the black hole. For any black hole with a mass like that of a star, the tidal forces would crush you laterally and stretch you longitudinally long before you reached the event horizon (Figure 11-15). The friction from such severe distortions of your body would heat you to millions of degrees, and you would emit X-rays and gamma rays. (Needless to say, this would render you inoperative as a thoughtful observer.)

Some years ago a popular book suggested that you could travel through the universe by jumping into a black hole in one place and popping out of another somewhere far across space. That might make for good science fiction, but tidal forces would make it an unpopular form of transportation even if it worked. You would certainly lose your luggage.

Your imaginary leap into a black hole is not frivolous. You now know how to find a black hole: Look for a strong source of X-rays. It may be a black hole into which matter is falling.

The Search for Black Holes

Do black holes really exist? The first X-ray telescopes reached orbit in the 1970s, and that allowed astronomers to begin searching for evidence of black holes. They tried to find one or more objects that were obviously black holes. That very difficult search

How Do We Know? 11-2

Checks on Fraud in Science

How do you know scientists aren't just making stuff up? The unwritten rules of science make fraud difficult, and the way scientists publish their research makes it almost impossible. Scientists depend on each other to be honest, but they also double-check everything.

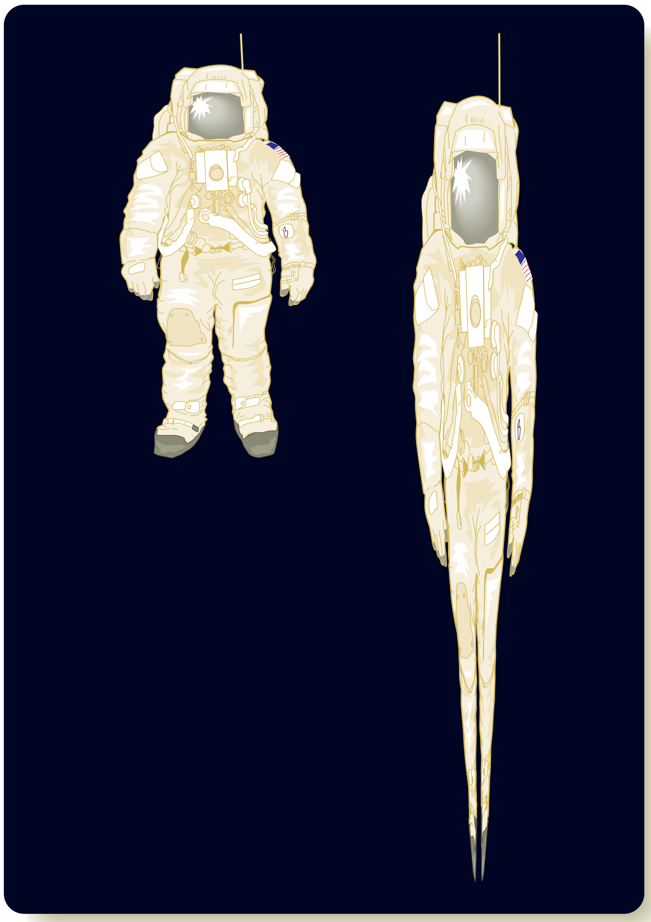
For example, all across North America, black-capped chickadees sing the same quick song. Some people say it sounds like *Chick-a-dee-dee-dee,* but others say it sounds like *Hey-sweetie-sweetie-sweetie.* You could invent tables of data and publish a paper reporting that you had recorded chickadees around Ash Lake in northern Minnesota that sing a backward song: *Sweetie-sweetie-sweetie-hey.* Experts of brain development and animal learning would be amazed, and your research might secure you praise from your colleagues, a job offer at a prestigious university, or a generous grant—but only if you could get away with it.

The first step in your scheme would be to publish your results in a scientific journal. Because the journal's reputation rests on the accuracy of the papers it publishes, the editor sends all submitted papers to two or three experts for peer review. Those world experts on chickadees would almost certainly notice things wrong with your made-up data tables. On their recommendation, the editor would probably refuse to publish your paper.

Even if your faked data fooled the peer reviewers, you would probably be found out once the paper was published. Experts on bird song would read your paper and flock to Ash Lake to study the bird songs themselves. By the next spring, you would be found out—and the journal would be forced to publish an embarrassing retraction of your article.

One of the rules of science is that good results must be repeatable. Scientists routinely repeat the work of others, not only to check the results, but as a way to start a new research topic. When someone calls a news conference and announces a new discovery, other scientists begin asking, "How does this fit with other observations? Has this been checked? Has this been peer reviewed?" Until a result has been published in a peer-reviewed journal, scientists treat it with caution. Fraud isn't unheard of in science. But because of peer review and the requirement of repeatability in science, bad research, whether the result of carelessness or fraud, is usually exposed quickly.

Chickadees always sing the same song. Hey-Sweety-Sweety-Sweety. (Steve and Dave Maslowski/Photo Researchers, Inc.)



is a good illustration of how the unwritten rules of science help scientists understand nature (■ How Do We Know? 11-2).

A black hole alone is totally invisible because nothing can escape from the event horizon, but if matter flows into a black hole, the matter will become very hot and can emit X-rays before it reaches the event horizon. An isolated black hole in space will not have much matter flowing into it, but a black hole in a binary system might receive a steady flow of matter transferred from the companion star. This suggests you can search for black holes by searching among X-ray binaries.

Some X-ray binaries such as Hercules X-1 contain a neutron star, and they will emit X-rays much as would a binary containing a black hole. You can tell the difference in two ways. If the compact object emits pulses, you know it is a neutron star. Otherwise, you must depend on the mass of the object. If the mass of the compact object is greater than 3 solar masses, it cannot be a neutron star; it must be a black hole.

The first X-ray binary suspected of harboring a black hole was Cygnus X-1, the first X-ray object discovered in Cygnus. It contains a supergiant B0 star and a compact object orbiting each

■ **Figure 11-15**

Leaping feet first into a black hole. A person of normal proportions (left) would be distorted by tidal forces (right) long before reaching the event horizon around a typical black hole of stellar mass. Tidal forces would stretch the body lengthwise while compressing it laterally. Friction from this distortion would heat the body to high temperatures.

other with a period of 5.6 days. Astronomers suspected that the X-rays were emitted by matter from the star flowing into the compact object. The object is invisible, but Doppler shifts in the spectrum reveal the motion of the B0 star around the center of mass of the binary. From the geometry of the orbit, astronomers were able to calculate that the mass of the compact object had to be greater than 3.8 solar masses, well above the maximum for a neutron star.

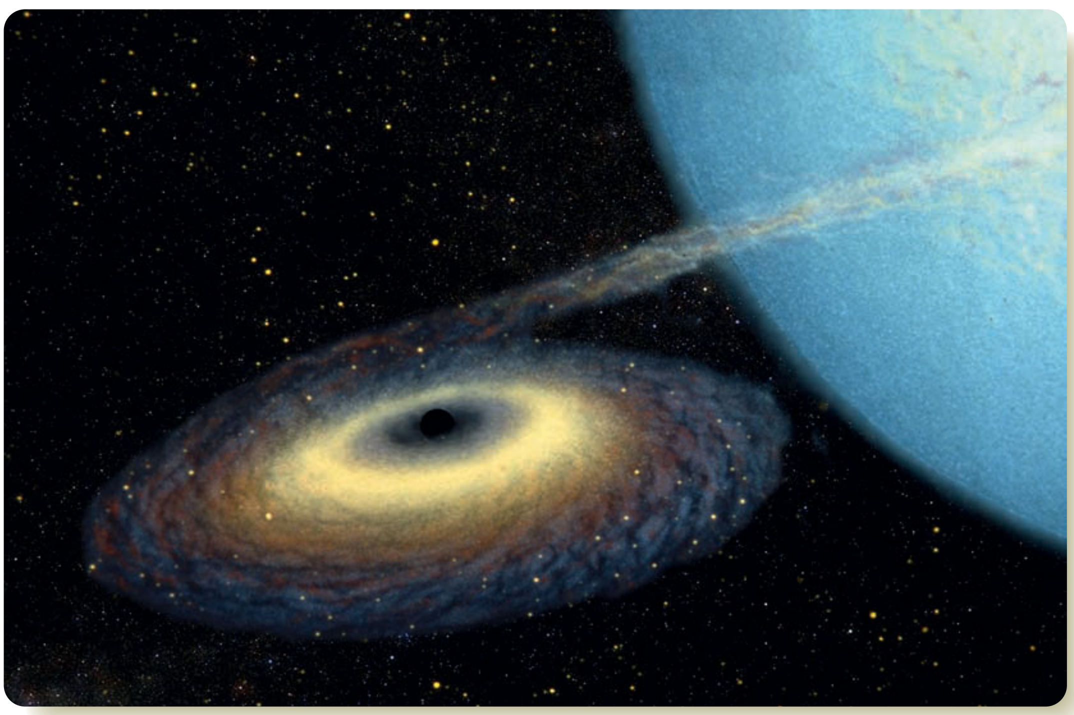
To confirm that black holes existed, astronomers needed a conclusive example, an object that couldn't be anything else. Cygnus X-1 didn't quite pass that test when it was first discovered. Perhaps the B0 star was not a normal star, or perhaps the system contained a third star. Either possibility would distort the analysis. At the time, astronomers could not conclusively show that Cygnus X-1 contained a black hole.

It took years of work to understand Cygnus X-1. Further observations and analysis show that the star has a mass of about 25 solar masses, and the compact object is about 10 times the mass of the sun. Astronomers conclude that matter flows from the B0 star as a strong stellar wind, and much of that matter gets caught in a hot accretion disk about five times larger in diameter than the orbit of Earth's moon. The inner few hundred kilometers of the disk has a temperature of about 2 million Kelvin—hot enough to radiate X rays (■ Figure 11-16). The evidence is now strong that Cyg X-1 contains a black hole.

As X-ray telescopes have found many more X-ray objects, the list of black hole candidates has grown to dozens. A few of these objects are shown in ■ Table 11-2. Each candidate is a compact object surrounded by a hot accretion disk in a close X-ray binary system. Some of the binary systems are easier to

■ Table 11-2 | Nine Black Hole Candidates

Object	Location	Companion Star	Orbital Period	Mass of Compact Object
Cygnus X-1	Cygnus	B0 supergiant	5.6 days	10 $M_\odot$
LMC X-3	Dorado	B3 main-sequence	1.7 days	$\sim$10 $M_\odot$
A0620-00	Monocerotis	K main-sequence	7.75 hours	10 ± 5 $M_\odot$
V404 Cygni	Cygnus	K main-sequence	6.47 days	12 ± 2 $M_\odot$
J1655-40	Scorpius	F–G main-sequence	2.61 days	6.9 ± 1 $M_\odot$
QZ Vul	Vulpecula	K main-sequence	8 hours	10 ± 4 $M_\odot$
4U 1543-47	Lupus	A main-sequence	1.123 days	2.7–7.5 $M_\odot$
V4641 Sgr	Sagittarius	B supergiant	2.81678 days	8.7–11.7 $M_\odot$
XTEJ1118+480	Ursa Major	K main-sequence	0.170113 days	$>$6 $M_\odot$

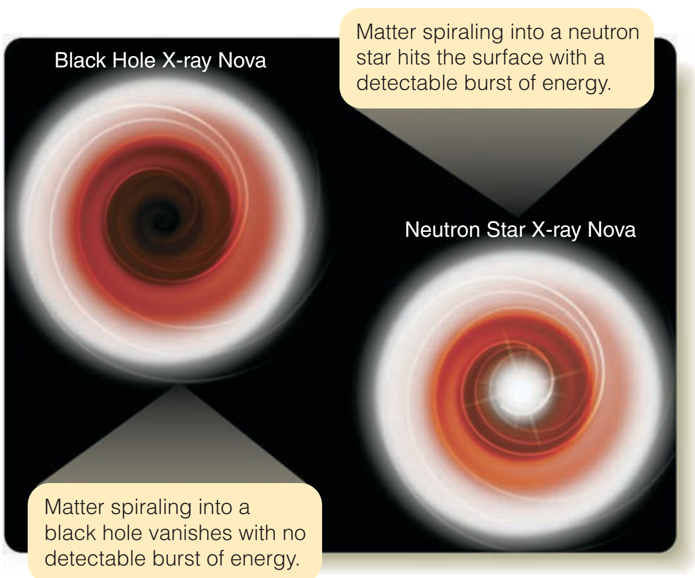


■ Figure 11-16

The X-ray source Cygnus X-1 consists of a supergiant B0 star and a compact object orbiting each other. Gas from the B0 star's stellar wind flows into the hot accretion disk around the compact object, and the X-rays astronomers detect come from the disk. (Don Dixon/cosmographica.com) analyze than others, but, in the end, it has become clear that some of these objects are too massive to be neutron stars and must be black holes. As you saw in How Do We Know? 11-1, absolute proof is not possible in science, but the evidence is now overwhelming: black holes really do exist.

Another way to confirm that black holes are real is to search for evidence of their distinguishing characteristic—event horizons—and that search too has been successful. In one study, astronomers selected 12 X-ray binary systems, six of which seemed to contain neutron stars and six of which were thought to contain black holes. Using X-ray telescopes, the astronomers monitored the systems, watching for telltale flares of energy as blobs of matter fell into the accretion disks and spiraled inward. In the six systems thought to contain neutron stars, the astronomers could also detect bursts of energy when the blobs of matter finally fell onto the surfaces of the neutron stars. In the six systems suspected of containing black holes, however, the blobs of matter spiraled inward through the accretion disks and vanished without final bursts of energy. Evidently, those blobs of matter had vanished as they approached the event horizons (■ Figure 11-17). This is dramatic evidence that event horizons are real.

The evidence shows that black holes really do exist. The problem now is to understand how these objects interact with the matter flowing into them through accretion disks to produce high-energy jets and outbursts.



■ **Figure 11-17**

Gas spiraling into an accretion disk grows hot; and, as it nears the central object, a strong gravitational redshift makes it appear redder and dimmer. Systems containing a neutron star emit bursts of energy when the gas hits the surface of the neutron star, but such bursts are not seen for systems containing black holes. In those systems, the matter vanishes as it approaches the event horizon. This is direct observational evidence of an event horizon around black holes. (NASA/CXC/SAO)

◄ **SCIENTIFIC ARGUMENT** ►

If relativistic effects slow time and prevent you from seeing matter cross the event horizon, how can infalling matter disappear without a trace?

This argument brings together observations and theory. Astronomers saw flares when matter hit the surfaces of neutron stars but saw nothing when matter fell into a black hole. Although time slows near the event horizon, remember the gravitational redshift. Hot matter flowing into a black hole can emit X-rays, but as the matter nears the event horizon, the gravitational redshift lengthens the wavelengths dramatically. The matter vanishes, not because you see it cross the event horizon, but because its photons are shifted to undetectably long wavelengths.

Now build a new argument to review a basic principle. **Why does matter become hot as it spirals into a black hole?**

◄ ►

11-3 Compact Objects with Disks and Jets

Matter flowing into a neutron star or a black hole forms an accretion disk, and that can produce some surprising phenomena. Astronomers are just beginning to understand these peculiar effects.

Jets of Energy from Compact Objects

Observations show that some compact objects are emitting jets of gas and radiation in opposite directions. These jets are similar to the bipolar outflows ejected by protostars but much more powerful. You have seen in the X-ray images on page 215 show that some young pulsars, including the Crab Nebula pulsar, are ejecting jets of highly excited gas. The Vela pulsar does the same (Figure 11-4). Systems containing black holes can also eject jets. The black hole candidate J1655-40 has been observed at radio wavelengths sporadically ejecting oppositely directed jets at 92 percent the speed of light.

One of the most powerful examples of this process is an X-ray binary called SS433. Its optical spectrum shows sets of spectral lines that are Doppler-shifted by about one-fourth the speed of light, with one set shifted to the red and one set shifted to the blue. Furthermore, the two sets of lines shift back and forth across each other with a period of 164 days. At the time of its discovery, news media reported, with a smirk, that astronomers had discovered an object that was both approaching and receding at a fantastic speed, but astronomers recognized the Doppler shifts as evidence of oppositely directed jets.

Apparently, SS433 is a binary system in which a compact object (probably a black hole) pulls matter from its companion star and forms an extremely hot accretion disk. Jets of high-temperature gas blast away in beams aimed in opposite directions. As the disk precesses, it sweeps these beams around the sky once every 164 days, and telescopes on Earth detect light from gas carried outward in both beams. One beam produces a redshift, and the other produces a blueshift.

It's not clear how the accretion disk can produce jets. Accretion disks around neutron stars and black holes are very small, spin very fast, and grow very hot. Somehow the hot gas in the disk can emit powerful beams of gas and radiation along its axis of rotation (■ Figure 11-18). The exact process isn't well understood, but it seems to involve magnetic fields that get caught in the accretion disk and are twisted into tightly wound tubes that squirt gas and radiation out of the disk and confine it in narrow beams. You can recognize the geometry of SS433 in Figure 11-18.

Such pairs of jets are a prototype that illustrates how the gravitational field around a compact object can produce powerful beams of radiation and matter. You will meet this phenomenon again when you study peculiar galaxies in a later chapter.

Gamma-Ray Bursts

The Cold War plays a minor part in the story of neutron stars and black holes. In 1963, a nuclear test ban treaty was signed, and by 1968, the United States was able to put a series of satellites in orbit to watch for nuclear tests that were violations of the treaty. A nuclear detonation emits gamma rays, so the satellites were designed to watch for bursts of gamma rays coming from Earth. The experts were startled when the satellites detected about one gamma-ray burst a day coming from space. When those data were finally declassified in 1973, astronomers realized that the bursts might be coming from neutron stars and black holes. These bursts are now known as **gamma-ray bursts.**

The Compton Gamma Ray Observatory reached orbit in 1991 and immediately began detecting gamma-ray bursts at the rate of a few a day. Its observations showed that the intensity of the gamma rays rises to a maximum in seconds and then fades away quickly; a burst is usually over in a few seconds to a minute. Data from the Compton Observatory also showed that the gamma-ray bursts were coming from all over the sky and not from any particular region. This helped astronomers eliminate some theories.

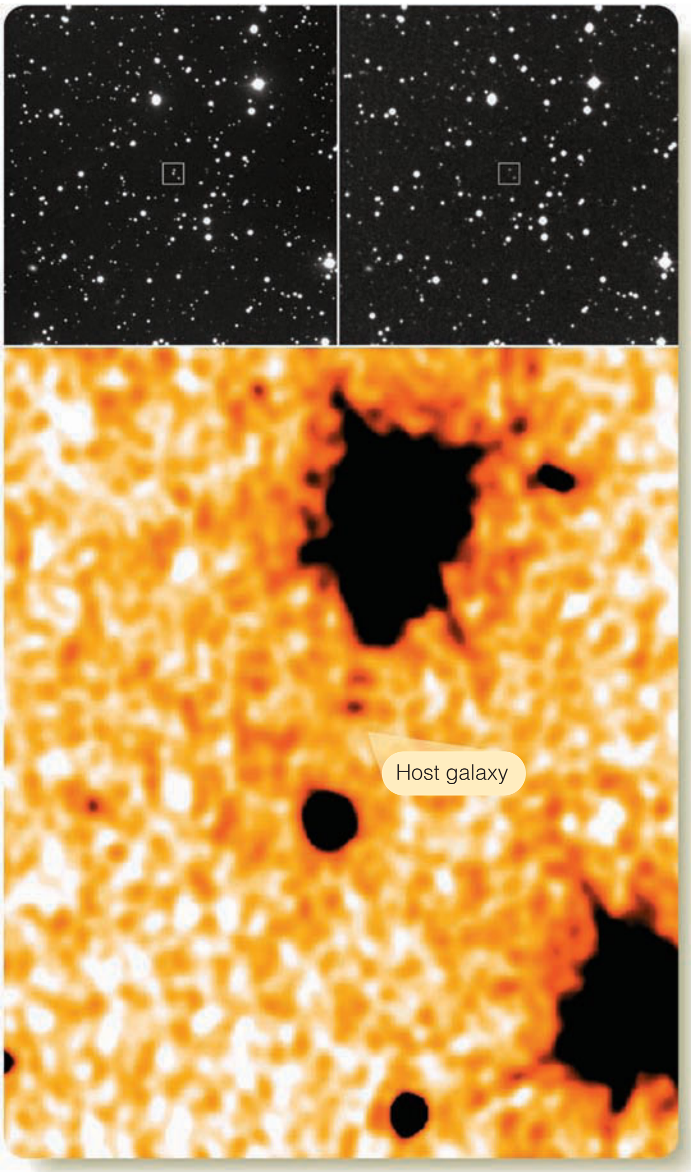
For example, some theories held that the gamma-ray bursts were being produced by relatively common events involving the stars in our galaxy, but the results from the Compton Observatory eliminated that possibility. If the gamma-ray bursts were produced among stars in our galaxy, you would expect to see them most often along the Milky Way where there are lots of stars. That the bursts occurred all over the sky meant that they were being produced by rare events in distant galaxies.

Gamma-ray bursts are hard to study because they occur without warning and fade so quickly; but, starting in 1997, new satellites in orbit were able to detect gamma-ray bursts. Such studies showed that there are two kinds of gamma-ray bursts. Short bursts last less than 2 seconds, but longer bursts can go on for many seconds. Specialized satellites could detect bursts, quickly determine their location in the sky, and immediately alert astronomers on the ground. When telescopes on Earth swiveled to image the locations of the bursts, they detected fading glows that resembled supernovae (■ Figure 11-19), suggesting that long gamma-ray bursts are produced by a certain kind of supernova explosion.

Theory proposes that a star more massive than some upper limit of about 20 solar masses can exhaust its nuclear fuel and collapse directly into a black hole. Models show that the collapsing star would conserve angular momentum and spin very rap-

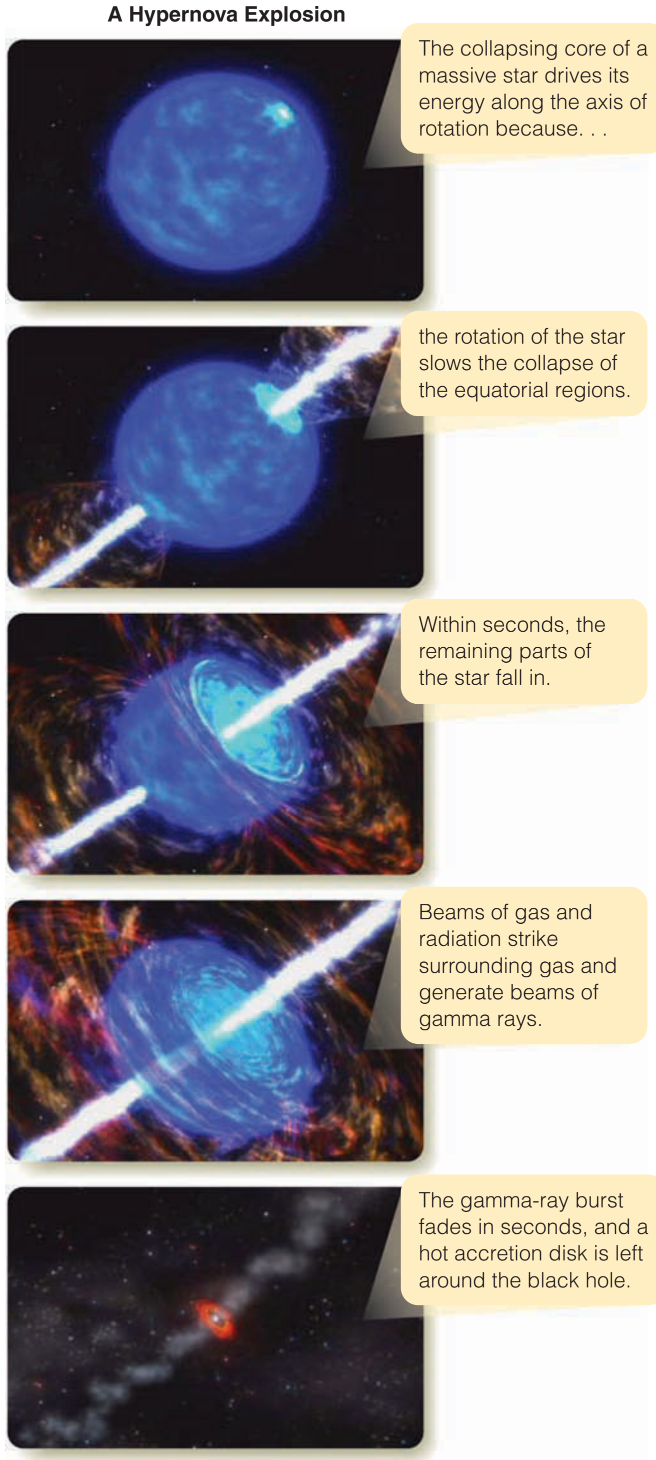
■ **Figure 11-18**

In this artist's impression, matter from a normal star flows into an accretion disk around a compact object. Processes in the spinning disk eject gas and radiation in jets perpendicular to the disk. (© 2005, Fahad Sulehria, www.novacelestia.com) **Animated!**



■ **Figure 11-19**

Alerted by gamma-ray detectors on satellites, observers used one of the VLT 8.2-meter telescopes on a mountaintop in Chile to image the location of a gamma-ray burst only hours after the burst. The image at top left shows that fading glow of the eruption. The image at top right, recorded 13 years before, reveals no trace of an object at the location of the gamma-ray burst. The Hubble Space Telescope image at bottom was recorded a year later and reveals a very faint galaxy at the location of the gamma-ray burst. (ESO and NASA)



■ **Figure 11-20**

The collapse of the cores of extremely massive stars can produce hypernova explosions, which are thought to be the source of gamma-ray bursts longer than 2 seconds. (NASA/Skyworks Digital) **Animated!**

idly, slowing the collapse of the equatorial parts of the star. The poles of the star would fall in quickly, and that would focus beams of intense radiation and ejected gas that would blast out along the axis of rotation. Such an eruption has been called a **hypernova** (■ Figure 11-20). If one of those beams were pointed at Earth, it could produce a powerful gamma-ray burst. The evidence seems clear that the long gamma-ray bursts are produced by hypernovae.

Short gamma-ray bursts don't seem to be associated with hypernovae. Some repeat, and these repeating bursts seem to be produced by neutron stars with magnetic fields 100 times stronger than that in a normal neutron star. Dubbed **magnetars,** these objects can produce bursts of gamma rays when shifts in the magnetic field break the crust of the neutron stars (■ Figure 11-21).

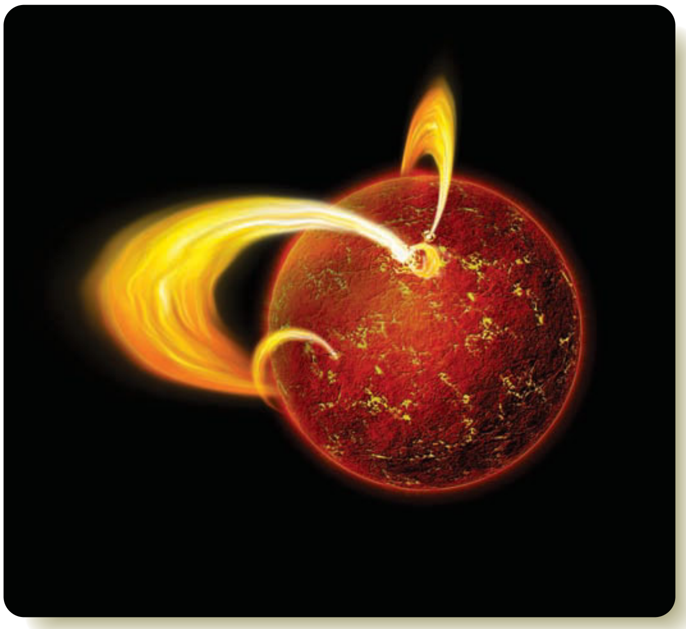


Figure 11-21

Some neutron stars appear to have magnetic fields up to 1000 times stronger than those in a normal neutron star. These magnetars can produce bursts of gamma rays when shifts in the magnetic field rupture the rigid crust of the neutron star. (NASA/CXC/M. Weiss)

One of these objects produced a burst of gamma rays that reached Earth on August 27, 1998, and temporarily ionized Earth's upper atmosphere, disrupting radio communication worldwide.

Not all short gamma-ray bursts are produced by magnetars. Some bursts have occurred in parts of distant galaxies where you would not expect to find the young, massive stars that produce magnetars or hypernovae, and the afterglows don't resemble fading supernovae. These bursts may be produced by the merger of two neutron stars that orbited around each other, radiated orbital energy as gravitational radiation, and spiraled into each other. The catastrophic merger would produce a violent explosion as the two neutron stars merged to form a black hole.

Some other short gamma-ray bursts are evidently produced by the merger of a neutron star and a black hole. As these objects spiral into each other, the neutron star is ripped apart and swallowed by the black hole. That could produce a gamma-ray burst, but the burst and its fading afterglow should be different from that produced by the merger of two neutron stars. Astronomers are now working to distinguish between these two kinds of short gamma-ray bursts.

The brightest gamma-ray burst ever recorded occurred on March 19, 2008. It originated in a galaxy 7.5 billion light years from Earth, but it was so powerful that for about 1 minute, it was bright enough to see with the unaided eye. No one knows what caused that burst, but it is clear that gamma-ray bursts are among the most powerful events in nature.

A nearby binary pulsar is only 1600 ly from Earth. If a gamma-ray burst occurred at this distance, the gamma rays would shower Earth with radiation equivalent to a 10,000-megaton nuclear blast. (The largest bombs ever made were a few megatons.) The gamma rays could create enough nitric oxide in the atmosphere to produce intense acid rain and would destroy the ozone layer and expose life on Earth to deadly levels of solar ultraviolet radiation. Gamma-ray bursts can occur relatively near the Earth as often as every few hundred million years and could be one of the causes of the mass extinctions that show up in the fossil record.

Does it surprise you that such rare events as merging neutron stars and hypernovae produce something so common that gamma-ray telescopes observe one or more every day? Remember that these events are so powerful they can be detected over very great distances. There may be 30,000 neutron star binaries in each galaxy, and there are billions of galaxies within range of gamma-ray telescopes.

What Are We? Abnormal

Look around. What do you see? A table, a chair, a tree? It's all normal stuff. The world we live in is familiar and comfortable, but astronomy reveals that "normal" isn't normal at all. The universe is, for the most part, utterly unlike anything you have ever experienced.

Throughout the universe, gravity makes clouds of gas form stars, and in turn the stars generate energy through nuclear fusion in their cores, which delays gravity's final victory. But gravity always wins. You have learned that stars of different masses die in different ways, but you have also discovered that they always reach one of three end states: white dwarfs, neutron stars, or black holes. However strange these compact objects are, they are common. They are normal.

The physics of compact objects is extreme and violent. You are not accustomed to objects as hot as the surface of a neutron star, and you have never experienced a black hole, where gravity is so strong it would pull you to pieces.

The universe is filled with things that are so violent and so peculiar they are almost unimaginable, but they are so common they deserve the label "normal." Next time you are out for a walk, look around and notice how beautiful Earth is and recall how unusual it is compared to the rest of the universe.

Study and Review

Summary

- When a supernova explodes, the core collapses to very small size. Theory predicts that the collapsing core cannot support itself as a white dwarf if its mass is greater than 1.4 solar masses, the Chandrasekhar limit. If its mass lies between 1.4 solar masses and about 3 solar masses, it can halt its contraction and form a **neutron star (p. 211).**
- A neutron star is supported by the pressure of the degenerate neutrons. Theory predicts that a neutron star should be about 10 km in radius, spin very fast because it conserves angular momentum as it contracts, and have a powerful magnetic field.
- **Pulsars (p. 212),** rapidly pulsing radio sources, were discovered in 1967. The **lighthouse model (p. 213)** explains pulsars as spinning neutron stars that emit beams of radiation from their magnetic poles. As they spin, they sweep the beams around the sky like lighthouses; if the beams sweep over Earth, astronomers detect pulses. The short pulses and the discovery of a pulsar in the supernova remnant called the Crab Nebula were key evidence that pulsars are neutron stars.
- A spinning neutron star slows as it radiates its energy into space. Most of the energy emitted by a pulsar is carried away as a **pulsar wind (p. 216).**
- Theory predicts that a neutron star cannot have a mass greater than about 3 solar masses. Dozens of pulsars have been found in binary systems, and those objects allow astronomers to estimate the masses of the neutron stars. Such masses are consistent with the predicted masses of neutron stars.
- Observations of the first binary containing two neutron stars revealed that the system is losing orbital energy by emitting **gravitational radiation (p. 217).**
- In some binary systems, mass flows into a hot accretion disk around the neutron star and causes the emission of X-rays. **X-ray bursters (p. 219)** are systems in which matter accumulates on the surface of the neutron star and explodes.
- The fastest pulsars, the **millisecond pulsars (p. 219),** appear to be old pulsars that have been spun up to high speed by mass flowing from binary companions.
- Planets have been found orbiting at least one neutron star. They may be the remains of a companion star that was mostly devoured by the neutron star.
- If the collapsing core of a supernova has a mass greater than 3 solar masses, then it must contract to a very small size—perhaps to a **singularity (p. 222),** an object of zero radius. Near such an object, gravity is so strong that not even light can escape, and the region is called a **black hole (p. 233).**
- The outer boundary of a black hole is the **event horizon (p. 233);** no event inside is detectable. The radius of the event horizon is the **Schwarzschild radius (p. 233),** amounting to only a few kilometers for a black hole of stellar mass.
- If you were to leap into a black hole, your friends who stayed behind would see two relativistic effects. They would see your clock slow relative to their own clock because of **time dilation (p. 224).** Also, they would see your light redshifted to longer wavelengths because of the **gravitational redshift (p. 224).**
- You would not notice these effects, but you would feel powerful tidal forces that would deform and heat your mass until you grew hot enough to emit X-rays. Any X-rays you emitted before reaching the event horizon could escape.
- To search for black holes, astronomers must look for binary star systems in which mass flows into a compact object and emits X-rays. If the mass of the compact object is greater than about 3 solar masses, then the object is presumably a black hole. A number of such objects have been located.
- Black holes and neutron stars at the center of accretion disks can eject powerful beams of radiation and gas. Such beams have been detected.
- **Gamma-ray bursts (p. 228)** appear to be related to violent events involving neutron stars and black holes. Bursts longer than 2 seconds appear to arise during **hypernovae (p. 229),** the collapse of massive stars to form black holes.
- The short gamma-ray bursts are produced by shifts in the powerful magnetic fields in **magnetars (p. 229)** or the merger of binary compact objects such as neutron-star pairs or neutron-star/black-hole pairs.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. How are neutron stars and white dwarfs similar? How do they differ?
2. Why is there an upper limit to the mass of neutron stars?
3. Why do you expect neutron stars to spin rapidly?
4. If neutron stars are hot, why aren't they very luminous?
5. Why do you expect neutron stars to have a powerful magnetic field?
6. Why did astronomers conclude that pulsars could not be pulsating stars?
7. What does the short length of pulsar pulses tell you?
8. How does the lighthouse model explain pulsars?
9. What evidence can you cite that pulsars are neutron stars?
10. Why would astronomers at first assume that the first millisecond pulsar was young?
11. How can a neutron star in a binary system generate X-rays?
12. If the sun has a Schwarzschild radius, why isn't it a black hole?
13. How can a black hole emit X-rays?
14. What evidence can you cite that black holes really exist?
15. How can mass transfer into a compact object produce jets of high-speed gas? X-ray bursts? Gamma-ray bursts?
16. Discuss the possible causes of gamma-ray bursts.
17. **How Do We Know?** How would you respond to someone who said, "Oh, that's only a theory."
18. **How Do We Know?** Why can't scientists prove a theory is conclusively correct?
19. **How Do We Know?** How does peer review make fraud rare in science?

Discussion Questions

1. In your opinion, has the link between pulsars and neutron stars been sufficiently tested to be called a theory, or should it be called a hypothesis? What about the existence of black holes?
2. Why wouldn't an accretion disk orbiting a giant star get as hot as an accretion disk orbiting a compact object?

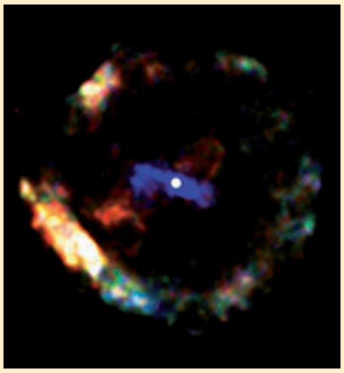
Problems

1. If a neutron star has a diameter of 10 km and rotates 642 times a second, what is the speed of the surface at the neutron star's equator in terms of the speed of light? (*Hint:* The circumference of a circle is $2\pi r$.)
2. A neutron star and a white dwarf have been found orbiting each other with a period of 11 minutes. If their masses are typical, what is the average distance between them? (*Hint:* Use Newton's version of Kepler's third law.)

3. If Earth's moon were replaced by a typical neutron star, what would the angular diameter of the neutron star be, as seen from Earth? (*Hint:* See Reasoning with Numbers 3-1.)
4. What is the Schwarzschild radius of Jupiter (mass $= 2 \times 10^{27}$ kg)? Of a human adult (mass = 75 kg)? (*Hint:* See Appendix A for the values of G and c.)
5. If the inner accretion disk around a black hole has a temperature of 1,000,000 K, at what wavelength will it radiate the most energy? What part of the spectrum is this in? (*Hint:* Use Wien's law.)
6. What is the orbital period of a bit of matter in an accretion disk 2×10^5 km from a 10-solar-mass black hole? (*Hint:* Use Newton's version of Kepler's third law.)
7. If a 20-solar-mass star and a neutron star orbit each other every 13.1 days, then what is the average distance between them? (*Hint:* Use Newton's version of Kepler's third law.)
8. What is the orbital velocity at a distance of 7400 meters from the center of a 5-solar-mass black hole? What kind of particles could orbit at this distance? (*Hint:* See Reasoning with Numbers 4-1.)
9. Compare the orbit in Problem 8 with an orbit having the same velocity around a 2-solar-mass neutron star. Why is this orbit impossible? (*Hint:* See Reasoning with Numbers 4-1.)

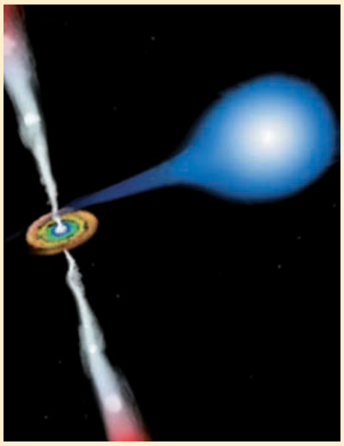
Learning to Look

1. The X-ray image at the right shows the supernova remnant G11.2–0.3 and its central pulsar in X-rays. The blue nebula near the pulsar is caused by the pulsar wind. How old do you think this system is? Discuss the appearance of this system a million years from now.



NASA/McGill, V. Kaspi et al.

2. What is happening in the artist's impression at the right? How would you distinguish between a neutron star and a black hole in such a system?



CXC/M. Weiss

12 The Milky Way Galaxy

The stars of our home galaxy, the Milky Way, rise behind a telescope dome and the highly polished surface of a submillimeter telescope at the La Silla European Southern Observatory in Chile. (ESO and Nico Housen)

Guidepost

You have traced the life story of the stars from their birth in clouds of gas and dust to their deaths as white dwarfs, neutron stars, or black holes. Now you are ready to see stars in their vast communities called galaxies. This chapter discusses our home galaxy, the Milky Way Galaxy, and attempts to answer four essential questions:

How do astronomers know what our galaxy is like?

How did our galaxy form and evolve?

What lies at the very center of our galaxy?

What are the spiral arms?

In this chapter you will see more examples of how scientists use evidence and theory to understand nature. If in some cases the evidence seems contradictory and the theories incomplete, do not be disappointed. The adventure of discovery is not yet over.

In the chapters that follow, you will meet some of the billions of galaxies that fill the depths of the universe. Understanding the Milky Way Galaxy is only one step in understanding the universe as a whole.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

The Stars Are Yours

JAMES S. PICKERING

THE STARS ARE Yours is the title of a popular astronomy book written by James S. Pickering in 1948. You would have liked that book back in 1948, although it has gotten a bit out of date over the last six decades. The point of its title is that the stars belong to everyone equally, and you can enjoy the wonder of the night sky as if you owned it.

The stars may indeed be yours. You live inside a galaxy—one of the large star systems that fill the universe. Our Milky Way Galaxy contains over 100 billion stars. If Earth is the only inhabited planet in the galaxy, then the galaxy belongs to all of us. Sharing equally, each person on Earth owns 15 stars plus any associated planets, moons, comets, and so on. Even if Earth must share the galaxy with a few other inhabited planets, you are rich beyond any dream. You lack only the transportation to visit your dominions.

As you read this chapter, you will learn about our home galaxy, but you will also learn how the stars in the Milky Way Galaxy have, generation after generation, made the atoms in your body. As you begin this chapter, it may seem that the stars belong to you; but, by the end of this chapter, you may decide that you belong to the stars.

12-1 The Discovery of the Galaxy

IT SEEMS ODD to say that astronomers discovered something that is all around us, but it isn't obvious that we live in a galaxy. We are inside, and we see nearby stars scattered all over the sky, while the more distant clouds of stars in our galaxy make a faint band of light circling the sky (■ Figure 12-1a). The ancient Greeks named that band *galaxies kuklos,* the "milky circle." The Romans changed the name to *via lactea,* "milky road" or "milky way." It was not until early in the 20th century that astronomers understood that we are inside a great wheel of stars and that the universe is filled with other such star systems. Drawing on the Greek word for milk, they called them galaxies.

Almost every celestial object visible to your naked eyes is part of the Milky Way Galaxy. The only exception visible from Earth's Northern Hemisphere is the Andromeda Galaxy (cataloged as M31), just visible to your unaided eyes as a faint patch of light in the constellation Andromeda.* Seen from a distance, our galaxy probably looks much like the Andromeda Galaxy (Figure 12-1b).

*Consult the star charts at the end of this book to locate the Milky Way and the Andromeda Galaxy (labeled M31 on the chart).

The Great Star System

Galileo's telescope revealed that the glowing Milky Way is made up of stars, and later astronomers realized that the sun must be located in a great wheel-shaped cloud of stars, which they called the star system. Only a wheel shape could produce the band of the Milky Way encircling the sky. In 1750, Thomas Wright (1711–1786), drawing on the technology of the time, referred to the wheel-shaped star system as the grindstone universe, by analogy with the thick disks of stone used in mills.

The English astronomer Sir William Herschel (1738–1822) and his sister Caroline Herschel (1750–1848), also a talented astronomer, attempted to gauge the true shape of the star system by counting stars in 683 different directions in the sky. Where they saw more stars, they assumed the star system extended farther into space. They plotted their results to create a diagram showing an irregular disk shape with the sun located near the center. That confirmed the grindstone model of the universe (■ Figure 12-2).

In some directions in the sky, the Herschels saw very few stars, and these "holes in the sky" produced great irregularities along the edge of their diagram, as shown in Figure 12-2. Modern astronomers know that these empty spots are caused by dense clouds of gas and dust inside the Milky Way Galaxy that block the view of more distant stars, but the Herschels, like many early astronomers, did not understand the importance of gas and dust in space. Although they thought they were seeing to the edge of the star system, they were actually seeing only as far into the Milky Way as the gas and dust permitted. Because they counted similar numbers of stars all around the Milky Way, they concluded that the sun was near the center of the grindstone universe.

The model proposed by the Herschels was widely accepted and studied by other astronomers. The Herschels were not able to measure the size of the star system, but later astronomers were able to estimate distances to stars in statistical ways, and they concluded that the star system the Herschels had mapped was only about 15,000 ly in diameter.

As the 20th century began, astronomers believed that the sun was located near the center of a rather small, wheel-shaped star system. How the human race realized the truth about their location in the universe is an adventure that begins with a woman studying stars that pulsate and leads to a man studying star clusters.

Cepheid Variable Stars

It is a **Common Misconception** of poets that the stars are eternal and unchanging; astronomers have known for centuries that some stars change in brightness. You already know that novae and supernovae appear, grow brighter, and then fade, but many other stars change periodically, growing brighter, then fainter, then brighter again. Some of these **variable stars** are eclipsing

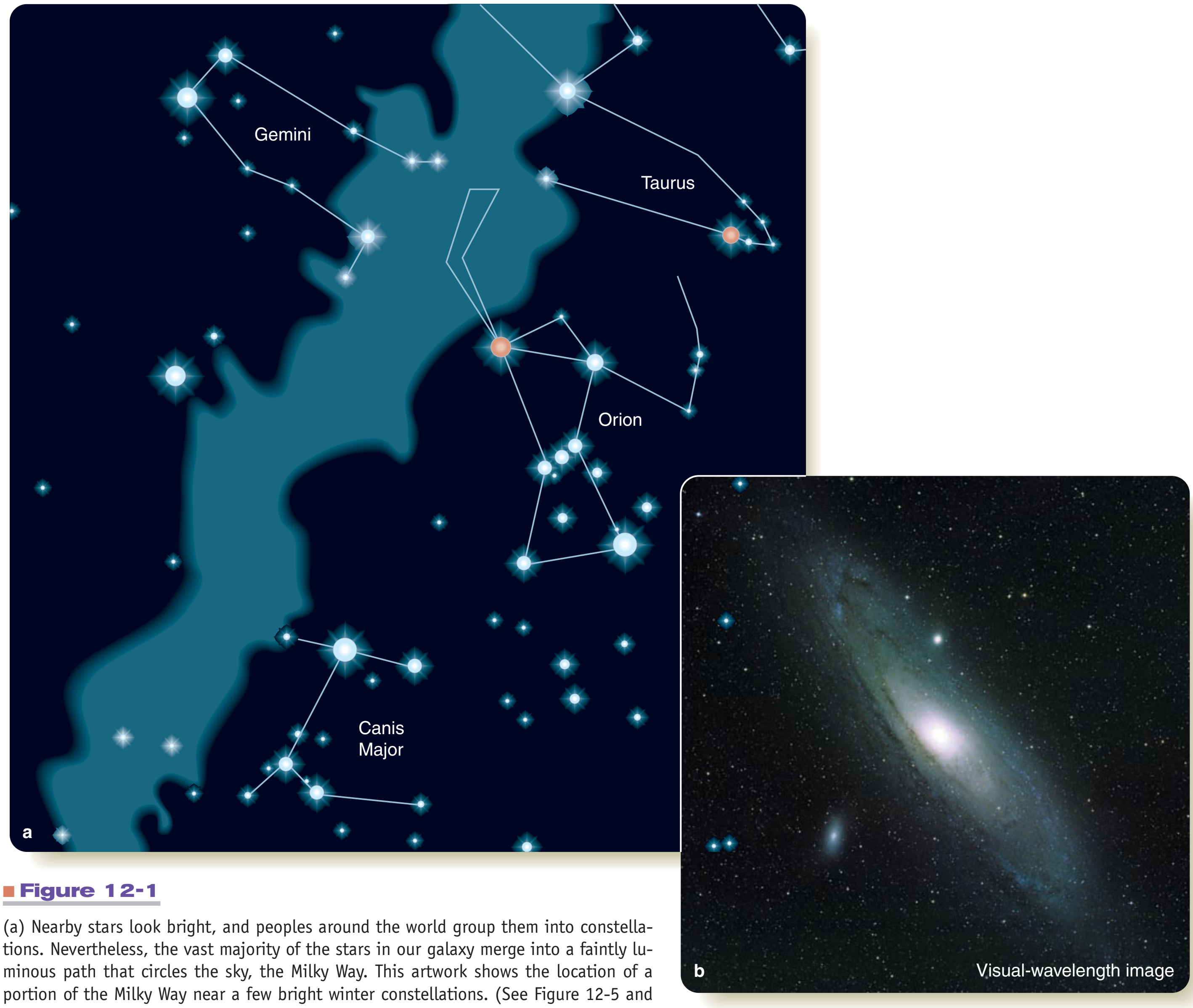


Figure 12-1

(a) Nearby stars look bright, and peoples around the world group them into constellations. Nevertheless, the vast majority of the stars in our galaxy merge into a faintly luminous path that circles the sky, the Milky Way. This artwork shows the location of a portion of the Milky Way near a few bright winter constellations. (See Figure 12-5 and the star charts at the end of this book to further locate the Milky Way in your sky.) (b) This photograph of the Andromeda Galaxy, a spiral galaxy about 2.3 million ly from Earth, shows what our own galaxy would look like if you could view it from a distance. (AURA/NOAO/NSF)

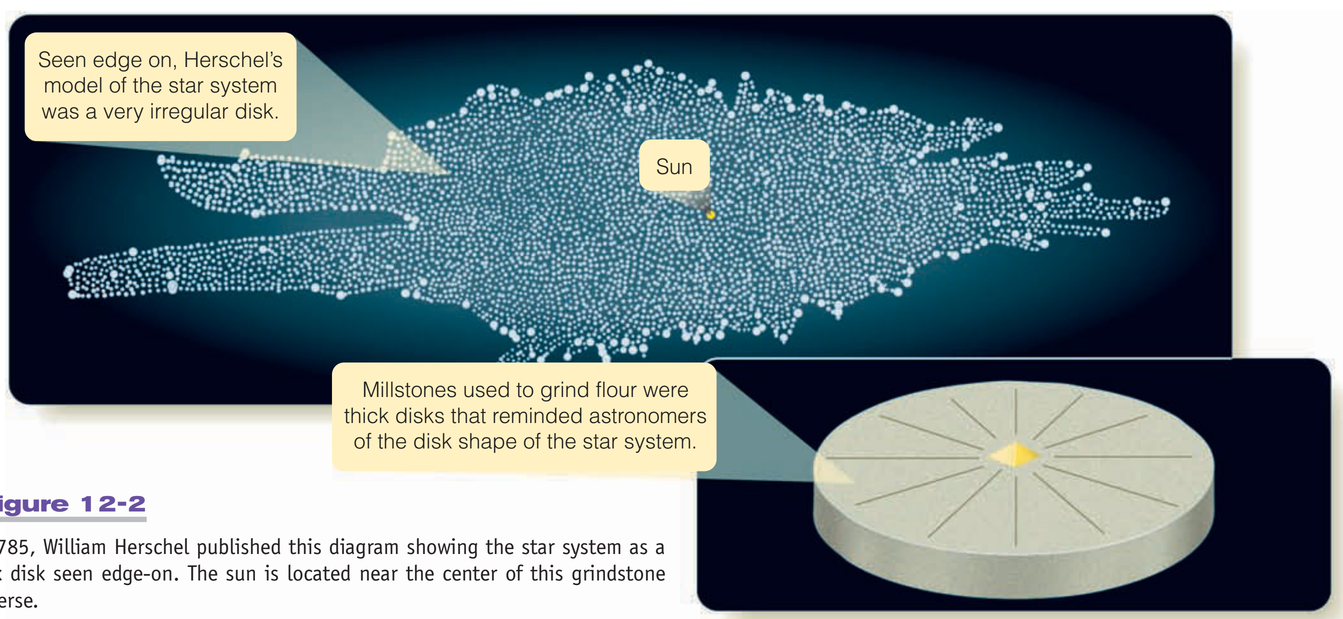


Figure 12-2

In 1785, William Herschel published this diagram showing the star system as a thick disk seen edge-on. The sun is located near the center of this grindstone universe.

binaries, but some are single stars that pulsate like beating hearts.

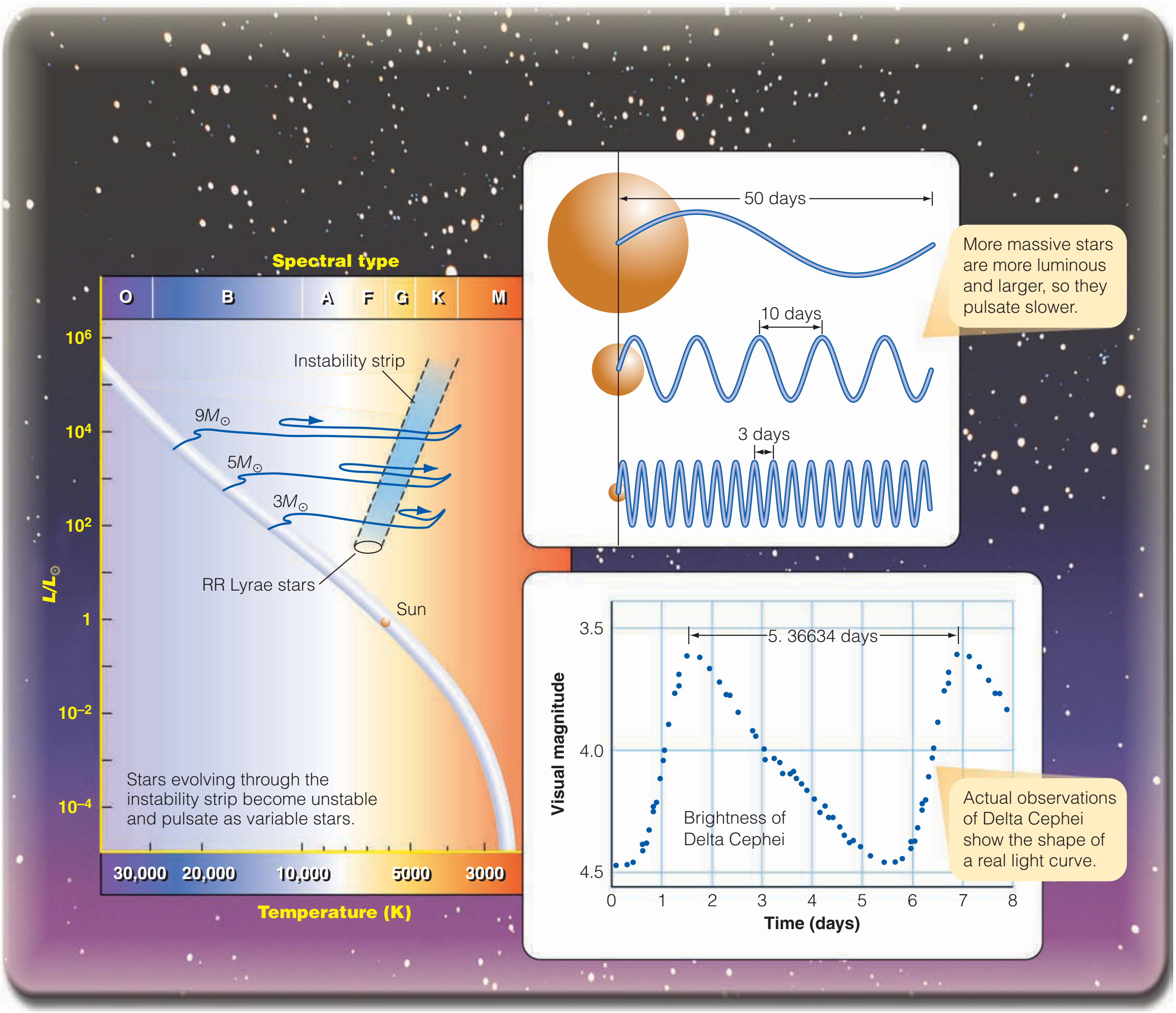
In 1912, Henrietta Leavitt (1868–1921) was studying a star cloud in the southern sky known as the Small Magellanic Cloud. On her photographic plates, she found many variable stars, and she noticed that the brightest had the longest periods—the time it takes a star to complete a cycle from bright to faint to bright. She didn't know the distance to the cloud, so she couldn't calculate the absolute magnitudes of the stars, but because all of the variables were at nearly the same distance she concluded that there was a relationship between period and luminosity.

The stars Leavitt saw, **Cepheid variable stars,** are named after the first such star discovered, δ Cephei. Many Cepheid variables are known today; they have periods from 1 to 60 days and lie in a region of the H–R diagram known as the **instability strip** (■ Figure 12-3). From their position in the diagram, you can tell that Cepheids are giant and supergiant stars. A related kind of variable star, an **RR Lyrae variable star,** named after the variable star RR in the constellation Lyra, has a period of about half a day and lies at the bottom of the instability strip.

Stars in the instability strip are unstable and pulsate because a layer of partially ionized helium is in just the right place in the star's atmosphere to absorb and release energy like a spring. In hotter stars the layer is too high to make the star unstable, and in cooler stars the layer is too low. Stars in the instability strip hap-

■ **Figure 12-3**

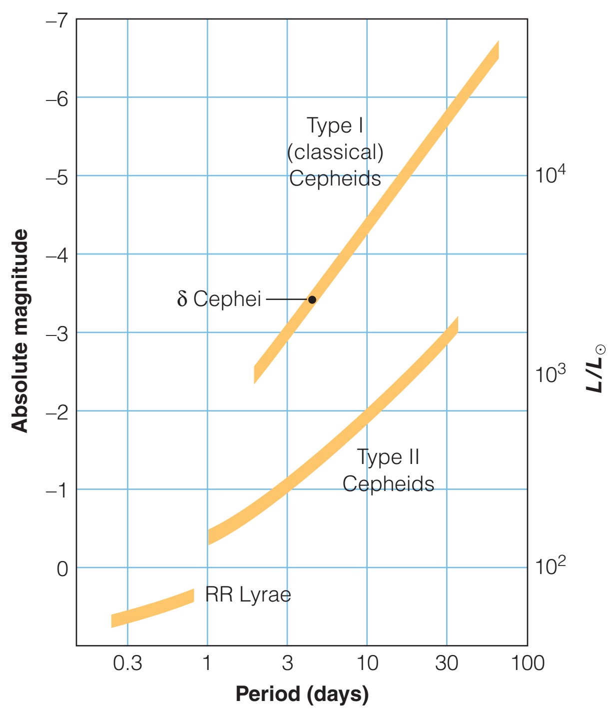
The more massive a star is, the more luminous it is. Massive stars become larger when they leave the main sequence, and those larger stars pulsate with longer periods when they pass through the instability strip. Because both luminosity and period depend on mass, there is a relationship between period of pulsation and luminosity.



pen to have this layer in just the right place to make them pulsate as energy flows out of their interiors. As stars evolve into the instability strip, they become unstable and pulsate; when they evolve out of the strip, they stop pulsating. Massive stars are very luminous, and they cross the instability strip higher in the H–R diagram. Because these massive stars are larger, they pulsate more slowly, just as large bells vibrate more slowly and have deeper tones. You will remember from an earlier chapter that Favorite Star Polaris is a supergiant; it lies high in the instability strip and pulsates as a Cepheid variable. Lower-mass stars are less luminous, cross the instability strip lower in the H–R diagram, and, because they are smaller, pulsate faster. This explains why the long-period Cepheids are more luminous than the short-period Cepheids. Leavitt first noticed this in 1912, and it is now known as the **period–luminosity relation** (■ Figure 12-4).

The physics of pulsating stars was unknown in 1912, but when Leavitt noticed that a Cepheid's period of pulsation was related to its luminosity, she found the key that unlocked the secret of the Milky Way.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Cepheid Variable."



■ **Figure 12-4**

The period–luminosity diagram is a graph of the brightness of variable stars versus their periods of pulsation. You could plot brightness as luminosity; but, because the diagram is used in distance calculations, it is more convenient with absolute magnitude on the vertical axis. Modern astronomers know that there are two types of Cepheids, something that astronomers in the early 20th century could not recognize in their limited data.

The Size of the Milky Way

Early in the 20th century, when Henrietta Leavitt was studying her photographic plates, astronomers believed that they lived near the center of a disk-shaped cloud of stars that they called the star system. Many believed that the star system was isolated in an otherwise empty universe.

A young astronomer named Harlow Shapley (1885–1972) began the discovery of the true nature of the Milky Way when he thought about different kinds of star clusters. In an earlier chapter, you met open clusters and globular clusters. Open clusters are concentrated along the Milky Way, but globular clusters are widely scattered. Shapley noticed that the globular clusters were more common toward the constellations Sagittarius and Scorpius (■ Figure 12-5).

Shapley assumed that this great cloud of globular clusters was controlled by the combined gravitational field of the entire star system. In that case, he realized, he could study the size and extent of the star system by studying the globular clusters. To do that, he needed to measure the distances to as many globular clusters as possible.

Globular clusters are much too far away to have measurable parallaxes, but they do contain variable stars. Shapley knew of Henrietta Leavitt's work on these stars, and he knew that she had been unable to find their absolute magnitudes. He also knew that knowing the absolute magnitudes would tell him the distances to the star clusters, so he turned his attention to variable stars.

All stars move through space, and over periods of a few years these motions, known as **proper motions,** can be detected as small shifts in the positions of the stars in the sky. The more distant stars have undetectably small proper motions, but the nearer stars have larger proper motions. Clearly, proper motions contain clues to distance. Although no Cepheids were close enough to have measurable parallaxes, Shapley searched catalogs of proper motion observations and found 11 Cepheids with measured proper motions. Through a statistical process, he was able to find their average distance and from that their average absolute magnitude. That meant he could replace Leavitt's apparent magnitudes with absolute magnitudes on the period–luminosity diagram (as shown in Figure 12-4). That is, he knew how bright the variable stars really were.

Finally, he was ready to find the distance to the globular clusters. He could identify the variable stars he found in the clusters, and he could measure their apparent magnitude from his photographs. Their short periods of pulsation alerted Shapley that the variables in his clusters were RR Lyrae variables, and comparing their apparent and absolute magnitude gave him the distance to the star cluster (Reasoning with Numbers 8-2). Notice how Shapley proceeded step-by-step in his research; astronomers say he **calibrated** the variable stars for distance determination (■ How Do We Know? 12-1).

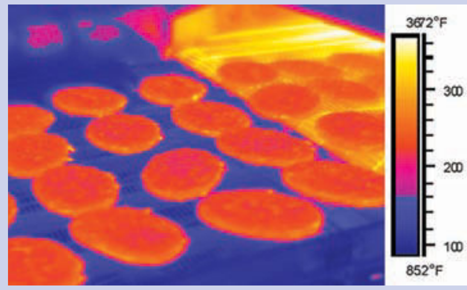
How Do We Know? 12-1

Calibration

How do scientists simplify difficult measurements? Astronomers say that Shapley "calibrated" the Cepheids for the determination of distance, meaning that he did all the detailed background work so that the Cepheids could be used to find distances. Once this was done, other astronomers could then use Shapley's calibrated diagram to find the distance to other Cepheids without repeating the detailed analysis.

Calibration is useful because it saves a lot of time and effort. For example, engineers in steel mills must monitor the temperature of molten steel, but they can't dip in a thermometer. Instead, they can use handheld devices that measure the color of molten steel. Recall from Chapter 6 that the color of blackbody radiation is determined by its temperature. Molten steel emits visible and infrared radiation that is nearly perfect blackbody radiation, so the manufacturer can calibrate the engineer's devices to convert the measured color to a temperature displayed on digital readouts. The engineers don't have to repeat the calibration every time; they just point their instrument at the molten steel and read off the temperature. Astronomers have made the same kind of color–temperature calibration for stars.

As you read about any science, notice how calibrations are used to simplify common measurements. But notice, too, how important it is to get the calibration right. An error in calibration can throw off every measurement made with that calibration.



By calibrating infrared color against temperature, bakers can monitor the operation of their ovens. (Courtesy of FLIR Systems, Inc.)

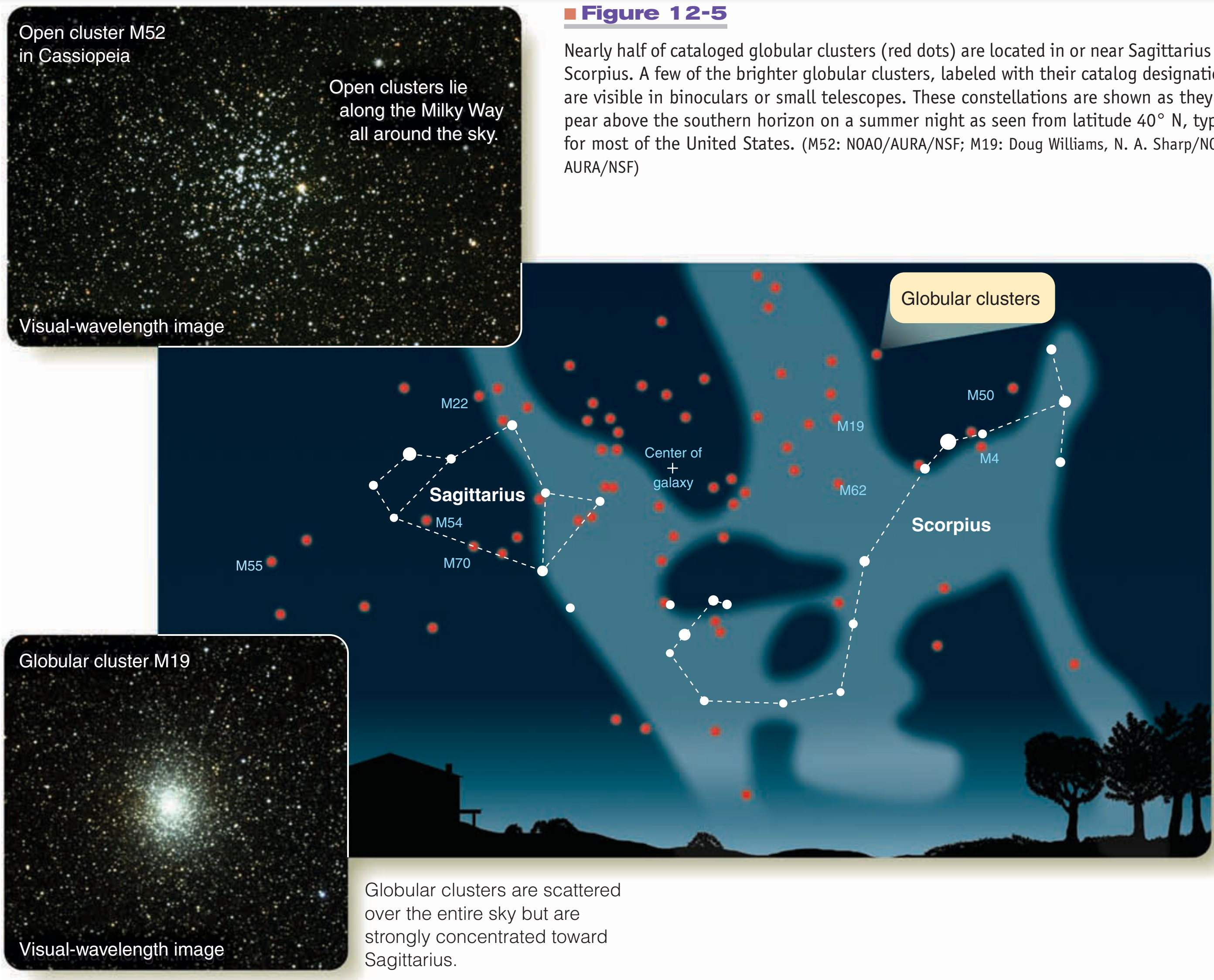


Figure 12-5

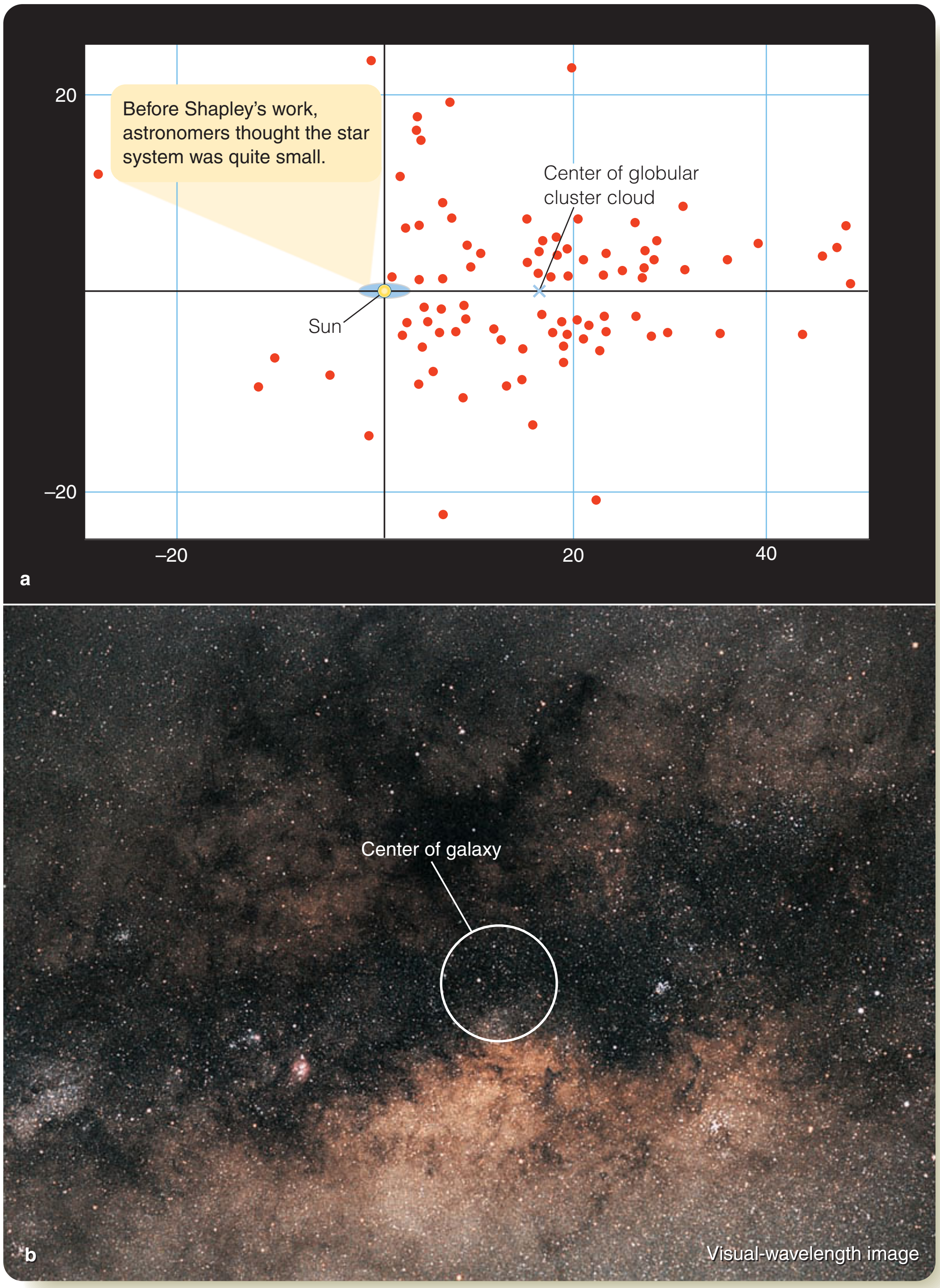
Nearly half of cataloged globular clusters (red dots) are located in or near Sagittarius and Scorpius. A few of the brighter globular clusters, labeled with their catalog designations, are visible in binoculars or small telescopes. These constellations are shown as they appear above the southern horizon on a summer night as seen from latitude 40° N, typical for most of the United States. (M52: NOAO/AURA/NSF; M19: Doug Williams, N. A. Sharp/NOAO/AURA/NSF)

Shapley later wrote that it was late at night when he plotted the direction and distance to the globular clusters and found that, just as he had supposed, they formed a great swarm. But the swarm was not centered on the sun. The center lay many thousands of light-years away in the direction of Sagittarius. The star system was much bigger than anyone had suspected (■ Figure 12-6). He found the only other person in the building, a cleaning lady, and the two stood looking at his graph as he explained that they were the only two people on Earth who understood that humanity lives, not at the center of a small star system, but in the suburbs of a vast wheel of stars, a galaxy.

It is interesting to note that Shapley made two mistakes. First, he assumed that space was empty. Later astronomers realized that the interstellar medium dims the more distant stars and makes them look farther away than they really are. Also, Shapley didn't know that there are two types of Cepheids. Those he saw in his clusters were the fainter type, but those he used for calibration were brighter. That meant he overestimated the luminosity of his cluster stars and overestimated the distance to the clusters. Consequently his estimate for the size of our galaxy shown in Figure 12-6 was bigger than the modern value. The point is not that he made a few errors of calibration, but that he got the main point right. We live not at the center of a small star system but in the suburbs of a very big wheel of stars.

It is not surprising that astronomers at first thought that the star system was small. When you look toward the band of the Milky Way, you can see only the neighborhood near the sun. Gas and dust make most of the star system invisible and, like a traveler in a fog, you seem to be at the center of a small region. Shapley was able to observe the globular clusters at greater distances because they lie outside the plane of the star system and are not dimmed very much by gas and dust (■ Figure 12-7).

Building on Shapley's work, other astronomers began to suspect that some of the faint patches of light visible through telescopes



■ Figure 12-6

(a) Shapley's study of globular clusters showed that they were not centered on the sun, at the origin of this graph, but rather formed a great cloud centered far away in the direction of Sagittarius. Distances on this graph are given in thousands of parsecs. (b) Looking toward Sagittarius, you see nothing to suggest this is the center of the galaxy. Gas and dust block your view. Only the distribution of globular clusters told Shapley the sun lay far from the center of the star system. (Daniel Good)

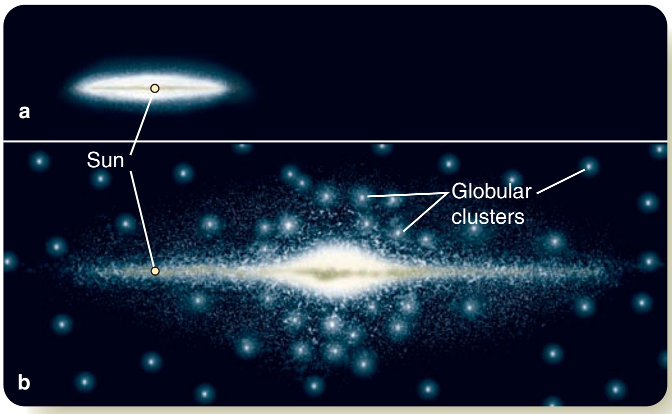


Figure 12-7

(a) Early studies of the star system concluded it was a small cloud of stars and was centered on the sun. (b) Shapley's study of globular clusters revealed that the galaxy was much larger and that the sun did not lie at the center.

were other galaxies like our own. Shapley and astronomer Heber Curtis met in 1920 in a famous debate known as the **Shapley–Curtis debate.** Curtis claimed the faint objects were other galaxies, and Shapley argued they were not other galaxies but were swirls of gas and stars in our star system. It isn't clear who won the debate; but, in 1923, Edwin Hubble photographed individual stars in the Andromeda Galaxy, and in 1924 he identified Cepheids there. Clearly the faint patches were other star systems like our own.

An Analysis of the Galaxy

Our galaxy, like many others, contains two primary components—a disk and a sphere. ■ Figure 12-8 shows these components and other features that you will discover as you analyze the structure of our galaxy.

The **disk component** consists of all matter confined to the plane of the galaxy's rotation—that is, everything in the disk itself. This includes stars, open star clusters, and nearly all of the galaxy's gas and dust. Because the disk contains lots of gas and dust, it is the site of most of the star formation and is illuminated by brilliant, blue, massive stars. Consequently, the disk of the galaxy tends to be blue.

The dimensions of the disk are uncertain for a number of reasons. Its thickness is uncertain because the disk does not have sharp boundaries. Stars become less crowded as you move away from the plane. Also, the thickness depends on what kind of object you study—O stars lie within a narrow disk only about 300 ly thick, but sunlike stars are more widely spread. The diameter of the disk and the position of the sun are also difficult to determine. Gas and dust block the view in the plane of the galaxy so you cannot see to the center or to the edge. The best studies suggest the sun is about 8 kpc from the center, where 1 kpc is a **kiloparsec,** or 1000 pc. Earth seems to be about two-thirds of the way from the center to the edge, so the diameter of our galaxy

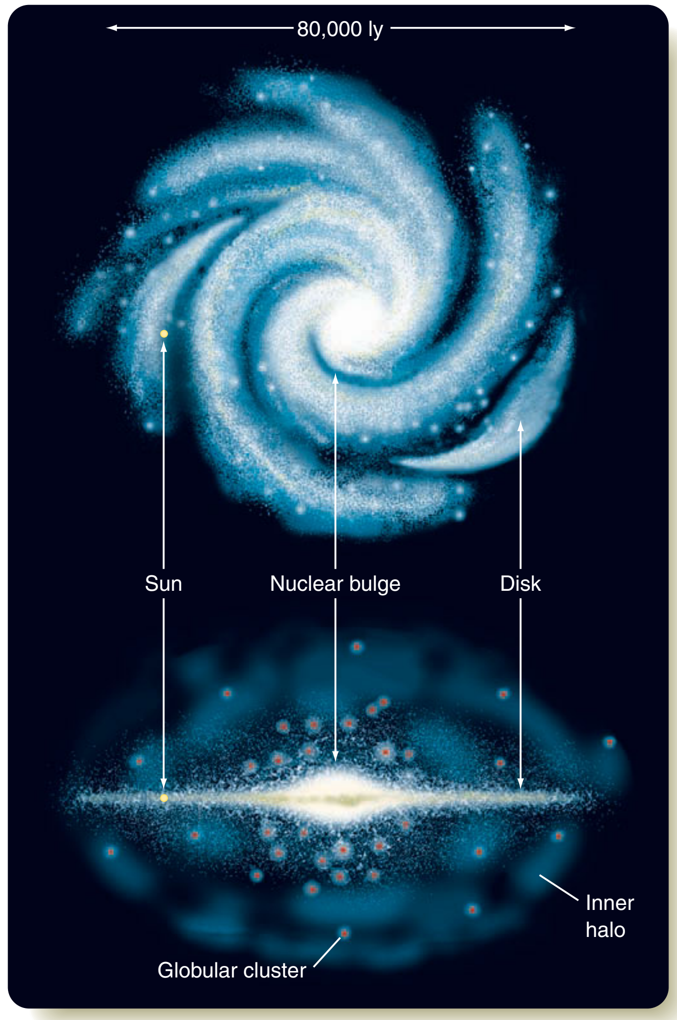


Figure 12-8

An artist's conception of our Milky Way Galaxy, seen face-on and edge-on. Note the position of the sun and the distribution of globular clusters in the halo. Hot blue stars light up the spiral arms. Only the inner halo is shown here. At this scale, the entire halo would be larger than a dinner plate.

appears to be about 25 kpc, which is about 80,000 ly. This is the diameter of the luminous part of our galaxy, the part you would see from a distance. You will learn later that strong evidence suggests that our galaxy is much larger than this but that the outer parts are not luminous.

One way to explore our galaxy is to use radio telescopes to map the distribution of un-ionized hydrogen. Fortunately for astronomers, neutral hydrogen in space is capable of radiating photons with a wavelength of 21 cm. This happens because the spinning proton and electron act like tiny magnets, as shown in ■ Figure 12-9. Depending on whether or not the electron and the proton are spinning in the same or opposite directions, the atom can be in one of two energy levels. One level has slightly higher energy than the other. If left undisturbed, an electron in the higher-energy level will eventually flip its spin and drop to the lower energy level. When that happens, the atom is left with

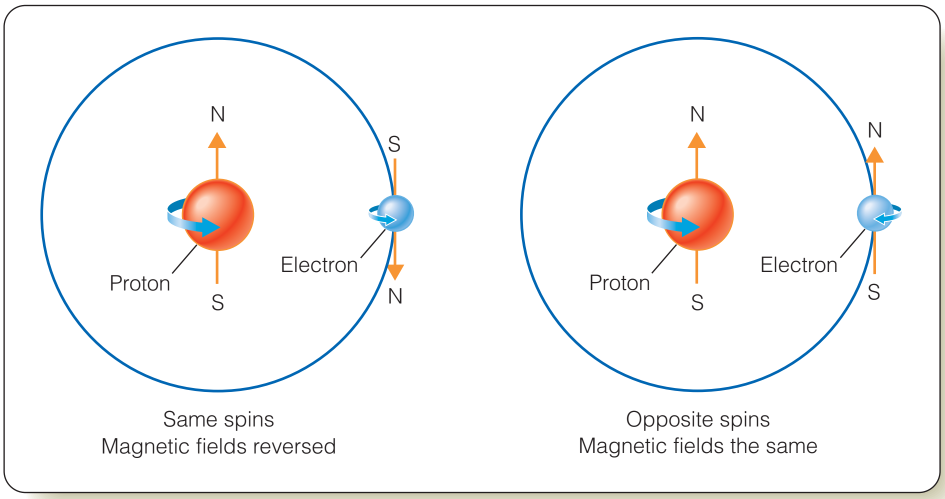


Figure 12-9

Both the proton and the electron in a neutral hydrogen atom spin and consequently have small magnetic fields. Because they have opposite electrostatic charges, they have opposite magnetic fields when they spin in the same direction. When they spin in opposite directions, their magnetic fields are aligned. As explained in the text, this allows cold, neutral hydrogen in space to emit radio photons with a wavelength of 21 cm.

a tiny amount of excess energy, which it radiates as a photon with a wavelength of 21 cm. That means that cold hydrogen gas floating in space naturally emits radio energy, and radio astronomers can map our entire galaxy at a wavelength of 21 cm because these long-wavelength photons are unaffected by the microscopic dust grains that scatter photons of light. Thus, radio telescopes can "see" our entire galaxy, while optical telescopes cannot.

Observations made at other wavelengths can also see through the dust and gas. Infrared photons have wavelengths long enough to be unaffected by the dust. Thus, a map of the sky at long infrared wavelengths reveals the disk of our galaxy (■ Figure 12-10).

The most striking features of the disk component are the **spiral arms**—long curves of bright stars, star clusters, gas, and dust. Such spiral arms are easily visible in other galaxies, and you will see later that our own galaxy has a spiral pattern.

The second component of our galaxy is the **spherical component,** which includes all matter in our galaxy scattered in a roughly spherical distribution around the center. This includes a large halo and the nuclear bulge.

The **halo** is a spherical cloud of thinly scattered stars and globular clusters. It contains only about 2 percent as many stars as the disk of the galaxy and has very little gas and dust. Because of this, no new stars are forming in the halo. In fact, the vast majority of the detectable halo stars are old, cool giants. Detailed studies, however, suggest that most halo stars are lower-main-sequence stars and old white dwarfs that are much too dim to be easily detected. The halos of other galaxies are generally too faint to detect.

The **nuclear bulge** is the dense cloud of stars that surrounds the center of our galaxy. It has a radius of about 2 kpc and is slightly flattened. It is hard to observe because the thick dust in the disk scatters radiation of visible wavelengths, but observations at longer wavelengths can penetrate the dust. The bulge seems to contain little gas and dust, and there is thus little star formation. Most of the stars are old, cool stars like those in the halo.

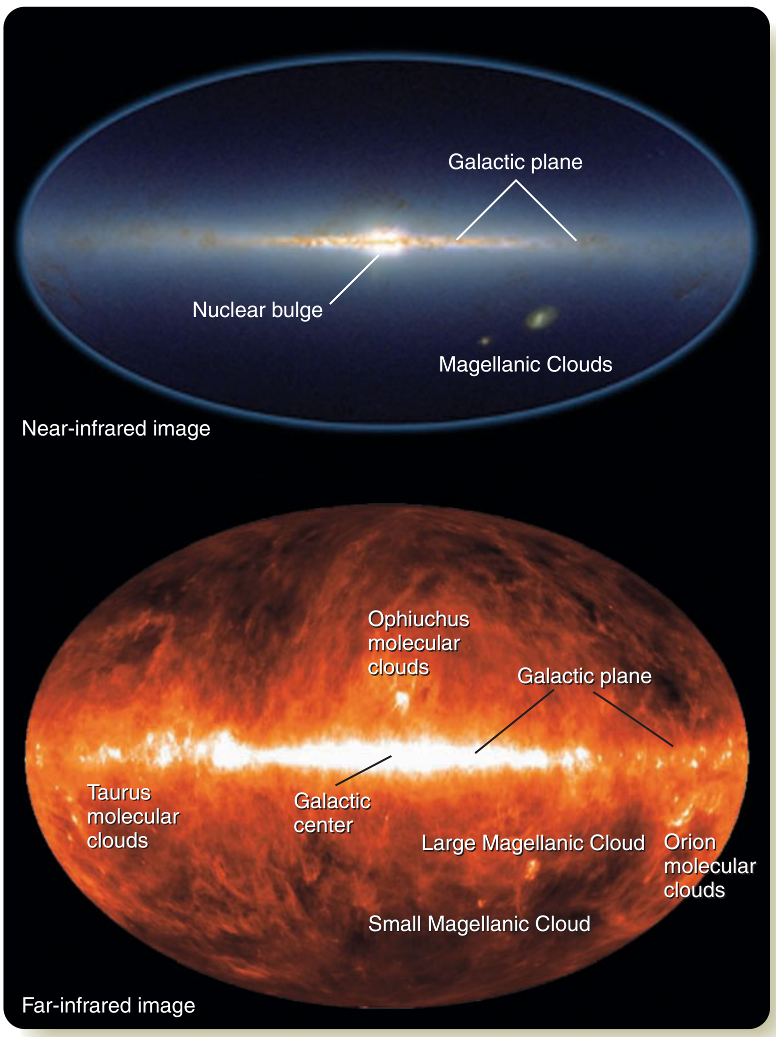


Figure 12-10

In these infrared images, the entire sky has been projected onto ovals with the center of the galaxy at the center of each oval. The Milky Way extends from left to right. In the near-infrared, the nuclear bulge is prominent, and dust clouds block the view along the Milky Way. At longer wavelengths, the dust emits significant blackbody radiation and glows brightly. (Near-IR: 2MASS; Far-IR: DIRBE image courtesy Henry Freudenreich) **Animated!**

The vast numbers of stars in the disk, halo, and nuclear bulge lead to a basic question: How massive is the galaxy?

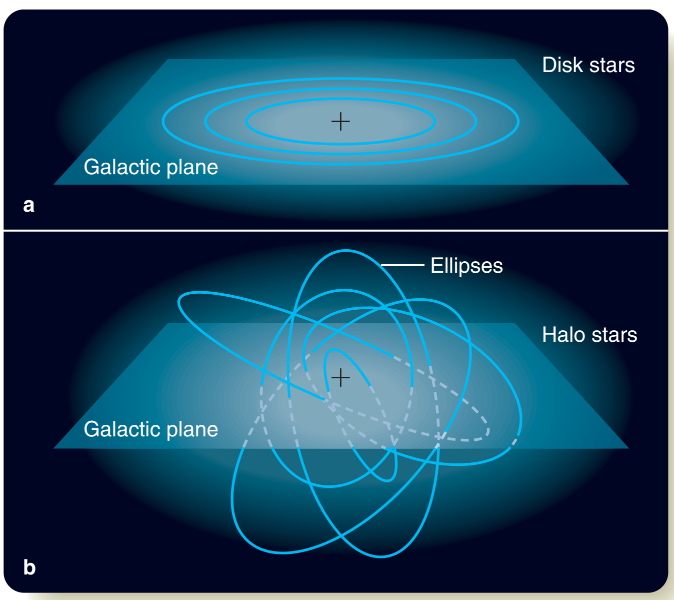
The Mass of the Galaxy

When you needed to find the masses of stars, you studied the orbital motions of pairs of stars in binary systems. To find the mass of the galaxy, you must look at the orbital motions of the stars within the galaxy. Every star in the galaxy follows an orbit around the center of mass of the galaxy.

Astronomers can figure out the orbits of stars by observing how they move. The Doppler effect reveals a star's radial velocity (see Reasoning with Numbers 6-2), which can be combined with the distance to a star and its proper motion, to reveal the size and shape of the star's orbit.

In the halo, each star and globular cluster follows its own randomly tipped elliptical orbit (■ Figure 12-11). These orbits carry the stars and clusters far out into the spherical halo, where they move slowly, but when they fall back into the inner part of the galaxy, their velocities increase. Motions in the halo are like the random motions of a swarm of bees.

In the disk, the stars follow concentric, circular orbits, and those motions can reveal the mass of the galaxy. The sun is a disk star and follows a nearly circular orbit around the galaxy that never carries it out of the disk. By observing the radial velocity of other galaxies in various directions around the sky, astronomers can tell that the sun is moving about 220 km/s in the direction of Cygnus, carrying Earth and the other planets of our solar system along with it. Because its orbit is a circle with a radius of 8 kpc, you can divide the circumference of the orbit by the velocity and find that the sun completes a single orbit in about 220 million years.



■ **Figure 12-11**

(a) Stars in the galactic disk have nearly circular orbits that lie in the plane of the galaxy. (b) Stars in the halo have randomly oriented, elliptical orbits.

If you think of the sun and the center of mass of our galaxy as two objects orbiting each other, you can find the mass of the galaxy (see Reasoning with Numbers 8-4) by dividing the separation in AU cubed by the period in years squared to find the mass in solar masses. This calculation tells you that the galaxy must have a mass of at least 100 billion solar masses.

This estimate is uncertain for a number of reasons. First, astronomers don't know the radius of the sun's orbit with great certainty. Astronomers estimate the radius as 8 kpc, but they could be wrong by 10 percent or more, and this radius gets cubed in the calculation, so it makes a big difference. Second, this estimate of the mass includes only the mass inside the sun's orbit. Mass spread uniformly outside the sun's orbit will not affect its orbital motion. Thus 100 billion solar masses is a lower limit for the mass of the galaxy, but no one knows exactly how much mass lies outside the sun's orbit.

Another thing that makes this mass estimate uncertain is the complex motion of the rotating galaxy. The motions of the stars near the sun show that the disk does not rotate as a solid body. Each star follows its own orbit, and stars at different distances from the center have different orbital periods. This is called differential rotation**.** (Recall that you met the term *differential rotation* when you studied the sun.) Differential rotation is different from the rotation of a solid body—a carousel, for instance. Three wooden horses side-by-side on a carousel will stay together as the carousel turns, but three stars lined up in the galaxy will draw apart because they travel at different orbital velocities and have different orbital periods.

A graph of the orbital velocity of stars at various orbital radii in the galaxy is called a **rotation curve** (■ Figure 12-12). If all of the mass in the galaxy were concentrated at its center, then orbital velocity would be high near the center and would decline as you moved outward. This kind of motion has been called Keplerian motion because it follows Kepler's third law. A good example is our own solar system, where nearly all of the mass is in the sun. Of course, the galaxy's mass is not all concentrated at its center, but if most of the mass were inside the orbit of the sun, then you would expect to see orbital velocities decline at greater distances. Many observations confirm, however, that velocities do not decline and may actually increase at greater distance; this observation shows that these larger orbits are enclosing more mass. Although it is difficult to determine a precise edge to the visible galaxy, it seems clear that large amounts of matter are located beyond the traditional edge of the galaxy—the edge you would see if you journeyed into space and looked back at our galaxy.

The evidence is clear that extra mass lies in an extended halo sometimes called a **dark halo** or **galactic corona.** It may extend up to ten times farther than the edge of the visible disk and could contain up to two trillion solar masses. Some small fraction of this mass is made up of low-luminosity stars and white dwarfs,

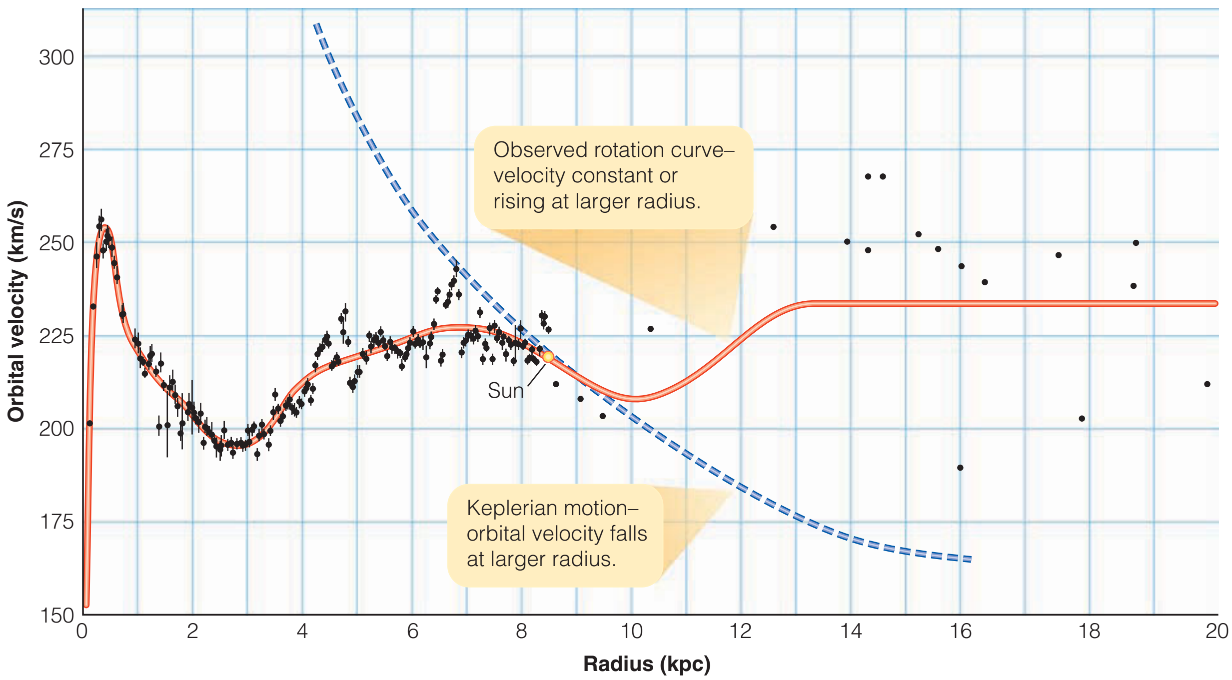


Figure 12-12

The rotation curve of our galaxy is plotted here as orbital velocity versus radius. Data points show measurements made by radio telescopes. Observations outside the orbit of the sun are much more uncertain, and the data points scatter widely. Orbital velocities do not decline outside the orbit of the sun, as you would expect if most of the mass of the galaxy were concentrated toward the center (Keplerian motion). Rather, the curve is approximately flat at great distances, suggesting that the galaxy contains significant mass outside the orbit of the sun. (Adapted from a diagram by Françoise Combes)

but most of the matter is not producing any light. Astronomers call it **dark matter** and conclude that it must be some as yet unknown form of matter. You will continue to learn about dark matter in the following three chapters. It is one of the fundamental problems of modern astronomy.

12-2 The Origin of the Milky Way

Just as paleontologists reconstruct the history of life on Earth from the fossil record, astronomers try to reconstruct our galaxy's past from the fossil it left behind as it formed and evolved. That fossil is the spherical component of the galaxy. The stars in the halo formed when the galaxy was young, so their chemical composition and distribution provide clues to the birth of the galaxy.

The Age of the Milky Way

To begin, you should ask yourself how old our galaxy is. That question is easy to answer because you already know how to find the age of star clusters. But there are uncertainties that make that easy answer hard to interpret.

The oldest open clusters in our galaxy have ages of about 9 billion years. These ages are determined by analyzing the turnoff point in the cluster H–R diagram (see Chapter 10), but three things make these ages uncertain. First, finding the age of a star cluster becomes more difficult for older clusters because they change more slowly. Also, the locations of their turnoff points and giant regions depend on their chemical composition, which can differ slightly from cluster to cluster. Finally, open clusters are not strongly bound by their own gravity, so there may have been older open clusters that dissipated as their stars wandered away. In any case, from open clusters, you can get a rough age for the disk of our galaxy of about 9 billion years.

Globular clusters have faint turnoff points in their H–R diagrams and are clearly old, but finding their ages is difficult. Clusters differ slightly in chemical composition, which must be accounted for when calculating the stellar models from which ages are determined. Also, to find the age of a cluster, astronomers must know the distance to the cluster, and globular clusters are too far away to have measurable parallaxes. The HIPPARCOS satellite has, however, measured precise parallaxes for Cepheid variable stars, and that has improved the calibration of these distance indicators. In addition, careful studies with the newest large telescopes have refined the H–R diagrams of globular clusters. Globular cluster ages seem to average about 11 billion years. There is still some uncertainty in the measurements of ages, but some clusters do seem to be older than others. The oldest are a bit over 13 billion years old, so the halo of our galaxy must be at least 13 billion years old.

The ages of star clusters tell you that the disk is younger than the halo and show that our galaxy is at least 13 billion years old. Now you can combine these ages with subtle differences in the chemical compositions of stars to tell the story of our galaxy.

Stellar Populations

In the 1940s, astronomers realized that there are two types of stars in the galaxy. The stars that are closest, brightest, and most easily studied are located in the disk. These are called **population I** stars. The second type are much more distant, fainter, and are found in the halo, in globular clusters, or in the central bulge. They are called **population II** stars. It is significant that the two stellar populations are associated with the two components of the galaxy.

The stars of the two populations are very similar. They fuse the same nuclear fuels and evolve in nearly identical ways. They differ only in their abundance of atoms heavier than helium, atoms that astronomers refer to collectively as **metals.** (Note that this is not the way the word *metal* is commonly used by nonastronomers.) Population I stars are metal rich, containing 2 to 3 percent metals, while population II stars are metal poor, containing only about 0.1 percent metals or less. The population membership of a star is defined by its metal content and not by its location in the disk or spherical component.

Population I stars belong to the disk component of the galaxy and are sometimes called disk population stars. They have circular orbits in the plane of the galaxy and are relatively young stars that formed within the last few billion years. The sun is a population I star.

Population II stars belong to the spherical component of the galaxy and are sometimes called halo population stars. These stars have randomly tipped orbits with a wide range of shapes. A few follow circular orbits, but most follow elliptical orbits. The population II stars are all lower-mass main-sequence stars or giants. They are old stars. The metal-poor globular clusters are part of the halo population.

You can now understand why there are two kinds of Cepheid variables shown in the period–luminosity diagram (Figure 12-4). Type I Cepheids are population I stars with metal abundance similar to that of the sun. The type II Cepheids are population II stars. The difference in metal abundance affects the opacity of the gas and thus the ease with which radiation can get through, and that changes the balance between gravity and energy flowing outward through the star. A few percent difference in metal abundance may seem like a small detail, but it makes a big difference to a star.

Since the discovery of stellar populations, astronomers have realized that there is a gradation between populations (■ Table 12-1). The most metal-rich stars are called extreme population I stars. Slightly less metal-rich population I stars are called intermediate population I stars. The sun is such a star. Stars even less metal rich, such as stars in the nuclear bulge, are intermediate population II stars. The most metal-poor stars are those in the halo and in globular clusters. They are extreme population II stars.

Why do the disk and halo stars have different metal abundances? They must have formed at different stages in the life of the galaxy, at times when the chemical composition of the galaxy differed. This is a clue to the history of our galaxy; but, to use the clue, you must discuss the cycle of element building.

The Element-Building Cycle

The atoms of which you are made were created in a process that spanned a number of generations of stars. Natural processes are all around you, and you must learn to recognize and understand them if you are to understand nature. The process that built the chemical elements may be one of the most important processes in the history of our galaxy (■ How Do We Know? 12-2).

When you studied the evolution of giant stars, you learned how small amounts of elements heavier than helium—the elements astronomers call metals—are cooked up during helium fusion. Additional heavy atoms are made by short-lived nuclear reactions that occur during a supernova explosion. When stars die, small amounts of these elements are spread back into the interstellar medium. Lower-mass atoms like carbon, nitrogen, and oxygen are common, but atoms significantly more massive than iron—such as gold, silver, platinum, and uranium—are rare because they are made only during supernovae. ■ Figure 12-13a shows the abundance of the chemical elements. This graph is usually drawn with an exponential scale as in part *a* of the figure. To get a feeling for the true abundance of the elements, you should draw the graph

■ **Table 12-1** | **Stellar Populations**

	Population I		Population II	
	Extreme	Intermediate	Intermediate	Extreme
Location	Spiral arms	Disk	Nuclear bulge	Halo
Metals (%)	3	1.6	0.8	Less than 0.8
Shape of orbit	Circular	Slightly elliptical	Moderately elliptical	Highly elliptical
Average age (yr)	100 million and younger	0.2–10 billion	2–10 billion	10–13 billion

How Do We Know? 12-2

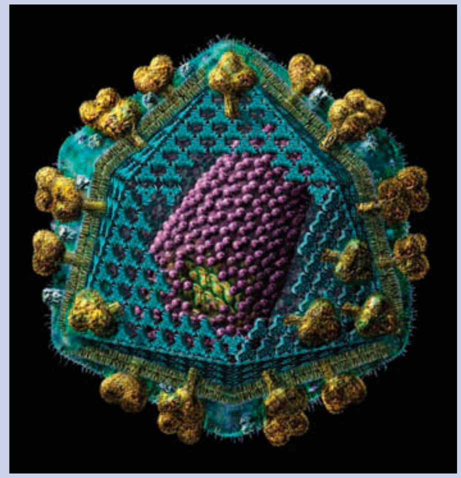
Nature as Processes

How does understanding a natural process unify facts and theories? Science, at first glance, seems to be nothing but facts, but in many cases you can organize the facts into the story of a process. For example, astronomers assemble the sequence of events that led to the formation of the chemical elements. If you understand that process, you have command over a lot of important facts and theories in astronomy.

A process is a sequence of events that leads to some result or condition, and much of science is focused on understanding how natural processes work. Biologists, for example, might try to understand how a virus reproduces. They must figure out how the virus tricks the immune system into leaving it alone, penetrates the wall of a healthy cell, injects its viral DNA, commandeers the cell's resources to make new viruses, and finally destroys the cell to release the new virus copies. A biologist may spend a lifetime studying a specific step, but the ultimate goal of science is to tell the entire story of the process.

As you study any science, be alert for the way processes organize scientific knowledge. When you see a process in science, ask yourself a few basic questions. What conditions prevailed at the beginning of the process? What sequence of steps occurred? Can some steps occur simultaneously, or must one step occur before another? What is the final state that this process produces?

Recognizing a process and learning to tell its story will help you remember a lot of details, but that is not its real value. Identifying a process and learning to tell its story helps you understand how nature works and explains why the universe is the way it is.



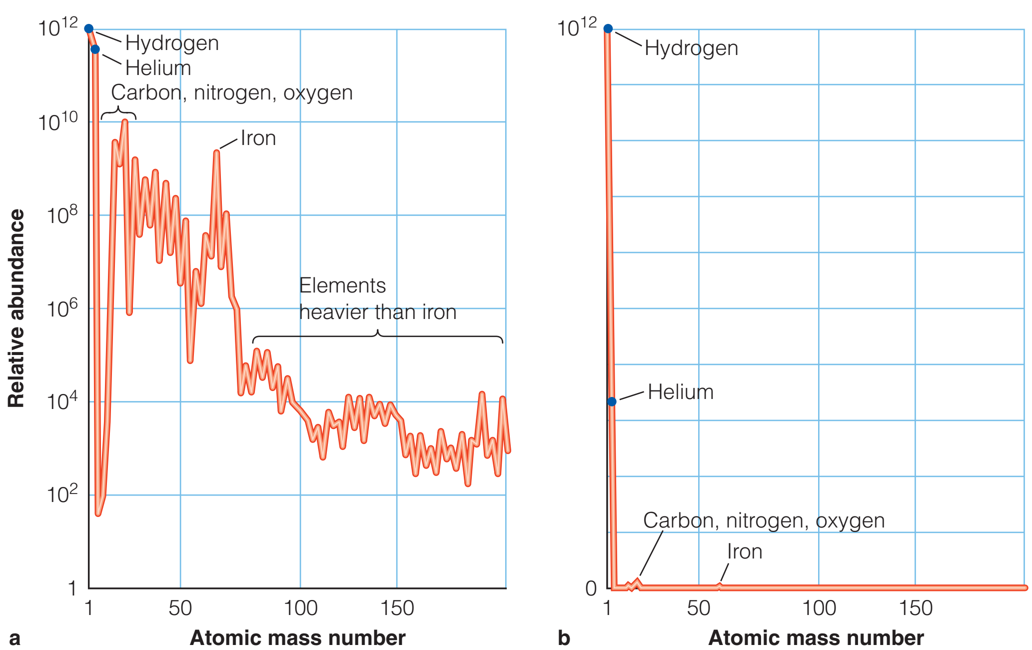
A virus is a collection of molecules that cannot reproduce until they penetrate into a living cell. The virus shown in this artist's conception causes HIV-AIDS. (Russell Knightly Media, rkm.com.au)

using a linear scale as in Figure 12-13b. Then you see how rare the elements heavier than helium are.

When the galaxy first formed, there should have been no metals because stars had not yet manufactured any. The gas from which the galaxy condensed must have contained about 90 percent hydrogen atoms and 10 percent helium atoms. (The hydrogen and helium came from the big bang that began the universe, which you will study in a later chapter.)

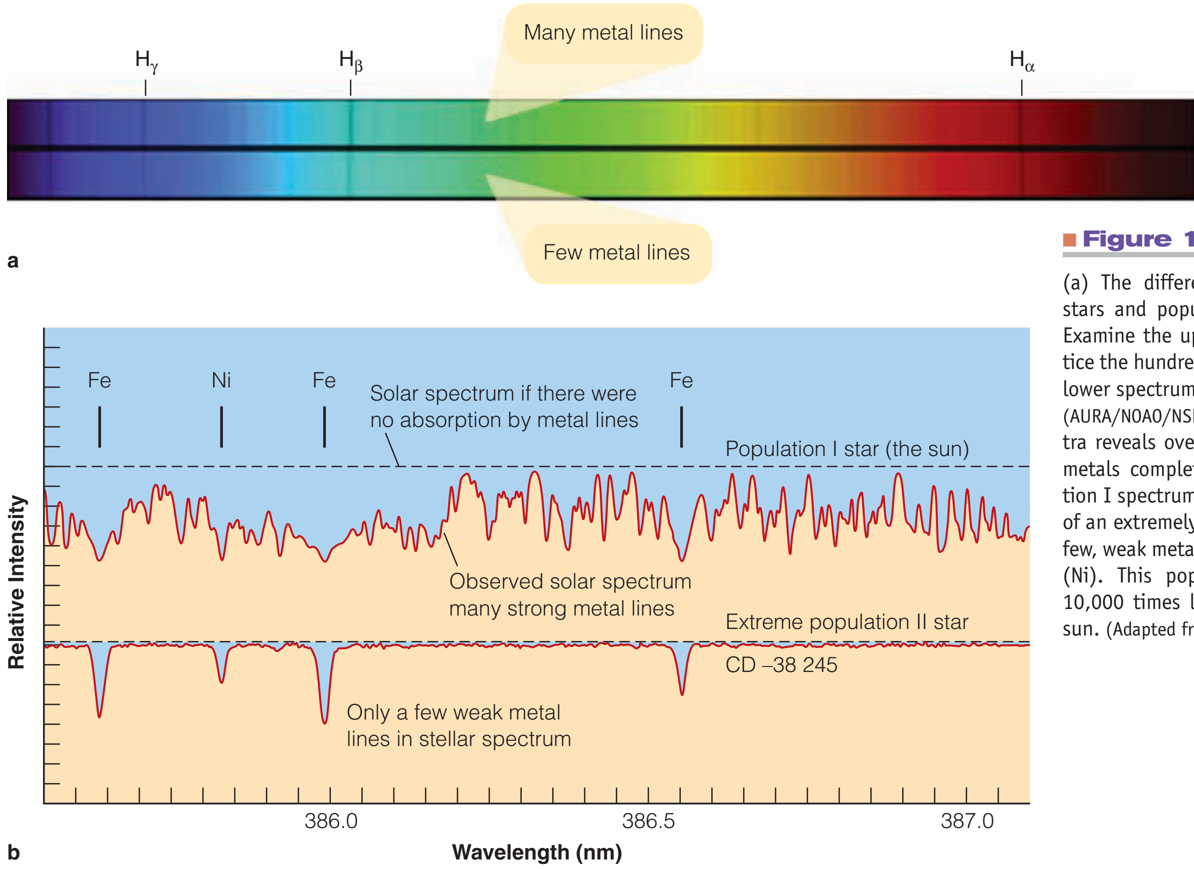
The first stars to form from this gas were metal poor. The more massive stars died, and now, 13 billion years later, only lower-mass stars are left with spectra that show few metal lines. Of course, they may have manufactured some atoms heavier than helium, but because the stars' interiors are not mixed, those heavy atoms stay trapped at the centers of the stars where they were produced and do not affect the spectra (■ Figure 12-14). The extreme population II stars in the halo are the survivors of an early generation of stars to form in the galaxy.

Most of the first stars evolved and died enriching the interstellar gas with metals. Succeeding generations of stars formed from gas clouds that were enriched, and each generation added to the enrichment. By



■ Figure 12-13

The abundance of the elements in the universe. (a) When the elements are plotted on an exponential scale, you see that elements heavier than iron are about a million times less common than iron and that all elements heavier than helium (the "metals") are quite rare. (b) The same data plotted on a linear scale provide a more realistic impression of how rare the metals are. Carbon, nitrogen, and oxygen make small peaks near atomic mass 15, and iron is just visible in the graph.



■ **Figure 12-14**

(a) The difference between population I stars and population II stars is dramatic. Examine the upper spectrum here and notice the hundreds of faint spectral lines. The lower spectrum has fewer and weaker lines. (AURA/NOAO/NSF) (b) A graph of such spectra reveals overlapping absorption lines of metals completely blanketing the population I spectrum. The lower spectrum is that of an extremely metal-poor star with only a few, weak metal lines of iron (Fe) and nickel (Ni). This population II star has about 10,000 times less metal content than the sun. (Adapted from an ESO illustration)

the time the sun formed, roughly 5 billion years ago, the element-building process had added about 1.6 percent metals to the interstellar medium. Since then, the metal abundance has increased further, and stars forming now incorporate 2 to 3 percent metals and become extreme population I stars. Thus metal abundance varies between populations because of the accumulation of heavy atoms produced in successive generations of stars.

The lack of metals in the spherical component of the galaxy tells you it is very old, a fossil left behind by the galaxy when it was young and drastically different from its present disk shape. The study of element building and stellar populations leads to the fundamental question, "How did our galaxy form?"

The History of the Milky Way Galaxy

In the 1950s, astronomers began to develop a hypothesis to explain the formation of our galaxy. Recent observations, however, are forcing a reevaluation of that traditional hypothesis.

The traditional hypothesis says that the galaxy formed from a single large cloud of gas over 13 billion years ago. As gravity pulled the gas together, the cloud began to fragment into smaller clouds, and because the gas was turbulent the smaller clouds had random velocities. Stars and star clusters that formed from these fragments went into randomly shaped and randomly tipped orbits. These first stars were metal poor because no earlier stars had existed to enrich the gas with metals. Some of these stars remain as the population II stars observed today in the halo.

The contracting gas cloud was roughly spherical at first, but the turbulent motions canceled out, as do eddies in recently stirred coffee, leaving the cloud with uniform rotation. A rotating, low-density cloud of gas cannot remain spherical. A star is spherical because its high internal pressure balances its gravity; but, in a low-density cloud, the pressure cannot support the weight. Like a blob of pizza dough spun in the air, the cloud had to flatten into a disk (■ Figure 12-15).

This contraction into a disk took billions of years, with the metal abundance slowly increasing as generations of stars were born from the gradually flattening gas cloud. The stars and globular clusters of the halo formed by the cloud when it was spherical were left behind, and subsequent generations of stars formed in flatter distributions. The gas distribution in the galaxy now is so flat that the youngest stars are confined to a thin disk only about 100 parsecs thick. These stars are metal rich and have nearly circular orbits.

This traditional hypothesis accounts for many of the Milky Way's properties. Advances in technology, however, have improved astronomical observation, and, beginning in the 1980s, contradictions between theory and observation arose. For example, improved observations show that globular clusters have a range of ages, but the traditional hypothesis suggests that the

Origin of the Halo and Disk

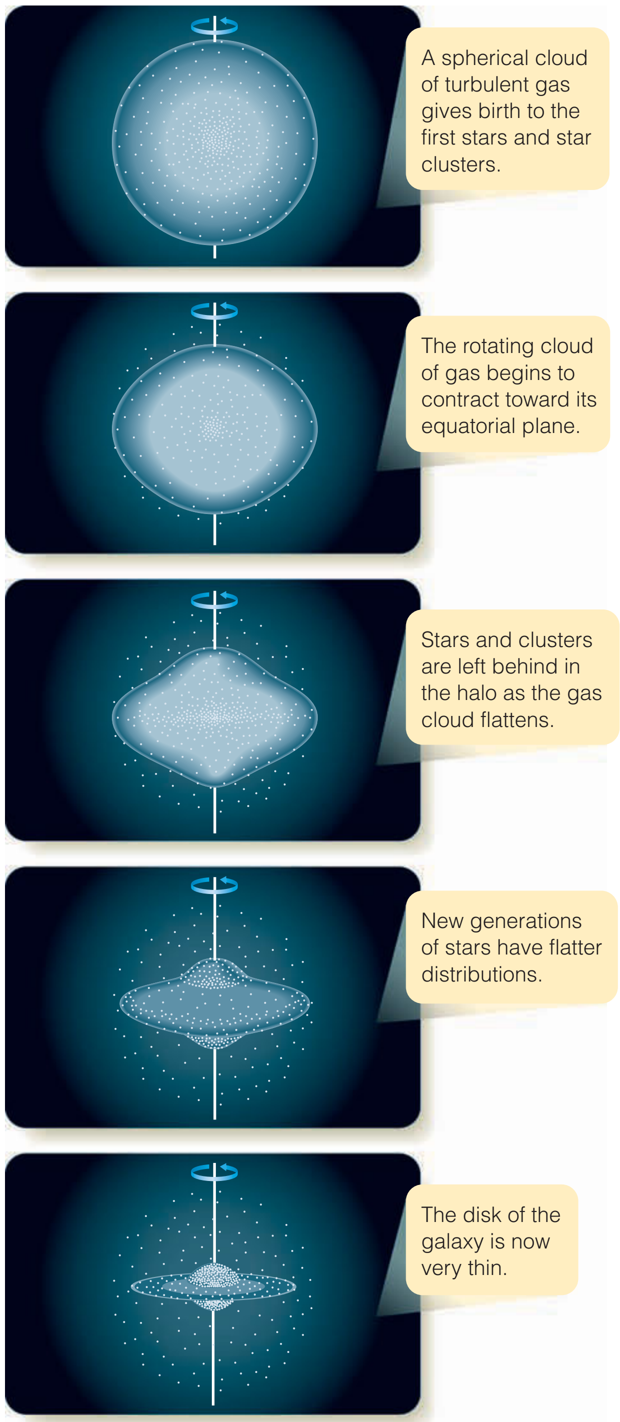


Figure 12-15

The traditional theory for the origin of our galaxy begins with a spherical gas cloud that flattens into a disk.

globular clusters should have roughly similar ages. Also, some of the oldest stars in the galaxy are in the nuclear bulge, not in the halo, but the theory says the halo formed first. Furthermore, the oldest clusters in the disk are much younger than the halo, but the traditional hypothesis suggests that star formation was continuous as the galaxy flattened. New observations show that some halo stars are even more metal poor than the globular clusters, so those stars must have formed before the globular clusters. Also, even the most metal-poor stars in the halo are not metal free, so there must have been a generation of stars that manufactured those metals before the halo formed. The traditional hypothesis says nothing about those first stars.

Can the traditional hypothesis be modified to account for these observations? Perhaps the galaxy began with the contraction of a gas cloud to form the central bulge, and the halo formed slightly later from infalling clouds of gas and stars that had been slightly enriched in metals by an early generation of massive stars. That first generation of stars would have formed from almost pure hydrogen and helium gas, which would have been so transparent to energy that the stars would have formed with high masses. Those stars would have lived very short lives, so none would be left today. The metals they made during their short lives would explain the metals observed in the oldest stars.

Observations made as part of the Sloan Digital Sky Survey have allowed astronomers to compute the motions of over 20,000 halo stars, and those motions reveal that the inner part of the halo, closer than 50,000 ly to the center, has a gentle rotation in the same direction the sun orbits. But the outer halo tends to rotate in the opposite direction. The stars in these two regions also have slight chemical differences. That suggests that the halo was built gradually as smaller clouds of gas and stars fell together.

The disk could have formed later as the gas that was already in the galaxy flattened into a disk and as more gas fell into the galaxy and settled into the disk. Perhaps entire galaxies were captured by the growing Milky Way Galaxy. (You will see in the next chapter that such mergers do occur.) If our galaxy absorbed a few small but partially evolved galaxies, then some of the globular clusters in the halo may be hitchhikers. This would explain the range of globular cluster ages.

Astronomers are slowly puzzling out the mystery of our galaxy's origin, and they are also solving the mystery at the center of the galaxy.

◄ SCIENTIFIC ARGUMENT ►

Why do metal-poor stars have the most elongated orbits?

A good argument makes connections between ideas, and sometimes those connections are not obvious. Certainly, the metal abundance of a star cannot affect its orbit, so an analysis must not confuse cause and effect with the relationship between these two factors. Both chemical composition and orbital shape depend on a third factor—age. The oldest stars are metal poor because they formed before many stars could create and scatter metals into the interstellar medium. Also, the oldest stars formed long ago when the galaxy was young and motions were not organized into a disk, and they tended to take up randomly shaped orbits, many of which are quite elongated. Consequently, today, the most metal-poor stars tend to follow the most elongated orbits.

Nevertheless, even the oldest stars known in our galaxy contain some metals. They are metal poor, not metal free. Adjust your argument. **Where did these metal-poor stars get their metals?**

◄ ►

12-3 The Nucleus

The most mysterious region of our galaxy is its very center, the nucleus. At visual wavelengths, this region is totally hidden by gas and dust that dim the light it emits by 30 magnitudes (Figure 12-6b). If a trillion (10^{12}) photons of light left the center of the galaxy on a journey to Earth, only one would make it through the gas and dust. The longer-wavelength infrared photons are scattered much less often; one in every ten makes it to Earth. Consequently, visual-wavelength observations reveal nothing about the nucleus, but it can be observed at longer wavelengths such as in the infrared and radio parts of the spectrum.

The Center of the Galaxy

Harlow Shapley's study of globular clusters placed the center of our galaxy in Sagittarius, and the first radio maps of the sky showed a powerful radio source located in Sagittarius. Infrared surveys detected an intense flood of radiation coming from the same source, and higher-resolution radio maps revealed a complex collection of radio sources, with one, **Sagittarius A*** (abbreviated Sgr A* and usually pronounced "sadge A-star"), lying at the expected location of the galactic core.

High-resolution observations made with the VLBA, the radio interferometer that stretches from Hawaii in the Pacific to the Virgin Islands in the Atlantic, show that Sgr A* is no more than one astronomical unit in diameter but is a powerful source of radio energy. What could be as small as Sgr A* and produce so much energy?

Study ■ **Sagittarius A*** on pages 250–251 and notice three important points:

1. First, observations at radio wavelengths reveal complex structures near Sgr A* caused by magnetic fields and by rapid star formation. Supernova remnants show that massive stars have formed there recently and died supernova deaths.
2. The center is crowded. Tremendous numbers of stars heat the dust and produce strong infrared radiation.
3. Finally, there is evidence that Sgr A* is a supermassive black hole into which gas is flowing.

A supermassive black hole is an exciting idea, but scientists must always be aware of the difference between adequacy and necessity. A supermassive black hole is adequate to explain the observations, but is it necessary? Could there be some other way to explain what is observed? For example, astronomers have suggested that gas flowing inward could trigger tremendous bursts of star formation. Such theories have been considered and tested against the evidence, but none appears to be adequate to explain the observations. So far, the only theory that seems adequate is that our galaxy is home to a supermassive black hole.

Meanwhile, observations are allowing astronomers to improve their models. For instance, Sgr A* is not as bright in X-rays as it should be if it has a hot accretion disk with matter constantly flowing into the black hole. Observations of X-ray and infrared flares lasting only a few hours suggest that mountain-size blobs of matter may occasionally fall into the black hole and be ripped apart and heated by tides. But the black hole may be mostly dormant and lack a fully developed hot accretion disk because little matter is flowing into it at the present time.

Such a supermassive black hole could not be the remains of a single dead star. It contains much too much mass. It probably formed when the galaxy first formed over 13 billion years ago.

The center of our galaxy is mysterious because it is hard to observe and hard to understand. Similarly, the spiral arms of our galaxy have given astronomers a challenging puzzle.

12-4 Spiral Arms and Star Formation

The most striking features of galaxies like the Milky Way are the beautiful spiral arms that wind outward and contain swarms of hot, blue stars; clouds of dust and gas; and young star clusters. These young objects hint that the spiral arms are regions of star formation. As you try to understand the spiral arms, you face two problems. First, how can you be sure our galaxy has spiral arms when our view is obscured by gas and dust? Second, why doesn't the differential rotation of the galaxy destroy the arms? The solution to both problems involves star formation.

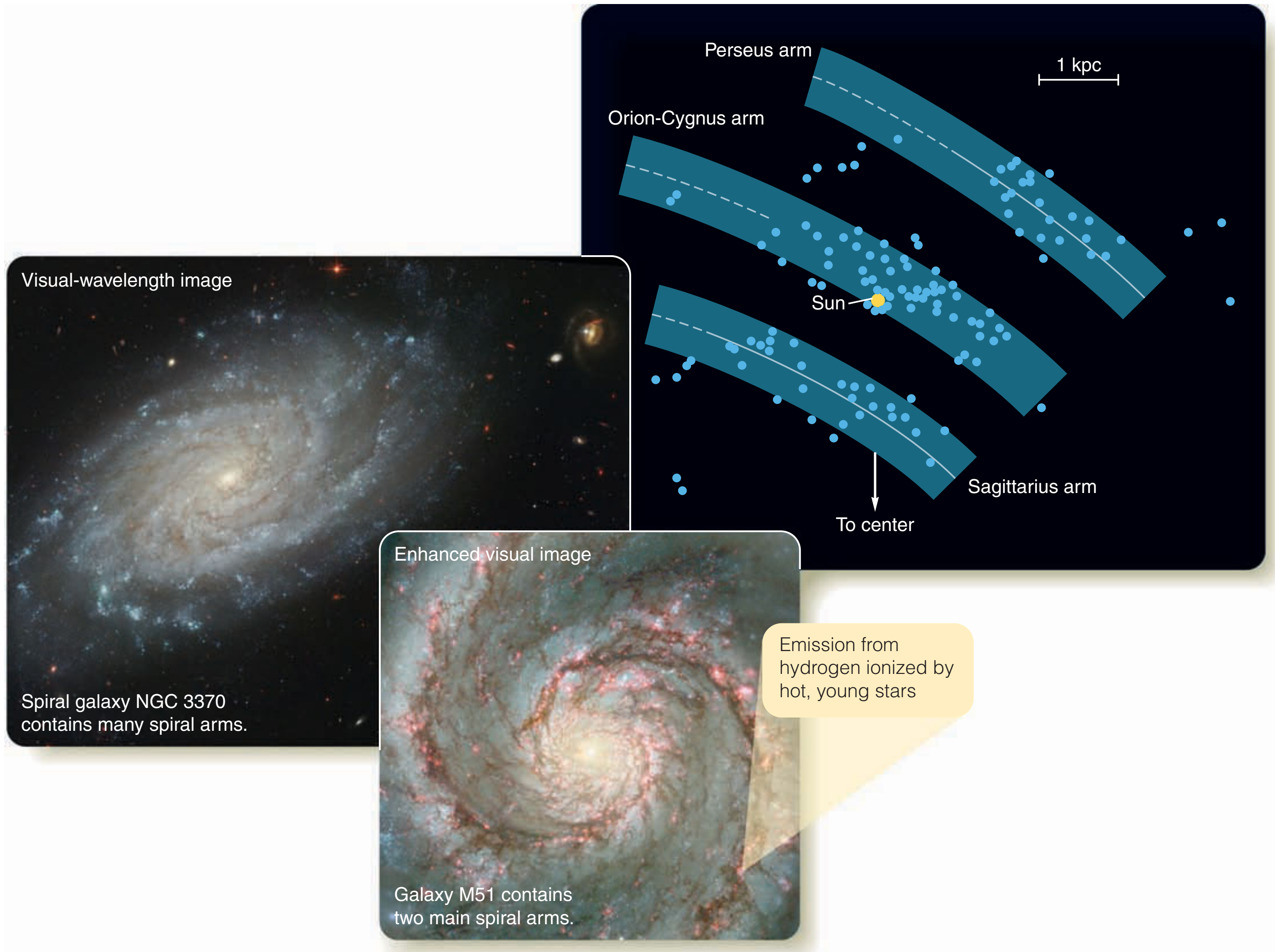
Tracing the Spiral Arms

Studies of other galaxies show that spiral arms contain hot, blue stars. Thus, one way to study the spiral arms of our own galaxy is to locate these stars. Fortunately, this is not difficult, since O and B stars are often found in associations and, being very bright, are easy to detect across great distances. Unfortunately, at these great distances their parallax is too small to measure, so their distances must be found by other means, usually spectroscopic parallax.

The O and B associations visible in the sky are not located randomly but fall along parts of three spiral arms near the sun, which have been named for the prominent constellations through which they pass (■ Figure 12-16). If you could penetrate the gas and dust, you could locate other O and B associations and trace the spiral arms farther; but, like a traveler in a fog, you see only the region nearby.

Objects used to map spiral arms are called **spiral tracers.** O and B associations are good spiral tracers because they are bright and easy to see at great distances. Other tracers include young open clusters, clouds of hydrogen ionized by hot stars (emission nebulae), and certain Cephid variable stars.

Notice that all spiral tracers are young objects. O stars, for example, live for only a few million years. If their orbital velocity



■ **Figure 12-16**

Many of the galaxies in the sky are disk shaped, and most of those galaxies have spiral arms. You can suspect that our own disk-shaped galaxy also has spiral arms. Images of other galaxies show that spiral arms are marked by hot, luminous stars that must be very young, and this should make you suspect that spiral arms are related to star formation. Gas and dust block our view of most of the disk of our galaxy, but near the sun in space young O and B stars fall along bands that appear to be segments of spiral arms. (Images: NASA, Hubble Heritage Team)

is about 200–250 km/s, they cannot move more than about 500 pc in their lifetimes. This is less than the width of a spiral arm, which means that the O and B stars must have formed in the spiral arms.

O and B associations trace out the nearby segments of spiral arms, but other methods allow astronomers to map spiral arms across the entire galaxy.

Go to **academic.cengage.com astronomy/seeds** to see Astronomy Exercise "Milky Way Galaxy" and see a 3-D model of your home star system.

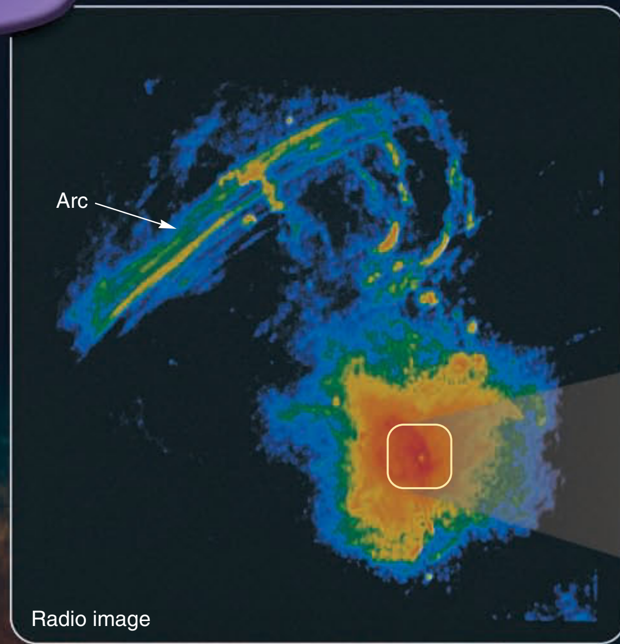
Radio Maps of Spiral Arms

The dust clouds that block the view at visual wavelengths are transparent at radio wavelengths because radio waves are much longer than the diameter of the dust particles. If you pointed a radio telescope at a section of the Milky Way, you would receive 21-cm radio signals coming from cool hydrogen in spiral arms at various distances across the galaxy. Fortunately, the signals can be unscrambled by measuring the Doppler shifts of the 21-cm radiation. In the direction toward the nucleus, however, the orbital motions of gas clouds are perpendicular to the line of sight, and all the radial velocities are zero. That is why the radio map shown in ■ Figure 12-17 reveals spiral arms throughout the disk of the galaxy but not in the wedge-shaped region toward the center.

The analysis of the 21-cm radial velocity data requires that astronomers estimate the orbital velocity of the clouds at different distances from the center of the galaxy. Because these velocities are not precisely known and because turbulent motions in the gas distort the radial velocities, you can't trace a perfect spiral pattern in Figure 12-17a. The radio map does, however, show

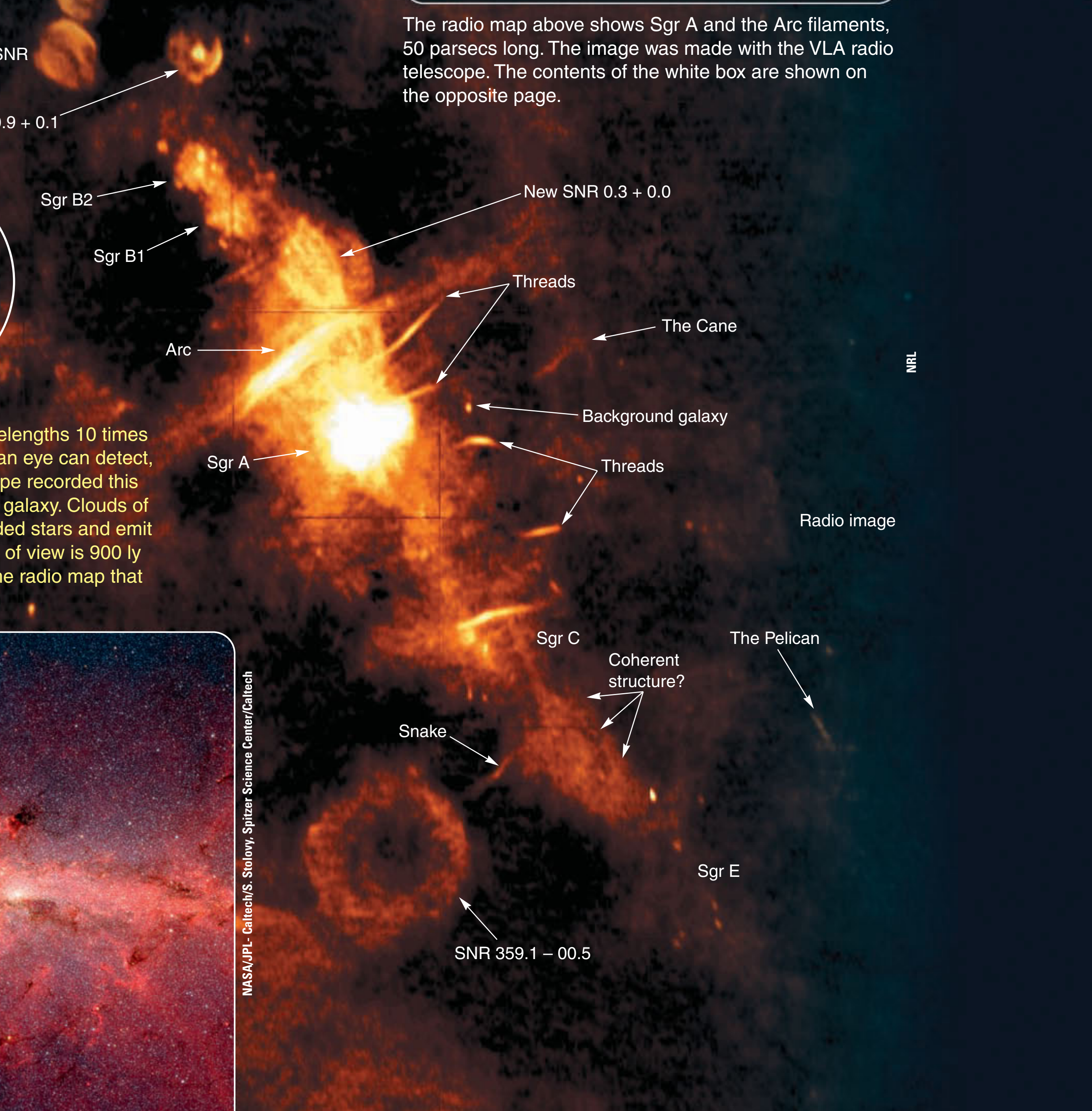
1 The constellation Sagittarius is so filled with stars and with gas and dust you can see nothing at visual wavelengths of the center of our galaxy.

The image below is a wide-field radio image of the center of our galaxy. Many of the features are supernova remnants (SNR), and a few are clouds of star formation. Peculiar features such as threads, the Arc, and the Snake may be gas trapped in magnetic fields. At the center lies Sagittarius A, the center of our galaxy.



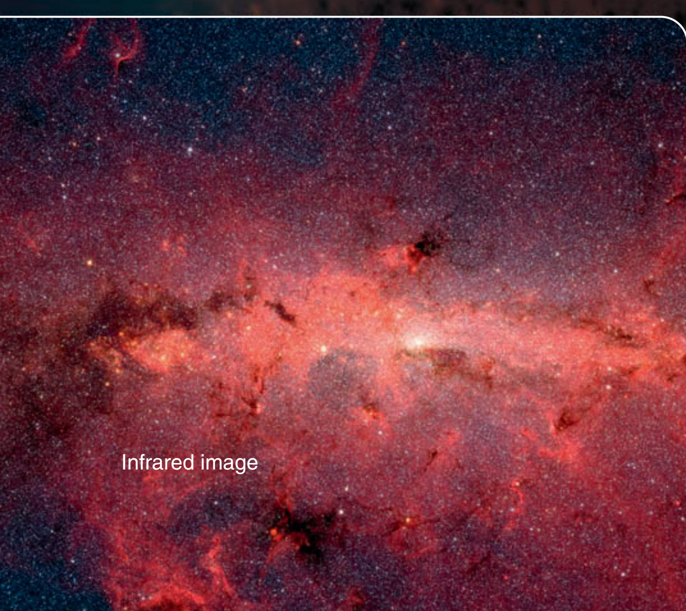
NRAO/AUI/NSF

The radio map above shows Sgr A and the Arc filaments, 50 parsecs long. The image was made with the VLA radio telescope. The contents of the white box are shown on the opposite page.



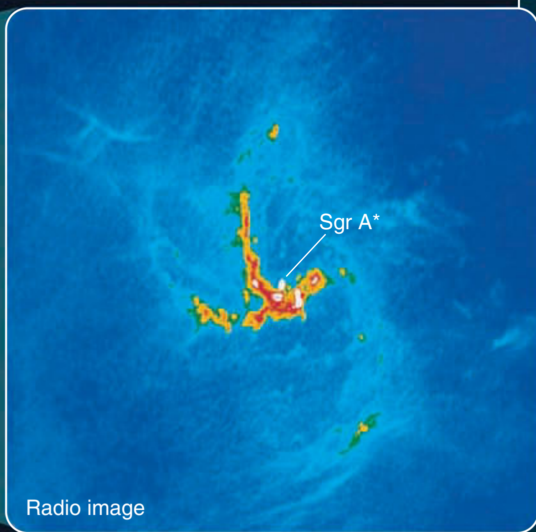
NRL

2 Observing at wavelengths 10 times longer than the human eye can detect, the Spitzer Space Telescope recorded this image of the center of our galaxy. Clouds of dust are warmed by crowded stars and emit infrared photons. The field of view is 900 ly wide—much larger than the radio map that fills most of this page.

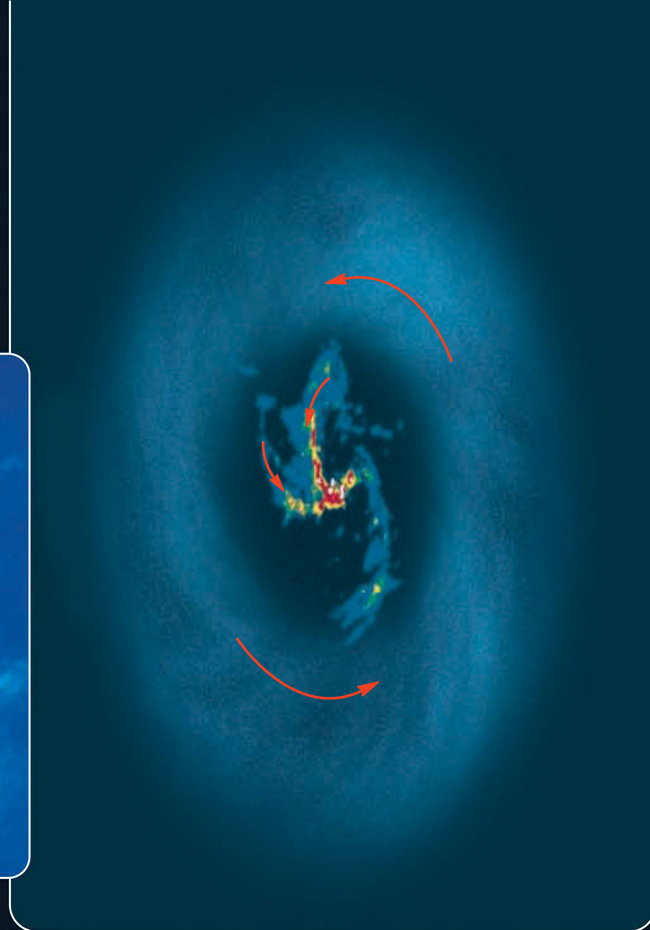


NASA/JPL- Caltech/S. Stolovy, Spitzer Science Center/Caltech

1a This high-resolution radio image of Sgr A (the white boxed area on the opposite page) reveals a spiral swirl of gas around an intense radio source known as Sgr A*, the presumed central object in our galaxy. About 3 pc across, this spiral lies in a low-density cavity inside a larger disk of neutral gas. The arms of the spiral are thought to be streams of matter flowing into Sgr A* from the inner edge of the larger disk (drawing at right).



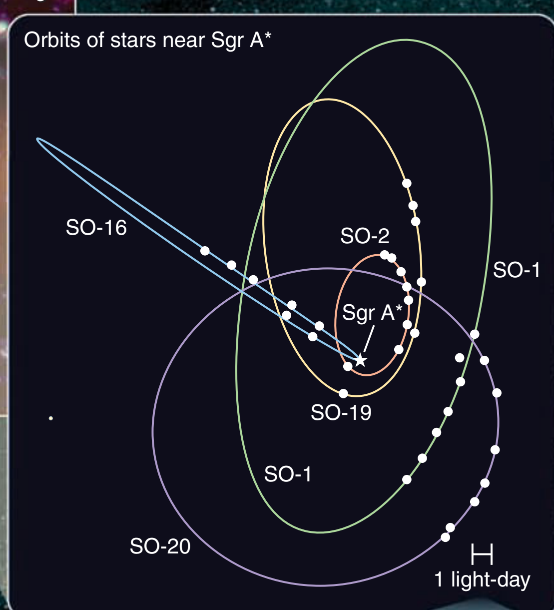
N. Killeen and Kwok-Yung Lo



Evidence of a Black Hole at the Center of Our Galaxy

3 Since the middle 1990s, astronomers have been able to use large infrared telescopes and active optics to follow the motions of stars orbiting around Sgr A*. A few of those orbits are shown here. The size and period of the orbit allows astronomers to calculate the mass of Sgr A* using Kepler's third law. The orbital period of the star SO-2, for example, is 15.2 years and the semimajor axis of its orbit is 950 AU. The combined motions of the observed stars suggest that Sgr A* has a mass of 3.7 million solar masses.

ESO



Our solar system is half a light-day in diameter.

Stars orbiting Sgr A* come within only a few light-hours of Sgr A*. This eliminates theories that Sgr A* is a cluster of stars, of neutron stars, or of stellar-mass black holes. Only a single black hole could contain so much mass in so small a region.

The Chandra X-ray Observatory has imaged Sgr A* and detected over 2000 other X-ray sources in the area.

NASA/CXC/MIT/F.K. Baganoff et al.

3a A black hole with a mass of 3.7 million solar masses would have an event horizon smaller than the smallest dot in this diagram. A slow dribble of only one ten millionth of a solar mass per year could produce the observed energy. An occasional star falling in could produce a violent eruption.

The evidence of a massive black hole at the center of our galaxy seems conclusive. It is much too massive to be the remains of a dead star, so it probably formed as the galaxy first took shape. It may have gained further mass as it absorbed gas, stars, and star clusters.

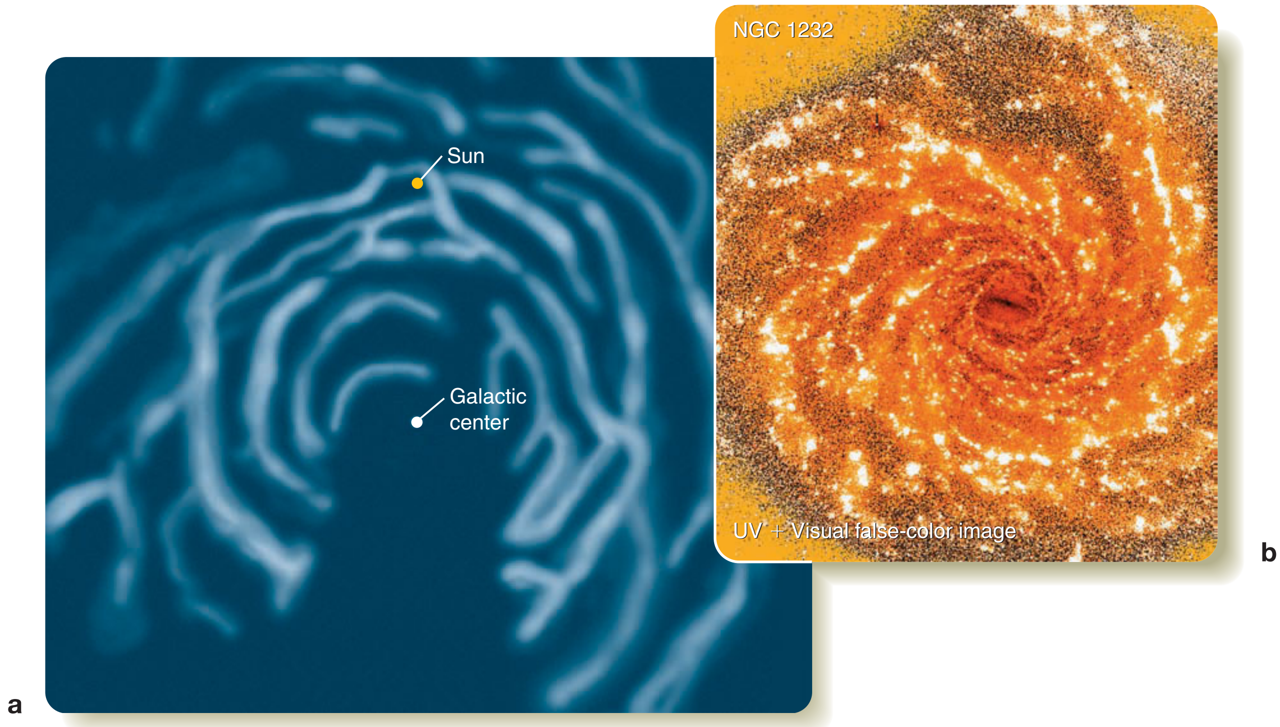


Figure 12-17

(a) This 21-cm radio map of our galaxy confirms that it has spiral arms, but the pattern is complex and suggests branches and spurs. (Adapted from a radio map by Gart Westerhout) (b) Many spiral galaxies have complex spiral patterns. In this image the brightest regions are the most active regions of star formation, and they outline spiral arms, branches, and spurs. (ESO)

that our galaxy has spiral arms, although it is hard to make out the overall pattern.

From radio and visual observations, astronomers conclude that the spiral arms are rather irregular and are interrupted by branches, spurs, and gaps. The stars in Orion, for example, appear to be a detached segment of a spiral arm, a spur. There are significant sources of error in the radio mapping method, but many of the irregularities along the arms seem real, and images of nearby spiral galaxies show similar features (Figure 12-17b). Studies comparing all the available data on our galaxy's spiral pattern with patterns seen in other galaxies do not necessarily agree with each other, although the newest models suggest that the nuclear bulge in our galaxy is elongated into a bar (■ Figure 12-18) and that the spiral arms spring from the ends of the bar. You will see in the next chapter that such bars are common in spiral galaxies.

The most important feature in the radio maps is easy to overlook—spiral arms are regions of higher gas density. You have seen earlier that the arms contain young objects suggesting active star formation. Radio maps confirm this suspicion by showing that the material needed to make stars is abundant in spiral arms.

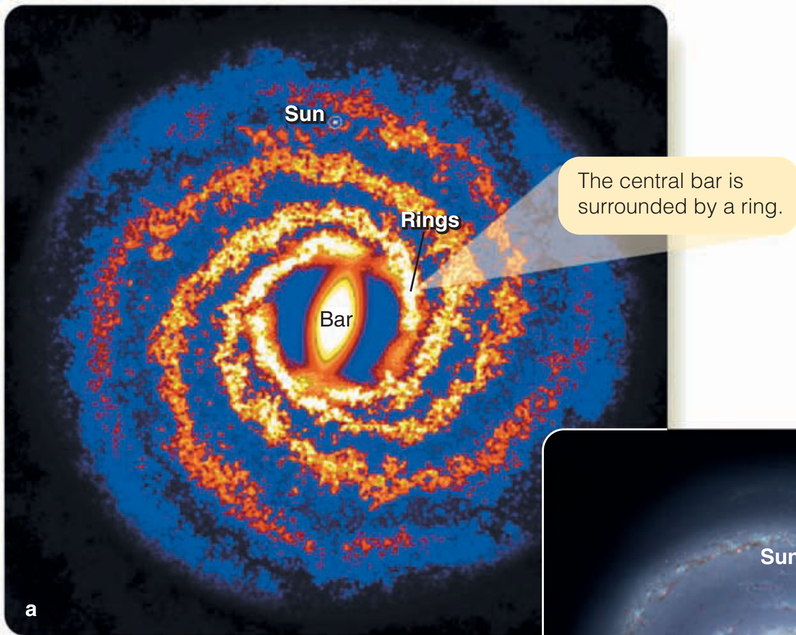
The Density Wave Theory

Having mapped the spiral pattern, you can ask, "What are spiral arms?" You can be sure that they are not physically connected structures, such as bands of a magnetic field holding the gas in place. Like a kite string caught on a spinning hubcap, such arms would be wound up and pulled apart by differential rotation within a few tens of millions of years. Yet spiral arms are common in disk-shaped galaxies and must be reasonably permanent features.

The most prominent theory since the 1950s is called the **density wave theory,** which proposes that spiral arms are waves of compression, rather like sound waves, that move around the galaxy, triggering star formation. Because these waves move slowly, orbiting gas clouds overtake the spiral arms from behind and create a moving traffic jam within the arms.

A density wave is a bit like a traffic jam behind a truck moving slowly along a highway. Seen from an airplane overhead, the jam seems a permanent, though slow-moving, feature. But individual cars overtake the jam from behind, slow down, move up through the jam, wait their turn, pass the truck, and resume speed along the highway. So too do clouds of gas overtake the spiral density wave, become compressed in the "traffic jam," and eventually move out in front of the arm, leaving the slower-moving density wave behind.

Mathematical models of this process have been very successful at generating spiral patterns that look like those seen in the Milky Way Galaxy and other spiral galaxies. In each case, the density wave takes on a regular two-armed spiral pattern that winds outward from the nuclear bulge to the edge of the disk.



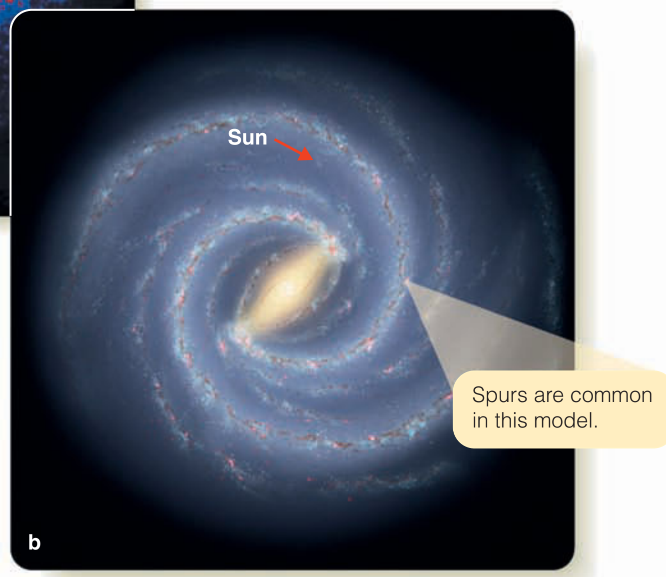


Figure 12-18

(a) This model of our galaxy is based on far-infrared observations of dust and includes four spiral arms with no branches. (Courtesy Henry Freudenreich) (b) This artist's impression of a two-armed model is based on observations with the Spitzer Infrared Space Telescope. Notice the larger central bar. (NASA/JPL-Caltech)

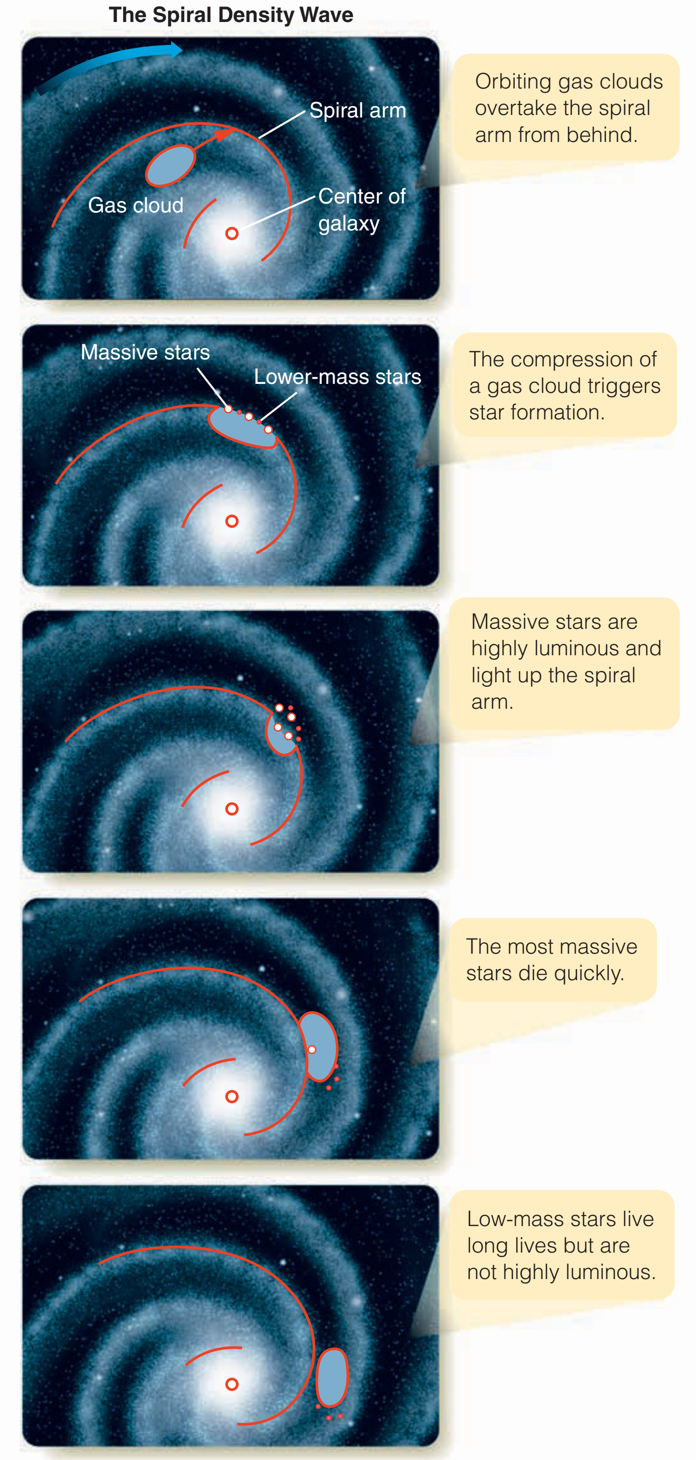
Such spiral arms will be stable for a billion years or so. The differential rotation does not wind them up because they are not physically connected structures.

As you would expect, star formation occurs where the gas clouds are compressed. Stars pass through the spiral arms unaffected, like bullets passing through a wisp of fog, but large clouds of gas slam into the spiral density wave from behind and are suddenly compressed (■ Figure 12-19). You saw in an earlier chapter that sudden compression of a cloud could trigger the formation of stars. Thus, new stars form along the spiral arms.

The brightest stars, the O and B stars, live such short lives that they never travel far from their birthplace and are found only along the arms. Their presence is what makes the spiral arms glow so brightly (■ Figure 12-20). Lower-mass stars, like the sun, live longer and have time to move out of the arms and continue their journey around the galaxy. The sun may have formed in a star cluster about 5 billion years ago when a gas cloud smashed into a spiral arm. Since that time, the sun has escaped from its cluster and made about 20 trips around the galaxy, passing through spiral arms many times.

The density wave theory is very successful in explaining the properties of spiral galaxies, but it has two problems. First, what stimulates the formation of the spiral pattern? Some process must generate the spiral arms in the first place, but the pattern must also be restimulated, or it would die away over a few billion years. Mathematical models show that the galaxy is naturally unstable to certain disturbances, just as a guitar string is unstable to certain vibrations. A sudden disturbance—the rumble of a passing truck, for example—can set the string vibrating at its natural frequencies. Similarly, minor fluctuations in the galaxy's disk or gravitational interactions with passing galaxies might generate a density wave.

The second problem with the density wave theory is that it does not account for the branches and spurs observed in the spiral arms of some galaxies. Computer models of density waves produce regular, two-armed spiral patterns. Some galaxies, called grand-design galaxies, do indeed have symmetric two-armed patterns, but others do not. Some galaxies have a great many short spiral segments, giving them a fluffy appearance. These galaxies have been termed **flocculent,** meaning "woolly." Our galaxy is



■ **Figure 12-19**

According to the density wave theory, star formation occurs as gas clouds pass through spiral arms.

probably intermediate between these two types. How can astronomers explain these variations? Perhaps the solution lies in a process that sustains star formation once it begins.

Star Formation in Spiral Arms

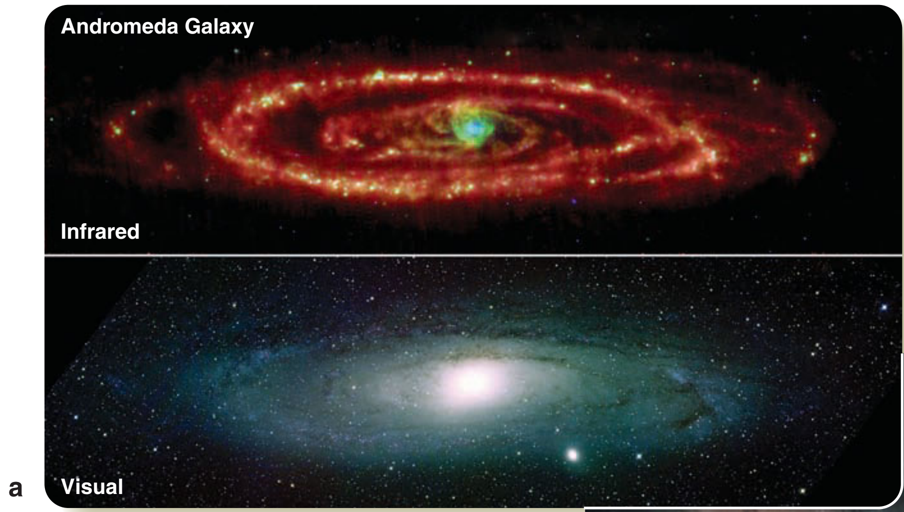
Star formation is a critical process in the creation of the spiral patterns. It not only makes spiral arms visible, but it may also shape the spiral pattern itself.

The spiral density wave creates spiral arms by the gravitational attraction of the stars and gas flowing through the arms. Even if there were no star formation at all, rotating disk galaxies could form spiral arms. But without star formation to make young, hot, luminous stars, the spiral arms would be difficult to see. It is the star formation that lights up the spiral arms and makes them so prominent. Thus, star formation helps determine what you see when you look at a spiral galaxy.

But star formation can also control the shape of spiral patterns if the birth of stars in a cloud of gas can trigger the formation of more new stars. Consider a newly formed star cluster with a single massive star. The intense radiation from that hot star can compress nearby parts of the gas cloud and trigger further star formation (■ Figure 12-21). Also, massive stars evolve so quickly that their lifetimes last only an instant in the history of a galaxy. When they explode as supernovae, the expanding gases can compress neighboring clouds of gas and trigger more star formation. Examples of such **self-sustaining star formation** have been found. You saw in the chapter on star formation that the Orion complex, consisting of the Great Nebula in Orion and the star formation buried deep in the dark interstellar clouds behind the nebula, is such a region.

Self-sustaining star formation can produce growing clumps of new stars, and the differential rotation of the galaxy can drag the inner edge of the clump ahead and let the outer edge lag behind to produce a cloud of star formation shaped like a segment of a spiral arm, a spur. Mathematical models of galaxies filled with such segments of spiral arms do have a spiral appearance, but they lack the bold, two-armed spiral that astronomers refer to as the grand-design spiral pattern (■ Figure 12-22). Astronomers suspect that self-sustaining star formation can produce the branches and spurs so prominent in flocculent galaxies, but only the spiral density wave can generate the beautiful two-armed spiral patterns.

This discussion of star formation in spiral arms illustrates the importance of natural processes. The spiral density wave creates the graceful arms, but it is star formation in the arms that makes them stand out so prominently. Self-sustaining star formation can act in some galaxies to modify the spiral arms and produce branches and spurs. By searching out and understanding the details of such natural processes, you can begin to understand the overall structure and evolution of the universe around us.



■ **Figure 12-20**

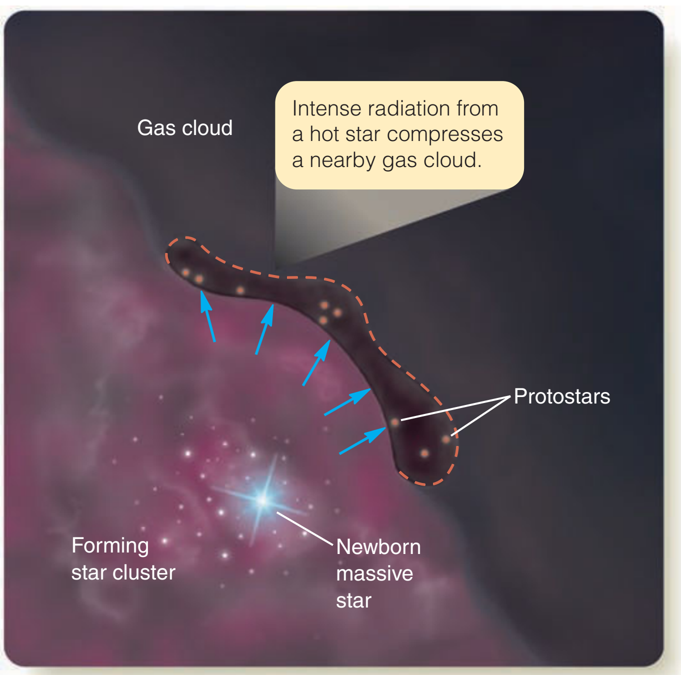
Besides bright stars, spiral arms also contain clouds of gas and dust. (a) In the Andromeda Galaxy, dust clouds glow brightly in the infrared and outline the spiral arms. (NASA/JPL-Caltech/K. Gordon, University of Arizona and NOAO/AURA/NSF) (b) This chance alignment of a small galaxy in front of a larger, more distant galaxy silhouettes clouds of dust and gas against the glare of the distant galaxy. (NASA and The Hubble Heritage Team)

SCIENTIFIC ARGUMENT

Why can't astronomers use solar-type stars as spiral tracers?

In some cases, the timing of events is the critical factor in a scientific argument. In this case, you need to think about the evolution of stars and their orbital periods around the galaxy. Stars like the sun live about 10 billion years, but the sun's orbital period around the galaxy is over 200 million years. The sun almost certainly formed when a gas cloud passed through a spiral arm, but since then the sun has circled the galaxy many times and has passed through spiral arms often. That means the sun's present location has nothing to do with the location of any spiral arms. An O star, however, lives only a few million years. It is born in a spiral arm and lives out its entire lifetime before it can leave the spiral arm. Short-lived stars such as O stars make good spiral tracers, but G stars are too long lived.

The spiral arms of our galaxy must make it beautiful in photographs taken from a distance, but we are trapped inside it. Create an argument based on evidence: **How do astronomers know that the spiral arms mapped out near us by spiral tracers actually extend across the disk of our galaxy?**



■ **Figure 12-21**

Most stars in a star cluster are too small and too cool to affect nearby gas clouds; but, once a massive star forms, it becomes so hot and so luminous its radiation can push gas away and compress a nearby gas cloud. In the densest regions of the cloud, new stars can begin forming.

Figure 12-22

(a) Some galaxies are dominated by two spiral arms; but, even in these galaxies, minor spurs and branches are common. The spiral density wave can generate the two-armed, grand-design pattern, but self-sustained star formation may be responsible for the irregularities. (b) Many spiral galaxies do not appear to have two dominant spiral arms. Spurs and branches suggest that star formation is proceeding rapidly in such galaxies. (a and b, Anglo-Australian Observatory/David Malin Images)

What Are We? Children of the Milky Way

Hang on tight. The sun, with Earth in its clutch, is ripping along at high velocity as it orbits the center of the Milky Way Galaxy. We live on a wildly moving ball of rock in a large galaxy that some call our home galaxy, but the Milky Way is more than just our home. Perhaps "parent galaxy" would be a better name.

Except for hydrogen atoms, which have survived unchanged since the universe began, you and Earth are made of metals—atoms heavier than helium. There is no helium in your body, but there is plenty of carbon, nitrogen, and oxygen. There is calcium in your bones and iron in your blood. All of those atoms and more were cooked up inside stars or in their supernova deaths over the history of our galaxy.

Stars are born when clouds of gas orbiting the center of the galaxy slam into the gas in spiral arms and are compressed. That process has given birth to generations of stars, and each generation has produced elements heavier than helium and spread them back into the interstellar medium. The abundance of metals has grown slowly in the galaxy. About 4.6 billion years ago a cloud of gas enriched in those heavy atoms slammed into a spiral arm and produced the sun, the Earth, and you. You have been cooked up by the Milky Way Galaxy—your parent galaxy.

Study and Review

Summary

- ► The hazy band of the Milky Way is our wheel-shaped galaxy seen from within, but its size and shape are not obvious. William and Caroline Herschel counted stars at many locations over the sky to show that the star system seemed to be shaped like a grindstone with the sun near the center, but they could not see very far into space because of gas and dust.
- ► In the early 20th century, Harlow Shapley **calibrated (p. 237)** certain **variable stars (p. 234)** called **Cepheids.** These stars fall within an **instability strip (p. 236)** in the H–R diagram with **RR Lyrae varible stars (p. 236)** at the bottom of the strip. Shapley was able to calibrate the **period–luminosity relation (p. 237)** by using the **proper motions (p. 237)** of a few Cepheid variable stars. That allowed him to find the distance to globular clusters and demonstrate that our galaxy is much larger than what humans can see and that the sun is not at the center.
- ► The **Shapley–Curtis debate (p. 240)** concerned the nature of certain faint nebulae seen in the sky. Curtis argued that they were other galaxies, and Shapley argued that they were nebulae within our galaxy. Better observations soon showed that they were other galaxies like our Milky Way galaxy.
- ► Modern observations reveal that our galaxy contains a **disk component (p. 240)** about 25 **kiloparsecs (p. 240)** in diameter and that the sun is two-thirds of the way from the center to the visible edge. The **spiral arms (p. 241)** are located in the disk. The **spherical component (p. 241)** includes the **nuclear bulge (p. 241)** around the center and an extensive **halo (p. 241)** containing old stars and little gas and dust.
- ► The mass of the galaxy can be found from its **rotation curve (p. 242).** Kepler's third law tells astronomers the galaxy contains over 100 billion solar masses, and the rising rotation curve at great distance from the center shows that the halo must contain much more mass in a **dark halo** or **galactic corona (p. 242).** Because that mass is not emitting detectable electromagnetic radiation, astronomers call it **dark matter (p. 243).**
- ► The oldest open star clusters indicate that the disk of our galaxy is only about 9 billion years old. The oldest globular clusters appear to be at least 13 billion years old, so our galaxy must have formed at least 13 billion years ago.
- ► Stellar populations are an important clue to the formation of our galaxy. The first stars to form in our galaxy, termed **population II stars (p. 244),** were poor in elements heavier than helium — elements that astronomers call **metals (p. 244).** As generations of stars manufactured metals and spread them back into the interstellar medium, the metal abundance of more recent generations increased. **Population I stars (p. 244)**, including the sun, are richer in metals.
- ► Because the halo is made up of population II stars and the disk is made up of population I stars, astronomers conclude that the halo formed first and the disk later. A hypothesis that the galaxy formed from a single, roughly spherical cloud of turbulent gas and gradually flattened into a disk has been amended to include mergers with other galaxies and infalling gas contributing to the disk.
- ► The nucleus of the galaxy is not visible at visual wavelengths, but radio, infrared, and X-ray radiation can penetrate the gas and dust in space. These wavelengths reveal crowded central stars and warmed dust.
- ► The very center of the Milky Way Galaxy is marked by a radio source, **Sagittarius A* (p. 248).** The object must be no more than an astronomical unit in diameter, but the motion of stars around the center shows that it must contain roughly 3.7 million solar masses. A supermassive black hole is the only object that could contain so much mass in such a small space.
- ► The most massive stars live such short lives that they don't have time to move from their place of birth. Because they are scattered along the spiral arms, astronomers conclude that the spiral arms are sites of star formation.
- ► The spiral arms can be traced through the sun's neighborhood by using **spiral tracers (p. 248)** such as O and B stars; but, to extend the map over the entire galaxy, astronomers must use radio telescopes to see through the gas and dust.
- ► The **density wave theory (p. 252)** suggests that the spiral arms are regions of compression that move around the disk. When an orbiting gas cloud overtakes the compression wave, the gas cloud is compressed and forms stars. Another process, **self-sustaining star formation (p. 254),** may act to modify the arms with branches and spurs as the birth of massive stars triggers the formation of more stars by compressing neighboring gas clouds. This may explain the appearance of galaxies that are described as **flocculent (p. 253).**

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why isn't it possible to locate the center of our galaxy from the appearance of the Milky Way?
2. Why is there a period–luminosity relation?
3. How can astronomers use variable stars to find distance?
4. Why is it difficult to specify the thickness or diameter of the disk of our galaxy?
5. Why didn't astronomers before Shapley realize how large the galaxy is?
6. How do astronomers know how old our galaxy is?
7. Why do astronomers conclude that metal-poor stars are older than metal-rich stars?
8. How can astronomers find the mass of the galaxy?
9. What evidence is there that our galaxy has an extended corona of dark matter?
10. How do the orbits of stars around the Milky Way Galaxy help explain its origin?
11. What evidence contradicts the traditional theory for the origin of our galaxy?
12. What is the evidence that the center of our galaxy contains a massive black hole?
13. Why must astronomers use infrared telescopes to observe the motions of stars around Sgr A*?
14. Why do spiral tracers have to be short lived?
15. What evidence is there that the density wave theory is not fully adequate to explain spiral arms in our galaxy?
16. **How Do We Know?** Calibration simplifies complex measurements, but how does that make the work of later astronomers dependent on the expertise of the astronomer who did the calibration?
17. **How Do We Know?** The story of a process makes the facts easier to remember, but that is not the true goal of the scientist. What is the real value of understanding a scientific process?

Discussion Questions

1. How would this chapter be different if interstellar matter didn't absorb starlight?
2. Are there any observations you could make with the Hubble Space Telescope that would allow you to better understand Sgr A*?

Problems

1. Make a scale sketch of our galaxy in cross section. Include the disk, sun, nucleus, halo, and some globular clusters. Try to draw the globular clusters to scale size.
2. Because of dust, you can see only about 5 kpc into the disk of the galaxy. What percentage of the galactic disk can you see? (*Hint:* Consider the area of the entire disk and the area you can see.)
3. If the fastest passenger aircraft can fly 1600 km/hr (1000 mph), how many years would it take to reach the sun? The galactic center? (*Hint:* 1 pc $= 3.1 \times 10^{13}$ km.)
4. If the RR Lyrae stars in a globular cluster have apparent magnitudes of 14, how far away is the cluster? (*Hint:* See Reasoning with Numbers 8-2.)
5. If interstellar dust makes an RR Lyrae star look 1 magnitude fainter than it should, by how much will astronomers overestimate its distance? (*Hint:* See Reasoning with Numbers 8-2.)
6. If a globular cluster is 10 arc minutes in diameter and 8.5 kpc away, what is its diameter? (*Hint:* See Reasoning with Numbers 3-1.)
7. If you assume that a globular cluster 4 minutes of arc in diameter is actually 25 pc in diameter, how far away is it? (*Hint:* See Reasoning with Numbers 3-1.)
8. If the sun is 5 billion years old, how many times has it orbited the galaxy?
9. If the true distance to the center of our galaxy is found to be 7 kpc and the orbital velocity of the sun is 220 km/s, what is the minimum mass of the galaxy? (*Hints:* Find the orbital period of the sun and then see Reasoning with Numbers 8-4.)
10. Infrared radiation from the center of our galaxy with a wavelength of about 2×10^{-6} m (2000 nm) comes mainly from cool stars. Use this wavelength as λ_{max} and find the temperature of the stars. (*Hint:* See Reasoning with Numbers 6-1.)

Learning to Look

1. Why does the galaxy shown at the right have so much dust in its disk? How big do you suppose the halo of this galaxy really is?

NASA/Hubble Heritage Team/ STScI/AURA

2. Why are the spiral arms in the galaxy at the right blue? What color would the halo be if it were bright enough to see in this photo?

NASA/Hubble Heritage Team and A. Riess, STScI

13 Galaxies

Spiral galaxy NGC 1672 is 60 million light-years from Earth. Light from the galaxy departed on its journey to Earth soon after the extinction of the dinosaurs. (NASA, ESA, Hubble Heritage Team, STScI/AURA-ESA Hubble Collaboration)

Guidepost

Our Milky Way Galaxy is large and complex, but it is only one among billions. This chapter will expand your horizon to discuss the different kinds of galaxies and their complex histories. Here you can expect answers to four essential questions:

- **What do galaxies look like?**
- **How do astronomers find the distance, size, luminosity, and mass of galaxies?**
- **Do other galaxies contain supermassive black holes and dark matter as does our own galaxy?**
- **Why are there different kinds of galaxies?**

These are perfectly reasonable questions, but the answers will lead you to an astonishing insight into how galaxies are born and how they evolve.

This chapter's discussion of normal galaxies will carry you through the next two chapters, where you will study violently erupting galaxies powered by supermassive black holes and the evolution of the universe as a whole.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

The ability to theorize is highly personal; it involves art, imagination, logic, and something more.

EDWIN HUBBLE
THE REALM OF THE NEBULAE

Science fiction heroes flit effortlessly between the stars, but almost none voyages between the galaxies. As you leave the Milky Way Galaxy behind, you voyage out into the vast depths of the universe, out among the galaxies, into space so deep it is unexplored even in fiction.

Less than a century ago, astronomers did not understand that there were galaxies. Nineteenth-century telescopes revealed faint nebulae scattered among the stars, and some of the nebulae were spiral. Astronomers argued about the nature of these nebulae, but it was not until the 1920s that they understood that some were other galaxies much like our own, and it was not until recent decades that astronomical telescopes could reveal the tremendous beauty and intricacy of the galaxies (■ Figure 13-1).

Before you can begin to theorize, you must gather some basic data. What are galaxies like? What are their diameters, luminosities, and masses? Once you know what galaxies are like, you will be ready to wonder about their origin and evolution.

13-1 The Family of Galaxies

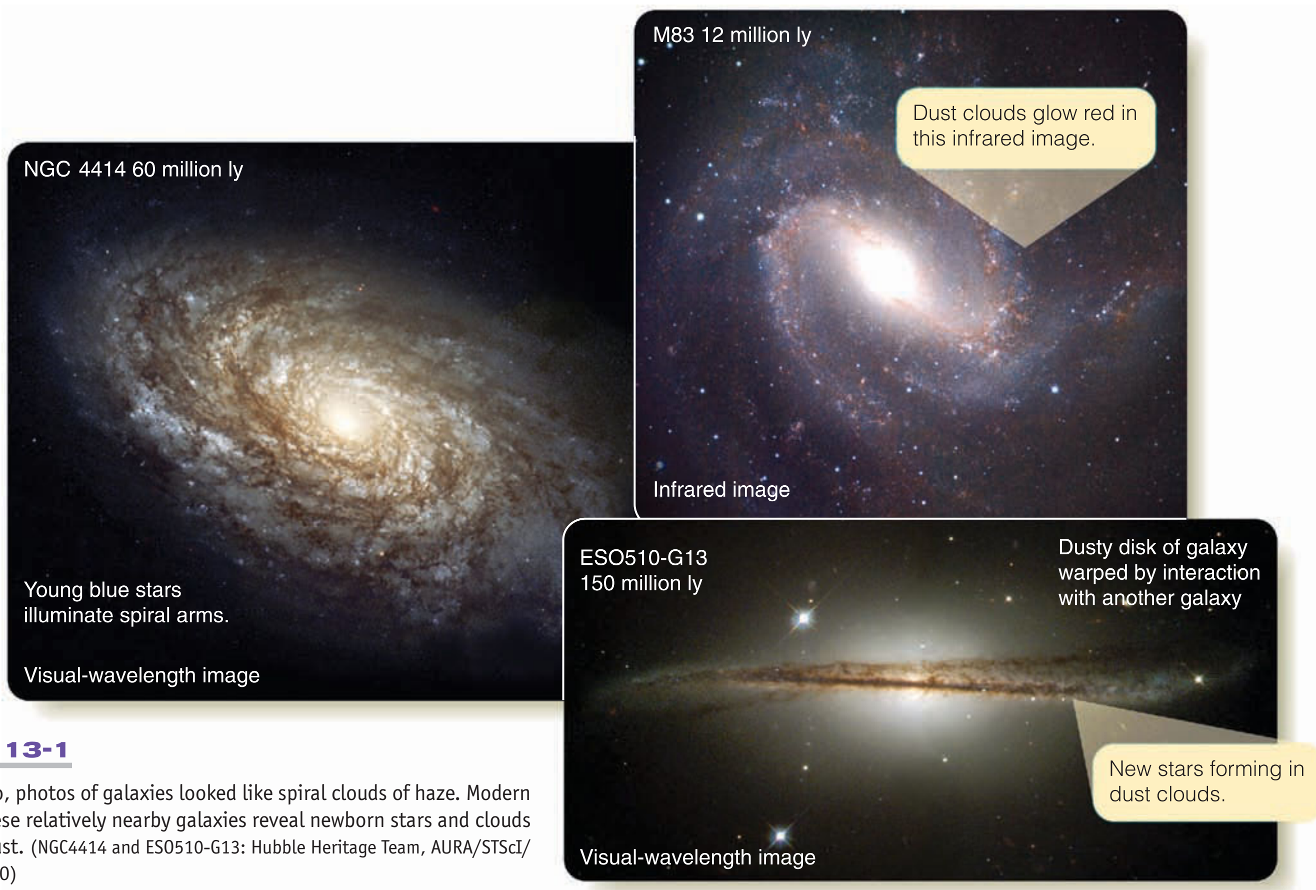
Astronomy books often include pictures of spiral galaxies. Like movie stars, spiral galaxies get a lot of attention because they are beautiful; but many galaxies are nearly featureless clouds of stars, and others are distorted muddles of gas and dust. You can begin by sorting out this jumble of galaxies.

The Shapes of Galaxies

Astronomers classify galaxies according to their shape using a system developed in the 1920s by Edwin Hubble (after whom the Hubble Space Telescope is named). Creating a system of classification is a fundamental technique in science (■ How Do We Know? 13-1).

Read ■ **Galaxy Classification** on pages 262–263 and notice three important points and four new terms:

1. Many galaxies have no disk, no spiral arms, and almost no gas and dust. These *elliptical galaxies* range from huge giants to small dwarfs.
2. Disk-shaped galaxies usually have spiral arms and contain gas and dust. Many *spiral galaxies* have a nucleus shaped like an elongated bar and are called *barred spiral galaxies*. A few disk galaxies contain little gas and dust.
3. *Irregular galaxies* are highly irregular in shape and tend to be rich in gas and dust.



■ **Figure 13-1**

A century ago, photos of galaxies looked like spiral clouds of haze. Modern images of these relatively nearby galaxies reveal newborn stars and clouds of gas and dust. (NGC4414 and ESO510-G13: Hubble Heritage Team, AURA/STScI/NASA; M83: ESO)

How Do We Know? 13-1

Classification in Science

What does classification tell a scientist? Classification is one of the most basic and most powerful of scientific methods. Establishing a system of classification is often the first step in studying a new aspect of nature, and it can produce unexpected insights.

Between 1831 and 1836, Charles Darwin sailed around the world with a scientific expedition aboard the ship HMS *Beagle*. Everywhere he went, he studied the living things he saw and tried to classify them. For example, he classified different types of finches he saw on the Galapagos Islands according to the shapes of their beaks. He found that those that fed on seeds with hard shells had thick, powerful beaks, whereas those that picked insects out of deep crevices had long, thin beaks. His classifications of these and other animals led him to think about how natural selection shapes creatures to survive in their environment, which helped him understand how living things evolve.

Years after Darwin's work, paleontologists classified dinosaurs into two orders, lizard-hipped and bird-hipped. This classification, based on the shapes of the dinosaurs' hip joints, helped the scientists understand patterns of dinosaur evolution. It also led them to the conclusion that modern birds, including the finches that Darwin saw on the Galapagos, evolved from dinosaurs.

Astronomers use classifications of galaxies, stars, moons, and many other objects to help them see patterns, trace evolution, and generally make sense of the astronomical world. Whenever you encounter a scientific discussion, look for the classifications on which it is based. Classifications are the orderly framework on which much of science is built.

The careful classification of living things can reveal surprising relationships. The birds, including this flamingo, are descended from dinosaurs. (M. Seeds)

You might also wonder what proportions of all galaxies are elliptical, spiral, and irregular, but that is a difficult question to answer. In some catalogs of galaxies, about 70 percent are spiral, but that is the result of a selection effect (■ How Do We Know? 13-2). Spiral galaxies contain hot, bright stars and are consequently very luminous and easy to see. Most ellipticals are fainter and harder to notice. Small galaxies, such as dwarf ellipticals and dwarf irregulars, may be very common, but they are hard to detect. From careful studies, astronomers can conclude that ellipticals are more common than spirals and that irregulars make up only about 25 percent of all galaxies. Among spiral galaxies, about two-thirds are barred spirals.

How Many Galaxies Are There?

Like leaves on the forest floor, galaxies carpet the sky. Look at any spot on the sky away from the dust and gas of the Milky Way, and you are looking deep into space. Photons that have traveled many billions of years enter your eye, but they are too few to register on your retina. Only the largest telescopes can gather enough light to detect the most distant galaxies.

In 1995, astronomers picked a seemingly empty spot in the sky near the Big Dipper and used the Hubble Space Telescope to record a time exposure that lasted an astonishing 10 days. This became known as a Hubble Deep Field; it was deep in that it recorded very faint objects. The image revealed that the "empty spot" in the sky was filled with galaxies.

More recent deep fields have recorded even fainter objects. A research effort called GOODS (Great Observatories Origins Deep Survey) has used the Hubble Space Telescope, the Chandra X-ray Observatory, the (infrared) Spitzer Space Telescope, the XXM-Newton X-ray Telescope, and the largest ground-based telescopes, to make multiwavelength studies of two selected areas in the northern and southern sky.

The GOODS deep images reveal tremendous numbers of galaxies (■ Figure 13-2). There is no reason to believe that the regions of the sky chosen for deep fields are unusual, so it seems likely that the entire sky is carpeted with galaxies. At least 100 billion would be visible with today's telescopes, and surely there are other galaxies too distant or too faint to see.

◄ SCIENTIFIC ARGUMENT ►

What color are galaxies?

Scientific arguments must be based on evidence, and evidence means observations. But you must be careful to analyze even the simplest observations with care. Different kinds of galaxies have different colors, depending mostly on how much gas and dust they contain. If a galaxy contains large amounts of gas and dust, it probably contains lots of young stars, and a few of those young stars will be massive, hot, luminous O and B stars. They will produce most of the light and give the galaxy a distinct blue tint. In contrast, a galaxy that contains little gas and dust will probably contain few young stars. It will lack O and B stars, and the most luminous stars in such a galaxy will be red giants. They will give the galaxy a red tint. Because the light from a galaxy is a blend of the light from billions of stars, the colors are only tints. Nevertheless, the most luminous stars in a galaxy determine its color. From this you can conclude that elliptical galaxies tend to be red and the disks of spiral galaxies tend to be blue.

Now create a new scientific argument and analyze a different kind of observation. **Why are most galaxies in catalogs spiral in spite of the fact that the most common kind of galaxy is elliptical?**

◄ ►

Galaxy Classification

1 **Elliptical galaxies** are round or elliptical, contain no visible gas and dust and have few or no bright stars. They are classified with a numerical index ranging from 1 to 7; E0s are round, and E7s are highly elliptical. The index is calculated from the largest and smallest diameter of the galaxy used in the following formula and rounded to the nearest integer.

$$\frac{10(a-b)}{a}$$

Outline of an E6 galaxy

AURA/NOAO/NSF

M87 is a giant elliptical galaxy classified E1. It is at least 50% larger in diameter than our own galaxy and is surrounded by a swarm of over 500 globular clusters.

Anglo-Australian Observatory/David Malin Images

The Leo 1 dwarf elliptical galaxy is not many times bigger than a globular cluster.

2 **Spiral galaxies** contain a disk and spiral arms. Their halo stars are not visible, but presumably all spiral galaxies have halos. Spirals contain gas and dust and hot, bright O and B stars, as shown at right. The presence of short-lived O and B stars alerts us that star formation is occurring in these galaxies. Sa galaxies have larger nuclei, less gas and dust, and fewer hot, bright stars. Sc galaxies have small nuclei, lots of gas and dust, and many hot, bright stars. Sb galaxies are intermediate.

Anglo-Australian Observatory/David Malin Images

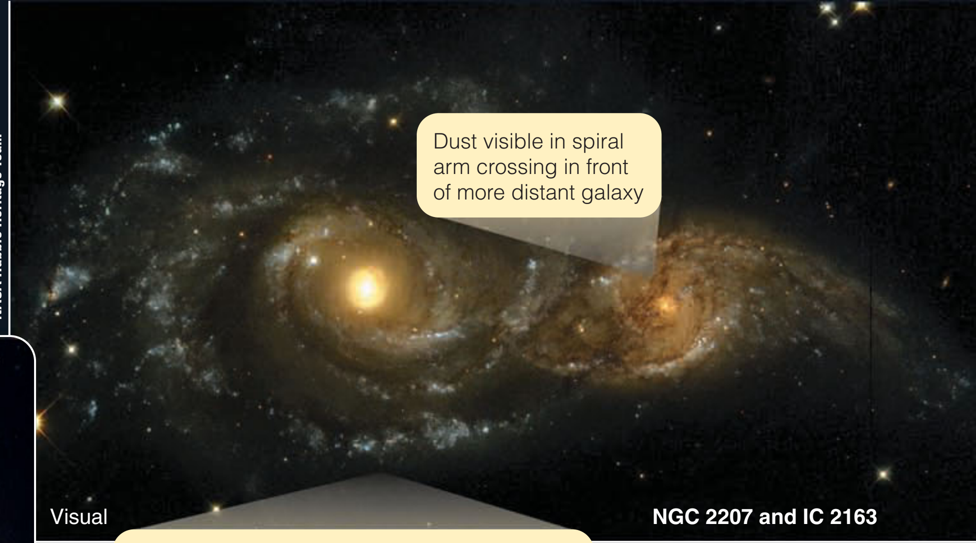
Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Galaxy Types" and review the classification of galaxies.

Anglo-Australian Observatory/David Malin Images

2a Roughly 2/3 of all spiral galaxies are **barred spiral galaxies** classified SBa, SBb, and SBc. They have an elongated nucleus with spiral arms springing from the ends of the bar, as shown at left. Our own galaxy is a barred spiral.

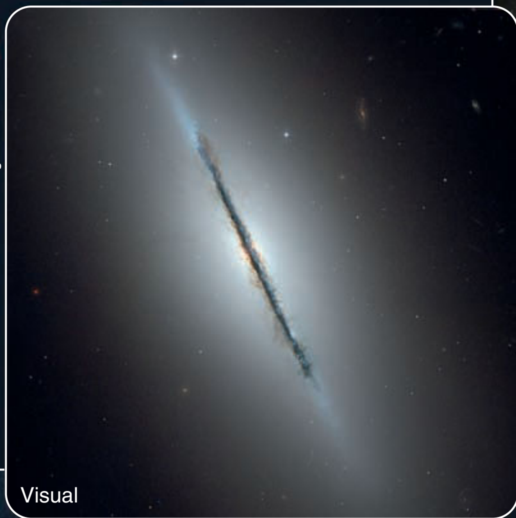
2b Dust in spiral galaxies is concentrated in the disk along the spiral arms. Even in a galaxy that has little dust such as NGC 5866 shown below the dust is prominent when the galaxy is seen edge on.

NASA Hubble Heritage Team

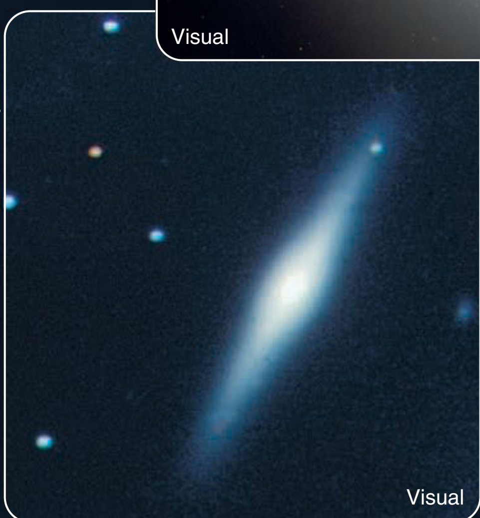


Dust in spiral galaxies is most common in the spiral arms. Here the spiral arms of one galaxy are silhouetted in front of a more distant galaxy.

NASA, ESA, and the Hubble Heritage Team, STScI/AURA



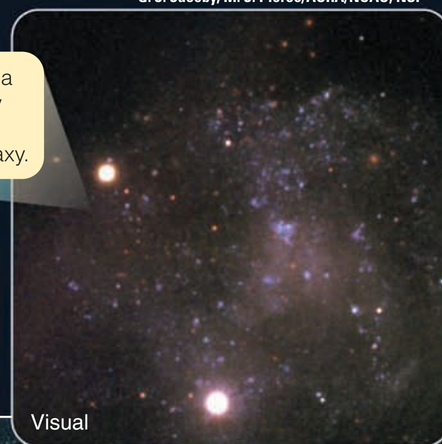
Rudolph E. Schild



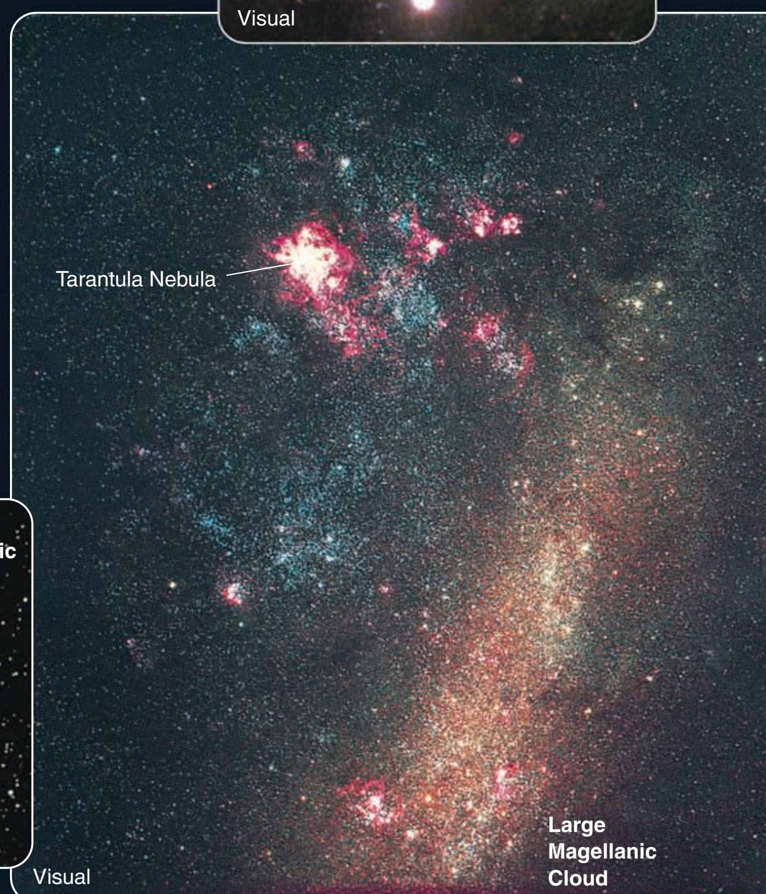
2c Galaxies with an obvious disk and nuclear bulge but no visible gas and dust and few or no hot bright stars are classified as S0 (pronounced “Ess Zero”). Compare this galaxy with the edge-on spiral above.

G. J. Jacoby, M. J. Pierce/AURA/NOAO, NSF

The galaxy IC4182 is a dwarf irregular galaxy only about 4 million parsecs from our galaxy.



3 **Irregular galaxies** (classified Irr) are a chaotic mix of gas, dust, and stars with no obvious nuclear bulge or spiral arms. The Large and Small Magellanic Clouds are visible to the unaided eye as hazy patches in the southern hemisphere sky. Telescopic images show that they are irregular galaxies that are interacting gravitationally with our own much larger galaxy. Star formation is dramatic in the Magellanic Clouds. The bright pink regions are emission nebulae excited by newborn O and B stars. The brightest nebula in the Large Magellanic Cloud is called the Tarantula Nebula.



Copyright R. J. Dufour, Rice University

AURA/NOAO/NSF

How Do We Know? 13-2

Selection Effects

How can selecting what to study be misleading? Scientists must design their research projects with great care. Biologists studying insects in the rain forest, for example, must choose which ones to catch. Because they can't catch every insect they see, they might decide to catch and study any insect that is red. But if they are not careful as they design their research plan, a selection effect could bias their data and lead them to incorrect conclusions without their ever knowing it.

To see how this could happen, suppose you needed to measure the speed of cars on a highway. There are too many cars to measure every one, so you might decide to measure only red cars. It is quite possible that this selection criterion will mislead you because people who buy red cars may be more likely to be younger and drive faster. Should you measure only brown cars? No, because older, more sedate people might tend to buy brown cars. Only by very carefully designing your experiment can you be certain that the cars you measure are traveling at representative speeds.

Astronomers understand that what you see through a telescope depends on what you notice, and that is powerfully influenced by selection effects. The biologists in the rain forest, for example, should not catch and study only red insects. Often, the most brightly colored insects are poisonous or at least taste bad to predators. Catching only red insects could lead the scientists to false conclusions about the kinds of insects that live in the forest.

Things that are bright and beautiful, such as spiral galaxies, may attract a disproportionate amount of attention. Scientists must be aware of such selection effects. (Hubble Heritage Team/STScI/AURA/NASA)

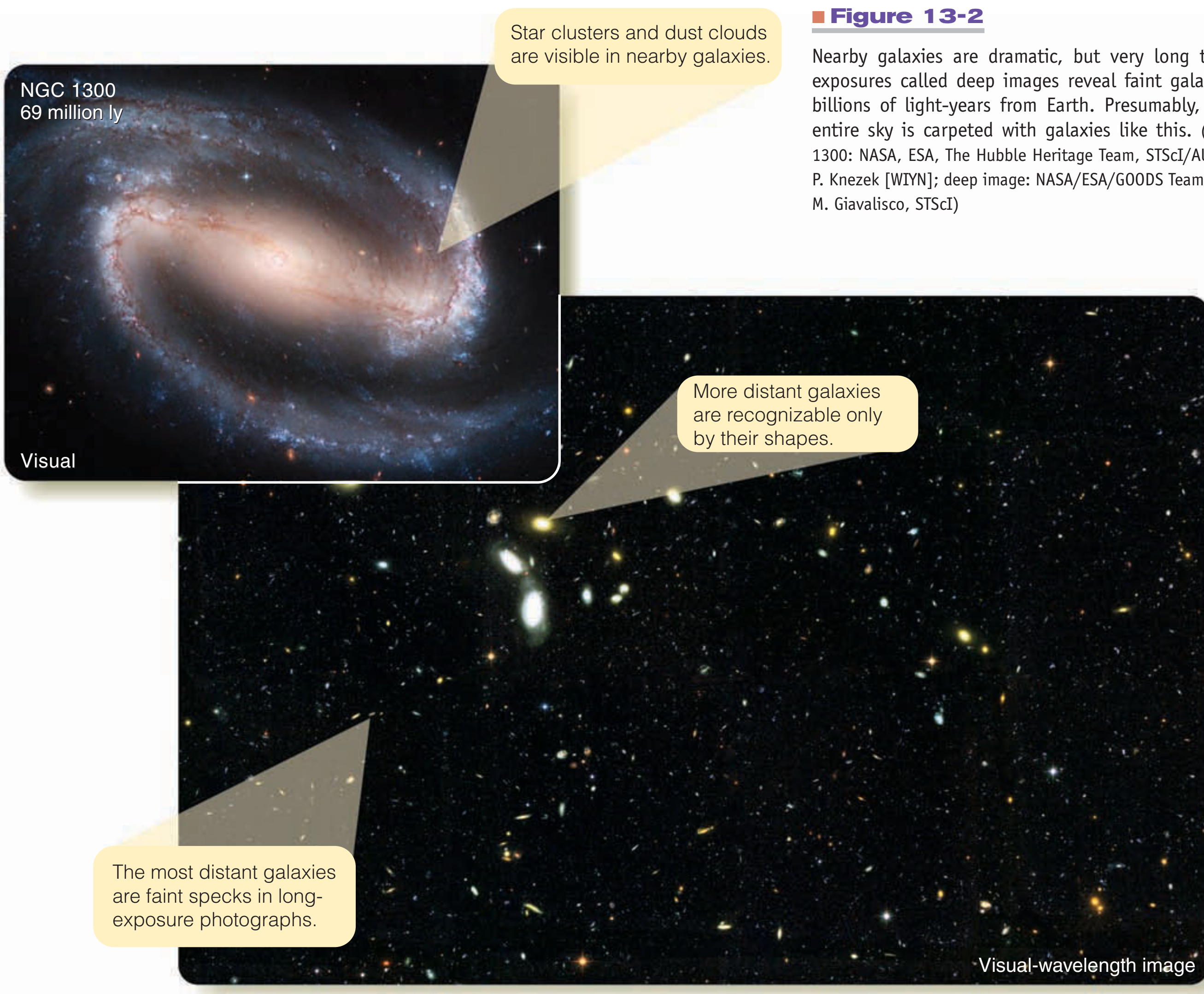


Figure 13-2

Nearby galaxies are dramatic, but very long time exposures called deep images reveal faint galaxies billions of light-years from Earth. Presumably, the entire sky is carpeted with galaxies like this. (NGC 1300: NASA, ESA, The Hubble Heritage Team, STScI/AURA, P. Knezek [WIYN]; deep image: NASA/ESA/GOODS Team and M. Giavalisco, STScI)

13-2 Measuring the Properties of Galaxies

Looking beyond the edge of our Milky Way Galaxy, astronomers find many billions of galaxies. Great clusters, some containing thousands of galaxies, fill space as far as telescopes can see. What are the properties of these star systems? What are the diameters, luminosities, and masses of galaxies? Just as in your study of stellar characteristics, the first step in determining the properties of galaxies is to find out how far away they are. Once you know a galaxy's distance, its size and luminosity are relatively easy to find. Later in this section, you will see that finding mass is more difficult.

Distance

The distances to galaxies are so large that it is not convenient to express them in light-years, parsecs, or even kiloparsecs. Instead, astronomers use the unit **megaparsec (Mpc),** or 1 million pc. One Mpc equals 3.26 million ly, or approximately 3×10^{19} km.

To find the distance to a galaxy, astronomers must search among its stars, nebulae, and star clusters for familiar objects whose luminosity or diameter they know. Such objects are called **distance indicators** because they can be used to find the distance to a galaxy. Distance indicators whose luminosity is known are often referred to by astronomers as **standard candles.** If you can find a standard candle in a galaxy, you can judge its distance.

Cepheid variable stars are reliable distance indicators because, as you learned in the previous chapter, their period is related to their luminosity. If you know the period of the star's variation, you can use the period–luminosity diagram to learn its absolute magnitude. Then, by comparing its absolute and apparent magnitudes, you can find its distance. ■ Figure 13-3 shows a galaxy in which the Hubble Space Telescope detected Cepheids.

Even the Hubble Space Telescope cannot detect Cepheids in galaxies much beyond 100 million ly (30 Mpc), so astronomers must search for less common but brighter distance indicators and calibrate them using nearby galaxies containing both distance indicators. For example, by studying nearby galaxies with distances known from Cepheids, astronomers have found that the brightest globular clusters have absolute magnitudes of about −10. If they find globular clusters in a more distant galaxy, they can assume that the brightest have similar absolute magnitudes and thus calculate the distance.

Astronomers can use globular clusters as distance indicators in a different way. Studies of nearby globular clusters with known distances show that they are about 25 pc in diameter. If astronomers can detect globular clusters in a distant galaxy, they can assume the clusters are about 25 pc in diameter and use the small-angle formula to find the distance to the galaxy.

When a supernova explodes in a distant galaxy, astronomers rush to observe it. Studies show that type Ia supernovae, those caused by the collapse of a white dwarf, all reach about the same absolute magnitude at maximum. By finding Cepheids and other distance indicators in nearby galaxies where type Ia supernovae have occurred, astronomers have been able to calibrate these supernovae. For example, a type Ia supernova was seen in 2002 in the galaxy shown in ■ Figure 13-4. Astronomers were able to find Cepheid variables in that galaxy and so could find its distance. Then they could calculate the absolute magnitude of the supernova at its brightest. From studies of many such supernovae, astronomers conclude that type Ia supernovae always reach the same maximum brightness—absolute magnitude equal to about –19.5.

Astronomers like to refer to distance indicators like Cepheids as standard candles, but one astronomer commented that type Ia supernovae are more like standard bombs. As you learned in an earlier chapter, white dwarfs accumulating mass explode as they reach the Chandrasekhar limit; consequently they reach the same maximum luminosity and make good distance indicators. When type Ia supernovae are seen in more distant galaxies, astronomers can measure the apparent magnitude at maximum and compare that with the known absolute magnitude of these supernovae to find the distance to the galaxy. Because type Ia supernovae are much brighter than Cepheids, they can be seen at great distances. The drawback is that supernovae are rare, and none may occur during your lifetime in a particular galaxy.

At the greatest distances, astronomers must calibrate the total luminosity of the galaxies themselves. For example, studies of nearby galaxies show that an average galaxy like the Milky Way Galaxy has a luminosity about 16 billion times that of the sun. If astronomers see a similar galaxy far away, they can measure its apparent magnitude and calculate its distance. Of course, it is important to recognize the different types of galaxies, and that is difficult to do at great distances (■ Figure 13-5). Averaging the distances to the brightest galaxies in a cluster can reduce the uncertainty in this method.

Notice how astronomers use calibration (How Do We Know? 12-1) to build a **distance scale** reaching from the nearest galaxies to the most distant visible galaxies. Often astronomers refer to this as the distance ladder because each step—parallax, Cepheids, globular clusters, supernovae, and so on—depends on the steps below it. Of course, the foundation of the distance scale rests on parallax measurements and astronomers' understanding of the luminosities of the stars in the H–R diagram.

Looking Back in Time

The most distant visible galaxies are a little over 10 billion ly (3000 Mpc) away, and at such distances you see an effect akin to time travel. When you look at a galaxy that is a few million light-years away, you do not see it as it is now but rather as it was mil-

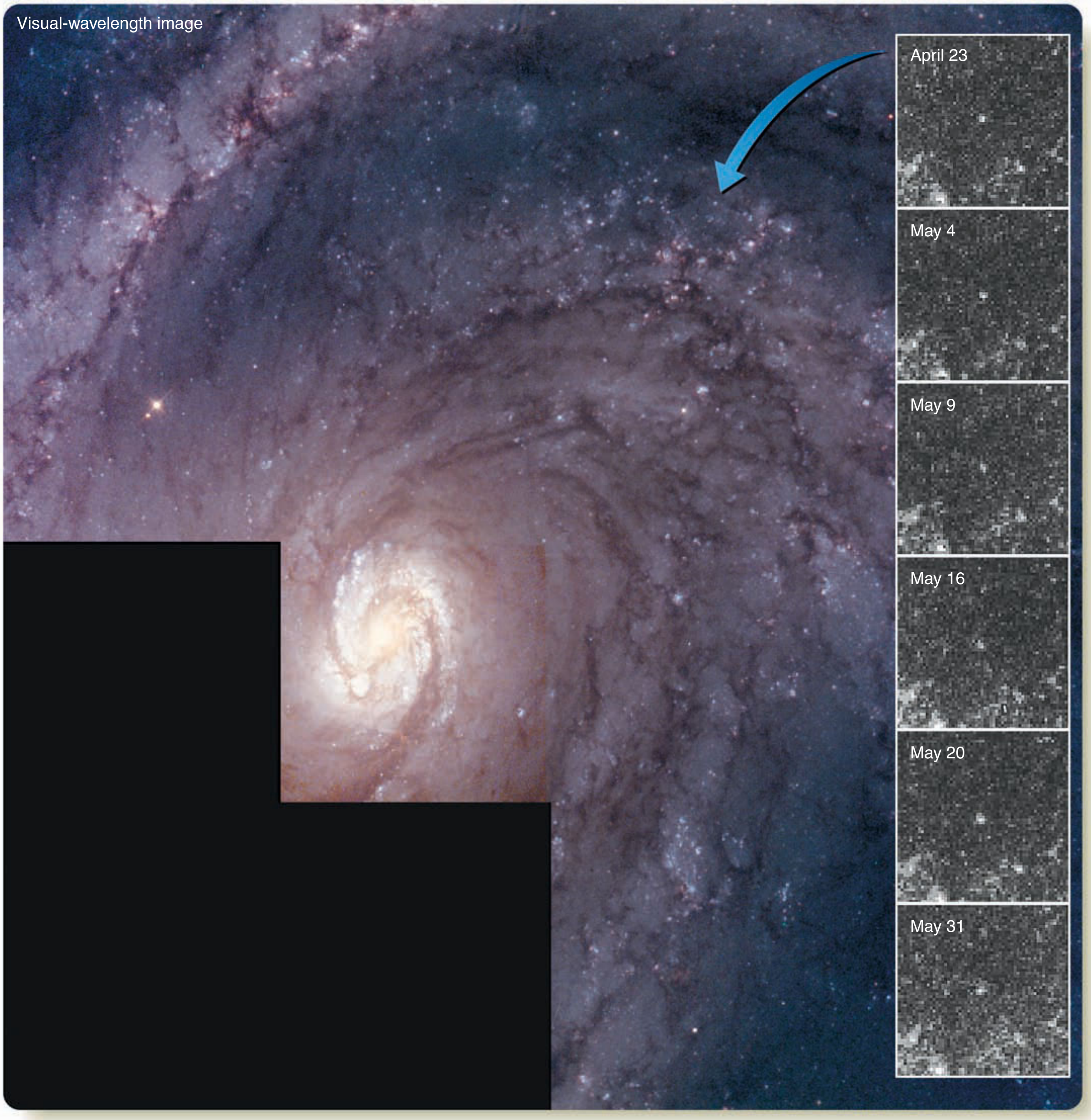


Figure 13-3

The vast majority of spiral galaxies are too distant for Earth-based telescopes to detect Cepheid variable stars. The Hubble Space Telescope, however, can locate Cepheids in some of these galaxies, as it has in the bright spiral galaxy M100. From a series of images taken on different dates, astronomers can locate Cepheids (inset), determine the period of pulsation, and measure the average apparent brightness. They can then deduce the distance to the galaxy—51 million ly for M100. (J. Trauger, JPL; Wendy Freedman, Observatories of the Carnegie Institution of Washington; and NASA)

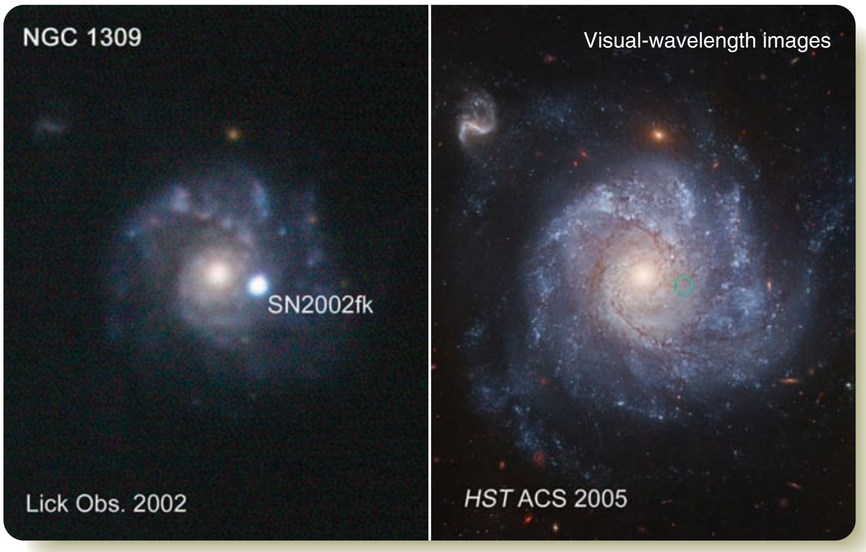
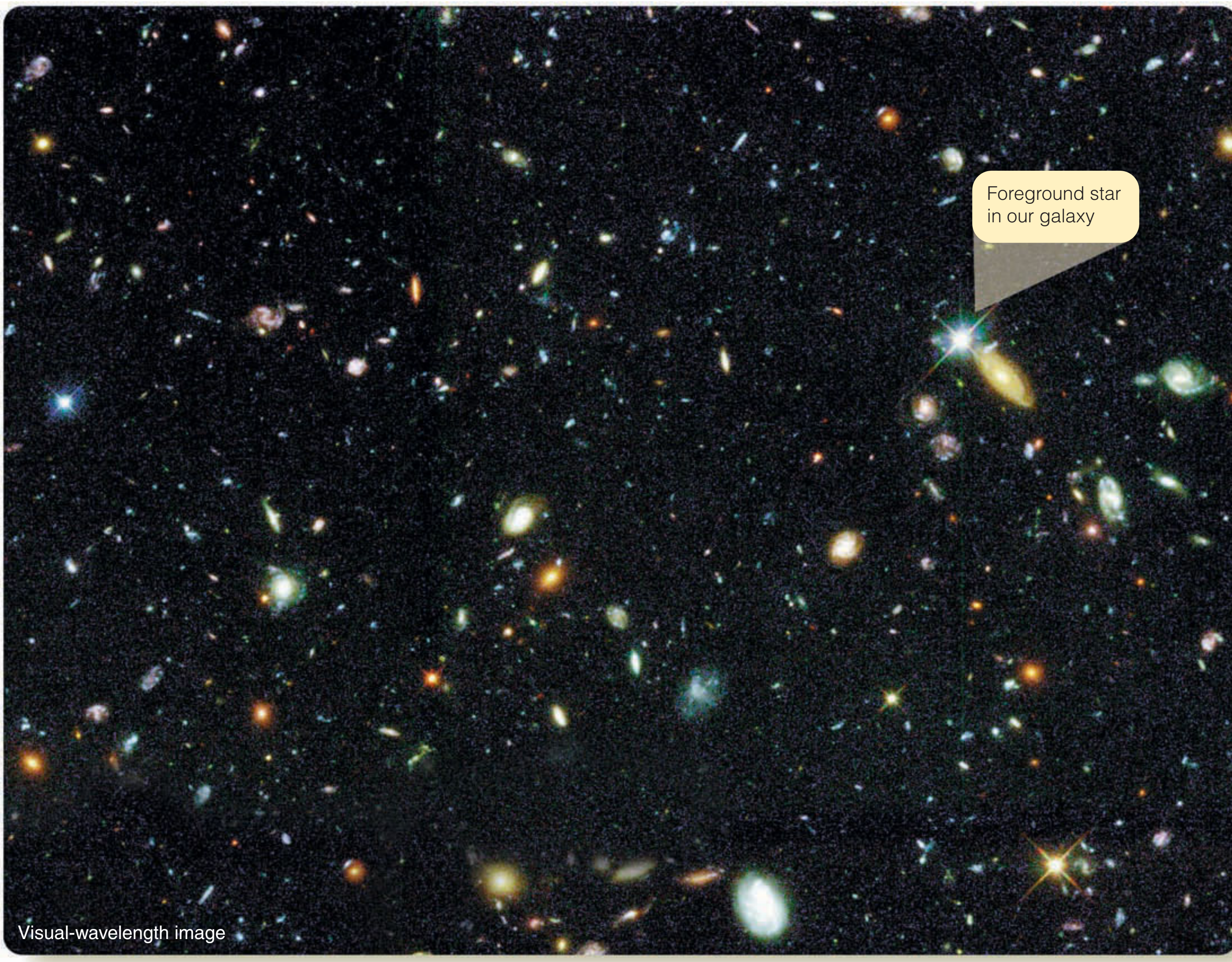


Figure 13-4

While dinosaurs roamed Earth, a white dwarf in the galaxy NGC 1309 collapsed and exploded as a type Ia supernova, and the light from that explosion reached Earth in 2002. Astronomers could find Cepheid variable stars in the galaxy, so they could determine that it was 100 million light-years away, and that allowed them to find the absolute magnitude of the supernova at its brightest. By combining observations of many supernovae, astronomers were able to calibrate type Ia supernovae as distance indicators. (Lick Observatory, W. Li and A. V. Filippenko, U. C. Berkeley; NASA, ESA, The Hubble Heritage Team, STScI/AURA and A. Riess, STScI)

Figure 13-5

Distant galaxies can be distinguished from stars by their shape. In this segment of the northern Hubble Deep Field, only four stars are visible; they are sharp points of light with diffraction spikes produced by the telescope optics. (Can you find them?) A few galaxies are obviously spirals, but most are only elongated shapes and cannot be classified at this distance. (R. Williams and the Hubble Deep Field Team, STScI, NASA)



lions of years ago when its light began the journey toward Earth. When you look at a more distant galaxy, you look further back into the past. For any given galaxy, this **look-back time** in years equals the distance its light traveled to reach us, in light-years.

You may have experienced look-back time if you have ever made a long-distance phone call carried by satellite or watched a TV newscaster interview someone on the other side of the world via satellite. A half-second delay occurs as a radio signal carries a question 23,000 miles out to a satellite, then back to Earth, and then carries the answer out to the satellite and again back to Earth. That half-second look-back delay can make people hesitate on long distance phone calls and produces an awkward delay in intercontinental TV interviews.

The look-back time to nearby objects is usually not significant. The look-back time across a football field is a tiny fraction of a second. The look-back time to the moon is 1.3 seconds, to the sun only 8 minutes, and to the nearest star about 4 years. The Andromeda Galaxy has a look-back time of about 2 million years, a mere eye blink in the lifetime of a galaxy. But when you look at more distant galaxies, the look-back time becomes an appreciable part of the age of the universe. Astronomers have strong evidence that the universe began about 13.7 billion years ago. When you look at the most distant visible galaxies, you are looking back over 10 billion years to a time when the universe may have been significantly different. This effect will be important as you think about the origin and evolution of galaxies and the universe as a whole.

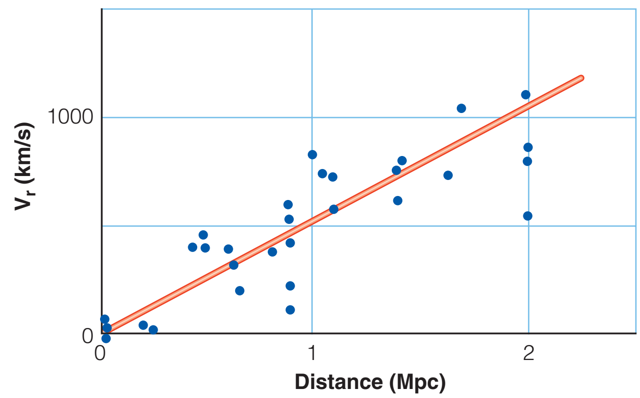
The Hubble Law

Although astronomers sometimes find it difficult to measure the distance to a galaxy, they often estimate such distances using a simple relationship.

Early in the 20th century, astronomers noticed that the spectral lines in galaxy spectra were shifted slightly toward longer wavelengths — the lines were redshifted. Interpreted as a consequence of the Doppler shift, these redshifts implied that the galaxies had large radial velocities and were receding from Earth. (Review the Doppler effect in Chapter 6.)

In 1929, the American astronomer Edwin Hubble published a graph that plotted the apparent velocity of recession versus distance for a number of galaxies. The points in the graph fell along a straight line (■ Figure 13-6). This relation between apparent velocity of recession and distance is known as the **Hubble law,** and the slope of the line is known as the **Hubble constant, *H*.**

The Hubble law is important in astronomy for two reasons. First, it is taken as evidence that the universe is expanding, a subject discussed in a later chapter. This chapter discusses another use of the Hubble law — estimating the distances to galaxies. As shown in ■ Reasoning with Numbers 13-1, the distance to a galaxy can be found by dividing its apparent velocity of recession by the Hubble constant. This is a very useful calculation, because it is usually possible to obtain a spectrum of a galaxy and measure its redshift even if it is too far away to have observable distance indicators. Obviously, knowing the Hubble constant is important.



■ **Figure 13-6**

Edwin Hubble's first diagram of the apparent velocities of recession and distances of galaxies did not probe very deeply into space. It did show, however, that the galaxies are receding.

Edwin Hubble's original measurement of *H* was too large because of errors in his distance measurements. Later astronomers have struggled to measure this important constant. The most precise measurements of the Hubble constant yield a value of about 70 km/s/Mpc with an uncertainty of just a few percent.

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Hubble Relations."

Diameter and Luminosity

Learning the distance to a galaxy is the key to finding its diameter and its luminosity. With even a modest telescope and a CCD camera, you can photograph a galaxy and measure its angular diameter in seconds of arc. If you know the distance to the galaxy, you could use the small-angle formula (Reasoning with Numbers 3-1) to find its linear diameter. Also, if you measured the apparent magnitude of the galaxy, you could use the distance to find its absolute magnitude and from that its luminosity, as shown in Reasoning with Numbers 8-2.

The results of such observations show that galaxies differ dramatically in size and luminosity. Irregular galaxies tend to be small, 1 to 25 percent the size of our galaxy, and have low luminosities. Although they are common, they are easy to overlook. Our Milky Way Galaxy is large and luminous compared with most spiral galaxies, though astronomers know of a few spiral galaxies that are nearly four times larger in diameter and about 10 times more luminous. Elliptical galaxies cover a wide range of diameters and luminosities. The largest, called giant ellipticals, are five times the size of our Milky Way Galaxy. But many elliptical galaxies are the very small dwarf ellipticals, which are only 1 percent the diameter of our galaxy.

To put galaxies in perspective, you can use an analogy. If our galaxy were an the size of an 18-wheeler, the smallest dwarf galaxies would be as small as pocket-size toy cars, and the largest giant ellipticals would be the size of jumbo jets. Even among spiral galaxies that looked similar, the sizes could range from that of a go-cart to a few times larger than a very big bus. Clearly, the diameter and luminosity of a galaxy do not determine its type. Other factors must influence the origin and evolution of galaxies.

Of the three basic parameters that describe a galaxy, you have found two—diameter and luminosity. The third, as was the case for stars, is more difficult to measure.

Mass

Although the mass of a galaxy is difficult to determine, it is an important quantity. It tells you how much matter the galaxy contains, which provides clues to the galaxy's origin and evolution. This section will describe two fundamental ways to find the masses of galaxies.

To begin, you should eliminate a **Common Misconception**. Some people think astronomers can see galaxies rotating. Some galaxies definitely look like spinning pinwheels, but you know that the orbital period of the sun around the Milky Way Galaxy is over 200 million years. Galaxies rotate very slowly, and no change is visible in a human lifetime. Nevertheless, the rotation of galaxies can give you a clue to their masses.

One way to find the mass of a galaxy is to measure its rotation. You know that the stars in the outer parts of the galaxy are in orbit around the center of mass, so you can use Kepler's third law to find the mass contained within their orbits. For the outermost stars, this is the mass of the visible part of the galaxy. All you need to know to make this calculation is the size of the galaxy and the orbital period with which its outer stars circle the galaxy.

Recall from the previous section that it is easy to find the size of the galaxy if you know the distance. You can measure its angular size and use the small-angle formula (Reasoning with Numbers 3-1). That gives you the size of the orbits of the outer stars, the semimajor axis of their orbits, a.

You can't observe the orbital period, P, directly because humans don't live long enough to see a galaxy rotate, but you can find the orbital velocity of the stars using the Doppler effect. Once you know the orbital velocity, you can find the orbital period by dividing the circumference of the orbit by the velocity. After finding a in astronomical units and P in years, you can use Kepler's third law to find the mass of the galaxy.

Measuring the orbital velocities of the stars in a galaxy requires the precise use of a spectrograph. Light enters a spectrograph through a narrow slit. If you focused the image of the galaxy on the slit of a spectrograph, you would see a bright spectrum formed by the nucleus of the galaxy, but you would also see fainter emission lines produced by ionized gas in the disk of the galaxy. Because the galaxy rotates, one side moves away from Earth and one side moves toward Earth, so the emission lines would be redshifted on one side of the galaxy and blueshifted on the other side. You could measure those changes in wavelength, use the Doppler formula to find the velocities, and plot a diagram showing the velocity of rotation at different distances from the center of the galaxy—a diagram called a **rotation curve.** The artwork in ■ Figure 13-7a shows the process of creating a rotation curve, and Figure 13-7b shows a real galaxy, its spectrum, and its rotation curve.

This way of measuring the mass of a galaxy is called the **rotation curve method,** and it is the most accurate way to find the masses of galaxies. But it works only for the nearer galaxies, whose rotation curves can be observed. Images of more distant galaxies are so small that astronomers cannot measure the radial velocity at different points along the galaxy. Also, careful studies of our own galaxy and others show that the outer parts of the rotation curve do not decline to lower velocities (Figure 13-7b). This shows that the outermost visible parts of some galaxies do not travel more slowly and reveals that these galaxies contain large amounts of mass outside this radius, perhaps in extended galactic coronas like the one that seems to surround our own galaxy. Because the rotation curve method can be applied only to nearby galaxies and because it cannot determine the masses of galactic coronas, you must consider a second way to find the masses of galaxies.

The **cluster method** of finding the mass of galaxies depends on the motions of galaxies within a cluster. If you measured the

Reasoning with Numbers | 13-1

The Hubble Law

The apparent velocity of recession of a galaxy, V_r, in kilometers per second is equal to the Hubble constant, H, multiplied times the distance to the galaxy, d, in megaparsecs:

$$V_r = Hd$$

Astronomers use this as a way to estimate distance from a galaxy's apparent velocity of recession.

Example: If a galaxy has a radial velocity of 700 km/s, and H is 70 km/s/Mpc,* then the distance to the galaxy equals the velocity divided by the Hubble constant, which is

$$d = \frac{700 \text{ km/s}}{70 \text{ km/s/Mpc}} = 10 \text{ Mpc}$$

Notice how the units km/s cancel out to leave the distance in Mpc.

*H has the units of a velocity divided by a distance. These are usually written as km/s/Mpc, meaning km/s per Mpc.

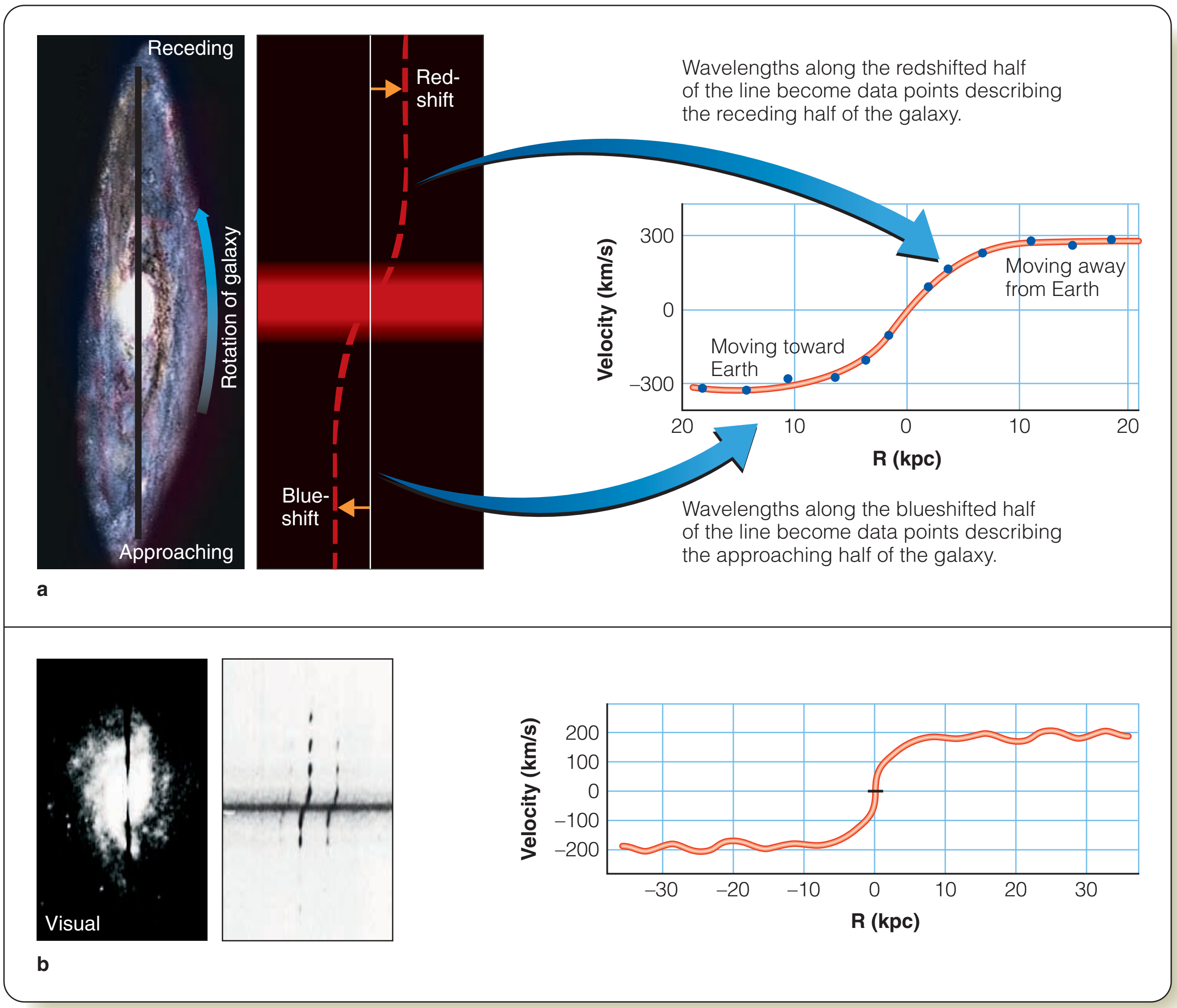


Figure 13-7

(a) In the artwork in the upper half of this diagram, the astronomer has placed the image of the galaxy over a narrow slit so that light from the galaxy can enter the spectrograph and produce a spectrum. A very short segment of the spectrum shows an emission line redshifted on the receding side of the rotating galaxy and blueshifted on the approaching side. Converting these Doppler shifts into velocities, the astronomer can plot the galaxy's rotation curve (right). (b) Real data are shown in the bottom half of this diagram. Galaxy NGC 2998 is shown over the spectrograph slit, and the segment of the spectrum includes three emission lines. (Courtesy Vera Rubin)

radial velocities of many galaxies in a cluster, you would find that some of the velocities are larger than others. Given the range of velocities and the size of the cluster, you could ask how massive a cluster of this size must be to hold itself together with this range of velocities. Dividing the total mass of the cluster by the number of galaxies in the cluster yields the average mass of a galaxy. This method contains the built-in assumption that the cluster is not flying apart. If it is, your result will be too large. Because it seems likely that most clusters are held together by their own gravity, the cluster method is probably valid.

A related way of measuring a galaxy's mass is called the **velocity dispersion method.** It works the same as the cluster method, but instead of observing the motions of galaxies in a cluster, astronomers observe the motions of matter within a galaxy. In the spectra of some galaxies, broad spectral lines indicate that stars and gas are moving at high velocities. If astronomers assume the galaxy is bound by its own gravity, they can ask how massive it must be to hold this moving matter within the galaxy. This method, like the one before, assumes that the system is not coming apart.

The masses of galaxies cover a wide range. The smallest contain about 10^{-6} as much mass as the Milky Way, and the largest contain as much as 50 times more than the Milky Way. The structure and evolution of a star is determined by its mass, but this is clearly not so for galaxies. You must search further for an explanation of the different types of galaxies.

Supermassive Black Holes in Galaxies

Rotation curves reveal the motions of the outer parts of a galaxy, but it is also possible to detect the Doppler shifts of stars orbiting close to the galaxy's center. Although these motions are not usually shown on rotation curves, they reveal something astonishing.

Measurements show that the stars near the centers of most galaxies are orbiting very rapidly. To hold stars in such small, short-period orbits, the centers of galaxies must contain masses of a million to a few billion solar masses. Yet no objects are visible, suggesting that the nuclei of galaxies contain supermassive black holes. You saw in the previous chapter that the Milky Way contains a supermassive black hole at its center. Evidently that is typical of galaxies.

Such a supermassive black hole cannot be the remains of a dead star. That would produce a black hole of only a few solar masses. A few dozen central black holes have measured masses, and those supermassive black holes typically contain 0.5 percent the mass of the nuclear bulges. A galaxy with a large nuclear bulge has a supermassive black hole whose mass is greater than the supermassive black hole in a galaxy with a small nuclear bulge. This implies that the supermassive black holes formed long ago as the galaxies began forming. Of course, matter has continued to drain into them, but they do not appear to have grown dramatically since they formed.

A billion-solar-mass black hole sounds like it contains a lot of mass, but note that it is only about 1 percent of the mass of a galaxy. The 3.7-million-solar-mass black hole at the center of the Milky Way Galaxy contains only a thousandth of a percent of the mass of the galaxy. In the next chapter you will discover that these supermassive black holes can produce titanic eruptions.

Dark Matter in Galaxies

Given the size and luminosity of a galaxy, astronomers can make a rough guess as to the amount of matter it should contain. Astronomers know how much light stars produce, and they know about how much matter there is between the stars, so it is quite possible to estimate very roughly the mass of a galaxy from its luminosity. But when astronomers measure the masses of galaxies, they find that the measured masses are much larger than expected given their luminosities. In the previous chapter, you studied the rotation curve of our own galaxy and concluded that it must contain large amounts of dark matter. This seems to be true of most galaxies. Measured masses of galaxies detect about ten times more mass than you would expect given the brightness of the galaxies.

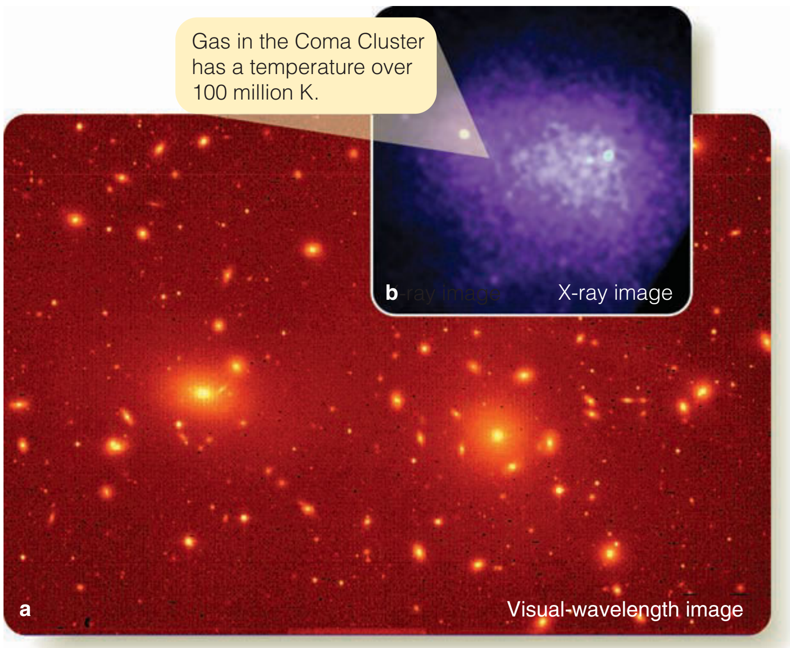
X-ray observations reveal more evidence of dark matter. X-ray images of galaxy clusters show that many of them are filled with very hot, low-density gas. The amount of gas present is much too small to account for the dark matter. Rather, the gas is important because it is very hot and its rapidly moving atoms have not leaked away. Evidently the gas is held in the cluster by a very strong gravitational field. To have enough gravity to hold such hot gas, the cluster must contain much more matter than what is visible. The detectable galaxies in the Coma cluster, for instance, amount to only a small fraction of the total mass of the cluster (■ Figure 13-8).

Gravitational Lensing and Dark Matter

When you solve a math problem and get an answer that doesn't seem right, you check your work for a mistake. Astronomers using the rotation curve method and the cluster method to measure the masses of galaxies have found evidence of large amounts of dark matter, so they look for ways to check their work. Fortunately, there is another way to detect dark matter that does not depend on observing orbital motion. You can understand how it works by following a light beam.

Albert Einstein described gravity as a curvature of space-time. The presence of mass actually distorts space-time around it, and that is what you feel as gravity. Einstein predicted that a light beam traveling through a gravitational field would be deflected by the curvature of space-time much as a golf ball is deflected as it rolls over a curved putting green. That effect has been observed and is a strong confirmation of Einstein's theories.

Gravitational lensing occurs when light from a distant object passes a nearby massive object and is deflected by the gravitational field. The gravitational field of the nearby object is actually a region of curved space-time that acts as a lens to deflect the passing light. Astronomers have used gravitational lensing to detect dark



■ **Figure 13-8**

(a) The Coma cluster of galaxies contains at least 1000 galaxies and is especially rich in E and S0 galaxies. Two giant galaxies lie near its center. Only the central area of the cluster is shown in this image. If the cluster were visible in the sky, it would span eight times the diameter of the full moon. (Gregory Bothun, University of Oregon) (b) In false colors, this X-ray image of the Coma cluster shows it filled and surrounded by hot gas. Note that the two brightest galaxies are visible in the X-ray image. (NASA/CXC/SAO/A.Vikhlinin et al.)

matter. When light from very distant galaxies passes through a cluster of galaxies on its way to Earth, it can be deflected by the strong curvature. That distorts the images of the distant galaxies into arcs, and the amount of the distortion depends on the mass of the cluster of galaxies (■ Figure 13-9). Observations of gravitational lensing made with very large telescopes confirm that clusters of galaxies contain far more matter than what is seen. That is, they contain large amounts of dark matter. This confirmation of the existence of dark matter is independent of orbital motion and gives astronomers much greater confidence that dark matter is real.

Dark matter is difficult to detect, and it is even harder to explain. Some astronomers have suggested that dark matter consists of low-luminosity white dwarfs and brown dwarfs scattered through the halos of galaxies. Both observation and theory support the idea that galaxies have massive extended halos, and searches for white dwarfs and brown dwarfs in the halo of our galaxy have been successful. Nevertheless, the searches have not turned up enough of these low-luminosity objects to make up all the dark matter.

Recall from the previous chapter that dark matter can't be hidden in vast numbers of black holes and neutron stars, because astronomers don't see the X-rays these objects would emit. There is roughly 10 times more dark matter than visible matter, and that many black holes would produce enough X-rays to be easily detected.

Because observations imply that the dark matter can't be hidden in normal objects, theorists suspect that the dark matter is made up of an unknown form of matter. Until recently, neutrinos were thought to be massless, but studies now suggest that they have a very small mass. Thus they could be part of the dark matter, but their masses are much too low to make up all of it. There must be some other undiscovered form of matter in the universe that is detectable only by its gravitational field. Dark matter is not a small issue. Observations show that 85 percent of the matter in the universe is dark matter. The universe you see—the kind of matter that you and the stars are made of—has been compared to the foam on an invisible ocean.

■ **Figure 13-9**

(a) The gravitational lens effect is visible in galaxy cluster 0024+1654 as its mass bends the light of a much more distant galaxy to produce arcs that are actually distorted images of the distant galaxy. This reveals that the galaxy cluster must contain large amounts of dark matter. (W. N. Colley and E. Turner, Princeton, and J. A. Tyson, Bell Labs, and NASA) (b) When two galaxy clusters passed through each other, normal matter (pink) collided and was swept out of the clusters, but the dark matter (purple), detected by gravitational lensing, was not affected. (NASA/CXC/CfA/STScI/Magellan/ESO WFI)

Dark matter remains one of the fundamental unresolved problems of modern astronomy. You will return to this problem in a later chapter when you try to understand how dark matter affects the nature of the universe, its past, and its future.

◄ SCIENTIFIC ARGUMENT ►

Why do you have to know the distance to a galaxy to find its mass?

A scientific argument must proceed step-by-step with no gaps. This is a good example. To find the mass of a galaxy, you need to know the size of the orbits and the orbital periods of stars at the galaxy's outer edge and then use Kepler's third law to find the mass inside the orbits. Measuring the orbital velocity of the stars as the galaxy rotates is easy if you can obtain a rotation curve from a spectrum. But you must also know the radii of the stars' orbits in meters or astronomical units. That is where the distance comes in. To find the radii of the orbits, you must know the distance to the galaxy. Then you can use the small-angle formula to convert the radius in arc seconds into a radius in parsecs or AU. If your measurement of the distance to the galaxy isn't accurate, you will get inaccurate radii for the orbits and will compute an inaccurate mass for the galaxy.

Many different measurements in astronomy depend on the calibration of the distance scale. Build a step-by-step argument to analyze the following problem. **What would happen to your measurements of the diameters and luminosities of the galaxies if astronomers discovered that the Cepheid variable stars were slightly more luminous than had been thought?**

◄ ►

13-3 The Evolution of Galaxies

YOUR GOAL IN this chapter is to build a theory to explain the evolution of galaxies. In the previous chapter, you developed a theory that described the origin of our own Milky Way Galaxy; presumably, other galaxies formed similarly. But why did some galaxies become spiral, some elliptical, and some irregular? Clues to that mystery lie in the clustering of galaxies.

Clusters of Galaxies

The distribution of galaxies is not entirely random. Galaxies tend to occur in clusters containing anywhere from a few galaxies to thousands. Deep photographs made with the largest telescopes reveal clusters of galaxies scattered out to the limits of detectability.

For this discussion, you can sort clusters of galaxies into two groups, rich and poor. **Rich galaxy clusters** contain over a thousand galaxies, mostly elliptical, scattered through a spherical volume about 3 Mpc (10^7 ly) in diameter. Such a cluster is very crowded, with the galaxies more concentrated toward the center. Rich clusters often contain one or more giant elliptical galaxies at their centers.

The Coma cluster (located in the constellation Coma Berenices) is an example of a rich cluster (Figure 13-8). It lies over 100 Mpc from Earth and contains at least 1000 galaxies, mostly E and S0 galaxies. Its galaxies are highly crowded around a central giant elliptical galaxy and a large S0.

Poor galaxy clusters contain fewer than 1000 galaxies, are irregularly shaped, and are less crowded. Our own **Local Group,** which contains the Milky Way Galaxy, is a good example of a poor cluster (■ Figure 13-10a). It contains a few dozen members scattered irregularly through a volume slightly over 1 Mpc in diameter. Of the brighter galaxies, 14 are elliptical, 3 are spiral, and 4 are irregular.

The total number of galaxies in the Local Group is uncertain because some lie in the plane of the Milky Way Galaxy and are difficult to detect. For example, a small dwarf galaxy, known as the Sagittarius Dwarf, has been found on the far side of our own galaxy, where it is almost totally hidden behind the star clouds of Sagittarius (Figure 13-10b). Even closer to the center of the Milky Way Galaxy is the Canis Major Dwarf Galaxy (Figure 13-10c). The galaxy was found by mapping the distribution of red supergiants detected by the 2MASS infrared survey. Other small galaxies in our Local Group have been found hidden behind the stars, gas, and dust of our Milky Way Galaxy.

Classifying galaxy clusters into rich and poor clusters reveals a fascinating and suggestive clue to the evolution of galaxies. In general, rich clusters tend to contain 80 to 90 percent E and S0 galaxies and few spirals. Poor clusters contain a larger percentage of spirals. Among isolated galaxies, those that are not in clusters, 80 to 90 percent are spirals. This suggests that a galaxy's environment is important in determining its structure and has led astronomers to suspect that the secrets to galaxy evolution lie in collisions between galaxies.

Colliding Galaxies

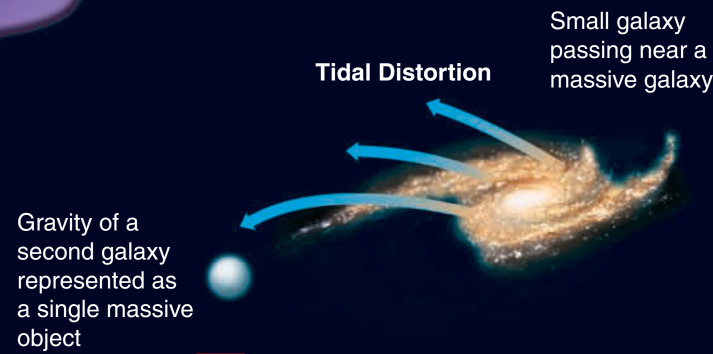
You should not be surprised that galaxies collide with each other. The average separation between galaxies is only about 20 times their diameter. Like two blindfolded elephants blundering about under a circus tent, galaxies should bump into each other once in a while. Stars, on the other hand, almost never collide. In the region of our galaxy near the sun, the average separation between stars is about 10^7 times their diameter. A collision between two stars is about as likely as a collision between two blindfolded gnats flitting about in a baseball stadium.

Study ■ **Interacting Galaxies** on pages 274–275 and notice four important points and three new terms:

1. Interacting galaxies can distort each other with tides, producing *tidal tails* and shells of stars. Encounters between galaxies may trigger the formation of spiral arms. Large galaxies can even absorb smaller galaxies, a process called *galactic cannibalism.*
2. Interactions between galaxies can trigger star formation.
3. Evidence left inside galaxies in the form of motions and multiple nuclei reveals that they have suffered past interactions and mergers.
4. Finally, the beautiful *ring galaxies* are bull's-eyes left behind by high-speed collisions.

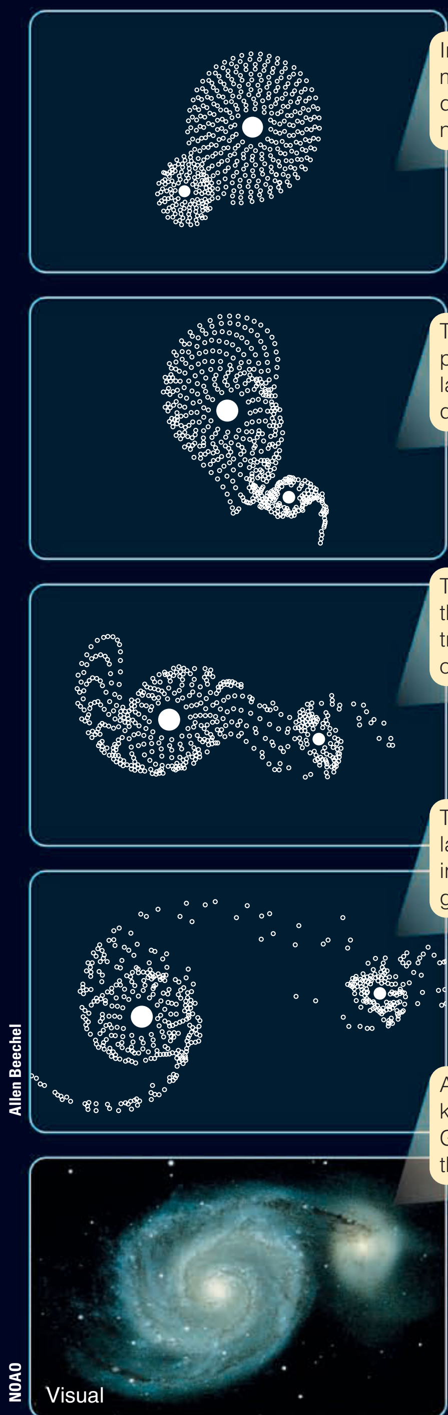
Interacting Galaxies

1 When two galaxies collide, they can pass through each other without stars colliding because the stars are so far apart relative to their sizes. Gas clouds and magnetic fields do collide, but the biggest effects may be tidal. Even when two galaxies just pass near each other, tides can cause dramatic effects, such as long streamers called **tidal tails.** In some cases, two galaxies can merge and form a single galaxy.



1a When a galaxy swings past a massive object such as another galaxy, tides are severe. Stars near the massive object try to move in smaller, faster orbits while stars farther from the massive object follow larger, slower orbits. Such tides can distort a galaxy or even rip it apart.

Galaxy interactions can stimulate the formation of spiral arms



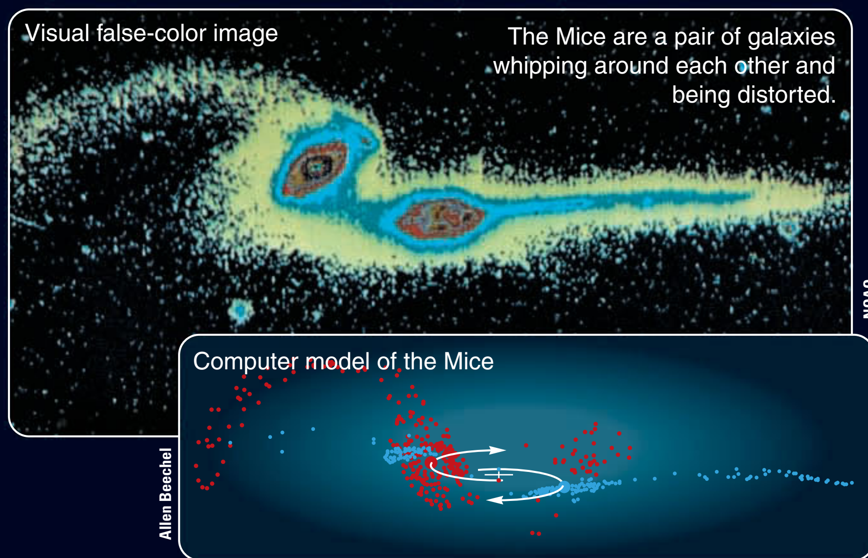
In this computer model, two uniform disk galaxies pass near each other.

The small galaxy passes behind the larger galaxy so they do not actually collide.

Tidal forces deform the galaxies and trigger the formation of spiral arms.

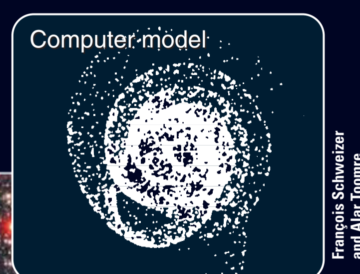
The upper arm of the large galaxy passes in front of the small galaxy.

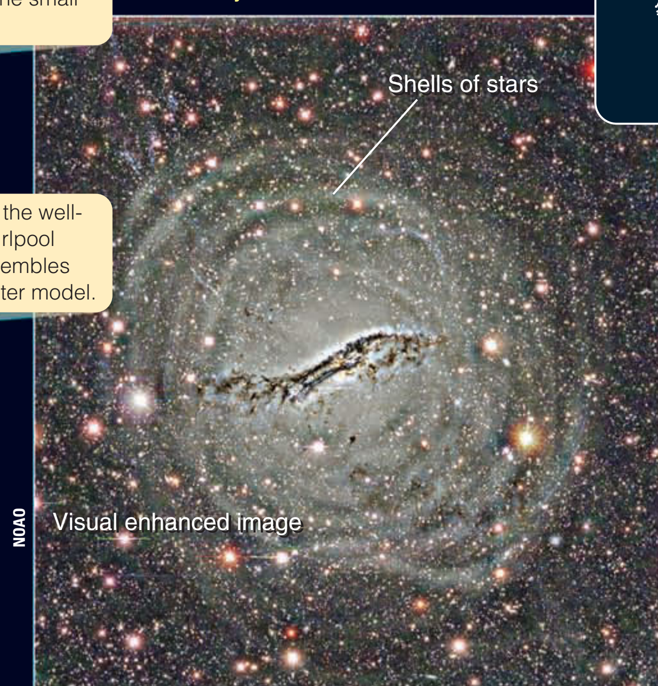
A photo of the well-known Whirlpool Galaxy resembles the computer model.



The Mice are a pair of galaxies whipping around each other and being distorted.

1b The merger of galaxies is called **galactic cannibalism.** Models show that merging galaxies spiral around their common center of mass while tides rip stars away and form shells.

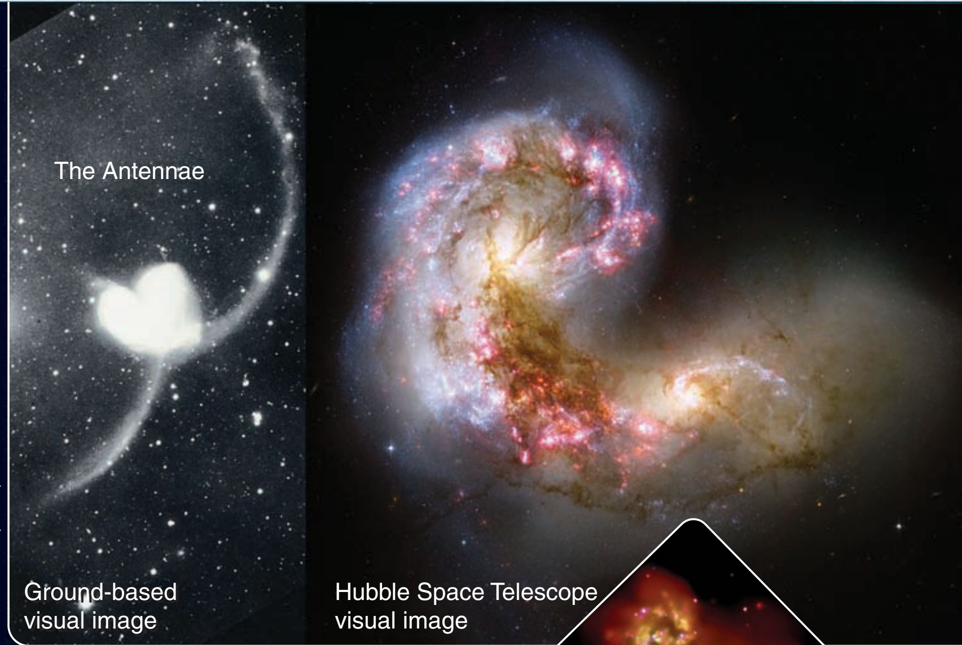


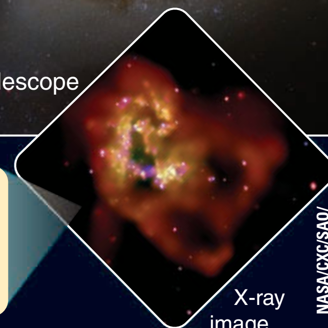


Such shells have been found around elliptical galaxies such as NGC 5128. It is peculiar in many ways and even has a belt of dusty gas. The shells revealed in this enhanced image are evidence that the giant galaxy has cannibalized at least one smaller galaxy. The giant galaxy itself may be the result of the merger of two large galaxies.

2 The collision of two galaxies can trigger firestorms of star formation as gas clouds are compressed. Galaxies NGC 4038 and 4039 have been known for years as the Antennae because the long tails visible in Earth-based photos resemble the antennae of an insect. Hubble Space Telescope images reveal that the two galaxies are blazing with star formation. Roughly a thousand massive star clusters have been born.

Spectra show that the galaxy is 10 to 20 times richer in elements like magnesium and silicon. Such metals are produced by massive stars and spread by supernova explosions.





An X-ray image of the Antennae shows clouds of very hot gas heated by supernovae exploding 30 times more often than in our own galaxy.

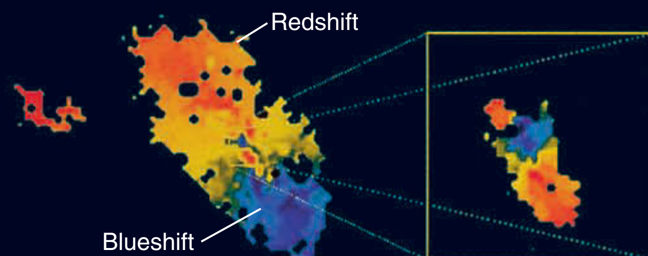
3 Evidence of past galaxy mergers shows up in the motions inside some galaxies. NCG 7251 is a highly distorted galaxy with tidal tails in this ground-based image.

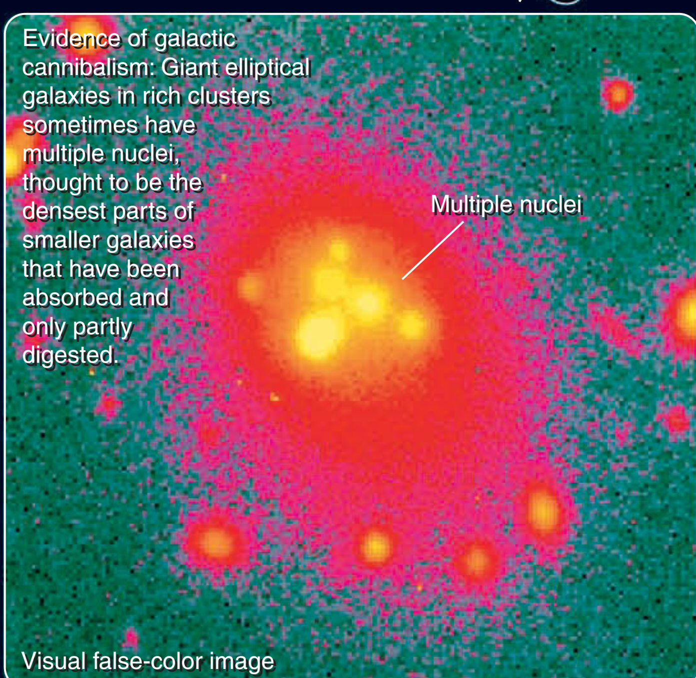
This Hubble Space Telescope image of the core of the galaxy reveals a small spiral spinning backward in the heart of the larger galaxy.

This counter rotation suggests that NCG 7251 is the remains of two oppositely rotating galaxies that merged about a billion years ago.

3a Radio evidence of past mergers: Doppler shifts reveal the rotation of the spiral galaxy M64. The upper part of the galaxy has a redshift and is moving away from Earth, and the bottom part of the galaxy has a blueshift and is approaching. A radio map of the core of the galaxy reveals that it is rotating backward. This suggests a merger long ago between two galaxies that rotate in opposite directions.

Rotation of galaxy M64





Evidence of galactic cannibalism: Giant elliptical galaxies in rich clusters sometimes have multiple nuclei, thought to be the densest parts of smaller galaxies that have been absorbed and only partly digested.

4 The Cartwheel galaxy below was once a normal galaxy but is now a **ring galaxy.** One of its smaller companions has plunged through at high speed almost perpendicular to the Cartwheel's disk. That has triggered a wave of star formation, and the more massive stars have exploded leaving behind black holes and neutron stars. Some of those are in X-ray binaries, and that makes the outer ring bright in X rays.

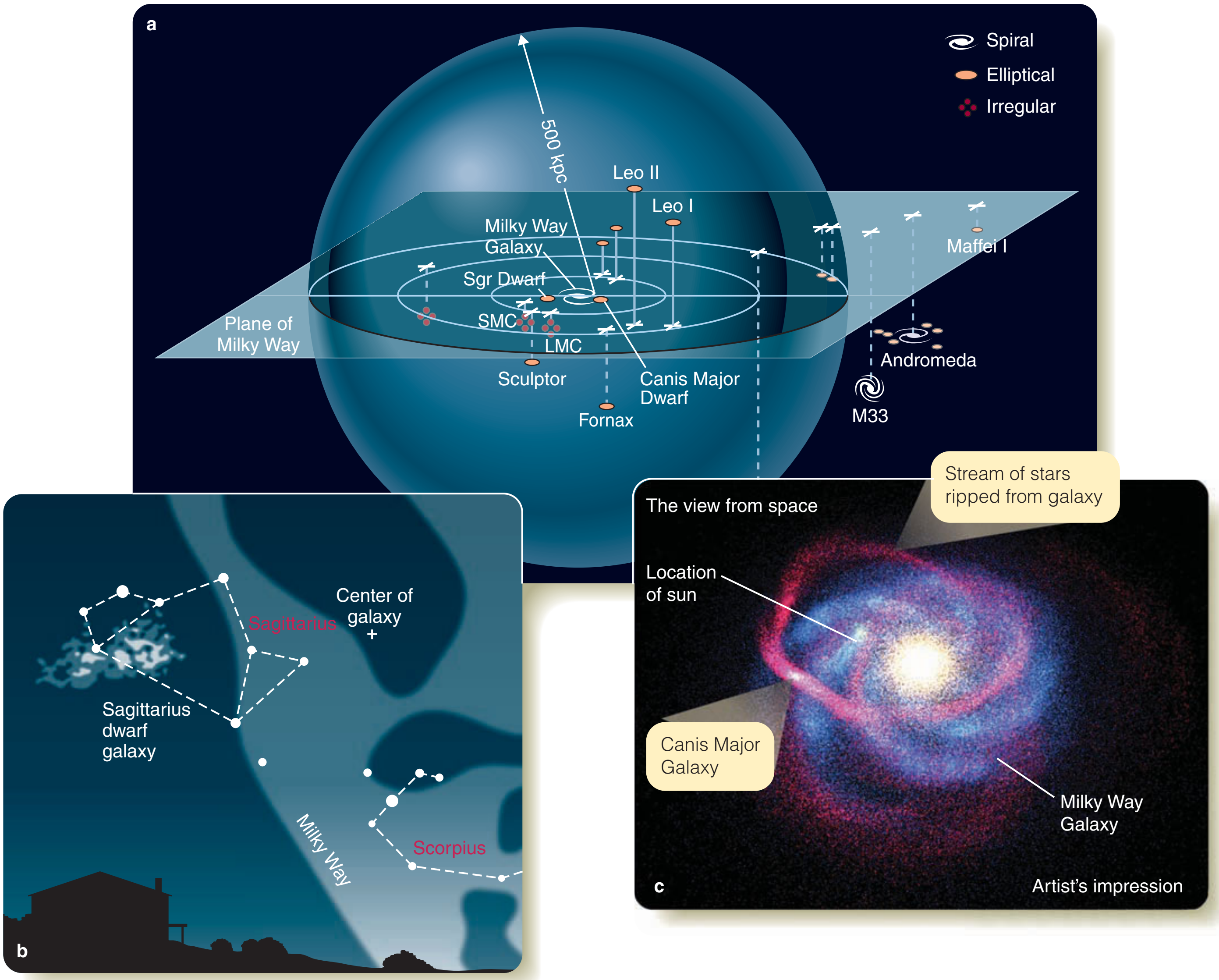


Figure 13-10

(a) The Local Group. Our galaxy is located at the center of this diagram. The vertical lines giving distances from the plane of the Milky Way are solid above the plane and dashed below. (b) The Sagittarius Dwarf Galaxy (Sgr Dwarf) lies on the other side of our galaxy. If you could see it in the sky, it would be 17 times larger than the full moon. (c) The Canis Major Dwarf Galaxy is even closer and is being ripped apart by tides to form streamers of stars. (Micolas Martin and Rodrigo Ibata, Strasbourg Observatory)

Evidence of galaxy mergers is all around you. Our Milky Way Galaxy is a cannibal galaxy, snacking on the Magellanic Clouds as they orbit past. Its tides are pulling the Sagittarius Dwarf Galaxy apart, and the Canis Major Dwarf galaxy has been almost completely digested as tides pulled stars away to form great streamers wrapped around the Milky Way Galaxy. Almost certainly, our galaxy has dined on other small galaxies. The nearby Andromeda Galaxy appears to be doing the same thing. Its halo contains streams of stars and star clusters that it has stripped from its swarm of small, nearby galaxies.

The Origin and Evolution of Galaxies

The test of any scientific understanding is whether you can put all the evidence and theory together to tell the history of the objects studied. Can you describe the origin and evolution of the galaxies? Just a few decades ago, it would have been impossible, but the evidence from space telescopes and new-generation telescopes on Earth combined with advances in computer modeling and theory allow astronomers to outline the story of the galaxies.

As you begin, you should eliminate a few older ideas immediately. It is easy to imagine that galaxies evolve from one type to another. But an elliptical galaxy cannot become a spiral galaxy or an irregular galaxy because ellipticals contain almost no gas and dust from which to make new stars. Elliptical galaxies can't be young. You can also argue that a single spiral or a single irregular galaxy cannot evolve into an elliptical galaxy because spiral and irregular galaxies contain both young and old stars. The old stars mean that spiral and irregular galaxies can't be young. The galaxy classes tell you something important, but a single galaxy does not change from one class to another any more than a cat can change into a dog.

Another old idea was that each galaxy formed from a single cloud of gas that contracted and formed stars. Astronomers call such a proposal a top-down theory. The evidence is quite clear that a galaxy does not form by a top-down contraction, but rather by a bottom-up accumulation of smaller clouds of gas and stars, the infall of gas, and the absorption of small galaxies.

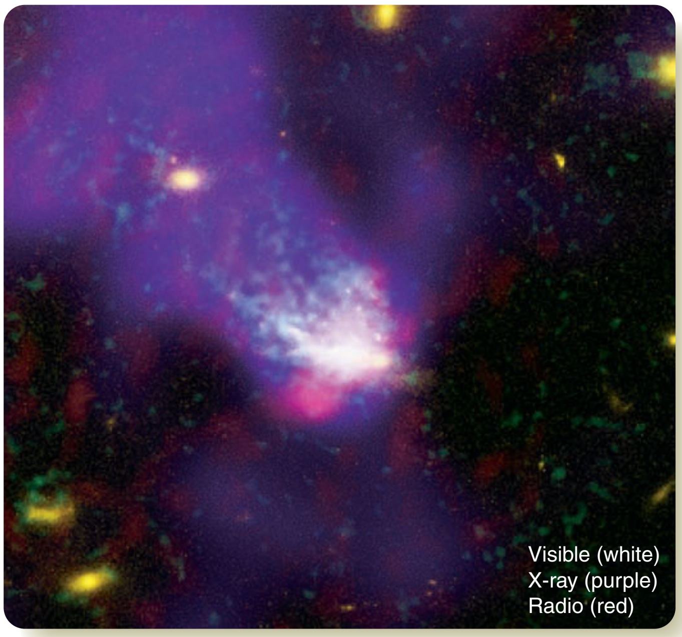
The ellipticals appear to be the products of galaxy mergers that triggered star formation and used up the gas and dust. In fact, astronomers see star formation being stimulated to high levels in many galaxies (■ Figure 13-11). **Starburst galaxies** are very luminous in the infrared because a collision has triggered a burst of star formation that heats the dust. The warm dust reradiates the energy in the infrared. The Antennae (page 275) contain over 15 billion solar masses of hydrogen gas and will become a starburst galaxy as the merger triggers rapid star formation. Supernovae in a starburst galaxy may eventually blow away much of the remaining gas and dust that doesn't get used up making stars. A galaxy that merges with a galaxy of similar size could be left with no gas and dust from which to make new stars. Astronomers now suspect that most ellipticals are formed by the merger of at least two or three galaxies. Galaxies crowded in rich clusters collide often, and that explains why elliptical and S0 galaxies are more common in such clusters.

In contrast, spirals seem never to have suffered major collisions. Their thin disks are delicate and would be destroyed by tidal forces during a collision with a similar size galaxy. Also, they retain plenty of gas and dust and continue making stars. Galaxies in poor clusters or isolated galaxies not in any cluster are unlikely to have collided with similar-sized galaxies, so they remain spirals. Our own Milky Way Galaxy has evidently never merged with another large galaxy, but it has clearly cannibalized smaller galaxies. Astronomers have found streams of stars in the halo of our galaxy that are too metal rich for their location. Another stream contains globular clusters with similar ages. These streams are evidently the remains of smaller galaxies that were absorbed.

Barred spiral galaxies may be the products of tidal interactions in which galaxies pass near each other but do not merge. Mathematical models show that bars are not stable and eventually dissipate. It may require tidal interactions with other galaxies to regenerate the bars. Because about two-thirds of all spiral gal-

■ **Figure 13-11**

Rapid star formation: (a) NGC 1569 is a starburst galaxy filled with clouds of young stars and supernovae. At least some starbursts are triggered by interactions between galaxies. (ESA/NASA/P. Anders) (b) The dwarf irregular galaxy NGC 1705 began a burst of star formation about 25 million years ago. (NASA/ESA/Hubble Heritage Team/AURA/STScI) (c) The inner parts of M64, known as the "black eye galaxy," are filled with dust produced by rapid star formation. Radio observations (page 275) show that the inner part of the galaxy rotates backward compared to the outer part of the galaxy, a product of a merger. Where the counterrotating parts of the galaxy collide, star formation is stimulated. (NASA/Hubble Heritage Team/AURA/STScI)



■ **Figure 13-12**

The distorted galaxy C153 is orbiting through the thin gas in its home cluster of galaxies at 2000 km/s (4.5 million miles per hour). At that speed it feels a tremendous wind stripping gas out of the galaxy in a trail 200,000 ly long. Such galaxies could quickly lose almost all of their gas and dust. (NASA, W. Keel, F. Owen, M. Ledlow, and D. Wang)

axies have bars, you can suspect that these tidal interactions are common.

Other processes can alter galaxies. The S0 galaxies may have lost much of their gas and dust in a burst of star formation but still managed to remain disk shaped. Also, galaxies moving through the gas trapped in dense galaxy clusters may have their own gas and dust blown away. For example, X-ray observations show that the Coma cluster contains thin, hot gas between the galaxies (Figure 13-8). A galaxy moving through that gas would encounter a tremendous wind that would blow its gas and dust away (■ Figure 13-12).

The dwarf ellipticals are too small to be made of merged spirals. Some dwarf ellipticals may have formed from small clouds of gas that were never incorporated into a larger galaxy, but other dwarf ellipticals, along with the irregular galaxies, may be small, scrambled fragments ripped from larger galaxies during collisions.

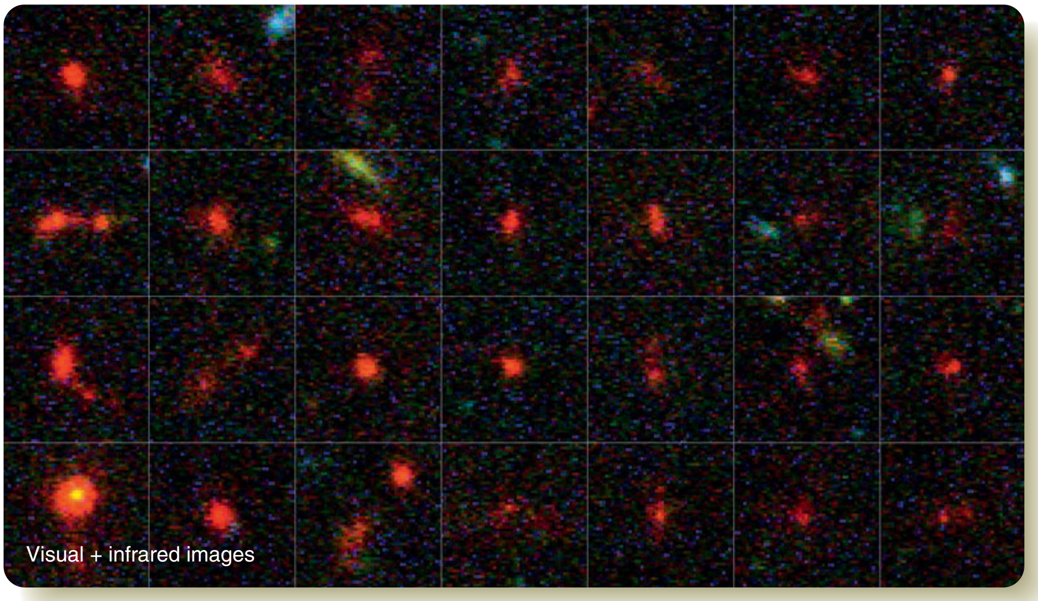
A good theory helps you understand how nature works, and astronomers are just beginning to understand the exciting and complex story of the galaxies. Nevertheless, it is already clear that galaxy evolution is much like a pie-throwing contest and just about as neat.

The Farthest Galaxies

Observations with the largest and most sophisticated telescopes are taking astronomers back to the age of galaxy formation. At great distances the look-back time is so large that they see the universe as it was soon after the galaxies began to form.

One way to search for distant galaxies is to use gravitational lensing. When light from very distant galaxies passes through a nearby cluster of galaxies, the lensing effect can focus the light from the distant galaxy, concentrate it, and make the galaxy look brighter. Using this technique, astronomers have found a few galaxies that are so far away they have enormous look-back times, and we see them as they were roughly half a billion years after the universe began. These galaxies are small and blue with active star formation. They appear to be the first clouds of gas and stars that merged to begin forming galaxies.

In the deep fields recorded by the Hubble Space Telescope, astronomers see faint, red galaxies (■ Figure 13-13). They look red because of their great redshifts. The floods of ultraviolet light emitted by these star-forming galaxies have been shifted into the far red part of the spectrum. Their look-back times are so great that they appear as they were when the universe was only



■ **Figure 13-13**

The Hubble Space Telescope was able to detect these distant galaxies because gravitational lensing has magnified and brightened them. Their redshifts are so large that we see them from Earth as they were only 900 million years after the universe began. These are some of the first galaxies to begin making stars. (NASA, ESA, R. Bouwens, and G. Illingworth UC Santa Cruz)

900 million years old. These galaxies are smaller, bluer, and of lower luminosity than galaxies later in the history of the universe.

At slightly smaller red shifts, astronomers see galaxies slightly later in the history of the universe and find that there were more spirals and fewer ellipticals than there are now. On the whole, galaxies long ago were more compact and more irregular. The observations show that galaxies were closer together then; about 33 percent of all distant galaxies are in close pairs, but only 7 percent of nearby galaxies are in pairs. This observational evidence clearly supports the bottom-up hypothesis that galaxies have evolved by merger.

◄ SCIENTIFIC ARGUMENT ►

Why can't spiral galaxies be produced by mergers?

This argument draws on both theory and observation. Observations show that spiral galaxies have thin disks in which the stars move in regular, concentric orbits. Computer models show that mergers with similar-sized galaxies produce strong tides that destroy the disk and scramble the orbits of the stars. Such a merger could produce an elliptical galaxy with its random orbits, but it could not preserve the thin disk of a spiral galaxy.

Now revise your argument. **What happens to the gas and dust in galaxies when they merge to form an elliptical galaxy?**

◄ ►

What Are We? Small and Proud

Do you feel insignificant yet? You are riding a small planet orbiting a humdrum star that is just one of at least 100 billion in the Milky Way Galaxy. You have just learned that there are at least 100 billion galaxies visible with existing telescopes, and that each of these galaxies contains roughly 100 billion stars. We humans fight wars over politics, religion, and economics; we do our work, play our games, and wash our laundry. It's all important stuff, but next time you are frantically rushing to a meeting, glance up at the sky. The sky is very deep.

When you look at galaxies you are looking across voids deeper than human imagination. You can express such distances with numbers and say a certain galaxy is 5 billion light-years from Earth, but the distance is truly beyond human comprehension. Furthermore, looking at distant galaxies also leads you back in time. You see distant galaxies as they were billions years ago. The realm of the galaxies is deep space and deep time.

Some people say astronomy makes them feel humble, but before you agree, consider that you can feel small without feeling humble. We humans live out our little lives on our little planet, but we are figuring out some of the biggest mysteries of the universe. We are exploring deep space and deep time and coming to understand what galaxies are and how they evolve. Most of all, we humans are beginning to understand what we are. That's something to be proud of.

Study and Review

Summary

- Astronomers divide galaxies into three classes—**elliptical (p. 262)**, **spiral (p. 262)**, and **irregular (p. 263)**—with subclasses specifying the galaxy's shape.
- Elliptical galaxies contain little gas and dust and cannot make new stars. Consequently, they lack hot, blue stars and have a reddish tint.
- Spiral galaxies contain more gas and dust in their disks and support active star formation, especially along the spiral arms. Some of the newborn stars are massive, hot, and blue, and that gives the spiral arms a blue tint. About two-thirds of spirals are **barred spiral galaxies (p. 262).**
- The halo and nuclear bulge of a spiral galaxy usually lack gas and dust and contain little star formation. The halos and nuclear bulges have a reddish tint because they lack hot, blue stars.
- Irregular galaxies have no obvious shape but contain gas and dust and support star formation.
- Galaxies are so distant astronomers measure their distances in **megaparsecs (p. 265)**—millions of parsecs.
- Astronomers find the distance to galaxies using **distance indicators (p. 265)**, sometimes called **standard candles (p. 265)**, objects of known luminosity. The most accurate distance indicators are the Cepheid variable stars. Globular clusters and type Ia supernovae explosions have also been calibrated as distance indicators.
- By calibrating additional distance indicators using galaxies of known distance, astronomers have built a **distance scale (p. 265).** The Cepheid variable stars are the most dependable.
- When astronomers look at a distant galaxy, they see it as it was when it emitted the light now reaching Earth. The **look-back time (p. 268)** to distant galaxies can be a significant fraction of the age of the universe.
- According to the **Hubble law (p. 268)**, the apparent velocity of recession of a galaxy equals its distance times the **Hubble constant (p. 268).** Astronomers can estimate the distance to a galaxy by observing its redshift, calculating its apparent velocity of recession, and then dividing by the Hubble constant.
- Once the distance to a galaxy is known, its diameter can be found from the small-angle formula and its luminosity from the magnitude–distance relation.
- Astronomers measure the masses of galaxies in two basic ways. The **rotation curve (p. 269)** of a galaxy shows the orbital motion of its stars, and astronomers can use the **rotation curve method (p. 269)** to find the galaxy's mass.
- The **cluster method (p. 269)** uses the velocities of the galaxies in a cluster to find the total mass of the cluster. The **velocity dispersion method (p. 270)** uses the velocities of the stars in a galaxy to find the total mass of the galaxy.
- Galaxies come in a wide range of sizes and masses. Some dwarf ellipticals and dwarf irregular galaxies are only a few percent the size and luminos-

ity of our galaxy, but some giant elliptical galaxies are five times larger than the Milky Way Galaxy.

- Stars near the centers of galaxies are following small orbits at high velocities, which suggests the presence of supermassive black holes in the centers of most galaxies.
- The mass of a galaxy's supermassive black hole is proportional to the mass of its nuclear bulge. That shows that the supermassive black holes must have formed when the galaxy formed.
- Observations of individual galaxies show that galaxies contain roughly 10 times more dark matter than visible matter.
- The hot gas held inside some clusters of galaxies and the **gravitational lensing (p. 271)** caused by the mass of galaxy clusters reveal that the clusters must be much more massive than can be accounted for by the visible matter—further evidence of dark matter.
- **Rich clusters (p. 273)** of galaxies contain thousands of galaxies with fewer spirals and more ellipticals. Our Milky Way Galaxy is located in a **poor cluster (p. 273)** of galaxies called the **Local Group (p. 273).** Poor clusters contain only a few galaxies with a larger proportion of spirals. This is evidence that galaxies evolve by collisions and mergers.
- When galaxies collide, tides twist and distort their shapes and can produce **tidal tails (p. 276).**
- Large galaxies can absorb smaller galaxies in what is called **galactic cannibalism (p. 276).** You can see clear evidence that our own Milky Way Galaxy is devouring some of the small galaxies that orbit nearby and that our galaxy has consumed other small galaxies in the past.
- Shells of stars, counterrotating parts of galaxies, streams of stars in the halos of galaxies, and multiple nuclei are evidence that galaxies can merge.
- **Ring galaxies (p. 277)** are produced by high-speed collisions in which a small galaxy plunges through a larger galaxy perpendicular to its disk.
- The compression of gas clouds can trigger bursts of star formation, producing **starburst galaxies (p. 275).** The rapid star formation can produce lots of dust, which is warmed by the stars to emit infrared radiation, making the galaxy bright in the infrared.
- The merger of two larger galaxies can scramble star orbits and drive bursts of star formation to use up gas and dust. Most larger ellipticals have evidently been produced by past mergers.
- Spiral galaxies have thin, delicate disks and appear not to have suffered mergers with large galaxies.
- A galaxy moving through the gas in a cluster of galaxies can be stripped of its own gas and dust and may become an S0 galaxy.
- Rare isolated galaxies tend to be spirals, which suggests that they have not suffered collisions with similar sized galaxies.
- At great distance and great look-back times, the largest telescopes reveal that galaxies were smaller, more irregular, and closer together. There were more spirals and fewer ellipticals long ago.
- At the largest distances, astronomers find small irregular clouds of stars that may be the objects that fell together to begin forming galaxies when the universe was very young.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why didn't astronomers at the beginning of the 20th century recognize galaxies for what they are?
2. What is the difference between an E0 galaxy and an E1 galaxy?
3. What is the difference between an Sa and an Sb galaxy? Between an SBb and an Sb?
4. Why can't a galaxy evolve from elliptical to spiral?
5. Why can't a galaxy evolve from spiral to elliptical?
6. Which galaxies are poorest in gas and dust? Which have the fewest bright stars?
7. Why are Cepheid variable stars good distance indicators? What about supernovae?
8. Why is it difficult to measure the Hubble constant?
9. How is the rotation curve method related to binary stars and Kepler's third law?
10. What evidence is there that galaxies contain dark matter? Supermassive black holes?
11. What evidence is there that galaxies collide and merge?
12. Why are the shells visible around some elliptical galaxies significant?
13. Ring galaxies often have nearby companions. What does that suggest?
14. Propose an explanation for the lack of gas, dust, and young stars in elliptical galaxies.
15. How do deep images by the Hubble Space Telescope confirm the bottom-up theory of galaxy formation?
16. **How Do We Know?** Classification helped Darwin understand how creatures evolve. How has classification helped you understand how galaxies evolve?
17. **How Do We Know?** How might you be misled if you studied only the brightest galaxies in a cluster?

Discussion Questions

1. Why do astronomers believe that galaxy collisions are likely, but star collisions are not?
2. Should an orbiting infrared telescope find irregular galaxies bright or faint in the far-infrared? Why? What about elliptical galaxies?

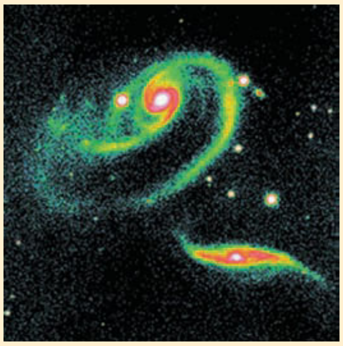
Problems

1. If a galaxy contains a type I (classical) Cepheid with a period of 30 days and an apparent magnitude of 20, what is the distance to the galaxy?
2. If you find a galaxy that contains globular clusters that are 2 arc seconds in diameter, how far away is the galaxy? (*Hints:* Assume that a globular cluster is 25 pc in diameter and see Reasoning with Numbers 3-1.)
3. If a galaxy contains a supernova that at its brightest has an apparent magnitude of 17, how far away is the galaxy? (*Hints:* Assume that the absolute magnitude of the supernova is -19 and see Reasoning with Numbers 8-2.)
4. If you find a galaxy that is the same size and mass as our Milky Way Galaxy, what orbital velocity would a small satellite galaxy have if it orbited 50 kpc from the center of the larger galaxy? (*Hint:* See Reasoning with Numbers 4-1.)
5. Find the orbital period of the satellite galaxy described in Problem 4. (*Hint:* See Reasoning with Numbers 8-4.)
6. If a galaxy has a radial velocity of 2000 km/s and the Hubble constant is 70 km/s/Mpc, how far away is the galaxy? (*Hint:* Use the Hubble law.)
7. If you find a galaxy that is 20 arc minutes in diameter, and you measure its distance to be 1 Mpc, what is its diameter? (*Hint:* See Reasoning with Numbers 3-1.)
8. You have found a galaxy in which the outer stars have orbital velocities of 150 km/s. If the radius of the galaxy is 4 kpc, what is the orbital period of the outer stars? (*Hints:* 1 pc $= 3.08 \times 10^{13}$ km, and 1 yr $= 3.16 \times 10^{7}$ s.)
9. A galaxy has been found that is 5 kpc in radius and whose outer stars orbit the center with a period of 200 million years. What is the mass of the galaxy? On what assumptions does this result depend? (*Hint:* See Reasoning with Numbers 8-4.)

Learning to Look

1. Study Figure 13-5. Which galaxies do you suppose are nearest, and which are farthest away. How is your estimate based on your calibration of galaxy size and luminosity?
2. Hubble's first determination of the Hubble constant was too high because the calibration of Cepheid variable stars was not very good. Use his original diagram, shown in Figure 13-6, to estimate his first determination of the Hubble constant.
3. This image of M33, the Pinwheel Galaxy, has emission by ionized hydrogen enhanced as bright pink. Discuss the location of these clouds of gas and explain how that provides important evidence toward understanding spiral arms.

4. In the image at right you see two interacting galaxies; one is nearly face on and the other is nearly edge on. Discuss the shapes of these galaxies and describe what is happening.



14 Active Galaxies and Supermassive Black Holes

Like cosmic fireworks, the galaxy Fornax A has erupted to expel two lobes visible at radio wavelengths. (Image courtesy of Barbara Harris)

Guidepost

In the last few chapters, you have explored our own and other galaxies, and now you are ready to stretch your scientific imagination and study some of the most powerful objects in nature. Supermassive black holes at the centers of galaxies are common but extreme. To study them, you will be combining many of the ideas you have discovered so far to answer three essential questions:

What makes some galaxy cores active?

How do supermassive black holes erupt?

How did supermassive black holes form and evolve?

The formation and evolution of supermassive black holes leads your imagination outward into space and backward in time to the era of galaxy formation. In the next chapter, you will take the next step and try to understand the birth and evolution of the entire universe.

Animated! This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.

Somewhere something incredible is waiting to be known.

CARL SAGAN

SUPERMASSIVE BLACK HOLES containing millions to billions of solar masses lurk at the centers of most galaxies. Such monsters can release tremendous energy and cause eruptions that dwarf the largest supernovae. Yet most galaxies that contain supermassive black holes are not erupting. Why? The study of supermassive black holes will introduce you to some of the most extreme conditions in the universe.

14-1 Active Galactic Nuclei

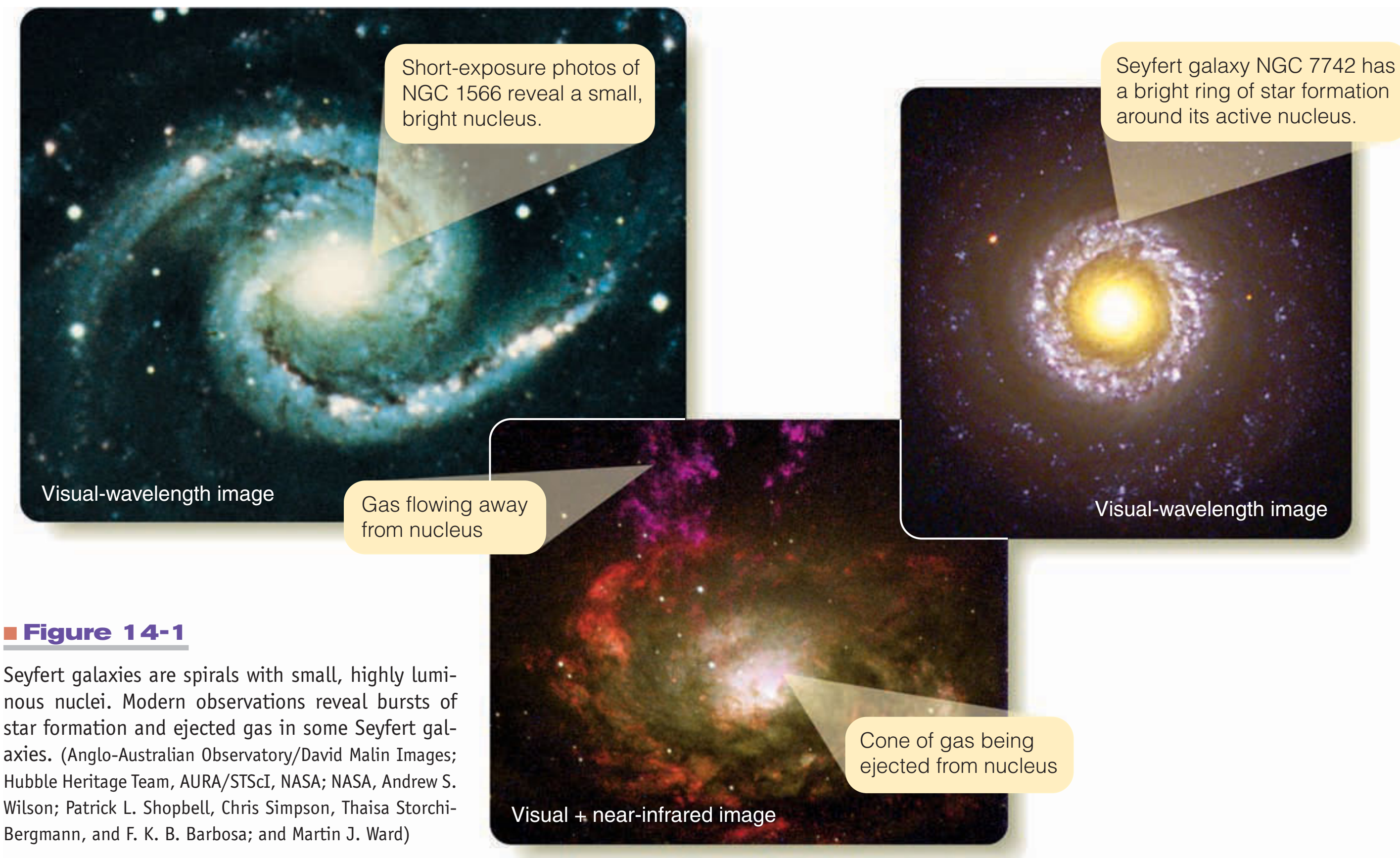
WITH THE CONSTRUCTION of the first large radio telescopes in the 1950s, astronomers discovered that some galaxies, dubbed **radio galaxies,** were bright at radio wavelengths. Later, when telescopes went into orbit, these galaxies were found to be emitting energy at other wavelengths as well, and they became known as **active galaxies.** Modern observations show that the energy comes from the nuclei of the galaxies, which are now known as **active galactic nuclei (AGN).** Only a few percent of galaxies are active, so your first step is to meet these peculiar galaxies.

Seyfert Galaxies

In 1943, Mount Wilson astronomer Carl K. Seyfert published his study of spiral galaxies. Observing at visual wavelengths, Seyfert found that some spiral galaxies have small, highly luminous nuclei with peculiar spectra (■ Figure 14-1). These galaxies are now known as **Seyfert galaxies.**

The spectrum of a normal galaxy is the combined spectra of many billions of stars, so galaxy spectra contain only a few of the strongest absorption lines found in stellar spectra. But the spectra of Seyfert galaxy nuclei contain broad emission lines of highly ionized atoms. Emission lines suggest a hot, low-density gas, and the presence of ionized atoms suggests that the gas is very excited. The width of the spectral lines suggests large Doppler shifts produced by high velocities in the nuclei; gas approaching Earth produces blueshifted spectral lines, and gas going away produces redshifted lines. The combined light, therefore, contains broad spectral lines showing that the velocities at the center of Seyfert galaxies are roughly 10,000 km/s, about 30 times greater than velocities at the center of normal galaxies. Something violent is happening in the cores of Seyfert galaxies.

About 2 percent of spiral galaxies appear to be Seyfert galaxies, and they are classified into two categories. Type 1 Seyfert galaxies are very luminous at X-ray and ultraviolet wavelengths and have the typical broad emission lines with sharp, narrow cores. Rapidly moving gas produces the broad part of the lines,



■ **Figure 14-1**

Seyfert galaxies are spirals with small, highly luminous nuclei. Modern observations reveal bursts of star formation and ejected gas in some Seyfert galaxies. (Anglo-Australian Observatory/David Malin Images; Hubble Heritage Team, AURA/STScI, NASA; NASA, Andrew S. Wilson; Patrick L. Shopbell, Chris Simpson, Thaisa Storchi-Bergmann, and F. K. B. Barbosa; and Martin J. Ward)

How Do We Know? 14-1

Statistical Evidence

How can statistics be useful if they can't be specific? Some scientific evidence is statistical. Observations suggest, for example, that Seyfert galaxies are more likely to be interacting with a nearby companion than a normal galaxy is. This is statistical evidence, so you can't be certain that any specific Seyfert galaxy will have a companion. How can scientists use statistical evidence to learn about nature when statistics contain built-in uncertainty?

Meteorologists use statistics to determine how frequently storms of a certain size are likely to occur. Small storms happen every year, but medium-sized storms may happen on average only every ten years. Hundred-year storms are much more powerful but occur much less frequently—on average only once in a hundred years.

Those meteorological statistics can help you make informed decisions — as long as you understand the powers and limitations of statistics. Would you buy a house protected from a river by a levee that was not designed to withstand a hundred-year storm? In any one year, the chance of your house being destroyed would be only 1 in 100. You know the storm will hit eventually, but you don't know when. If you buy the house, a storm might destroy the levee the next year, but you might own the house for your whole life and never see a hundred-year storm. The statistics can't tell you anything about a specific year.

Before you buy that house, there is an important question you should ask the meteorologists. "How much data do you have on storms?" If they only have 10 years of data, then they don't really know much about hundred-year storms. If they have three centuries of data, then their statistical data are significant.

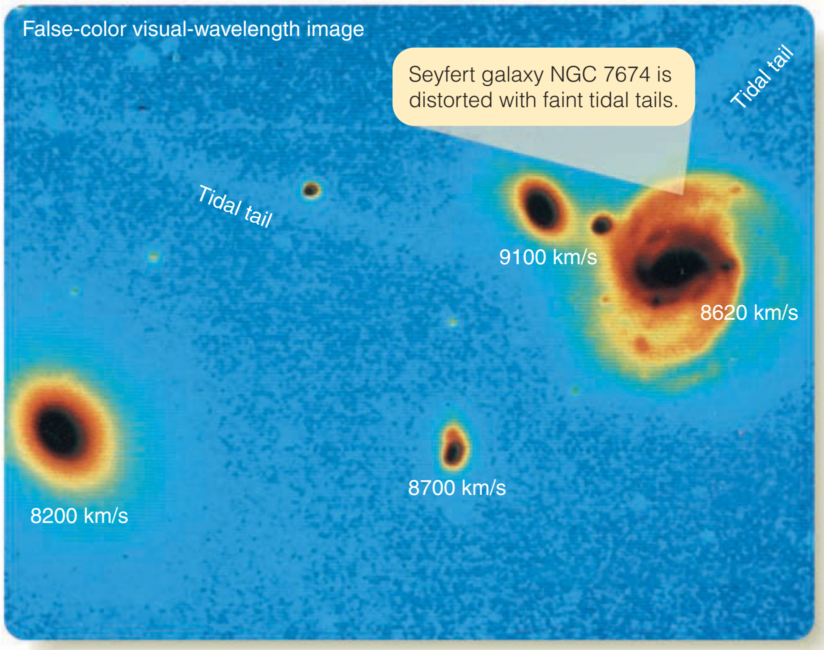
Sometimes people dismiss important warnings by saying, "Oh, that's only statistics." Scientists can use statistical evidence if it passes two tests. It cannot be used to draw conclusions about specific cases, and it must be based on large enough data samples so the statistics are significant. With these restrictions, statistical evidence can be a powerful scientific tool.

Statistics can tell you that a bad storm will eventually hit, but it can't tell you when. (Marko Georgiev/Getty Images)

but some lower-velocity gas must also be present to produce the narrow cores. Type 2 Seyfert galaxies have much weaker X-ray emission and have emission lines that are narrower than those of type 1 Seyfert galaxies but still broader than spectral lines produced by a normal galaxy.

The brilliant nuclei of Seyfert galaxies fluctuate rapidly, especially at X-ray wavelengths. A Seyfert nucleus can change its X-ray brightness by a significant amount in only minutes. As you saw when you studied neutron stars, an astronomical body cannot change its brightness in a time shorter than the time it takes light to cross its diameter. If the Seyfert nucleus can change in a few minutes, then it cannot be more than a few light-minutes in diameter. In spite of their small size, the cores of Seyfert galaxies produce tremendous amounts of energy. The brightest emit a hundred times more energy than the entire Milky Way Galaxy. All of this is consistent with supermassive black holes lurking in the centers of these galaxies orbited by hot accretion disks no bigger than Earth's orbit.

The shapes of Seyfert galaxies provide an important clue. Some astronomers argue that Seyfert galaxies are more common in interacting pairs of galaxies than in isolated galaxies. Also, about 25 percent have peculiar shapes suggesting tidal interactions with other galaxies (■ Figure 14-2a). This statistical evidence (■ How Do We Know? 14-1) hints that Seyfert galaxies may have been triggered into activity by collisions or interactions with companions. You will find more such evidence as you study other kinds of active galaxies.



■ **Figure 14-2**

The Seyfert galaxy NGC7674 appears to be distorted by interacting with its companion galaxies. You can be sure they are a group because they all have similar apparent velocities of recession and must be at about the same distance. (John W. Mackenty, Institute for Astronomy, University of Hawaii)

Double-Lobed Radio Sources

Beginning in the 1950s, radio astronomers found that some sources of radio energy in the sky consisted of pairs of radio-bright regions. When optical telescopes studied the locations of these **double-lobed radio galaxies,** they revealed galaxies located between the two regions emitting radio energy. Unlike Seyfert galaxies, where the energy comes from their cores, these radio galaxies were producing energy from two radio lobes.

Study ■ **Cosmic Jets and Radio Lobes** on pages 286–287 and notice four important points and two new terms:

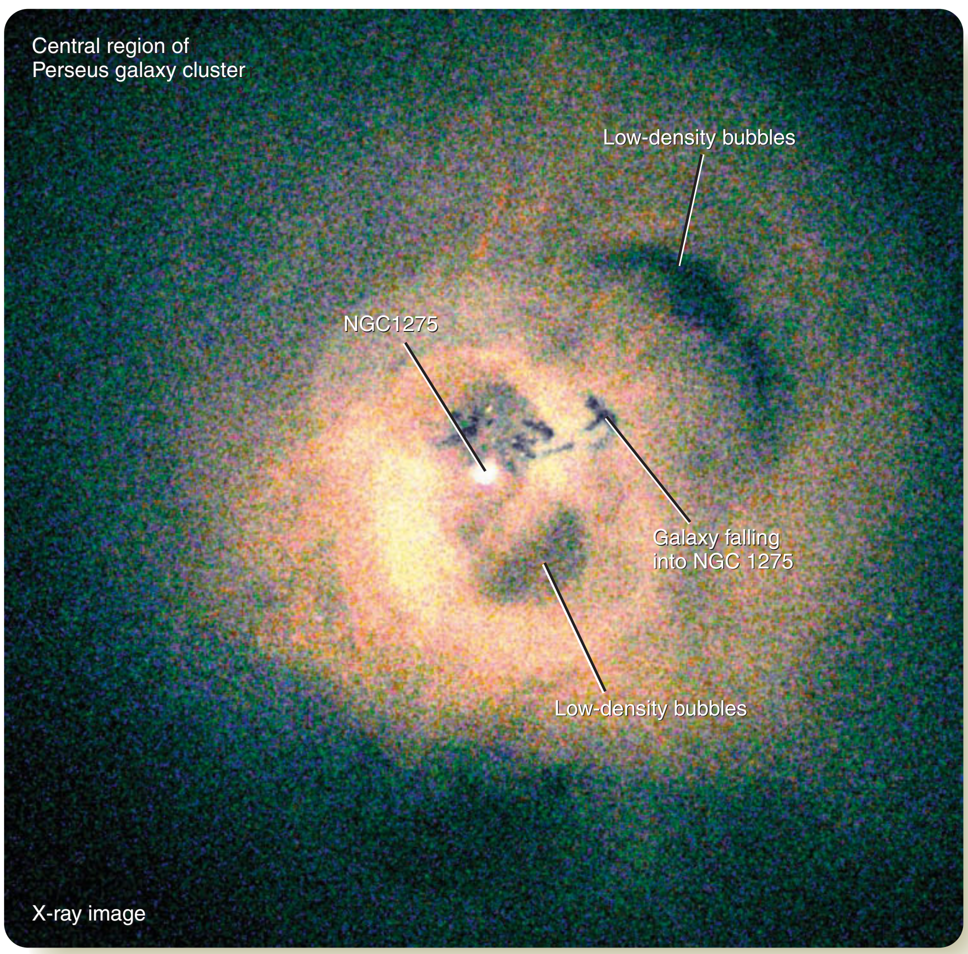
1. The shapes of radio lobes suggest that they are inflated by jets of excited gas emerging from the nucleus of the central galaxy. This has been called the *double-exhaust model,* and the presence of *hot spots* and synchrotron radiation shows that the jets are very powerful.
2. Acitive galaxies that have jets and radio lobes are often deformed or interacting with other galaxies.
3. The complex shapes of some jets and radio lobes can be explained by the motions of the active galactic nuclei. A good example of this is 3C31 (the 31st source in the *Third Cambridge Catalog of Radio Sources*) with its twisting radio lobes.
4. These jets are consistent with matter falling into a central supermassive black hole. You have seen similar jets produced by accretion disks around protostars, neutron stars, and stellar mass black holes; although the details are not understood, the same process seems to be producing all of these jets.

The evidence shows that the cores of most galaxies are occupied by supermassive black holes and that matter flowing into the black holes can produce tremendous energy and eject jets in opposite directions. In fact, the violence of these active galaxies is so great it can influence entire clusters of galaxies. The Perseus galaxy cluster contains thousands of galaxies and is one of the largest objects in the universe. One of its galaxies, NGC 1275, is one of the largest galaxies known. It is pumping out jets of high-energy particles, heating the gas in the galaxy cluster, and inflating low-density bubbles that distort the huge gas cloud (■ Figure 14-3). The hot gas observed in galaxy clusters is heated to multimillion-degree temperatures as galaxy after galaxy goes through eruptive stages that can last for hundreds of millions of years. Astronomers observe NGC 1275 erupting now, and it is so powerful and has heated the surrounding gas so hot that it has probably limited its own growth.

Quasars

Distant galaxies look like fuzzy blobs on photographs, so astronomers were surprised in the early 1960s when some radio sources turned out to look like stars in visual-wavelength photographs. First called quasi-stellar objects, they were soon referred to as **quasars.** Many more quasars have been found over the years, and most are radio silent. Nevertheless, they have been a puzzle to astronomers.

The spectra of quasars were strange in that they contained a few unidentified emission lines superimposed on a continuous spectrum. In 1963 Maarten Schmidt at Hale Observatories tried

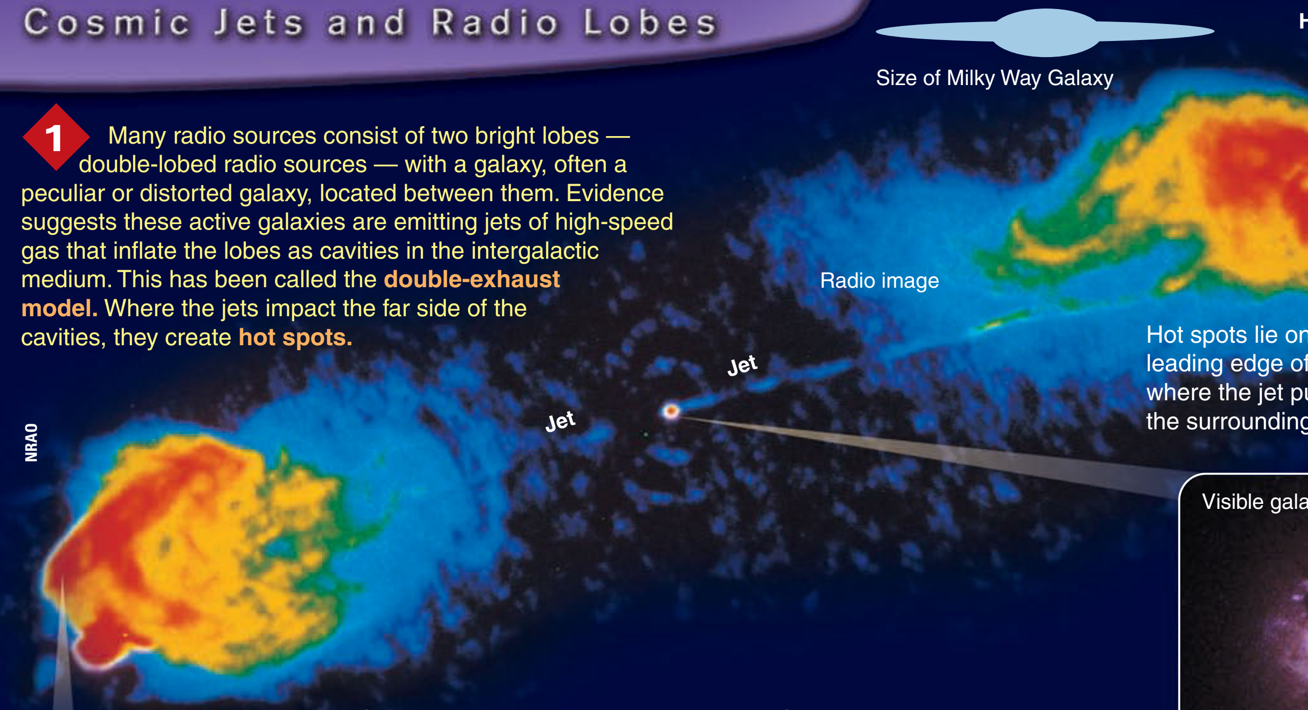


■ **Figure 14-3**

Active galaxy NGC 1275, also known as radio source Perseus A, lies in the Perseus galaxy cluster and is spewing out jets and streamers of high-energy particles that are inflating low-density cavities in the hot gas. Galaxy clusters commonly contain such hot gas, and active galaxies may be the main source of energy that heats the gas. The entire galaxy cluster is roughly 50 times larger in diameter than this image. (NASA/CXC/IoA/A. Fabian et al.)

Cosmic Jets and Radio Lobes

1 Many radio sources consist of two bright lobes — double-lobed radio sources — with a galaxy, often a peculiar or distorted galaxy, located between them. Evidence suggests these active galaxies are emitting jets of high-speed gas that inflate the lobes as cavities in the intergalactic medium. This has been called the **double-exhaust model.** Where the jets impact the far side of the cavities, they create **hot spots.**



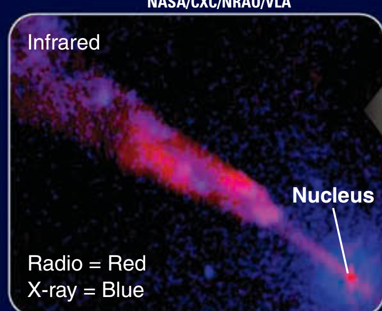
NRAO

Used with permission, Fosbury R. A. E., Vernet, J., Villar-Martin, M., Cohen, M. H., Ogle, P. M., Tran, H. D. & Hook, R. N. 1998, Optical continuum structure of Cygnus A. "KNAW colloquium on: The most distant radio galaxies," Amsterdam, 15–17 October 1997, Roettgering, H., Best, P., and Lehnert, M. eds (Reidel).

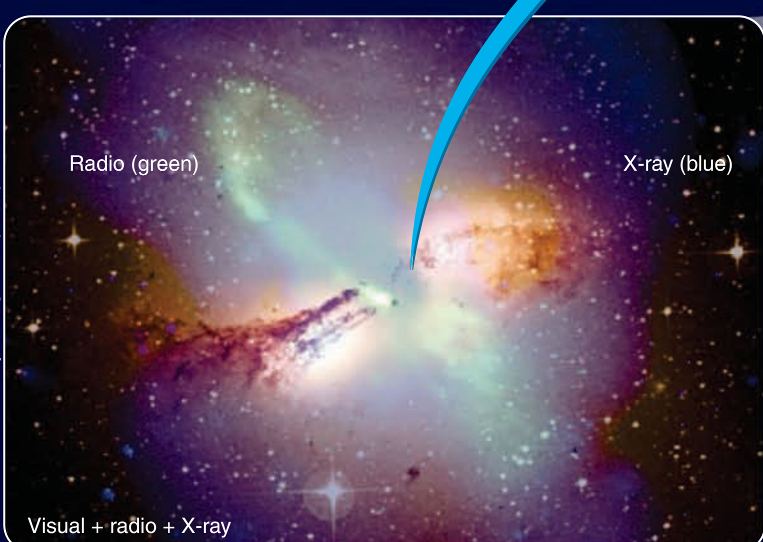
1a Cygnus A, the brightest radio source in Cygnus, is a pair of lobes with jets leading from the nucleus of a highly disturbed galaxy. In this false-color image, the areas of strongest radio signals are shown in red and the weakest in blue. Because the radio energy detected is synchrotron radiation, astronomers conclude that the jets and lobes contain very-high-speed electrons, usually called relativistic electrons, spiraling through magnetic fields about 1000 times weaker than Earth's field. The total energy in a radio lobe is about 10^{53} J — what you would get if you turned the mass of a million suns directly into energy.

2 The radio galaxy NGC 5128 lies between two radio lobes, and, like many active galaxies, is strangely distorted. The dust ring rotates about an axis perpendicular to the ring, but the spherical cloud of stars rotates about an axis that lies in the plane of the ring. NGC 5128 appears to be two galaxies, a giant elliptical and a spiral, passing through each other. This has triggered multiple eruptions. An earlier eruption has produced a large outer pair of lobes, and a more recent eruption has produced an inner pair.

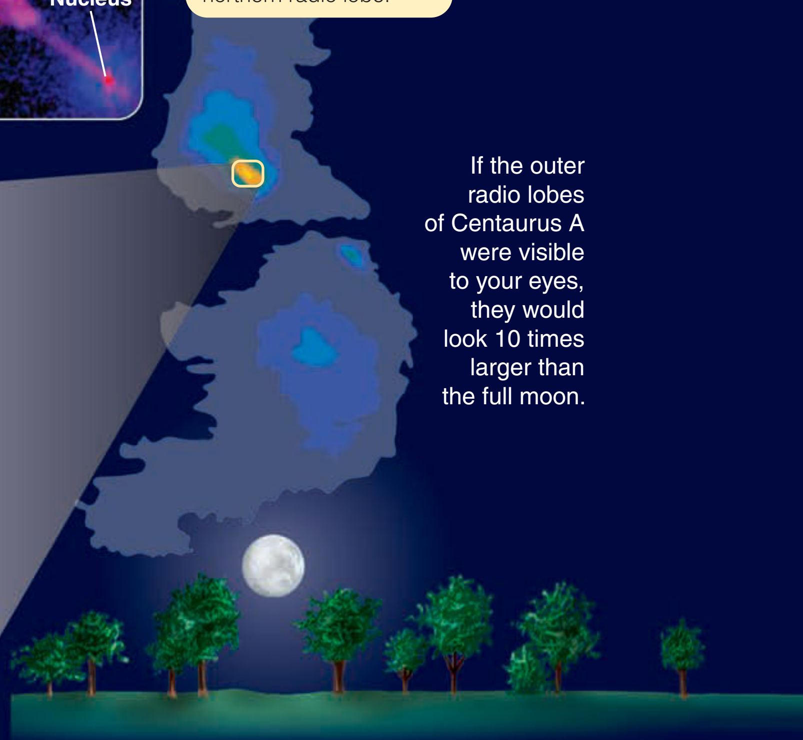
NASA/CXC/NRAO/VLA



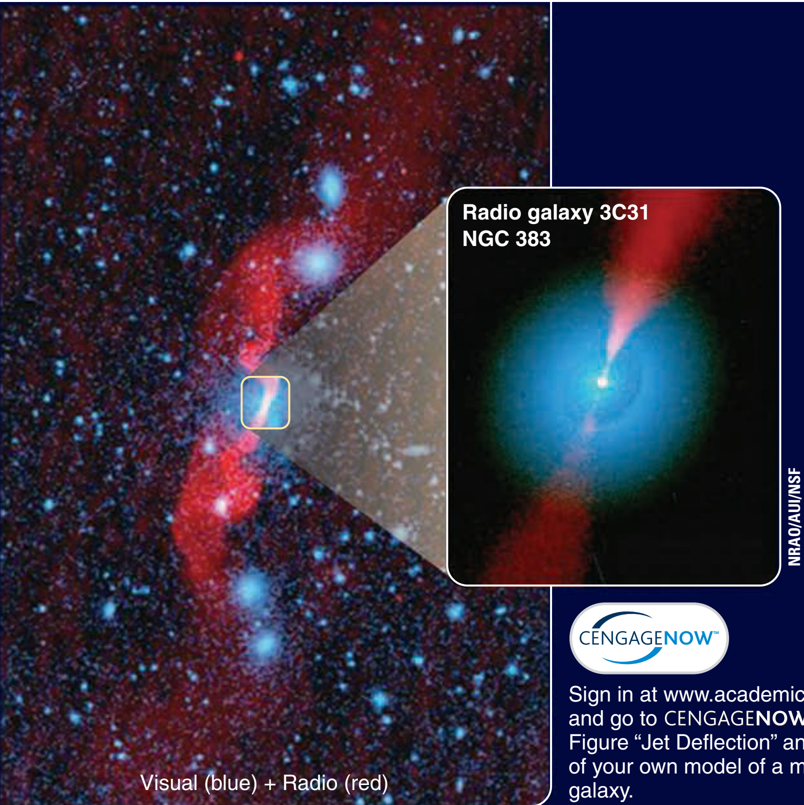
The combined radio and X-ray image at the left shows a high-energy jet at the very center of the galaxy pointing to the upper left into the northern radio lobe.



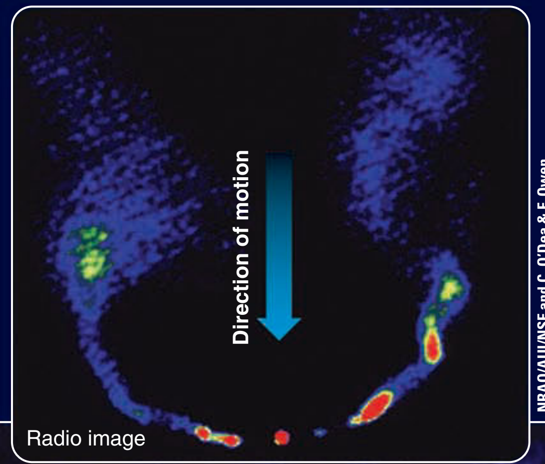
X-ray: NASA/CXC/M. Karovska et al., Radio: NRAO/VLA/Schiminovich, et al., Radio: NRAO/VLA/J. Condon et al., Optical: Digitized Sky Survey U.K. Schmidt Image/STScI.



If the outer radio lobes of Centaurus A were visible to your eyes, they would look 10 times larger than the full moon.



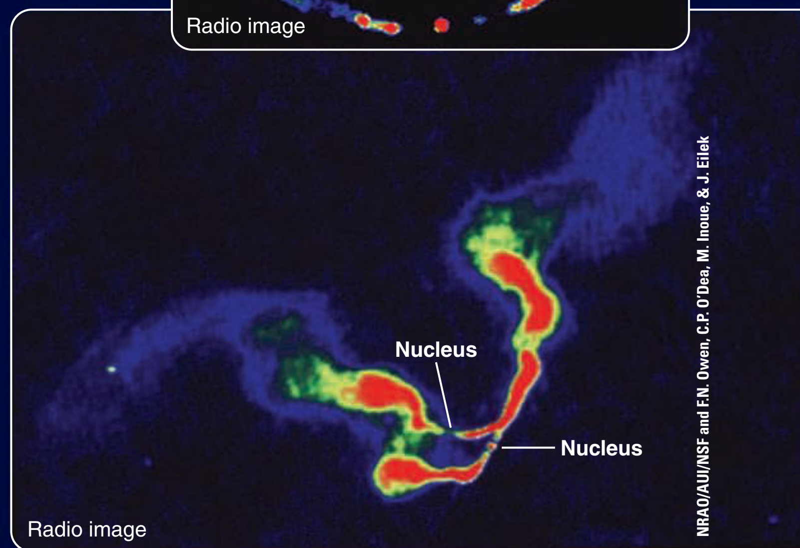
3 The radio jets from NGC 1265 are being left behind as the galaxy moves rapidly through the gas of the intergalactic medium. Twists in the tails are presumably caused by motions of the active nucleus.



Sign in at www.academic.cengage.com and go to CENGAGE**NOW**™ to see Active Figure "Jet Deflection" and take control of your own model of a moving radio galaxy.

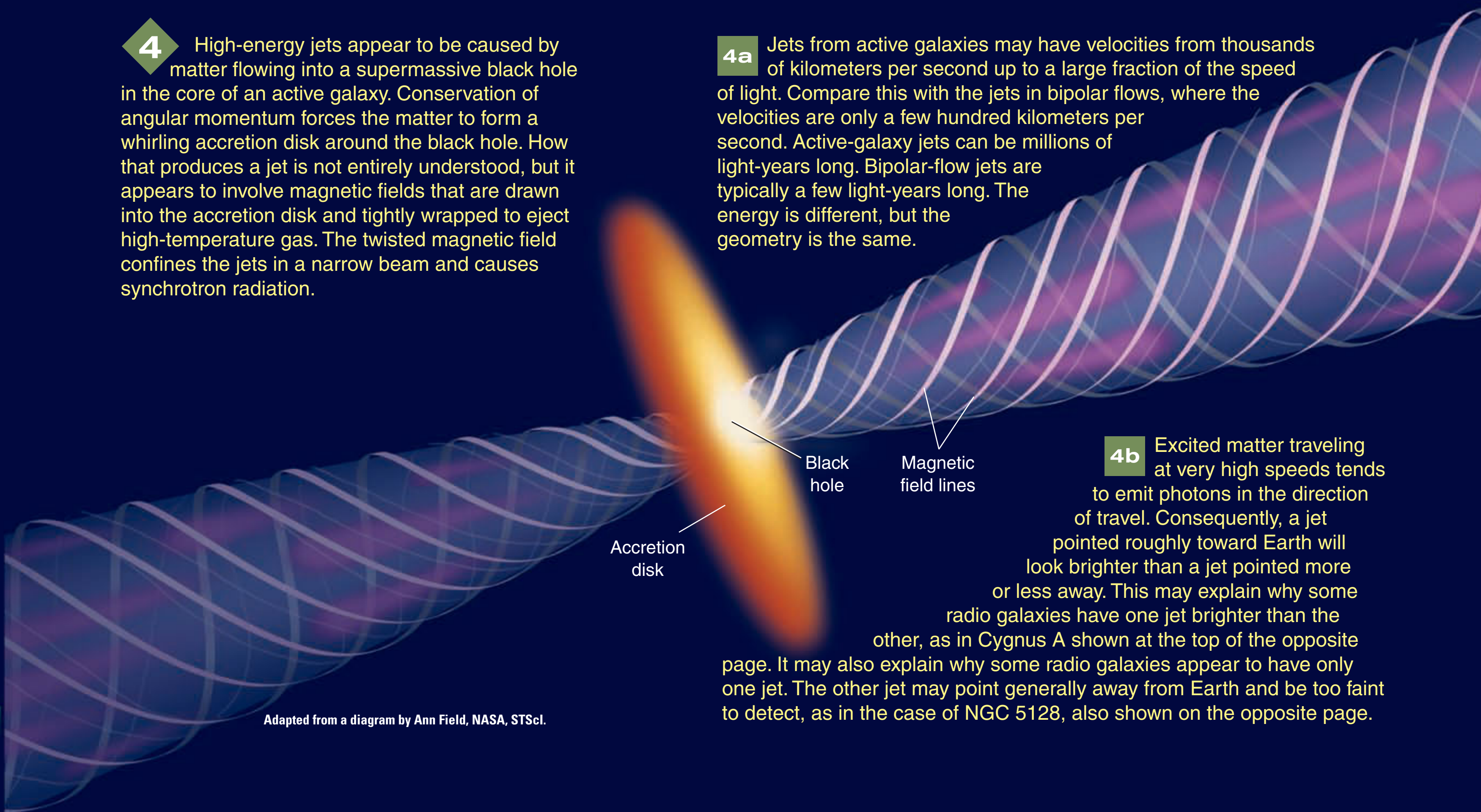
3a Radio galaxy 3C31 is one of a chain of galaxies. It has ejected jets from its core that twist, presumably because the active nucleus is orbiting another object such as the nucleus of a recently absorbed galaxy.

3b The radio source 3C75 is produced by two galaxies experiencing a close encounter. As the active nuclei whip around each other, their jets twist and turn. The size of the visible galaxies would be about the size of cherries at the scale of this image.



4 High-energy jets appear to be caused by matter flowing into a supermassive black hole in the core of an active galaxy. Conservation of angular momentum forces the matter to form a whirling accretion disk around the black hole. How that produces a jet is not entirely understood, but it appears to involve magnetic fields that are drawn into the accretion disk and tightly wrapped to eject high-temperature gas. The twisted magnetic field confines the jets in a narrow beam and causes synchrotron radiation.

4a Jets from active galaxies may have velocities from thousands of kilometers per second up to a large fraction of the speed of light. Compare this with the jets in bipolar flows, where the velocities are only a few hundred kilometers per second. Active-galaxy jets can be millions of light-years long. Bipolar-flow jets are typically a few light-years long. The energy is different, but the geometry is the same.



4b Excited matter traveling at very high speeds tends to emit photons in the direction of travel. Consequently, a jet pointed roughly toward Earth will look brighter than a jet pointed more or less away. This may explain why some radio galaxies have one jet brighter than the other, as in Cygnus A shown at the top of the opposite page. It may also explain why some radio galaxies appear to have only one jet. The other jet may point generally away from Earth and be too faint to detect, as in the case of NGC 5128, also shown on the opposite page.

redshifting the hydrogen Balmer lines to see if they could be made to agree with the lines in 3C273's spectrum. At a redshift of 15.8 percent, three lines clicked into place (■ Figure 14-4). Other quasar spectra quickly yielded to this approach, revealing even larger redshifts.

Numerically the redshift z is the change in wavelength $\Delta\lambda$ divided by the unshifted wavelength λ_o:

$$\text{redshift} = \frac{\Delta\lambda}{\lambda_o}$$

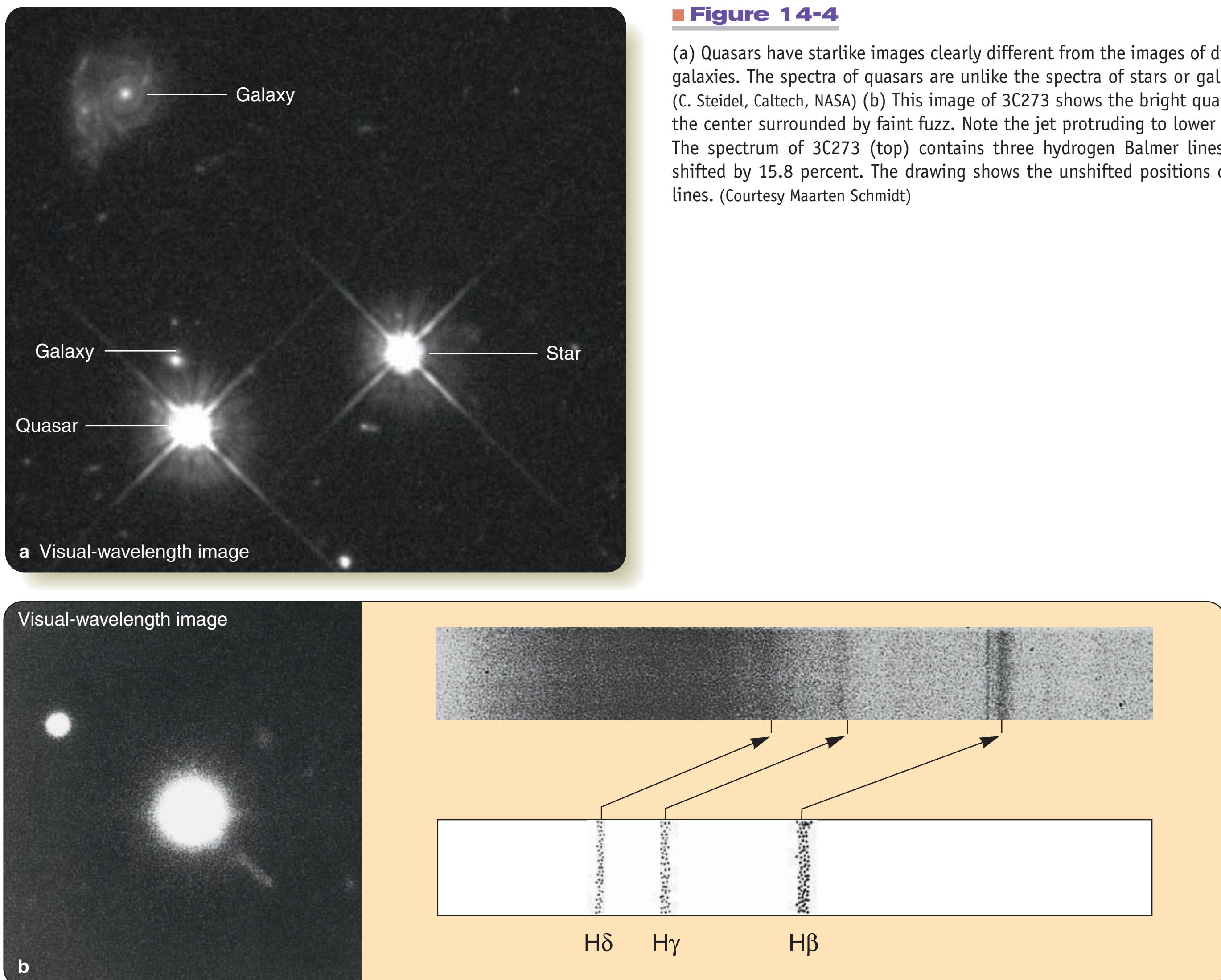
According to the Hubble law, these large redshifts imply very large distances. The first quasars studied were the brighter ones, but surveys have found lots. The Sloan Digital Sky Survey found 90,000. Most of these quasars have very high redshifts and lie very far from Earth.

Many quasar redshifts are greater than one, and that may strike you as impossible. The Doppler formula implies that such objects must have velocities greater than the speed of light. But the Doppler formula you have studied applies only to low velocities. At high velocities, relativistic effects come into play, so quasar redshifts greater than one are indeed possible for very distant quasars.

Although the quasars are far away, they are not very faint. A galaxy at such a distance would be faint and extremely difficult to detect, but quasars show up on photographs as little points of light. If you put the apparent brightnesses and huge distances of quasars into the magnitude-distance relation, you will discover that quasars must have 10 to 1000 times the luminosity of a large galaxy. Quasars must be ultraluminous.

Soon after quasars were discovered, astronomers detected fluctuations in brightness over times as short as a few hours. Because an object cannot change its brightness appreciably in less time than it takes light to cross its diameter, the rapid fluctuations in quasars showed that they are small objects, not more than a few light-hours in diameter.

By the late 1960s astronomers faced a problem: How could quasars be ultraluminous but also very small? What could make



■ **Figure 14-4**

(a) Quasars have starlike images clearly different from the images of distant galaxies. The spectra of quasars are unlike the spectra of stars or galaxies. (C. Steidel, Caltech, NASA) (b) This image of 3C273 shows the bright quasar at the center surrounded by faint fuzz. Note the jet protruding to lower right. The spectrum of 3C273 (top) contains three hydrogen Balmer lines redshifted by 15.8 percent. The drawing shows the unshifted positions of the lines. (Courtesy Maarten Schmidt)

10 to 1000 times more energy than a galaxy in a region as small as our solar system? New, large telescopes in space and on Earth's surface have revealed that quasars are often surrounded by hazy features whose spectra resemble those of normal galaxies. In addition, radio telescopes have revealed that some quasars are ejecting jets and inflating radio lobes. The evidence is now overwhelming that quasars are the active cores of very distant galaxies, and many of those galaxies are distorted or interacting with other galaxies. Quasars have turned out to be the most extreme of active galactic nuclei.

14-2 Supermassive Black Holes

In the previous two chapters, you learned that most galaxies contain supermassive black holes at their centers. How do such objects produce eruptions, and how did they form?

Disks and Jets

Matter flowing inward in a galaxy must conserve angular momentum and spin faster forming a flattened disk around the central black hole. Even a supermassive black hole is quite small. A ten-million-solar-mass black hole would be only one-fifth the diameter of Earth's orbit. That means matter in an accretion disk can get very close to the black hole, orbit very fast, and grow very hot. Theoretical calculations predict that the central cavity in the disk around the black hole is very small but that the disk there is "puffed up" and thick. This means the black hole may be hidden deep inside this central well. The outer part of the disk, according to calculations, is a fat, cool torus (doughnut shape) of dusty gas.

Astronomers can't see the black hole, but in some active galaxies the Hubble Space Telescope can detect the outer parts of the central disks (■ Figure 14-5). Spectra reveal the speed of rotation, and Kepler's third law yields the mass of the central object. Most supermassive black holes have masses of a few million solar masses, but the most massive found contains 18 billion solar masses.

As the matter in the disk drifts inward, it must orbit faster, and friction heats it to high temperatures. The outer disk may be cool, thick and dusty, but the innermost disk is very thin and very hot. Near the black hole, the gas can have temperatures of millions of degrees and emit X-rays before it falls toward the event horizon and vanishes.

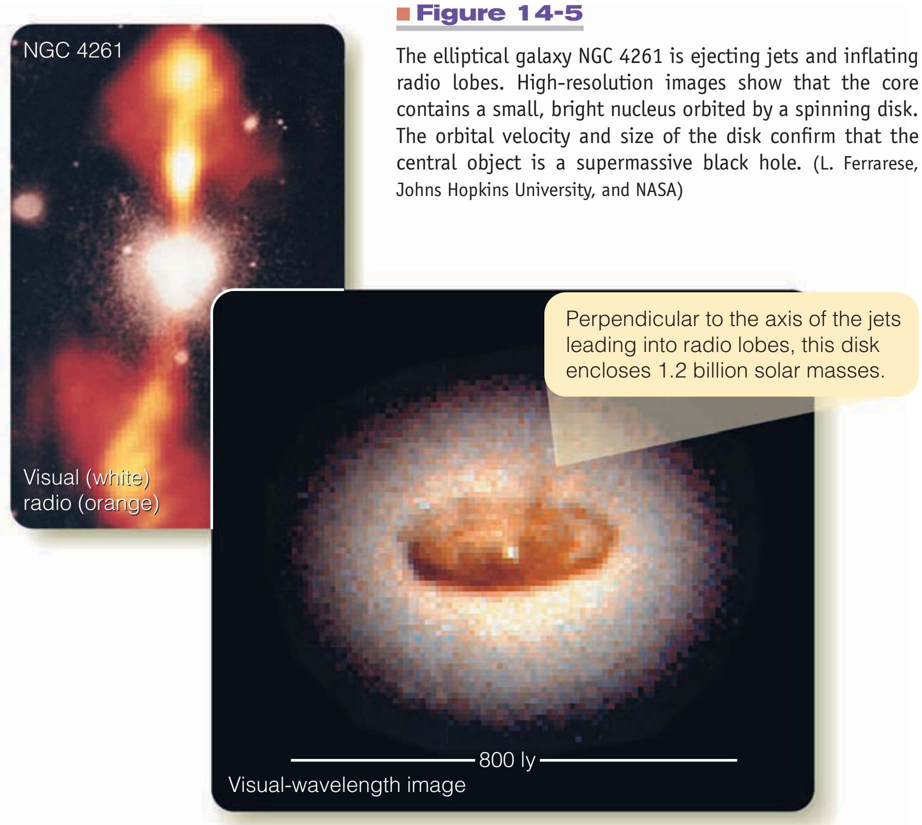
No one knows exactly how the supermassive black hole and its disk produce jets of gas and radiation, but magnetic fields are a factor. Because the disk is at least partially ionized, magnetic fields are trapped in the gas of the disk and drawn inward and wound up. Theorists suggest that this creates powerful magnetic tubes extending along the axis of rotation, channeling hot gas outward in opposite directions. The jets seem to originate very close to the supermassive black hole and are then focused and confined by the enclosing magnetic tubes.

The mechanism that produces jets is understood in only a general way, but astronomers are now trying to work out the details. How do supermassive black holes produce all of the different kinds of active galaxies that are observed?

The Search for a Unified Model

When a field of research is young, scientists find many seemingly different phenomena, such as Seyfert galaxies, double-lobed radio galaxies, quasars, cosmic jets, and so on. As the research matures, the scientists begin seeing similarities and eventually are able to unify the different phenomena as different aspects of a single process. This organization of evidence and theory into logical arguments that explain how nature works is the real goal of science. Astronomers studying active galaxies are now developing a **unified model** of active galaxy cores. A monster black hole is the centerpiece.

According to the unified model, what you see when you view the core of an active galaxy depends on how this accretion



■ **Figure 14-5**

The elliptical galaxy NGC 4261 is ejecting jets and inflating radio lobes. High-resolution images show that the core contains a small, bright nucleus orbited by a spinning disk. The orbital velocity and size of the disk confirm that the central object is a supermassive black hole. (L. Ferrarese, Johns Hopkins University, and NASA) disk is tipped with respect to your line of sight. You should note that the accretion disk may be tipped at a steep angle to the plane of its galaxy, so just because you see a galaxy face-on doesn't mean you are looking at the accretion disk face-on.

If you view the accretion disk from the edge, you cannot see the central area at all because the thick dusty torus blocks your view. Instead, you see radiation emitted by gas lying above and below the central disk. Because this gas is farther from the center, it is cooler, orbits more slowly, and has smaller Doppler shifts and thus narrower spectral lines (■ Figure 14-6). This might account for the type 2 Seyfert galaxies.

If the accretion disk is tipped slightly, you may be able to see some of the intensely hot gas in the central cavity. This broad line region emits broad spectral lines because the hot gas is orbiting at high velocities and the high Doppler shifts smear out the lines. Type 1 Seyfert galaxies may be explained by this phenomenon.

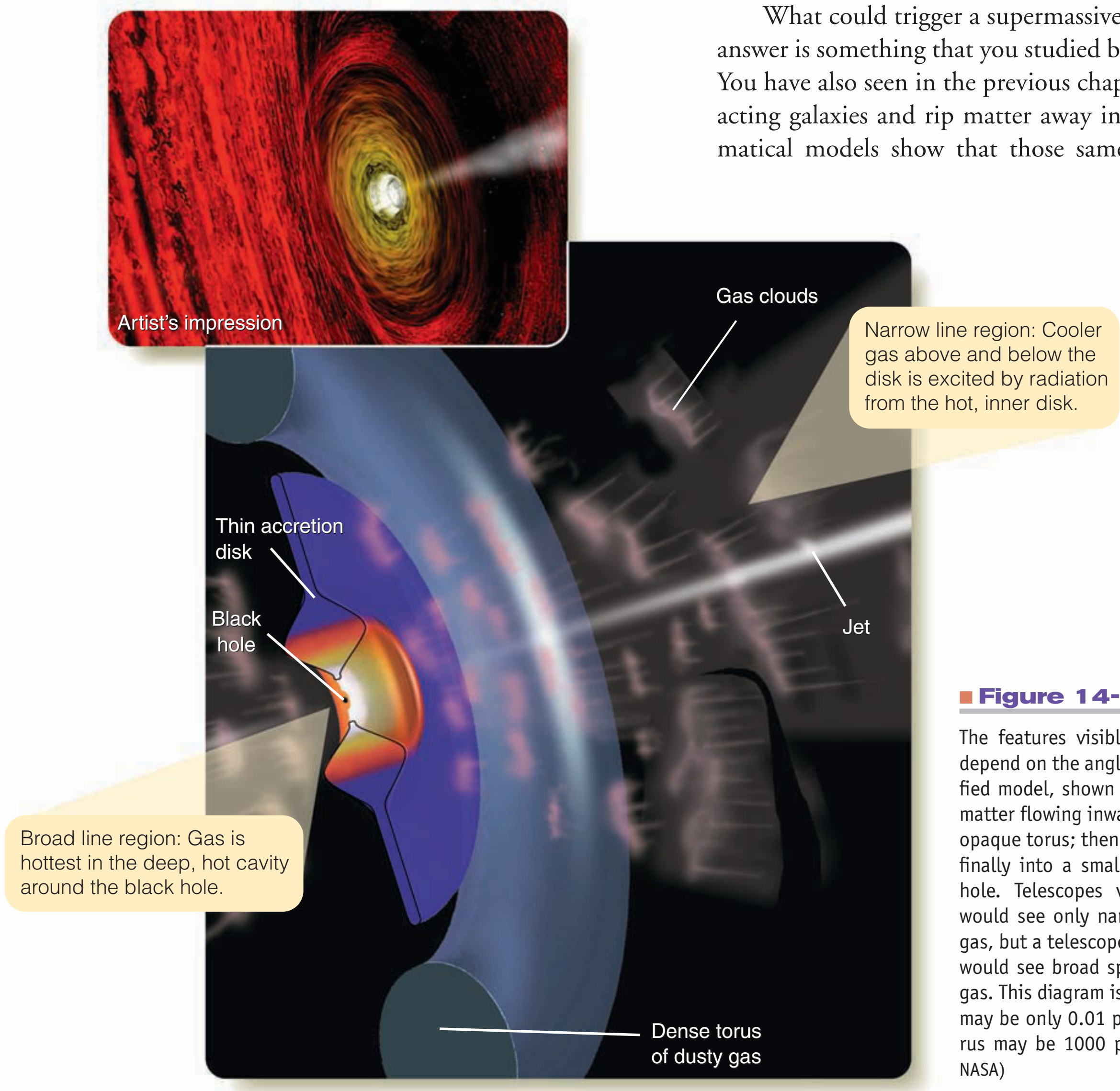
What happens if you look directly into the central cavity? According to the unified model, you would be looking directly into the high energy jet—down the dragon's throat. A few such objects have been found; they have a featureless spectrum and fluctuate in brightness in only hours.

The unified model is far from complete. The actual structure of accretion disks is poorly understood, as is the process by which the disks produce jets. Unification does not explain all of the differences among active galaxies. Rather, it is a model that provides some clues to what is happening in the cores of active galaxies.

The Origins of Supermassive Black Holes

Most galaxies contain a supermassive black hole at their centers, but only a few percent of galaxies have active galactic nuclei. That must mean that most of the supermassive black holes are dormant. They are sleeping. Presumably they are not being fed large amounts of matter. A slow trickle of matter flowing into the supermassive black hole at the center of our galaxy could explain the mild activity seen there. But it would take a larger meal to trigger an eruption such as those seen in active galactic nuclei.

What could trigger a supermassive black hole to erupt? The answer is something that you studied back in Chapter 4—tides. You have also seen in the previous chapter how tides twist interacting galaxies and rip matter away into tidal tails, but mathematical models show that those same interactions can throw



■ **Figure 14-6**

The features visible in the spectrum of an AGN depend on the angle at which it is viewed. The unified model, shown in cross section, suggests that matter flowing inward passes first through a large, opaque torus; then into a thinner, hotter disk; and finally into a small, hot cavity around the black hole. Telescopes viewing such a disk edge-on would see only narrow spectral lines from cooler gas, but a telescope looking into the central cavity would see broad spectral lines formed by the hot gas. This diagram is not to scale. The central cavity may be only 0.01 pc in radius, while the outer torus may be 1000 pc in radius. (Artist's impression: NASA) matter inward. A sudden flood of matter flowing into the accretion disk around a supermassive black hole would trigger it into eruption. This explains why active galaxies are often distorted; they have been twisted by tidal forces as they interacted or merged with another galaxy. A quasar has been found at the center of an elliptical galaxy that is surrounded by shells of stars. In the previous chapter you learned that such shells are evidence of a recent encounter with another galaxy. Some active galaxies have nearby companions, and you can suspect that the companions are guilty of tidally distorting the other galaxy and triggering an eruption.

Tidal forces between galaxies span distances of 100,000 ly or more, but the same bit of physics becomes important when matter comes very close to a supermassive black hole. ■ Figure 14-7 shows how a passing star would be ripped apart and, at least partially, consumed by the black hole. Inflowing gas, dust, and an occasional star would be an energy feast for a supermassive black hole.

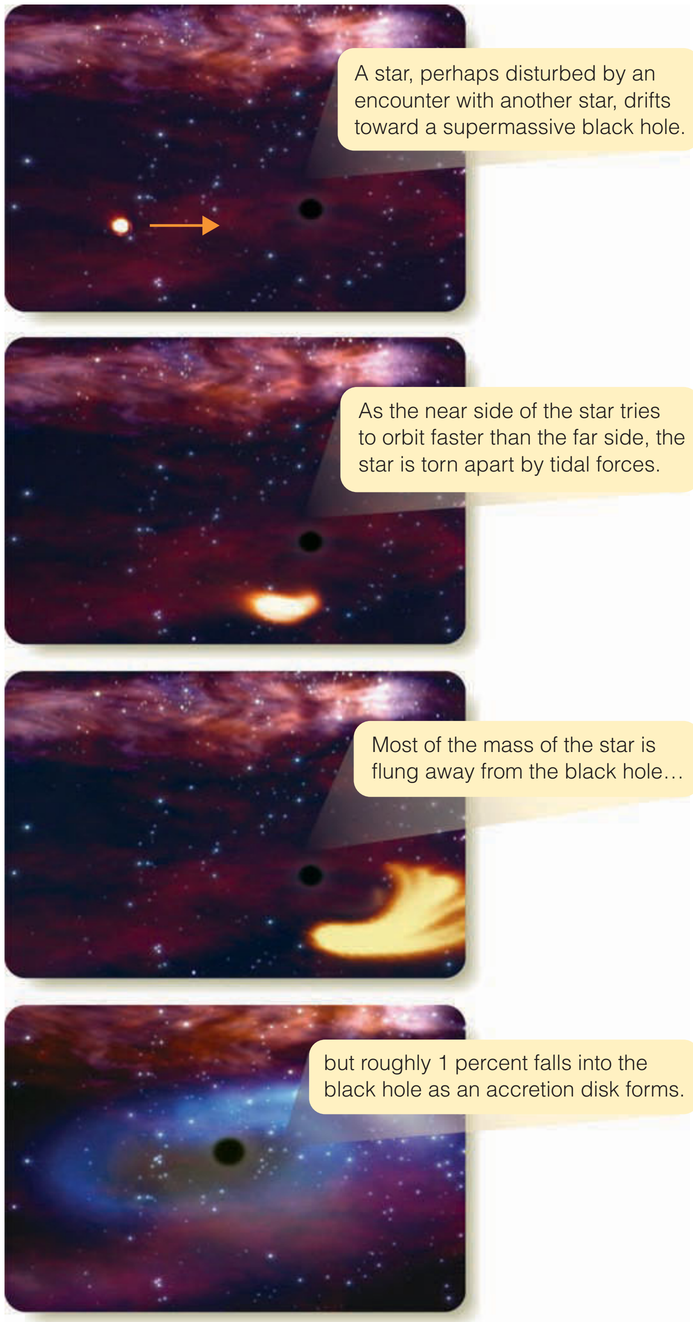
A few dozen supermassive black holes have measured masses, and their masses are correlated with the masses of the host galaxies' nuclear bulges. In each case, the mass of the black hole is about 0.5 percent the mass of the surrounding nuclear bulge. But there is no relationship between the masses of the black holes and the masses of the disks of galaxies. This bit of statistics provides an exciting insight into how galaxies form.

Apparently a galaxy forms its nuclear bulge first and its disk later. As the nuclear bulge forms, a small fraction of the mass, lacking orbital momentum, sinks to the middle where it forms a supermassive black hole. All of that matter flowing together to form the black hole would release a tremendous amount of energy and trigger a violent eruption. Long ago, when galaxies were actively forming, the birth of the nuclear bulges must have triggered active galactic nuclei. The formation of a nuclear bulge was evidently a violent process.

Recall from Chapter 12 that the disk of the Milky Way Galaxy formed late as matter settled into the galaxy. By that stage, the nuclear bulge and central black hole were formed, and the gradual development of the disk didn't trigger a violent eruption. "Disk formation is wimpy," said one astronomer.

If supermassive black holes formed with the nuclear bulges, then the black holes have been in galaxies since the beginning, and it should be possible to trace their history.

Star Falling into a Black Hole



■ **Figure 14-7**

Orbiting X-ray telescopes observing active galaxies sometimes detect X-ray flares equaling the energy of a supernova explosion. Such flares are evidently caused when a star wanders too close to the supermassive black hole at the center of the galaxy and tidal forces rip the star apart. (ESA)

Supermassive Black Holes Through Time

When you look at a photo of galaxies with large redshifts, the look-back time is large, and you see the universe as it was long ago. The light journeying from such a great distance carries information about how the galaxies formed and developed.

In the next chapter, you will see evidence that the universe began 13.7 billion years ago and that it has been expanding ever since. The first clouds of gas began forming stars and falling together to form great star clouds that became the nuclear bulges of galaxies, and the supermassive black holes formed at the same time. Matter flooding into these black holes should have triggered powerful outbursts.

Some of the most distant quasars could be caused by the formation of supermassive black holes, but these extremely distant objects are difficult to image with existing telescopes. Most of the active galaxies that astronomers see with today's telescopes have been triggered into eruption by the interaction and collision of galaxies, a process that throws matter into the central black holes. You should also note that galaxies were closer together when the universe was young and had not expanded very much. Because they were closer together, the forming galaxies collided more often. Also, as small galaxies fell inward and formed halos around the nuclear bulges, matter could have been fed inward to the supermassive black holes and triggered more eruptions.

Quasars are most common with redshifts of a little over 2 and less common with redshifts above 2.7. The largest quasar redshifts are over 6, and such high-redshift quasars are quite rare. Evidently astronomers see few quasars at high redshifts because they are looking back to an age when the universe was so young it had not yet formed many galaxies. At redshifts between about 2 and 2.7, galaxies were actively growing, colliding, and merging (■ Figure 14-8), and quasars were about 1000 times more common than they are now. Nevertheless, even during the age of quasars, quasar eruptions must have been unusual. At any one time, only a small fraction of galaxies had quasars erupting in their cores. As expansion carried the galaxies apart, and as galaxy formation became less active, quasars became less common.

Then where are all the dead quasars? An astronomer commented recently, "There is no way to get rid of supermassive black holes, so all of the galaxies that had short-lived quasars must still have those supermassive black holes." Where are all those dead quasars today? You know where to look—the cores of galaxies.

Most galaxies contain supermassive black holes at their centers, but they have gobbled up most of the nearby gas, dust, and stars, and they are dormant. Our own Milky Way Galaxy could have had a quasar at its core long ago, but now its central black hole is on a strict diet. Occasional interactions between galaxies can throw matter inward and nudge a sleeping supermassive black hole into eruption to become a Seyfert galaxy or a double-lobed radio galaxy.

Active galaxies are not some rare kind of galaxy. They are normal galaxies passing through a stage that many galaxies experience. They are not really peculiar. They are an important part of the story of the formation and evolution of galaxies.

◄ SCIENTIFIC ARGUMENT ►

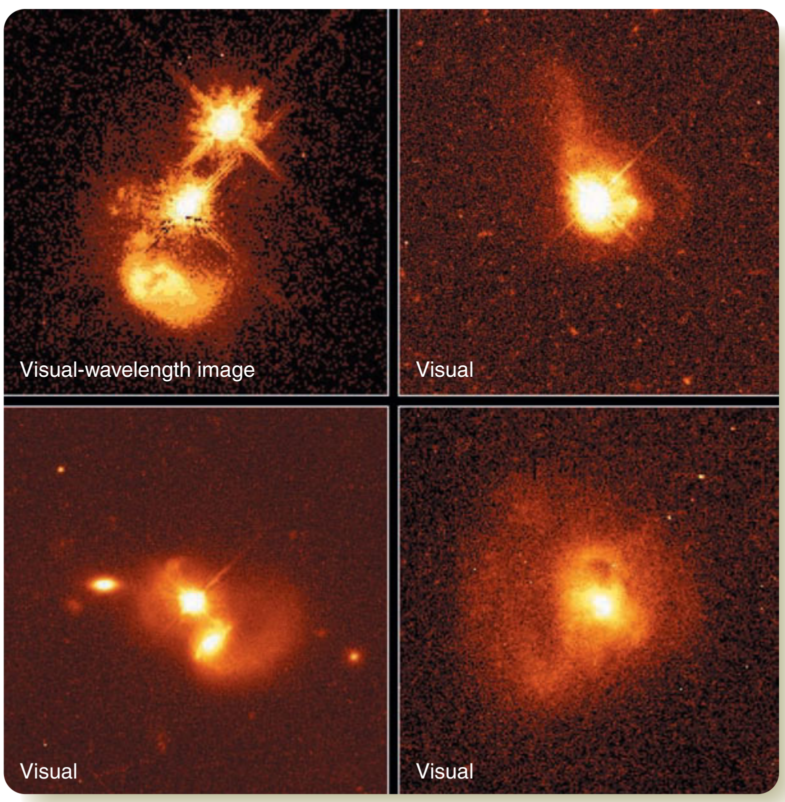
Why are most quasars so far away?

Sometimes a scientific argument hinges on simple geometry, but in this case, you need to add a second factor. First, quasars are the active cores of galaxies, and only a small percentage of galaxies contain active cores. To sample a large number of galaxies in your search, you must extend your search to great distances. Most galaxies lie far away because the amount of space searched increases rapidly with distance. Just as most seats in a baseball stadium are far from home plate, most galaxies are far from our Milky Way Galaxy. Consequently, most of the galaxies that might contain quasars lie at great distances.

But a second factor is much more important. The farther you look into space, the farther back in time you look. It seems that there was a time in the distant past when quasars were more common, and consequently you see most of those quasars at large look-back times, meaning at large distances. For these two reasons, most quasars lie at great distances.

Now expand your argument. **What observational evidence makes you think that quasars must be triggered into eruption?**

◄ ►



■ **Figure 14-8**

The bright object at the center of each of these images is a quasar. Fainter objects around and near the quasars are galaxies distorted by collisions and tidal interactions. (J. Bahcall, Institute for Advanced Study, Mike Disney, University of Wales, NASA)

What Are We? Changed

Next time you are at your local shopping mall, glance at the people around you. How many of them, do you suppose, know that galaxies can erupt or that there was an age of quasars? The vast majority of people have no idea how their lives fit into the story of the universe. Most people don't know what they are. They eat pizza and watch TV without understanding that they are part of a universe in which galaxy collisions trigger supermassive black holes to erupt in titanic explosions.

Astronomy is changing you. As you learn more about stars and galaxies and quasars, you are learning more about yourself and your connection with nature. "Perspective" can mean a view of things in their true relationships. As you study astronomy, you are gaining perspective. Our galaxy, our sun, our planet, and the local shopping mall take on new meaning when you think astronomically.

Study and Review

Summary

- First called **radio galaxies (p. 283)**, **active galaxies (p. 283)** are now known to emit energy at many wavelengths. Because the energy comes from their cores, they are known as **active galactic nuclei (AGN) (p. 283)**.
- **Seyfert galaxies (p. 283)** are spirals with small, highly luminous cores and spectra that show the nuclei contain highly excited gas.
- **Double-lobed radio galaxies (p. 285)** emit radio energy from areas on either side of the galaxies. As described by the **double-exhaust model (p. 286)**, these lobes appear to be inflated by jets ejected from the nuclei of the galaxies. Where the jets push into surrounding gas, they form **hot spots (p. 286)**.
- Jets from active galaxies are at least in part responsible for heating the gas trapped in clusters of galaxies.
- The **quasars (p. 285)** have a starlike appearance but very large redshifts that imply they are very distant. To be visible at such distances, quasars must be ultraluminous. Yet rapid fluctuations in brightness show that quasars must be small.
- The best images show that quasars are embedded in fuzzy galaxies that are often distorted or have close companions. From this, astronomers conclude that quasars are the highly luminous cores of very distant active galaxies.
- Matter flowing into a supermassive black hole must conserve angular momentum and form an accretion disk that is thick and dusty at great distance but thin and hot closer to the black hole. A few such disks have been imaged in the cores of active galaxies, and their rotations permit the measurement of the mass of the black hole.
- The spinning disk pulls in and winds up the magnetic field, and by a process not yet understood, this magnetic field ejects and focuses two jets of radiation and high-speed gas in opposite directions along the axis of rotation.
- Supermassive black holes erupt only when large amounts of matter flow inward, so most galaxies have dormant cores. During interactions and collisions, tides can throw matter inward and trigger eruptions. This explains why active galaxies are often distorted or have nearby companions.
- Because quasars lie at great distances, astronomers see them as they were long ago—over 10 billion years ago—when the universe was young and just forming galaxies.
- According to the **unified model (p. 289)**, what you see depends on the tilt of the accretion disk. If you see into the core, you see broad spectral lines. If you see the disk edge-on, you see only narrow spectral lines. If the jet from the black hole points directly at you, you see a featureless spectrum.
- Because the mass of a supermassive black hole is related to the mass of the nuclear bulge around it, astronomers conclude that the black holes formed soon after the beginning of the universe when the first gas clouds fell together to form the nuclear bulges.
- Mergers with smaller galaxies are thought to have built the halos of galaxies and may have triggered eruptions of the cores, but gas falling in to form the disks was not violent and didn't trigger eruptions.
- Some of the most distant quasars may be erupting because of the formation of supermassive black holes, but most activity is triggered by interactions and collisions with other galaxies. Collisions were more common in the past before the universe had expanded very far.
- The age of quasars at redshifts from about 2 to 2.7 was a time when galaxies were actively growing and merging.
- The supermassive black holes that produced quasar eruptions are now mostly dormant because very little mass is flowing into them. A galaxy can be triggered to become an active galaxy if tides throw matter inward during an interaction with another galaxy.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What is the difference between the terms radio galaxy and active galaxy?
2. What statistical evidence suggests that Seyfert galaxies have suffered recent interactions with other galaxies?
3. What evidence shows that the energy source in a double-lobed radio galaxy lies at the center of the galaxy?
4. How does the peculiar rotation of NGC 5128 help explain the origin of this active galaxy?
5. What evidence shows that quasars are ultraluminous but must be very small?
6. What evidence is there that quasars occur in distant galaxies?
7. How does the unified model explain the two kinds of Seyfert galaxies?

8. Why are there few quasars at low redshifts and at high redshifts but many at redshifts between about 2 and 2.7?
9. Why did galaxies collide more often in the distant past than they do now?
10. **How Do We Know?** How would you respond to someone who said, "Oh, that's only statistics"?

Discussion Questions

1. Why do quasars, active galaxies, SS433, and protostars have similar geometry?
2. By custom, astronomers refer to the unified model of AGN and not to the unified hypothesis or unified theory. In your opinion, which of the words seems best?
3. Do you think that our galaxy has ever been an active galaxy? Could it have hosted a quasar when it was young?

Problems

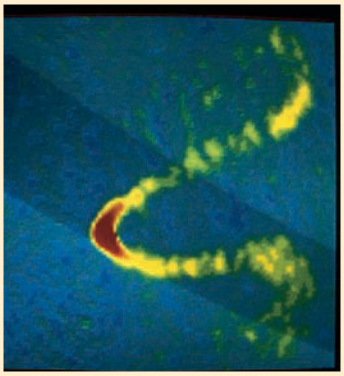
1. The total energy stored in a radio lobe is about 10^{53} J. How many solar masses would have to be converted to energy to produce this energy? (*Hints:* Use $E = mc^2$. One solar mass equals 2×10^{30} kg.)
2. If the jet in NGC 5128 is traveling at 5000 km/s and is 40 kpc long, how long will it take for gas to travel from the core of the galaxy to the end of the jet? (*Hint:* 1 pc equals 3×10^{13} km.)
3. Cygnus A is roughly 225 Mpc away, and its jet is about 50 arc seconds long. What is the length of the jet in parsecs? (*Hint:* See Reasoning with Numbers 3-1.)
4. What is the linear diameter of a radio source with an angular diameter of 0.0015 arc seconds and a distance of 3.25 Mpc. (*Hint:* See Reasoning with Numbers 3-1.)
5. If the active core of a galaxy contains a black hole of 10^6 solar masses, what will the orbital period be for matter orbiting the black hole at a distance of 0.33 AU? (*Hint:* See Reasoning with Numbers 8-4.)
6. If a quasar is 1000 times more luminous than an entire galaxy, what is the absolute magnitude of such a quasar? (*Hint:* The absolute magnitude of a bright galaxy is about −21.)
7. If the quasar in Problem 6 were located at the center of our galaxy, what would its apparent magnitude be? (*Hints:* See Reasoning with Numbers 8-2 and ignore dimming by dust clouds.)
8. If the Hubble constant is 70 and a quasar has an apparent velocity or recession of 45,000 km/s, how far away is it?
9. The hydrogen Balmer line H_β has a wavelength of 486.1 nm. It is shifted to 563.9 nm in the spectrum of 3C273. What is the redshift of this quasar? (*Hint:* What is $\Delta\lambda$?)

Learning to Look

1. The image at right combines visual (blue) with radio (red) to show the galaxy radio astronomers call Fornax A. Explain the features of this image. Is it significant that the object is a distorted elliptical galaxy in a cluster?

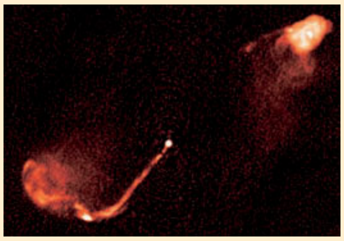
NRAO/AUI

2. Explain the features of this radio image of the galaxy IC708.



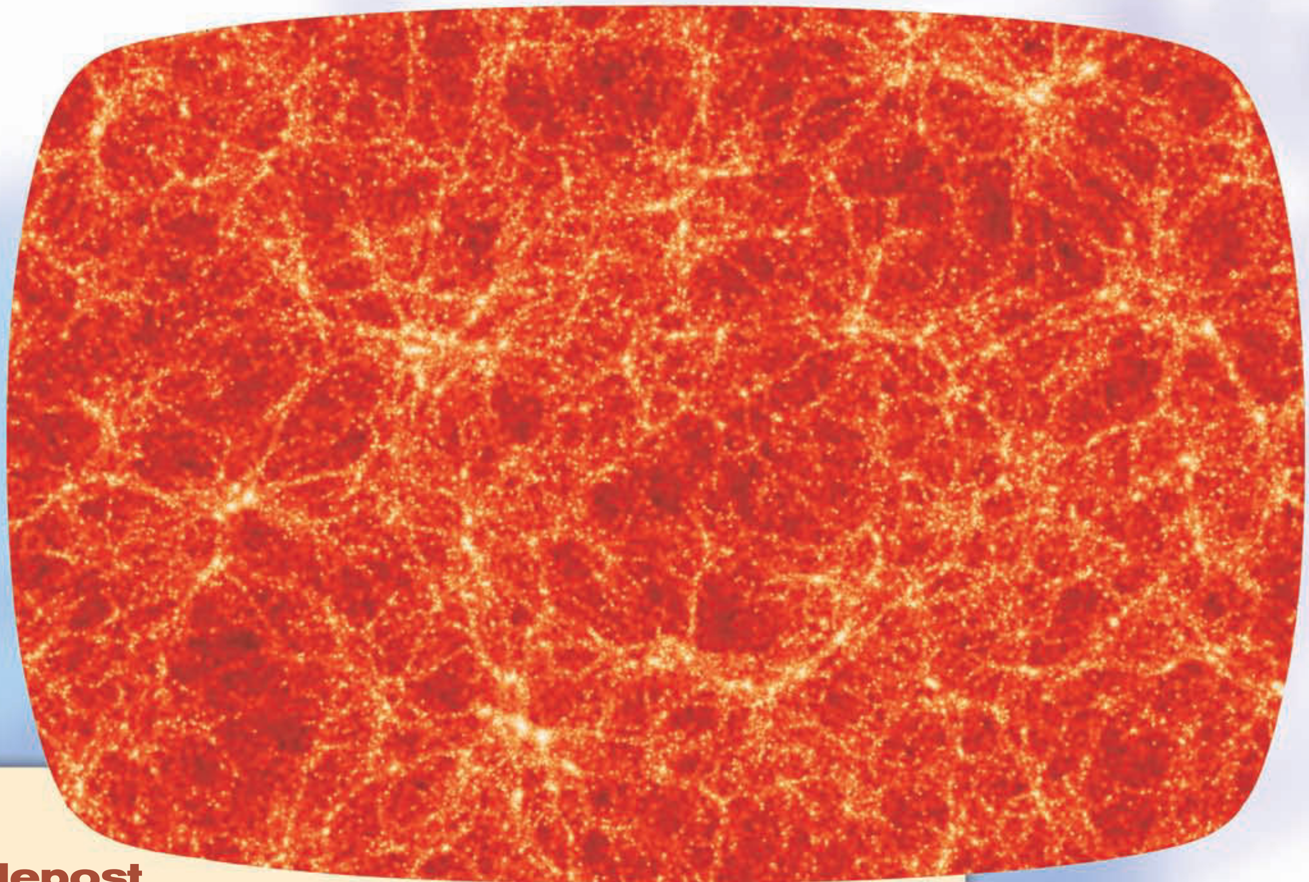
NRAO/AUI

3. A radio image of quasar 3C334 is shown at the right. Why do you see only one jet? What does it mean that this looks so much like a radio image of a double-lobed radio galaxy?



NRAO/AUI

15 Modern Cosmology



All of the energy and matter in the universe, including the matter in your body, began in the big bang. In this computer model of the evolution of the universe soon after the big bang, matter is being drawn together to form great clouds of galaxies. (MPA and Jörg Colberg)

Guidepost

Since Chapter 1, you have been on an outward journey through the universe. Now you have reached the limit of your travels in space and time and can study the universe as a whole. The ideas in this chapter are among the biggest and most difficult in all of science. Can you imagine a limitless universe, or the first instant of time?

As you explore cosmology, you will find answers to three essential questions:

Does the universe have an edge and a center?

How do you know that the universe began with a big bang?

How has the universe evolved, and what will be its fate?

These questions will lead you to a new understanding of where you are and what you are. Once you have finished this chapter, you will have a modern insight into the nature of the universe, and it will be time to focus back on your local neighborhood—the subject of the rest of this book.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

The Universe, as has been observed before, is an unsettlingly big place, a fact which for the sake of a quiet life most people tend to ignore.

DOUGLAS ADAMS
THE RESTAURANT AT THE END OF THE UNIVERSE

LOOK AT YOUR thumb. The matter in your body was present in the fiery beginning of the universe. **Cosmology,** the study of the universe as a whole, can tell you where your matter came from, and it can tell you where your matter is going.

Cosmology is a mind-bendingly weird subject, and you can enjoy it for its strange ideas. It is fun to think about space stretching like a rubber sheet, invisible energy pushing the universe to expand faster and faster, and the origin of vast walls of galaxy clusters. Notice that this is better than speculation—it is all supported by evidence. Cosmology, however strange it may seem, is a serious and logical attempt to understand the structure and evolution of the entire universe.

This chapter will help you climb the cosmology pyramid (■ Figure 15-1) one step at a time. You already have some ideas about what the universe is like. Start with those, test them against observations, and also compare them with scientific theories. Step-by-step you can build a modern understanding of cosmology. Each step in the pyramid is small, but it leads to some astonishing insights into how the universe works, and how you came to be a part of it.

15-1 Introduction to the Universe

MOST PEOPLE HAVE an impression of the universe as a vast ocean of space filled with stars and galaxies (■ Figure 15-2), but, as you begin exploring the universe, you need to become aware of your expectations so they do not mislead you. The first step is to deal with an expectation so obvious that most people, for the sake of a quiet life, don't think about it.

■ **Figure 15-1**

Climbing the cosmology pyramid step-by-step isn't very difficult, and it leads to some fascinating ideas about the origin and evolution of the universe.

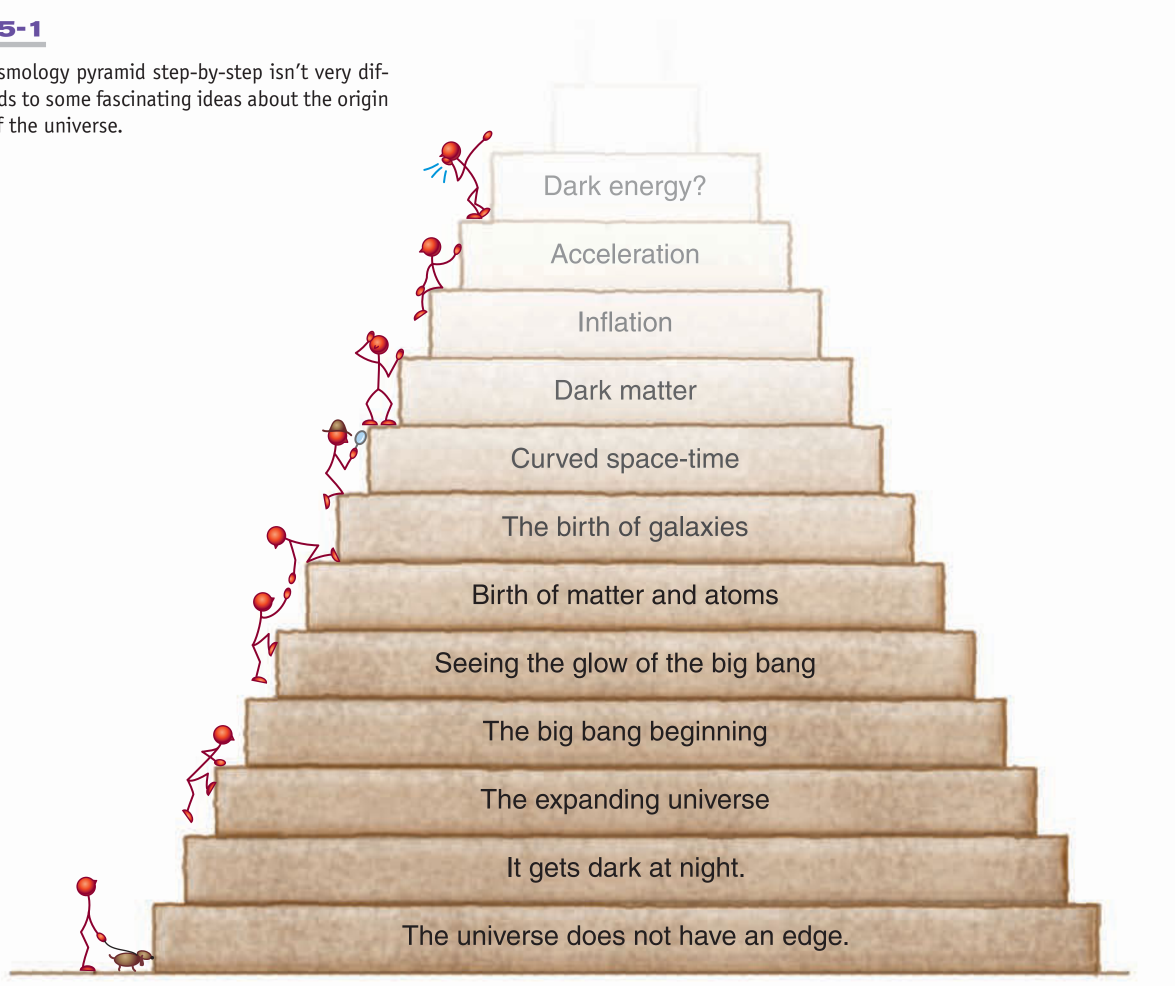


Figure 15-2

The entire sky is filled with galaxies. Some lie in clusters of thousands, and others are isolated in nearly empty voids between the clusters. In this image of a typical spot on the sky, bright objects are nearby stars; their "spikes" are caused by diffraction in the telescope. All other objects are galaxies, ranging from the relatively nearby face-on spiral at upper right to the most distant galaxies, visible only in the infrared, shown as red in this composite image. (R. Williams, STScI HDF-South Team, NASA)

The Edge–Center Problem

In your daily life you are accustomed to boundaries. Rooms have walls, athletic fields have boundary lines, countries have borders, oceans have shores. It is natural to think of the universe also as having an edge, but that idea can't be right.

If the universe had an edge, imagine going to that edge. What would you find there: A wall of some type? A great empty space? Nothing? Even a child can ask: If there is an edge to the universe, what's beyond it? A true edge would have to be more than just an end of the distribution of matter. It would have to be an end of space itself. But, then, what would happen if you tried to reach past, or move past, that edge?

An edge to the universe violates common sense, and modern observations (which you will study later in this chapter) indicate that the universe could be infinite and would therefore have no edge. Perhaps even more important, if the universe has no edge, then it cannot have a center. You find the centers of things—pizzas, football fields, oceans, galaxies—by referring to their edges. If the universe has no edge, then it cannot have a center.

It is a **Common Misconception** to imagine that the universe has a center, but, as you just realized, that is impossible. As you study cosmology, you should take care to avoid thinking that there is a center of the universe.

The Idea of a Beginning

Of course you have noticed that the night sky is dark. That is an important observation because, you may be surprised to learn, seemingly reasonable assumptions about the universe can lead to the conclusion that the night sky actually should glow blindingly bright. This conflict between observation and theory is called **Olbers's paradox** after Heinrich Olbers, a physician and astronomer who publicized the problem in 1826. (*Problem* or *question* might be a more accurate word than *paradox*.) Olbers was not the first person to pose this question; the problem of the dark night sky was first discussed by Thomas Digges in 1576 and was further analyzed by astronomers such as Johannes Kepler in 1610 and Edmund Halley in 1721. Olbers gets the credit through an accident of scholarship on the part of modern scientists who were not aware of the earlier discussions.

The point made by Olbers seems simple. Suppose you assume, as did most scientists in Olbers's time, that the universe is infinite in size, infinite in age, **static** (a fancy word for unchanging overall), and filled with stars. If you look in any direction, your line of sight must eventually run into the surface of a star. (The clumping of stars into galaxies and galaxies into clusters can be shown mathematically to make no difference.) Look at ■ Figure 15-3, which uses the analogy of lines of sight in a forest. (The use of analogies in science is discussed in ■ How Do We Know? 15-1). When you are deep in a forest, every line of sight ends at a tree trunk, and you cannot see out of the forest.

By analogy to the view from inside a forest, every line of sight from Earth into space should eventually end at the surface of a star. Of course, the more distant stars would be fainter than nearby stars because of the inverse square law. However, the farther you look into space, the larger the volume you are viewing and the more stars are included; the two effects cancel out. The result should be that the entire sky should be as bright as the surface of an average star—like suns crowded "shoulder to shoulder," covering the sky from horizon to horizon. It should not get dark at night.

Astronomers and physicists who study cosmology, called **cosmologists,** now believe they understand why the sky is dark. Olbers's paradox makes an incorrect prediction because it is based on incorrect assumptions. The universe may be infinite in size, but it is neither infinitely old nor static. The essence of modern cosmologists' answer to Olbers's question was suggested first by Edgar Allan Poe in 1848. Poe proposed that the night sky is dark because the universe is not infinitely old but came into existence at some finite time in the past. The more distant stars are so far away that light from them has not yet reached Earth. That is, if you look far enough away, the look-back time approaches the age of the universe, and you see to a time before the first stars began to shine. The night sky is dark because the universe had a beginning. You can now answer Olbers's question, and understand why the night sky is dark, by revising your original assumptions about the universe.

How Do We Know? 15-1

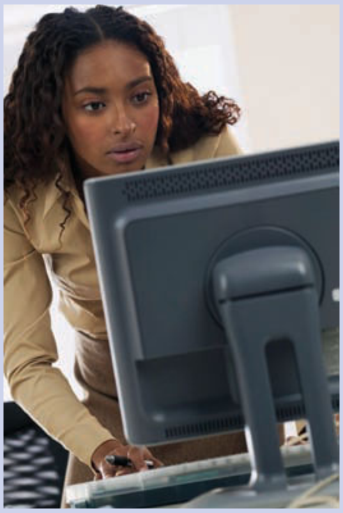
Reasoning by Analogy

How do scientists use analogies? "The economy is overheating, and it may seize up," an economist might say. Economists like to talk in analogies because economics is often abstract, and one of the best ways to think about abstract problems is to find a more approachable analogy. Rather than discussing details of the national economy, you might be able to make conclusions about how the economy works by thinking about how a gasoline engine works.

Much of astronomy is abstract, and cosmology is the most abstract subject in astronomy. Furthermore, cosmology is highly mathematical, and unless you are prepared to learn some difficult mathematics, you instead have to use analogies, such as lines of sight in a forest.

Reasoning by analogy is a powerful technique. An analogy can reveal unexpected insights and lead you to further discoveries. Carrying an analogy too far, however, can be misleading. You might compare the human brain to a computer, and that would help you understand how data flow in and are processed and how new data flow out, but the analogy is flawed. For example, although data in computers are stored in specific locations, memories are stored in the brain in a distributed form. No single brain cell holds a specific memory. So, if you carry the analogy too far, it can mislead you. Whenever you reason using analogies, you should be alert for their limitations.

As you study any science, be alert for analogies. They are tremendously helpful, but you have to be careful not to carry them too far.

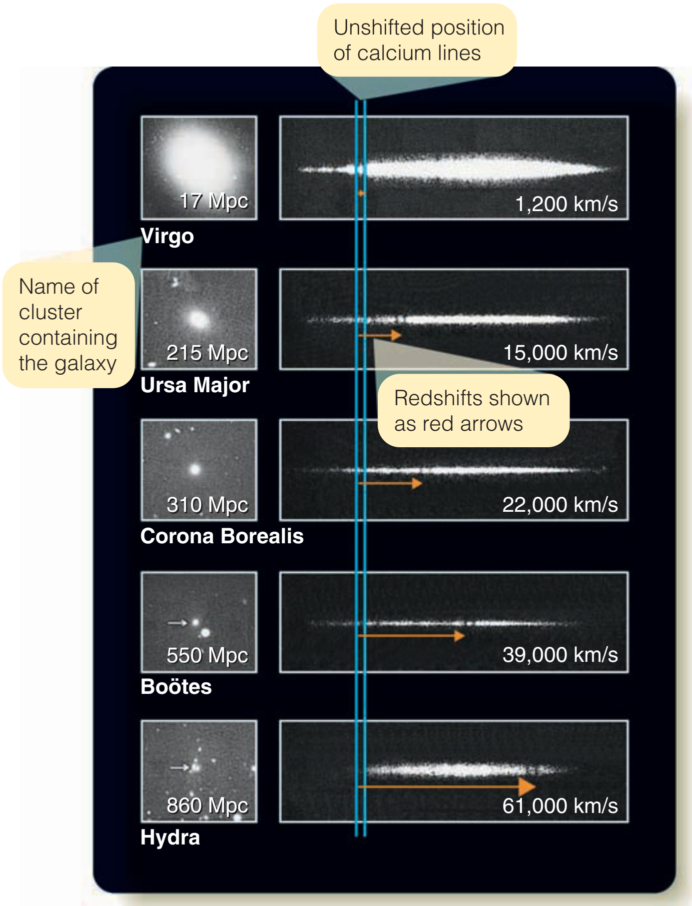


Blend Images/SuperStock

The analogy between a human brain and a computer is useful in some ways but not in others.

Figure 15-3

(a) Every direction you look in a forest eventually reaches a tree trunk, and you cannot see out of the forest. (Photo courtesy Janet Seeds) (b) If the universe is infinite and filled with stars, then any line from Earth will eventually reach the surface of a star. This assumption leads to a prediction that the night sky should glow as brightly as the surface of the average star, a puzzle commonly referred to as Olbers's paradox.



■ Figure 15-4

These galaxy spectra extend from the near-ultraviolet at left to the blue part of the visible spectrum at right. The two dark absorption lines of once-ionized calcium are prominent in the near-ultraviolet. The redshifts in galaxy spectra are expressed here as apparent velocities of recession. Note that the apparent velocity of recession is proportional to distance, which is known as the Hubble law. (Caltech) **Animated!**

The answer to Olbers's question is a powerful idea because it clearly illustrates the difference between the universe and the **observable universe.** The universe is everything that exists, and it could be infinite. The observable universe, in contrast, is the part (maybe a very small part) that you can see from Earth using the most powerful telescopes. You will learn later in this chapter about evidence that the universe is around 14 billion years old. That means you can't observe objects farther away than a look-back time of around 14 billion years. Do not confuse the observable universe, which is finite, with the universe as a whole, which might be infinite.

There is a **Common Misconception** that the universe is unchanging overall and not evolving. In the next section you will discover that the universe is actually changing. Observations reveal an **expanding universe.**

Cosmic Expansion

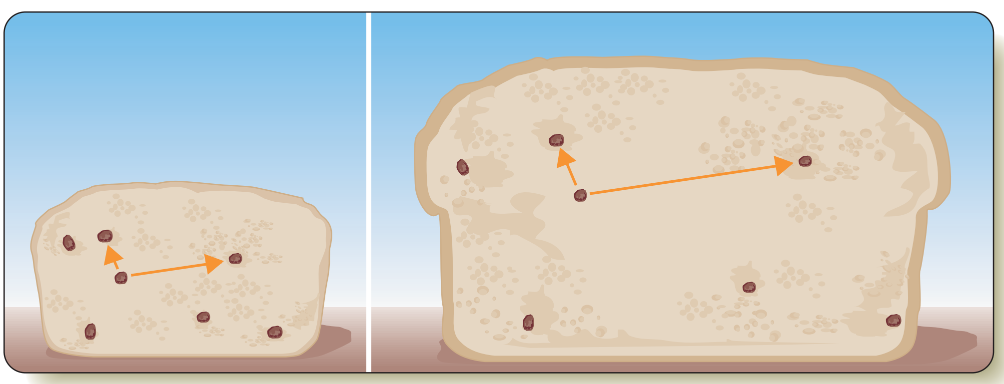
In 1929, Edwin Hubble published his discovery that the sizes of galaxy redshifts are proportional to galaxy distances. Nearby galaxies have small redshifts, and more distant galaxies have large redshifts. You learned this as the Hubble law in Chapter 13, when you used it to estimate the distances to galaxies. Those galaxy redshifts, interpreted as Doppler shifts, imply that galaxies are receding from each other and that the universe is expanding.

■ Figure 15-4 shows spectra of galaxies in galaxy clusters at various distances. The Virgo cluster is relatively nearby, and its redshift is small. The Hydra cluster is very distant, and its redshift is so large that the two dark spectral lines formed by ionized calcium are shifted from near-ultraviolet wavelengths well into the visible part of the spectrum.

The expansion of the universe does not imply that Earth has a special location. To see why, look at ■ Figure 15-5, which shows an analogy to baking raisin bread. As the dough rises, it pushes the raisins away from each other at speeds that are proportional to distance. Two raisins that were originally close to each other are pushed apart slowly, but two raisins that were far apart, having more dough between them, are pushed apart faster. If bacterial astronomers lived on a raisin in your raisin bread, they could observe the redshifts of the other raisins and derive a bacterial Hubble law. They would conclude that their universe was expanding uniformly. It does not matter which raisin the bacterial astronomers lived on, they would get the same Hubble law—no raisin has a special viewpoint. Similarly, astronomers in any galaxy will see the same law of expansion—no galaxy has a special viewpoint.

■ Figure 15-5

An illustration of the raisin-bread analogy for the expansion of the universe. As the dough rises, raisins are pushed apart with velocities proportional to distance. A colony of bacteria living on any raisin will find that the redshifts of the other raisins are proportional to their distances. **Animated!**



When you look at Figure 15-5, you see the edge of the loaf of raisin bread, and you can identify a center to the loaf. The raisin bread analogy to the universe stops working when you consider the crust—the edge—of the bread. Remember that the universe cannot have an edge or a center, so there is no center to the expansion. The raisin bread analogy is useful but also imperfect.

15-2 The Big Bang Theory

Now YOU ARE ready to take a historic step up the cosmology pyramid. The expansion of the universe led cosmologists to conclude that the universe must have begun with an event of cosmic intensity.

Necessity of the Big Bang

Imagine that you have a video of the expanding universe and run it backward. You would see the galaxies moving toward each other. There is no center to the expansion of the universe, so you would not see galaxies approaching a single spot. Rather, you would see the distances between all galaxies decreasing. Eventually, as your video ran back toward the beginning, galaxies would begin to merge. If you ran the video far enough back, you would see the matter and energy of the universe compressed into a high-density, high-temperature state. The expanding universe must have begun from this moment of extreme conditions which cosmologists call the **big bang.**

How long ago did the universe begin? You can estimate the age of the universe with a simple calculation. If you need to drive to a city 100 miles away, and you can travel 50 miles per hour, you divide distance by rate of travel and learn the travel time—in this example, 2 hours. To find the age of the universe, you can divide the distance between two galaxies by the speed with which they are moving away from each other and find out how long they have taken to reach their present separation. The time you calculate by this procedure, an estimate of the universe's age, is called the **Hubble time** and is the same no matter which two galaxies you pick. The details of this simple calculation are given in ■ Reasoning with Numbers 15-1.

The Hubble time is an estimate of the age of the universe. You will fine-tune your estimate later in this chapter, but for the moment you can conclude that basic observations of the universe, especially the recession of the galaxies, require that the universe began with a big bang about 14 billion years ago.

The phrase *big bang* was invented by early critics of that theory, and the label gives a misimpression. Do not think of an edge or a center when you think of the big bang. It is a very **Common Misconception** that the big bang was an explosion and that the galaxies are flying away from the location of that explosion. Instead, try to imagine the entire universe. At the time of the big bang, all of the matter that now composes galaxies, stars, people, and atoms was much closer together. But, like the universe now, the early universe also had no center and no edge. The big bang filled the entire universe, happening everywhere at the same moment as the matter of the universe began to draw apart. A more accurate term might be the *big stretch.*

Although your imagination tries to visualize the big bang as a localized event, you should keep firmly in mind the correct picture that the big bang did not occur at a single place but filled the entire volume of the universe. The matter of which you are made was part of that big bang, so you are inside the remains of that event, and the universe continues to expand around you. You cannot point to any particular place and say, "The big bang occurred over there." This is a brain-straining (stretching?) idea, but you will find more information to help you understand it better later in this chapter when you study the true nature of space and time.

Your instinct also is to think of the big bang as an event that happened long ago and can no longer be observed, like the Gettysburg Address. Amazingly, the look-back time makes it possible to observe the big bang now, directly. The look-back time to nearby galaxies is "only" a few million years; the look-back time to more distant galaxies is a large fraction of the age of the universe. Suppose you look between the distant galaxies, seeing even farther away and farther back in time. You should be able to detect the hot gas that filled the universe long ago, right after the big bang, before the first stars and galaxies formed. The big bang occurred everywhere, and, in whatever direction you look, at great distance you can see back to the age when the universe was filled with hot gas (■ Figure 15-6).

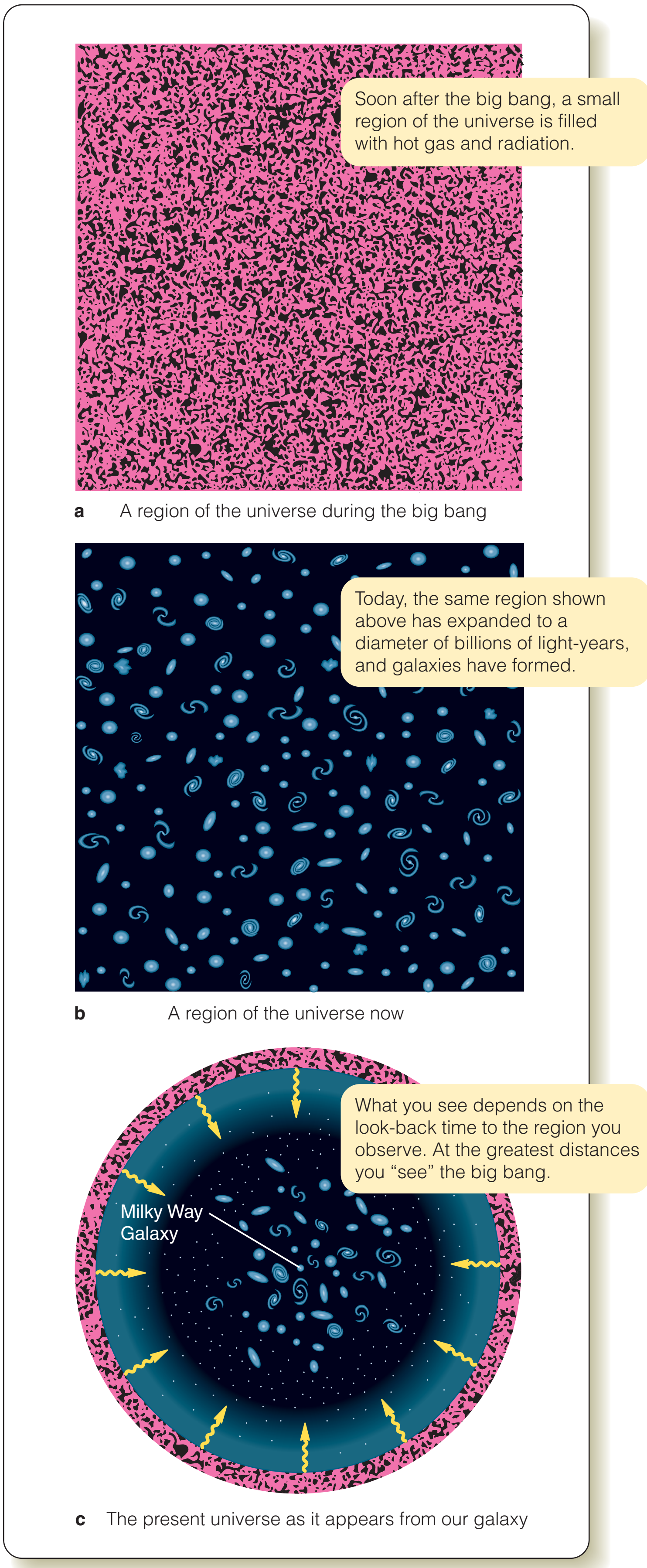
Reasoning with Numbers | 15-1

The Age of the Universe

Dividing the distance to a galaxy by the apparent velocity with which it recedes gives you an estimate of the age of the universe, and the Hubble constant simplifies your task further. The Hubble constant H has the units km/s per Mpc, which is a velocity divided by a distance. If you calculate $1/H$, you have a distance divided by velocity. To finish the division and get an age, you need to convert megaparsecs into kilometers, and then the distances will cancel out and leave you with an age in seconds. To get years, you divide by the number of seconds in a year. If you make these simple changes in units, the age of the universe in years is approximately 10^{12} divided by H in its normal astronomical units, km/s/Mpc:

$$\tau_H = \frac{10^{12}}{H} \text{ years}$$

This estimate of the age of the universe is known as the Hubble time. For example, if H is 70 km/s/Mpc, that leads to an estimated age for the universe of $10^{12}/70$, or 14 billion, years.



The radiation that comes from such a great distance has a large redshift. The most distant visible objects are faint galaxies and quasars, with redshifts around 10. In contrast, the radiation from the hot gas right after the big bang has a redshift of about 1100. That means the light emitted as visible and near-infrared light by the hot gas in the early universe arrives at Earth as far-infrared, microwave, and radio waves. You can't see it with your eyes, but it can be detected with infrared and radio telescopes. Unlike the Gettysburg Address, the big bang can still be observed by the radiation it emitted. That amazing discovery is the subject of the next section.

The Cosmic Background Radiation

The story of the discovery of radiation from the time of the big bang begins in the mid-1960s when two Bell Laboratories physicists, Arno Penzias and Robert Wilson, were measuring the brightness of the sky at radio wavelengths (■ Figure 15-7). Their measurements showed a strange extra signal in the system, which they first attributed to the infrared glow from pigeon droppings inside the antenna. Perhaps they would have enjoyed scraping out the antenna more if they had known they would win the 1978 physics Nobel Prize for the discovery they were about to make.

When the antenna was cleaned, they again measured the radio brightness of the sky and found the low-level noise was still there. The pigeons were innocent, but what was causing the extra signal?

The explanation for the noise goes back decades earlier. In 1939, astronomers noticed that spectra of some molecules in the interstellar medium showed they were bathed in radiation from a source with a temperature of 2 to 3 K. In 1948, physicist George Gamow predicted that the gases in the universe right after the big bang would have been hot and should have emitted strong blackbody radiation (Chapter 6). A year later, physicists Ralph Alpher and Robert Herman pointed out that the large redshift of the big bang material relative to us would lengthen the wavelengths of that radiation into the microwave part of the spectrum. In the mid-1960s Robert Dicke at Princeton concluded the radiation should be just strong enough to detect with newly developed techniques. Dicke and his team began building a receiver. When Penzias and Wilson heard of Dicke's work, they recognized the mysterious extra signal they had detected as radiation from the big bang, the **cosmic microwave background** radiation.

■ Figure 15-6

This diagram shows schematically the expansion of a small part of the universe. Although the universe is now filled with galaxies, the look-back time distorts what you see. Nearby you see galaxies, but at greater distances, the look-back time reveals the universe at earlier stages, before galaxies formed. At very great distances, the big bang is detectable as infrared, microwave, and radio energy arriving from the hot gas that filled the universe soon after the big bang.

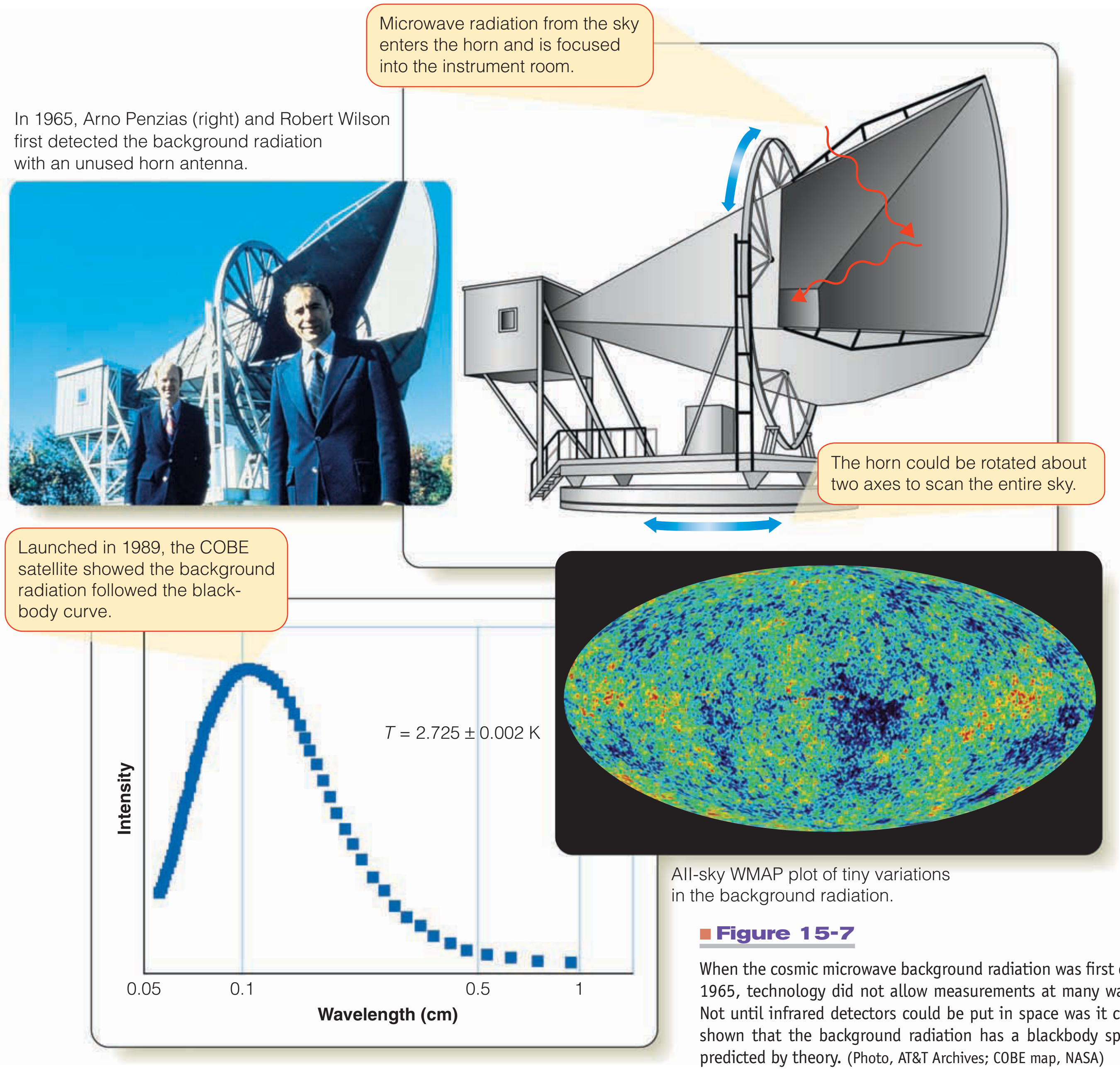


Figure 15-7

When the cosmic microwave background radiation was first detected in 1965, technology did not allow measurements at many wavelengths. Not until infrared detectors could be put in space was it conclusively shown that the background radiation has a blackbody spectrum, as predicted by theory. (Photo, AT&T Archives; COBE map, NASA)

The detection of the background radiation was tremendously exciting, but cosmologists wanted confirmation. Theory predicted that the radiation should have a spectrum like blackbody radiation coming from a very cool source, but the critical observations could not be made from the ground because Earth's atmosphere is opaque at the predicted blackbody peak wavelength. It was not until January 1990 that satellite measurements confirmed that the background radiation has exactly a blackbody spectral distribution, with an apparent temperature of 2.725 +/− 0.002 K—close to the original prediction.

It may seem strange that the hot gas of the big bang seems to have a temperature of just 2.7 degrees above absolute zero, but recall the tremendous redshift. Observers on Earth receive radiation that has a redshift of about 1100—that is, the wavelengths of the photons are about 1100 times longer than when they were emitted. The gas clouds that originally emitted the photons had a temperature of about 3000 K, and they emitted blackbody radiation with a λ_{max} of about 1000 nm (Wien's law, Chapter 6). Although that wavelength is in the near-infrared, the gas would also have emitted enough visible light to be seen glowing orange-red if there had been a human eye present at the time. The redshift has made the wavelengths of the background radiation about 1100 times longer on arrival at Earth than when they were when emitted, so λ_{max} is about 1 million nm, or 1 mm. That is why the hot gas of the big bang seems to be about 2.7 K, 1100 times cooler than it actually was.

The Story of the Big Bang

The first few steps up the cosmology pyramid have not been very difficult. Simple observations of the darkness of the night sky and the redshifts of the galaxies tell you that the universe must have had a beginning, and you have seen that the cosmic microwave background radiation is clear evidence of the big bang, that the early universe was hot and dense. Theorists can combine these observations with modern physics to add some details to the story of how the big bang occurred.

Cosmologists cannot begin their history of the big bang at time zero, because no one understands the physics of matter and energy under such extreme conditions, but they can come surprisingly close. If you could visit the universe when it was only 10 millionths of a second old, you would find it filled with high-energy photons having a temperature well over 1 trillion (10^{12}) K and a density of 5×10^{18} kg/m^3, greater than the density of an atomic nucleus. When cosmologists say the photons have a given temperature, they mean the photons have the same spectrum as blackbody radiation emitted by an object of that temperature. Consequently, the photons in the early universe were gamma rays, with very short wavelength and therefore very high energy. When cosmologists say that the radiation had a certain density, they refer to Einstein's equation $E = mc^2$. Using that equation, you can express a given amount of radiation energy per volume as if it were matter of a given density.

If photons have enough energy, two photons can interact and convert their energy into a pair of particles—a particle of normal matter and a particle of **antimatter.** When an antimatter particle meets its matching particle of normal matter—when an antiproton meets a normal proton, for example—the two particles annihilate each other and convert their mass into energy in the form of two gamma rays. In the early universe, the photons were gamma rays and had enough energy to produce proton–antiproton pairs or neutron–antineutron pairs. When these particles collided with their antiparticles, they converted their mass back into photons. Thus, the early universe was filled with a dynamic soup of energy flickering from photons into particles and back again.

While all this went on, the expansion of the universe caused the temperature of the radiation to drop, reducing the energy of the photons. When the universe was 0.0001 second old its temperature had fallen to 10^{12} K. By that time, the average energy of the radiation photons had fallen below the energy equivalent to the mass of a proton or a neutron, so the gamma rays could no longer produce such heavy particles. Those particles that did exist combined with their antiparticles and quickly converted their mass into photons.

It would seem that all of the protons and neutrons should have been annihilated with their antiparticles; but, for reasons that are poorly understood, a small excess of normal particles existed. For every billion protons annihilated by antiprotons, one survived with no antiparticle to destroy it. Consequently, you live in a world of normal matter, and antimatter is very rare.

Although the gamma ray photons did not have enough energy when the universe was older than about 0.0001 second to produce any more protons and neutrons, they could still produce electrons and anti-electrons (called positrons), which are about 1800 times less massive than protons and neutrons. That process continued until the universe was about 1 minute old, at which time the expansion had cooled to the point at which there were no remaining radiation photons with enough energy to create electron–positron pairs. Again, most of the electrons and positrons combined to form photons, and only one in a billion electrons survived. Cosmologists can calculate, based on the known properties of subatomic particles and also the characteristics of the universe as a whole, that the protons, neutrons, and electrons of which our universe is now made were produced during the first minute of its history.

The universal soup of hot gas and radiation continued to cool. Radiation photons with high enough energy can break up an atomic nucleus, so the formation of stable nuclei could not occur until the universe had cooled enough. By the time the universe was about 2 minutes old, protons and neutrons could link to form deuterium, the nucleus of a heavy hydrogen atom, without being immediately broken apart. By the end of the third minute, further reactions began converting deuterium into helium. Almost no atoms heavier than helium could be built in the big bang, however, because there are no stable nuclei with atomic weights of 5 or 8 (in units of the hydrogen atom). Nuclei of atomic weights 5 and 8 are radioactive and decay almost instantly back into smaller particles. Cosmic element building during the big bang had to proceed step-by-step, like someone hopping up a flight of stairs (■ Figure 15-8). The lack of stable nuclei at atomic weights of 5 and 8 meant there were missing steps in the stairway, and the step-by-step reactions had great difficulty jumping over these gaps during the few minutes of the big bang. As a result, cosmologists can calculate that only a tiny amount of lithium (atomic weight 7) would have been produced during the big bang, and no heavier elements.

By the time it was 3 minutes old, the universe had become so cool that almost all nuclear reactions had stopped. By the time it was 30 minutes old, the nuclear reactions had ended completely. At that point, about 75 percent of the mass of the universe was in the form of protons—hydrogen nuclei. The rest was helium nuclei. That composition, set during the first minute of the universe, is the composition observed now for the oldest stars. Formation of elements with atomic weights greater than lithium had to wait for relatively slow-cooking nucelosynthesis processes in stars (Chapter 10), beginning many millions of years after the big bang.

At first, the universe was so hot that the gas was totally ionized, and the electrons were not attached to nuclei. The free electrons interacted with photons so easily that a photon could

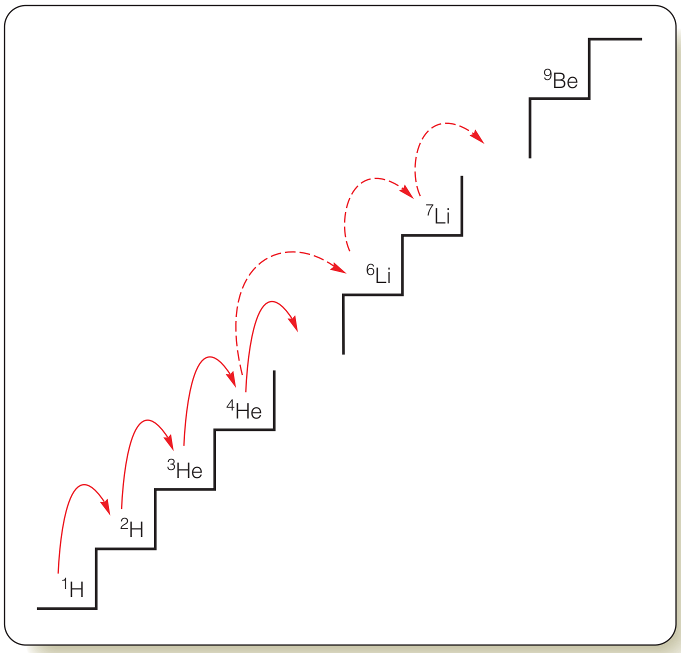


Figure 15-8

Cosmic element building. During the first few minutes of the big bang, temperatures and densities were high, and nuclear reactions built heavier elements. Because there are no stable nuclei with atomic weights of 5 or 8, the process built very few atoms heavier than helium.

not travel very far before it encountered an electron and was deflected. The photons interacted continuously with the matter, and the radiation and matter cooled together at a rate set by the expansion rate of the universe. In those circumstances, the universe was dominated by the radiation.

As the young universe continued to expand and cool, it went through three important changes. First, when the universe reached an age of roughly 50,000 years, the density of energy in the form of photons became less than the density of the gas. Before that time, ordinary matter could not clump together because the intense sea of photons smoothed the gas out. As you will learn in the next section, dark matter, which does not interact with photons, could start clumping much earlier. Once the density of the radiation fell below that of ordinary matter, the ordinary matter could begin to draw together under the influence of the gravitational attraction of dark matter to form the clouds that eventually became galaxies.

The expansion of the universe spread the particles of the still-ionized gas farther and farther apart. As the universe reached the age of about 400,000 years, the second important change began. The free electrons were spread so far apart that the photons could travel for thousands of parsecs before being deflected by an electron. That is, the universe began to become transparent. At about the same time, the third change happened. As the falling temperature of the universe reached 3000 K, protons were able to capture and hold free electrons to form neutral hydrogen, a process called **recombination.** (This term is a little misleading, because the particles had never been together stably before; *combination* would be more accurate.) As the free electrons were gobbled up, the gas finally became almost completely transparent, and the photons could travel through the gas without being deflected.

After recombination, although the gas continued to cool, the photons no longer interacted with the gas, and consequently the photons retained the blackbody temperature that the gas had at recombination. Those photons, which started their journey with a blackbody temperature of 3000 K, are observed now as the cosmic microwave background radiation. Remember that the large redshift—the expansion of the universe—has stretched the wavelengths of the background radiation so that it appears now to have a temperature of about 2.7 K.

Recombination resulted in the gas of the big bang being neutral, warm, and transparent. As the universe expanded and cooled, the glow from the warm gas faded into infrared wavelengths, and the universe entered what cosmologists call the **dark age,** a period lasting about 400 million years until the formation of the first stars. During the dark age the universe expanded in darkness.

The dark age ended as the first stars began to form. The gas from which the first stars formed contained almost no metals and was therefore highly transparent. Mathematical models show that stars formed from this metal-poor gas would have been very massive, very luminous, and very short lived. That first violent burst of massive star formation produced enough ultraviolet light to begin ionizing the gas, and astronomers now, looking back to the most distant visible quasars and galaxies, can see traces of that **reionization** era in the universe (■ Figure 15-9). Reionization marks the end of the dark age and the beginning of the age of stars and galaxies in which you live now.

Look carefully at ■ Figure 15-10; it summarizes the story of the big bang, from the formation of helium in the first 3 minutes through energy–matter equality, recombination, and finally reionization of the gas. It may seem amazing that mere humans limited to Earth can draw such a diagram, but remember that it is based on evidence and on the best understanding of how matter and energy interact (■ How Do We Know? 15-2).

◄ SCIENTIFIC ARGUMENT ►

How do you know there was a big bang?

A good scientific argument combines evidence and theory to describe how nature works, and this question calls for a detailed argument. The cosmic microwave background radiation consists of photons emitted by the hot gas of the big bang, so when astronomers detect those photons, they are "seeing" the big bang. Of course, all scientific evidence must be interpreted, so you need to understand how the big bang could produce radiation all around the sky before you can accept the background radiation as evidence. First, you must remember that the big bang event filled the entire universe with hot, dense gas. The big bang didn't happen in a single place; it happened everywhere. At recombination, the expansion of the universe reached the stage where the matter became transparent, and the radiation that had previously been trapped, bouncing between matter particles, was freed to

How Do We Know? | 15-2

Science: A System of Knowledge

What is the difference between believing in the big bang and understanding it? If you ask a scientist, "Do you believe in the big bang?" she or he may hesitate before responding. The question implies something incorrect about the way science works.

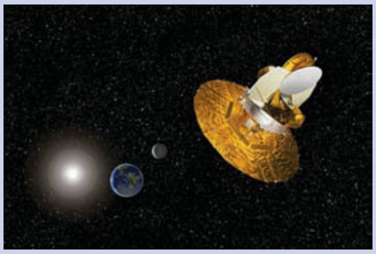
Science is a system of knowledge and process but not a system of belief. Science is an attempt to understand logically how nature works and is based on observations and experiments used to test and confirm hypotheses and theories. A scientist does not really believe in even a well-confirmed theory in the way people normally use the word *believe*. Rather, the scientist understands the theory and recognizes how different pieces of evidence support or contradict the theory.

There are other ways to know things, and there are many systems of belief. Religions, for example, are systems of belief that are not entirely based on observation. In some cases, a political system is also a system of belief; many people believe that democracy is the best form of government and do not ask for, or expect, evidence supporting that belief. A system of belief can be powerful and lead to deep insights, but it is different from science.

Scientists try to be careful with words, so thoughtful scientists would not say they *believe* in the big bang. They would say that the evidence is overwhelming that the big bang really did occur and that they are compelled by a logical analysis of both the observations and the theory, to conclude that the theory is very likely correct. In this way scientists try to be objective and reason without distortion by personal feelings and prejudices.

A scientist once referred to "the terrible rule of evidence." Sometimes the evidence forces a scientist to a conclusion she or he does not like, but science is not a system of belief, so the personal preferences of each scientist must take second place to the rule of evidence.

Do you believe in the big bang? Or, instead, do you have confidence that the theory is right because of your analysis of the evidence? There is a big difference.



Scientific knowledge is based objectively on evidence, such as that gathered by spacecraft. (NASA/WMAP Science Team)

travel through space. Now that radiation from the age of recombination arrives from all over the sky. It is all around you because you are part of the big bang event, and as you look out into space to great distance, you look back in time and see the hot gas in any direction you look. You can't see the radiation as visible light because the large redshift has lengthened the wavelengths by a factor of 1100 or so, but you can detect the radiation as photons with infrared, microwave, and radio wavelengths.

With this interpretation, the cosmic microwave background radiation is powerful evidence that there was a big bang. That tells you how the universe began, but your argument includes an important point. **Why do you think the universe cannot have a center or an edge?**

◀ ▶

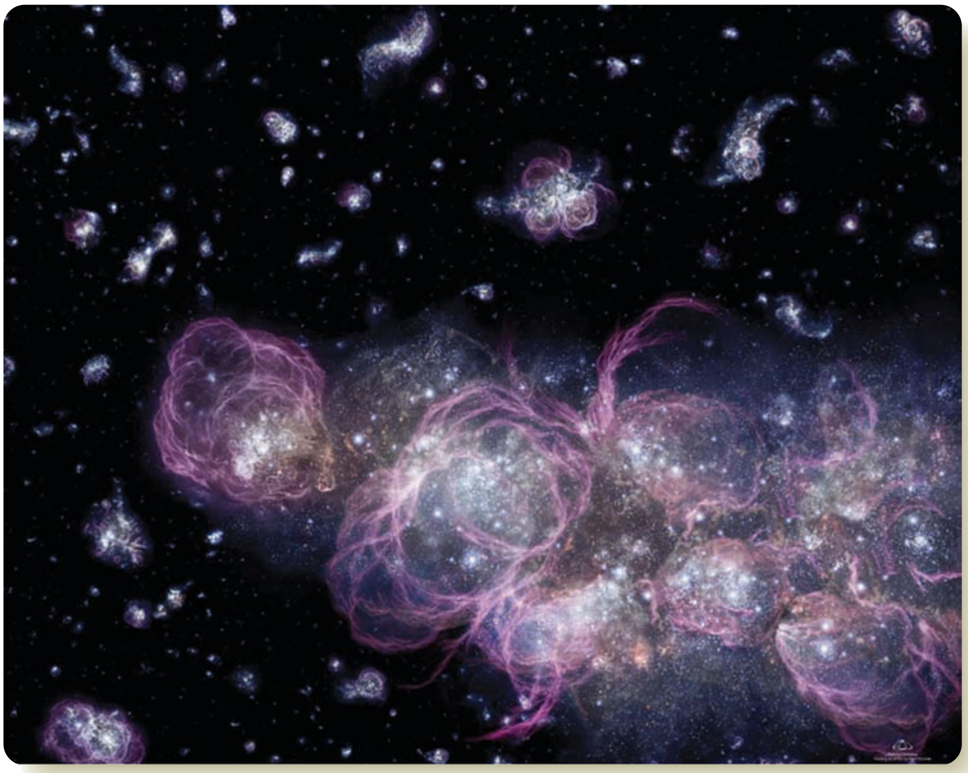
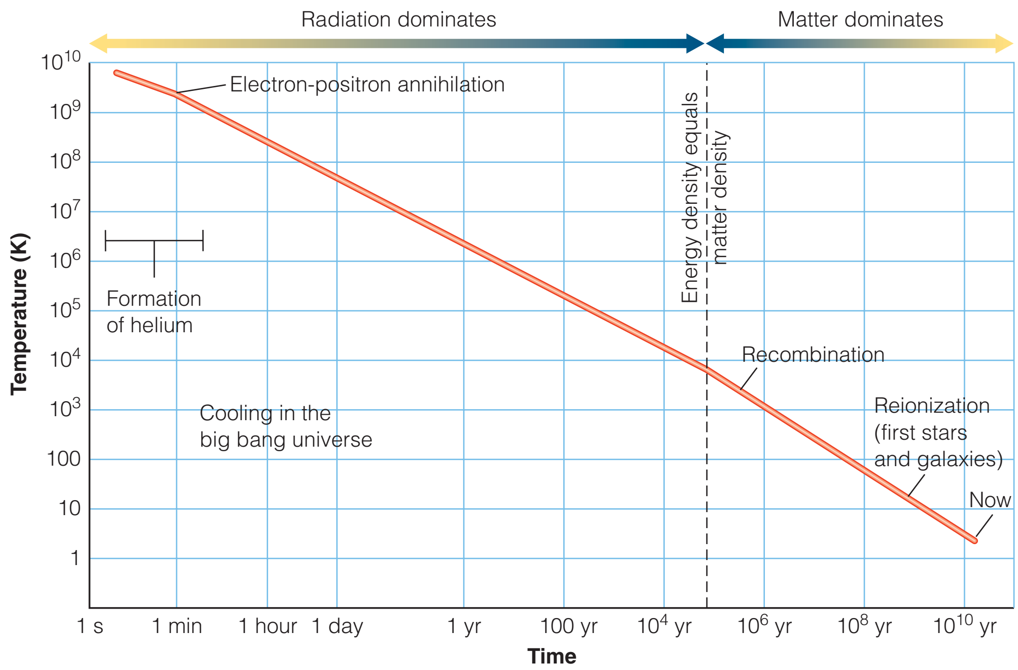


Figure 15-9

In this artist's conception of reionization, the first stars produce floods of ultraviolet photons that ionize the gas in expanding bubbles. Such a storm of star formation ended the so-called dark age, during which the universe expanded in darkness. Spectra of the most distant quasars reveal that those first galaxies were surrounded by neutral gas that had not yet been fully ionized. Thus the look-back time allows modern astronomers to observe the age of reionization. (K. Lanzetta, SUNY, A. Schaller, STScI, and NASA)

Figure 15-10

During the first few minutes of the big bang, some hydrogen was fused to produce helium, but the universe quickly became too cool for such fusion reactions to continue. The rate of cooling increased as matter began to dominate over radiation. Recombination freed the radiation from the influence of the gas, and reionization was caused by the birth of the first stars. Note how the exponential scale for the graph's time axis stretches early history and compresses recent history.



15-3 Space, Time, Matter, and Energy

How can the big bang have happened everywhere? To solve the puzzle, you have to put seemingly reasonable expectations aside and look carefully at how space and time behave on cosmic scales.

Looking at the Universe

The universe looks about the same in whichever direction you observe. The property of being the same in all directions is called **isotropy.** Of course, there are local differences. If you look toward a galaxy cluster you see more galaxies, but that is only a local variation. On the average, you see similar numbers of galaxies in every direction. Furthermore, the background radiation is almost perfectly uniform across the sky. The universe is observed to be highly isotropic.

The universe also seems about the same everywhere. The property of being the same in all places is called **homogeneity.** Of course, there are local variations; some regions contain more galaxies and some less. Also, because the universe evolves, then at large look-back times you see galaxies at an earlier developmental stage. But, if you account for these well-understood variations, then the universe seems to be, on average, the same everywhere.

Isotropy and homogeneity lead to the **cosmological principle,** which says that any observer in any galaxy sees the same general properties for the universe. As you just learned, this overall principle intentionally overlooks minor local and evolutionary variations. The cosmological principle implies there are no special places in the universe. What you see from the Milky Way Galaxy is typical of what all intelligent creatures see from their respective galaxies. Furthermore, the cosmological principle states, in a new form, the idea that the universe can have no center or edge. Such locations would be special places, and the cosmological principle says there are no special places.

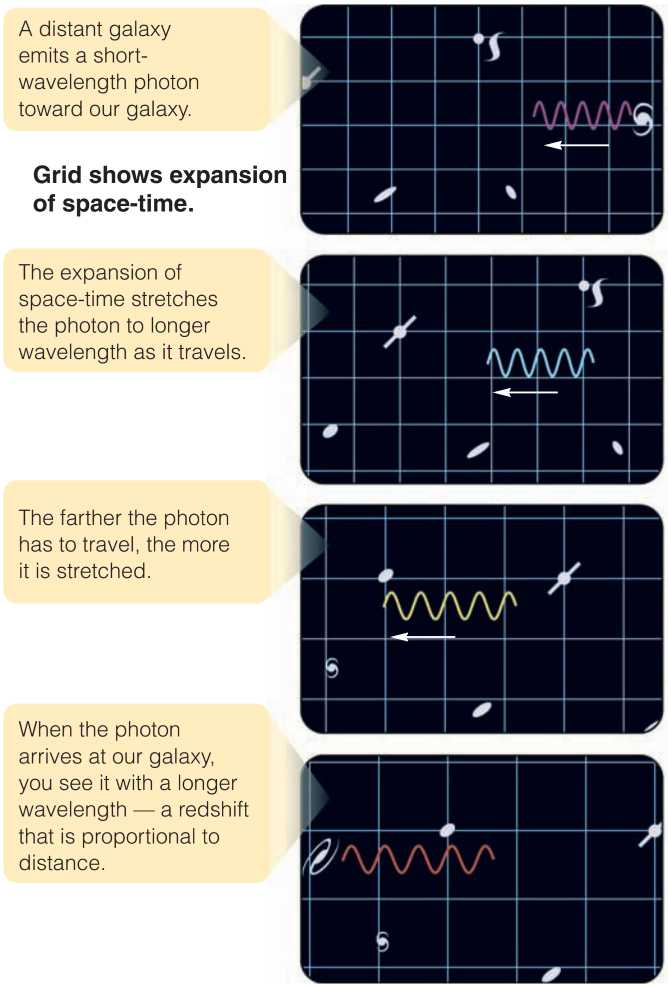
The Cosmic Redshift

Einstein's theory of general relativity, published in 1916, describes space and time together as the fabric of the universe, called space-time. That idea will give you a new insight into the meaning of the phrase *expanding universe.*

General relativity describes space-time as if it is made of stretching rubber, and that explains one of the most important observations in cosmology—the redshifts. It is a **Common Misconception** that cosmological redshifts are Doppler shifts of galaxies flying away through space. Instead, except for small, local motions within clusters of galaxies, the galaxies are at rest. They are being *separated* from each other as space-time expands. Also, as space-time expands, it stretches any photon traveling through space so that its wavelength increases. Photons from distant galaxies spend more time traveling through space and are stretched more, in other words have larger redshifts, than photons from nearby galaxies. That is why redshift depends on distance (■ Figure 15-11).

Astronomers often express redshifts as if they were actual radial velocities, but the redshifts of the galaxies are not Doppler shifts. That is why this book is careful to refer to a galaxy's *apparent* velocity of recession. Some textbooks convert cosmological redshifts to velocities using Einstein's relativistic Doppler formula, but that formula applies to motion through space and not

What are the cosmological redshifts?



■ **Figure 15-11**

Like a rubber sheet, space-time stretches, moving the galaxies away from each other and increasing the wavelength of photons as they travel through space-time.

to the behavior of space-time itself, so that formula should not be used in a cosmological context. The Hubble law still applies: Redshifts can be used to find the distance to galaxies because redshifts show how much the universe has expanded, and thus how much time has elapsed, since the photons started on their journey.

Model Universes

Almost immediately after Einstein published his general relativity theory, other theorists were able to solve the sophisticated mathematics and compute simplified descriptions of the behavior of space-time and matter. The resulting model universes dominated cosmology throughout the 20th century.

The equations allowed three possibilities. Space-time might be curved in either of two different ways, or it might have no curvature at all. Most people find these curved models difficult to imagine, and modern observations have shown that the simplest model, without curvature, is almost certainly correct, so you don't have to wrap your brain around the curved models. You might, however, like to know a few of their most important properties.

Some models predicted that space-time is curved back on itself to form a **closed universe.** You would not notice this curvature in daily life; it would only be evident in measurements involving very distant objects. Throughout the 20th century, observations could not eliminate closed models, so they were considered a real possibility. Closed universe models have finite volumes, but, because space-time is curved back on itself, a closed universe nevertheless would have no edge and no center. Such closed models predicted that the expansion of the universe would eventually become a contraction that would bring all of the matter and energy back to the high-density big bang state, which was sometimes called the "big crunch."

Other models predicted an uncurved, **flat universe.** That is the kind of space-time you would expect from your daily life, with the same rules of geometry, even for cosmological distances, that you learned in high school. Flat models are infinite, have no edge or center, and expand forever.

A third kind of model predicted an **open universe.** Such models contain curved space-time but are not curved back on themselves. Like the flat models, open models are infinite, have no center or edge, and expand forever.

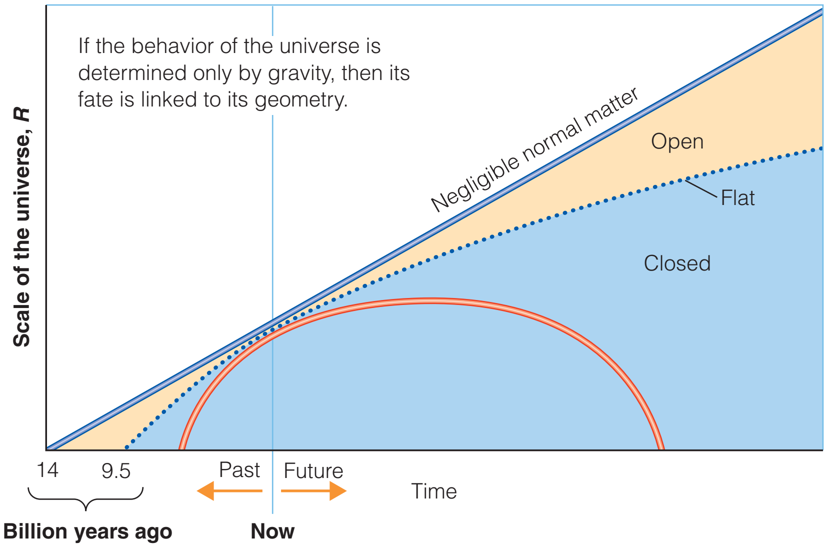
Cosmologists of the 20th century struggled to find reasons or observations that would allow them to choose among these three kinds of models. The main clue was density. According to general relativity, the overall curvature of space-time is determined by the average density of matter plus energy in the universe. Cosmologists could calculate that space-time is flat if the average density of the universe equals a **critical density** of 9×10^{-27} kg/m^3. If the average density of the universe is more than the critical density, the universe must be closed; and if it is less, the universe must be open. Attempts to measure the density, however, were inconclusive.

These three kinds of universe models are illustrated in ■ Figure 15-12, which compares the expansions versus time of the different models. The parameter R on the vertical axis is a measure of the extent to which the universe has expanded. You can think of R, in a highly simplified way, as the average distance between galaxies. In the figure you can see that closed universes expand and then contract while both flat and open universes expand forever.

Notice that you can't find the actual age of the universe unless you know whether it is open, closed, or flat. The Hubble time is the age the universe would have if it were totally open, which means it would have to contain almost no matter at all. If the universe contains enough matter to make it flat, then its age would be $^2/_3$ of the Hubble time. If H is about 70 km/sec/Mpc, as astronomers have measured, and the universe is flat, then it

Figure 15-12

Illustrations of some simple universe models that depend only on the effect of gravity, without including dark energy. Open universe models expand without end, and the corresponding curves fall in the region shaded orange. Closed models expand and then contract again (red curve). A flat universe (dotted line) marks the boundary between open and closed universe models. The relationship between estimated age and actual age of the universe (shown on the graph as the horizontal distance from "Now" back to the time when the cosmic scale factor *R* equaled 0) depends on the curvature of space-time.



should be only 9.5 billion years old. Astronomers know of globular star clusters that are older than that, so this was known as the age problem: How can the universe be younger than some of the objects in it?

Modern observations show that the universe is flat, and later in this chapter you will see how the evidence eliminated curved models and also solved the age problem. For now, however, you can think about a different puzzle. If the universe is flat, then its average density must equal the critical density. Yet when cosmologists added up the matter that could be detected, they found only a few percent of the critical density. They wondered if the dark matter made up the rest.

Dark Matter in Cosmology

There is more to the universe than meets the eye. In Chapter 13, you discovered that galaxies have much stronger gravitational fields than you would predict based on the amount of visible matter. Their gravitational fields are much stronger than expected even when you add in the gravity of the gas and dust that would be contained in the galaxies. Galaxies must contain dark matter.

This is not a small issue. Look at ■ Figure 15-13. Gravitational lensing is dramatic evidence of dark matter. The galaxy cluster in the figure contains so much dark matter that it warps space-time and focuses the images of more distant galaxies into short arcs of light. The universe contains much more dark matter than normal matter. Looking at the universe of visible matter in galaxies is like looking at a tree and seeing only the leaves.

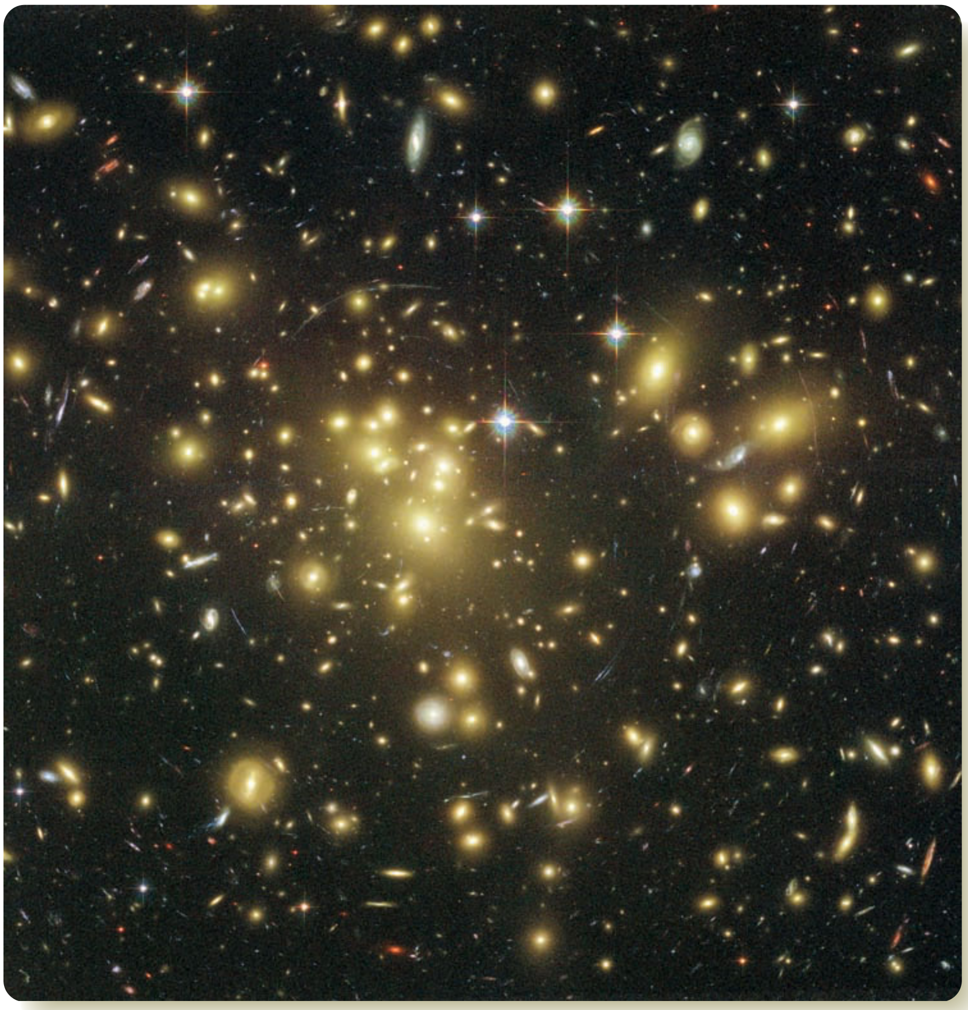
Modern evidence shows that the dark matter is not the same as the normal matter you are made of. To follow that argument, you need to estimate the density of normal matter, and you can start with what is known about conditions right after the big-bang.

As you have learned, during the first few minutes of the big bang, nuclear reactions converted some protons into helium and a very small amount into other elements. How much of elements heavier than helium was produced depended on the density of normal matter during the big bang. For example, if the density of matter were high so that there were lots of normal particles such as protons and neutrons flying around, they would have collided with the deuterium nuclei and converted most of them into helium. On the other hand, if the density of normal particles had been relatively low, more deuterium would have survived intact without being converted to helium.

Lithium is another nucleus that could have been made in small amounts during the big bang. Figure 15-8 shows that there is a gap between helium and lithium; there is no stable nucleus with atomic mass 5, so regular nuclear reactions during the first few minutes of the big bang had difficulty converting helium into lithium. If, however, the density of protons and neutrons was high enough, a few nuclear reactions could have leaped the gap and produced a small amount of the isotope lithium-7.

Both deuterium and lithium are destroyed in stars, so astronomers attempt the difficult task of studying clouds of gas at large look-back times that had not yet been altered by nuclear reactions in stars. As shown in ■ Figure 15-14, the amount of deuterium observed now in the universe sets a lower limit on the density of the universe relative to the critical density, and the observed abundance of lithium-7 sets an upper limit. The surprising conclusion from these observations and calculations is that normal matter from which you, Earth and the stars are made cannot make up much more than 4 percent of the critical density. Yet, observations of gravitational effects show that galaxies and galaxy clusters contain much larger amounts of matter than that, which can't be normal matter but instead must be dark matter.

The protons and neutrons that make up normal matter belong to a family of subatomic particles called *baryons,* so cosmologists think that most of the dark matter cannot be baryons. Only a small amount of the mass in the universe can be baryonic; the dark matter must be **nonbaryonic matter.**



■ **Figure 15-13**

Gravitational lensing shows that galaxy clusters contain much more mass than what is visible. The yellowish galaxies in the image above are members of a relatively nearby cluster of galaxies. Most of the objects in this image are blue or red images of very distant galaxies focused by the gravitational field of the nearby foreground cluster. Some of the distant imaged galaxies may be over 13 billion light-years away. That they can be seen at all is evidence that the foreground galaxy cluster contains large amounts of dark matter. (NASA, Benites, Broadhurst, Ford, Clampin, Hartig, Illingworth, ACS Science Team, and ESA)

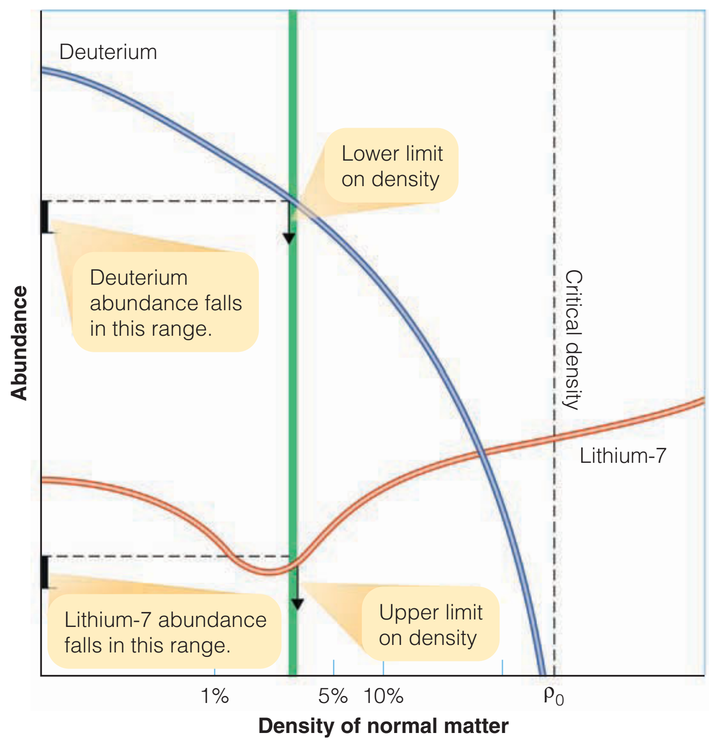
Some theorists thought that neutrinos, particles which are predicted to be very abundant in the universe yet are not baryons, might have enough mass to make up the dark matter, but modern measurements show that neutrinos are not massive enough. Some particle physics theories predict the existence of **WIMPs** (weakly interacting massive particles), but they have not been detected with certainty in the laboratory. The true nature of dark matter remains one of the mysteries of astronomy.

Theoretical models of galaxy formation in the early universe provide more hints about the nature of dark matter. If dark matter is composed of particles that were moving at, or near, the speed of light when matter and energy had equal density, it is called **hot dark matter.** Such fast-moving particles do not clump together easily and could not have stimulated the formation of objects as small as galaxies and clusters of galaxies. Instead, the most successful models of galaxy formation require that the dark matter actually be made up of **cold dark matter,** meaning the particles move slowly and could clump into relatively small structures.

Nonbaryonic dark matter does not interact significantly with normal matter or with photons, which is why you can't see it. This also means that dark matter was not affected by the intense radiation that dominated the universe when it was very young. So long as radiation dominated the universe, it prevented normal matter from contracting to begin forming galaxies. But the dark matter was

■ **Figure 15-14**

The diagram below compares observation with theory. Theory predicts how much deuterium (blue curve) and lithium-7 (red curve) you would observe for different possible densities of normal matter. The observed density of deuterium falls in a narrow range shown at upper left and sets a lower limit on the possible density of normal matter. The observed density of lithium-7 sets an upper limit. This means the true density of normal matter must fall in a narrow range represented by the green column. Certainly, the density of normal matter is much less than the critical density.



immune to the radiation, and the dark matter could contract to form clumps while the universe was very young. In other words, dark matter could have given galaxy formation a head start soon after the big bang. Models with cold, nonbaryonic dark matter are most successful at predicting the formation of galaxies and clusters of galaxies with the right size and at the right time in the history of the universe to match the observations.

Although the evidence is very strong that plenty of dark matter exists, it is still not abundant enough to make the universe flat. Dark matter appears to constitute less than 30 percent of the critical density. As you will see later in this chapter, there is more to the universe than meets the eye, more even than dark matter.

◄ SCIENTIFIC ARGUMENT ►

Why do cosmologists think that dark matter can't be baryonic?

Good scientific arguments always fall back on evidence. In this case the evidence is very compelling. Small amounts of isotopes like deuterium and lithium-7 were produced in the first minutes of the big bang, and the resulting abundance of those elements depends strongly on how many protons and neutrons were available at that time. Because those particles belong to the family of particles called baryons, physicists refer to normal matter as baryonic. Measurements of the abundances of deuterium and lithium-7 show that the universe cannot contain more baryons than about 4 percent of the critical density. Yet, observations of galaxies and galaxy clusters show that dark matter must make up almost 30 percent of the critical density. Consequently, cosmologists conclude that the dark matter must be made up of nonbaryonic particles.

Finding the dark matter is important because the density of matter in the universe determines the curvature of space-time. Now build a new argument. **How does the modern understanding of space-time explain cosmic redshifts?**

◄ ►

15-4 21st-Century Cosmology

If you are a little dizzy from the weirdness of expanding space-time and dark matter, make sure you are sitting down before you read further. As the 21st century began, astronomers made a discovery that startled all cosmologists: The expansion of the universe is actually accelerating. You'll have to go back a couple of decades to understand how that amazing new discovery nevertheless fits well with some of the things you have been learning in this chapter.

Inflation

By 1980, accumulating evidence had made the big bang theory widely accepted by cosmologists, but it faced two problems that led to the development of a new theory—a revised big bang with an important addition.

One of the problems is called the **flatness problem.** The properties of the universe appear to be close to the dividing line between being open or closed; that is, the geometry of the universe seems nearly flat. Given the vast range of possibilities, from zero to infinity, it seems peculiar that the density of the universe is close to the critical density that would make it exactly flat. If dark matter is as common as the measurements indicate, the universe's density must be within a factor of three of the critical density.

Even a small departure from critical density when the universe was young would be magnified by subsequent expansion. To be so near critical density now, the density of the universe during its first moments must have been within 1 part in 10^{49} of the critical density. So the flatness problem can be stated as: Why is the universe so close to perfectly flat?

The second problem with the original big bang theory is called the **horizon problem.** This is related to the observed isotropy of the cosmic microwave background radiation. When astronomers make a correction for the motion of Earth, they see the same intensity and temperature of background radiation in all directions to a precision of better than 1 part in 1000. Yet when you observe background radiation coming from two points in the sky separated by an angle larger than about 1 degree, you are looking at two bits of matter that seemingly should never have been connected from the big bang up to the time when the radiation was emitted. That is, when recombination occurred and the gas of the big bang became transparent to the background radiation, the universe was not old enough for any form of energy to have traveled from one of those two regions to the other. Thus, according to the standard big bang theory, the two spots you observe had not had enough time, since the start of the big bang, to exchange energy and even out their temperatures. Then, how did every part of the entire big bang universe get to be so nearly the same temperature at the time of recombination? This term *horizon* is used because the two spots are said to lie beyond their respective light-travel horizons.

The key to these two problems and to others involving subatomic physics may lie with the theory called the **inflationary universe** because it predicts a sudden rapid expansion, called inflation, when the universe was very young, a violent expansion even more extreme than that predicted by the original big bang theory.

To understand the inflationary universe, you need to recall that physicists know of only four forces—gravity, the electromagnetic force, the strong force, and the weak force (Chapter 7). You are familiar with gravity; you fought it to get out of bed this morning. The electromagnetic force is responsible for making magnets stick to refrigerator doors and cat hair stick to sweaters charged with static electricity, as well as holding electrons in orbit around atomic nuclei and being intimately connected with processes that make light and radiation. The strong force holds atomic nuclei together, and the weak force is involved in certain kinds of radioactive decay.

For many years, theorists have tried to unify these forces; that is, they have tried to describe all four forces as aspects of a single mathematical law. A century ago, James Clerk Maxwell showed that the electric force and the magnetic force were really the same effect, and physicists now count them as a single electromagnetic force. Similarly, in the 1960s, theorists succeeded in unifying the electromagnetic force and the weak force in what is now called the electroweak force, but those two forces operate effectively as a single unified force only in very high energy processes. At lower energies, the electromagnetic force and the weak force behave differently. Now, theorists have proposed ways of unifying the electroweak force with the strong force at even higher energies. These new theories are called **grand unified theories,** or **GUTs.**

According to the inflationary universe, the universe expanded and cooled until about 10^{-35} second after the big bang, when it became cool enough that the electroweak force and the strong force began to disconnect from each other; that is, they began to behave in different ways (■ Figure 15-15). This change released tremendous energy, which suddenly inflated the universe by a size factor between 10^{20} and 10^{30}. At the start of inflation, the part of the universe that is now observable from Earth is calculated to have been no larger than the volume of an atom, but it suddenly inflated to the volume of a cherry pit and then continued its slower expansion to its present extent.

Early rapid inflation of the universe can solve both the flatness problem and the horizon problem. The sudden inflation of the universe would have forced whatever curvature it had before that moment toward a value of zero, just as inflating a balloon makes a small spot on its surface flatter. Consequently, you now live in a universe that is almost perfectly flat because of that long-ago moment of inflation. In addition, because the entire observable part of the universe was once no larger in volume than an atom, it would have had enough time to equalize its temperature before inflation occurred. As a result, you now live in a universe in which the background radiation has the same temperature in all directions.

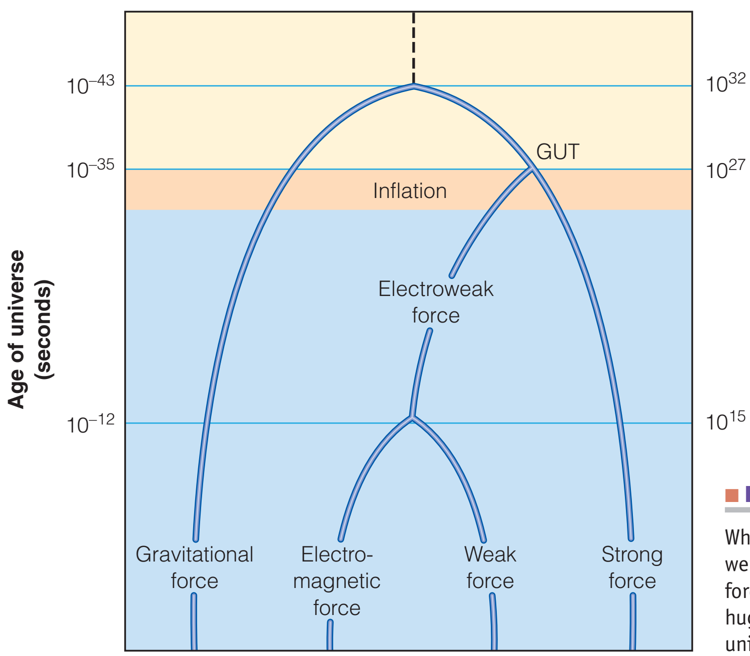
The inflationary theory predicts that the universe is almost perfectly flat, which means that the actual density equals the critical density. A theory can never be used as evidence, of course, but the beauty and simplicity of the inflationary theory has given many cosmologists confidence that the universe must be flat. Observations, however, seem to say that the universe does not contain enough matter (ordinary baryonic matter plus nonbaryonic dark matter) to be flat. Can there be more to the universe than baryonic matter and dark matter? What could be weirder and more obscure than dark matter? Read on.

The Acceleration of the Universe

Both common sense and the theory of general relativity suggest that as the galaxies recede from each other, the expansion should be slowed by gravity trying to pull the galaxies toward each other. How much the expansion is slowed should depend on the amount of matter in the universe. If the density of matter and energy in the universe is less than the critical density, the expansion should be slowed only somewhat, and the universe should expand forever. If the density of matter and energy is greater than the critical density, the expansion should be slowing down dramatically, and the universe should eventually begin contracting. Notice that this is the same as saying a low-density universe should be open and a high-enough-density universe should be closed.

For decades, several teams of astronomers struggled to measure the redshifts and distances to very distant galaxies and detect the slowing of the expansion. Detecting a change in the rate of expansion is difficult because it requires accurate measurements of the distances to very remote galaxies. The competing research teams used the same technique of calibrating type Ia supernovae as distance indicators. As you learned in Chapter 11, a type Ia supernova occurs when a white dwarf gains matter from a companion star, exceeds the Chandrasekhar limit, and collapses in a supernova explosion. Because all such white dwarfs should collapse at the same mass threshold, they should all produce explosions of the same size and luminosity, which makes them good distance indicators.

The two teams calibrated type Ia supernovae by locating such supernovae occurring in nearby galaxies, with



■ **Figure 15-15**

When the universe was very young and hot (top), the four forces of nature were indistinguishable in behavior. As the universe began to expand and cool, the forces "separated" (began to have different characteristics), which released a huge amount of energy and triggered a sudden rapid inflation in the size of the universe.

distances known from Cepheid variables and other reliable distance indicators. Once the peak luminosity of type Ia supernovae had been determined, they could be used to find the distances of much more distant galaxies.

Both teams announced their results in 1998. They agree that the expansion of the universe is not slowing down. Contrary to expectations, it is speeding up! That is, the expansion of the universe is accelerating (■ Figure 15-16).

The announcement that the expansion of the universe is accelerating was totally unexpected, and astronomers immediately began testing it. This conclusion depends critically on the calibration of type Ia supernovae as distance indicators, and some astronomers suggested that the calibration might be wrong (look back to How Do We Know? 12-1). Problems with the calibration have been ruled out by subsequently observed supernovae at even greater distances. The universe really does seem to be expanding faster and faster, rather than slowing down.

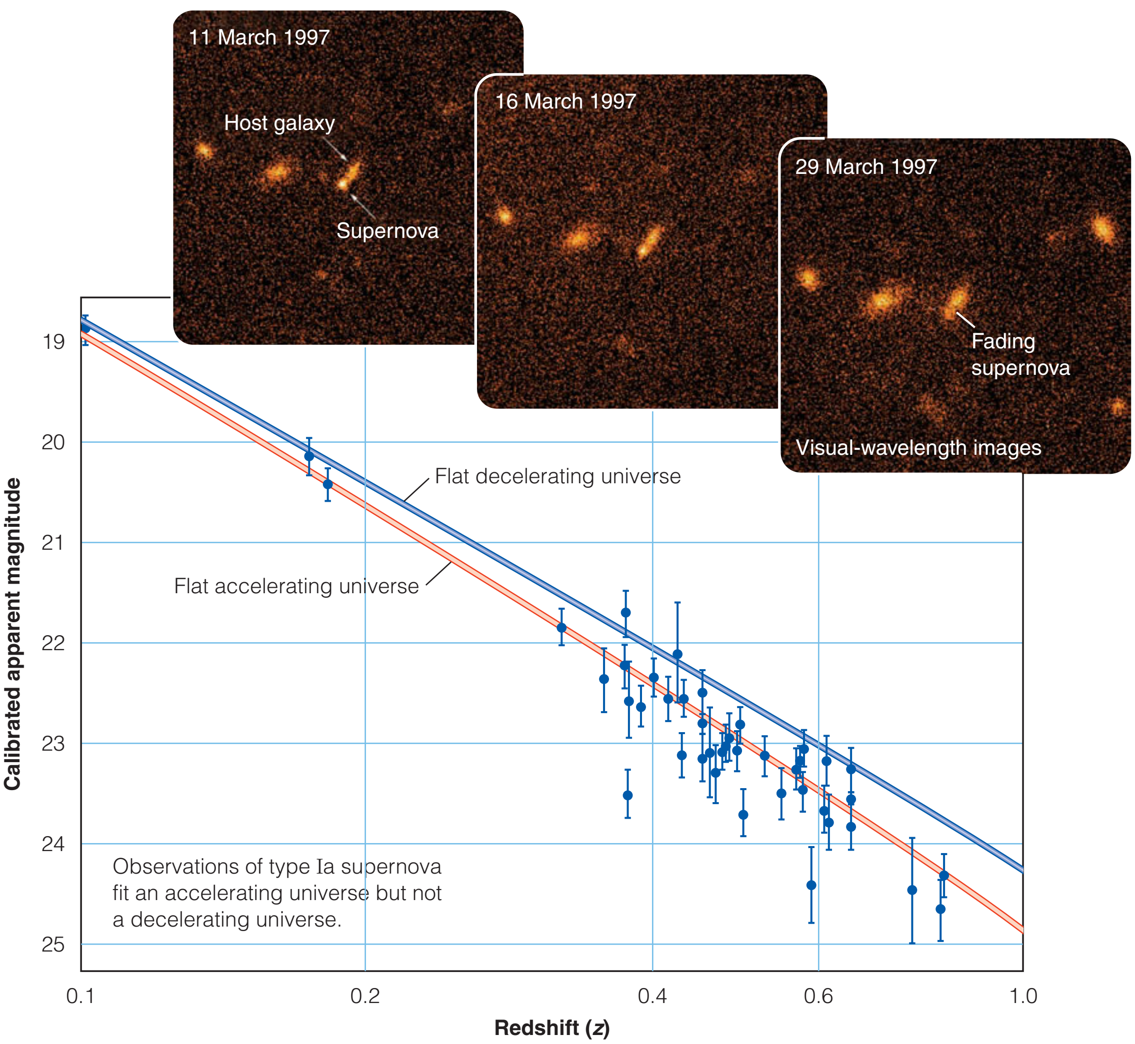
Dark Energy and Acceleration

If the expansion of the universe is accelerating, then there must be a force of repulsion in the universe that counteracts gravity, and cosmologists are struggling to understand what it could be. One possibility leads back to 1916.

When Albert Einstein published his theory of general relativity in 1916, he recognized that his equations describing spacetime implied that the galaxies could not float unmoving in space because their gravity would pull them toward each other. The only solutions seemed to be either a universe that was contracting under the influence of gravity or a universe in which the galaxies were rushing away from each other so rapidly that gravity could not pull them together. In 1916, cosmologists did not yet know that the universe was expanding, so Einstein made what he later thought was a mistake.

To balance the attractive force of gravity, Einstein added a constant to his equations called the **cosmological constant,** represented by an upper-case lambda (**Λ**). That constant represents a force of repulsion that balances the gravitational attraction between galaxies so the universe would not contract or expand. Thirteen years later, in 1929, Edwin Hubble announced his observations indicating that the universe is expanding, and Einstein said introducing the cosmological constant was his biggest blunder. Modern cosmologists think he may have been right after all.

One explanation for the acceleration of the universe is that the cosmological constant represents a force that causes a continuing acceleration in the expansion of the universe. Because the cosmological constant by definition remains constant with time,



■ **Figure 15-16**

From the way supernovae fade over time, astronomers can identify those that are type Ia. Once calibrated, the peak brightness of each of those supernovae could be compared with their respective redshifts, revealing that distant type Ia supernovae were about 25 percent fainter than expected. That must mean they are farther away than expected, given their redshifts. This is strong evidence that the expansion of the universe is accelerating. (ESO) the universe would have experienced this acceleration throughout its history.

Another solution is to suppose that totally empty space, the vacuum, contains energy, which drives the acceleration. This is an interesting possibility because for years theoretical physicists have discussed the idea that empty space contains energy. Cosmologists have begun referring to this energy of the vacuum as **quintessence.** Unlike the cosmological constant, quintessence would not necessarily remain constant over time.

Whatever the correct explanation might be, the observed acceleration of the universe's expansion is evidence that some form of energy is spread throughout space. Cosmologists refer to it as **dark energy,** an energy that drives the acceleration of the universe but does not contribute to the formation of starlight or the cosmic microwave background radiation.

You will recall that acceleration and dark energy were first discovered when astronomers found that supernovae a few billion light-years away were slightly fainter than expected based on their redshifts. The acceleration of the expansion has made those supernovae a bit farther away than implied by their redshifts, and so they look fainter. Since then, astronomers have continued to find even more distant type Ia supernovae, some as distant as 12 billion light-years. The more distant of those supernovae are not too faint; they are too bright! That reveals even more about dark energy.

Unlike the medium-distant supernovae, the very distant supernovae are a bit too bright because they are not as far away as expected, and that confirms a theoretical prediction based on dark energy. When the universe was young, galaxies were closer together, and their gravitational pull on each other was stronger than the effect of dark energy and slowed the expansion. That makes the very distant supernovae a bit too bright. As space-time expanded, it moved the galaxies farther apart, their gravitational pull on each other became weaker than dark energy, and acceleration began. That makes the medium-distant supernovae a bit too faint. In other words, the observations show that sometime about 6 billion years (equal to 40 percent of the age of the universe) ago, the universe shifted gears from deceleration to acceleration. The calibration of type Ia supernovae allows astronomers to observe this change from deceleration to acceleration.

Furthermore, dark energy can help you understand the curvature of the universe. The theory of inflation makes the specific prediction that the universe is flat. Recent observations you will learn about in the rest of this chapter confirm that prediction. Dark energy seems to explain how the universe can have enough matter plus energy to be flat. As you have already learned, $E=mc^2$ means that energy and matter are equivalent. Thus, the dark energy is equivalent to an amount of mass spread through space. Baryonic matter plus dark matter makes up about one-third of the critical density, and dark energy appears to make up the remaining two-thirds. That is, when you include dark energy, the total density of the universe equals the critical density, which makes the universe flat.

Step by step, you have been climbing the cosmological pyramid. Each step has been small and logical, but look where it has led you. You now know some of nature's deepest secrets, but there are yet more steps above and more secrets to explore.

The Age and Fate of the Universe

Acceleration helps with another problem. The Hubble constant equals 70 km/s/Mpc, and earlier in this chapter you calculated the Hubble time, an estimate of the age of the universe, and found it was about 14 billion years. Further, you calculated a more precise estimate of the universe's age on the assumption that its geometry is flat, which gives an age of two-thirds of the Hubble time, or 2/3 times 14 billion years, which equals about 9 billion years. That was a puzzle because the globular star clusters are older than that. Now you are really ready to solve the age problem and then consider the fate of the universe.

If the expansion of the universe has been accelerating, then it must have been expanding slower in the past, and the galaxies would have taken longer to arrive at the average separations you can observe now. That means the universe can be flat but nevertheless older than two-thirds of the Hubble time. The latest estimates suggest that acceleration makes the age of a flat universe almost 14 billion years, coincidentally about the same as the Hubble time estimate. That age is clearly older than the oldest known star clusters, which solves the age problem.

For many years, cosmologists have enjoyed saying, "Geometry is destiny." Thinking of models of open, closed, and flat universes, they concluded that the density of a model universe determines its geometry, and its geometry determines its fate. By this they meant that if the universe is open it must therefore expand forever, whereas if it is closed, it must eventually begin contracting. But that is true only if the behavior of the universe as a whole is ruled completely by gravity. If dark energy causes an acceleration that can dominate gravity, then geometry is NOT destiny, and, depending on the precise properties of dark energy, even a closed universe might expand forever.

The ultimate fate of the universe depends on the nature of dark energy. If dark energy is described by the cosmological constant, then the force driving acceleration does not change with time, and our flat universe will expand forever with the galaxies getting farther and farther apart and using up their gas and dust making stars, and stars dying, until each galaxy is isolated, burnt out, dark, and alone. If, however, dark energy is described by quintessence, then the force may be increasing with time, and the universe would accelerate faster and faster as space pulls the galaxies away from each other, eventually pulling the galaxies apart and then pulling the stars apart and finally ripping individual atoms apart. This possibility has been called the **big rip.**

Probably there will be no big rip; critically important observations made by the Chandra X-ray Observatory have been used to measure the amount of hot gas and dark matter in almost

30 galaxy clusters. Because the distance of each cluster is one of the variables in the calculations, the X-ray astronomers could compare the amount of hot gas and dark matter and solve for distance. The farthest of the clusters is 8 billion light-years away. These observations are important for two reasons. First, the redshifts and distance of these galaxies independently confirm the conclusion from the supernova observations that the expansion of the universe first decelerated but changed to accelerating about 6 billion years ago. Second, the Chandra results almost completely rule out quintessence. If dark energy is described by the cosmological constant and not by quintessence, then there will be no big rip (■ Figure 15-17).

The Origin of Large-Scale Structure

On the largest scales, the universe is isotropic. That is, it looks the same in all directions. But, on smaller scales, there are irregularities. The sky is filled with galaxies and clusters of galaxies that seem to be related to their neighbors in even larger aggregations that astronomers call **large-scale structure.** Studies of large-scale structure lead to insights about how the universe has evolved.

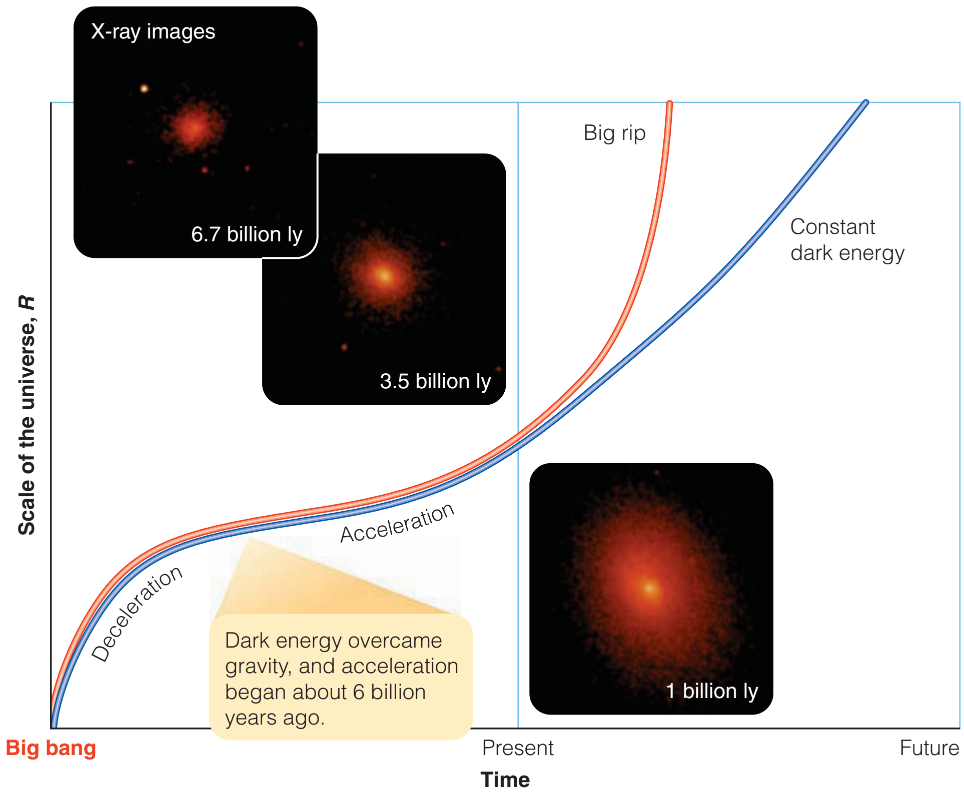
When you look at galaxies in the sky you see them in clusters ranging from a few galaxies to thousands, and those clusters appear to be grouped into **superclusters.** The Local Supercluster, in which we live, is a roughly disk-shaped swarm of galaxy clusters 50 to 75 Mpc in diameter. By measuring the redshifts and positions of over 100,000 galaxies in great slices across the sky, astronomers have been able to create maps revealing that superclusters are not scattered at random. They are distributed in long, narrow **filaments** and thin walls that outline great **voids** nearly empty of galaxies (■ Figure 15-18).

This large-scale structure is a problem because the cosmic microwave background radiation is very uniform, and that means the gas of the big bang must have been extremely uniform at the time of recombination. Yet the look-back time to the farthest galaxies and quasars is about 93 percent of the way back to the big bang. How did the uniform gas at the time of recombination coagulate so quickly to form galaxies? How did it make clusters of galaxies and the supermassive black holes we observe as quasars (Chapter 14) so early in the history of the universe?

Baryonic matter is so rare in the universe that cosmologists can calculate that it did not have enough gravity to pull itself together quickly after the big bang. As you have read earlier, cosmologists propose that dark matter is nonbaryonic and therefore immune to the intense radiation that prevented normal matter from contracting. Dark matter, in contrast, was able to collapse into clouds and then pull in the normal matter to begin the formation of galaxies, clusters, and superclusters. Mathematical models have been made to describe this process, and cold dark matter does seem capable of jump-starting the formation of structure of the right sizes and in the right amount of time (■ Figure 15-19).

But what started the clumping of the dark matter? Theorists say that space is filled with tiny, random quantum mechanical fluctuations smaller than the smallest atomic particles. At the moment of inflation, those tiny fluctuations would have been stretched to very large but very subtle variations in gravitational fields that could have stimulated the formation of clusters, filaments, and walls. The structure you see in Figure 15-18 may be the present-day ghostly traces of microscopic random fluctuations in the infant universe.

The Two-Degree-Field Redshift Survey mapped the position and redshift of 250,000 galaxies and 30,000 quasars. As



■ Figure 15-17

X-ray observations of hot gas in galaxy clusters confirm that in its early history the universe was decelerating because gravity was stronger than the dark energy. As expansion weakened the influence of gravity, dark energy began to cause acceleration. The evidence is not conclusive, but it most directly supports the cosmological constant form of dark energy and weighs against quintessence, which means the universe will probably not undergo a big rip. This diagram is only schematic, and the two curves are drawn separated for clarity; at the present time, the two curves have not diverged from each other. (NASA/CXC/IoA/S. Allen et al.)

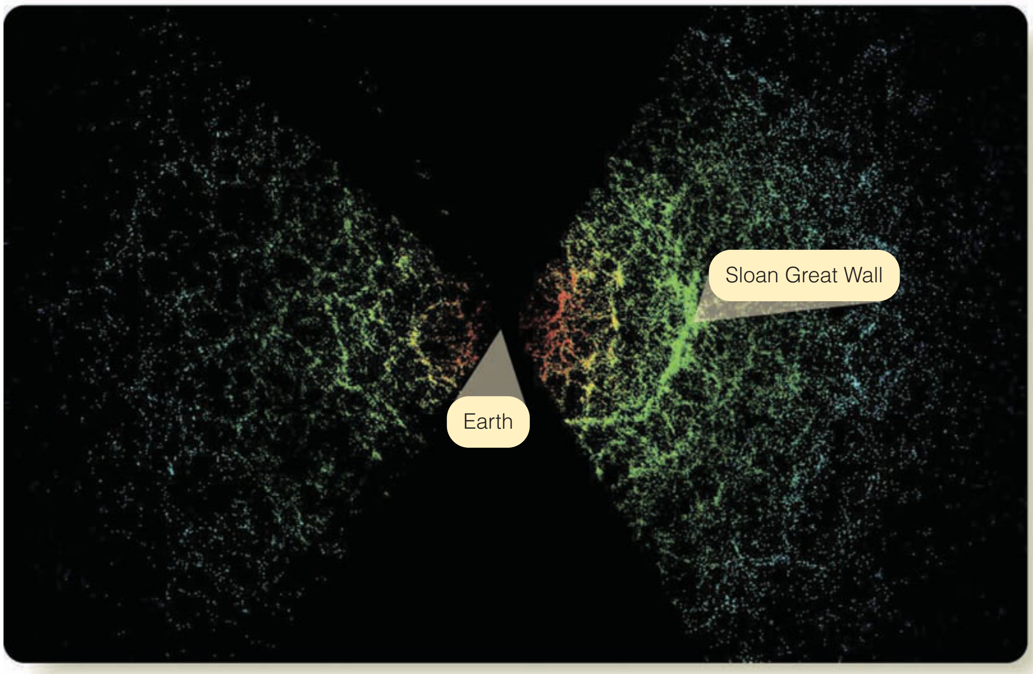


Figure 15-18

Nearly 70,000 galaxies are plotted in this double slice of the universe extending outward in the plane of Earth's equator. The nearest galaxies are shown in red and the more distant in green and blue. The galaxies form filaments and walls enclosing mostly empty voids. The Sloan Great Wall is almost 1.4 billion light-years long and is the largest known structure in the universe. The most distant galaxies in this diagram are roughly 3 billion light-years from Earth. (Sloan Digital Sky Survey)

expected, the galaxies are spread in filaments and walls, and a statistical analysis of the distribution matches the prediction of inflation and shows that the current expansion is accelerating. This is an important result because it is yet another confirmation that is independent of the brightness of type Ia supernovae or X-ray-emitting gas in galaxy clusters. When a theory is confirmed by observations of many different types, scientists have much more confidence that it is a true description of nature.

Curvature of the Universe

Astronomers continue to search for more ways to confirm that the universe is flat and that its expansion is accelerating. One way to measure the curvature of space-time is to measure the size of things at great look-back times; for example, observing tiny irregularities in the background radiation. That type of measurement is yielding impressive results.

The background radiation is nearly isotropic—it looks almost exactly the same in all directions after effects of Earth's motion and also emission by material in the foreground are taken into account. After those corrections are made to a map of the background radiation and the average intensity is subtracted from each spot on the sky, small irregularities are evident. That is, some spots on the sky look a tiny bit hotter and brighter, or cooler and fainter, than the average (■ Figure 15-20; also look back at Figure 15-7). Those irregularities contain lots of information.

The inflation-modified big bang theory of the universe makes specific predictions about the angular sizes of the variations you should find in the background radiation. Observations made by the COBE satellite in 1992 detected the variations with the largest diameter, but detecting smaller variations is critical for testing the theory. Six teams of astronomers have built specialized telescopes to make these observations. Some have flown under balloons high in the atmosphere, and others have observed from the ice of Antarctica. The most extensive observations are made by the Wilkinson Microwave Anisotropy Probe (WMAP), a space infrared telescope.

It is a **Common Misconception** that explosions in space produce sound. Science-fiction movies imply that sound can travel through a vacuum and that exploding space ships make satisfying *kabooms.* That's not true now, of course, but the early universe was dense enough that sound could travel through the gas, and the big bang did make a noise. A theorist described it as "a descending scream, building to a deep rasping roar, and ending in a deafening hiss." The pitch of the sound was about 50 octaves too low for you to hear, but those powerful sound waves did have an effect on the universe. They determined the size of the irregularities now detectable in the cosmic microwave background radiation.

Cosmologists can analyze those irregularities using sophisticated mathematics to find out how commonly the spots of different angular sizes recur. The mathematics confirm that spots about 1 degree in diameter are the most common, but spots of other sizes occur as well, and it is possible to plot a graph such as ■ Figure 15-21 to show how common different-size irregularities are.

Theory predicts that most of the irregularities in the hot gas of the big bang should be about one degree in diameter if the universe is flat. If the universe were open, the most common irregularities would be smaller. Careful measurements of the size of the irregularities in the cosmic background radiation show that the observations fit the theory very well for a flat universe, as you can see in Figure 15-20. Not only is the theory of inflation confirmed, an exciting result itself, but these data show that the universe is flat, meaning space-time has no overall curvature, which indirectly confirms the existence of dark energy and the acceleration of the universe.

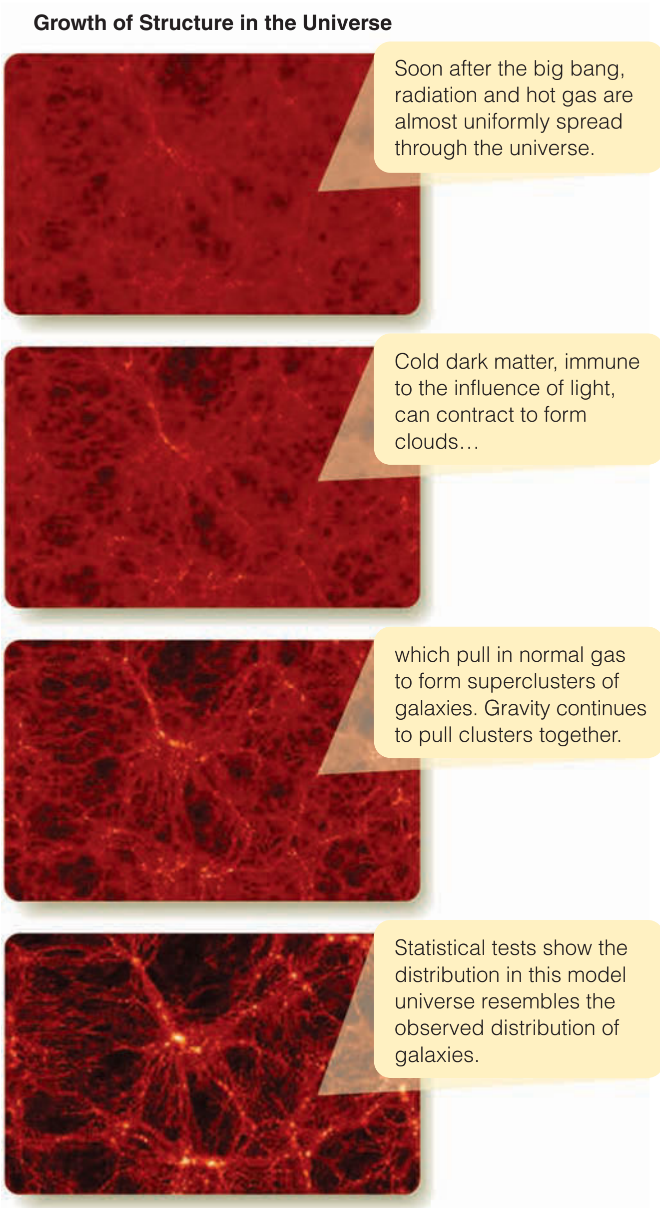


Figure 15-19

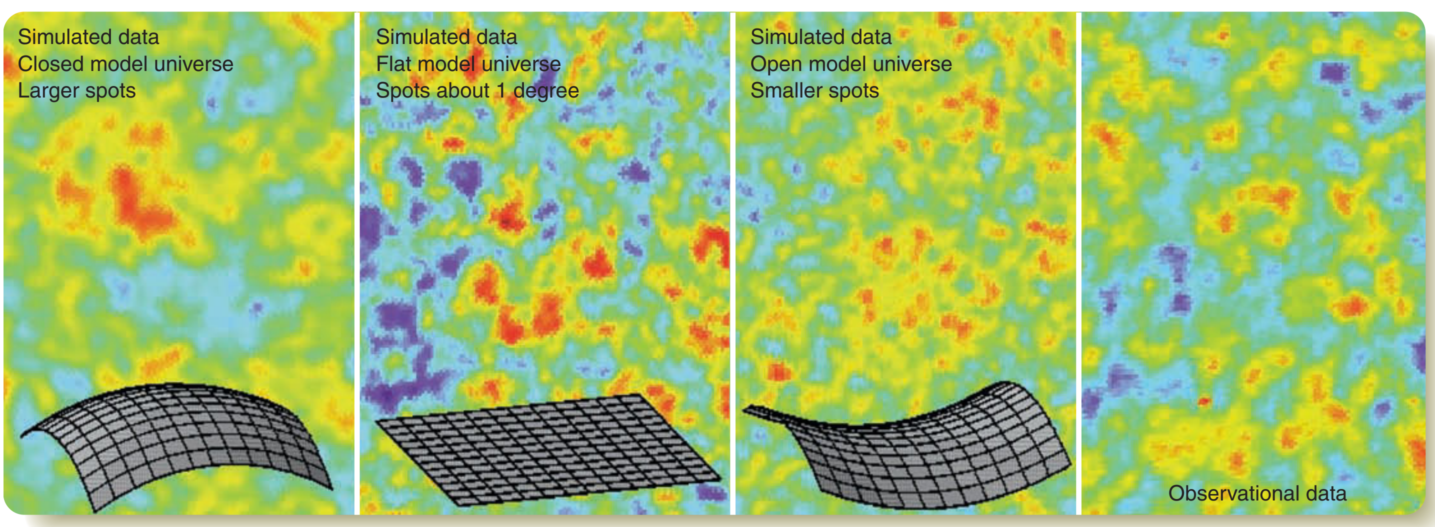
This computer model traces formation of structure in the universe from soon after the big bang to the present. (Adapted from a model by Kauffmann, Colberg, Diaferio, & White: Max-Planck Institute für Astronomie)

The data points from the WMAP observations follow a wiggling line in Figure 15-21, and the size and positions of those wiggles tell cosmologists a great deal about the universe. The details of the curve show that universe is flat, accelerating, and will expand forever. The age of the universe derived from the data is 13.7 billion years. Furthermore, the smaller peaks in the curve reveal that the universe contains 4 percent baryonic (normal) matter, 23 percent nonbaryonic dark matter, and 73 percent dark energy. The Hubble constant is confirmed to be 71 km/s/Mpc. The inflationary theory is confirmed, and the data support the cosmological constant version of dark energy, although quintessence is not quite ruled out. Hot dark matter is ruled out. The dark matter needs to be cold dark matter to clump together rapidly enough after the big bang to make the galaxy clusters and superclusters we observe. In fact, a model based on these data predicts that the first stars began to produce light when the universe was only about 400 million years old.

Please reread the preceding paragraph. Especially to people who have been working in the field for years, that collection of firm facts is mind-blastingly amazing. WMAP and other studies of the cosmic microwave background radiation and the distribution of galaxies have revolutionized cosmology. At last, astronomers have accurate observations against which to test theories. The values of the basic cosmic parameters are known to a precision of 1 percent or so.

Figure 15-20

You can see the difference yourself. Compare the observations of the irregularities in the background radiation in the far right panel with three simulations starting from the left. The observed size of the irregularities fits best with cosmological models having flat geometry. Detailed mathematical analysis confirms your visual impression: The universe is flat. (Courtesy of the BOOMERANG Collaboration)



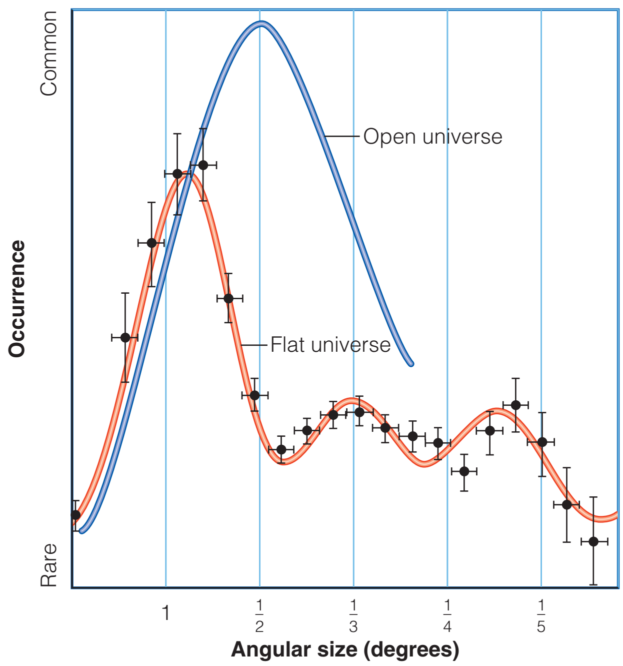


Figure 15-21

This graph shows how commonly irregularities of different sizes occur in the cosmic microwave background radiation. Irregularities of about 1 degree in diameter are most common. Models of the universe that are open or closed are ruled out. The data fit a flat model of the universe very well. Crosses on the data points show uncertainty in the measurements.

On reviewing these results, one cosmologist announced that "Cosmology is solved!" but that might be premature. Cosmologists don't understand dark matter or the dark energy that drives the acceleration, so over 95 percent of the universe is not understood. Hearing this, another cosmologist suggested a better phrase was "Cosmology in crisis!" Certainly there are further mysteries to be explored, but cosmologists are growing more confident that they can describe the origin and evolution of the universe.

SCIENTIFIC ARGUMENT

How does inflation theory solve the flatness problem?

This question requires a carefully constructed argument. The flatness problem can be stated as a question: Why is the universe so flat? After all, the density of matter in the universe could be anything from zero to infinite, but the observed sizes of variations in the cosmic background radiation show that the universe is flat, and therefore the average density of the universe is close to the critical density (Figure 15-21). Furthermore, the density must have been astonishingly close to the critical density when the universe was very young, or it would not be so close now. The inflationary theory solves this problem by proposing that the universe had a moment of rapid inflation when it was a tiny fraction of a second old. That inflation drove the universe toward flatness just as blowing up a balloon makes a spot on the balloon flatter and flatter.

Understanding theory in cosmology is critically important, but science depends on evidence. Now build another argument based on evidence. **What evidence can you cite that the expansion of the universe is accelerating?**

What Are We? The Big Picture

As you climbed the cosmology pyramid, you negotiated many steps. Most were easy, and all were logical. They have carried you up to some sweeping views, and you have traced the origin of the universe, the creation of the chemical elements, the birth of galaxies, and the births and deaths of stars. You now have a perspective that few humans share.

We are by-products of two cosmic processes, gravitational contraction and nuclear fusion. Gravity created instabilities in the hot matter of the big bang and triggered the formation of clusters of galaxies. Further instabilities caused the contraction of that material to form individual galaxies and then stars within the galaxies. As stars began to shine through the universe, nuclear fusion in their cores began to cook hydrogen and helium into the heavier atoms from which humans are made.

As you review the history and structure of the universe, it is wise to recognize the mysteries that remain; but note that they are mysteries that may be solved and not mysteries that are unknowable. Only a century ago, humanity didn't know there were other galaxies, or that the universe was expanding, or that stars generate energy by nuclear fusion. Human curiosity has solved many of the mysteries of cosmology and will solve more during your lifetime.

Study and Review

Summary

- **Cosmology (p. 296)** is the study of the structure and history of the universe as a whole. Astronomers and physicists who do research in cosmology are called **cosmologists (p. 297).**
- Cosmologists conclude that it is impossible for the universe to have an edge because an edge would introduce logical inconsistencies. In other words, an edge to the universe does not make sense.
- If the universe has no edge, then it cannot have a center.
- The darkness of the night sky leads to the conclusion that the universe is not infinitely old. If the universe were infinite in extent, infinite in age, and **static (p. 297),** then every spot on the sky would glow as brightly as the surface of a star. This problem, commonly labeled **Olbers's paradox (p. 297),** implies that the universe had a beginning.
- The **observable universe (p. 299)**, the part you can see from Earth with a large telescope, is limited in size. It is possibly only a tiny portion of the entire universe, which could be infinite.
- Edwin Hubble's 1929 discovery that the redshift of a galaxy is proportional to its distance, known as the Hubble law, shows that the galaxies are moving away from each other. That phenomenon is called the **expanding universe (p. 299).**
- Tracing the expansion of the universe backward in time brings you to imagine an initial high-density, high-temperature state commonly called the **big bang (p. 300).**
- A rough estimate of the age of the universe based on the presently observed expansion rate is called the **Hubble time (p. 300).**
- Although the expanding universe began from the big bang, it has no center. The galaxies do not move away from a single point. Cosmologists understand that galaxies remain approximately in their respective positions and become farther away from each other as the space between them expands.
- The **cosmic microwave background (p. 301)** radiation is blackbody radiation with a temperature of about 2.73 K, spread nearly uniformly over the entire sky. This radiation is the light from the big bang, freed from the gas at the time of **recombination (p. 304),** and now redshifted by a factor of 1100.
- The background radiation is clear evidence that the universe began with a big bang.
- During the earliest moments of the universe, matter and **antimatter (p. 303)** particles continually flashed in and out of existence. A slight excess of ordinary matter remained after most of the matter and antimatter particles annihilated each other.
- During the first three minutes of the big bang, nuclear fusion converted some of the hydrogen into helium but was unable to make many other heavy atoms because no stable nuclei exist with weights of 5 or 8. Now, hydrogen and helium are common in the universe, but heavier atoms are rare.
- After recombination, for a period of hundreds of millions of years called the **dark age (p. 304),** the universe expanded in darkness until the first stars came into existence. Astronomers have observed signs of **reionization (p. 304)** of the universe caused by that first generation of stars.
- The chemical composition of the oldest stars is about 75 percent hydrogen and 25 percent helium, which is what models of the big bang nuclear processes would predict. This is further evidence supporting the big bang theory.
- The universe is **isotropic (p. 306)** and **homogeneous (p. 306).** In other words, in its major features, the universe looks the same in all directions and in all locations.
- Isotropy and homogeneity lead to the **cosmological principle (p. 306)**, the idea that there are no special places in the universe. Except for minor local differences, every place is the same, and the view from every place is the same.
- Einstein's theory of general relativity explains that cosmic redshifts are caused by wavelengths of photons stretching as they travel through expanding space-time.
- **Closed universe (p. 307)** models are finite in size, but their space-time is curved back on itself so they have no edge or center.
- **Open universe (p. 307)** models have curved space-time, but it is not curved back on itself. Such universes are infinite.
- **Flat universe (p. 307)** models have uncurved space-time and are infinite. Modern observations show that the universe is probably flat.
- Open model universes contain less than the **critical density (p. 307)** and closed model universes more. If the universe is flat, then its density must equal the critical density.
- The amounts of deuterium and lithium-7 in the universe shows that normal baryonic matter can make up only about 4 percent of the critical density. Dark matter must be **nonbaryonic (p. 308)** and appears to make up less than 30 percent of the critical density.
- One hypothesis for what dark matter might be is a type of subatomic particle labeled **WIMPs (p. 309).** The observed ranges of sizes and masses for galaxy clusters disagrees with models that assume rapidly moving **hot dark matter (p. 309)** particles that would not clump together easily enough, and agrees with models that assume slowly moving **cold dark matter (p. 309).**
- The **inflationary universe (p. 310),** a modification to the big bang theory, proposes that the universe briefly expanded dramatically, just a tiny fraction of a second after the big bang.
- The energy to drive inflation would have been released when the four forces of nature changed their respective properties as the universe cooled in its earliest moments. This "separation" of forces is predicted by **grand unified theories (GUTs) (p. 311)** that explain the forces of nature as being aspects of a single force, unified in particle interactions with very high energies.
- Inflation explains the **flatness problem (p. 310)** because the large expansion forced the universe to become flat, as a spot on an inflating balloon becomes flatter as the balloon inflates.
- Inflation explains the **horizon problem (p. 310)** because what is now the observable part of the universe was so small before inflation that energy could move and equalize the temperature everywhere in that volume.
- Observations of type Ia supernovae reveal the surprising fact that the expansion of the universe is speeding up. Cosmologists propose that this acceleration is caused by energy present in empty space labeled **"dark energy" (p. 313).**
- The nature of dark energy is unknown. It may be described by Einstein's **cosmological constant (Λ) (p. 312),** or its strength may change with time, which cosmologists label **"quintessence."** Some models of quintessence predict an ever-accelerating expansion leading to a **"big rip" (p. 313)** that eventually would tear all the objects, even the atoms, of the universe apart.
- Corrected for acceleration, the observed value of the Hubble constant implies that the universe is 13.7 billion years old.
- The sudden inflation of the universe is thought to have magnified tiny quantum mechanical fluctuations in the density of matter and energy. These very wide but very weak differences in density caused dark matter, followed by baryonic matter, to draw together and produce the present-day **large-scale structure (p. 314)** consisting of galaxy **superclusters**

(p. 314) arranged in great walls and **filaments (p. 314)** outlining enormous **voids (p. 314).**

- Statistical observations of the large-scale structure of the universe confirm that it is flat and contains 4 percent baryonic matter, 23 percent dark matter, and 73 percent dark energy.
- The mass equivalent of dark energy added to dark matter and baryonic matter makes the observed density of the universe equal to the critical density, thereby confirming the prediction made by inflation theory that the universe is flat.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds.**

1. How does the darkness of the night sky tell you something important about the universe?
2. How can Earth be located at the center of the observable universe if you accept the cosmological principle?
3. Why can't an open universe have a center? How can a closed universe not have a center?
4. What evidence shows that the universe is expanding? What evidence shows that it began with a big bang?
5. Why couldn't atomic nuclei exist when the age of the universe was less than 2 minutes?
6. Why are measurements of the present density of the universe important?
7. How does the inflationary universe theory resolve the flatness problem? How does it resolve the horizon problem?
8. If the Hubble constant were really 100 instead of 70 km/s/Mpc, much of what astronomers understand about the evolution of stars and star clusters would have to be wrong. Explain why. (*Hint:* What would the age of the universe be?)
9. What is the evidence that the universe was very uniform during its first million years?
10. What is the difference between hot dark matter and cold dark matter? What difference does that make to cosmology?
11. What evidence can you cite that the expansion of the universe is accelerating?
12. What evidence can you cite that the universe is flat?
13. **How Do We Know?** Reasoning by analogy often helps make complicated systems or abstract ideas easier to understand. Why do you have to be careful when using analogies?
14. **How Do We Know?** The word *believe* has a meaning in normal conversation that does not really apply to scientific work. What is the difference between understanding the big bang and believing in it?

Discussion Questions

1. Do you think Copernicus would have accepted the cosmological principle? Why or why not?
2. If you reject any model of the universe that has an edge in space because you can't comprehend such a thing, shouldn't you also reject any model of the universe that has a beginning? Isn't a beginning like an "edge" in time, or is there a difference?

Problems

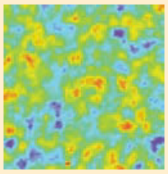
1. Use the data in Figure 15-4 to plot a velocity–distance diagram, calculate the Hubble constant *H,* and estimate the Hubble time.
2. If a galaxy is 9.0 Mpc away from Earth and recedes at 513 km/s, what is *H?* What is the Hubble time? How old would the universe be assuming space-time is flat and there has been no acceleration of the universe's expansion? How would acceleration change your answer?
3. What was the wavelength of maximum intensity for radiation from the gas of the big bang at the time of recombination? By what factor is that different from the wavelength of maximum intensity of the cosmic microwave background radiation observed now?
4. Pretend that galaxies are spaced evenly, 2 Mpc apart, and the average mass of a galaxy is 10^{11} solar masses. What is the average density of matter in the universe? (*Hints:* The volume of a sphere is $\frac{4}{3}\pi r^3$, and the mass of the sun is 2×10^{30} kg.)
5. Figure 15-12 is based on an assumed Hubble constant of 70 km/s/Mpc. How would you change the diagram to fit a Hubble constant of 50 km/s/Mpc?
6. Hubble's first estimate of the Hubble constant was 530 km/s/Mpc. His distances were too small by a factor of about 7 because of a calibration error. If he had not had that calibration problem, what value for *H* would he have obtained?
7. What was the maximum age of the universe predicted by Hubble's first estimate of the Hubble constant (in Problem 6)?

Learning to Look

1. Explain why some of the galaxies in this photo have elongated, slightly curved images. What do such observations tell you about the universe?

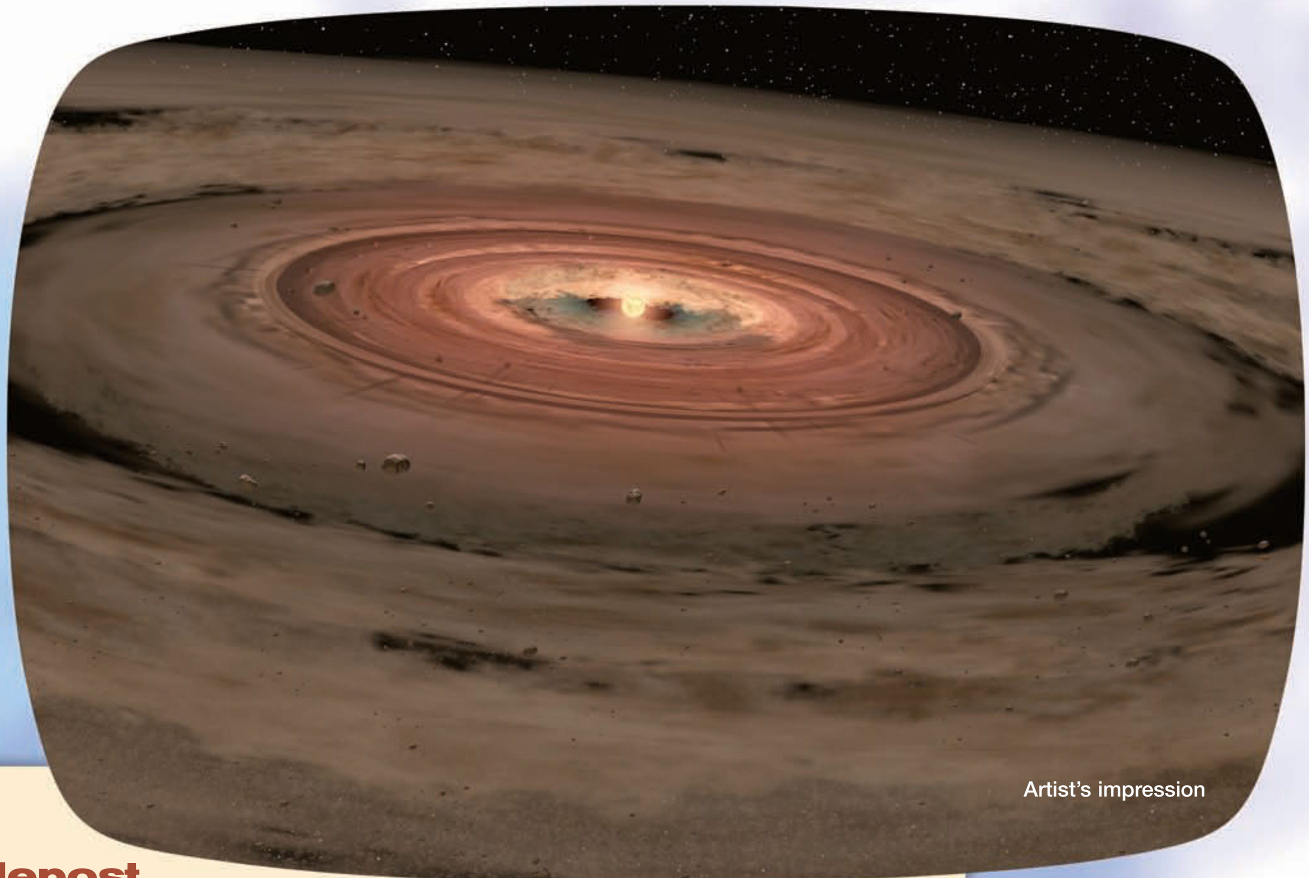
NASA, ACS Science Team, and ESA

2. The image at the right shows irregularities in the background radiation. Why isn't the background radiation perfectly uniform? What does the size of these irregularities tell you?



BOOMERANG Collaboration

16 The Origin of the Solar System



The human race lives on a planet in a planetary system that formed in a nebula around the protostar that became the sun. This artist's impression shows such a nebular disk around a forming star. (NASA/JPL)

Guidepost

You have studied the appearance, origin, structure, and evolution of stars, galaxies, and the universe itself. So far, though, your studies have left out one important type of object—planets. Now it is time for you to correct that omission. In this chapter, once you learn the evidence for how the solar system formed, you can understand how the processes you have been studying produced Earth, your home planet.

As you explore our solar system in space and time, you will find answers to three essential questions:

What is the theory for the origin of the solar system?

What are the observed properties of the solar system that the theory of its origin can explain?

What do astronomers know about other planetary systems?

In the next three chapters you will explore in more detail each of the planets, plus asteroids, comets, and meteoroids. By studying the origin of the solar system before studying the individual objects in it, you give yourself a better framework for understanding these fascinating worlds.

Animated! This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.

What place is this?
Where are we now?

CARL SANDBURG, "GRASS"

MICROSCOPIC CREATURES LIVE in the roots of your eyelashes. Don't worry. Everyone has them, and they are harmless.* They hatch, fight for survival, mate, lay eggs, and die in the tiny spaces around the roots of your eyelashes without doing any harm. Some live in renowned places—the eyelashes of a glamorous movie star, for example—but the tiny beasts are not self-aware; they never stop to say, "Where are we?"

You can study the solar system for many reasons. You can study Earth and its sibling planets because, as you are about to discover, there are almost certainly more planets in the universe than stars. Above all, you can study the solar system because it is your home in the universe. Because humans are an intelligent species, we have the ability and the responsibility to wonder where we are and what we are. Our kind have inhabited this solar system for at least a million years, but only within the last few hundred years have we begun to understand what the solar system is.

16-1 The Great Chain of Origins

YOU ARE LINKED through a great chain of origins that leads backward through time to the instant when the universe began, 13.7 billion years ago. The gradual discovery of the links in that chain is one of the most exciting adventures of the human intellect. In earlier chapters, you studied some of that story: the origin of the universe in the big bang, the formation of galaxies, the origin of stars, and the growth of the chemical elements. Now you will explore further to consider the origin of planets.

The History of the Atoms in Your Body

The universe began in the big bang (Chapter 15). By the time the universe was 3 minutes old, the protons, neutrons, and electrons in your body had come into existence. You are made of very old matter.

Although those particles formed quickly, they were not linked together to form many of the atoms that are common today. Most of the matter in the early universe was hydrogen, and about 25 percent was helium. Although your body does not contain helium, it does contain many of those ancient hydrogen atoms unchanged since the universe began. Evidence indicates that almost no atoms heavier than helium were made in the big bang.

**Demodex folliculorum* has been found in 97 percent of individuals and is a characteristic of healthy skin.

Within a few hundred million years after the big bang, matter began to collect to form galaxies containing billions of stars. You have learned how nuclear reactions inside stars combine low-mass atoms such as hydrogen to make heavier atoms (Chapter 7). Generation after generation of stars cooked the original particles, fusing them into atoms such as carbon, nitrogen, and oxygen that are common in your body (Chapters 9 and 12). Even the calcium atoms in your bones were assembled inside stars.

Massive stars produce iron in their cores, but much of that iron is lost when the core collapses and the star explodes as a supernova. Most of the iron on Earth and in your body was produced instead by carbon fusion in type Ia supernova explosions and by the decay of radioactive atoms in the expanding matter ejected by type II supernovae. Atoms heavier than iron such as gold, silver, and iodine are created by rapid nuclear reactions that can occur only during supernova explosions (Chapter 10). Iodine is critical to the function of your thyroid gland, and you probably have gold and silver jewelry or dental fillings. Realize that these types of atoms, which are part of your life on Earth, were made during the violent deaths of massive stars long ago.

Our galaxy contains at least 100 billion stars, of which the sun is one. Astronomers have a variety of evidence that the sun formed from a cloud of gas and dust about 5 billion years ago, and the atoms in your body were part of that cloud. How the sun took shape, how the cloud gave birth to the planets, how the atoms in your body found their way onto Earth and into you, is the story of this chapter. As you explore the origin of the solar system, keep in mind the great chain of origins that created the atoms. As the geologist Preston Cloud remarked, "Stars have died that we might live."

The Origin of the Solar System

Over the last two centuries, astronomers have proposed two kinds of hypotheses for the origin of the planets. Catastrophic hypotheses proposed that the planets formed from some improbable event such as the collision of the sun and another star. Evolutionary hypotheses proposed that the planets formed gradually and naturally as the sun formed. Since about 1940, the evidence has become overwhelming for the evolutionary scenario (■ How Do We Know 16-1). In fact, the evolutionary hypothesis is so comprehensive and explains so many of the observations that it can be considered to have "graduated" from being just a hypothesis to being properly called a theory. Today, astronomers are continuing to refine the details of that theory.

The **solar nebula theory** supposes that planets form in the rotating disks of gas and dust around young stars (■ Figure 16-1). You already have seen clear evidence that disks of gas and dust are common around young stars (Chapter 9). Bipolar flows from protostars were the first evidence of such disks, but modern techniques can image the disks directly (■ Figure 16-2; also, see page 171). Our own planetary system formed in such a disk-

How Do We Know? | 16-1

Two Kinds of Theories: Catastrophic and Evolutionary

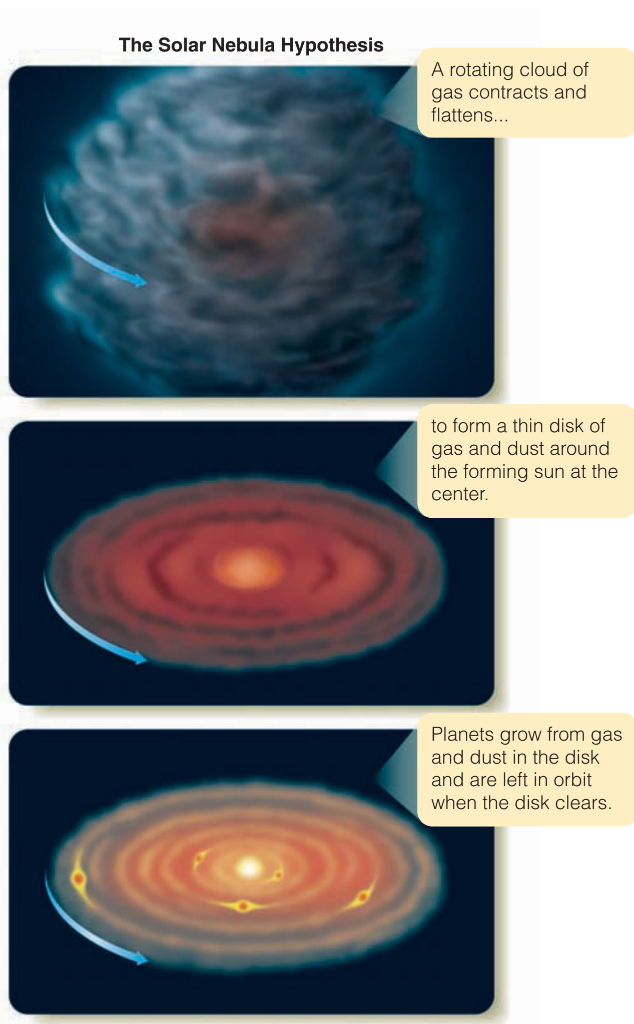
How big a role have sudden, catastrophic events played in the history of the solar system? Many theories in science can be classified as either evolutionary, in that they involve gradual processes, or catastrophic, in that they depend on specific, unlikely events. Scientists have generally preferred evolutionary theories. Nevertheless, catastrophic events do occur.

Some people prefer catastrophic theories, perhaps because they like to see spectacular violence from a safe distance, which may explain the success of movies that include lots of car crashes and explosions. Also, catastrophic theories resonate with scriptural accounts of cataclysmic events and special acts of creation. Thus, many people have an interest in catastrophic theories.

Nevertheless, most scientific theories are evolutionary. Such theories do not depend on unlikely events or special acts. For example, geologists study theories of mountain building that are evolutionary and describe mountains being pushed up slowly as millions of years pass. The evidence of erosion and the folded rock layers show that the process is gradual. Because most such natural processes are evolutionary, scientists sometimes find it difficult to accept any theory that depends on catastrophic events.

You will see in this and later chapters that catastrophes do occur. The planets, for example, are bombarded by debris from space, and some of those impacts are very large. As you study astronomy or any other natural science, notice that most theories are evolutionary but that you need to allow for the possibility of unpredictable catastrophic events.

Mountains ascend to great heights by rising slowly, not catastrophically. (Janet Seeds)



shaped cloud around the sun. When the sun became luminous enough, the remaining gas and dust were blown away into space, leaving the planets orbiting the sun.

According to the solar nebula theory, Earth and the other planets of the solar system formed billions of years ago as the sun condensed from the interstellar medium. If planet formation is a natural part of star formation, most stars should have planets.

◄ SCIENTIFIC ARGUMENT ►

Why does the solar nebula theory imply planets are common?

Often, the implications of a theory are more important in building a scientific argument than the theory itself. The solar nebula theory is an evolutionary theory; and, if it is correct, the planets of our solar system formed from the disk of gas and dust that surrounded the sun as it condensed from the interstellar medium. That suggests it is a common process. Most stars form with disks of gas and dust around them, and planets should form in such disks. Planets should therefore be very common in the universe.

Now build a new scientific argument. **Why would a catastrophic hypothesis for the formation of the solar system suggest that planets are not common?**

◄ ►

Figure 16-1

The solar nebula hypothesis proposes that the planets formed with the sun from the same spinning cloud of interstellar material.

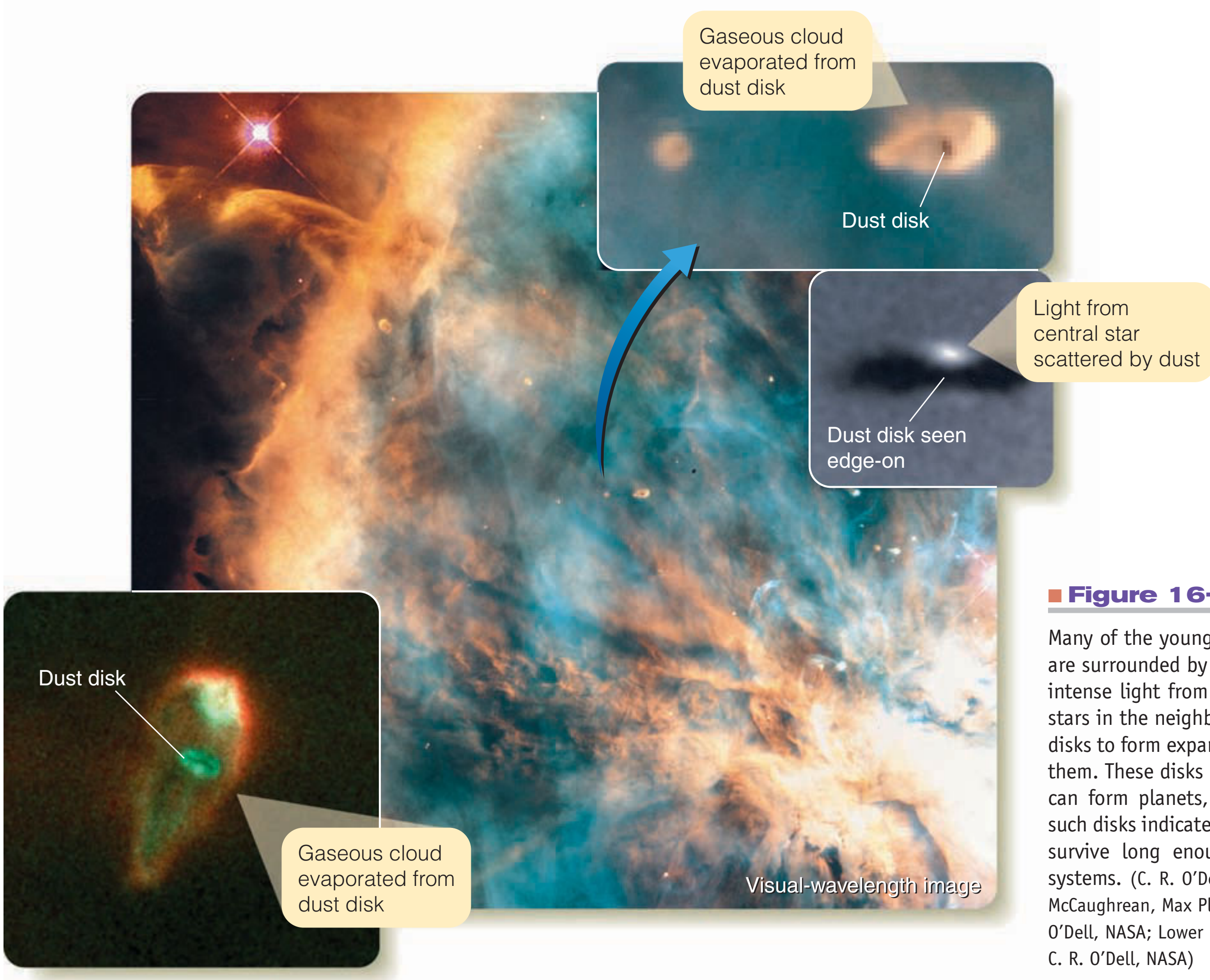


Figure 16-2

Many of the young stars in the Orion nebula are surrounded by disks of gas and dust, but intense light from the brightest and hottest stars in the neighborhood is evaporating the disks to form expanding clouds of gas around them. These disks may evaporate before they can form planets, but the large number of such disks indicates that some disks probably survive long enough to become planetary systems. (C. R. O'Dell, Rice, NASA; Dark Disk: M. McCaughrean, Max Plank Inst. for Astronomy, C. R. O'Dell, NASA; Lower left inset: J. Bally, H. Throop, C. R. O'Dell, NASA)

16-2 A Survey of the Solar System

To TEST THEIR hypotheses about how the solar system was born, astronomers searched the present solar system for evidence of its past. In this section you will survey the solar system and compile a list of its most significant characteristics that are potential clues to how it formed.

You can begin with the most general view of the solar system. It is, in fact, almost entirely empty space (look back to Figure 1-7). Imagine reducing the scale of the solar system until Earth is the size of a grain of table salt, about 0.3 mm (0.01 in.) in diameter. The moon is a speck of pepper about 1 cm (0.4 in.) away, and the sun is the size of a small plum located 4 m (13 ft) from Earth. Jupiter is an apple seed 20 m (66 ft) from the sun. Neptune, at the edge of the planetary zone, is a large grain of sand over 120 m (400 ft) from the central plum. Although your model solar system would be larger than a football field, you would need a powerful microscope to detect the asteroids orbiting between Mars and Jupiter. The planets are tiny specks of matter scattered around the sun—the last remains of the solar nebula.

Revolution and Rotation

The planets revolve* around the sun in orbits that lie close to a common plane. The orbit of Mercury, the closest planet to the sun, is tipped 7° to Earth's orbit. The rest of the planets' orbital planes are inclined by no more than 3.4°. As you can see, the solar system is basically flat and disk shaped.

The rotation of the sun and planets on their axes also seems related to the rotation of the disk. The sun rotates with its equator inclined only 7° to Earth's orbit, and most of the other planets' equators are tipped less than 30°. The rotations of Venus and Uranus are peculiar, however. Venus rotates backward compared with the other planets, whereas Uranus rotates on its side (with its equator almost perpendicular to its orbit). You will explore these planets in detail in Chapters 17 and 18, but later in this chapter you will be able to understand how they might have acquired their peculiar rotations.

*Recall from Chapter 2 that the words *revolve* and *rotate* refer to different types of motion. A planet revolves around the sun but rotates on its axis. Cowboys in the old west didn't carry revolvers. They actually carried rotators. And you don't rotate your tires every 6 months—you actually revolve them.

There is a preferred direction of motion in the solar system—counterclockwise as seen from the north, like the curl of the fingers of your right hand if you point your thumb toward your eyes. All the planets revolve counterclockwise around the sun; and, with the exception of Venus and Uranus, they rotate counterclockwise on their axes. Furthermore, nearly all of the moons in the solar system, including Earth's moon, orbit around their respective planets counterclockwise. With only a few exceptions, most of which are understood, revolution and rotation in the solar system follow a single theme. Apparently, these motions today are related to the original rotation of a disk of solar system construction material.

Two Kinds of Planets

Perhaps the most striking clue to the origin of the solar system comes from the obvious division of the planets into two groups: the small Earth-like worlds and the giant Jupiter-like worlds. The difference is so dramatic that you are led to say, "Aha, this must mean something!" Study ■ **Terrestrial and Jovian Planets** on pages 326–327 and notice three important points and two new terms:

1. The two kinds of planets are distinguished by their mass and location. The four inner *Terrestrial planets* are quite different from the four outer *Jovian planets.*
2. Craters are common. Almost every solid surface in the solar system is covered with craters.
3. The two groups of planets are also distinguished by properties such as presence or absence of rings and numbers of moons. A theory of the origin of the planets needs to explain these properties.

The division of the planets into two groups is a clue to how our solar system formed. The present properties of individual planets, however, don't tell everything you need to know about their origins. The planets have all evolved since they formed. For further clues about the origin of the planets, you can look at smaller objects that have remained largely unchanged since soon after the birth of the solar system.

Space Debris

The sun and planets are not the only products of the solar nebula. The solar system is littered with three kinds of space debris: asteroids, comets, and meteoroids. Although these objects represent a tiny fraction of the mass of the system, they are a rich source of information about the origin of the planets.

The **asteroids,** sometimes called minor planets, are small rocky worlds, most of which orbit the sun in a belt between the orbits of Mars and Jupiter. More than 100,000 asteroids have orbits that are charted, of which about 2000 follow paths that bring them into the inner solar system where they can potentially collide with a planet. Earth has been struck many times in its history. Some asteroids share Jupiter's orbit, while others have been found beyond the orbit of Saturn.

About 200 asteroids are more than 100 km (60 mi) in diameter, and tens of thousands are estimated to be more than 10 km (6 mi) in diameter. There are probably a million or more that are larger than 1 km (0.6 mi) and billions that are smaller. Because the largest are only a few hundred kilometers in size, Earth-based telescopes can detect no details on their surfaces, and even the Hubble Space Telescope can image only the largest features.

Spectroscopic observations indicate that asteroid surfaces are a variety of rocky and metallic materials. Photographs returned by robotic spacecraft show that asteroids are generally irregular in shape and covered with craters (■ Figure 16-3). These observations will be discussed in detail in Chapter 19, but in this quick survey of the solar system you can conclude that the solar nebula included elements that compose rock and metals and also that collisions have played an important role in the solar system's history.

Astronomers recognize the asteroids as debris left over from the failure of a planet to form at a distance of about 3 AU from the sun. A good theory should explain why a planet didn't form there but instead left behind a belt of rubble.

Since 1992, astronomers have discovered more than a thousand small, dark, icy bodies orbiting in the outer fringes of the solar system beyond Neptune. This collection of objects is called the **Kuiper belt** after astronomer Gerard Kuiper (KYE-per), who predicted their existence in the 1950s. There are probably 100 million bodies larger than 1 km in the Kuiper belt, many more than in the asteroid belt, and a successful theory should also explain how they came to be where they are.

In contrast to the rocky asteroids and dark Kuiper belt objects, the brightest **comets** can be seen with the naked eye and are impressively beautiful objects (■ Figure 16-4). Most comets are faint, however, and are difficult to locate even at their brightest. A comet may take months to sweep through the inner solar system, during which time it appears as a glowing ball with an extended tail of gas and dust.

The beautiful tail of a comet can be longer than an AU, but it is produced by a nucleus only a few tens of kilometers in diameter. The nuclei of comets are ice-rich bodies, sometimes described as "dirty snowballs," left over from the origin of the planets. From this you can conclude that at least some parts of the solar nebula had abundant icy material.

A comet nucleus remains frozen and inactive while it is far from the sun. As the nucleus moves along its elliptical orbit into the inner solar system, the sun's heat begins to vaporize the ices, releasing gas and dust. The pressure of sunlight and the solar wind push the gas and dust away, forming a long tail. The motion of the nucleus along its orbit, the effects of sunlight, and the outward flow of the solar wind can create comet tails that are long and straight or gently curved, but in any case the tail of a comet always points approximately away from the sun (Figure 16-4b), no matter what direction the comet itself is moving.

Unlike the stately comets, **meteors** flash across the sky in momentary streaks of light (■ Figure 16-5). They are commonly

■ **Figure 16-3**

(a) Over a period of three weeks, the NEAR spacecraft approached the asteroid Eros and recorded a series of images arranged here in an entertaining pattern showing the irregular shape and 5-hour rotation of the asteroid. Eros is 34 km (21 mi) long. (b) This close-up of the surface of Eros shows an area about 11 km (7 mi) from top to bottom. (Johns Hopkins University, Applied Physics Laboratory, NASA)

called "shooting stars." Of course, they are not stars but small bits of rock and metal colliding with Earth's atmosphere and bursting into incandescent vapor because of friction with the air about 80 km (50 mi) above the ground. This vapor condenses to form dust that settles slowly to Earth, adding about 40,000 tons per year to the planet's mass.

Technically, the word *meteor* refers to the streak of light in the sky. In space, before its fiery plunge, the object is called a **meteoroid,** and any part of it that survives its fiery passage to Earth's surface is called a **meteorite.** Most meteoroids are specks of dust, grains of sand, or tiny pebbles. Almost all the meteors you see in the sky are produced by meteoroids that weigh less than 1 gram. Only rarely is a meteoroid massive and strong enough to survive its plunge and reach Earth's surface.

Thousands of meteorites have been found, and you will learn more about their various types in Chapter 19. Meteorites are mentioned here for one specific clue they can give you concerning the solar nebula: Meteorites can reveal the age of the solar system.

The Age of the Solar System

According to the solar nebula theory, the planets should be about the same age as the sun. The most accurate way to find the age of a rocky body is to bring a sample into the laboratory and analyze the radioactive elements it contains.

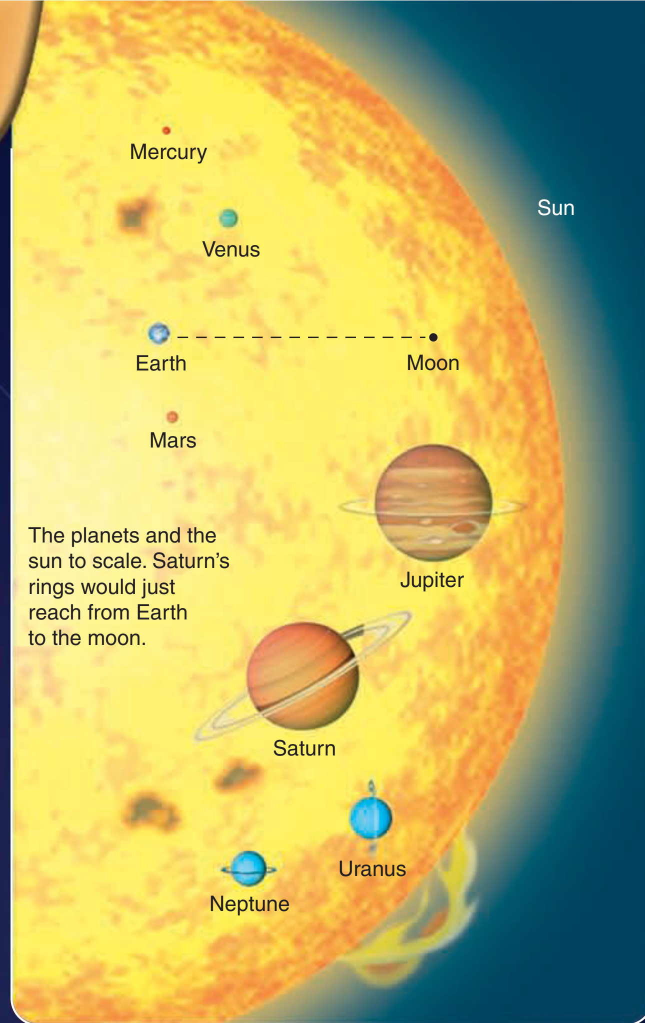
When a rock solidifies, it incorporates known percentages of the chemical elements. A few of these elements have forms, called isotopes (Chapter 6), that are radioactive, meaning they gradually decay into other isotopes. For example, potassium-40, called a parent isotope, decays into calcium-40 and argon-40, called daughter isotopes. The **half-life** of a radioactive substance is the time it takes for half of the parent isotope atoms to decay into daughter isotope atoms. The abundance of a radioactive substance gradually decreases as it decays, and the abundances of the daughter substances gradually increase (■ Figure 16-6). The half-life of potassium-40 is 1.3 billion years. If you also have information about the abundances of the elements in the original rock, you can measure the present abundances and find the age of the rock. For example, if you study a rock and find that only 50 percent of the potassium-40 remains and the rest has become a mixture of daughter isotopes, you could conclude that one half-life must have passed and that the rock is 1.3 billion years old.

Potassium isn't the only radioactive element used in radioactive dating. Uranium-238 decays with a half-life of 4.5 billion years to form lead-206 and other isotopes. Rubidium-87 decays to strontium-87 with a half-life of 47 billion years. Any of these substances can be used as a radioactive clock to find the age of mineral samples.

Of course, to find a radioactive age, you need to get a sample into the laboratory, and the only celestial bodies of which scientists have samples are Earth, the moon, Mars, and meteorites. The oldest Earth rocks so far discovered and dated are tiny zircon crystals from Australia that are 4.3 billion years old. That does not mean that Earth formed 4.3 billion years ago. The surface of Earth is active, and the crust is continually destroyed and re-formed with material welling up from beneath the crust (see Chapter 17). Those types of processes tend to dilute the daughter atoms and spread them away from the parent atoms, effectively causing the radioactive clocks to reset to zero. The radioactive age of a rock is actually the length of time since the material in that rock was last melted. Consequently, the dates of these oldest rocks tells you only a lower limit to the age of Earth, in other words, that Earth is *at least* 4.3 billion years old.

Terrestrial and Jovian Planets

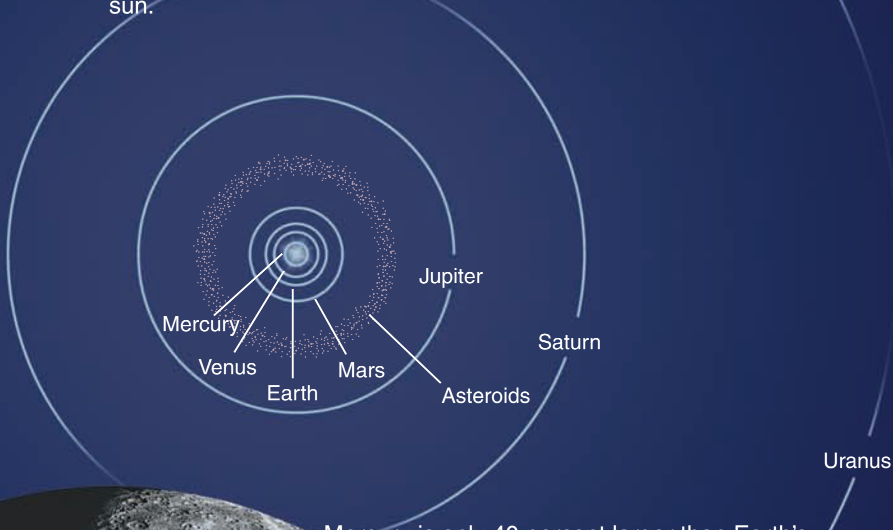
1 The distinction between the Terrestrial planets and the Jovian planets is dramatic. The inner four planets, Mercury, Venus, Earth, and Mars, are **Terrestrial planets,** meaning they are small, dense, rocky worlds with little or no atmosphere. The outer four planets, Jupiter, Saturn, Uranus, and Neptune, are **Jovian planets,** meaning they are large, low-density worlds with thick atmospheres and liquid interiors.

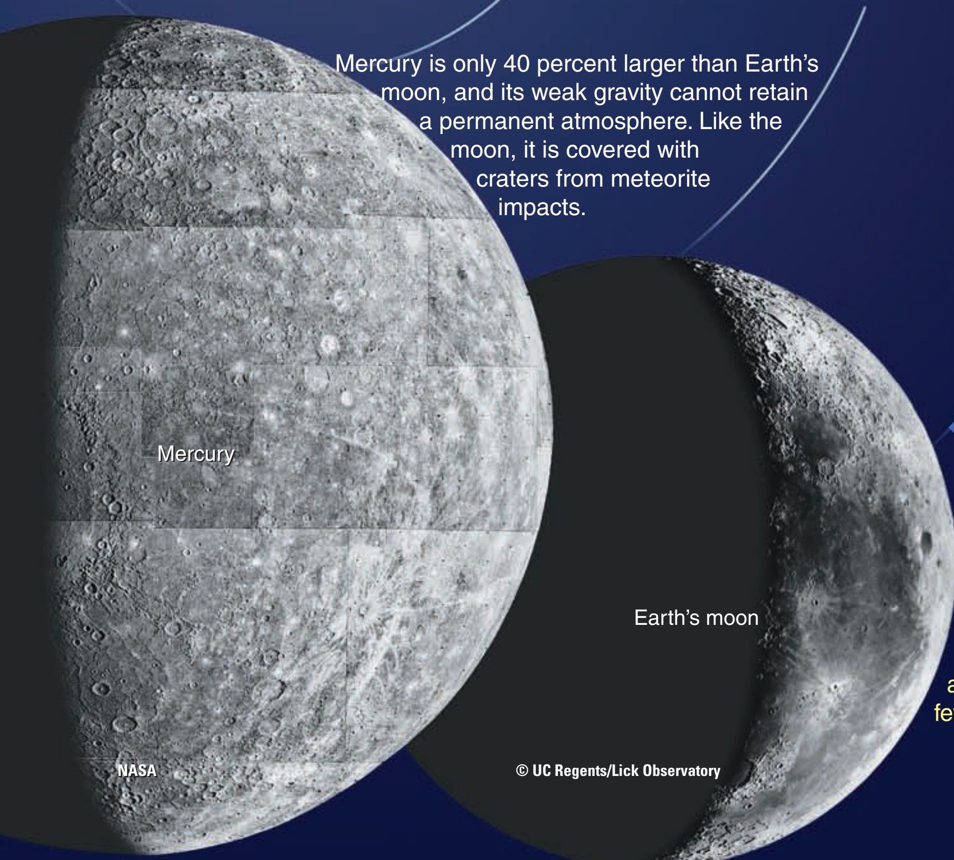


The planets and the sun to scale. Saturn's rings would just reach from Earth to the moon.

1a Of the Terrestrial planets, Earth is most massive, but the Jovian planets are much more massive. Jupiter is over 300 Earth masses, and Saturn is nearly 100 Earth masses. Uranus and Neptune are 15 and 17 Earth masses.

Planetary orbits to scale. The Terrestrial planets lie quite close to the sun, whereas the Jovian planets are spread far from the sun.



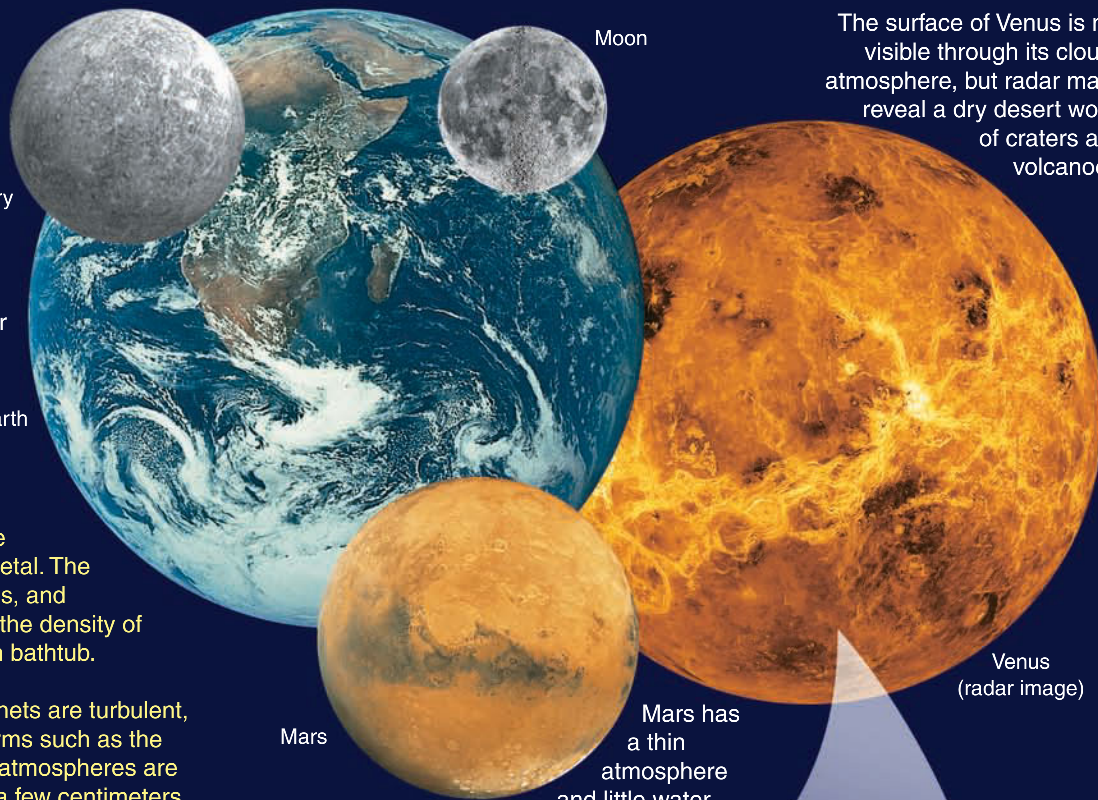


Mercury is only 40 percent larger than Earth's moon, and its weak gravity cannot retain a permanent atmosphere. Like the moon, it is covered with craters from meteorite impacts.

2 Craters are common on all of the surfaces in the solar system that are strong enough to retain them. Earth has about 150 impact craters, but many more have been erased by erosion. Besides the planets, the asteroids and nearly all of the moons in the solar system are scarred by craters. Ranging from microscopic to hundreds of kilometers in diameter, these craters have been produced over the ages by meteorite impacts. When astronomers see a rocky or icy surface that contains few craters, they know that the surface is young.

Mercury is so close to the sun it is difficult to study from Earth. The Mariner 10 and MESSENGER spacecraft flew past Mercury in 1974 and 2008, respectively, and were able to take detailed close-up photos of the planet's surface.

These five worlds are shown in proper relative size.



The surface of Venus is not visible through its cloudy atmosphere, but radar maps reveal a dry desert world of craters and volcanoes.

Moon, © UC Regents/Lick Observatory; all planets, NASA

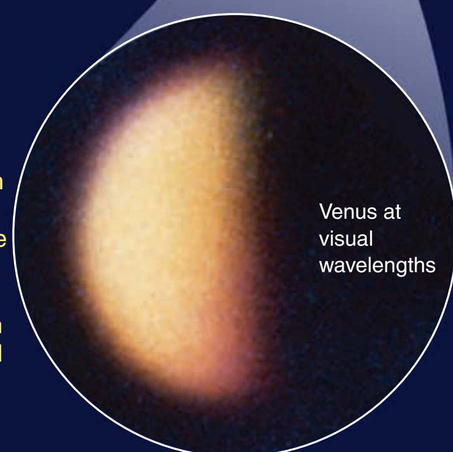
3 The Terrestrial planets have densities like that of rock or metal. The Jovian planets all have low densities, and Saturn's density is only 70 percent the density of water. It would float in a big-enough bathtub.

The atmospheres of the Jovian planets are turbulent, and some are marked by great storms such as the Great Red Spot on Jupiter, but the atmospheres are not deep. If Jupiter were shrunk to a few centimeters in diameter, its atmosphere would be no deeper than the fuzz on a badly worn tennis ball.

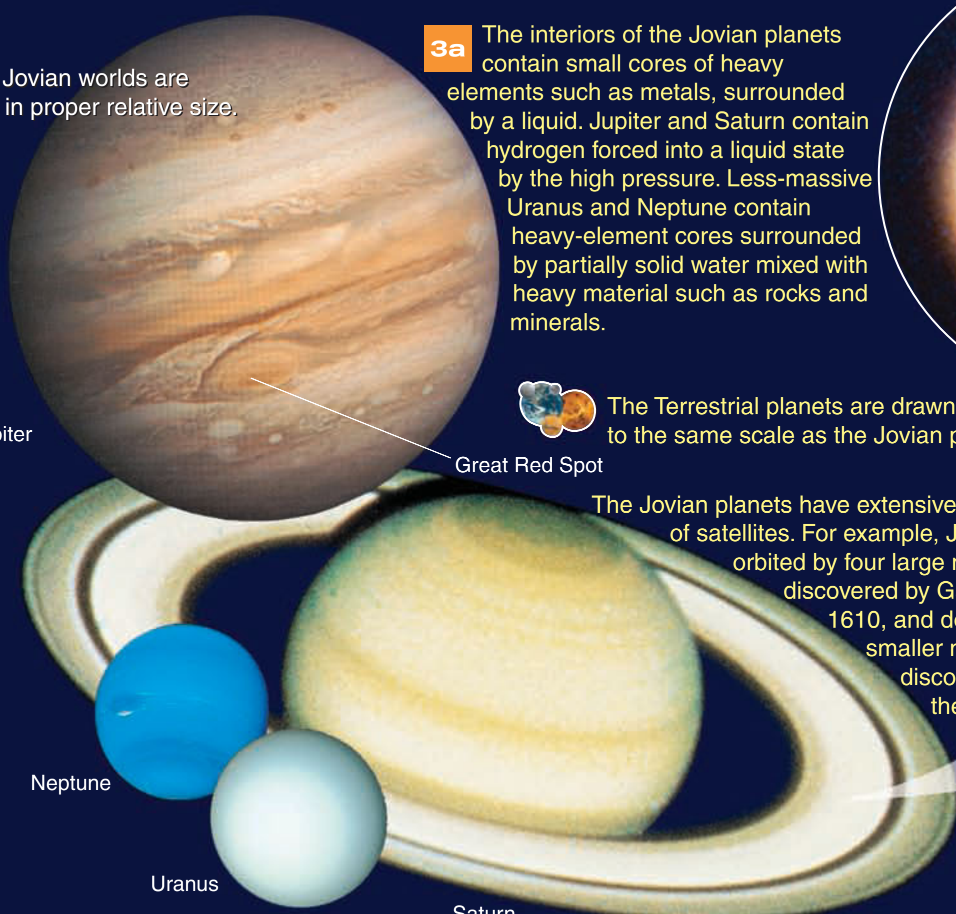
Mars has a thin atmosphere and little water. Craters and volcanoes are common on its desert surface.

These Jovian worlds are shown in proper relative size.

3a The interiors of the Jovian planets contain small cores of heavy elements such as metals, surrounded by a liquid. Jupiter and Saturn contain hydrogen forced into a liquid state by the high pressure. Less-massive Uranus and Neptune contain heavy-element cores surrounded by partially solid water mixed with heavy material such as rocks and minerals.



The Terrestrial planets are drawn here to the same scale as the Jovian planets.

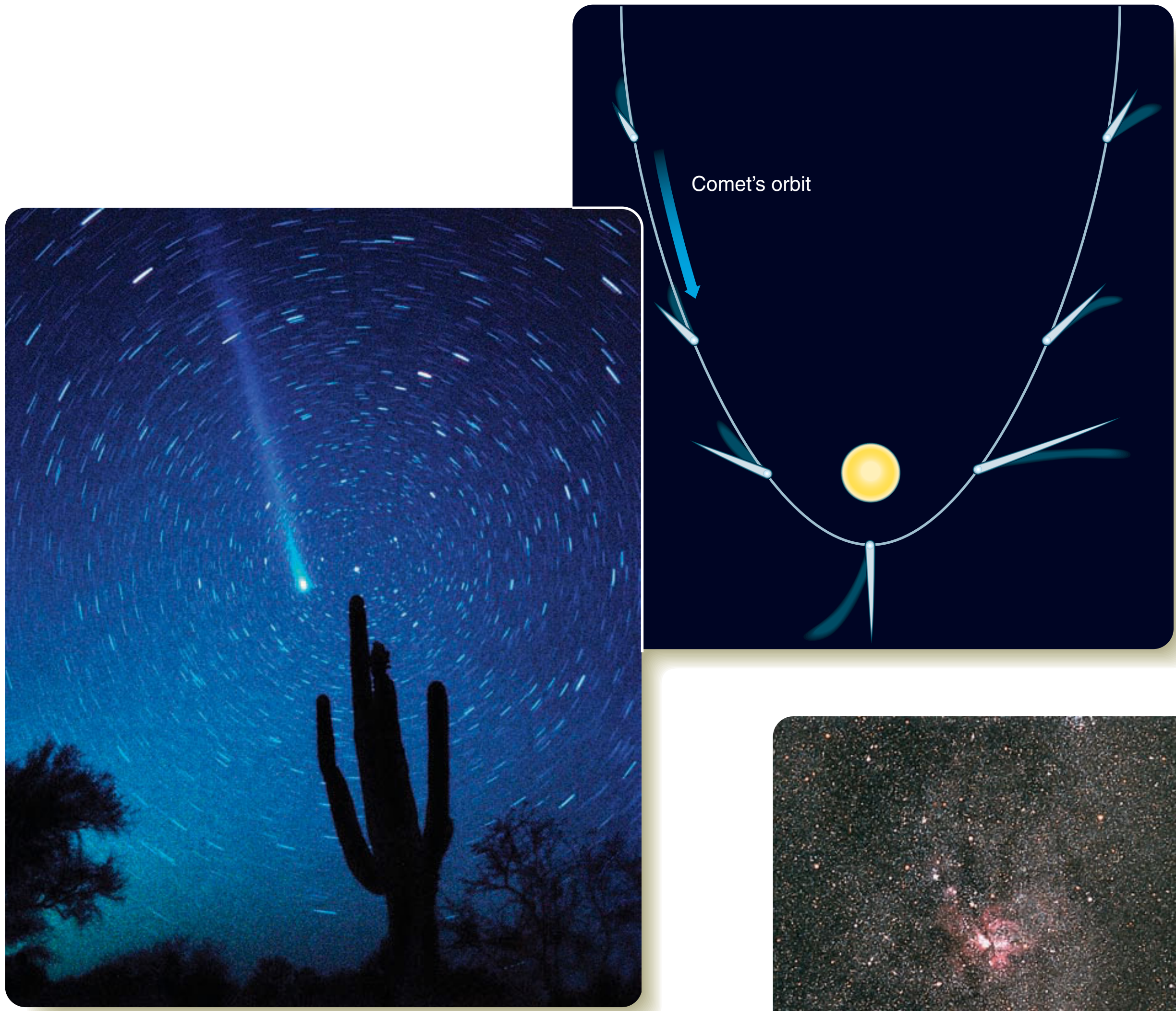


NASA

The Jovian planets have extensive systems of satellites. For example, Jupiter is orbited by four large moons discovered by Galileo in 1610, and dozens of smaller moons discovered up to the present day.

Grundy Observatory

3b All four Jovian planets have ring systems. Saturn's rings are made of ice particles. The rings of Jupiter, Uranus, and Neptune are made of dark rocky particles. Terrestrial planets have no rings.



a

b

■ **Figure 16-4**

(a) A comet may remain visible in the evening or morning sky for weeks as it moves through the inner solar system. Although comets are moving rapidly along their orbits, they are so distant that, at any particular moment, a comet seems to hang motionless in the sky. Comet Hyakutake is shown here near Polaris in 1996. (Kent Wood/Photo Researchers, Inc.) (b) A comet in a long, elliptical orbit becomes visible when the sun's heat vaporizes its ices and pushes the gas and dust away in a tail that points away from the sun.

One of the most exciting goals of the Apollo lunar landings was to bring lunar rocks back to Earth's laboratories where their ages could be measured. Because the moon's surface is not geologically active like Earth's surface, some moon rocks might have survived unaltered since early in the history of the solar system. In fact, the oldest moon rocks are 4.48 billion years old. That means the moon must be *at least* 4.48 billion years old.

Although no one has yet been to Mars, over a dozen meteorites found on Earth have been identified by their chemical composition as having come from Mars. Most of these have ages of only a billion years or so, but one has an age of approximately 4.5 billion years. Mars must be at least that old.

■ **Figure 16-5**

A meteor is the streak of glowing gases produced by a small bit of solid material colliding with Earth's atmosphere. Friction with the air vaporizes the material about 80 km (50 mi) above Earth's surface. (Daniel Good)

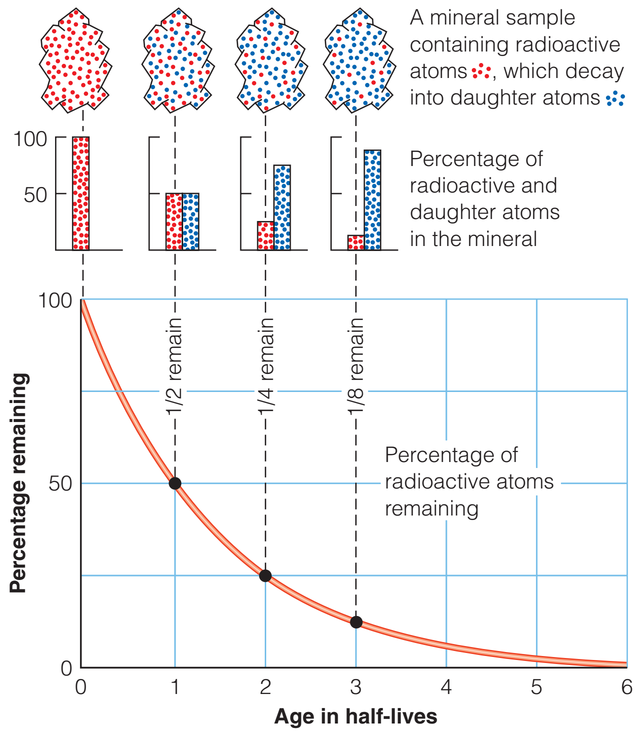


Figure 16-6

The radioactive atoms in a mineral sample (red) decay into daughter atoms (blue). Half the radioactive atoms are left after one half-life, a fourth after two half-lives, an eighth after three half-lives, and so on. **Animated!**

The most important source for determining the age of the solar system is meteorites. Radioactive dating of meteorites yields a range of ages, but there is a fairly precise upper limit — many meteorite samples have ages of 4.56 billion years old, and none are older. That figure is widely accepted as the age of the solar system and is often rounded to 4.6 billion years. The true ages of Earth, the moon, and Mars are also assumed to be 4.6 billion years, although no rocks from those bodies have yet been found that have remained unaltered for that entire stretch of time.

One last celestial body deserves mention: the sun. Astronomers estimate the age of the sun to be about 5 billion years, but that is not a radioactive date because we have no samples of radioactive material from the sun. Instead, an independent estimate for the age of the sun can be made using helioseismological observations and mathematical models of the sun's interior (Chapters 7 and 9). This yields a value of about 5 billion years, plus or minus 1.5 billion years, a number that is in agreement with the age of the solar system derived from the age of meteorites.

Apparently, all the bodies of the solar system formed at about the same time, some 4.6 billion years ago. You can add this as the final item to your list of characteristic properties of the solar system (■ Table 16-1).

Table 16-1 | Characteristic Properties of the Solar System

1. Disk shape of the solar system
   - Orbits in nearly the same plane
   - Common direction of rotation and revolution
2. Two planetary types
   - Terrestrial — inner planets; high density
   - Jovian — outer planets; low density
3. Planetary rings and large satellite systems
   - Yes for Jupiter, Saturn, Uranus, and Neptune
   - No for Mercury, Venus, Earth, and Mars
4. Space debris — asteroids, comets, and meteors
   - Composition, orbits
   - Asteroids in inner solar system, composition like Terrestrial planets
   - Comets in outer solar system, composition like Jovian planets
5. Common age of about 4.6 billion years measured or inferred for Earth, the moon, Mars, meteorites, and the sun

◄ SCIENTIFIC ARGUMENT ►

In what ways does the solar system resemble a disk?

Notice that this argument is really a summary of pieces of evidence. First, the general shape of the solar system is that of a disk. The orbit of Mercury is inclined 7° to the plane of Earth's orbit, and the rest of the planets are in orbits inclined less than that. In other words, the planets are confined to a thin disk with the sun at its center.

Second, the motions of the sun and planets also follow this disk theme. The sun and most of the planets rotate in the same direction, counterclockwise as seen from the north, with their equators near the plane of the solar system. Also, all of the planets revolve around the sun in that same direction. The objects in our solar system mostly move in the same direction, which further reflects a disk theme.

One of the basic characteristics of our solar system is its disk shape, but another dramatic characteristic is the division of the planets into two groups. Build an argument to detail that evidence. **What are the distinguishing differences between the Terrestrial and Jovian planets?**

◄ ►

16-3 The Story of Planet Building

The challenge for modern planetary scientists is to compare the characteristics of the solar system with predictions of the solar nebula theory, so they can work out details of how the planets formed (■ How Do We Know? 16-2).

The Chemical Composition of the Solar Nebula

Everything astronomers know about the solar system and star formation suggests that the solar nebula was a fragment of an interstellar gas cloud. Such a cloud would have been mostly hydrogen with some helium and small amounts of the heavier elements.

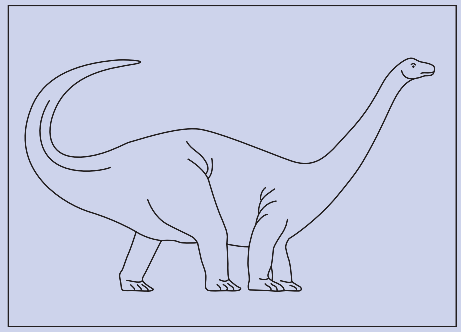
How Do We Know? | 16-2

Reconstructing the Past from Evidence and Hypothesis

How can we know how the solar system formed? Scientists often solve problems in which they must reconstruct the past. Some of these reconstructions are obvious, such as an archaeologist excavating the ruins of a burial tomb, but others are less obvious. In each case, success requires the interplay of hypotheses and evidence to re-create a past that no longer exists.

It is obvious that astronomers reconstruct the past when they use evidence gathered from meteorites to study the origin of the solar system, but a biologist studying a centipede is also reconstructing the past. How did this creature come to have a segmented body with so many legs? How did it evolve the metabolism that allows it to move quickly and hunt prey? Although the problem might at first seem to be one of mere anatomy, the scientist must reconstruct an environment that no longer exists.

The astronomer's problem is not just to understand what the planets are like but also to understand how they got that way. That means planetary scientists must look at the evidence they can see today and reconstruct a history of the solar system, a past that is quite different from the present. If you had a time machine, it would be a fantastic adventure to go back and watch the planets form. Time machines are impossible, but scientists can use the scientific method's grand interplay of evidence and hypothesis to journey back billions of years and reconstruct a past that no longer exists.



One way science can enrich and inform our lives is by re-creating a world that no longer exists.

That is precisely what you see in the composition of the sun (look back again at Table 6-2). Analysis of the solar spectrum shows that the sun is mostly hydrogen, with a quarter of its mass being helium and only about 2 percent being heavier elements. Of course, nuclear reactions have fused some hydrogen into helium, but this happens in the sun's core and has not affected the composition of its surface and atmosphere, which are the parts you can observe directly. That means the composition revealed in the sun's spectrum is essentially the composition of the gases from which the sun formed.

This must have been the composition of the solar nebula, and you can also see that composition reflected in the chemical compositions of the planets. The inner planets are composed of rock and metal, and the outer planets are rich in low-density gases such as hydrogen and helium. The chemical composition of Jupiter resembles the composition of the sun. Furthermore, if you allowed low-density gases to escape from a blob of stuff with the same overall composition as the sun or Jupiter, the relative proportions of the remaining heavier elements would resemble Earth's chemical composition.

The Condensation of Solids

An important clue to understanding the process that converted the nebular gas into solid matter is the variation in density among solar system objects. You have already noted that the four inner planets are small and have high density, resembling Earth, whereas the outermost planets are large and have low density, resembling Jupiter.

Even among the four Terrestrial planets, you will find a pattern of slight differences in density. Merely listing the observed densities of the Terrestrial planets does not reveal the pattern clearly because Earth and Venus, being more massive, have stronger gravity and have squeezed their interiors to higher densities. The **uncompressed densities**—the densities the planets would have if their gravity did not compress them, or to put it another way, the average densities of their original construction materials—can be calculated using the actual densities and masses of each planet (■ Table 16-2). In general, the closer a planet is to the sun, the higher its uncompressed density.

This density variation originated when the solar system first formed solid grains. The kind of matter that could condense in a particular region depended on the temperature of the gas there. In the inner regions, the temperature was evidently 1500 K or so. The only materials that can form grains at that temperature are compounds with high melting points, such as metal oxides and

■ **Table 16-2 | Observed and Uncompressed Densities**

Planet	Observed Density (g/cm^3)	Uncompressed Density (g/cm^3)
Mercury	5.44	5.30
Venus	5.24	3.96
Earth	5.50	4.07
Mars	3.94	3.73

pure metals, which are very dense. Farther out in the nebula it was cooler, and silicates (rocky material) could also condense, in addition to metal. These are less dense than metal oxides and metals. Mercury, Venus, Earth, and Mars are evidently composed of a mixture of metals, metal oxides, and silicates, with proportionately more metals close to the sun and more silicates farther from the sun. Even farther from the sun there was a boundary called the **ice line** beyond which water vapor could freeze to form ice particles. Yet a little farther from the sun, compounds such as methane and ammonia could condense to form other types of ice. Water vapor, methane, and ammonia were abundant in the solar nebula, so beyond the ice line the nebula would have been filled with a blizzard of ice particles, mixed with small amounts of silicate and metal particles that could also condense there. Those ices are low-density materials. The compositions of Jupiter and the other outer planets are a mix of ices plus relatively small amounts of silicates and metal.

The sequence in which the different materials condense from the gas as you move away from the sun toward lower temperature is called the **condensation sequence** (■ Table 16-3). It suggests that the planets, forming at different distances from the sun, should have accumulated from different kinds of materials in a predictable way.

People who have read a little bit about the origin of the solar system may hold the **Common Misconception** that the matter in the solar nebula was sorted by density, with the heavy rock and metal sinking toward the sun and the low-density gases being blown outward. That is not the case. The chemical composition of the solar nebula was originally roughly the same throughout the disk. The important factor was temperature: The inner nebula was hot, and only metals and rock could condense there, whereas the cold outer nebula could form lots of ices along with metals and rock. The ice line seems to have been between Mars and Jupiter, and it separates the region for formation of the high-density Terrestrial planets from that of the low-density Jovian planets.

■ **Table 16-3 | The Condensation Sequence**

Temperature (K)	Condensate	Planet (Estimated Temperature of Formation; K)
1500	Metal oxides	Mercury (1400)
1300	Metallic iron and nickel	
1200	Silicates	
1000	Feldspars	Venus (900)
680	Troilite (FeS)	Earth (600)
		Mars (450)
175	H_2O ice	Jovian (175)
150	Ammonia–water ice	
120	Methane–water ice	
65	Argon–neon ice	Pluto (65)

The Formation of Planetesimals

In the development of a planet, two processes operate to collect solid bits of matter—rock, metal, ice—into larger bodies called **planetesimals,** which eventually build the planets. The study of planet building is the study of these processes of condensation and accretion, described in this section, plus gravitational collapse, described in the next section.

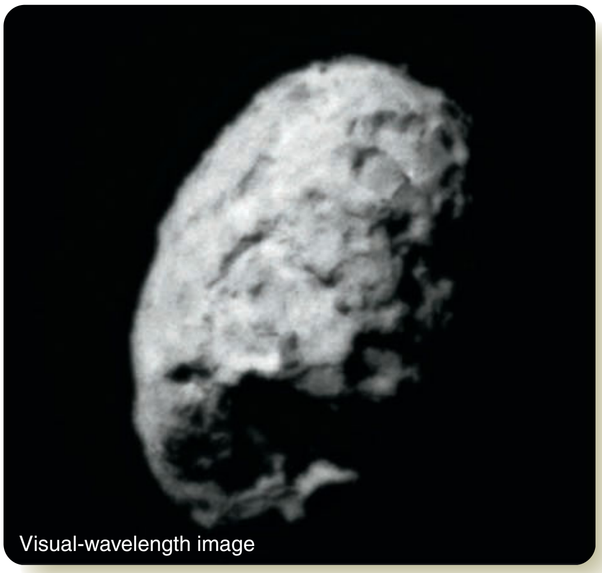
According to the solar nebula theory, planetary development in the solar nebula began with the growth of dust grains. These specks of matter, whatever their composition, grew from microscopic size first by condensation, then by accretion.

A particle grows by **condensation** when it adds matter one atom or molecule at a time from a surrounding gas. Snowflakes, for example, grow by condensation in Earth's atmosphere. In the solar nebula, dust grains were continuously bombarded by atoms of gas, and some of those stuck to the grains. A microscopic grain capturing a layer of gas molecules on its surface increases its mass by a much larger fraction than a gigantic boulder capturing a single layer of molecules. That is why condensation can increase the mass of a small grain rapidly, but, as the grain grows larger, condensation becomes less effective.

The second process is **accretion,** the sticking together of solid particles. You may have seen accretion in action if you have walked through a snowstorm with big, fluffy flakes. If you caught one of those "flakes" on your glove and looked closely, you saw that it was actually made up of many tiny, individual flakes that had collided as they fell and accreted to form larger particles. In the solar nebula, the dust grains were, on average, no more than a few centimeters apart, so they collided frequently and could accrete into larger particles.

When the particles grew to sizes larger than a centimeter, they would have been subject to new processes that tended to concentrate them. One important effect was that the growing solid objects would have collected into the plane of the solar nebula. Small dust grains could not fall into the plane because the turbulent motions of the gas kept them stirred up, but larger objects had more mass, and gas motions could not have prevented them from settling into the plane of the spinning nebula. Astronomers calculate this would have concentrated the larger solid particles into a relatively thin layer about 0.01 AU thick that would have made further growth more rapid. There is no clear distinction between a very large grain and a very small planetesimal, but you might consider an object to be a planetesimal when its diameter approaches a kilometer (0.6 mi) or so (■ Figure 16-7).

This concentration of large particles and planetesimals into the plane of the nebula is analogous to the flattening of a forming



■ **Figure 16-7**

What did the planetesimals look like? You can get a clue from this photo of the 5-km-wide nucleus of Comet Wild 2 (pronounced *Vildt-two*). Whether rocky or icy, the planetesimals must have been small, irregular bodies, scarred by craters from collisions with other planetesimals. (NASA)

galaxy, and a process also found in galaxies may have become important once the plane of planetesimals formed. Computer models show that the rotating disk of particles should have been gravitationally unstable and would have been disturbed by spiral density waves resembling the much larger ones found in spiral galaxies. Those waves could have further concentrated the planetesimals and helped them coalesce into objects up to 100 km (60 mi) in diameter.

Through these processes, the theory proposes, the nebula became filled with trillions of solid particles ranging in size from pebbles to tiny planets. As the largest began to exceed 100 km in diameter, additional accretion processes began to affect them, and a new stage in planet building began, the formation of protoplanets.

The Growth of Protoplanets

The coalescing of planetesimals eventually formed **protoplanets,** the name for massive objects destined to become planets. As these larger bodies grew, a new process helped them grow faster and altered their physical structure.

If planetesimals had collided at orbital velocities, it is unlikely that they would have stuck together. A typical orbital velocity in the solar system is about 10 km/s (22,000 mph). Head-on collisions at this velocity would vaporize the material. However, the planetesimals were all moving in the same direction in the nebular plane and didn't collide head on. Instead, they merely "rubbed shoulders," so to speak, at low relative velocities. Such gentle collisions would have been more likely to combine planetesimals than to shatter them.

Some adhesive effects probably helped. Sticky coatings and electrostatic charges on the surfaces of the smaller planetesimals probably aided formation of larger bodies. Collisions would have fragmented some of the surface rock; but, if the planetesimals were large enough, their gravity could have held on to some fragments to form a layer of soil composed of crushed rock. Such a relatively soft soil layer on the surfaces of larger planetesimals may have been effective in trapping smaller bodies.

The largest planetesimals would grow the fastest because they had the strongest gravitational field. Their stronger gravity could attract additional material, and they could also hold on to a cushioning layer to trap fragments. Astronomers calculate that the largest planetesimals would have grown quickly to protoplanetary dimensions, sweeping up more and more material.

Protoplanets had to begin growing by accumulating solid material because they did not have enough gravity to capture and hold large amounts of gas. In the warm solar nebula, the atoms and molecules of gas were traveling at velocities much larger than the escape velocities of modest-sized protoplanets. Therefore, in their early development, the protoplanets could grow only by attracting solid bits of rock, metal, and ice. Once a protoplanet approached a size of 15 Earth masses or so, however, it could begin to grow by **gravitational collapse,** the rapid accumulation of large amount of in-falling gas.

The theory of protoplanet growth into planets supposes that all the planetesimals had about the same chemical composition. The planetesimals accumulated to form a planet-sized ball of material with homogeneous composition throughout. Once the planet formed, heat would begin to accumulate in its interior from the decay of short-lived radioactive elements.

The violent impacts of in-falling particles would also have released energy called **heat of formation.** These two heating sources would eventually have melted the planet and allowed it to differentiate. **Differentiation** is the separation of material according to density. Once a planet melted, the heavy metals such as iron and nickel, plus elements chemically attracted to them, would settle to the core, while the lighter silicates and related materials floated to the surface to form a low-density crust. The story of planetesimals combining into planets that subsequently differentiated is shown in ■ Figure 16-8.

The process of differentiation depends partly on the presence of short-lived radioactive elements whose rapid decay would have released enough heat to melt the interior of planets. Astronomers know such radioactive elements were present because very old rock from meteorites contains daughter isotopes such as magnesium-26. That isotope is produced by the decay of aluminum-26 with a half-life of only 0.74 million years. The aluminum-26 and similar short-lived radioactive isotopes are gone now, but they must have been produced in a supernova explosion that occurred shortly before the formation of the solar nebula. In fact, some astronomers suspect that supernova explosions could have triggered the formation of the sun and other

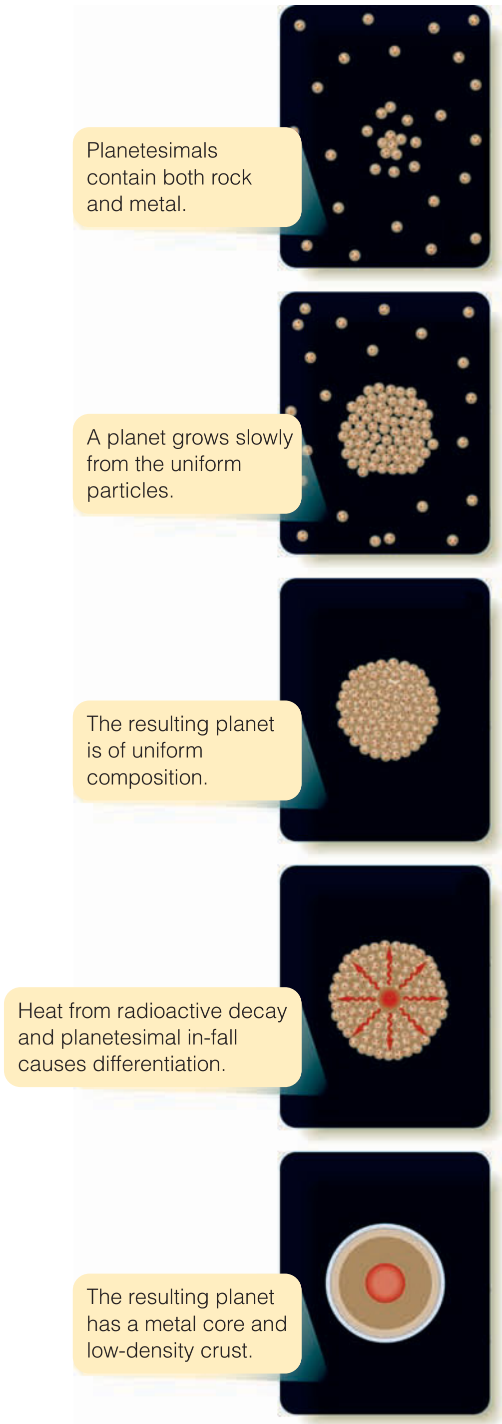


Figure 16-8

This simple model of planet building assumes planets formed from accretion and collision of planetesimals that were of uniform composition, containing both metals and rocky material, and the planets later differentiated, meaning they melted and separated into layers by density and composition.

stars by compressing interstellar clouds (see Figure 9-4). Thus, our solar system may exist because of a supernova explosion that occurred about 4.6 billion years ago.

If planets formed by accretion of planetesimals and were later melted by radioactive decay and heat of formation, then Earth's early atmosphere may have consisted of a combination of gases delivered by planetesimal impacts and released from the planet's interior during differentiation. The creation of a planetary atmosphere from a planet's interior is called **outgassing.** Given the location of Earth in the solar nebula, gases released from its interior during differentiation would not have included as much water as Earth now has. So, some astronomers think that Earth's water and even some of its present atmosphere and biosphere accumulated late in the formation of the planet as Earth swept up volatile-rich planetesimals. These icy planetesimals would have formed in the cool outer parts of the solar nebula and could have been scattered toward the Terrestrial planets by encounters with the Jovian planets, creating a comet bombardment.

According to the solar nebula theory, the Jovian planets could begin growing by the same processes that built the Terrestrial planets. However, in the inner solar nebula, only metals and silicates could form solids, so the Terrestrial planets grew slowly. In contrast, the outer solar nebula contained not just solid bits of metals and silicates but also plentiful ices. Astronomers calculate that the Jovian planets would have grown faster than the Terrestrial planets and quickly become massive enough to begin even faster growth by gravitational collapse, drawing in large amounts of gas from the solar nebula. The Terrestrial planet zone did not include ice particles, so those planets developed relatively slowly and never became massive enough to grow further by gravitational collapse.

The Jovian planets must have reached their present size in no more than about 10 million years, before the sun become hot and luminous enough to blow away the remaining gas in the solar nebula, removing the raw material for further Jovian growth. As you will learn in the next section, disturbances from outside the forming solar system may have reduced the time available for Jovian planet formation even more severely. The Terrestrial planets, in comparison, grew from solids and not from the gas, so they could have continued to grow by accretion from solid debris left behind after the gas was removed. Mathematical models indicate that the Terrestrial planets were at least half finished within 10 million years but could have continued to grow for another 20 million years or so.

The solar nebula theory has been very successful in explaining the formation of the solar system. But there are some problems, and the Jovian planets are the troublemakers.

The Jovian Problem

New information about the star formation process makes it hard to explain the formation of the Jovian planets, and this has caused astronomers to expand and revise the theory of planet formation.

The new information is that gas and dust disks around newborn stars don't last long. Earlier you saw images of dusty gas disks around the young stars in the Orion star-forming region (Figure 16-2 and Chapter 9). Those disks are being evaporated by intense ultraviolet radiation from hot O and B stars forming nearby. Astronomers have calculated that nearly all stars form in clusters containing O and B stars, so this evaporation may happen to most disks. Even if a disk did not evaporate quickly, the gravitational influence of the crowded stars in a cluster could strip away the outer parts of the disk. Those are troublesome observations because they seem to indicate that disks can't last longer than about 7 million years at most, and many evaporate within the astronomically very short span of 100,000 years or so. That's not long enough to grow a Jovian planet by the combination of condensation, accretion, and gravitational collapse proposed in the standard solar nebular theory

Yet, Jovian planets are common. In the final section of this chapter, you will see evidence that astronomers have found planets orbiting other stars, and almost all of those planets have the mass of Jovian planets. There may also be many Terrestrial planets orbiting those stars that are too small for astronomers to detect at present, but the important point is that there are lots of Jovian planets around.

Mathematical models of the solar nebula have been computed using specially built computers running programs that take weeks to finish a calculation. The results show that the rotating gas and dust of the solar nebula could have become unstable and formed Jovian planets by skipping straight to the step of gravitational collapse. That is, massive planets may have been able to form directly from the gas without first forming a dense core by the condensation and accretion of solid material. Jupiters and Saturns can form in these direct collapse models in only a few hundred years. If the Jovian planets formed in this way, they could have formed before the solar nebula disappeared, even if the nebula was eroded quickly by neighboring massive hot stars.

This new insight into the formation of the outer planets may also help explain a puzzle about the formation of Uranus and Neptune. Those planets are so far from the sun that accretion could not have built them rapidly. The gas and dust of the solar nebula must have been sparse out there, and Uranus and Neptune orbit so slowly they would not have swept up material very rapidly. The conventional view is that they grew by accretion so slowly that they never became quite massive enough to begin accelerated growth by gravitational collapse. In fact, it is hard to understand how they could have reached even their present sizes if they started growing by accretion so far from the sun. Theoretical calculations show that they might instead have formed closer to the sun, in the region of Jupiter and Saturn, and then could have been shifted outward by gravitational interactions with the bigger planets. In any case, the formation of Uranus and Neptune is part of the Jovian problem.

The traditional solar nebula theory proposes that the planets formed by accreting a core and then, if they became massive enough, accelerated growth by gravitational collapse. The proposed modification to the theory suggests that the outer planets could have skipped the core accretion phase.

Explaining the Characteristics of the Solar System

Now you have learned enough to put all the pieces of the puzzle together and explain the distinguishing characteristics of the solar system in Table 16-1.

The disk shape of the solar system is inherited from the motion of material in the solar nebula. The sun and planets and moons mostly revolve and rotate in the same direction because they formed from the same rotating gas cloud. The orbits of the planets lie in the same plane because the rotating solar nebula collapsed into a disk, and the planets formed in that disk.

The solar nebula hypothesis is evolutionary in that it calls on continuing processes to gradually build the planets. To explain the odd rotations of Venus and Uranus, however, you may need to consider catastrophic events. Uranus rotates on its side. This might have been caused by an off-center collision with a massive planetesimal when the planet was nearly formed. Two hypotheses have been proposed to explain the backward rotation of Venus. Theoretical models suggest that the sun can produce tides in the thick atmosphere of Venus that could have eventually reversed the planet's rotation—an evolutionary hypothesis. It is also possible that the rotation of Venus was altered by an off-center impact late in the planet's formation, and that is a catastrophic hypothesis. Both may be true.

The second item in Table 16-1, the division of the planets into Terrestrial and Jovian worlds, can be understood through the condensation sequence. The Terrestrial planets formed in the inner part of the solar nebula, where the temperature was high and only compounds such as the metals and silicates could condense to form solid particles. That produced the small, dense Terrestrial planets. In contrast, the Jovian planets formed in the outer solar nebula, where the lower temperature allowed the gas to form large amounts of ices, perhaps three times more ices than silicates. That allowed the Jovian planets to grow rapidly and became massive, low-density worlds. Also, Jupiter and Saturn are so massive they were able to grow by drawing in cool gas directly from the solar nebula. The Terrestrial planets could not do this because they never became massive enough.

The heat of formation (the energy released by in-falling matter) was tremendous for these massive planets. Jupiter must have grown hot enough to glow with a luminosity of about 1 percent that of the present sun, although it never got hot enough to generate nuclear energy as a star would. Nevertheless, Jupiter is still hot inside. In fact, both Jupiter and Saturn radiate more heat than they absorb from the sun, so they are evidently still cooling.

A glance at the solar system suggests that you should expect to find a planet between Mars and Jupiter at the present location of the asteroid belt. Mathematical models indicate that the reason asteroids are there, rather than a planet, is that Jupiter grew into such a massive body that it was able to gravitationally disturb the motion of nearby planetesimals. The bodies that could have formed a planet just inward from Jupiter's orbit instead collided at high speeds and shattered rather than combining, were thrown into the sun, or ejected from the solar system. The asteroids seen today are the last remains of those rocky planetesimals.

The comets, in contrast, are evidently the last of the icy planetesimals. Some may have formed in the outer solar nebula beyond Neptune and Pluto, but many probably formed among the Jovian planets where ices could condense easily. Mathematical models show that the massive Jovian planets could have ejected some of these icy planetesimals into the far outer solar system. In a later chapter, you will see evidence that some comets are icy bodies coming from those distant locations, falling back into the inner solar system.

The icy Kuiper belt objects appear to be ancient planetesimals that formed in the outer solar system but were never incorporated into a planet. They orbit slowly far from the light and warmth of the sun and, except for occasional collisions, have not changed much since the solar system was young. The gravitational influence of the planets deflects Kuiper belt objects into the inner solar system where they also are seen as comets.

The large satellite systems of the Jovian worlds may contain two kinds of moons. Some moons may have formed in orbit around forming planets in a miniature version of the solar nebula. Some of the smaller moons, in contrast, may be captured planetesimals, asteroids, and comets. The large masses of the Jovian planets would have made it easier for them to capture satellites.

In Table 16-1, you noted that all four Jovian worlds have ring systems, and you can understand this by considering the large mass of these worlds and their remote location in the solar system. A large mass makes it easier for a planet to hold onto orbiting ring particles; and, being farther from the sun, the ring particles are not as quickly swept away by the pressure of sunlight and the solar wind. It is hardly surprising, then, that the Terrestrial planets, low-mass worlds located near the sun, have no planetary rings.

The last entry in Table 16-1 is the common ages of solar system bodies, and the solar nebula hypothesis has no difficulty explaining that characteristic. If the hypothesis is correct, then the planets formed at the same time as the sun and should have roughly the same age.

Clearing the Nebula

The sun probably formed along with many other stars in a swirling nebula. Observations of young stars in Orion (Figure 16-2) suggest that radiation from nearby hot stars would have evaporated the disk of gas and dust around the sun and that the gravitational influence of nearby stars would have pulled gas away. Even without the external effects, four internal processes would have gradually destroyed the solar nebula.

The most important of these internal processes was **radiation pressure.** When the sun became a luminous object, light streaming from its photosphere pushed against the particles of the solar nebula. Large bits of matter like planetesimals and planets were not affected, but low-mass specks of dust and individual atoms and molecules were pushed outward and eventually driven from the system.

The second effect that helped clear the nebula was the solar wind, the flow of ionized hydrogen and other atoms away from the sun's upper atmosphere. This flow is a steady breeze that rushes past Earth at about 400 km/s (250 mi/s). Young stars have even stronger winds than stars of the sun's age and also irregular fluctuations in luminosity, like those observed in young stars such as T Tauri stars, which can accelerate the wind. The strong surging wind from the young sun may have helped push dust and gas out of the nebula.

The third effect that helped clear the nebula was the sweeping up of space debris by the planets. All of the old, solid surfaces in the solar system are heavily cratered by meteorite impacts (■ Figure 16-9). Earth's moon, Mercury, Venus, Mars, and most of the moons in the solar system are covered with craters. A few of these craters have been formed recently by the steady rain of meteorites that falls on all the planets in the solar system, but most of the craters appear to have been formed roughly 4 billion years ago in what is called the **heavy bombardment,** as the last of the debris in the solar nebula were swept up by the planets.

The fourth effect was the ejection of material from the solar system by close encounters with planets. If a small object such as a planetesimal passes close to a planet, the small object's path will be affected by the planet's gravity field. In some cases, the small object can gain energy from the planet's motion and be thrown out of the solar system. Ejection is most probable in encounters with massive planets, so the Jovian planets were probably very efficient at ejecting the icy planetesimals that formed in their region of the nebula.

Attacked by the radiation and gravity of nearby stars and racked by internal processes, the solar nebula could not survive very long. Once the gas and dust were gone and most of the planetesimals were swept up, the planets could no longer gain significant mass, and the era of planet building ended.

◄ SCIENTIFIC ARGUMENT ►

Why are there two kinds of planets in our solar system?

This is an opportunity for you to build an argument that closely analyzes the solar nebula theory. Planets begin forming from solid bits of matter, not from gas. Consequently, the kind of planet that forms at a given distance from the sun depends on the kind of substances that can condense out of the gas there to form solid particles. In the inner parts of the solar nebula, the temperature was so high that most of the gas could not condense to form solids. Only metals and silicates could form solid grains, and the

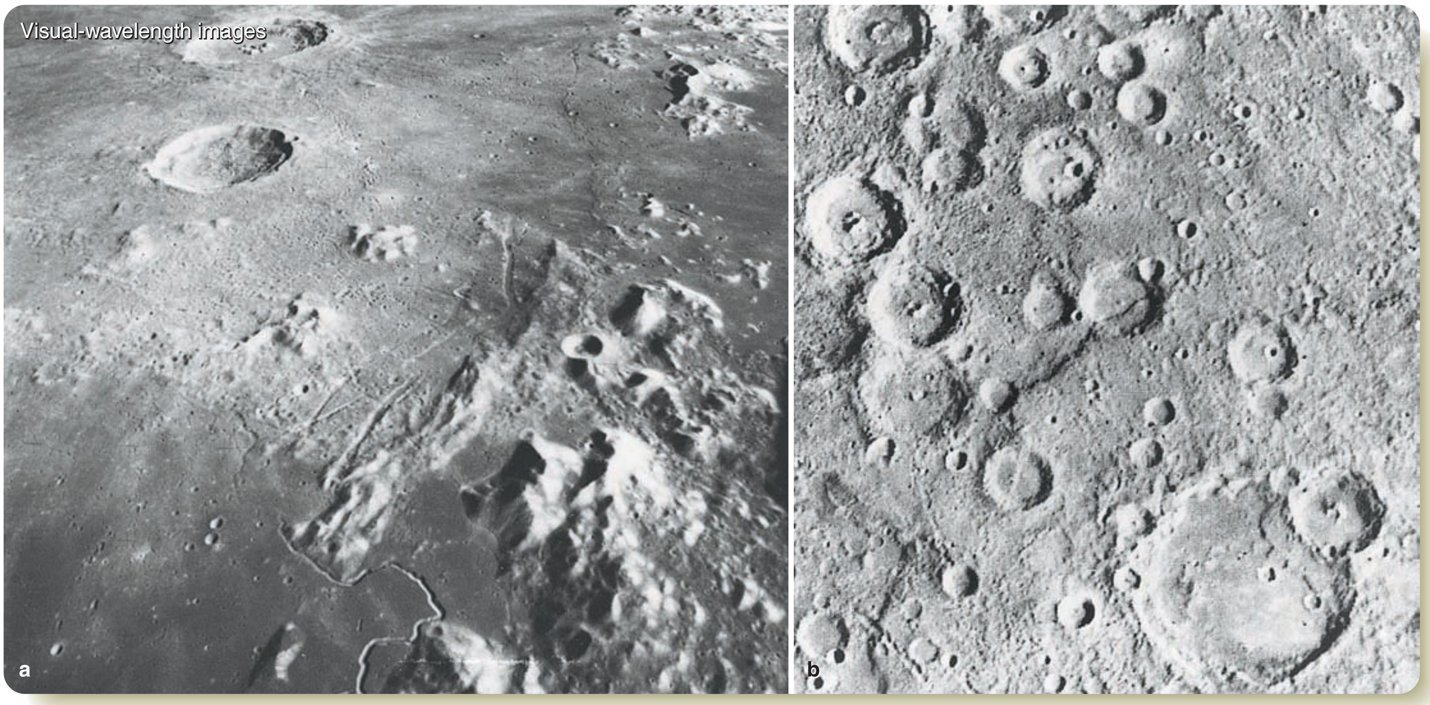


Figure 16-9

Every old, solid surface in the solar system is scarred by craters. (a) Earth's moon is scarred by craters ranging from basins hundreds of kilometers in diameter down to microscopic pits. (b) The surface of Mercury, as photographed by the Mariner 10 spacecraft, shows vast numbers of overlapping craters. (NASA)

innermost planets grew from this dense material. Much of the mass of the solar nebula consisted of hydrogen, helium, water vapor, and other gases, and they were present in the inner solar nebula but couldn't form solid grains. The small Terrestrial planets could grow only from the solids in their zone, not from the gases, so the Terrestrial planets are small and dense.

In the outer solar nebula, the composition of the gas was the same, but it was cold enough for water vapor, and other simple molecules containing hydrogen, to condense to form ice grains. Because hydrogen was so abundant, there was lots of ice available. The outer planets grew from large amounts of ice combined with small amounts of metals and silicates. Eventually the outer planets grew massive enough that they could begin to capture gas directly from the nebula, and they became the hydrogen- and helium-rich Jovian worlds.

The condensation sequence combined with the solar nebula theory gives you a way to understand the difference between the Terrestrial and Jovian planets. Now expand your argument: **Why do some astronomers argue that the formation of the Jovian planets is a problem that needs further explanation?**

◄ ►

16-4 Planets Orbiting Other Stars

ARE THERE OTHER planetary systems? The evidence says yes. Do they contain planets like Earth? The evidence so far is incomplete.

Planet-Forming Disks Around Other Suns

Both visible- and radio-wavelength observations detect dense disks of gas orbiting young stars. For example, at least 50 percent of the stars in the Orion nebula are surrounded by dense disks of gas and dust (Figure 16-2 and page 171). A young star is visible at the center of most disks, and astronomers can measure that the disks contain many Earth masses of material in a region a few times larger in diameter than our solar system. The Orion star-forming region is only a few million years old, so planets may not have formed in these disks yet. Furthermore, the intense radiation from nearby hot stars is evaporating the disks so quickly that planets may never have a chance to grow large. The important point for astronomers is that so many of these young stars have disks. Evidently, disks of gas and dust are a common feature around stars that are forming.

The Hubble Space Telescope can detect dense disks of gas and dust around young stars in a slightly different way. The disks show up in silhouette against the nebulae that surround the newborn stars (■ Figure 16-10). These disks are related to the formation of bipolar flows (page 169) in that they focus the gas flowing away from a young star into two jets shooting in opposite directions.

In addition to these dense, hot planet-forming disks around young stars, infrared astronomers have found cold, low-density

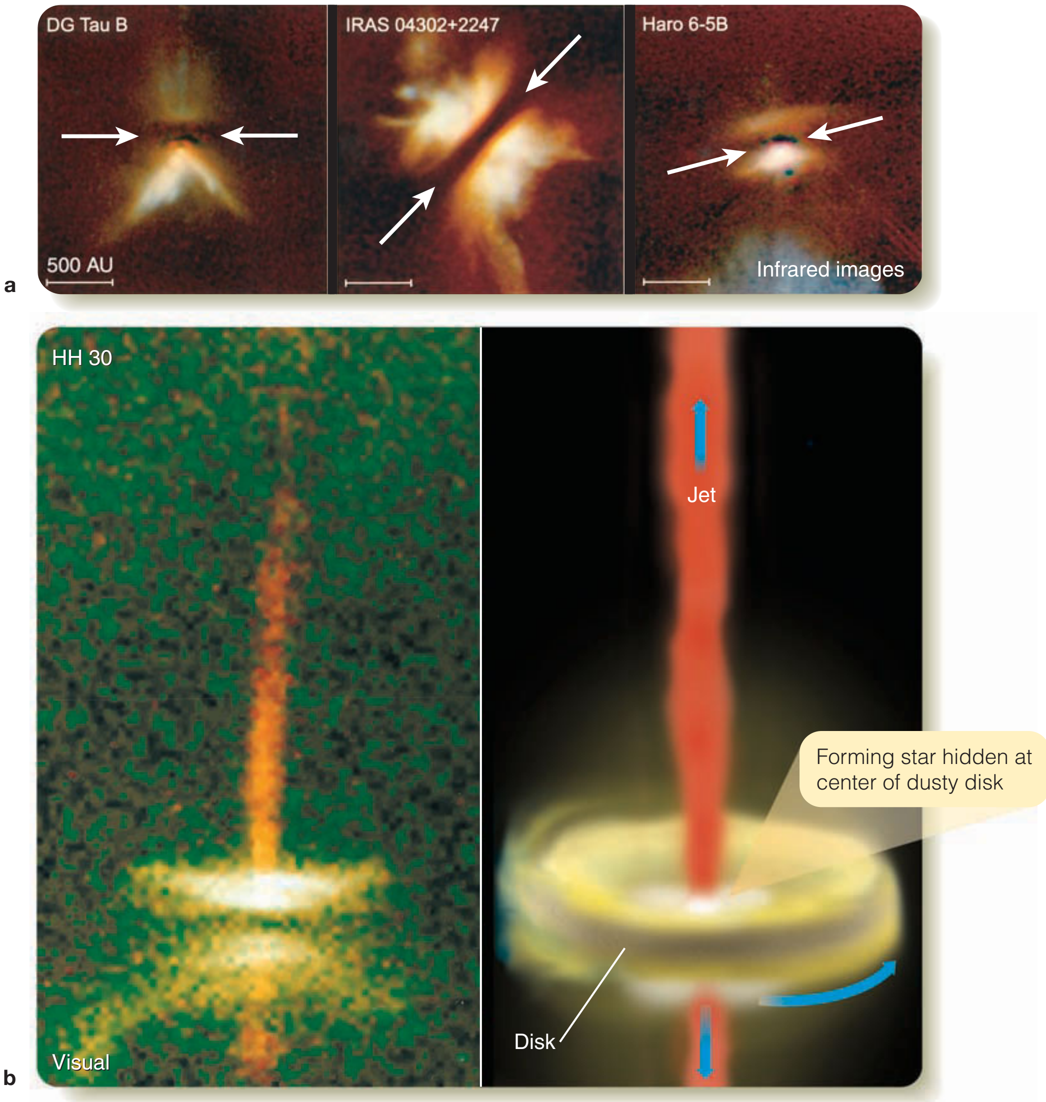


Figure 16-10

(a) Dark bands (indicated by arrows) are edge-on disks of gas and dust around young stars seen in Hubble Space Telescope near-infrared images. Planets may eventually form in these disks. These systems are so young that material is still falling inward and being illuminated by light from the stars. (D. Padgett, IPAC/Caltech; W. Brandner, IPAC; K. Stapelfeldt, JPL; and NASA) (b) The newborn star HH 30 with a dense dusty disk and bipolar jet seen in silhouette against background nebulosity, clarified in the artist's impression at right. In many cases like HH 30, the stars are so young that material is still falling inward, and interactions with the spinning star produce jets of gas being ejected in opposite directions. (C. Burrows, STScI & ESA, WFPC 2 Team, NASA)

dust disks around stars older than the newborn stars in Orion, old enough to have finished forming planets. These tenuous dust disks are sometimes called **debris disks** because they are evidently made of dusty debris produced in collisions among small bodies such as comets, asteroids, and Kuiper belt objects, rather than dust left over from an original protostellar disk. That conclusion is based on calculations showing that the observed dust would be removed by radiation pressure in a much shorter time than the ages of those stars, meaning the dust there now must have been created relatively recently. Our own solar system contains such "second-generation" dust, and astronomers have evidence that the solar system's Kuiper belt (Chapter 19) extending beyond the orbit of Neptune is an example of an old debris disk.

Some examples of debris disks are around the stars Beta Pictoris, HD 107146, and Epsilon Eridani (■ Figure 16-11). The dust disk around Beta Pictoris, an A-type star more massive and luminous than the sun, is about 20 times the diameter of our solar system. The dust disk around Epsilon Eridani, which is a K-type star somewhat smaller than the sun, is similar in size to the solar system's Kuiper belt. Like most of the other known low-density disks, both of these examples have central zones with even lower density. Those inner regions may be places where planets have finished forming and swept up most of the construction material.

Planets orbiting stars with debris disks have not yet been definitely detected, but the presence of dust with short lifetimes around old stars assures you that small bodies such as asteroids and comets must be present as sources of the dust. If those small objects are there, then it is likely that there are also planets orbiting those stars.

Infrared observations reveal that the star Vega, easily visible in the northern hemisphere summer sky, also has a debris disk, and detailed studies show that much of the dust in that disk is tiny. The pressure of the light from Vega should blow away small dust particles quickly, so astronomers conclude that the dust being observed now must have been produced by a big event like the collision of two large planetesimals within the last million years. Fragments from that collision are still smashing into each other now and then and producing more dust, continuing to enhance the debris disk. This effect has also been found in the disk around HIP 8920 and several other stars. Such smashups probably happen rarely in a dust disk, but when they happen, they make the disk very easy to detect.

Notice the difference between the two kinds of planet-related disks that astronomers have found. The low-density dust disks such as the ones around Beta Pictoris, Epsilon Eridani, and Vega are produced by dust from collisions among comets, aster-

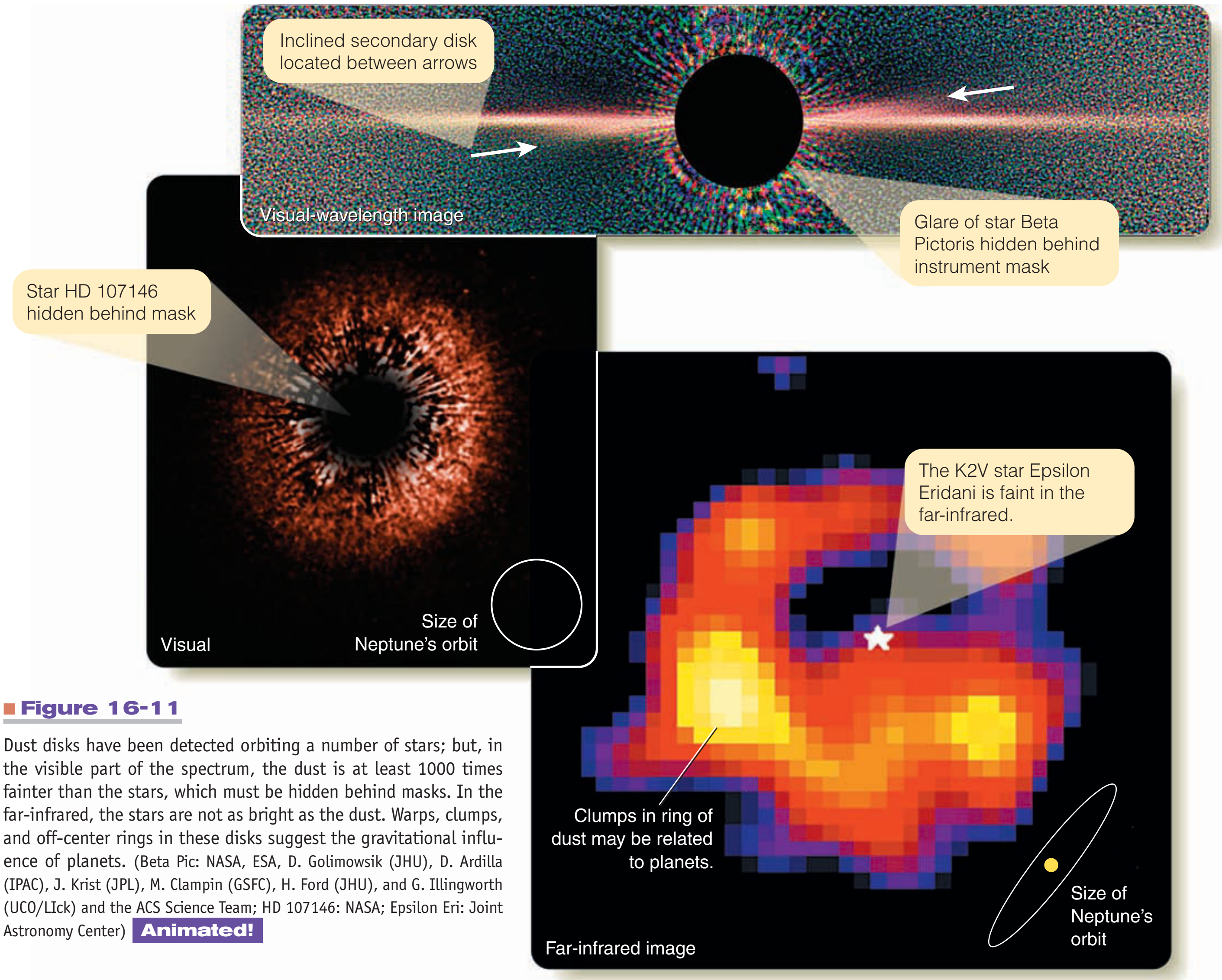


Figure 16-11

Dust disks have been detected orbiting a number of stars; but, in the visible part of the spectrum, the dust is at least 1000 times fainter than the stars, which must be hidden behind masks. In the far-infrared, the stars are not as bright as the dust. Warps, clumps, and off-center rings in these disks suggest the gravitational influence of planets. (Beta Pic: NASA, ESA, D. Golimowsik (JHU), D. Ardilla (IPAC), J. Krist (JPL), M. Clampin (GSFC), H. Ford (JHU), and G. Illingworth (UCO/LIck) and the ACS Science Team; HD 107146: NASA; Epsilon Eri: Joint Astronomy Center) **Animated!**

oids, and Kuiper belt objects. Such disks are evidence that planetary systems have already formed. In comparison, the dense disks of gas and dust such as those seen round the stars in Orion are sites where planets could be forming right now.

Observing Extrasolar Planets

A planet orbiting another star is called an **extrasolar planet.** Such a planet would be quite faint and difficult to see so close to the glare of its star. But there are ways to find these planets. To understand how, all you have to do is imagine walking a dog.

You will remember that Earth and the moon orbit around their common center of mass, and two stars in a binary system orbit around their center of mass. When a planet orbits a star, the star moves very slightly as it orbits the center of mass of the planet–star system. Think of someone walking a poorly trained dog on a leash; the dog runs around pulling on the leash, and even if it were an invisible dog, you could plot its path by watching how its owner was jerked back and forth. Astronomers can detect a planet orbiting another star by watching how the star moves as the planet tugs on it.

The first planet detected this way around a sunlike star was discovered in 1995. It orbits the star 51 Pegasi. As the planet circles the star, the star wobbles slightly, and that very small motion of the star is detectable by Doppler shifts in the star's spectrum (Figure 16-12a) (Chapters 6 and 8). From the motion of the star and estimates of the star's mass, astronomers can deduce that the planet has at least half the mass of Jupiter and orbits only 0.05 AU from the star. Half the mass of Jupiter amounts to 160 Earth masses, so this is a large planet. Note also that it orbits very close to its star, much closer than Mercury orbits around our sun.

Astronomers were not surprised by the announcement that a planet orbited 51 Pegasi; for years astronomers had assumed that many stars had planets. Nevertheless, they greeted the dis-

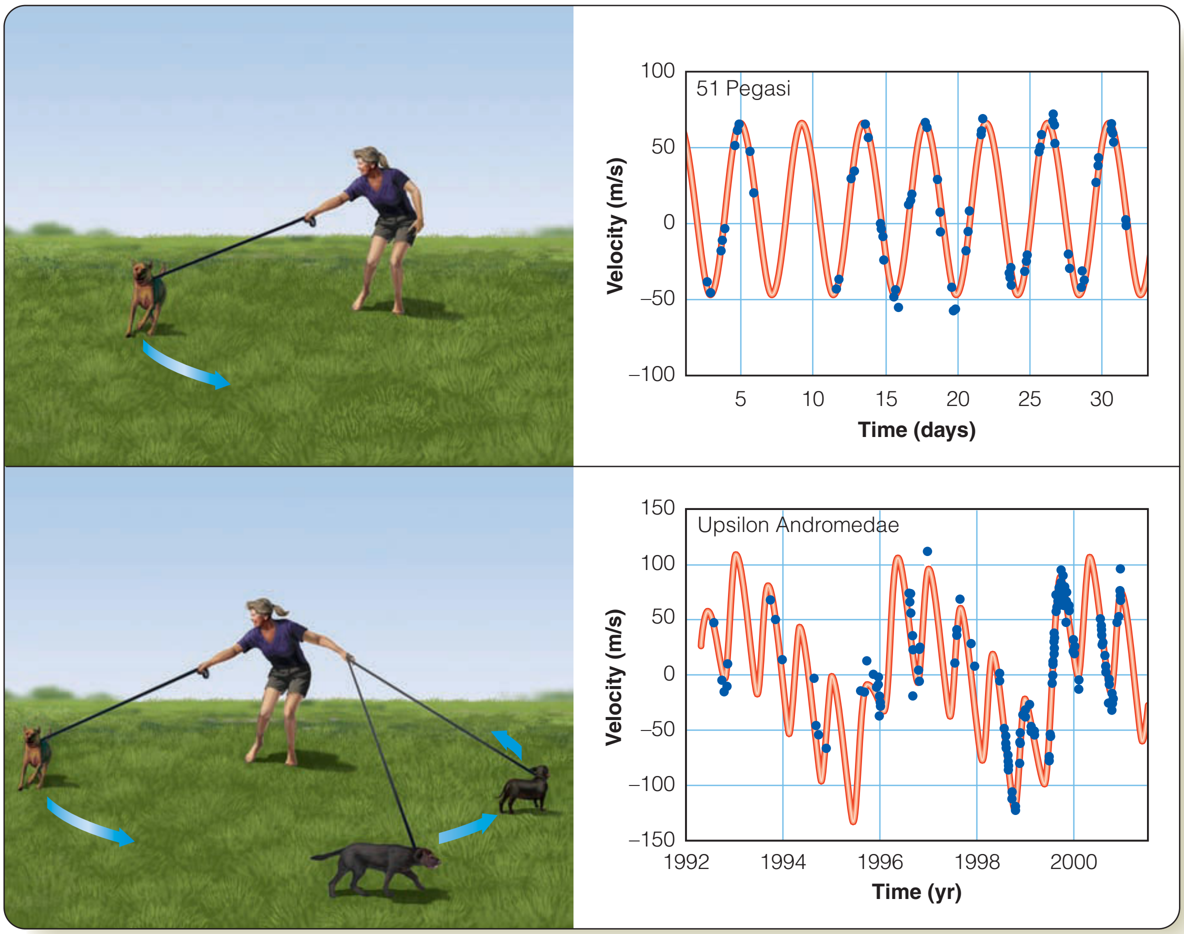


Figure 16-12

Just as someone walking a lively dog is tugged around, the star 51 Pegasi is pulled back and forth by the gravity of the planet that orbits it every 4.2 days. The wobble is detectable in precision observations of the star's Doppler shift. Someone walking three dogs is pulled about in a more complicated pattern, and you can see something similar in the Doppler shifts of the star Upsilon Andromedae, which is orbited by three planets detected so far.

covery with typical skepticism (■ How Do We Know? 16-3). That skepticism led to careful tests of the data and further observations that confirmed the discovery. In fact, more than 300 planets have been discovered in this way, including at least three planets orbiting the star Upsilon Andromedae (Figure 16-12b) and five orbiting 55 Cancri—true planetary systems. Over 25 such multiple-planet systems have been found.

Another way to search for planets is to look for changes in the brightness of the star when the orbiting planet crosses in front of the star, called a **transit**. The decrease in light during a planet transit is very small, but it is detectable, and astronomers have used this technique to find several planets. From the amount of light lost, astronomers can tell that all of these planets have Jovian sizes. The Spitzer Infrared Space Telescope has detected infrared radiation from two large hot planets already known from star wobble Doppler shifts. As these planets orbit their parent stars, the amount of infrared radiation from each system varies. When the planets pass behind their parent stars, the total infrared brightness of the systems noticeably decreases. These measurements confirm the existence of the planets and indicate their temperatures and sizes.

Notice how the techniques used to detect these planets resemble techniques used to study binary stars (Chapter 8). Most of the planets were discovered using the same observational methods used to study spectroscopic binaries, but a few were found by observing the stars as if they were eclipsing binaries.

The planets discovered so far tend to be massive and have short orbital periods because lower-mass planets or longer-period planets are harder to detect. Low-mass planets don't tug on their stars very much, and present-day spectrographs can't detect the very small velocity changes that these gentle tugs produce. Planets with longer periods are harder to detect because astronomers have not been making high-precision observations for a long enough time. Jupiter takes 11 years to circle the sun once, so it will take years for astronomers to see the longer-period wobbles produced by planets lying farther from their stars. You should not be surprised that the first planets discovered are massive and have short orbital periods.

The new planets may seem odd for another reason. In our own solar system, the large planets formed farther from the sun where the solar nebula was colder and ices could condense. How could big planets form so near their stars? Theoretical calculations indicate that planets forming in an especially dense disk of matter could spiral inward as they sweep up gas and planetesimals. That means it is possible for a few planets to become the massive, short-period objects that are detected most easily.

How Do We Know? | 16-3

Scientists: Courteous Skeptics

What does it mean to be skeptical, yet also open to new ideas? "Scientists are just a bunch of skeptics who don't believe in anything." That is a **Common Misconception** among people who don't understand the methods and goals of science. Yes, scientists are skeptical about new ideas and discoveries, but they do hold strong beliefs about how nature works. Scientists are skeptical not because they want to disprove everything but because they are searching for the truth and want to be sure that a new description of nature is reliable before it is accepted.

Another **Common Misconception** is that scientists automatically accept the work of other scientists. On the contrary, scientists skeptically question every aspect of a new discovery. They may wonder if another scientist's instruments were properly calibrated or whether the scientist's mathematical models are correct. Other scientists will want to repeat the work themselves using their own instruments to see if they can obtain the same results. Every observation is tested, every discovery is confirmed, and only an idea that survives many of these tests begins to be accepted as a scientific truth.

Scientists are prepared for this kind of treatment at the hands of other scientists. In fact, they expect it. Among scientists it is not bad manners to say, "Really, how do you know that?" or "Why do you think that?" or "Show me the evidence!" And it is not just new or surprising claims that are subject to such scrutiny. Even though astronomers had long expected to discover planets orbiting other stars, when a planet was finally discovered circling 51 Pegasi, astronomers were skeptical. This was not because they thought the observations were necessarily flawed but because that is how science works.

The goal of science is to tell stories about nature. Some people use the phrase "telling a story" to describe someone who is telling a fib. But the stories that scientists tell are exactly the opposite; perhaps you could call them antifibs, because they are as true as scientists can make them. Skepticism eliminates stories with logical errors, flawed observations, or misunderstood evidence and eventually leaves only the stories that best describe nature.

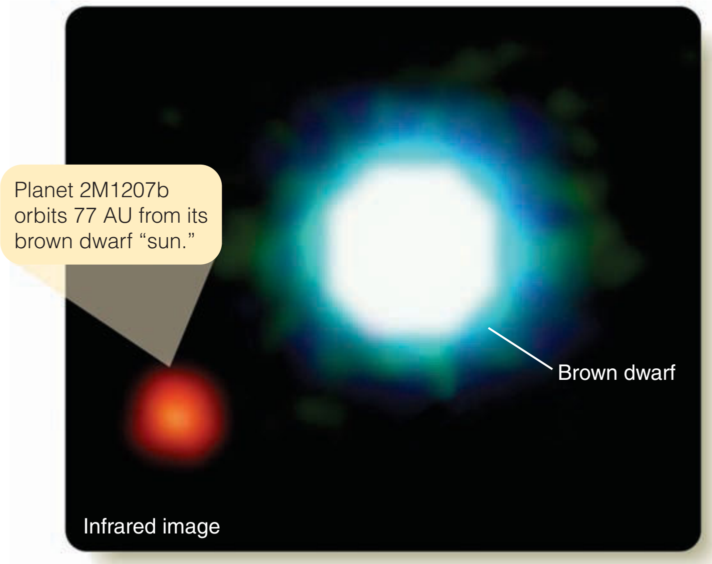
Skepticism is not a refusal to hold beliefs. Rather, it is a way for scientists to find and keep those natural principles that are worthy of belief.

Photographing a planet orbiting another star is about as easy as photographing a bug crawling on the bulb of a searchlight miles away. Planets are small and dim and get lost in the glare of the stars they orbit. Nevertheless, a few objects have been detected that appear to be planets (■ Figure 16-13). Searches for more are being conducted. For example, NASA's Kepler mission will monitor the brightness of 100,000 stars for 4 years searching for transits of Earth-sized planets. Other space observatories will be able eventually to image Jovian and Terrestrial planets around sunlike stars.

The discovery of extrasolar planets gives astronomers added confidence in the solar nebula theory. The theory predicts that planets are common, and astronomers are finding them orbiting many stars.

■ **Figure 16-13**

Infrared observations reveal a planet of about 5 Jupiter masses orbiting a brown dwarf in an orbit roughly twice as large Neptune's orbit around the sun. Spectra showing water vapor and the object's infrared colors suggest it is relatively cool and is probably a planet. (ESO)



◄ SCIENTIFIC ARGUMENT ►

Why are dust disks evidence that planets have already formed?

Sometimes a good scientific argument combines evidence, theory, and an astronomer's past experience, a kind of scientific common sense. Certainly the cold dust disks seen around stars like Vega are not places where planets are forming. They are not hot enough or dense enough to be young disks. Rather, the dust disks must be older, and the dust is being produced by collisions among comets, asteroids, and Kuiper belt objects. Small dust particles would be blown away or destroyed relatively quickly, so these collisions must be a continuing process. Astronomers have reason to believe that where you find comets, asteroids, and Kuiper belt objects, you should also find planets, so the dust disks seem to be evidence that planets have already formed in such systems.

Now build a new argument. **What direct evidence can you cite that planets orbit other stars?**

◄ ►

What Are We? Planet Walkers

The matter you are made of came from the big bang, and it has been cooked into a wide variety of atoms inside stars. Now you can see how those atoms came to be part of Earth. Your atoms were in the cloud of gas that formed the solar system 4.6 billion years ago, and nearly all of that matter contracted to form the sun, but a small amount left behind in a disk formed planets. In the process, your atoms became part of Earth.

You are a planet walker, and you have evolved to live on the surface of Earth. Are there other beings like you in the universe? Now you know that planets are common, and you can reasonably suppose that there are more planets in the universe than there are stars. However complicated the formation of the solar system was, it is a common process, so there may indeed be more planet walkers living on other worlds.

But what are those distant planets like? Before you can go very far in your search for life beyond Earth, you need to explore the range of planetary types. It is time to pack your spacesuit and voyage out among the planets of our solar system, visit them one by one, and search for the natural principles that relate planets to each other. That journey begins in the next chapter.

Study and Review

Summary

- Hypotheses for the origin of the solar system have been either catastrophic or evolutionary. Catastrophic hypotheses depend on a rare event such as the collision of the sun with another star. Evolutionary hypotheses propose that the planets formed by gradual, natural processes. The evidence now strongly favors the **solar nebula theory (p. 321)**, an evolutionary scenario.
- Modern astronomy reveals that all the matter in the universe, including our solar system, was originally formed as hydrogen and helium in the big bang. Atoms heavier than helium were cooked up in nuclear reactions in later generations of stars. The sun and planets evidently formed from a cloud of gas and dust in the interstellar medium.
- The solar nebula theory proposes that the planets formed in a disk of gas and dust around the protostar that became the sun. Observations show that these disks are common.
- The solar system is disk shaped, including the orbital revolution of the planets and their moons and the rotation of the planets on their axes.
- The planets are divided into two groups. The inner four planets are **Terrestrial planets (p. 326)**—small, rocky, dense Earth-like worlds. The next four outward are Jupiter-like **Jovian planets (p. 326)** that are large and low density.
- All four of the Jovian worlds have ring systems and large families of moons. The Terrestrial planets have no rings and few moons.
- Most of the **asteroids (p. 324)**, small, irregular rocky bodies, are located between the orbits of Mars and Jupiter.
- **Comets (p. 324)** are icy bodies that pass through the inner solar system along long elliptical orbits. As the ices vaporize and release dust, the comet develops a tail that points approximately away from the sun.
- **Meteoroids (p. 325)** that fall into Earth's atmosphere are vaporized by friction and are visible as streaks of light called **meteors (p. 324).** Larger and stronger meteoroids may survive to reach the ground, where they are called **meteorites (p. 325).**
- The **Kuiper belt (p. 324)** is composed of small, icy bodies that orbit the sun beyond the orbit of Neptune.
- The age of a rocky body can be found by radioactive dating, based on the decay **half-life (p. 325)** of radioactive atoms. The oldest rocks from Earth, the moon, and Mars have ages over 4 billion years. The oldest objects in our solar system are some meteorites that have ages of 4.6 billion years. This is taken to be the age of the solar system.
- **Condensation (p. 331)** in the solar nebula converted some of the gas into solid bits of matter, which **accreted (p. 331)** to form billions of **planetesimals (p. 331).**
- Planets begin growing by accretion of solid material into **protoplanets (p. 331).** Once a protoplanet approaches about 15 Earth masses, it can begin growing by **gravitational collapse (p. 332)** as it pulls in gas from the solar nebula.
- According to the **condensation sequence (p. 331),** the inner part of the solar nebula was so hot that only metals and rocky materials could form solid grains. The dense Terrestrial planets grew from those solid particles and did not include many ices or low-density gases.
- The outer solar nebula, beyond the **ice line (p. 331)** was cold enough for large amounts of ices as well as metals, rocky minerals to form solid particles. The Jovian planets grew rapidly and incorporated large amounts of low-density ices and gases.
- Evidence that the condensation sequence was important in the solar nebula can be found in the densities of the Terrestrial planets compared to the Jovian planets. Comparing the **uncompressed densities (p. 330)** of the Terrestrial planets shows that the innermost Terrestrial planets have the highest densities.
- The Terrestrial planets may have formed slowly from the accretion of planetesimals of similar composition and then **differentiated (p. 332)** later when radioactive decay plus **heat of formation (p. 332)** melted each planet's interior. In that scenario, Earth's early atmosphere was probably supplied by a combination of planetesimal impacts and **outgassing (p. 333)** from Earth's interior.
- Disks of gas and dust around protostars may not last long enough to form Jovian planets by accretion and then by gravitational collapse. Some models suggest the Jovian planets could have formed more rapidly by direct gravitational collapse, skipping the condensation and accretion steps.
- In addition to intense light from hot nearby stars and the gravitational influence of passing stars, the solar nebula was eventually cleared away by **radiation pressure (p. 335),** the solar wind, and the sweeping up or ejection of debris by the planets.

- All of the old surfaces in the solar system were heavily cratered in an early **heavy bombardment (p. 335)** by debris that filled the solar system when it was young.
- Hot disks of gas and dust have been detected in early stages of star formation and are thought to be the kind of disk in which planets could form.
- Cold dust disks, also known as **debris disks (p. 337)**, appear to be produced by dust released by collisions among comets, asteroids, and Kuiper belt objects. Such disks may be signs that planets have already formed.
- Planets orbiting other stars, called **extrasolar planets (p. 338)**, have been detected by the way they tug their stars about, creating small Doppler shifts in the stars' spectra. Planets have also been detected in **transits (p. 339)** as they cross in front of their star and partly block the star's light. A few planets have been detected when they orbited behind their star and their infrared radiation was cut off.
- Nearly all extrasolar planets found so far are massive, Jovian worlds. Lower-mass Terrestrial planets are harder to detect but are probably common.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What produced the helium now present in the sun's atmosphere? In Jupiter's atmosphere? In the sun's core?
2. What produced the iron and heavier elements like gold and silver in Earth's core and crust?
3. What evidence can you cite that disks of gas and dust are common around young stars?
4. According to the solar nebula theory, why is the sun's equator nearly in the plane of Earth's orbit?
5. Why does the solar nebula theory predict that planetary systems are common?
6. Why do astronomers think the solar system formed about 4.6 billion years ago?
7. If you visited another planetary system, would you be surprised to find planets older than Earth? Why or why not?
8. Why is almost every solid surface in our solar system scarred by craters?
9. What is the difference between condensation and accretion?
10. Why don't Terrestrial planets have rings like the Jovian planets?
11. How does the solar nebula theory help you understand the location of asteroids?
12. How does the solar nebula theory explain the dramatic density difference between the Terrestrial and Jovian planets?
13. What does the term *differentiated* mean when applied to a planet? Would you expect to find that planets are usually differentiated? Why?
14. What processes cleared the nebula away and ended planet building?
15. What is the difference between the dense hot disks seen around some stars and the low-density cold disks seen around some other stars?
16. What evidence can you cite that planets orbit other stars?
17. **How Do We Know?** What is the difference between a catastrophic theory and an evolutionary theory?
18. **How Do We Know?** How can scientists know anything about the formation of the solar system when there was nobody there to witness those events?
19. **How Do We Know?** Why don't scientists automatically accept the measurements and hypotheses of other scientists?

Discussion Questions

1. If you visited some other planetary system in the act of building planets, would you expect to see the condensation sequence at work, or was it probably unique to our solar system? How do the properties of the extrasolar planets discovered so far affect your answer?
2. In your opinion, do most planetary systems have asteroid belts? Would all planetary systems show evidence of an age of heavy bombardment?
3. If the solar nebula hypothesis is correct, then there are probably more planets in the universe than stars. Do you agree? Why or why not?
4. The human race has intelligence and consequently has both the ability and the responsibility to wonder about its origins. Do you agree?

Problems

1. If you observed the solar system from the nearest star (distance = 1.3 parsecs), what would the maximum angular separation be between Earth and the sun? (*Hint:* See Reasoning with Numbers 3-1.)
2. The brightest planet in our sky is Venus, which is sometimes as bright as apparent magnitude −4 when it is at a distance of about 1 AU. How many times fainter would it look from a distance of 1 parsec (206,265 AU)? What would its apparent magnitude be? (*Hints:* Remember the inverse square law, from Chapter 8; also see the definition of magnitudes in Chapter 2.)
3. What is the smallest-diameter crater you can identify in the photo of Mercury on page 326? (*Hint:* See Appendix A, "Properties of the Planets," to find the diameter of Mercury in kilometers.)
4. A sample of a meteorite has been analyzed, and the result shows that out of every 1000 nuclei of potassium-40 originally in the meteorite, only 200 have not decayed. How old is the meteorite? (*Hint:* See Figure 16-6.)
5. In Table 16-2, which object's observed density differs least from its uncompressed density? Why?
6. What composition might you expect for a planet that formed in a region of the solar nebula where the temperature was about 100 K?
7. Suppose that Earth grew to its present size in 1 million years through the accretion of particles averaging 100 grams each. On the average, how many particles did Earth capture per second? (*Hint:* See Appendix A to find Earth's mass.)
8. If you stood on Earth during its formation, as described in Problem 7, and watched a region covering 100 m^2, how many impacts would you expect to see in an hour? (*Hints:* Assume that Earth had its present radius. The surface area of a sphere is $4\pi r^2$.)
9. The velocity of the solar wind is roughly 400 km/s. How long does it take to travel from the sun to Earth?

Learning to Look

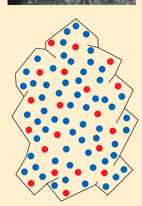
1. What do you see in the image at the right that indicates that the planet formed far from the sun?

NASA

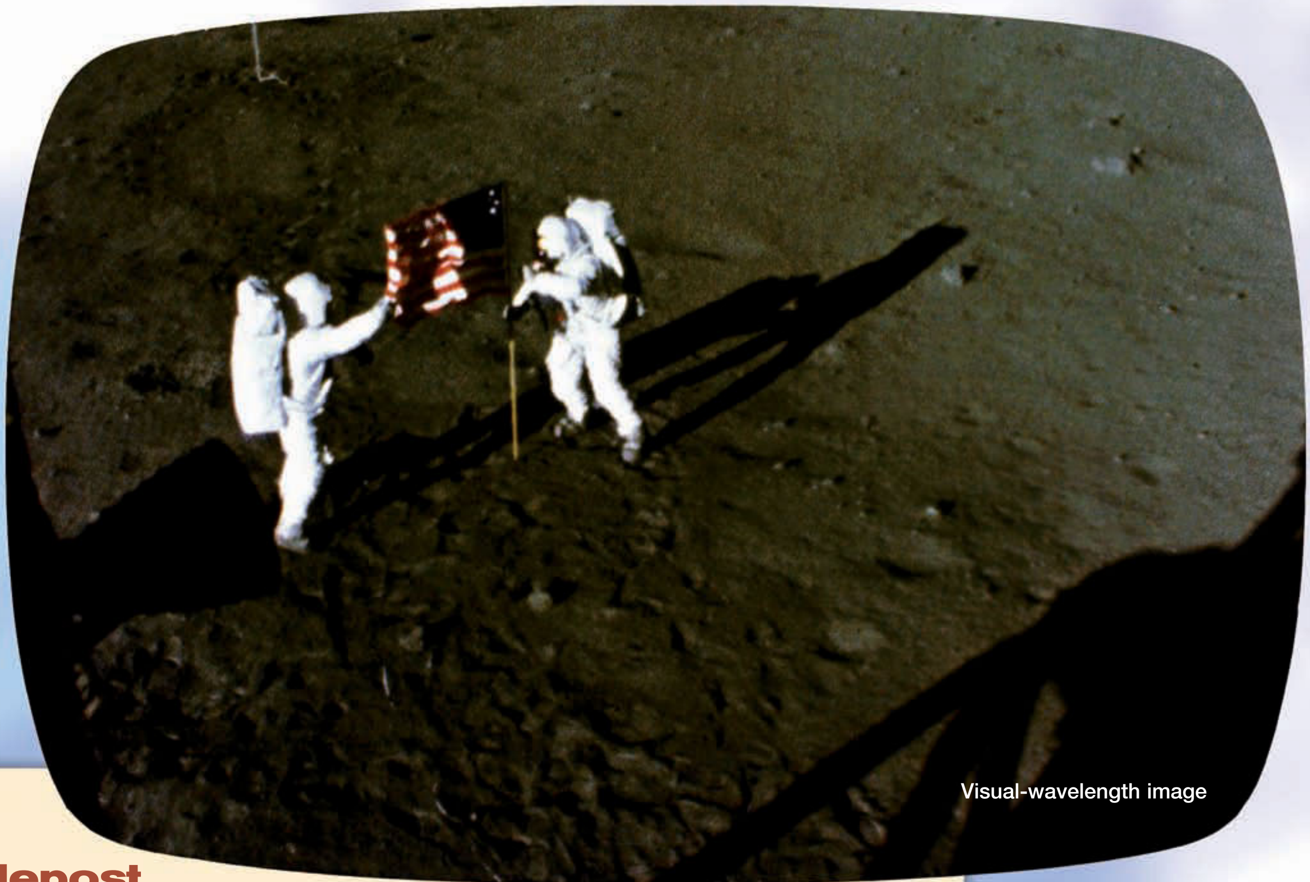
2. Why do astronomers conclude that the surface of Mercury, shown at right, is old? When did the majority of these craters form?

NASA

3. In the mineral specimen represented to the right, radioactive atoms (red) have decayed to form daughter atoms (blue). How old is this specimen in half-lives? (See Figure 16-6.)



17 The Terrestrial Planets



When astronauts stepped onto the surface of the moon, they found an unearthly world with no air, no water, weak gravity, and a dusty, cratered surface. Through comparative planetology, the moon reveals a great deal about our own beautiful Earth. (JSC/NASA)

Guidepost

In the preceding chapter, you learned how our solar system formed as a by-product of the formation of the sun. You also saw how distance from the sun determined the general character of each planet. In this chapter, as you begin to study the individual planets, you can continue comparing the planets with each other, searching for similarities and contrasts. Like people, the Terrestrial planets are more alike than they are different, but it is the differences that are most memorable.

As you explore, you will be searching for answers to four essential questions:

What are the main features of Earth when you view it as a planet?

How does distance from the sun affect the characteristics of a planet and its atmosphere?

How does size determine the geologic activity and evolution of a planet?

What is the evidence that surface conditions on Venus and Mars were originally more Earth-like than at present?

Once you are familiar with the family of the Terrestrial planets, you will be ready to meet a stranger group of characters in the next chapter, the worlds of the outer solar system.

Animated! This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.

That's one small step for [a] man . . .
one giant leap for mankind.

NEIL ARMSTRONG,
ON THE MOON

Beautiful, beautiful.
Magnificent desolation . . .

EDWIN ALDRIN,
ON THE MOON

If you had been the first person to step onto the surface of the moon, what would you have said? Neil Armstrong responded to the historic significance of the first human step on the surface of another world. Buzz Aldrin was second, and he responded to the moon itself. It *is* desolate, and it *is* magnificent. But it is not unusual. Many planets in the universe probably look like Earth's moon, and astronauts may someday walk on such worlds and compare them with Earth's moon.

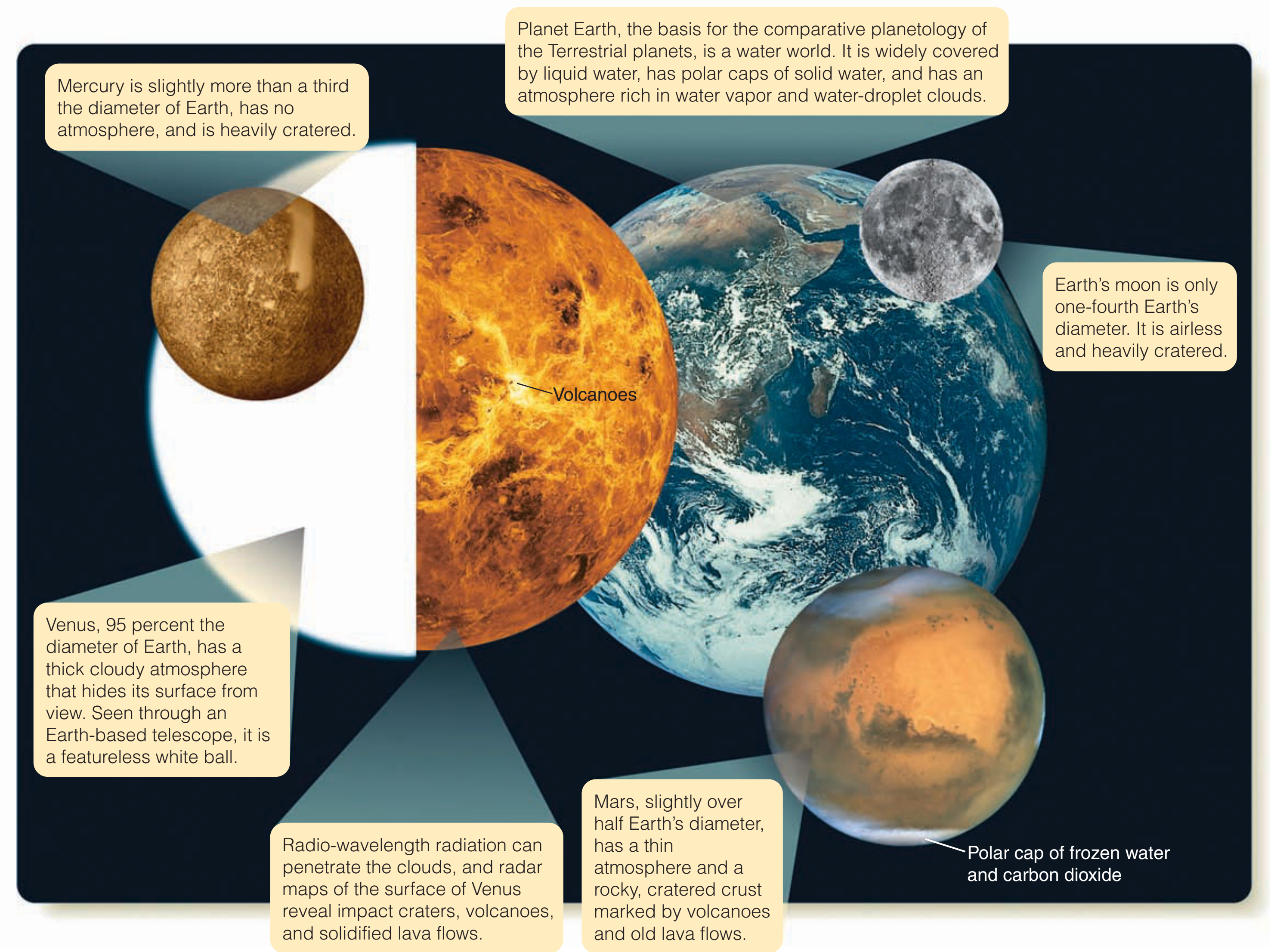
The comparison of one planet with another is called **comparative planetology,** and it is the best way to analyze the worlds in our solar system. You can learn much more by comparing planets than you could by studying them separately.

17-1 A Travel Guide to the Terrestrial Planets

If you visit the city of Granada in Spain, you will probably consult a travel guide, and if it is a good guide, it will do more than tell you where to find museums and restrooms. It will tell

Figure 17-1

Planets in comparison. Earth and Venus are similar in size, but their atmospheres and surfaces are very different. The moon and Mercury are much smaller, and Mars is intermediate in size. (Moon: © UC Regents/Lick Observatory; All planets: NASA)



you to look at the palace called the Alhambra and compare it with buildings in Morocco. Their similarities reveal the Moorish influence on Spain. In this chapter you are going to visit five Earth-like worlds, and this preliminary section will be your guide to important features and comparisons.

Five Worlds

You are about to visit Mercury, Venus, Earth, Earth's moon, and Mars. It may surprise you that the moon is on your itinerary. It is, after all, just a natural satellite orbiting Earth and isn't one of the planets. But the moon is a fascinating world of its own, it makes a striking comparison with the other worlds on your list, and its history gives you important information about the history of Earth and the other planets.

■ Figure 17-1 compares the five worlds you are about to study. The first feature you might notice is diameter. The moon is small, and Mercury is not much bigger. Earth and Venus are large and quite similar in size, but Mars is a medium-sized world. You will discover that size is a critical factor in determining a world's personality. Small worlds tend to be geologically inactive, while larger worlds tend to be active.

Follow the Heat

The Terrestrial worlds are made up of rock and metal. They are all differentiated, which means they are each separated into layers of different density, with high density materials on the inside and lower density materials on the outside.

As you learned in the previous chapter, when the planets formed, their surfaces were subjected to heavy bombardment by leftover planetesimals and debris in the young solar system. You will see lots of craters on these worlds, especially on Mercury and the moon, many of them dating back to the heavy bombardment era. Notice that cratered surfaces are old. For example, if a lava flow covered up some cratered landscape after the end of the heavy bombardment, few craters could be formed later on that surface because most of the debris in the solar system was gone. When you see a smooth plain on a planet, you can guess that surface is younger than the heavily cratered areas.

Another important way you can study a planet is by following the energy flow. In the preceding chapter you learned that the heat in the interior of a planet may be partly from radioactive decay and partly left over from the planet's formation, but in any case it must flow outward toward the cooler surface where it is radiated into space. In the process of flowing outward, the heat can cause convection currents, magnetic fields, plate motions, quakes, faults, volcanism, mountain building, and more. Heat flowing outward through the cooler crust makes a large world like Earth geologically active (■ How De We Know? 17-1). In contrast, the moon and Mercury, both small worlds, cooled quickly inside, so they have little heat flowing outward now and are relatively inactive.

Atmospheres

When you look at Mercury and the moon in Figure 17-1, you can see their craters, plains, and mountains clearly; they each have little or no atmosphere to obscure your view. In comparison, the surface of Venus is completely hidden by a cloudy atmosphere even thicker than Earth's. Mars, the medium-sized planet, has a relatively thin atmosphere.

You might ponder two questions. First, why do some worlds have atmospheres while some do not? You will discover that both size and temperature are important. The second question is more complex. Where did those atmospheres come from? To answer that question, you will have to study the geological history of these worlds.

◄ **SCIENTIFIC ARGUMENT** ►

Why do you expect the inner planets to be high-density worlds?

In the previous chapter, you saw how the inner planets formed from hot inner parts of the solar nebula. No ice solidified there, so the inner planets could grow only from particles of rock and metal able to condense from hot gas. So, you expect the inner planets to be made mostly of rock and metal, which are dense materials.

As you visit the Terrestrial planets, you will find craters almost everywhere. **What made all of those craters?**

◄ ►

17-2 Earth: The Active Planet

Earth is the basis for your comparative study of the Terrestrial planets, so you should pretend to visit it as if you didn't live here. It is geologically active, with a molten interior and heat flowing outward that powers volcanism, earthquakes, and moving crustal plates. Almost 75 percent of Earth's surface is covered by liquid water, unlike any other planet in our solar system, and the atmosphere contains a significant amount of oxygen, also unlike any other planet.

Four Stages of Planetary Development

There is evidence that Earth and the other Terrestrial planets, plus Earth's moon, passed through four developmental stages (■ Figure 17-2).

The first stage of planetary evolution is *differentiation,* the separation of material according to density. As you have already learned, Earth is differentiated: It has a dense metallic core, a less-dense rocky mantle, and a low-density crust. That differentiation is understood to have occurred due to melting of Earth's interior caused by heat from a combination of radioactive decay plus energy released by in-falling matter during the planet's formation. Once the interior of Earth melted, the densest materials were able to sink to the core.

How Do We Know? 17-1

Understanding Planets: Follow the Energy

What causes change? One of the best ways to think about a scientific problem is to follow the energy. According to the principle of cause and effect, every effect must have a cause, and every cause must involve energy. Energy moves from regions of high concentration to regions of low concentration and, in doing so, produces changes. For example, coal burns to make steam in a power plant, and the steam passes through a turbine and then escapes into the air. In flowing from the burning coal to the atmosphere, the heat spins the turbine and makes electricity.

Scientists commonly use energy as a key to understanding nature. A biologist might ask where certain birds get the energy to fly thousands of miles, and a geologist might ask where the energy comes from to power a volcano. Energy is everywhere, and when it moves, whether it is in birds or molten magma, it causes change. Energy is the "cause" in "cause and effect."

In earlier chapters, the flow of energy from the inside of a star to its surface helped you understand how stars, including the sun, work. You saw that the outward flow of energy supports the star against its own weight, drives convection currents that produce magnetic fields, and causes surface activity such as spots, prominences, and flares. You were able to understand stars because you could follow the flow of energy outward from their interiors.

You can also think of a planet by following the energy. The heat in the interior of a planet may be left over from the formation of the planet, or it may be heat generated by radioactive decay, but it must flow outward toward the cooler surface, where it is radiated into space. In flowing outward, the heat can cause convection currents in the mantle, magnetic fields, plate motions, quakes, faults, volcanism, mountain building, and more.

When you think about any world, be it a small asteroid or a giant planet, think of it as a source of heat that flows outward through the planet's surface into space. If you can follow that energy flow, you can understand a great deal about the world. A planetary astronomer once said, "The most interesting thing about any planet is how its heat gets out."

Heat flows out of Earth's interior and generates geological activity such as that at Yellowstone National Park. (M. Seeds)

The second stage, *cratering,* could not begin until a solid surface formed. The heavy bombardment of the early solar system made craters on Earth just as it did on the moon and other planets. As the debris in the young solar system cleared away, the rate of cratering impacts fell rapidly to its present low rate.

The third stage, *flooding,* began as radioactive decay continued to heat Earth's interior and caused rock to melt in the upper mantle, where the pressure was lower than in the deep interior. Some of that molten rock welled up through cracks in the crust and flooded the deeper impact basins. Later, as the environment cooled, water fell as rain and flooded the basins to form the first oceans. Note that on Earth, basin flooding was first by lava and later by water.

The fourth stage, *slow surface evolution,* has continued for at least the past 3.5 billion years. Earth's surface is constantly changing as sections of crust slide over and against each other, push up mountains, and shift continents. In addition, moving air and water erode the surface and wear away geological features. Almost all traces of the first billion years of Earth's history have been destroyed by the active crust and erosion.

Terrestrial planets pass through these four stages, but differences in mass, temperature, and composition between the planets can emphasize some of those stages over others and produce surprisingly different worlds.

Earth's Interior

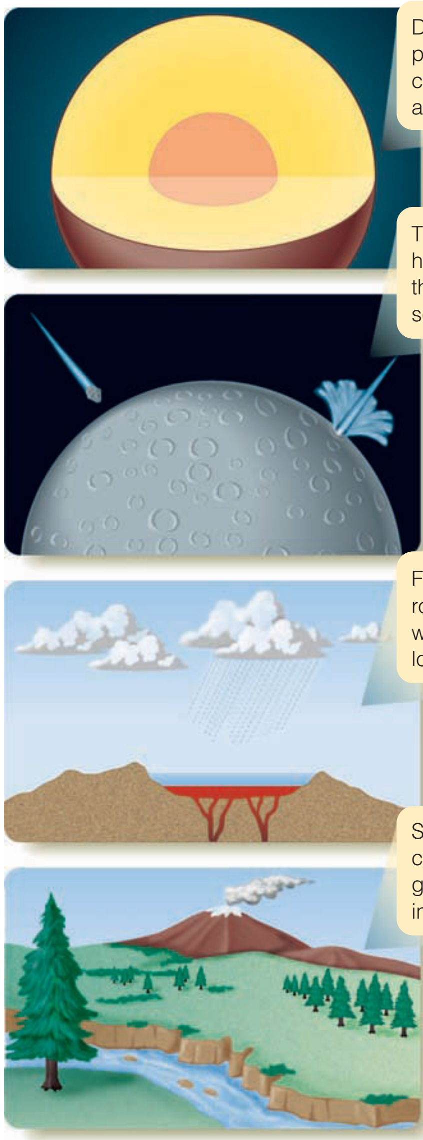
From what you know of the formation of Earth, you would expect it to have differentiated; but in science, evidence rules. What does the evidence reveal about Earth's interior?

Earth's mass divided by its volume tells you its average density, about 5.5 g/cm^3 (■ Celestial Profile 2). But the density of Earth's rocky crust is only about half that much, so a large part of Earth's interior must be made of material denser than crust rock.

Earth scientists have clear proof that Earth did differentiate, even though the deepest wells drilled do not even reach to the bottom of the crust. Each time an earthquake occurs, seismic waves travel through the interior and register on seismographs all over the world. Analysis of those waves shows that Earth's interior is divided into a metallic core, a dense rocky mantle, and a thin, low-density crust.

The core has a density of 14 g/cm^3, denser than lead, and is evidently composed of iron and nickel at a temperature of roughly 6000 K. That means the core of Earth is as hot as the surface of the sun, but the high pressure keeps the metal solid near the center of the core and liquid in its outer part. Two kinds of seismic waves show that the outer core is liquid. The ***P* waves** travel like sound waves and can penetrate a liquid, but ***S* waves** travel as a type of side-to-side vibration that can't pass through a

Four Stages of Planetary Development

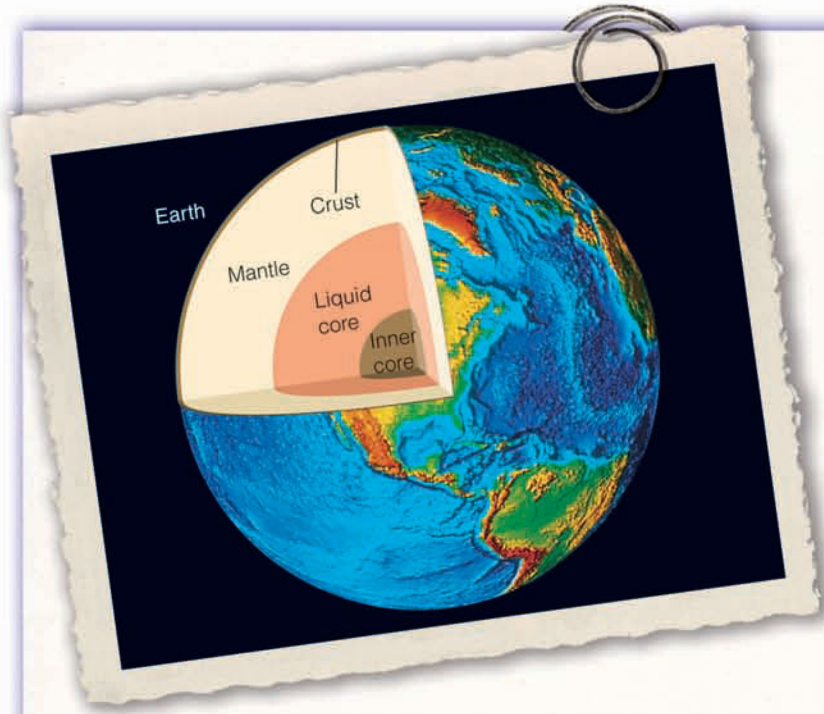


■ **Figure 17-2**

The four stages of Terrestrial planet development are illustrated for Earth.

liquid. Scientists deduce the size of the liquid core by observing where *S* waves get through Earth's interior and where they don't (■ Figure 17-3). The outer boundary of the core lies slightly more than halfway between Earth's center and surface.

Earth's magnetism gives you further evidence about the core. The presence of a magnetic field is a clue that part of Earth's core must be a liquid metal. Convection currents stir the liquid, and it also rotates as Earth rotates. Because of these motions, and because it is a very good conductor of electricity, the liquid outer core generates a magnetic field through the dynamo effect—a different version of the process that creates the sun's magnetic field (see Chapter 7). From traces of magnetic field retained by rocks that formed long ago, geologists conclude that Earth's magnetic field reverses itself every 700,000 years or so. Periodic rever-



Earth's surface has high continents and low sea floors. The crust is only 10 to 60 km thick. Below the crust is a deep mantle and an iron core. (NGDC)

Celestial Profile 2: Earth

Motion:

Average distance from the sun	1.00 AU (1.50 × 10^8 km)
Eccentricity of orbit	0.017
Inclination of orbit to ecliptic	0° [by definition]
Orbital period	1.000 y (365.25 days)
Period of rotation (with respect to the sun)	24.00 h
Period of rotation (sidereal = with respect to the stars)	23.93 h
Inclination of equator to orbit	23.4°

Characteristics:

Equatorial diameter	1.28 × 10^4 km
Mass	5.97 × 10^{24} kg
Average density	5.50 g/cm^3 (4.1 g/cm^3 uncompressed)
Surface gravity	1.00 Earth gravity (9.81 m/s^2)
Escape velocity	11.2 km/s
Surface temperature	−90° to 60°C (−130° to 140°F)
Average albedo	0.39
Oblateness	0.0034

Personality Point:

Earth comes, through Old English *eorthe* and Greek *Eraze,* from the Hebrew *erez,* which means ground. *Terra* comes from the Roman goddess of fertility and growth, *Terra Mater,* Mother Earth.

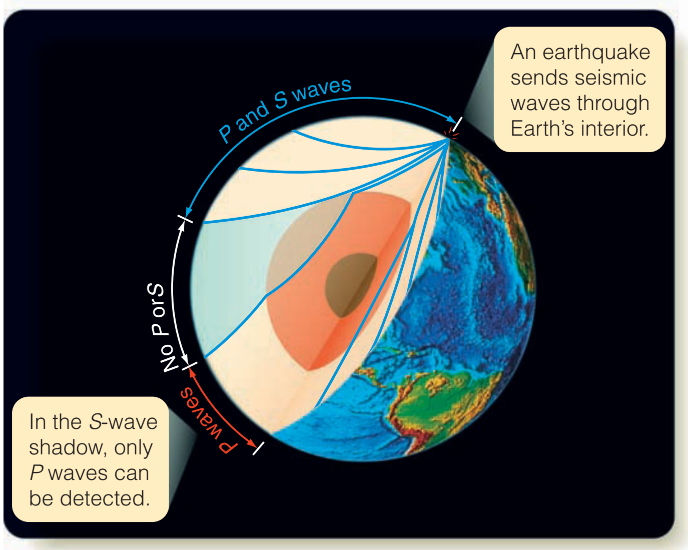


Figure 17-3

P and *S* waves give you clues to the structure of Earth's interior. No direct *S* waves from an earthquake reach the side of Earth opposite their source, showing that Earth's core is liquid. The size of the *S* wave "shadow" tells you the size of the liquid part of the core. **Animated!**

sals, though poorly understood, seem to be a characteristic of the dynamo effect.

As displayed in Celestial Profile 2, Earth's mantle is a thick layer of dense rock that lies between the molten core and the solid crust. Models based on seismic data indicate the mantle material is solid but capable of flowing slowly under pressure, like asphalt used in paving roads, which shatters if struck by a sledgehammer but bends slightly under the weight of a truck. The geologic term for material with those properties is **plastic.** Just below Earth's crust, where the pressure is less than at great depths, the mantle is more plastic, meaning it flows more easily.

Go to **academic.cengage.com astronomy/seeds** to see the Astronomy Exercise "Convection and Magnetic Fields."

Earth's Active Crust

Earth's rocky crust is made up of low-density rocks, and you can think of it as floating on the mantle. The image of a rock floating may seem odd, but recall that the rock underneath the crust, in the mantle, is very dense. Also, just below the crust, the mantle rock tends to be most plastic, so great sections of low-density crust do indeed float on the mantle like huge lily pads floating on a pond. The crust is thickest under the continents, up to 60 km thick, and thinnest under the oceans, where it is only about 10 km thick.

Tectonic motions and the erosive action of water make Earth's crust highly active. Look at **The Active Earth** on pages 350–351 and notice three important points plus six new terms:

1. The motion of crustal plates produces most, but not all, of the geologic activity on Earth. Earthquakes, volcanism, and mountain building are usually linked to motions in the crust and the locations of plate boundaries.

2. The continents on Earth's surface have moved and changed significantly over periods of hundreds of millions of years. A hundred million years is only 0.1 billion years, $^1/_{46}$ of the age of Earth, so sections of Earth's crust are in rapid motion viewed from the perspective of geologic time scales.

3. Most of the geologic features you know—mountain ranges, the Grand Canyon, and even the outline of the continents—are relatively *recent* products of Earth's active surface.

Earth's surface is constantly renewed. The oldest known Earth rocks, small crystals called zircons from western Australia, are 4.3 billion years old. Most of the crust is much younger than that. Most of the mountains and valleys you see around you are probably no more than a few tens of millions of years old. (Recall from the previous chapter that melting a rock resets its radioactive clock, so those ages always refer to the time span since the material was last molten.)

Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercise "Convection and Plate Tectonics."

Earth's Atmosphere

When you think about Earth's atmosphere, you should consider three questions: How did it form? How has it evolved? How are humans changing it? Answering these questions will help you understand other planets as well as our own.

Our planet's first atmosphere, its **primary atmosphere,** was once thought to have contained gases from the solar nebula such as hydrogen and methane. Modern studies, however, indicate that the Terrestrial planets formed hot, so gases such as carbon dioxide, nitrogen, and water vapor could have been outgassed from (cooked out of) the rock and metal as Earth grew. In addition, the final stages of planet building may have seen Earth and the other planets accreting planetesimals rich in volatile materials such as water, ammonia (which contains nitrogen), and carbon dioxide. Thus the primary atmosphere was probably rich in carbon dioxide, nitrogen, and water vapor, from both outgassing and from planetesimal impacts. The atmosphere you breathe today is a **secondary atmosphere** produced later in Earth's history partly by further outgassing and by green plants that produced oxygen.

Soon after Earth formed, it began to cool; once it cooled enough, oceans began to form, and carbon dioxide began to dissolve in the water. Carbon dioxide is highly soluble in water—which explains the easy manufacture of carbonated beverages. As the oceans removed carbon dioxide from the atmosphere, the carbon dioxide reacted with compounds dissolved in the water to form silicon dioxide, limestone, and other mineral sediments. Thus, the oceans transferred the carbon dioxide from the atmosphere to the seafloor and made the air correspondingly richer in other gases left behind, especially nitrogen.

This removal of carbon dioxide is critical to Earth because an atmosphere rich in carbon dioxide can trap heat by a process called the **greenhouse effect.** When visible-wavelength sunlight shines through the glass roof of a greenhouse, it heats the benches and plants inside. The warmed interior radiates heat in the form of infrared radiation, which can't get out through the glass. Heat is trapped in the greenhouse, and the temperature climbs until the glass itself grows warm enough to radiate heat away as fast as sunlight enters (■ Figure 17-4a). (Of course, a greenhouse also retains its heat because the walls prevent the warm air from mixing with the cooler air outside.) This is also called the "parked car effect" for obvious reasons.

Like the glass roof of a greenhouse, a planet's atmosphere can allow sunlight to enter and warm the surface. Carbon dioxide and other greenhouse gases such as water vapor and methane are opaque to infrared radiation, so an atmosphere containing enough greenhouse gases can trap heat and raise the temperature of a planet's surface (Figure 17-4b).

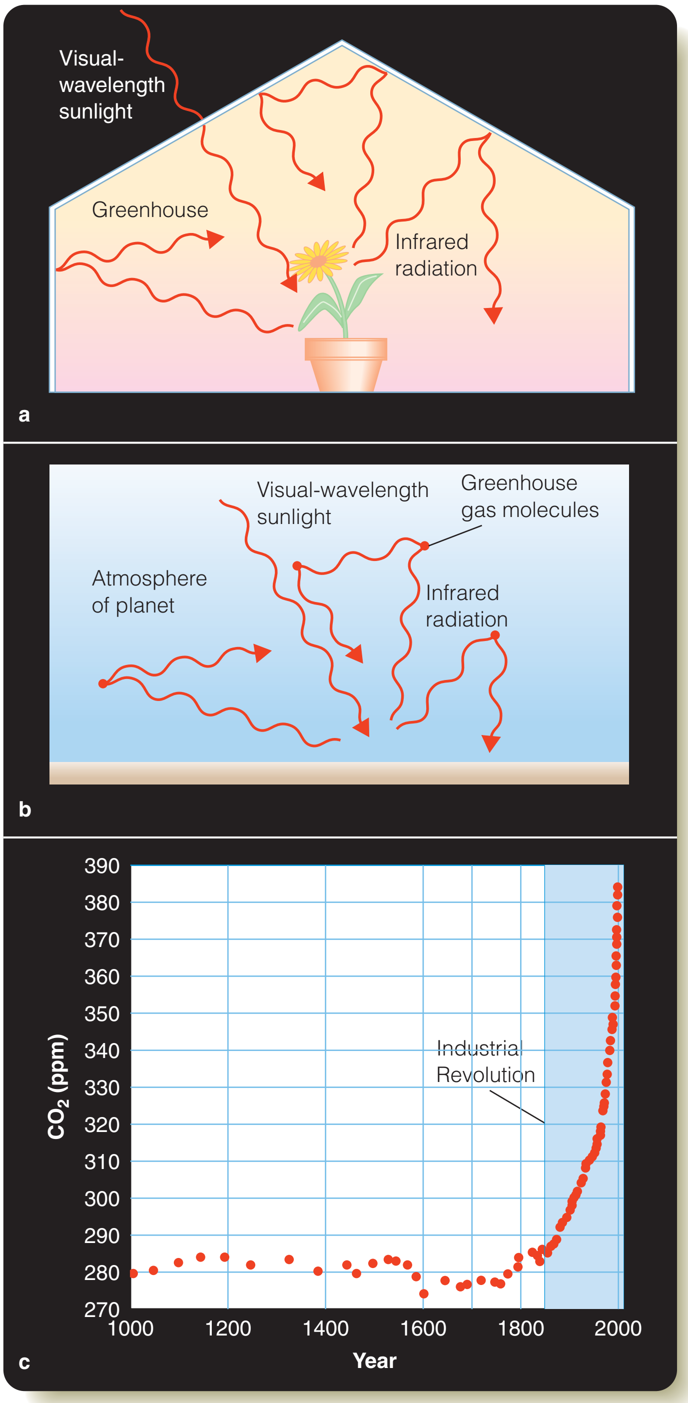
It is a **Common Misconception** that the greenhouse effect is entirely bad. Without the greenhouse effect, Earth would be at least 30 K (54°F) colder, with a planetwide average temperature far below freezing. The problem is that human civilization is rapidly adding greenhouse gases to those that were already in the atmosphere. For 4 billion years, Earth's oceans and plant life have been absorbing carbon dioxide and burying it in the form of carbonates such as limestone and in carbon-rich deposits of coal, oil, and natural gas. In the last century or so, human civilization has begun digging up those fuels, burning them for energy, and releasing the carbon back into the atmosphere as carbon dioxide (Figure 17-4c). This process is steadily increasing the carbon dioxide concentration in our atmosphere and warming Earth's climate in what is called **global warming.**

Global warming is a critical issue because it will change climate patterns, warming some areas and cooling others, directly affecting agriculture. It addition, global warming is melting the polar ice caps, causing sea levels to rise. A rise of just a few feet will flood major land areas where many people live. When you visit Venus, you will see a planet dominated by a runaway greenhouse effect that destroyed the planet's original environment.

Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercises "Primary Atmospheres" and "The Greenhouse Effect."

Oxygen in Earth's Atmosphere

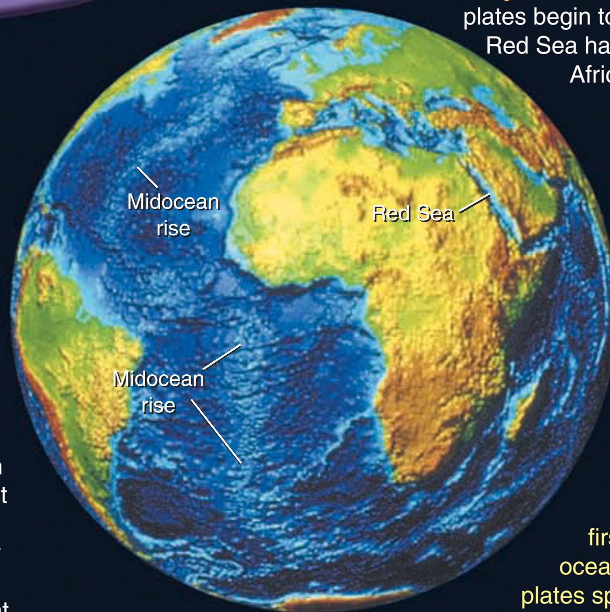
When Earth was young, its atmosphere had no free oxygen, that is, oxygen not combined with other elements. Oxygen is very reactive and quickly forms oxides in the soil. Plant life is what keeps a steady supply of oxygen in Earth's atmosphere: Photosynthesis makes energy for plants by absorbing carbon dioxide and releasing oxygen. Beginning about 2 to 2.5 billion years ago,



■ **Figure 17-4**

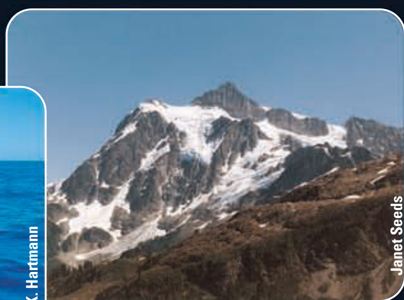
The greenhouse effect: (a) Visual-wavelength sunlight can enter a greenhouse and heat its contents, but the longer-wavelength infrared radiation cannot get out. (b) The same process can heat a planet's surface if its atmosphere contains greenhouse gases such as CO_2. (c) The concentration of CO_2 in Earth's atmosphere as measured in Antarctic ice cores remained roughly constant until the beginning of the Industrial Revolution. Since then it has increased by more than 30 percent. Evidence from proportions of carbon and oxygen isotopes proves that most of the added CO_2 is the result of humans burning fossil fuels. (Graph adapted from a figure by Etheridge, Steele, Langenfelds, Francey, Barnola, and Morgan.)

1 Our world is an astonishingly active planet. Not only is it rich in water and therefore subject to rapid erosion, but its crust is divided into moving sections called plates. Where plates spread apart, lava wells up to form new crust; where plates push against each other, they crumple the crust to form mountains. Where one plate slides over another, you see volcanism. This process is called **plate tectonics,** referring to the Greek word for "builder." (An architect is literally an arch builder.)



A **rift valley** forms where continental plates begin to pull apart. The Red Sea has formed where Africa has begun to pull away from the Arabian peninsula.

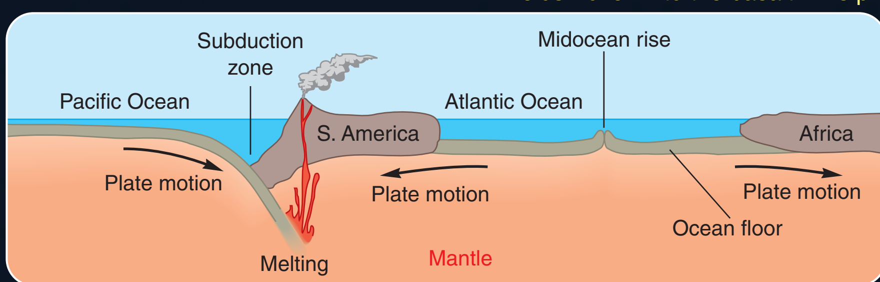
A typical view of planet Earth



Mountains are common on Earth, but they erode away rapidly because of the abundant water.

1a Evidence of plate tectonics was first found in ocean floors, where plates spread apart and magma rises to form **midocean rises** made of rock called **basalt,** a rock typical of solidified lava. Radioactive dating shows that the basalt is younger near the midocean rise. Also, the ocean floor carries less sediment near the midocean rise. As Earth's magnetic field reverses back and forth, it is recorded in the magnetic fields frozen into the basalt. This produces a magnetic pattern in the basalt that shows that the seafloor is spreading away from the midocean rise.

1b A **subduction zone** is a deep trench where one plate slides under another. Melting releases low-density magma that rises to form volcanoes such as those along the northwest coast of North America, including Mt. St. Helens.

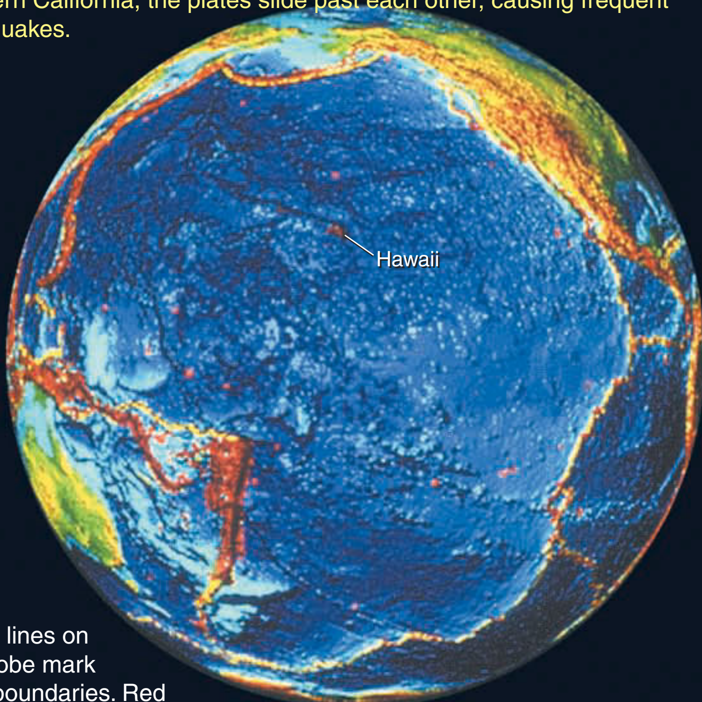


1c Hot spots caused by rising magma in the mantle can poke through a plate and cause volcanism such as that in Hawaii. As the Pacific plate has moved northwestward, the hot spot has punched through to form a chain of volcanic islands, now mostly worn below sea level. **Folded mountain ranges** can form where plates push against each other. For example, the Ural Mountains lie between Europe and Asia, and the Himalaya Mountains are formed by India pushing north into Asia. The Appalachian Mountains are the remains of a mountain range thrust up when North America was pushed against Africa.

CENGAGENOW™

Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Hot Spot Volcanoes." Notice how the moving plate can produce a chain of volcanic peaks, mostly under water in the case of Earth.

1d The floor of the Pacific Ocean is sliding into subduction zones in many places around its perimeter. This pushes up mountains such as the Andes and triggers earthquakes and active volcanism all around the Pacific in what is called the Ring of Fire. In places such as southern California, the plates slide past each other, causing frequent earthquakes.



National Geophysical Data Center

Yellow lines on this globe mark plate boundaries. Red dots mark earthquakes since 1980. Earthquakes within the plate, such as those at Hawaii, are related to volcanism over hot spots in the mantle.

Not long ago, Earth's continents came together to form one continent.

Pangaea broke into a northern and a southern continent.

Notice India moving north toward Asia.

The continents are still drifting on the "plastic" upper mantle.

Continental Drift

2 The floor of the Atlantic Ocean is not being subducted. It is locked to the continents and is pushing North and South America away from Europe and Africa at about 3 cm per year, a motion called *continental drift*. Radio astronomers can measure this motion by timing and comparing radio signals from pulsars using European and American radio telescopes. Roughly 200 million years ago, North and South America were joined to Europe and Africa. Evidence of that lies in similar fossils and similar rocks and minerals found in the matching parts of the continents. Notice how North and South America fit against Europe and Africa like a puzzle.

Mike Seeds

3 Plate tectonics pushes up mountain ranges and causes bulges in the crust, and water erosion wears the rock away. The Colorado River began cutting the Grand Canyon only about 10 million years ago when the Colorado plateau warped upward under the pressure of moving plates. That sounds like a long time ago, but it is only 0.01 billion years. A mile down, at the bottom of the canyon, lie rocks 0.57 billion years old, the roots of an earlier mountain range that stood as high as the Himalayas. It was pushed up, worn away to nothing, and covered with sediment long ago. Many of the geological features we know on Earth have been produced by relatively recent events.

photosynthetic plants in the oceans had multiplied to the point where they made oxygen at a rate faster than chemical reactions could remove it from the atmosphere. After that time, atmospheric oxygen increased rapidly. It is a **Common Misconception** that there is life on Earth because of oxygen. The truth is exactly the opposite: There is oxygen in Earth's atmosphere because of life. Most life forms on Earth do not need oxygen (except the minority of creatures that are animals), and some are even poisoned by it.

Because there is oxygen (O_2) in the atmosphere now, there is also a layer of ozone (O_3) at altitudes of 15 to 30 km. Many people have a **Common Misconception** that ozone is bad because they hear it mentioned as part of pollution. Breathing ozone is in fact bad for you, but you need the ozone layer in the upper atmosphere to protect you from harmful solar UV photons. Certain compounds called chlorofluorocarbons (CFCs), used in refrigeration and some industrial applications, can destroy ozone when they leak into the atmosphere. Since the late 1970s, the ozone concentration in the upper atmosphere has been falling (referred to as an "ozone hole"), and the intensity of harmful UV radiation at Earth's surface, especially at high latitudes, has been increasing.

There is yet another **Common Misconception** that global warming and ozone depletion are two names for the same thing. Take careful note that the ozone hole is a second Earth environmental issue that is basically separate from global warming. While ozone depletion poses an immediate problem for public health on Earth, it is also of interest astronomically. When you visit Mars, you will see the effects of an atmosphere without ozone.

◄ SCIENTIFIC ARGUMENT ►

What evidence indicates that Earth has a liquid metal core?

A good scientific argument focuses on evidence. In this case, the evidence is indirect because you can never visit Earth's core. Seismic waves from distant earthquakes pass through Earth, but a certain kind of wave, the *S* type, does not pass through the core. Because the *S* waves cannot move through a liquid, scientists conclude that Earth's core is partly liquid. Earth's magnetic field gives further evidence of a metallic core. The theory for the generation of magnetic fields, the dynamo effect, requires a moving, conducting liquid (for a planet) or gas (for a star) in the interior. If Earth's core were not partly a liquid metal, it would not be able to generate a magnetic field.

Two different kinds of evidence tell you that our planet has a liquid core. Can you build a new argument? **What evidence can you cite to support the theory of plate tectonics?**

◄ ►

17-3 The Moon

You can't go for a stroll on the moon without a spacesuit. There is no air, and the temperature difference from sunshine to shade is extreme (■ Celestial Profile 3).

Lunar Geology

You could visit two kinds of terrain on the moon. The dark gray areas visible from Earth by naked eye are the smooth lunar lowlands, which, using the Latin word for *seas,* earlier astronomers named **maria** (plural of **mare,** which is pronounced mah-ray). You could also visit the comparatively bright, rugged lunar highlands.

The color of moon rocks is dark gray, but Earth's moon looks quite bright in the night sky. In fact, the average **albedo** of Earth's moon, the fraction of the light that it reflects, is only 0.06. In other words, the moon reflects only 6 percent of the sunlight that hits it. In comparison, Earth, thanks mostly to its bright clouds, has an average albedo of 0.39. The moon looks bright only in contrast to the night sky. It is in reality a dark gray world.

Wherever you went on the moon, you would find craters. These craters look quite dramatic near the **terminator,** the name for the boundary between daylight and darkness on the moon where shadows are long. The highlands are marked heavily by craters, whereas the smooth lowlands contain relatively few craters.

The craters on the moon were formed by the impact of meteorites. Study ■ **Impact Cratering** on pages 354–355 and notice three important points plus four new terms:

1. Impact craters have certain distinguishing characteristics, such as their shape and the way the impacts ejected material across the lunar surface.
2. There is a great range of sizes from giant basins to microscopic pits.
3. Most of the craters on the moon are old; they were formed long ago when the solar system was young.

Twelve Apollo astronauts visited the lunar lowlands and highlands between 1969 and 1972 (■ Figure 17-5). Most of the rocks they found were typical of hardened lava, and some were **vesicular basalt,** which contains holes formed by bubbles in the molten rock (■ Figure 17-6). These bubbles are made when rock flows out onto the surface, and the lower pressure allows gases dissolved in the rock expand to form bubbles. The same thing happens when you open a bottle of carbonated beverage and bubbles form. The presence of vesicular basalts shows that much of the surface of Earth's moon has been covered by successive lava flows, and the dark flat plains of the lunar lowlands, the maria, are actually solidified ancient lava. The highlands, in contrast, are composed of rock containing minerals that have low density and would be among the first to solidify and float to the top of molten rock. For example, the highlands are rich in **anorthosite,** a light-colored and low-density rock that contributes to the highlands' bright contrast with the dark lowlands.

Many of the rocks all over the moon are **breccias,** rocks made up of fragments of broken rock cemented together under pressure (Figure 17-6). The breccias show how extensively the lunar surface has been pounded by meteorites. Nowhere did the astronauts find what could be called bedrock; the entire surface of Earth's moon is fractured by meteorite impacts.

As the astronauts bobbed across the lunar surface under its low gravity, their boots kicked up the powdery dust. This lunar dust is produced by the continuous bombardment of the lunar surface by tiny meteorites that slowly grind exposed rocks into fine gray grit with a consistency like talcum powder.

Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercise "The Moon's Craters."

The Origin of Earth's Moon

Over the last two centuries, astronomers developed three different hypotheses for the origin of Earth's moon. The *fission hypothesis* proposed that the moon broke from a rapidly spinning young Earth. The *condensation hypothesis* suggested that Earth and its moon condensed from the same cloud of matter in the solar nebula. The *capture hypothesis* suggested that the moon formed elsewhere in the solar nebula and was later captured by Earth. Each of these traditional ideas had problems and failed to survive comparison with all the evidence.

In the 1970s, a new hypothesis originated that combined some aspects of the three traditional hypotheses. The **large-impact hypothesis** proposes that the moon formed when a very large planetesimal, estimated to have been at least as massive as Mars, smashed into the proto-Earth. Model calculations indicate that this collision would have ejected a disk of debris into orbit around Earth that would have quickly formed the moon (■ Figure 17-7).

This hypothesis explains several phenomena. If the collision occurred off-center, it would have spun the Earth–moon system rapidly and would thus explain the present high angular momentum. If the proto-Earth and impactor had each already differentiated, the ejected material would have been mostly iron-poor mantle and crust, which would explain the moon's low density and iron-poor composition. Furthermore, the material would have lost its volatile components while it was in space, so the moon also would have formed lacking volatiles. Such an impact would have melted the proto-Earth, and the material falling together to form the moon would also have been heated hot enough to melt. This fits the evidence that the highland anorthosite in the moon's oldest rocks formed by differentiation of large quantities of molten material. The large-impact hypothesis survives comparison with the known evidence and is now considered likely to be correct.

The History of Earth's Moon

The four-stage history of Earth's moon is dominated by a single fact that makes its unfolding noticeably different from Earth's history in the later stages. The moon is small, only one-fourth the diameter of Earth. Its escape velocity is low, so the moon has been unable to hold any atmosphere, cannot have surface water,

Earth's moon has about a quarter the diameter of Earth. Its low density indicates that it does not contain much iron. The size of its core, if any, and the amount of remaining heat are unknown. (NASA)

Celestial Profile 3: The Moon

Motion:

Average distance from Earth	3.84×10^5 km (center to center)
Eccentricity of orbit	0.055
Inclination of orbit to ecliptic	5.1°
Orbital period (sidereal)	27.3 d
Orbital period (synodic = phase cycle)	29.5 d
Inclination of equator to orbit	6.7°

Characteristics:

Equatorial diameter	3.48×10^3 km (0.272 $D_\oplus$)
Mass	7.35×10^{22} kg (0.0123 $M_\oplus$)
Average density	3.36 g/cm^3 (3.3 g/cm^3 uncompressed)
Surface gravity	0.17 Earth gravity
Escape velocity	2.4 km/s (0.21 $V_\oplus$)
Surface temperature	−170° to 130°C (−275° to 265°F)
Average albedo	0.07

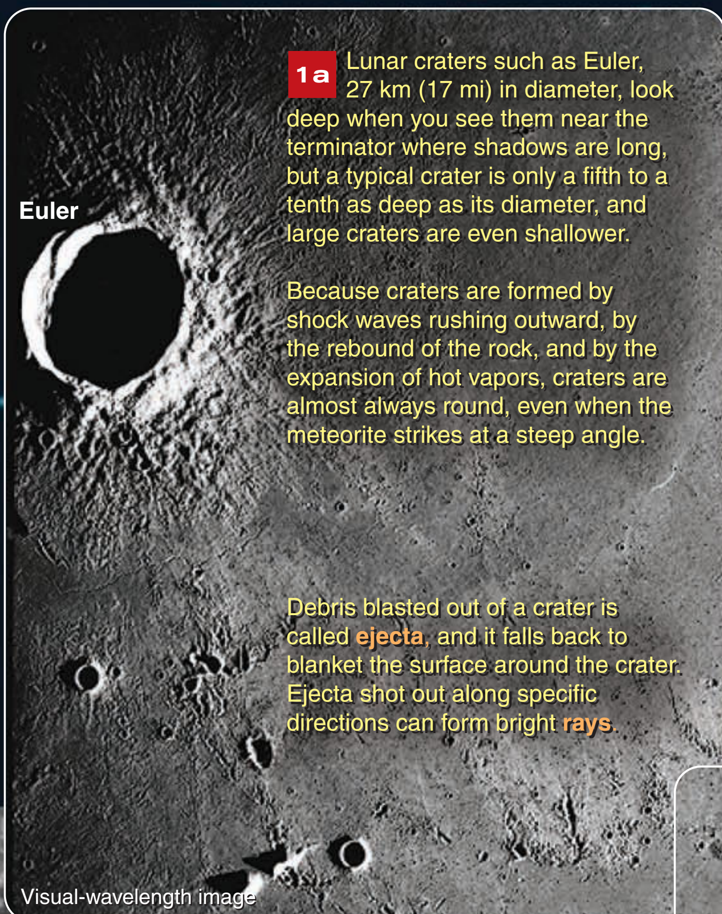
Personality Point:

Lunar superstitions are very common. *Lunatic* and *lunacy* come from *luna,* the moon. Someone who is *moonstruck* is supposed to be at least a bit nutty. Because the moon affects the ocean tides, many superstitions link the moon to water, to weather, and to women's cycle of fertility. Moonlight is supposed to be harmful to unborn children; but, on the plus side, moonlight rituals supposedly can remove warts.

Impact Cratering

1 The craters that cover the moon and many other bodies in the solar system were produced by the high-speed impact of meteorites of all sizes. Meteorites striking the moon travel 10 to 60 km/s and can hit with the energy of many nuclear bombs.

A meteorite striking the moon's surface can deliver tremendous energy and can produce an impact crater 10 or more times larger in diameter than the meteorite. The vertical scale is exaggerated at right for clarity.

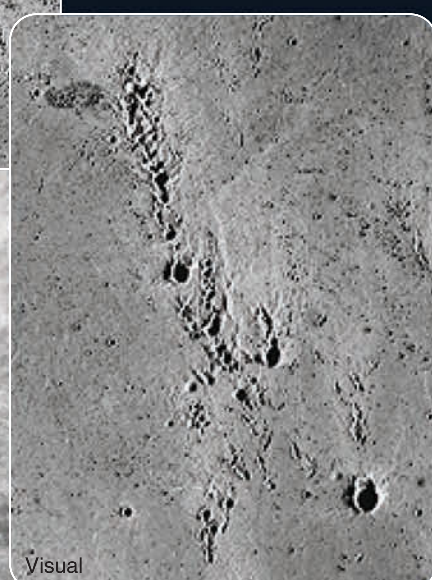


1a Lunar craters such as Euler, 27 km (17 mi) in diameter, look deep when you see them near the terminator where shadows are long, but a typical crater is only a fifth to a tenth as deep as its diameter, and large craters are even shallower.

Because craters are formed by shock waves rushing outward, by the rebound of the rock, and by the expansion of hot vapors, craters are almost always round, even when the meteorite strikes at a steep angle.

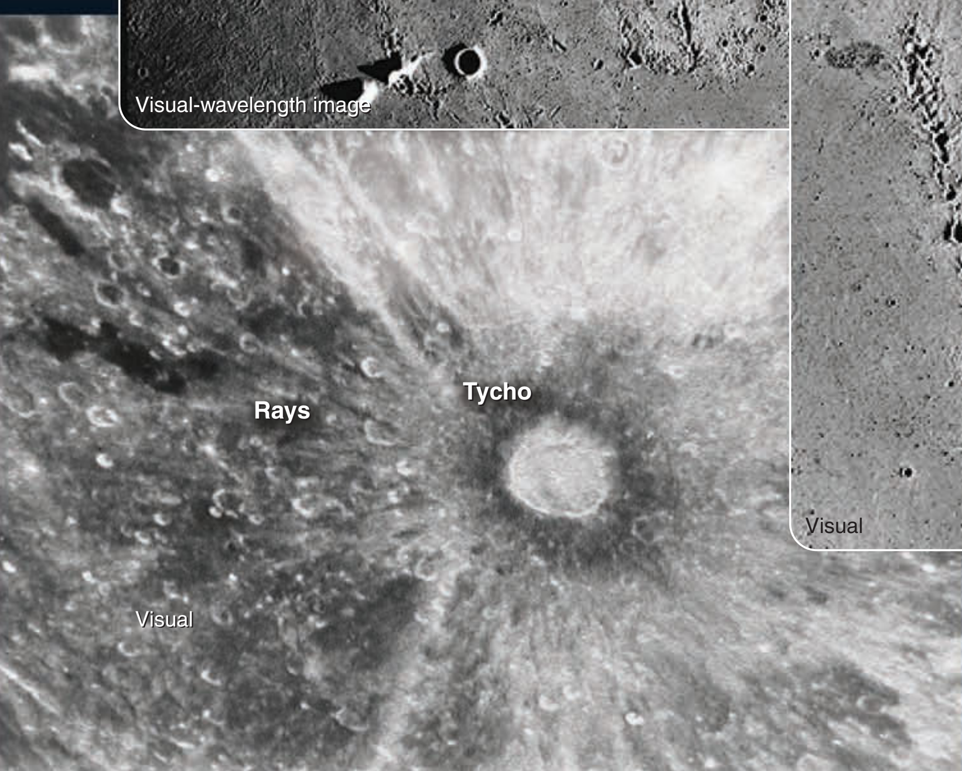
Debris blasted out of a crater is called **ejecta**, and it falls back to blanket the surface around the crater. Ejecta shot out along specific directions can form bright **rays**.

CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "The Moon's Craters." Notice that the structure of the craters depends on their size.

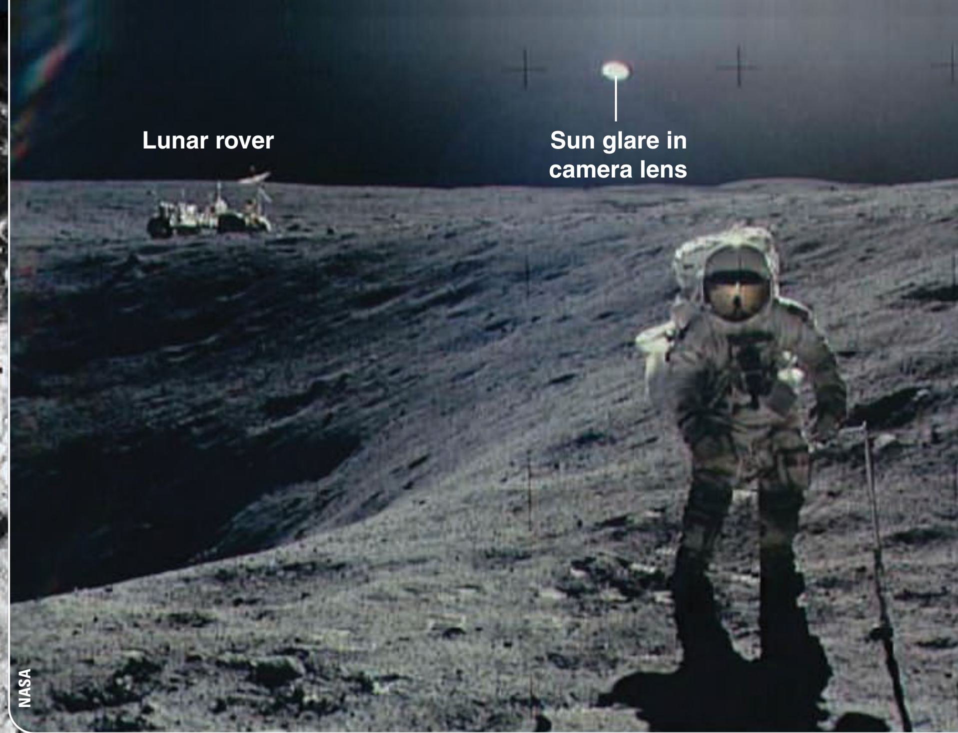


1b Rock ejected from distant impacts can fall back to the surface and form smaller craters called **secondary craters.** The chain of craters here is a 45-km-long chain of secondary craters produced by ejecta from the large crater Copernicus 200 km out of the frame to the lower right.

Bright ejecta blankets and rays gradually darken as sunlight darkens minerals and small meteorites stir the dusty surface. Bright rays are signs of youth. Rays from the crater Tycho, perhaps only 100 million years old, extend halfway around the moon.



2 Plum Crater, 40 m (130 ft) in diameter, was visited by Apollo 16 astronauts. Note the many smaller craters visible. Lunar craters range from giant impact basins to tiny pits in rocks struck by micrometeorites, meteorites of microscopic size.



Visual-wavelength images

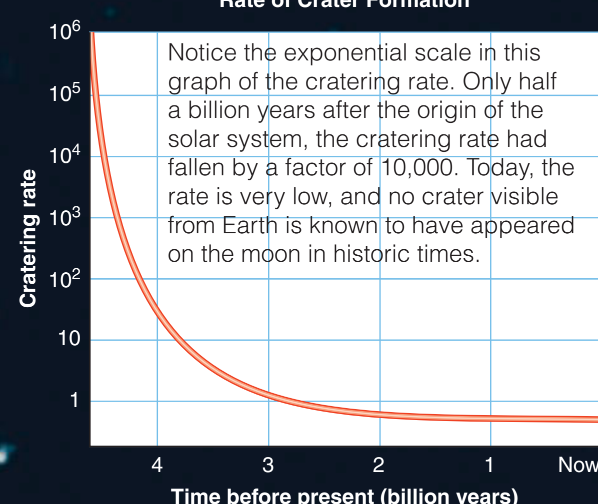
2a In larger craters, the deformation of the rock can form one or more inner rings concentric with the outer rim. The largest of these craters are called **multiringed basins.** In Mare Orientale on the west edge of the visible moon, the outermost ring is almost 900 km (550 mi) in diameter.

2b The energy of an impact can melt rock, some of which falls back into the crater and solidifies. When the moon was young, craters could also be flooded by lava welling up from below the crust.

A few meteorites found on Earth have been identified chemically as fragments of the moon's surface blasted into space by cratering impacts. The fragmented nature of these meteorites indicates that the moon's surface has been battered by impact craters.

3 Most of the craters on the moon were produced long ago when the solar system was filled with debris from planet building. As that debris was swept up, the cratering rate fell rapidly, as shown below.

Rate of Crater Formation



Meteorite from moon

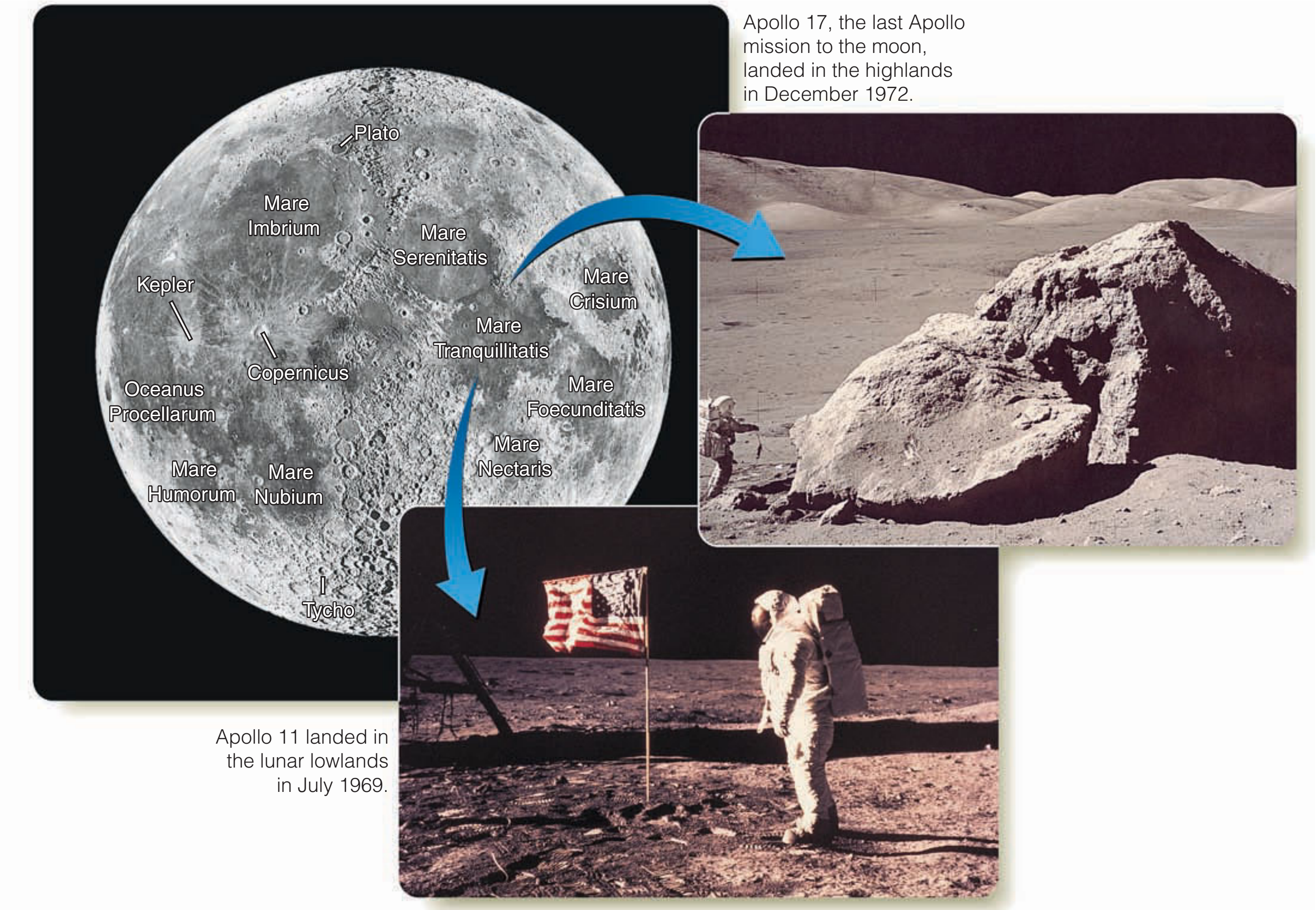


Figure 17-5

The lunar maria are dark smooth lowlands filled with solidified lava. Apollo 11 landed in Mare Tranquillitatis ("Sea of Tranquillity") in 1969, and the horizon was straight and level. The lunar highlands are comparatively bright and heavily cratered. When Apollo 17 landed in 1972 at Taurus-Litrow in the highlands, the astronauts found the horizon mountainous and the terrain rugged. (Moon: © UC Regents/Lick Observatory; Apollo images: NASA)

and its interior cooled rapidly as its internal heat flowed outward into space. Small worlds have less heat and lose it more rapidly, so the moon's small size has been critical in its history.

The Apollo moon rocks, especially anorthosite from the highlands, show that the moon must have formed in a molten state. Planetary geologists now refer to the exterior of the newborn moon as a **magma ocean.** (*Magma* is the term for molten rock in general, whereas *lava* means molten rock flowing on the surface of a world.) Denser materials sank to the bottom of the magma; and, as the magma cooled, low-density minerals floated to the top to form a low-density crust. In this way the moon partly differentiated. The radioactive ages of moon rocks brought back by the Apollo astronauts show that the surface solidified about 4.4 billion years ago.

The second stage, cratering, began as soon as the crust solidified, and the older highlands show that cratering was intense during the first 0.5 billion years—during the heavy bombardment at the end of planet building. The moon's crust was shattered, and the largest impacts formed giant multiringed crater basins hundreds of kilometers in diameter (■ Figure 17-8). The basin that became Mare Imbrium ("Sea of Rains"), for instance, was blasted out by the impact of an object about the size of Rhode Island. This Imbrium event occurred about 4 billion years ago and blanketed 16 percent of the moon with ejecta. Between 4.1 and 3.9 billion years ago, the cratering rate fell rapidly to almost the current rate.

Astronomers can calculate that the tremendous impacts that formed the lunar basins would have cracked the crust to depths of 10 kilometers or more and led to the third stage—flooding. Though Earth's moon cooled rapidly after its formation, radioactive decay continued to heat lunar subsurface material, as it did inside Earth. Parts of the lunar mantle and lower crust remelted, producing lava that followed the cracks up into the giant basins (■ Figure 17-9). The basins were flooded by successive lava flows of dark basalts from 3.8 to 3.2 billion years ago, thus forming the maria.

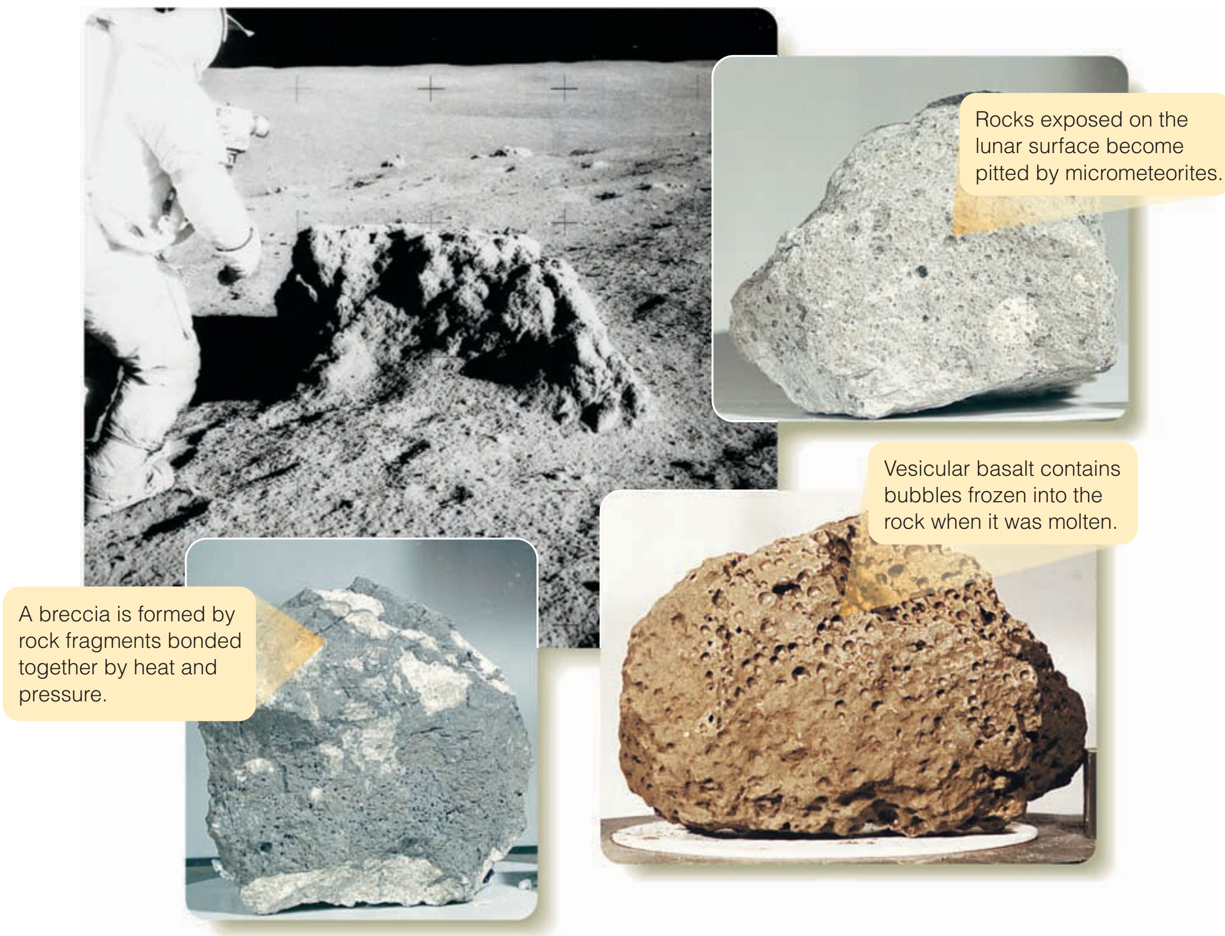


Figure 17-6

Rocks returned from the moon show that the moon once had a deep magma ocean of molten rock, that after the surface solidified it was heavily fractured by cratering, and that it is now affected mainly by micrometeorites grinding away at surface rock. (NASA)

Studies of the moon show that its crust is thinner on the side toward Earth, perhaps due to tidal effects. Consequently, while lava flooded the basins on the Earthward side, it was unable to rise through the thicker crust to flood the lowlands on the far side. The largest known impact basin in the solar system is the moon's South Pole–Aitken Basin (Figure 17-9b). It is about 2500 kilometers (1500 miles) in diameter and as deep as 13 kilometers (8 miles) in places, but flooding has never filled it with smooth lava flows to make an obvious mare.

The fourth stage, slow surface evolution, has been more limited on the moon than on Earth because the moon lacks water and has cooled rapidly. Flooding on Earth included water, but the moon has never had an atmosphere and thus has never had liquid surface water. With no air and no water, erosion is limited to the constant bombardment of micrometeorites and rare larger impacts. Indeed, a few meteorites found on Earth have been identified as moon rocks ejected from the moon by impacts within the last few million years. As the moon lost its internal heat, volcanism died down, and the moon became geologically dead. Its crust never divided into moving plates—evident from the fact that there are no folded mountain ranges—and the moon is now a one-plate object, frozen between stages 3 and 4.

◄ SCIENTIFIC ARGUMENT ►

Why are the maria nearly free of craters?

It's been said that timing is everything, and in this case your argument must carefully consider the sequence of events. The evidence from radioactive ages of moon rocks is that the moon's crust formed and was heavily cratered before about 4 billion years ago. The impacts of the heavy bombardment marked the entire lunar surface and created some very large crater basins. Later, after the end of that heavy bombardment, lava welled up and filled the lowlands of the largest crater basins with basalt to form the maria. Craters in the basins were covered over by lava, and there were few impacts afterward to form new craters. Thus, the maria are nearly free of craters, but the ancient highlands remain heavily cratered.

Now build a new argument. **How is timing important in explaining the formation of an iron-poor moon in the large-impact hypothesis?**

◄ ►

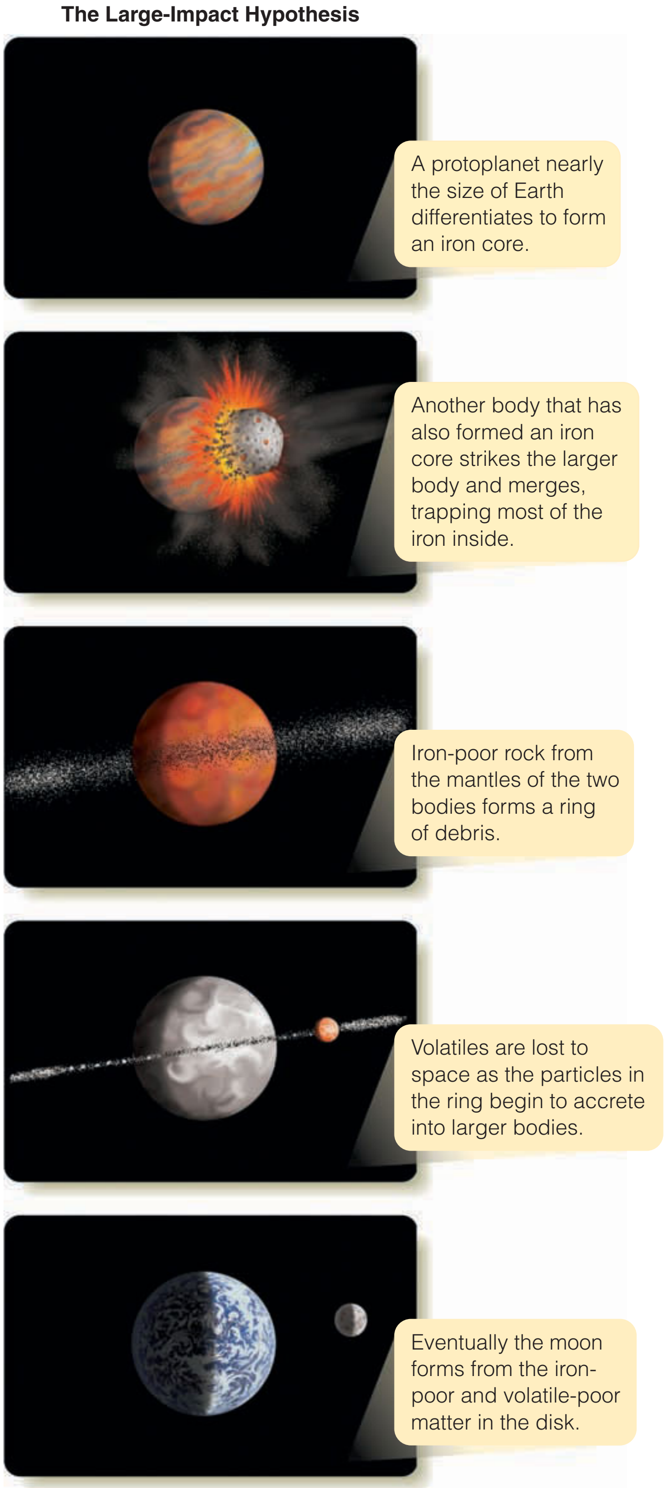


Figure 17-7

When the solar system was about 50 million years old, a massive collision produced the moon in its orbit inclined to Earth's equator.

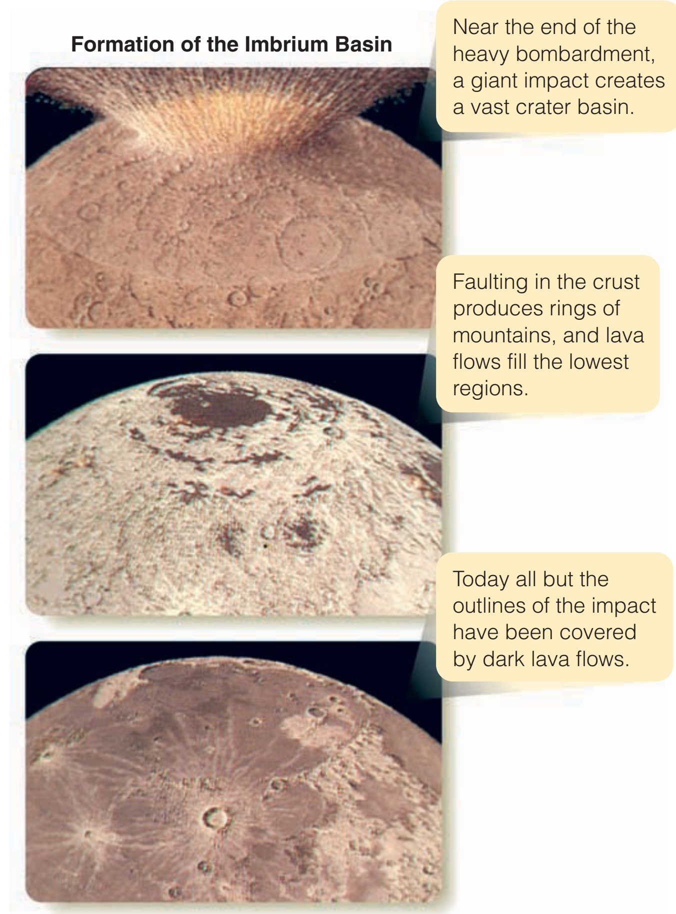


Figure 17-8

Mare Imbrium ("Sea of Rains") on the moon has a generally round outline, the consequence of its formation by a giant impact 4 billion years ago. (Courtesy Don Davis)

17-4 Mercury

YOUR SPACESUIT FOR strolling on Mercury will need to be an even more expensive model than the one you used on Earth's moon. The temperature extremes between sunshine and shade could be deadly (■ Celestial Profile 4). Like Earth's moon, Mercury is small and nearly airless, and it cooled too quickly to develop plate tectonics, so you will find it a cratered, dead world.

Spacecraft Visiting Mercury

Mercury orbits so close to the sun that it is difficult to observe from Earth, and little was known about it until 1974–1975, when the Mariner 10 spacecraft flew past Mercury three times and revealed a planet whose surface is heavily cratered, much like that of Earth's moon. Analysis of the Mariner 10 data showed that large areas have been flooded by lava and then cratered. New information is arriving now from the MESSENGER spacecraft that will fly by Mercury three times during 2008–2010 and then settle into orbit around the planet in 2011.

The largest impact feature on Mercury is the Caloris Basin, a ringed area that MESSENGER photos reveal as 1300 km (800 miles) in diameter (■ Figure 17-10), resembling the large ringed basin Mare Orientale ("Eastern Sea") on Earth's moon. The Caloris basin on Mercury and Mare Orientale on the moon both include concentric rings of cliffs formed by a large impact.

Though Mercury looks moonlike, it does have several features that Earth's moon lacks. Mariner 10 photos revealed long curving ridges called **lobate scarps** up to 3 km (2 mi) high and 500 km (300 mi) long (Figure 17-10). The scarps even cut through craters, indicating that they formed after most of the heavy bombardment. The lobate scarps are the kind of faults that form by compression, but there are no faults on Mercury that could have formed by extension or stretching. This suggests that the entire crust was compressed long ago. MESSENGER photos revealed a "spider" (Figure 17-10) of raised ridges appearing to extend from near a medium-sized crater; geologists are not sure what process could have caused the spider. Spectroscopic observations indicate that Mercury has an extremely thin atmosphere that may be partly outgassed from the crust and partly atoms captured from the solar wind.

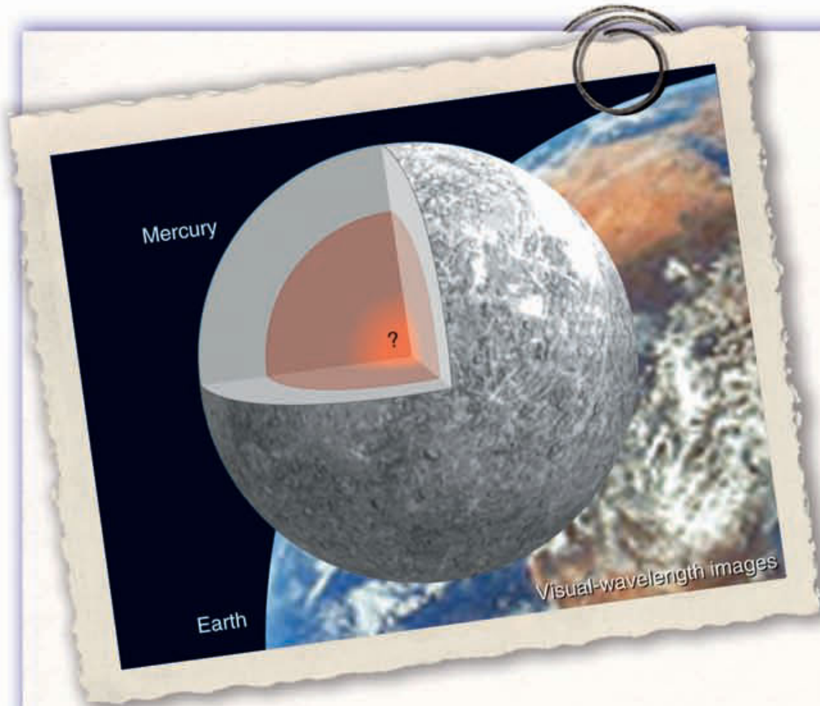
Mercury is quite dense, and models indicate that it must have a large metallic core (Celestial Profile 4). In fact, the metallic core occupies about 70 percent of the radius of the planet. In a sense, Mercury is a metal planet with a thin rock mantle and crust.

The History of Mercury

The accumulated facts about Mercury don't really help you understand the planet until you have a unifying hypothesis. Like a story, it must make sense and bring the known facts together in a logical argument that explains how Mercury got to be the way it is (■ How Do We Know? 17-2).

Mercury is small, and that fact has determined much of its history. Like Earth's moon, Mercury has lost much of its internal heat, and thus is no longer geologically active.

In the first stage of its planetary history, Mercury differentiated to form a metallic core and a rocky mantle. Mariner 10 discovered a magnetic field about 10^{-4} as strong as Earth's—further evidence of a metallic core. In the previous chapter, you saw that the condensation sequence could explain a high abundance of metals in Mercury, but detailed calculations show that Mercury contains even more iron than the condensation sequence would predict. Drawing on the large-impact hypothesis for the origin of Earth's moon, scientists have proposed that Mercury suffered a major impact soon after it differentiated, an impact so large it shattered the rocky mantle and drove much of it away. The remaining iron and rock then re-formed the present Mercury with



Mercury has slightly over a third the diameter of Earth. Its high density means it must have a very large iron core. The amount of heat that Mercury retains is unknown. (NASA)

Celestial Profile 4: Mercury

Motion:

Average distance from the sun	0.387 AU (5.79 × 10^7 km)
Eccentricity of orbit	0.206
Inclination of orbit to ecliptic	7.0°
Orbital period	0.241 y (88.0 d)
Period of rotation (sidereal)	58.6 d
Inclination of equator to orbit	0°

Characteristics:

Equatorial diameter	4.89 × 10^3 km (0.382 $D_{\oplus}$)
Mass	3.31 × 10^{23} kg (0.0558 $M_{\oplus}$)
Average density	5.44 g/cm^3 (5.4 g/cm^3 uncompressed)
Surface gravity	0.38 Earth gravity
Escape velocity	4.3 km/s (0.38 $V_{\oplus}$)
Surface temperature	−170°C to 430°C (−275°F to 800°F)
Average albedo	0.1
Oblateness	0

Personality Point:

Mercury lies very close to the sun and completes an orbit in only 88 Earth days. For this reason, the ancients named the planet after Mercury, the fleet-footed messenger of the gods. The name is also applied to the element mercury, which is known as quicksilver because it is a heavy, quick-flowing silvery liquid at room temperatures.

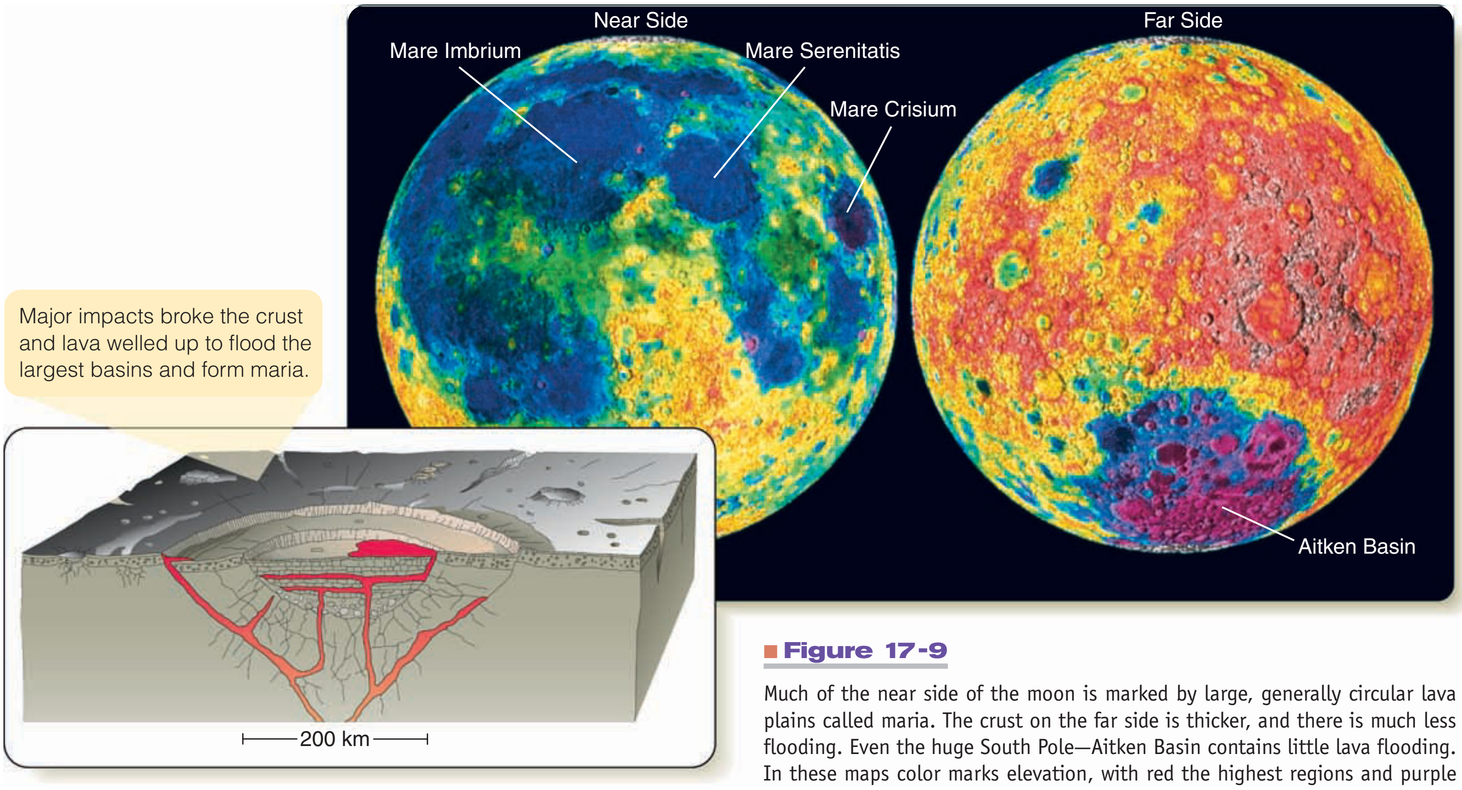


Figure 17-9

Much of the near side of the moon is marked by large, generally circular lava plains called maria. The crust on the far side is thicker, and there is much less flooding. Even the huge South Pole—Aitken Basin contains little lava flooding. In these maps color marks elevation, with red the highest regions and purple the lowest. (Adapted from a diagram by William Hartmann; NASA/Clementine)

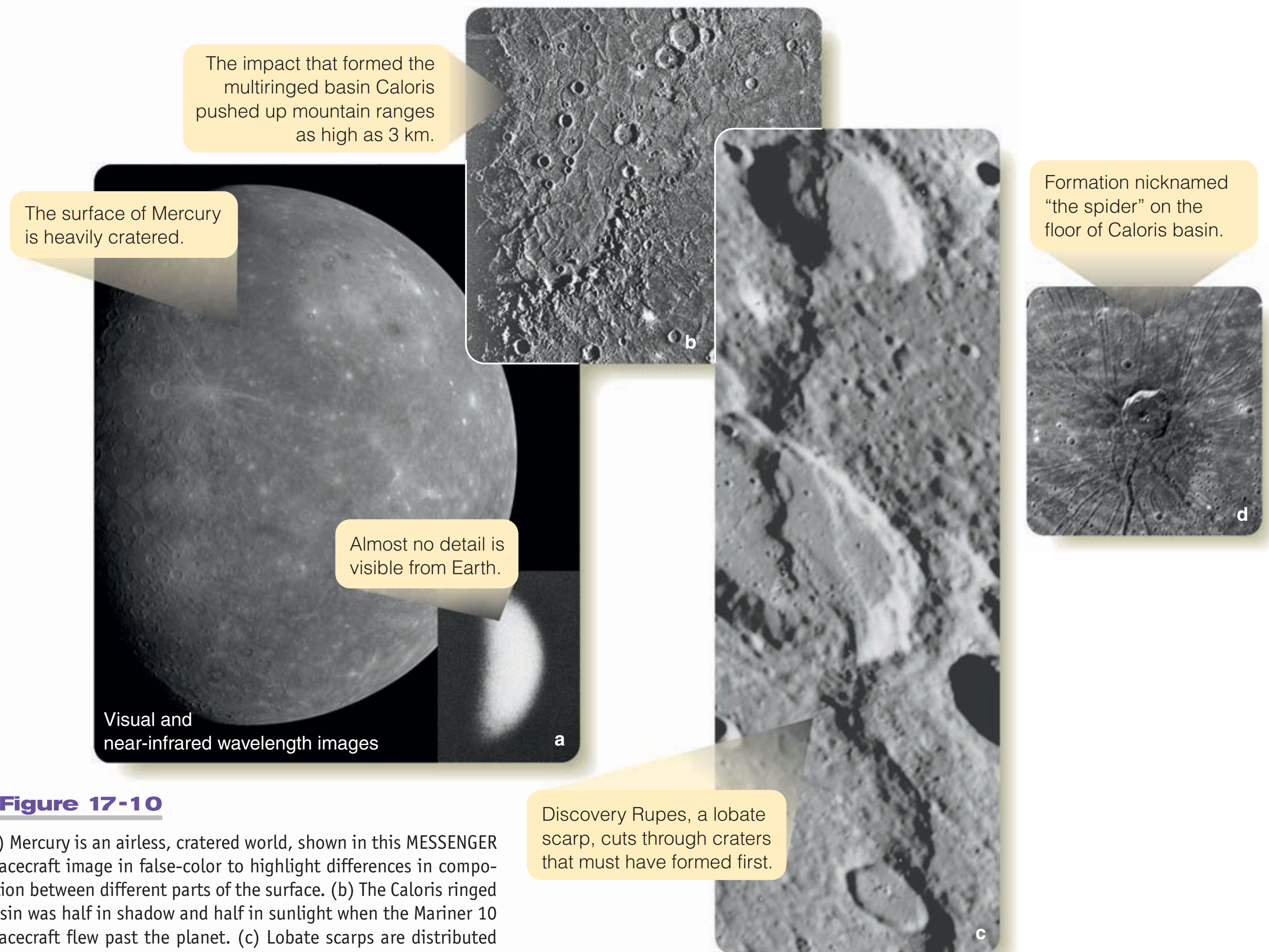


Figure 17-10

(a) Mercury is an airless, cratered world, shown in this MESSENGER spacecraft image in false-color to highlight differences in composition between different parts of the surface. (b) The Caloris ringed basin was half in shadow and half in sunlight when the Mariner 10 spacecraft flew past the planet. (c) Lobate scarps are distributed all around the planet. (d) The origin of the "spider" formation photographed by MESSENGER is a puzzle. (NASAJUH/APL; NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington; Inset: Lowell Observatory)

How Do We Know? 17-2

Hypotheses and Theories Unify the Details

How do scientists make sense out of all the details? Like any technical subject, science includes a mass of details, facts, figures, measurements, and observations. It is easy to be overwhelmed by the flood of details, but one of the most important characteristics of science comes to your rescue. The goal of science is not to discover more details but to explain the details with a unifying hypothesis or theory. A good theory is like a basket that makes it easier for you to carry a large assortment of details.

This is true of all the sciences. When a psychologist begins studying the way the human eye and brain respond to moving points of light, the data are a sea of detailed measurements and observations. Once the psychologist forms a hypothesis about the way the eye and brain interact, the details fall into place as parts of a logical story. If you understand the hypothesis, the details all fit together and make sense, and thus you can remember the details without blindly memorizing tables of facts and figures. The goal of science is understanding, not memorization.

Scientists are in the storytelling business. The stories are often called hypotheses or theories, but they are, in a sense, just stories to explain how nature works. The difference between scientific stories and works of fiction lies in the use of facts. Scientific stories are constructed to fit all known facts and are then tested over and over against new facts obtained by observation and experiment.

When you try to tell the story of each planet in our solar system, you pull together all the hypotheses and theories and try to make them into a logical history of how the planet got to be the way it is. Of course, your stories will be incomplete because scientists don't understand all the factors in planetary evolution. Nevertheless, your story of each planet will draw together the known facts and details and attempt to make them into a logical whole.

Memorizing a list of facts can give you a false feeling of security, just as when you memorize the names of things without understanding them. Rather than memorizing facts, you should search for the unifying hypothesis that pulls the details together into a single story. Your goal in studying science should be to understand nature, not just to remember facts or solve problems.

When scientists create a hypothesis, it draws together a great many observations and measurements. (Phyllis Leber)

its unusually large iron core. Such catastrophic events are rare in nature, but they do occur, so astronomers must be prepared to consider such hypotheses. (See "How Do We Know?" 16-1.)

In the second and third stages of planet formation, cratering battered Mercury's the crust, and lava flows welled up to fill the lowlands, just as they did on the moon. As Mercury lost internal heat, its large metal core contracted and its crust was compressed, breaking to form the lobate scarps much as the peel of a drying apple wrinkles.

Mercury is now a one-plate planet much like Earth's moon and, lacking a significant atmosphere to erode its surface, has changed little since the last lava hardened.

◄ SCIENTIFIC ARGUMENT ►

Why don't Earth and Earth's moon have lobate scarps, but Mercury does? At first glance, you might build an argument to propose that any world with a metallic interior should have lobate scarps, but other factors are also important. Earth has a fairly large metallic core; but, being a large world, it has not cooled very much, so it presumably hasn't shrunk much. Also, the geologic activity on Earth's surface would have erased such scarps if they formed long ago. On the other hand, Earth's moon is not geologically active, but it also does not contain a significant metallic core. Although Earth's moon has lost much of its internal heat, its interior is mostly rock and didn't shrink as much as metal would have.

Now expand your argument. **How do you know the lobate scarps formed after most of the heavy bombardment was over?**

◄ ►

17-5 Venus

YOU MIGHT EXPECT Venus to be much like Earth. Its diameter is 95 percent of Earth's (■ Celestial Profile 5), it has a similar average density and composition, and it is just 30 percent closer to the sun. The surface of Venus is perpetually hidden below thick clouds, and only in the past few decades have planetary scientists discovered that Venus is a deadly hot desert world of volcanoes, lava flows, and impact craters lying at the bottom of a deep ocean of hot gases. No spacesuit will allow you to visit the surface of Venus.

The Atmosphere of Venus

In composition, temperature, and density, the atmosphere of Venus is more Hades than Heaven. The air is unbreathable, very hot, and almost 100 times denser than Earth's air. How do astronomers know this? Because U.S. and Soviet space probes have descended into the atmosphere and, in a few cases, landed and reported back from the surface.

In composition, the atmosphere of Venus is roughly 96 percent carbon dioxide. The rest is mostly nitrogen, with some argon, sulfur dioxide, and small amounts of sulfuric acid, hydrochloric acid, and hydrofluoric acid. There is only a tiny amount of water vapor. On the whole, the composition is deadly unpleasant, and most certainly smells bad too. Spectra show that the impenetrable clouds that hide the surface are made up of droplets of sulfuric acid and microscopic crystals of sulfur (■ Figure 17-11).

This unbreathable atmosphere is 90 times denser than Earth's atmosphere. The air you breathe is 1000 times less dense than water, but on Venus the air is only 10 times less dense than water. If you could survive the unpleasant conditions, you could strap wings on your arms and fly in Venus's atmosphere.

The surface temperature on Venus (Celestial Profile 5) is hot enough to melt lead, and you can understand that because the thick atmosphere creates a severe greenhouse effect. Sunlight filters down through the clouds and warms the surface, but heat cannot escape easily because the atmosphere is opaque to infrared radiation. Traces of sulfur dioxide and water vapor help trap the infrared, but it is the overwhelming abundance of carbon dioxide that makes the greenhouse effect on Venus much more severe than on Earth.

The Surface of Venus

Although the thick clouds on Venus are opaque to visible light, they are transparent to radio waves, so astronomers have been able to map Venus using radar. As early as 1965, Earth-based radio telescopes made low-resolution maps, but later both U.S. and Soviet spacecraft orbited Venus and mapped its surface by radar. Maps made in the early 1990s by the Magellan spacecraft reveal objects as small as 100 meters (300 ft) in diameter.

Radar maps of Venus are reproduced using arbitrary colors. In some maps, scientists have chosen to give Venus an overall orange glow because sunlight filtering down through the clouds bathes the landscape in a perpetual sunset glow. Other radar maps have been colored gray, the natural color of the rocks. In yet other maps, lowlands are colored blue, but there are no oceans on Venus. When you look carefully at colored radar maps of Venus, recall that its surface is a deadly dry desert.

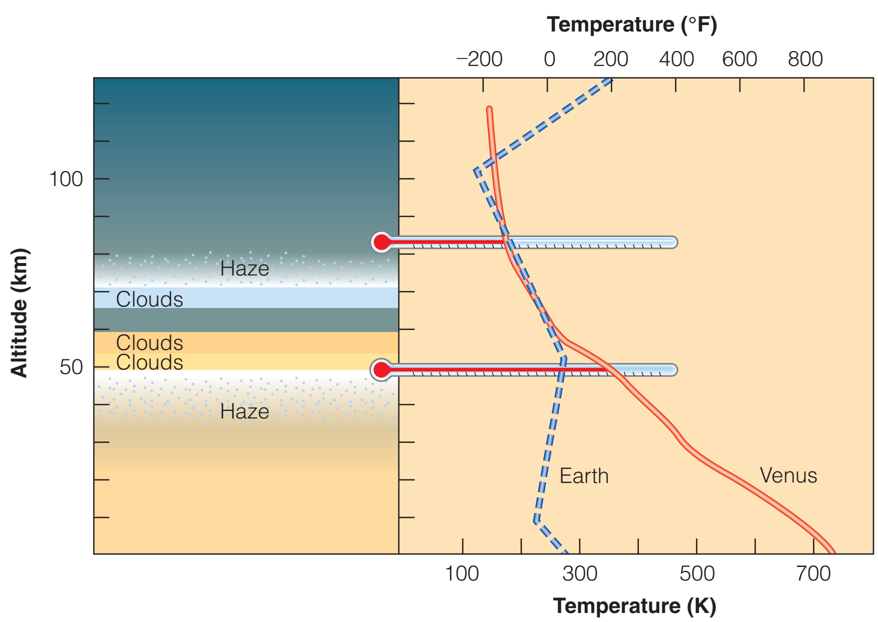
By international agreement, names on Venus are all female, with three exceptions—Maxwell, a high mountain, and Alpha Regio and Beta Regio, two high volcanic peaks—which were all named before the international naming convention for Venus was adopted.

Radar maps show that Venus is similar to Earth in one way but strangely different in other ways. Nearly 75 percent of Earth's surface is covered by low-lying, basaltic seafloors, and 85 percent of Venus's surface is covered by basaltic lowlands. There is no liquid water on Venus, however, so its lowlands are not really seafloors, and the remaining highlands are not like the well-defined continents you see on Earth. Whereas Earth is dominated by plate tectonics, something different is happening on Venus.

The highland area Ishtar Terra, named for the Babylonian goddess of love, is about the size of Australia (■ Figure 17-12). At its eastern edge, the mountain called Maxwell Montes rises to an altitude of 12 km, with the impact crater Cleopatra on its lower slopes (for comparison, Mt. Everest, the tallest mountain on Earth, is 8.8 km high). Bounded by mountain ranges in the north and west, the center of Ishtar Terra is occupied by Lakshmi Planum, a great plateau about 4 km above the surrounding plains. The collapsed calderas Colette and Sacajawea suggest that Lakshmi Planum is a great lava plain. The mountains bounding Ishtar Terra, including Maxwell, resemble folded mountain ranges, which suggests that limited horizontal motion in the crust as well as volcanism may have helped form the highlands.

As usual, you can learn more about other worlds by comparing them with each other and with Earth. Study ■ **Volcanoes** on pages 364–365 and notice three important ideas plus two new terms:

1. There are two main types of volcanoes found on Earth. *Composite volcanoes* are associated mostly with plate boundaries, and *shield volcanoes* are associated with hot spots that are not related to plate boundaries.
2. Volcanoes on Venus and Mars can be recognized by their shapes as being shield volcanoes, the kind produced by hot-spot volcanism and not by plate tectonics.
3. Some volcanoes on Venus and Mars are very large. They have grown to great sizes because of repeated eruptions at the same place in the crust. This is also evidence that neither Venus nor Mars has been dominated by horizontal plate tectonics like Earth's.



■ **Figure 17-11**

On Venus, three main cloud layers composed mostly of sulfuric acid droplets reflect much of the sunlight away. What reaches the surface is deeply reddened, like an intense sunset. If you could insert thermometers into the atmosphere, you would find that the lower atmosphere of Venus is much hotter than that of Earth. **Animated!**

While you are thinking about volcanoes, you can correct a **Common Misconception.** The molten rock that emerges from volcanoes comes from pockets of melted rock in the upper mantle and lower crust and not from a planet's molten core.

Many features on Venus testify to its volcanic history. Long, narrow lava channels meander for thousands of kilometers (■ Figure 17-13a). Radar maps reveal many smaller volcanoes, faults, and sunken regions produced when magma below the surface drained away. Other volcanic features include the **coronae,** circular bulges up to 2100 km (1300 mi) in diameter bordered by fractures, volcanoes, and lava flows (Figure 17-13b). These appear to be produced by rising convection currents of molten magma that push up under the crust. When the magma withdraws, the crust sinks back, but the circular fractures mark the edge of the corona.

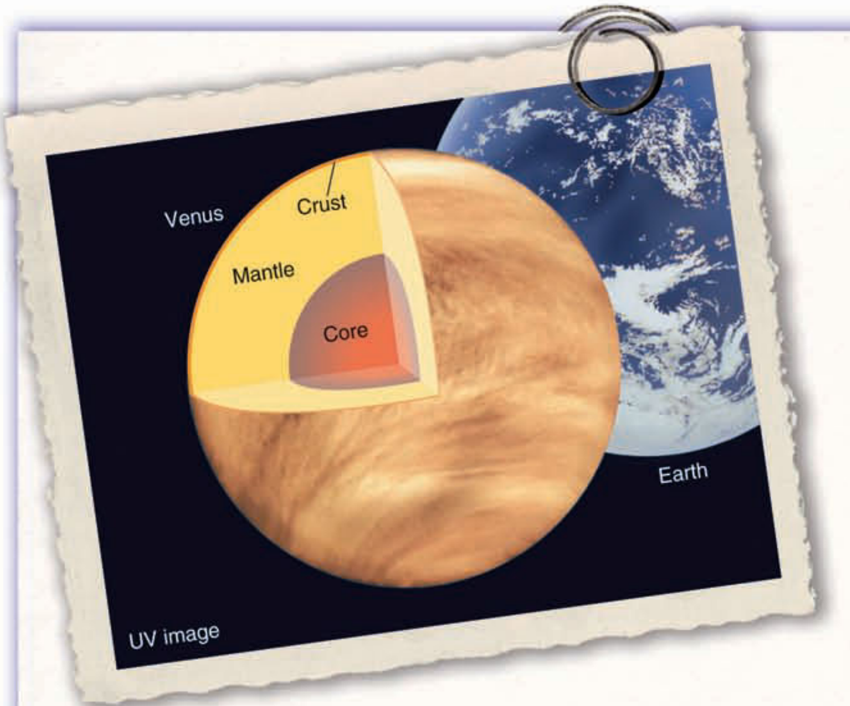
Radar images show that Venus is marked by numerous craters (Figure 17-13c). The atmosphere protects the surface from smaller meteorites that would produce craters smaller than 3 km in diameter. Larger meteorites penetrate the atmosphere and have formed about 10 percent as many craters on Venus as on the maria of Earth's moon. The number of craters shows that the crust is not as ancient as the lunar maria but also not as young as Earth's active surface. The average age of the surface of Venus is estimated from crater counts to be roughly half a billion years. Geologic processes are not renewing the surface of Venus as rapidly as Earth's surface, but no heavily cratered terrain or large impact basins remain on Venus from the heavy bombardment era.

No astronaut has ever stood on Venus, but a few spacecraft landed on the surface and survived the heat and pressure for a few hours. Some of those spacecraft analyzed nearby rocks and snapped a few photographs (■ Figure 17-14). The surface rocks on Venus are dark gray basalts much like those in Earth's ocean floors. This confirms the evidence that volcanism is important on Venus.

The History of Venus

To tell the story of Venus you must draw together all the evidence and find hypotheses to explain two things, the thick carbon dioxide atmosphere and the peculiar geology.

Calculations show that Venus and Earth should have outgassed about the same amount of carbon dioxide, but Earth's oceans have dissolved most of Earth's carbon dioxide and converted it to sediments such as limestone. If all of Earth's carbon were dug up and converted back to carbon dioxide, our atmosphere would be about as dense as the air on Venus and composed mostly of carbon dioxide, like Venus's atmosphere. This suggests that the main difference between Earth and Venus is the lack of water on Venus that would have removed carbon dioxide from the atmosphere. There is evidence that Venus had oceans when it was young; but, being closer to the sun, it was warmer, and the carbon dioxide in the atmosphere created a greenhouse effect that made



Venus is only 5 percent smaller than Earth. Its atmosphere is perpetually cloudy and its surface is hot enough to melt lead. It probably has a liquid metal core about the size of Earth's. (NASA)

Celestial Profile 5: Venus

Motion:

Average distance from the sun	0.723 AU (1.08×10^8 km)
Eccentricity of orbit	0.007
Inclination of orbit to ecliptic	3.4°
Orbital period	0.615 y (224.68 d)
Period of rotation (sidereal)	243.01 d
Inclination of equator to orbit	177° (retrograde rotation)

Characteristics:

Equatorial diameter	1.21×10^4 km (0.949 $D_\oplus$)
Mass	4.87×10^{24} kg (0.815 $M_\oplus$)
Average density	5.24 g/cm^3 (4.2 g/cm^3 uncompressed)
Surface gravity	0.90 Earth gravity
Escape velocity	10.3 km/s (0.92 $V_\oplus$)
Surface temperature	470°C (880°F)
Albedo (cloud tops)	0.76
Oblateness	0

Personality Point:

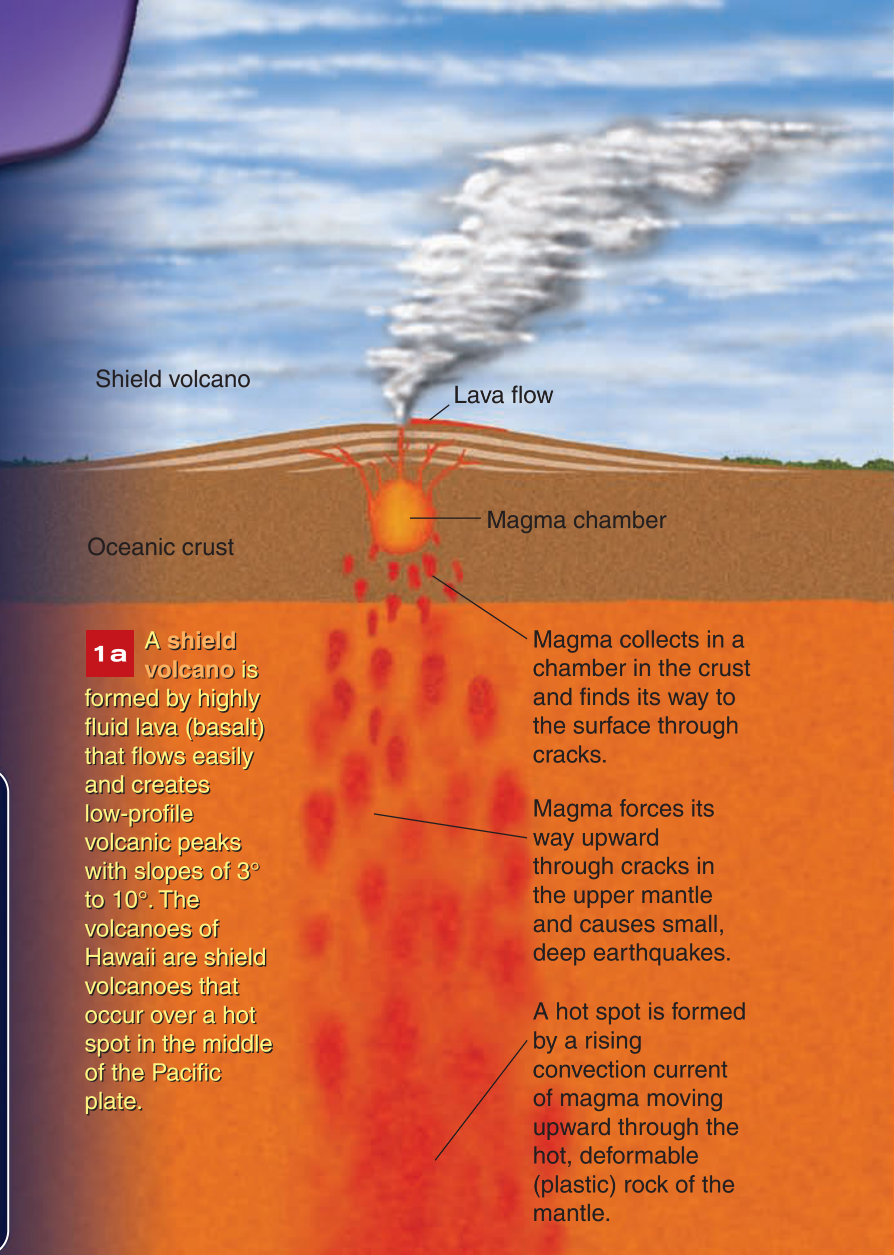
Venus is named for the Roman goddess of love, perhaps because the planet often shines so beautifully in the evening or dawn sky. In contrast, the ancient Maya identified Venus as their war god Kukulkan and sacrificed human victims to the planet when it rose in the dawn sky.

Volcanoes

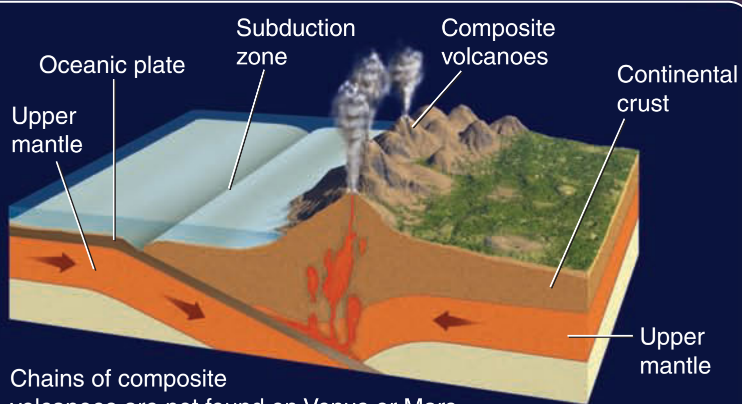
1 Molten rock (magma) is less dense than the surrounding rock and tends to rise. Where it bursts through Earth's crust, you see volcanism. The two main types of volcanoes on Earth provide good examples for comparison with those on Venus and Mars.

On Earth, **composite volcanoes** form above subduction zones where the descending crust melts and the magma rises to the surface. This forms chains of volcanoes along the subduction zone, such as the Andes along the west coast of South America.

Magma rising above subduction zones is not very fluid, and it produces explosive volcanoes with sides as steep as 30°.



1a A shield volcano is formed by highly fluid lava (basalt) that flows easily and creates low-profile volcanic peaks with slopes of 3° to 10°. The volcanoes of Hawaii are shield volcanoes that occur over a hot spot in the middle of the Pacific plate.



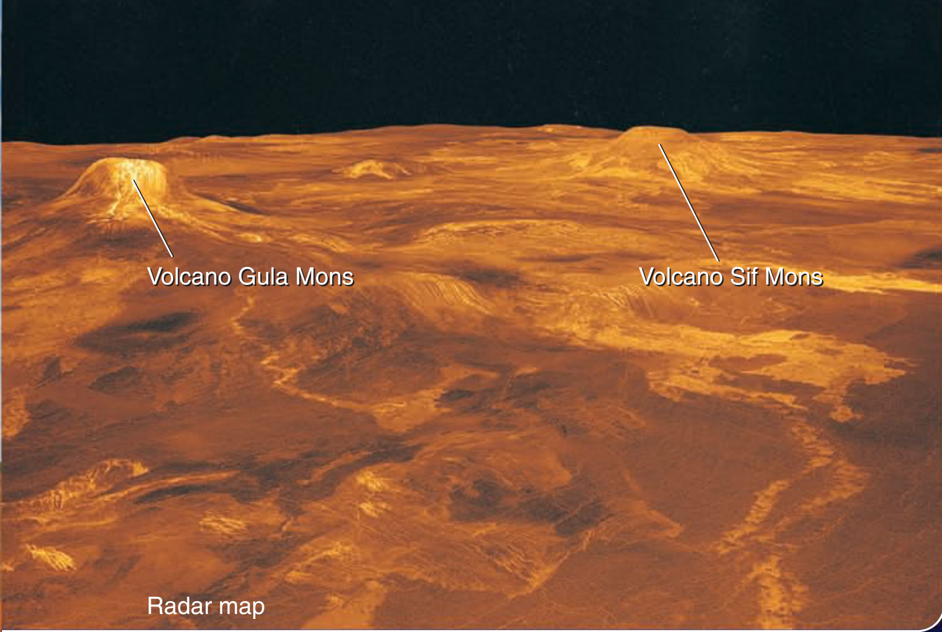
Chains of composite volcanoes are not found on Venus or Mars, which is evidence that subduction and plate motion does not occur on those worlds.

Based on *Physical Geology*, 4th edition, James S. Monroe and Reed Wicander, Wadsworth Publishing Company. Used with permission.

The Cascade Range composite volcanoes are produced by an oceanic plate being subducted below North America and partially melting.

Mount St. Helens exploded northward on May 18, 1980, killing 63 people and destroying 600 km^2 (230 mi^2) of forest with a blast of winds and suspended rock fragments that moved as fast as 480 km/hr (300 mph) and had temperatures as hot as 350°C (660°F). Note the steep slope of this composite volcano.

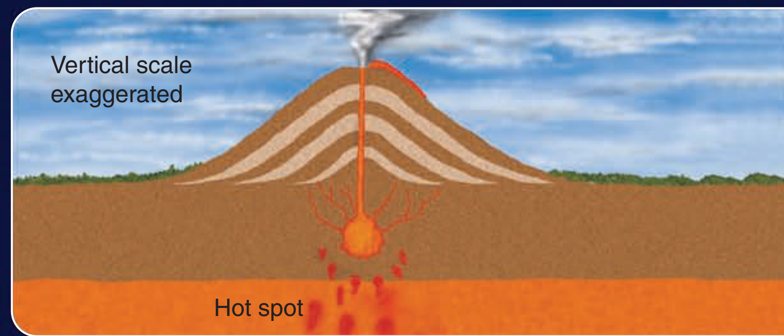
USGS



NASA

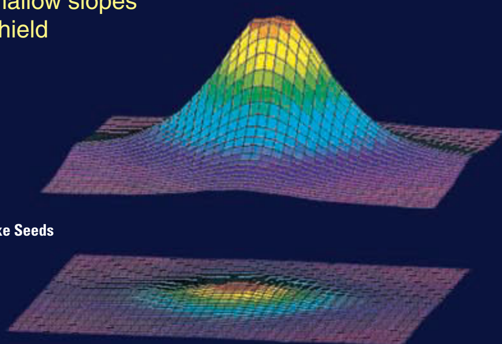
2 Volcanoes on Venus are shield volcanoes. They appear to be steep sided in some images created from Magellan radar maps, but that is because the vertical scale has been exaggerated to enhance detail. The volcanoes of Venus are actually shallow-sloped shield volcanoes.

3 Volcanism over a hot spot results in repeated eruptions that build up a shield volcano of many layers. Such volcanoes can grow very large.

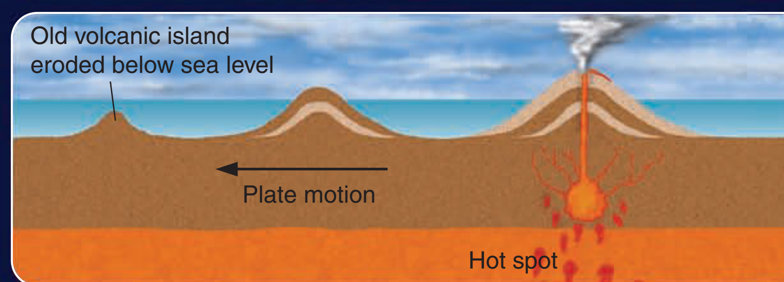


2a This computer model of a mountain with the vertical scale magnified 10 times appears to have steep slopes such as those of a composite volcano.

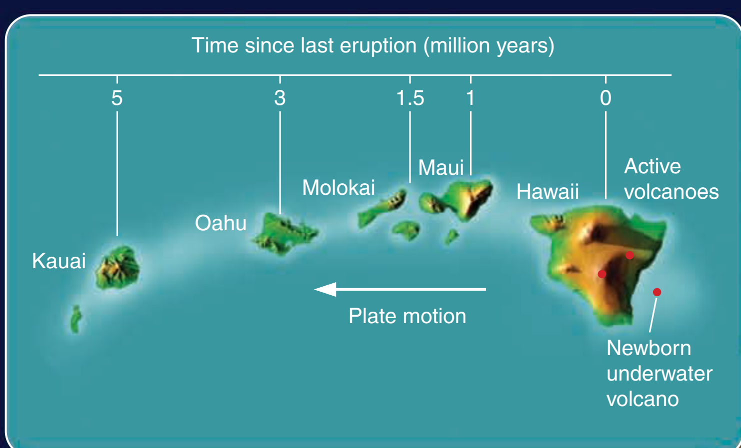
A true profile of the computer model shows the mountain has very shallow slopes typical of shield volcanoes.



Mike Seeds



If the crustal plate is moving, magma generated by the hot spot can repeatedly penetrate the crust to build a chain of volcanoes. Only the volcanoes over the hot spot are active. Older volcanoes slowly erode away. Such volcanoes cannot grow large because the moving plate carries them away from the hot spot.



CENGAGENOW™

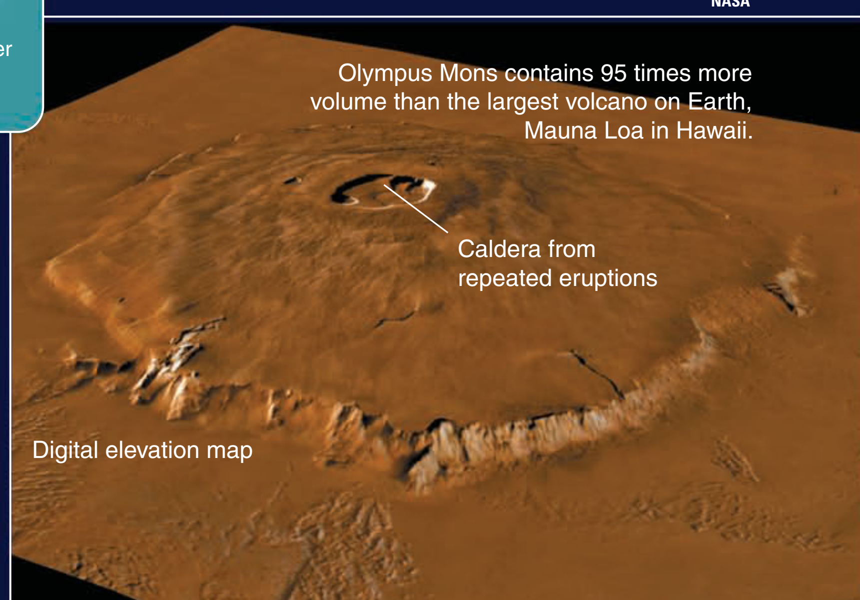
Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see the Active Figure "Hot Spot Volcanoes" and compare volcanism on Earth with that on Venus.

3a The volcanoes that make up the Hawaiian Islands as shown at left have been produced by a hot spot poking upward through the middle of the moving Pacific plate.

NASA

3b The plate moves about 9 cm/yr and carries older volcanic islands northwest, away from the hot spot. The volcanoes cannot grow extremely large because they are carried away from the hot spot. New islands form to the southeast over the hot spot.

Olympus Mons at right is the largest volcano on Mars. It is a shield volcano 25 km (16 mi) high and 700 km (440 mi) in diameter at its base. Its vast size is evidence that the crustal plate must have remained stationary over the hot spot. This is evidence that Mars has not had plate tectonics.



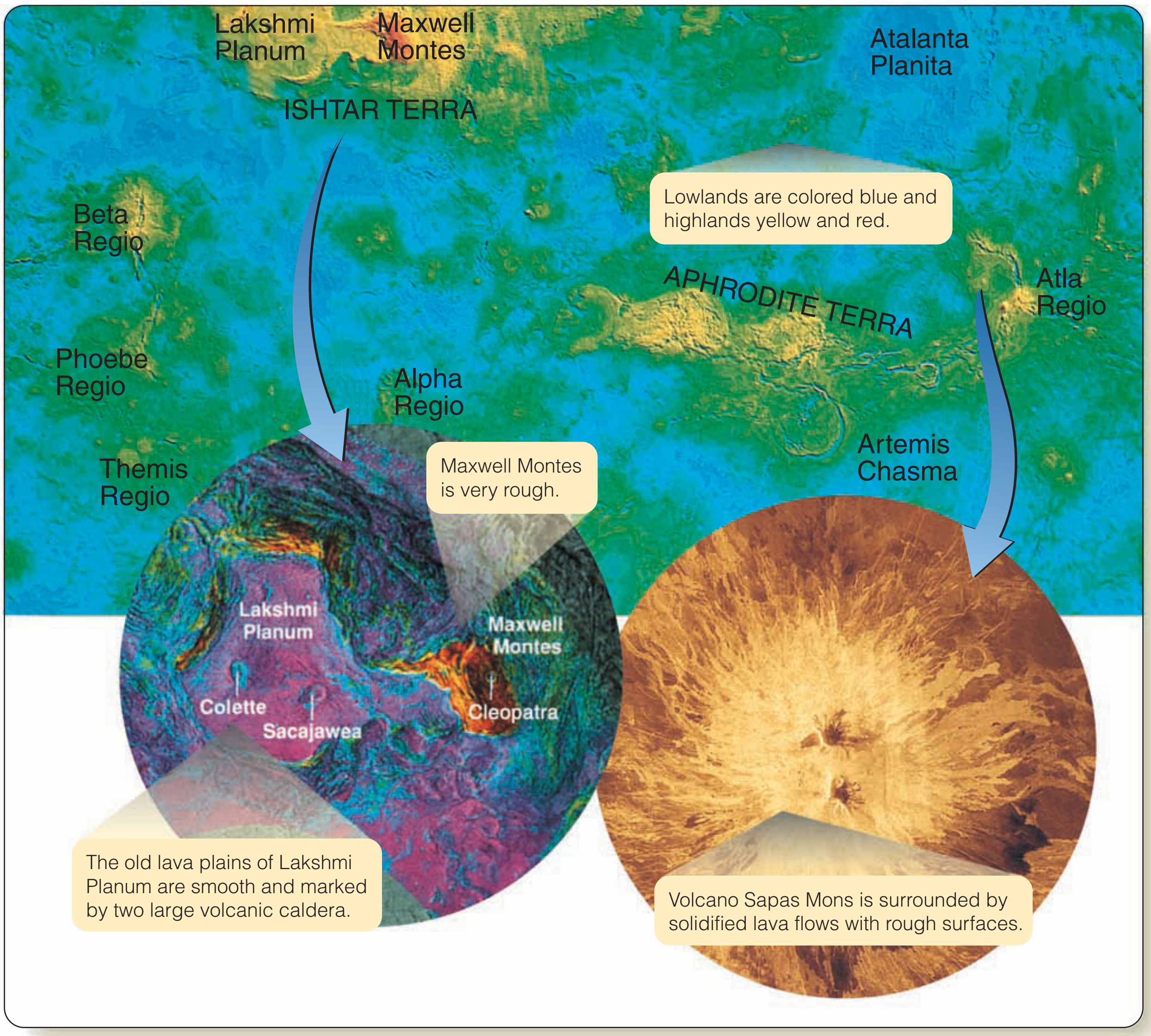


Figure 17-12

Three radar maps showing different aspects of Venus's surface. The main radar map here shows elevation over most of the surface, omitting the polar areas. The detailed map of Maxwell Montes and Lakshmi Planus are colored according to roughness, with orange representing the roughest terrain. The map of volcano Sapas Mons also shows roughness but is given an orange color to mimic the color of sunlight at the surface. (Maxwell and Lakshmi Planum map: USGS; Other maps: NASA)

the planet even warmer. That process could have vaporized any oceans that did exist and reduced the ability of the planet to purge its atmosphere of carbon dioxide. As more carbon dioxide was outgassed, the greenhouse effect grew even more severe. Thus, Venus was trapped in what is called a runaway greenhouse effect.

The intense heat at the surface may have affected the geology of Venus by making the crust more flexible so that it was unable to break into moving plates as on Earth. There are no signs of real global plate tectonics on Venus but rather evidence that convection currents below the crust are deforming the crust to create coronae and push up mountains such as Maxwell. Other mountains, like those around Ishtar Terra, appear to be folded mountains caused by limited horizontal motions in the crust, driven perhaps by convection in the mantle.

The small number of craters on the surface of Venus indicates that the entire crust has been replaced within the last half-billion years or so. The resurfacing may have occurred in a planetwide overturning as the old crust broke up and sank, and lava flows created a new crust. Hypothesizing such drama may not be necessary, however. Models of the climate on Venus show that an outburst of volcanism could increase the greenhouse effect and drive the surface temperature up by as much as 100°C. This could further soften the crust, increase the volcanism, and push the planet into a resurfacing episode. This type of catastrophe may happen periodically on Venus, or the planet may have had a single, geologically recent resurfacing event. In either case, un-Earthly Venus may eventually reveal more about how our own world works.

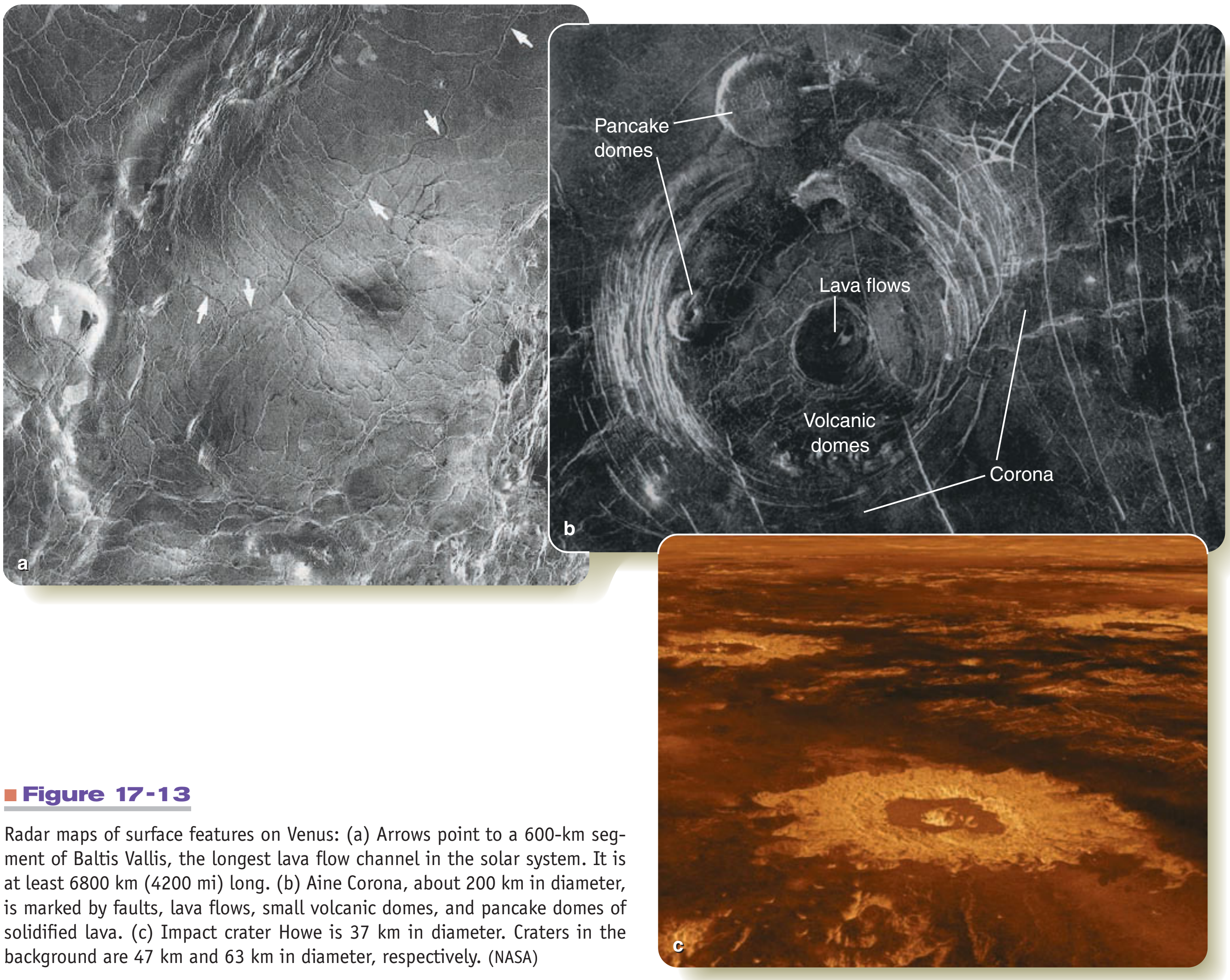


Figure 17-13

Radar maps of surface features on Venus: (a) Arrows point to a 600-km segment of Baltis Vallis, the longest lava flow channel in the solar system. It is at least 6800 km (4200 mi) long. (b) Aine Corona, about 200 km in diameter, is marked by faults, lava flows, small volcanic domes, and pancake domes of solidified lava. (c) Impact crater Howe is 37 km in diameter. Craters in the background are 47 km and 63 km in diameter, respectively. (NASA)

◄ SCIENTIFIC ARGUMENT ►

What evidence indicates that Venus does* not *have plate tectonics?

This argument must cite evidence and use comparison. On Earth, plate tectonics is identifiable by the worldwide network of faults, subduction zones, volcanism, and folded mountain chains that outline the plates. Although some of these features are visible on Venus, they do not occur in a planetwide network that outlines multiple plates. Volcanism is widespread, but folded mountain ranges occur in only a few places. Rather than being dominated by the horizontal motion of rigid crustal plates, Venus may have a more flexible crust dominated instead by vertical tectonics, for example, rising plumes of molten rock that strain the crust to produce coronae or that break through to form volcanoes and lava flows.

Earth and Venus are sibling worlds in some ways, but in other ways they seem to be no more than distant cousins. Now build an argument to compare atmospheres. **Why isn't Earth's atmosphere like that of Venus?**

◄ ►

17-6 Mars

Mars is a medium-sized world about half the diameter of Earth (■ Celestial Profile 6). The surface is old, cratered, and marked by volcanoes, but as you explore, watch for evidence that water once flowed there.

The Atmosphere of Mars

The Martian air contains 95 percent carbon dioxide, 3 percent nitrogen, and 2 percent argon. That is much like the chemical composition of the air on Venus, but the Martian atmosphere is very thin, less than 1 percent as dense as Earth's atmosphere, one ten-thousandth as dense as Venus's atmosphere.

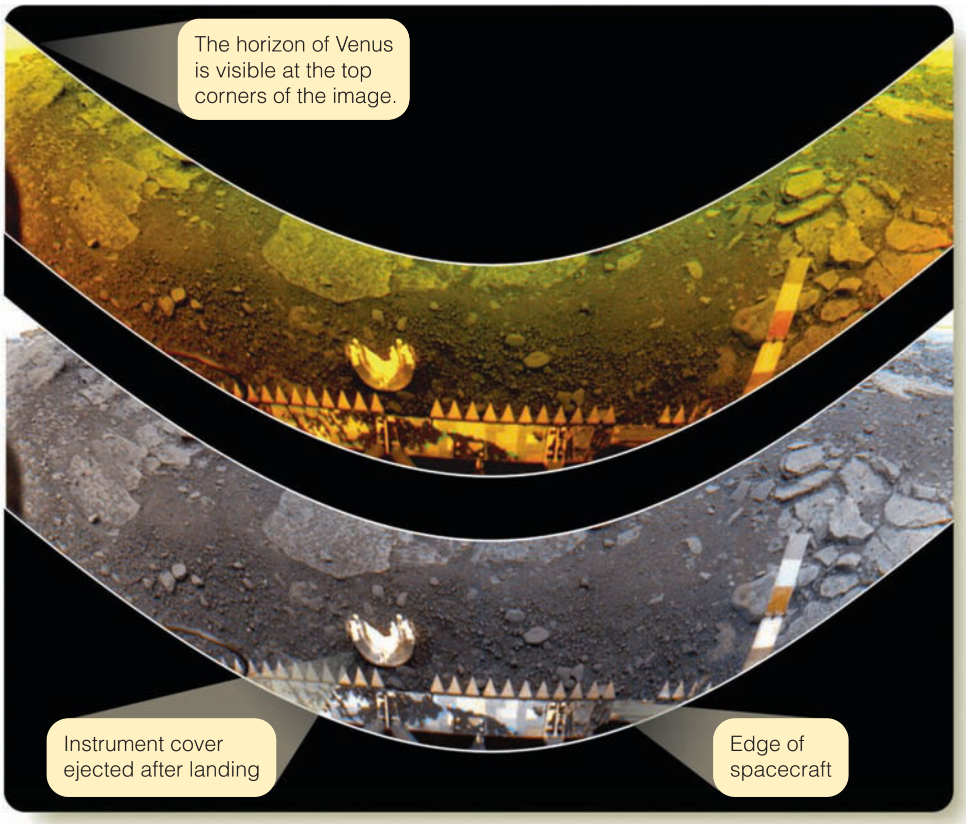


Figure 17-14

The Venera 13 lander touched down on Venus in 1982 and carried a camera that swiveled from side to side to photograph the surface. The orange glow is produced by the thick atmosphere; when that is corrected digitally, you can see that the rocks are dark gray. Isotopic analysis suggests they are basalts. (NASA)

There is very little water in the Martian atmosphere, and the polar caps are composed of frozen water ice coated over by frozen carbon dioxide ("dry ice"). As summer comes to a Martian hemisphere, planetary scientists observe the carbon dioxide in that polar cap turning from solid to vapor and adding carbon dioxide to the atmosphere, while winter in the opposite hemisphere is freezing carbon dioxide out of the atmosphere and adding it to that polar cap.

Liquid water cannot survive on the surface of Mars because the air pressure is too low. Any liquid water would immediately boil away; and if you stepped out of a spaceship on Mars without your spacesuit, your body heat would make your blood boil. Whatever water is present on Mars must be frozen in the polar caps or in the form of **permafrost** within the soil.

Although the present atmosphere of Mars is very thin, you will see evidence that the climate once permitted liquid water to flow over the surface, so Mars must have once had a thicker atmosphere. As a Terrestrial planet, it should have outgassed significant amounts of carbon dioxide, nitrogen, and water vapor; but because it was small, it could not hold onto its gases. The escape velocity on Mars is only 5 km/s, less than half of Earth's, so it was easier for rapidly moving gas molecules to escape into space. Another factor is the temperature of a planet. If Mars had been colder, the gas molecules in its atmosphere would have been traveling more slowly and would not have escaped as easily. You can see this in ■ Figure 17-15, which plots the escape velocity of each planet versus the temperature of the region from which molecules would escape. For Earth, the temperature is that of the upper atmosphere. For Mercury, the temperature is that of the hot rocky surface. Clearly, small worlds cannot keep atmospheric gases easily.

A further problem is that Mars has no ozone layer to protect its atmosphere from ultraviolet radiation. (Sunbathing on Mars would be a fatal mistake.) The ultraviolet photons can break atmospheric molecules up into smaller fragments, which escape more easily. Water, for example, can be broken up into hydrogen and oxygen. Thus, Mars is large enough to have had a substantial atmosphere when it was young, and may have had water falling as rain and collecting in rivers and lakes. It gradually lost much of its atmosphere and is now a cold, dry world.

Exploring the Surface of Mars

If you ever visit another world, Mars may be your best choice. You will need a heated, slightly pressurized spacesuit with air, water, and food, but Mars is much more hospitable than the moon or Venus. It is also more interesting, with weather, complex geology, and signs that water once flowed over its surface. You might even hope to find traces of ancient life hidden in the rocks.

Spacecraft have been visiting Mars for almost 40 years, but the pace has picked up recently. A small fleet of spacecraft has gone into orbit around Mars to photograph and analyze its surface, and five spacecraft have landed. Two Viking landers touched down in 1976, and three rovers have landed in recent years. Pathfinder and its rover Sojourner landed in 1997. Rovers Spirit and Opportunity landed in January 2004 and carried sophisticated instruments to explore the rocky surface (■ Figure 17-16) The Phoenix robot laboratory landed in the north polar region in 2008.

Data recorded by orbiting satellites show that the southern hemisphere of Mars is a heavily cratered highland region estimated to be at least 2 to 3 billion years old. The northern hemisphere is mostly a much younger lowland plain with few craters (■ Figure 17-17). This lowland plain may have been smoothed by lava flows, but growing evidence suggests that it was once filled with an ocean, a controversial hypothesis discussed in the next section.

Volcanism on Mars is dramatically evident in the Tharsis region, a highland region of volcanoes and lava flows bulging 10 km (6 mi) above the surrounding surface. A similar uplifted volcanic plain, the Elysium region, is more heavily cratered and eroded and appears to be older than the Tharsis bulge. The lack of many impact craters suggests that some volcanoes have been active within the last few hundred million years. There is no reason to think the volcanoes are completely dead.

All of the volcanoes on Mars are shield volcanoes, which are produced by hot spots penetrating upward through the crust. Shield volcanoes are not related to plate tectonics and are not evidence of plate motion on Mars. In fact, the largest volcano on Mars, Olympus Mons, provides clear evidence that plate tectonics has not been significant on Mars. Olympus Mons is 600 km (370 mi) in diameter at its base and rises 21 km (13 mi) high. The largest volcano on Earth is Mauna Loa in Hawaii, rising only 10 km (6 mi) above its base on the seafloor. Mauna Loa is so heavy that it has sunk into Earth's crust, producing an undersea moat around its base. In contrast, Olympus Mons, two times higher, has no moat and is supported entirely by the Martian crust (■ Figure 17-18). Evidently, the crust of Mars is much stronger than Earth's.

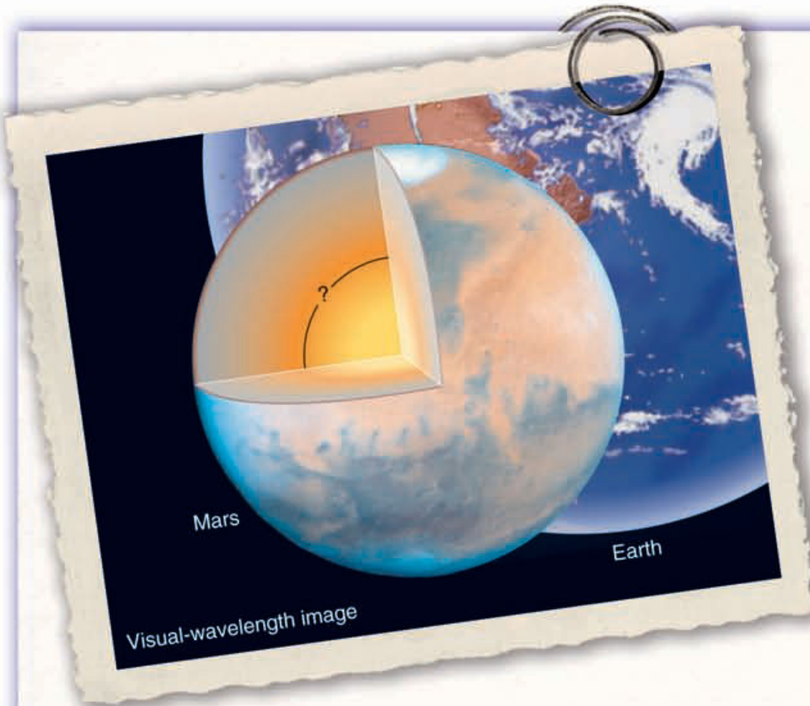
When the crust of a planet is strained, it may break, producing faults and rift valleys. Near the Tharsis region is a great valley, Valles Marineris (Figure 17-18), named after the Mariner spacecraft that first photographed it. This valley is a block of crust that has dropped downward along parallel faults. Erosion and landslides have further modified the valley into a great canyon stretching almost one-fifth of the way around the planet. It is four times deeper, nearly ten times wider, and over ten times longer than the Grand Canyon. The number of craters in the valley indicates that it is 1 to 2 billion years old, placing its origin sometime before the end of the most active volcanism in the Tharsis region.

Before you can tell the story of Mars, you should consider a difficult issue—water. How much water has Mars had, how much has been lost, and how much remains?

Searching for Water on Mars

You would hardly expect water on the surface of Mars. It is a cold, dry desert world. However, observations from orbiting spacecraft have revealed landforms that suggest there was once water on Mars, and rovers on the surface have turned up further traces of water.

In 1976, the two Viking spacecraft reached orbit around Mars and photographed its surface. Those photos revealed two kinds of water-related features. **Outflow channels** appear to have



Mars has half the diameter of Earth and probably retains some internal heat, but the size and composition of its core are not well known. (NASA)

Celestial Profile 6: Mars

Motion:

Average distance from the sun	1.52 AU (2.28 × 10^8 km)
Eccentricity of orbit	0.093
Inclination of orbit to ecliptic	1.9°
Orbital period	1.881 y (686.95 d)
Period of rotation (sidereal)	24.62 h
Inclination of equator to orbit	24.0°

Characteristics:

Equatorial diameter	6.80 × 10^3 km (0.531 $D_\oplus$)
Mass	6.42 × 10^{25} kg (0.108 $M_\oplus$)
Average density	3.94 g/cm^3 (3.3 g/cm^3 uncompressed)
Surface gravity	0.38 Earth gravity
Escape velocity	5.0 km/s (0.45 $V_\oplus$)
Surface temperature	−140° to 15°C (−220° to 60°F)
Average albedo	0.16
Oblateness	0.009

Personality Point:

Mars is named for the god of war. Minerva was the goddess of defensive war, but Bullfinch's *Mythology* refers to Mars's "savage love of violence and bloodshed." You can see the planet glowing reddish-orange from Earth, reminiscent of blood to cultures throughout history, because of iron oxides in its soil.

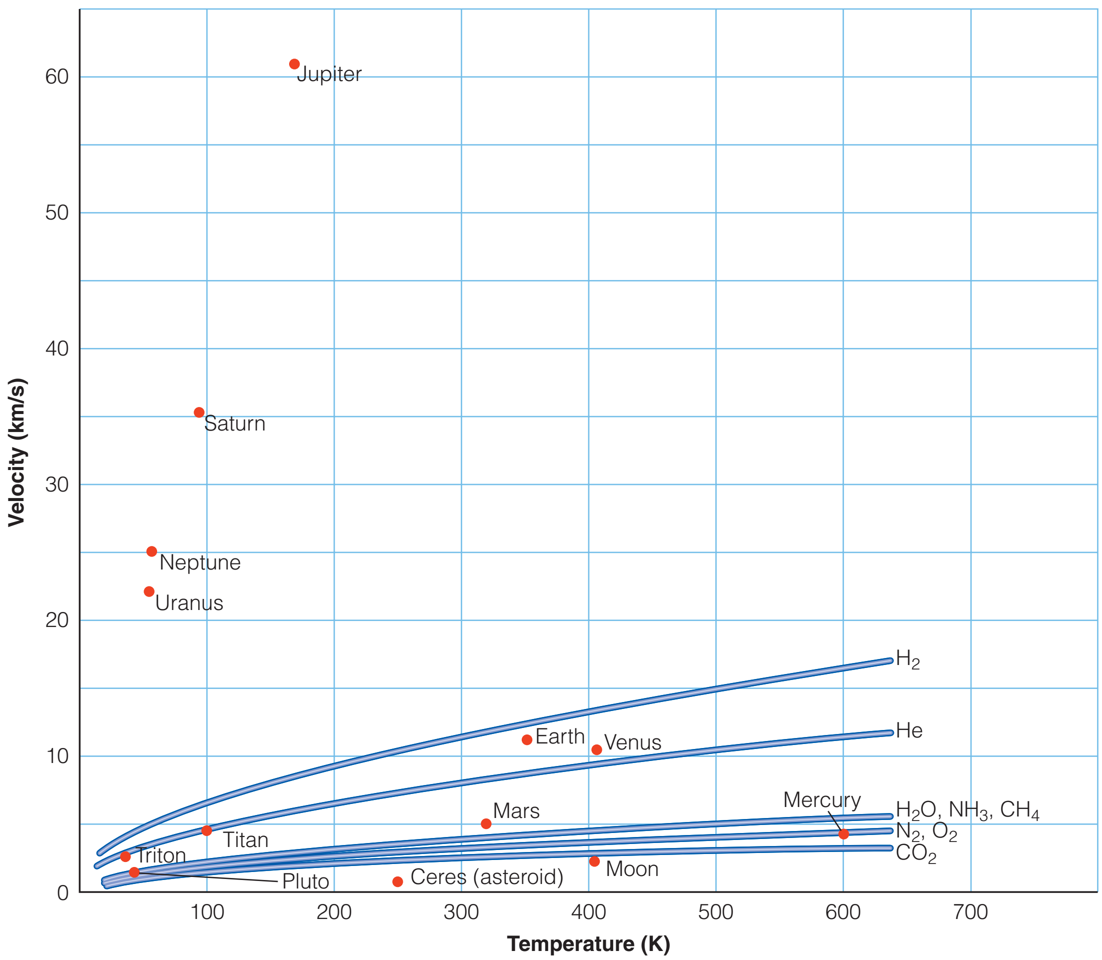


Figure 17-15

This plot shows the ability of planets to retain atmospheres. Dots represent the escape velocity and temperature of various solar system bodies. The lines represent the typical highest velocities of gas molecules of various masses. The Jovian planets have high escape velocities and can hold on to even the lowest-mass molecules. Mars can hold only the more massive molecules, and the moon has such a low escape velocity that all gas molecules can escape.

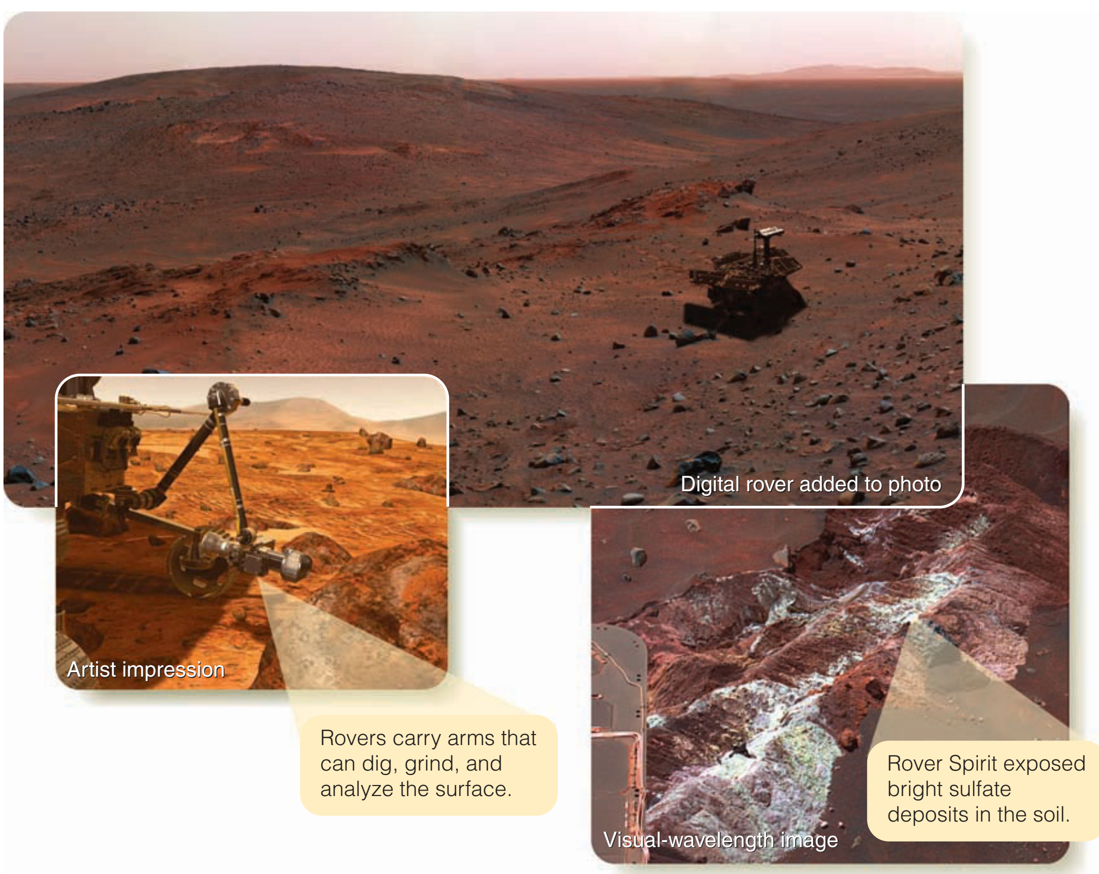


Figure 17-16

About the size of riding lawn mowers, rovers Spirit and Opportunity were directed from Earth to move across the surface of Mars, explore features, dig in the soil, grind the surfaces of rocks to expose their interiors and make spectroscopic analyses. Rover Spirit's discovery of sulfate deposits in the soil confirms other evidence that a body of salty water once covered the area and evaporated, leaving the sulfates behind. (NASA/JPL-Caltech/Cornell)

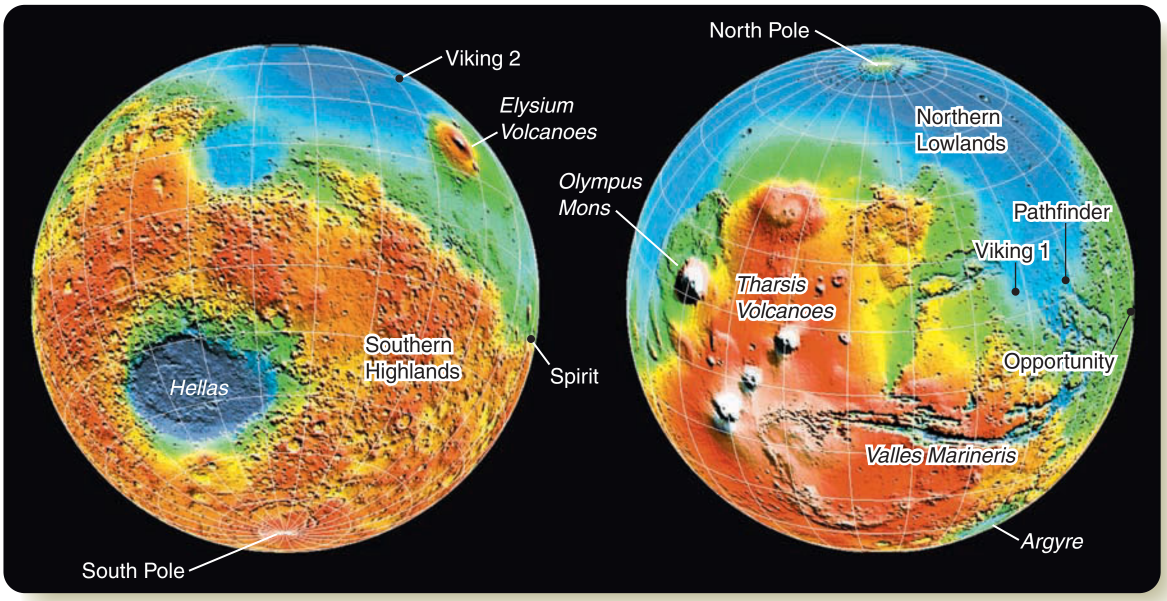
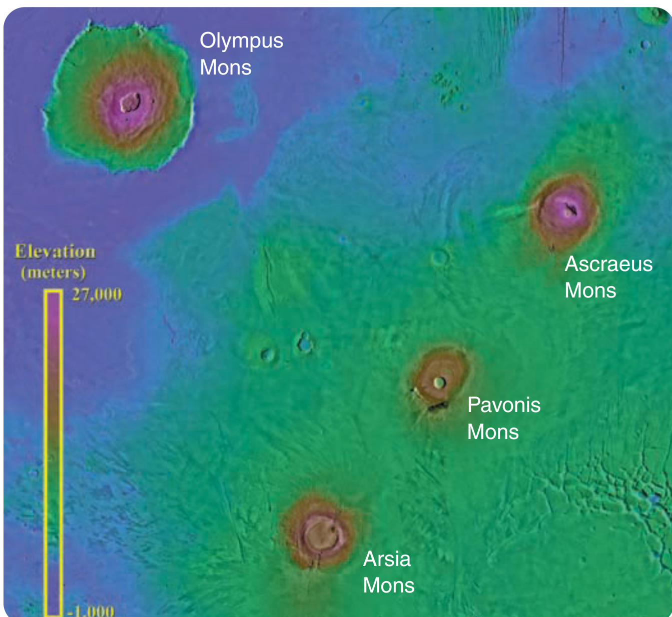
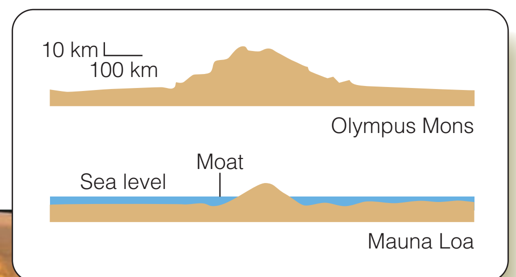


Figure 17-17

These hemisphere maps of Mars are color-coded to show elevation. The northern lowlands lie about 4 km below the southern highlands. Volcanoes are very high (white), and the giant impact basins, Hellas and Argyre, are low. Note the depth of the canyon Valles Marineris. (NASA)





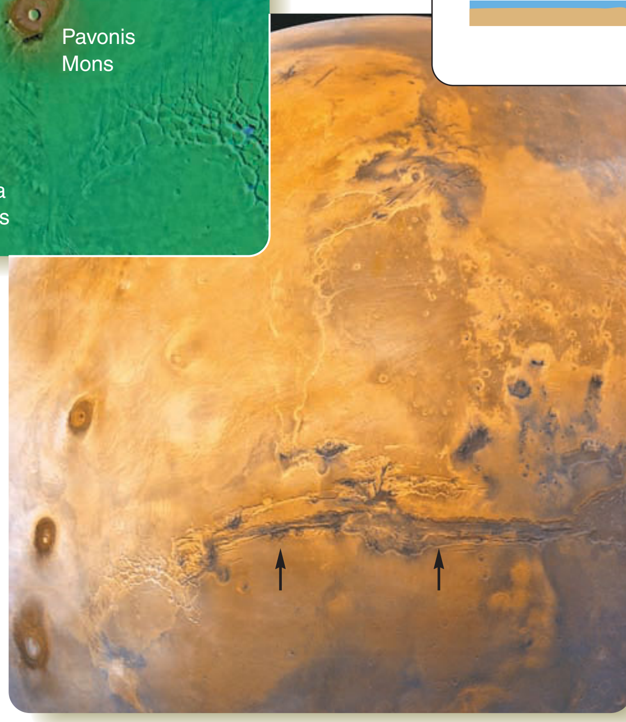


Figure 17-18

High volcanoes and deep canyons mark the surface of Mars. Olympus Mons, a shield volcano, is much larger than the largest volcano on Earth. In this false-color image, three other volcanoes are visible. Those three volcanoes are also visible in the photo along with the canyon Valles Marineris (indicated by black arrows), which stretches as far as the distance from New York to Los Angeles. (Four volcanoes: © Calvin J. Hamilton, Columbia, Maryland; Valles Marineris: NASA/USGS)

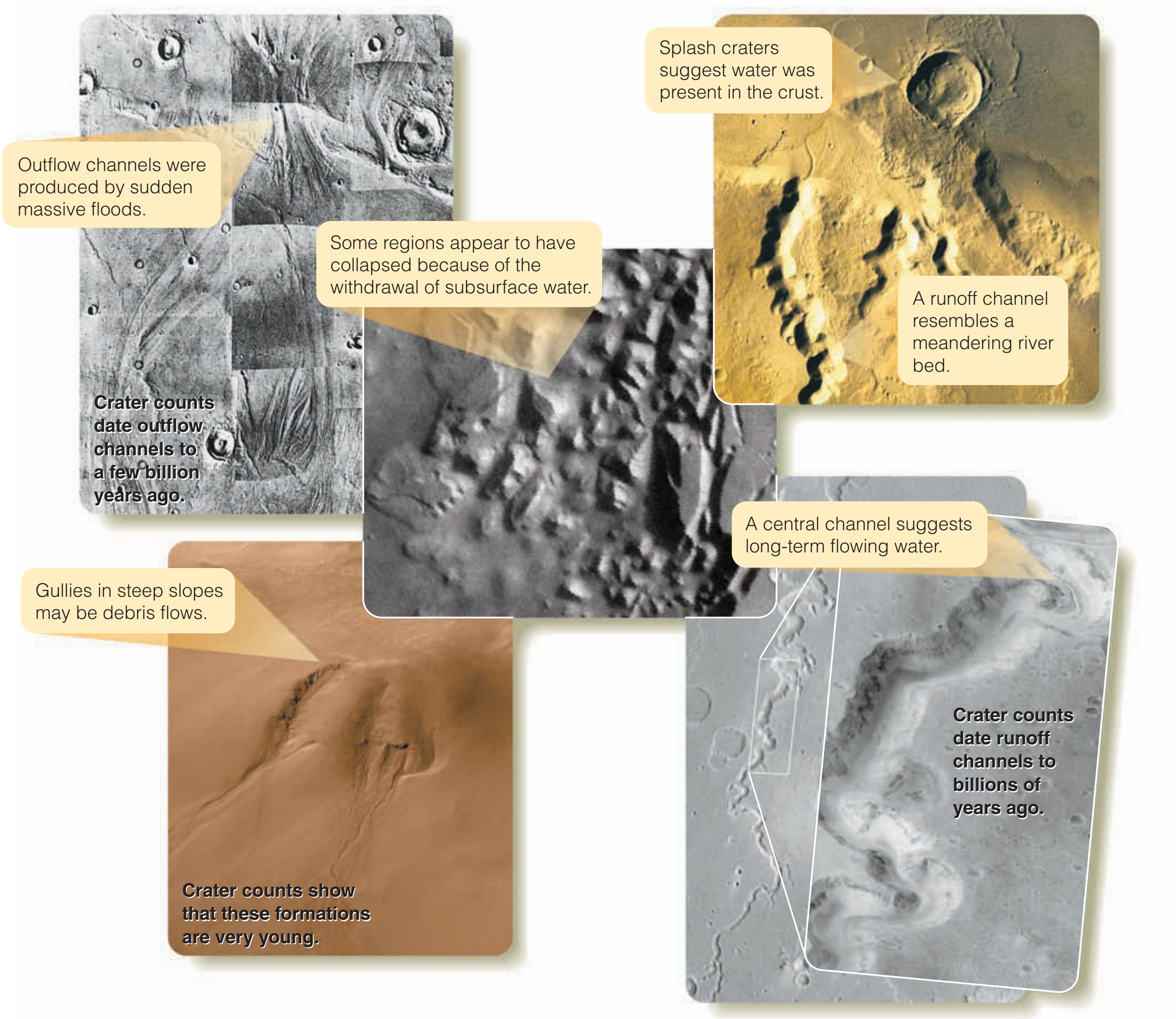


Figure 17-19

These visual-wavelength images made by the Viking orbiters and Mars Global Surveyor show some of the features that suggest liquid water on Mars. Outflow channels and runoff channels are old, but some gullies may be quite recent. (Malin Space Science Systems and NASA)

been cut by massive floods carrying as much as 10,000 times the volume of water flowing down the Mississippi River. In a matter of hours or days, such floods swept away geological features and left scarred land such as that shown in ■ Figure 17-19. In contrast, **valley networks** look like meandering riverbeds with sandbars, deltas, and tributaries typical of streams that flowed for extended periods of time. The number of craters on top of these features reveals that they are quite old.

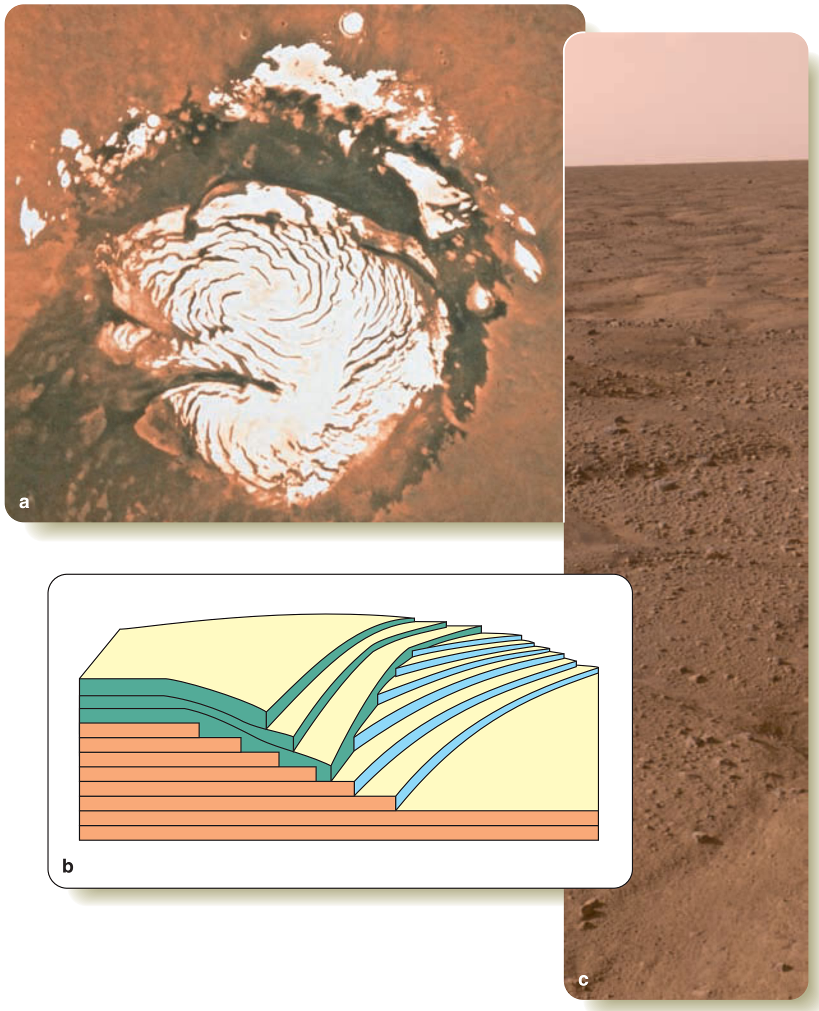
Images made from orbit also show regions of jumbled terrain, suggesting that subsurface ice may have melted and drained away. Gullies leading down slope suggest water seeping from underground sources. The terrain at the edges of the northern lowlands has been compared to shorelines, and some scientists suspect that the northern lowlands were filled with an ocean roughly 3 billion years ago. Look again at Figure 17-17, where the lowlands have been color-coded blue, and notice the major outflow channels leading from the highlands into the lowlands northwest of the Viking-1 landing site and southeast of the Pathfinder landing site, like rivers flowing into an ocean. Careful measurements of the location of the hypothetical ocean's shoreline indicate that it might originally have been an enormous impact basin.

Spacecraft in orbit around Mars have used remote instruments to detect large amounts of water frozen in the soil. A radar study has found frozen water extending at least a kilometer beneath both polar caps.

Rovers Spirit and Opportunity were both targeted to land in areas suspected of having had water on their surfaces, and they each made important discoveries. Using close-up cameras, they found small spherical concretions of the mineral hematite (nicknamed "blueberries") that must have formed in water. In other places, they found layers of sediments with ripple marks and crossed layers showing they were deposited in moving water (■ Figure 17-20). Chemical analysis revealed minerals in the soil such as sulfates that would have been left behind when standing water evaporated.

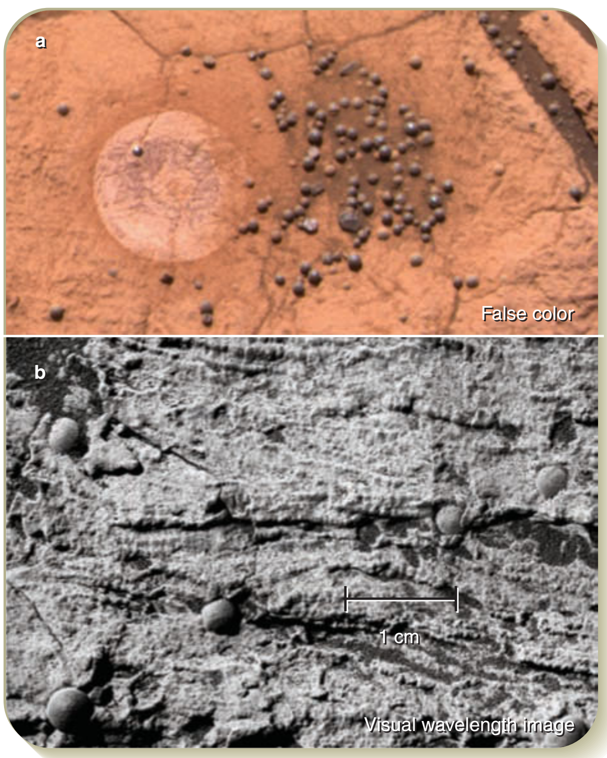
Orbiting spacecraft have photographed layered terrain near the polar caps (■ Figure 17-21). Year after year dust accumulates on the polar caps and is then left behind in a layer when the polar caps vaporize in the spring. Over periods of thousands of years, deep layers can develop. What is significant is that orbiters have photographed newer layers oriented differently from older underlying layers, showing that the climate and wind patterns on Mars have changed repeatedly. These layers suggest that the climate on Mars may vary because of cyclic changes in the axis orientation and orbital shape of the planet. Recall from Chapter 2 that Earth is affected by such cycles.

Mars has water, but it is hidden. The climate has changed time after time, but the atmosphere has gradually grown thinner. The oceans and lakes are gone. The last of the water on Mars is in the polar caps or frozen in the crust. Water is the first necessity of life, so the evidence for running water long ago on Mars is



■ **Figure 17-21**

(a) The north polar cap of Mars is made of many regions of ice separated by narrow valleys free of ice. (NASA) (b) In some regions of the polar cap, layers of ice plus dust with different orientations are superimposed, suggesting periodic changes in the Martian climate. (Adapted from a diagram by J. A. Cutts, K. R. Balasius, G. A. Briggs, M. H.Carr, R. Greeley, and H. Masursky) (c) This view from the Phoenix lander shows the landscape of Mars's north polar plains, including polygonal cracks believed to result from seasonal expansion and contraction of ice under the surface. (NASA/JPL-Caltech/University of Arizona)



■ **Figure 17-20**

(a) Rover Opportunity photographed hematite concretions ("blueberries") weathered from rock. The round mark is a spot cleaned by the rover. The spheres appear to have grown as minerals collected around small crystals in the presence of water. Similar concretions are found on Earth. (b) The layers in this rock were deposited as sand and silt in rapidly flowing water. From the way the layers curve and cross each other, geologists can estimate that the water was at least ten centimeters deep. A few "blueberries" are also visible in this image. (NASA/JPL/Cornell/USGS)

How Do We Know? | 17-3

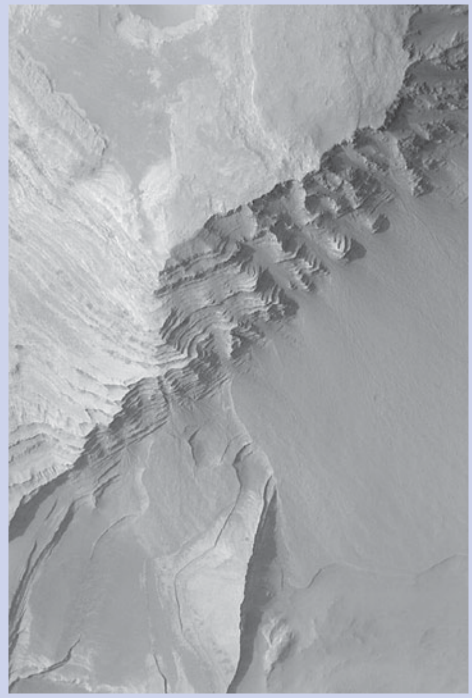
The Present Is the Key to the Past

How can we know what happened long ago if there weren't any witnesses? Geologists are fond of saying "The present is the key to the past." By that they mean that you can learn about the history of Earth by looking at the present condition of Earth's surface. The position and composition of various rock layers in the Grand Canyon, for example, tell you that the western United States was once at the floor of an ocean. This principle of geology is relevant today as you try to understand the history of other worlds such as Venus and Mars.

In the late 1700s, naturalists first recognized that the present gave them clues to the history of Earth. At the time that was astonishing, because most people assumed that Earth had no history. That is, they assumed either that Earth had been created in its present state as described in the Old Testament or that Earth was eternal. In either case, people commonly assumed that the hills and mountains they saw around them had always existed more or less as they were. By the 1700s, naturalists began to see evidence that the hills and mountains were not eternal but were the result of past processes and were slowly changing. That gave rise to the idea that Earth had a history.

As those naturalists made the first attempts to thoughtfully and logically explain the nature of Earth by looking at the evidence, they were inventing modern geology as a way of understanding Earth. What Copernicus, Kepler, and Newton did for the heavens in the 1500s and 1600s, the first geologists did for Earth beginning in the 1700s. Of course, the invention of geology as the study of Earth led directly to the modern attempts to understand the geology of other worlds.

Geologists and astronomers share a common goal: They are attempting to reconstruct the past (How Do We Know? 16-2). Whether you study Earth, Venus, or Mars, you are looking at the present evidence and trying to reconstruct the past history of the planet by drawing on observations and logic to test each step in the story. How did Venus get to be covered with lava, and how did Mars lose its atmosphere? The final goal of planetary astronomy is to draw together all of the available evidence (the present) to tell the story (the past) of how the planet got to be the way it is. Those first geologists in the 1700s would be fascinated by the stories planetary astronomers tell today.



Layers of rock in the Martian crater Terby hint at a time when the crater was filled with a lake. (NASA/JPL/Malin Space Science Systems)

exciting. Someday an astronaut may scramble down an ancient Martian streambed, turn over a rock, and find a fossil.

The History of Mars

Did Mars ever have plate tectonics? Where did the water go? These fundamental questions challenge you to assemble the evidence and hypotheses for Mars and tell the story of its evolution (■ "How Do We Know?" 17-3).

The four-stage history of Mars is a case of arrested development. The planet began by differentiating into a crust, mantle, and core. Studies of its rotation reveal that it has a dense core. The Mars Global Surveyor spacecraft detected no planetwide magnetic field, but it did find regions of the crust with fields a bit over 1 percent as strong as Earth's. Apparently, the young Mars had a molten iron core and generated a magnetic field, which became frozen into parts of the crust. The core must have cooled quickly and shut off the dynamo effect that was producing the planetwide field. The magnetic regions of the crust remain behind like fossils.

The crust of Mars is now quite thick, as shown by the mass of Olympus Mons, but it was thinner in the past. Cratering may have broken or at least weakened the crust, triggering lava flows that flooded some basins. Most of the northern hemisphere may have been a huge impact basin that was later filled with water, creating a martian ocean that has since vanished. Mantle convection may have pushed up the Tharsis and Elysium volcanic regions and broken the crust to form Valles Marineris, but moving crustal plates never dominated Mars. There are no folded mountain ranges on Mars and no signs of plate boundaries. As the small planet cooled rapidly, its crust grew thick and immobile.

The large size of volcanoes on Mars is evidence that the crust does not move. On Earth, volcanoes like those that formed the Hawaiian Islands occur over rising currents of hot material in the mantle. Because the plate moves, the hot material heats the crust in a string of locations and forms a chain of volcanoes instead of a single large feature. The Hawaiian Islands are merely the most recent of a series of volcanic islands called the Hawaiian–Emperor island chain (see page 350), which stretches nearly 3800 km (2400 mi) across the Pacific Ocean floor. A lack of plate motion on Mars would have allowed a rising current of magma to heat the crust repeatedly in the same place and build a very large volcanic cone.

At some point in the history of Mars, water was abundant enough to flow over the surface in great floods and may have filled lakes and oceans, but the age of liquid water must have ended over 3 billion years ago. The climate on Mars has changed as atmospheric gases and water were lost to space and as water was frozen into the soil as permafrost.

For Mars, the fourth stage of planetary development has been one of moderate activity and slow decline. Volcanoes may still occasionally erupt, but this medium-sized planet has lost much of its internal heat, and most volcanism occurred long ago. The atmosphere is thin, and the surface is a forbiddingly dry, cold desert.

The Moons of Mars

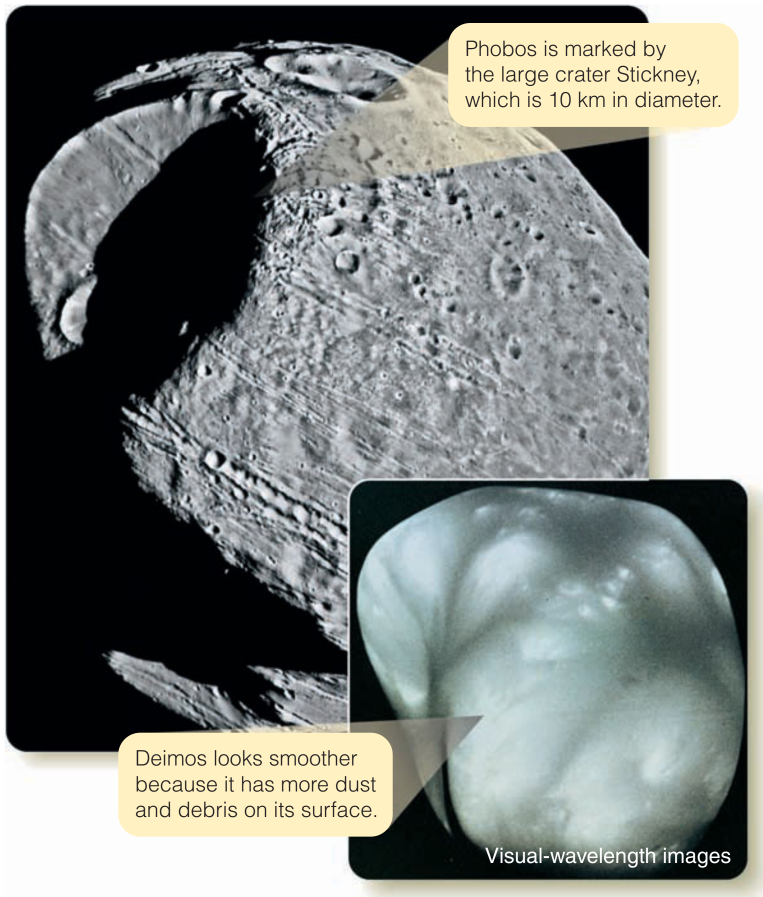
Unlike Mercury or Venus, Mars has moons. Small and irregular in shape, Phobos ($28 \times 23 \times 20$ km in diameter) and Deimos ($16 \times 12 \times 10$ km) are probably captured asteroids.

Photographs reveal a unique set of narrow, parallel grooves on Phobos (■ Figure 17-22a). Averaging 150 m (500 ft) wide and 25 m (80 ft) deep, the grooves run from Stickney, the largest crater, to an oddly featureless region on the opposite side of the satellite. One theory suggests that the grooves are deep fractures caused by the impact that formed Stickney.

Deimos not only has no grooves, but it also looks smoother because of a thicker layer of dust on its surface. This material partially fills craters and covers minor surface irregularities (Figure 17-22b). It seems likely that Deimos experienced many collisions in its past, so its fractures may be hidden below the debris.

The debris on the surfaces of the moons raises an interesting question: How can the weak gravity of small bodies hold any fragments from meteorite impacts? The escape velocity on Phobos is only about 12 m/s (40 ft/s). An athletic astronaut who could jump 2 m (6 ft) high on Earth could jump 2.8 km (1.7 mi) on Phobos. Most of the fragments from an impact should escape, but the slowest particles could fall back in the weak gravity and accumulate on the surface.

Because Deimos is smaller than Phobos, its escape velocity is smaller, so it seems surprising that it has more debris on its surface. This may be related to Phobos's orbit close to Mars. The Martian gravity is almost strong enough to pull loose material off of Phobos's surface, so Phobos may be able to retain less of its cratering impact debris.



■ **Figure 17-22**

The moons of Mars are too small to pull themselves into spherical shape. The two moons, shown here to scale, were named for the mythical attendants of the god of war, Mars. Phobos was the god of fear, and Deimos was the god of dread. (Phobos: Damon Simonelli and Joseph Ververka, Cornell University/NASA; Deimos: NASA)

◄ SCIENTIFIC ARGUMENT ►

Why would you be surprised to find volcanism on Phobos or Deimos?

This argument hinges on the principle that the larger a world is, the more slowly it loses its internal heat. It is the flow of that heat from the interior through the surface into space that drives geologic activity such as volcanism and plate motion. A small world, like Earth's moon, cools quickly and remains geologically active for a shorter time than does a larger world like Earth. Phobos and Deimos are not just small; they are tiny. However they formed, any interior heat would have leaked away very quickly; with no energy flowing outward, there can be no volcanism.

Some futurists suggest that the first human missions to Mars will not land on the planet's surface but instead will build a colony on Phobos or Deimos. These plans speculate that there may be water deep inside the moons that colonists could use. **What would happen to water released in the sunlight on the surface of such small worlds?**

◄ ►

What Are We? Comfortable

Many planets in the universe probably look like the moon and Mercury—small, airless, and cratered. Some are made of stone; and some, because they formed farther from their star, are made mostly of ices. If you randomly visited a planet anywhere in the universe, you would probably stand on a moonscape.

Earth is unusual but not rare. The Milky Way Galaxy contains over 100 billion stars, and over 100 billion galaxies are visible with existing telescopes. Most of those stars probably have planets, and although many planets may look like Earth's moon and Mercury, there also must be plenty of Earth-like worlds.

As you look around your planet, you should feel comfortable living on such a beautiful planet, but it was not always such a nice place. The craters on the moon and the moon rocks returned by astronauts show that the moon formed with a sea of magma. Mercury seems to have had a similar history, so Earth probably formed the same way. Its surface was once a seething ocean of liquid rock swathed in a hot, thick atmosphere, torn by eruptions of more rock, explosions of gas from the interior, and occasional impacts from space. The moon and Mercury assure you that that is the way Terrestrial planets begin. Earth has evolved to become your home world, but Mother Earth has had a violent past.

Study and Review

Summary

- The Terrestrial worlds differ mainly in size, but they all have low-density crusts, **mantles (p. 345)** of dense rock, and metallic cores.
- Earth's moon is included in this study because it is a complex world and illustrates important principles such as cratering and flooding by lava.
- **Comparative planetology (p. 344)** leads you to expect that cratered surfaces are old, that heat flowing out of a planet drives geological activity, and that the nature of a planet's atmosphere depends on both the size of the planet and its temperature.
- Earth has passed through four stages in its history: (1) differentiation, (2) cratering, (3) flooding by lava and water, and (4) slow surface evolution. The other Terrestrial planets plus Earth's moon have also passed through versions of the same four stages, although the relative importance of the stages differs from object to object.
- Seismology, especially tracking the paths of earthquake ***P* waves (p. 346)** and ***S* waves (p. 346)**, shows that Earth has differentiated to form a metallic core that is partly liquid. The metallic liquid outer core generates Earth's magnetic field.
- Earth is dominated by **plate tectonics (p. 350)** that breaks the crust into moving plates, driven by heat flowing upward from the interior. The crustal plates float on denser mantle rocks that are described as **plastic (p. 348)** because they slowly flow under pressure.
- Evidence of plate tectonics includes **rift valleys (p. 350)**, **midocean rises (p. 350)**, **subduction zones (p. 350)**, and **folded mountain ranges (p. 350).** New crust is mostly created at midocean rises as a type of volcanic rock called **basalt (p. 350).**
- Earth's **primary atmosphere (p. 348)** was probably mostly carbon dioxide, but that gas mostly dissolved in seawater, and plant life has released oxygen, creating the present **secondary atmosphere (p. 348).**
- Gases such as carbon dioxide, methane, and water vapor trap heat in the Earth's environment, called the **greenhouse effect (p. 349).** Without the greenhouse effect, Earth would be uninhabitable, with an average temperature well below freezing.
- Human burning of fossil fuels has significantly increased the amount of carbon dioxide in Earth's atmosphere, causing a noticeable rise in average temperatures called **global warming (p. 349).**
- Earth's moon formed in a mostly molten state and differentiated, but it contains little metal and has an overall low density.
- Earth's moon is a small world and so has lost most of its internal heat and is no longer geologically active. Its old highlands are heavily cratered, but the lowlands are filled by lava flows that formed smooth **maria** (singular, **mare**) **(p. 352)** soon after the end of the heavy bombardment.
- The overall reflectivity, or **albedo (p. 352)**, of the moon is very low, even in the relatively bright highlands.
- The moon's surface is fractured by impacts, producing craters that are especially easy to see near the **terminator (p. 352)**, the moving boundary between the sun-lit and unlit parts of the moon.
- Debris blasted out of craters is called **ejecta (p. 354)** and can produce **secondary craters (p. 354).** Impacts have resulted in prominent features such as crater **rays (p. 354)** and **multiringed basins (p. 355).**
- Lunar rocks brought back to Earth include **vesicular basalts (p. 352)** from parts of the moon where the surface was covered by successive lava flows; **anorthosite (p. 352)**, which shows that a large portion of the moon was once a **magma ocean (p. 356);** and **breccias (p. 352)**, providing evidence that the moon has been repeatedly pounded by impacts.
- The **large-impact hypothesis (p. 353)** suggests the moon formed when an impact between the proto-Earth and a very large planetesimal surrounded Earth with a disk of collision debris. The moon formed from that disk.
- Mercury is smaller than Earth but larger than Earth's moon. It is airless and has an old, heavily cratered surface.
- Mercury has a much higher density than Earth's moon and must have a large metallic core. Mercury may have suffered a major impact that blasted away low-density crustal rock and left it with a metallic core that is large in proportion to the diameter of the planet.
- **Lobate scarps (p. 359)** are long curving cliffs formed by compression on Mercury when its large metallic core solidified and contracted.
- Venus is almost as large as Earth. It has a thick, cloudy atmosphere of carbon dioxide that hides the surface from sight. It can be studied by radar mapping.
- Venus's carbon dioxide atmosphere drives an intense greenhouse effect and makes that planet's surface hot enough to melt lead.

- Venus is slightly closer to the sun than Earth, too warm for liquid water oceans to dissolve carbon dioxide from the atmosphere easily, and warm enough to start evaporating its oceans, leading to a runaway greenhouse effect.
- The hot crust of Venus is not dominated by plate tectonics but rather by volcanism and vertical tectonics, including **coronae (p. 363),** large circular uplifted regions. Crater counts show that the entire surface of Venus has been covered over or replaced in the past half billion years.
- **Composite volcanoes (p. 364)** are associated with subduction zones and plate boundaries on Earth, whereas **shield volcanoes (p. 364),** found on Earth, Venus, and Mars, are caused by rising columns of magma called hot spots that break through the crust from below.
- Mars is about half the size of Earth; it has a thin atmosphere and has lost much, but not all, of its internal heat.
- The loss of atmospheric gases depends on the size of a planet and its temperature. Mars is cold, but it is small and has a low escape velocity, and thus many of its lighter gases have leaked away.
- Some water may have leaked away from Mars as ultraviolet radiation from the sun broke it into hydrogen and oxygen, but some water is frozen in the polar caps, and as **permafrost (p. 368)** in the soil.
- **Outflow channels (p. 369)** and **valley networks (p. 372)** visible from Mars orbiters seem to have been cut by sudden floods or by longer-term drainage, but water cannot exist as a liquid on Mars now because of its low temperature and low atmospheric pressure. Therefore, conditions on Mars must have once been different, allowing liquid water to flow on the surface.
- The southern hemisphere of Mars is old cratered terrain, but some large volcanoes lie in the north. The size of these volcanoes strongly suggests that the crust does not move horizontally.
- Some volcanism may still occur on Mars, but because the planet is small, it has cooled and is not very active geologically.
- Orbiters have found evidence of large amounts of water frozen below the surface.
- Robot rovers have found clear signs that the Martian climate was different in the past and that liquid water flowed over the surface in at least some places. The northern lowlands may even have held an ocean at one time.
- The two moons of Mars are probably captured asteroids. They are small, airless, and cratered. They lost their internal heat long ago.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What are the four stages in the development of a Terrestrial planet?
2. Why would you expect planets to be differentiated?
3. How does plate tectonics create and destroy Earth's crust?
4. Why do astronomers suspect that Earth's original atmosphere was rich in carbon dioxide?
5. Why doesn't Earth have as many craters as Earth's moon or Venus?
6. What kind of erosion is now active on Earth's moon?
7. Discuss the evidence and hypotheses concerning the origin of Earth's moon.
8. Why do neither Earth nor Earth's moon have lobate scarps like the ones observed on Mercury?
9. How did Earth avoid the runaway greenhouse effect that made Venus so hot?
10. What evidence indicates that plate tectonics does not occur on Venus? On Mars?
11. What evidence suggests that Venus has been resurfaced within the last half-billion years?
12. Why is the atmosphere of Venus rich in carbon dioxide? Why is the atmosphere of Mars rich in carbon dioxide?
13. What evidence indicates that the climate of Mars has changed?
14. Why do astronomers conclude that the crust on Mars must be thicker than Earth's crust?
15. What evidence indicates that there has been liquid water on Mars?
16. **How Do We Know?** Why is heat flow the key to understanding a planet's surface activity?
17. **How Do We Know?** If memorizing facts is not the point of science, what is the point?
18. **How Do We Know?** How is the present the key to the past?

Discussion Questions

1. If you visited a planet in another solar system and discovered oxygen in its atmosphere, what might you guess about the environment there?
2. If liquid water is rare on the surface of planets, then most Terrestrial planets must have CO_2-rich atmospheres. Why?

Problems

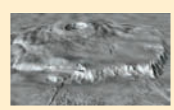
1. If the Atlantic seafloor is spreading at 30 mm/year and is now 6000 km wide, how long ago were Europe and North America in contact?
2. Earth is four times larger in diameter than its moon. How many times larger is it in surface area? In volume?
3. The smallest detail visible through Earth-based telescopes is about 1 second of arc in diameter. What size object would this represent on Earth's moon? What size object would this represent on Mars at closest approach to Earth? (*Hints:* See Reasoning with Numbers 3-1 and Appendix A.)
4. What is the maximum angular diameter of Phobos as seen from Earth? (*Hint:* See Reasoning with Numbers 3-1.)
5. How long would it take radio signals to travel from Earth to Venus and back if Venus were at its nearest approach to Earth? Repeat the calculation for Venus at its farthest distance from Earth.
6. Imagine that a spacecraft has landed on Mercury and is transmitting radio signals to Earth at a wavelength of 10.000 cm. When Mercury is seen from Earth in the evening sky, at its greatest angular distance east of the sun, it is moving toward Earth at its maximum possible relative speed of 47.9 km/s. To what wavelength must you tune your radio telescope to detect the signals? (*Hint:* See Reasoning with Numbers 6-2.)
7. Phobos orbits Mars at a distance of 9380 km from the center of the planet and has a period of 0.3189 day. Calculate the mass of Mars. (*Hint:* See Reasoning with Numbers 4-1; remember to use units of meters, kilograms, and seconds.)

Learning to Look

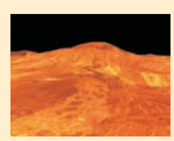
1. In this photo, Astronaut Alan Bean works at the Apollo 12 lander *Intrepid*. Describe the surface you see. What kind of terrain did they land on, for this, the second human landing on the moon?


NASA

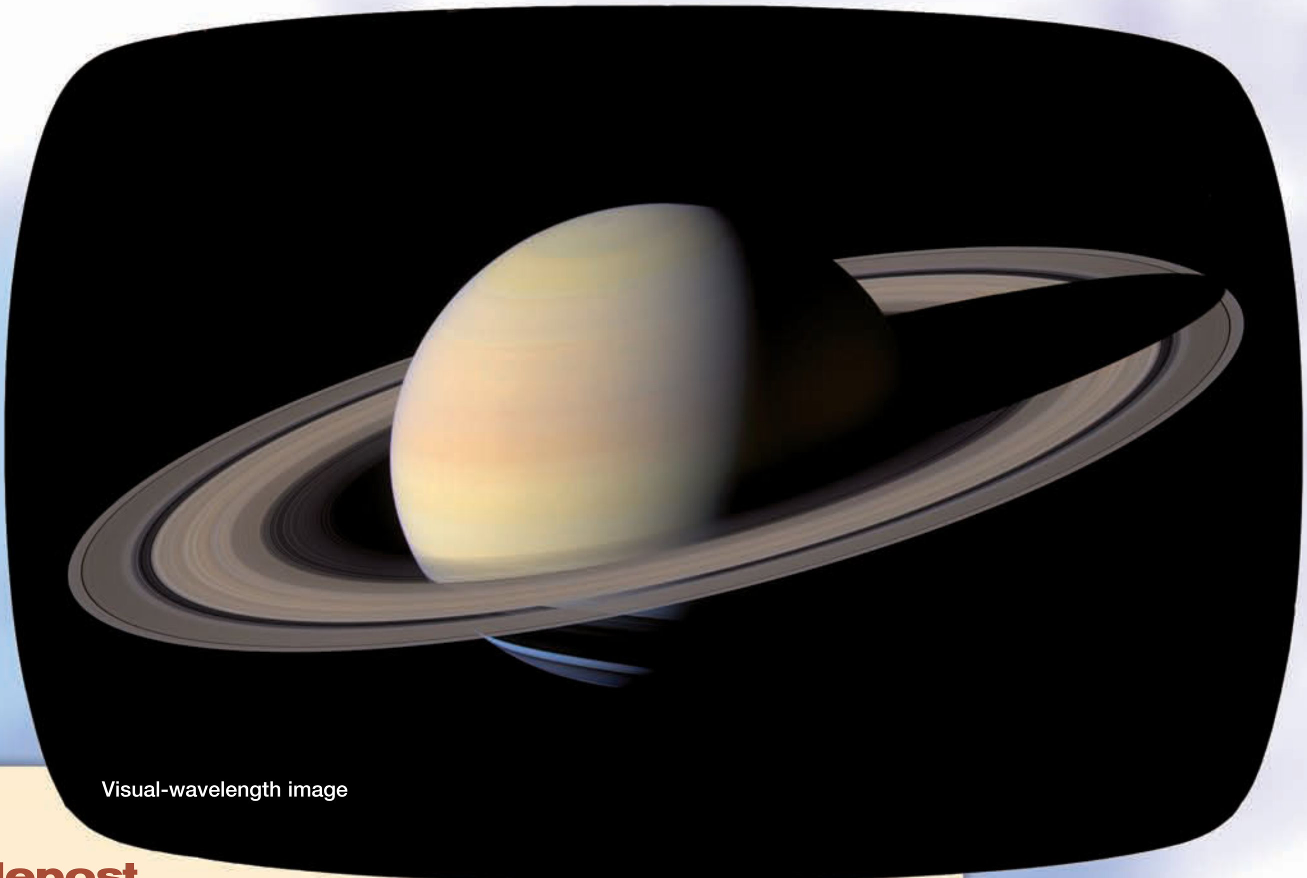
2. Olympus Mons on Mars is a very large volcano. In this image you can see multiple caldera superimposed at the top. What do those multiple caldera and the immense size of Olympus Mons indicate about the geology of Mars?


NASA

3. Volcano Sif Mons on Venus is shown in this radar image. What kind of volcano is it, and why is it orange in this image? What color would the rock be if you could hold it in your hands on Earth?


NASA

18 The Jovian Planets, Pluto, and the Kuiper Belt



Saturn is an intriguing world with beautiful rings. Its atmosphere is mostly hydrogen and helium, and it has no solid surface—the planet's interior is almost entirely liquid. The robot probe Cassini orbiting Saturn since 2004 has uncovered fascinating information about its rings and moons. (NASA/JPL)

Guidepost

In the last two chapters, you watched our solar system form and explored the Terrestrial planets. Now you can explore the outer solar system. The Jovian planets will be a challenge to your imagination; they are so un-Earthly, they would be unbelievable if you didn't have direct observational evidence to tell you what they are like. Nevertheless, the concepts of comparative planetology will continue to be faithful guides.

As you explore, you will find answers to four essential questions:

What are the properties of the Jovian (Jupiter-like) planets?

What evidence indicates that some moons in the outer solar system have been geologically active?

How are planetary rings formed and maintained?

What do Pluto and the Kuiper belt tell us about the formation of the solar system?

The planets of the solar system formed from small bodies in the solar nebula; to some extent, the planets are still affected by those small objects. In the next chapter, you will adapt your understanding of comparative planetology to study the last remains of the solar nebula: meteorites, asteroids, and comets.

Animated! This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.

There wasn't a breath in that land of death . . .

ROBERT SERVICE
THE CREMATION OF SAM MCGEE

THE SULFURIC acid clouds of Venus may seem totally alien to you, but compared with the planets of the outer solar system, Venus is almost like home. For example, the four Jovian planets have no solid surfaces. As you begin your study of these strange worlds, you will discover new principles of comparative planetology.

The worlds of the outer solar system can be studied from Earth, but much of what astronomers know has been radioed back to Earth from space probes. The Pioneer and Voyager probes flew past the outer planets in the 1970s and 1980s, the Galileo probe orbited Jupiter in the late 1990s, and the Cassini/Huygens orbiter and probe arrived at Saturn in 2004. The New Horizons probe will pass Pluto in 2015 and then sail on into the Kuiper belt. Throughout this discussion, you will find images and data returned by these robotic explorers (■ How Do We Know? 18-1).

18-1 A Travel Guide to the Outer Solar System

YOU ARE ABOUT to visit worlds that are truly un-Earthly. This travel guide will warn you what to expect.

How Do We Know? 18-1

Basic Science and Practical Technology

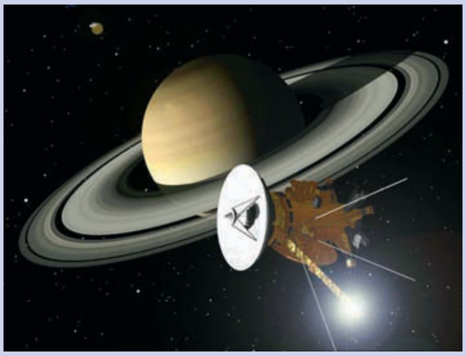
What practical value is there in sending a space probe to Saturn? Sending spacecraft to another world is expensive, and it may seem pointless when that world is hostile to human life. To resolve that question, you need to consider the distinctions among science, technology, and engineering.

Science is simply the logical study of nature. Although much scientific knowledge proves to have tremendous practical value, the only real goal of science is a better understanding of how nature works. Technology, in contrast, is the practical application of scientific knowledge to solve a specific problem. People trying to find a faster way to paint automobiles might use all the tools and techniques of science, but if their goal has some practical outcome, you would more properly call it technology rather than science. Engineering is the most practical form of technology. An engineer is likely to use well-understood technology to find a practical solution to a problem.

Of course, there are situations in which science and technology blur together. For example, humanity has a practical and urgent need to solve the HIV-AIDS problem; the world needs a cure and methods of prevention. Unfortunately, scientists don't understand HIV itself or viruses in general well enough to design a simple solution, and thus much of the HIV research work involves going back to basic science and trying to better understand in general terms how viruses interact with the human body. Is this technology or science? It's hard to decide.

You might describe science that has no immediate practical value as basic science or basic research. The exploration of worlds such as Saturn would be called basic science, and it is easy to argue that basic science is not worth the effort and expense because it has no known practical use. Of course, the problem is that you have no way of knowing what knowledge will be of use until you acquire that knowledge. In the middle of the 19th century, Queen Victoria is supposed to have asked physicist Michael Faraday what good his experiments with electricity and magnetism were. He answered, "Madam, what good is a baby?" Of course, Faraday's experiments were the beginning of the electronic age. Many of the practical uses of scientific knowledge that fill the world—computers, vaccines, plastics—began as basic research. Basic scientific research provides the raw materials that technology and engineering use to solve problems.

Basic scientific research has yet one more important use that is so valuable it seems an insult to refer to it as "merely practical." Science is the study of nature, and as we learn more about how nature works, we learn more about what our existence in this universe means for us. The seemingly impractical knowledge gained from space probes to other worlds tells us about Earth and our own role in the scheme of nature. Science tells us where we are and what we are, and that knowledge is beyond value.



Exploring other worlds is valuable. It helps we humans understand ourselves. (NASA/JPL/Space Science Institute)

The Outer Planets

The outermost planets in our solar system are Jupiter, Saturn, Uranus, and Neptune. They are often called the "Jovian planets," meaning they are like Jupiter. In fact, they are each individuals with separate personalities. ■ Figure 18-1 compares the four outer worlds to each other and to Earth. One striking feature is their size. Jupiter is the largest of the Jovian worlds, over 11 times the diameter of Earth. Saturn is slightly smaller, and Uranus and Neptune are quite a bit smaller than Jupiter and Saturn, but still four times the size of Earth. Pluto, not pictured in the figure, is smaller than Earth's moon but was considered a planet from the time of its discovery in 1930 until a decision by the International Astronomical Union (**IAU**) in 2006 that reclassified Pluto as a dwarf planet. You will learn about Pluto's characteristics, and the reasons for the IAU decision, in this chapter.

The other feature you will notice immediately when you look at Figure 18-1 is Saturn's rings. They are bright and beautiful and composed of billions of ice particles. Jupiter, Uranus, and Neptune also have rings, but they are not easily detected from Earth and are not visible in this figure. Nevertheless, as you visit these worlds you will be able to compare four different sets of planetary rings.

■ **Figure 18-1**

The principal worlds of the outer solar system are the four massive but low-density Jovian planets, each much larger than Earth. (NASA/JPL/Space Science Institute/University of Arizona)

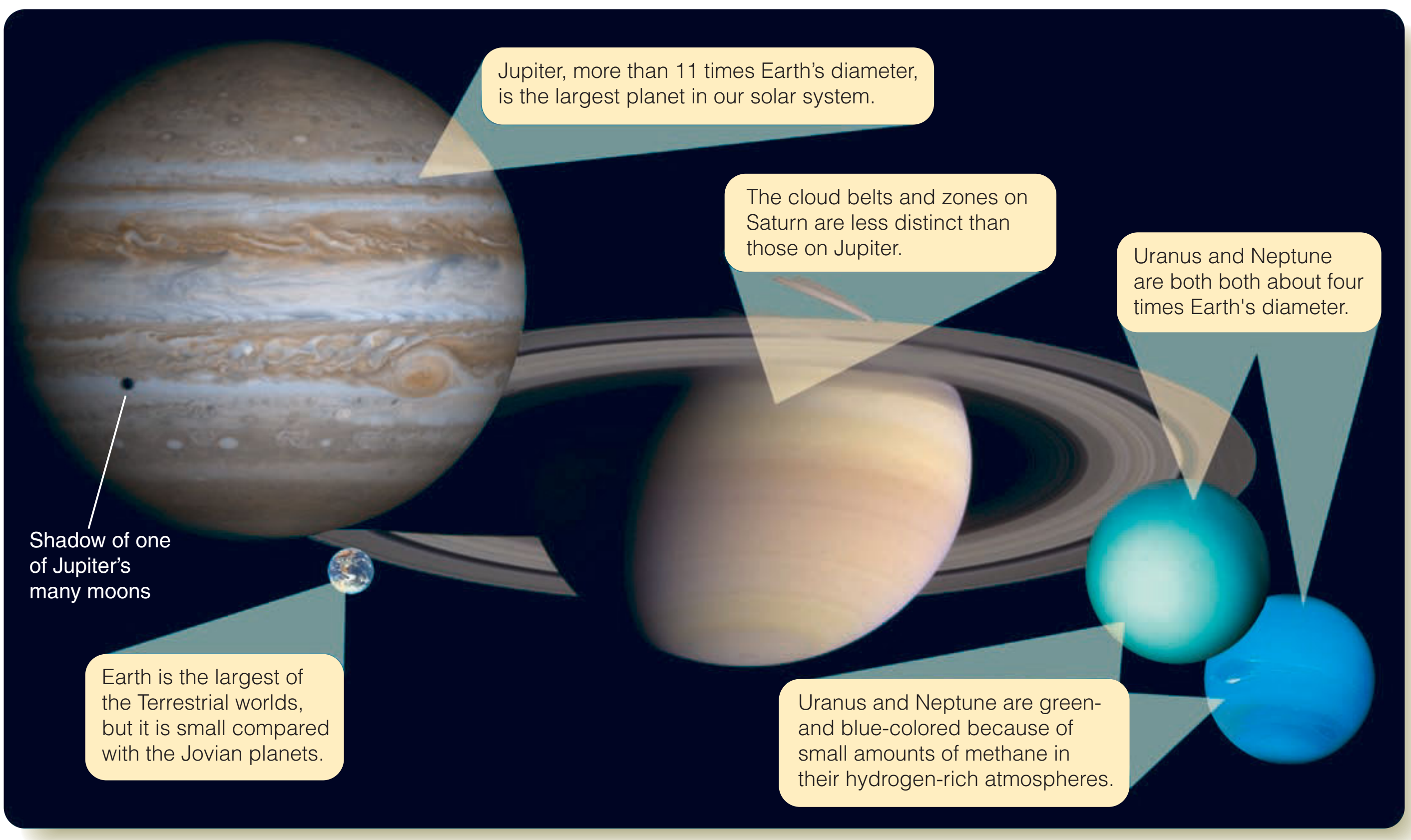
Atmospheres and Interiors

The four Jovian worlds have hydrogen-rich atmospheres filled with clouds. On Jupiter and Saturn, you can see that the clouds form stripes and bands that circle each planet. You will find traces of these same types of features on Uranus and Neptune, but less distinct.

Models based on observations indicate that the atmospheres of the Jovian planets are not very deep; for example, Jupiter's atmosphere makes up only about one percent of its radius. Below their atmospheres Jupiter and Saturn are mostly liquid, so the old-fashioned term for these planets, *gas giants,* should probably be changed to *liquid giants*. Uranus and Neptune are sometimes called *ice giants* because they contain abundant water in solid forms. Only near their centers do the Jovian planets have cores of dense material with the composition of rock and metal. None of the Jovian worlds has a definite solid surface on which you could walk.

Satellite Systems

You can't really land your spaceship on the Jovian worlds, but you might be able to land on one of their moons. All of the Jovian worlds have extensive satellite systems. In many cases, the moons interact gravitationally, mutually adjusting their orbits and also affecting the planetary ring systems. Some of the moons are geologically active now, while others show signs of past activity. Of course, geological activity depends on heat flow from the



interior, so you might ponder what could be heating the insides of these small objects.

◄ SCIENTIFIC ARGUMENT ►

Why do you expect the outer planets to be low-density worlds?

This should be a familiar argument. In an earlier chapter, you discovered that the inner planets could not incorporate ice when they formed because it was too hot at their locations near the sun. By contrast, in the outer solar nebula, water vapor could freeze to form ice particles. The icy particles accumulated rapidly into Jovian protoplanets with density lower than the rocky Terrestrial planets and asteroids. Consequently the Jovian planets grew massive enough to pull in even lower-density hydrogen and helium gas directly from the nebula by gravitational collapse. The ices and gas made the outer planets low-density worlds.

Now you can expand your argument. **Why do you expect the outer planets to have high-density cores?**

◄ ►

18-2 Jupiter

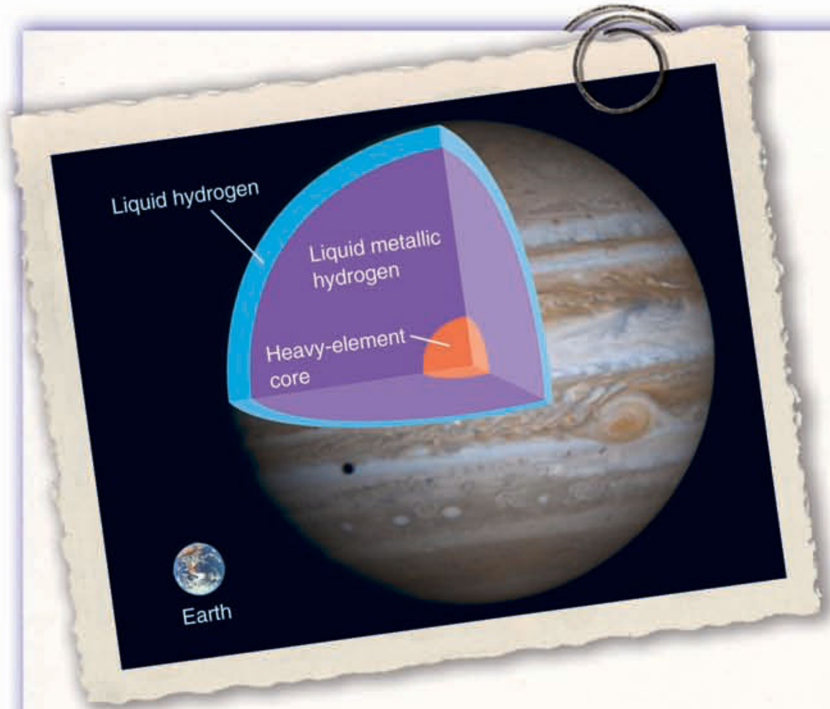
JUPITER IS THE largest and most massive of the Jovian planets, containing 71 percent of all the planetary matter in the entire solar system. Just as you used Earth, the largest of the Terrestrial planets, as the basis for comparison with the others, you can examine Jupiter in detail as a standard in your comparative study of the other Jovian planets.

The Interior

Although Jupiter is very large, it is only 1.3 times denser than water (■ Celestial Profile 7). For comparison, Earth is more than 5.5 times denser than water. As you have already learned, the density of a planet is an important clue about the average composition of the planet's interior. Jupiter's shape also gives information about its interior. Jupiter and the other Jovian planets are all slightly flattened. A world with a large rocky core and mantle would not be flattened much by rotation, but an all-liquid planet would flatten significantly. Thus Jupiter's **oblateness,** the fraction by which its equatorial diameter exceeds its polar diameter, combined with its average density, helps astronomers calculate what its insides are like.

Models show that the interior of Jupiter is mostly liquid hydrogen. However, if you jumped into Jupiter carrying a kayak, expecting an ocean, you would be disappointed. The base of the atmosphere is so hot and the pressure is so high that there is no sudden boundary between liquid and gas. As you fell deeper and deeper through the atmosphere, you would find the gas density increasing around you until you were sinking through a liquid, but you would never splash into a distinct liquid surface.

Under very high pressure, liquid hydrogen becomes **liquid metallic hydrogen**—a material that is a very good conductor of electricity. Model calculations indicate that most of Jupiter's interior is composed of this material. That large mass of conduct-



Jupiter is mostly a liquid hydrogen planet with a small core of heavy elements that is not much bigger than Earth. (NASA/JPL/University of Arizona)

Celestial Profile 7: Jupiter

Motion:

Average distance from the sun	5.20 AU (7.78×10^8 km)
Eccentricity of orbit	0.0484
Inclination of orbit to ecliptic	1.3°
Orbital period	11.9 y
Period of rotation	9.92 h
Inclination of equator to orbit	3.1°

Characteristics:

Equatorial diameter	1.43×10^5 km (11.2 $D_\oplus$)
Mass	1.90×10^{27} kg (318 $M_\oplus$)
Average density	1.34 g/cm^3
Gravity at base of clouds	2.5 Earth gravities
Escape velocity	61 km/s (5.4 $V_\oplus$)
Temperature at cloud tops	140°K (−200°F)
Albedo	0.51
Oblateness	0.064

Personality Point:

Jupiter is named for the Roman king of the gods (the Greek Zeus), and it is the largest planet in our solar system. It can be very bright in the night sky, and its cloud belts and four largest moons can be seen through even a small telescope. Its moons are even visible with a good pair of binoculars mounted on a tripod.

ing liquid, stirred by convection currents and spun by the planet's rapid rotation, drives the dynamo effect and generates a powerful magnetic field. Jupiter's field is over ten times stronger than Earth's.

A planet's magnetic field deflects the solar wind and dominates a volume of space around the planet called the **magnetosphere.** Jupiter's magnetosphere is 100 times larger than Earth's (■ Figure 18-2a). If you could see it in the sky, it would be six times larger than the full moon. Just as in the case of Earth (see Chapter 7), interactions between Jupiter's magnetic field and the solar wind generate powerful electric currents that flow around the planet's magnetic poles. These are visible at ultraviolet wavelengths as rings of auroral lights that are larger in diameter than Earth (Figure 18-2b).

The strong magnetic field around Jupiter traps charged particles from the solar wind in radiation belts a billion times more intense than the Van Allen belts that surround Earth. The spacecraft that have flown through these regions received over 4000 times the radiation that would have been lethal for a human.

At Jupiter's center, a so-called rocky core contains heavier elements, such as iron, nickel, silicon, and so on. With a temperature four times hotter than the surface of the sun and a pressure 50 million times Earth's sea level atmospheric pressure, this material is unlike any rock on Earth. The term *rocky core* refers to the chemical composition, not to the properties of the material.

Careful measurements of the heat flowing out of Jupiter reveal that it emits about twice as much energy as it absorbs from the sun. This energy appears to be heat left over from the formation of the planet. In Chapter 16 you saw that Jupiter should have grown very hot when it formed, and some of this heat remains in its interior, slowly leaking into space.

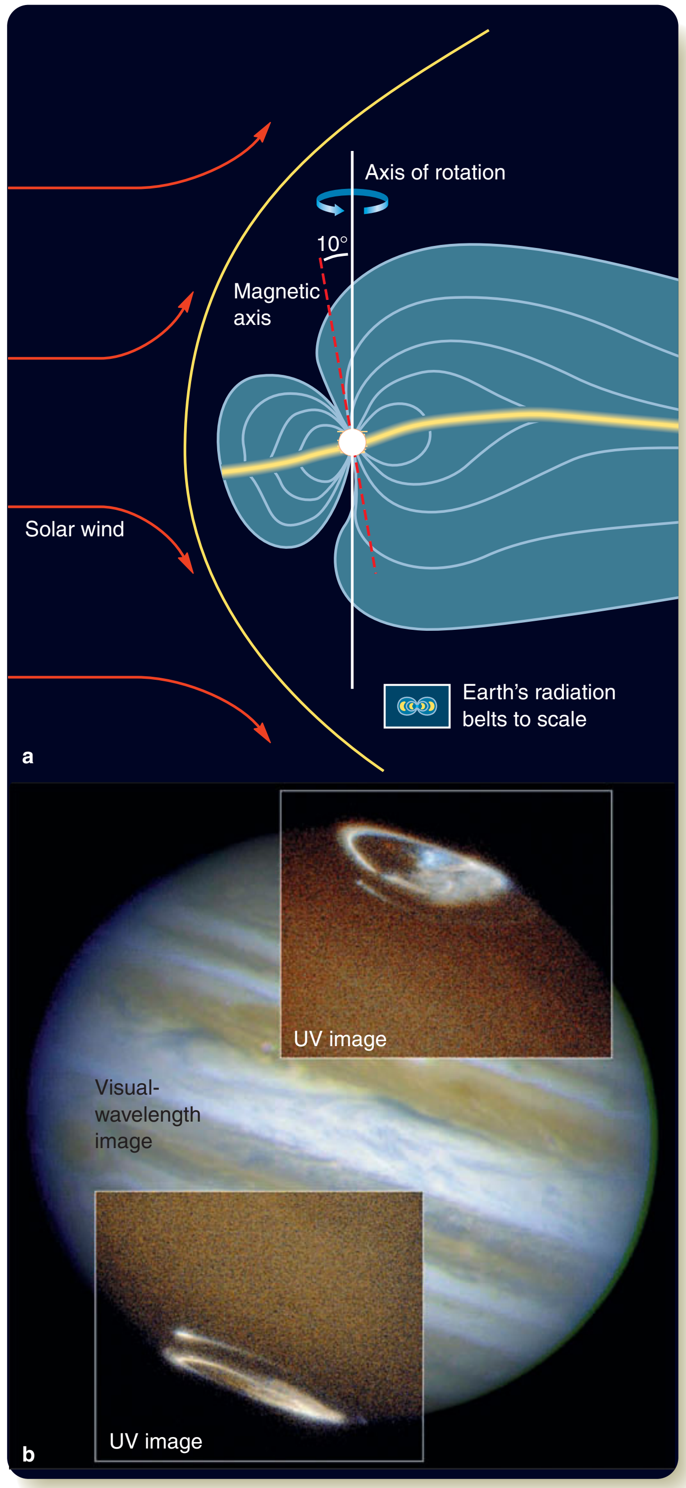
Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercises "Aurora," "Convection and Magnetic Fields," and "Convection and Turbulence."

Jupiter's Complex Atmosphere

It is a **Common Misconception** that Jupiter is a ball of gases. In fact, as you have just learned, Jupiter is almost entirely a liquid planet. Its atmosphere is only a thin outer skin of turbulent gases and clouds. The processes you will find there are repeated in slightly different ways on the other Jovian worlds.

Study ■ **Jupiter's Atmosphere** on pages 384–385 and notice four important ideas plus two new terms:

1. The atmosphere is hydrogen rich, and the clouds are confined to a shallow layer.
2. The positions of the cloud layers are at certain temperatures within the atmosphere where ammonia (NH_3), ammonium hydrosulfide (NH_4SH), and water (H_2O) can condense.
3. The pattern of colored cloud bands circling the planet like stripes on a child's ball is called **belt–zone circulation.** This



■ **Figure 18-2**

(a) Jupiter's large conducting core and rapid rotation create a powerful magnetic field that holds back the solar wind and dominates a region called the magnetosphere. High-energy particles trapped in the magnetic field form giant radiation belts. (b) Auroras on Jupiter are confined to rings around the north magnetic pole and the south magnetic pole, as shown in these ultraviolet images. Earth's auroras follow the same pattern. The small comet-shaped spots are caused by powerful electrical currents flowing from Jupiter's moon Io. (John Clarke, University of Michigan, and NASA) **Animated!**

pattern is related to the high- and low-pressure areas found in Earth's atmosphere.

4 The large circular or oval spots seen in Jupiter's clouds are circulating storms that can remain stable for decades or even centuries.

Photos of Jupiter lead to the **Common Misconception** that the clouds are at the top of the atmosphere. Notice that the atmosphere of transparent hydrogen and helium extends high above the cloud tops. What you see in photos is only the cloud layers.

Jupiter's Moons

Jupiter has four large moons and at least 25 smaller moons. Larger telescopes and modern techniques are rapidly finding more small moons orbiting all the Jovian planets. Each of the Jovian planets probably has many more small, undiscovered moons. Some of the small moons are probably captured asteroids. In contrast, the four largest moons of Jupiter (■ Figure 18-3), called the **Galilean moons** after their discoverer, Galileo, are clearly related to each other and probably formed with Jupiter.

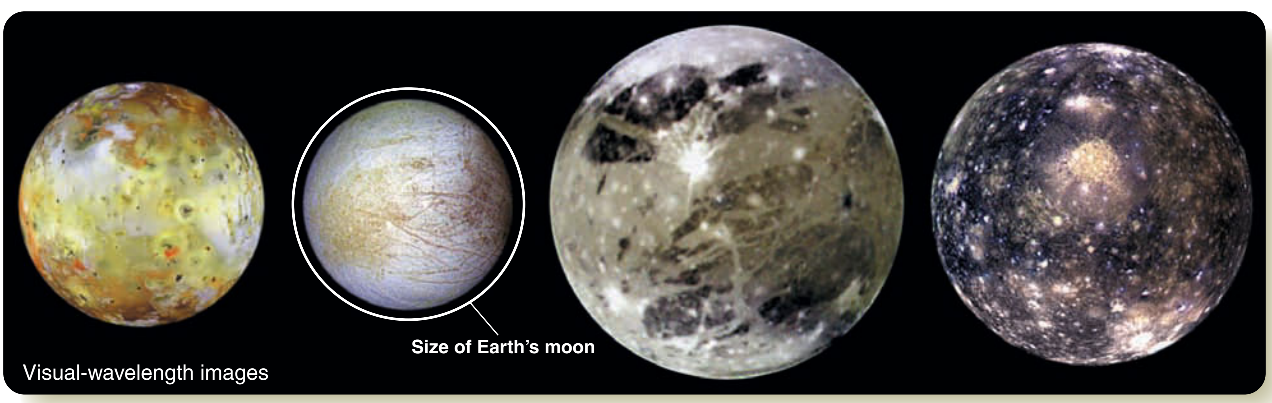
The outermost Galilean moons, Ganymede and Callisto, are about the size of Mercury, one and a half times the size of Earth's moon. In fact, Ganymede is the largest moon in the solar system. Ganymede and Callisto have low densities of only 1.9 and 1.8 g/cm^3, respectively, meaning they must consist roughly of half rock and half ice. Observations of their gravitational fields by the Galileo spacecraft reveal that both moons have rocky or metallic cores and lower-density icy exteriors, so they are differentiated. Both moons interact with Jupiter's magnetic field in a way that shows they probably have mineral-rich layers of liquid water 100 km or more below their icy crusts.

Callisto's surface and most of Ganymede's surface appear old because they are heavily cratered and very dark (■ Figure 18-4a). The continuous blast of meteoroids evaporates surface ice, leaving behind embedded minerals to form a dark skin like the grimy crust on an old snowbank. Thus, icy surfaces get darker with age. More recent impacts dig up cleaner ice and leave bright craters, as you can see on Callisto in Figure 18-3. Ganymede has some younger, brighter **grooved terrain** believed to be systems of faults in the brittle crust. Some sets of grooves overlap other sets of grooves, suggesting extended episodes of geological activity (Figure 18-4b).

The density of the next moon inward, Europa, is 3.0 g/cm^3, high enough to mean that Europa is mostly rock with a thin icy crust. The visible surface is very clean ice, contains very few craters, has long cracks in the icy crust, and includes complicated terrain that resembles blocks of ice in Earth's Arctic Ocean (■ Figure 18-5). The pattern of mountainlike folds on its surface suggests that the icy crust breaks as the moon is flexed by tides (see Chapter 4). Europa's gravitational influence on the Galileo spacecraft reveals that a liquid-water ocean perhaps 200 km deep lies below the 10- to 100-km-thick crust. The lack of craters tells you that Europa is an active world where craters are quickly erased.

Images from spacecraft reveal that Io, the innermost of the four Galilean moons, has over 100 volcanic vents on its surface (■ Figure 18-6). The active volcanoes throw sulfur-rich gas and ash high above the surface. That ash falls back to bury the surface at a rate of a few millimeters a year. This explains why you see no impact craters on Io—they are covered up as fast as they form. Io's density is 3.6 g/cm^3, showing that it is composed of rock and metal. Its gravitational influence on the passing Galileo spacecraft revealed that it is differentiated into a large metallic core, a rocky mantle, and a low-density crust.

The activity you see in the Galilean moons must be driven by energy flowing outward, yet these objects are too small to have remained hot from the time of their formation. Io's volcanism seems to be driven by **tidal heating.** Io follows a slightly elliptical orbit caused by its interactions with the other moons. As Io's distance from Jupiter varies, the planet's gravitational field flexes the moon with varying tidal force, and the resulting friction

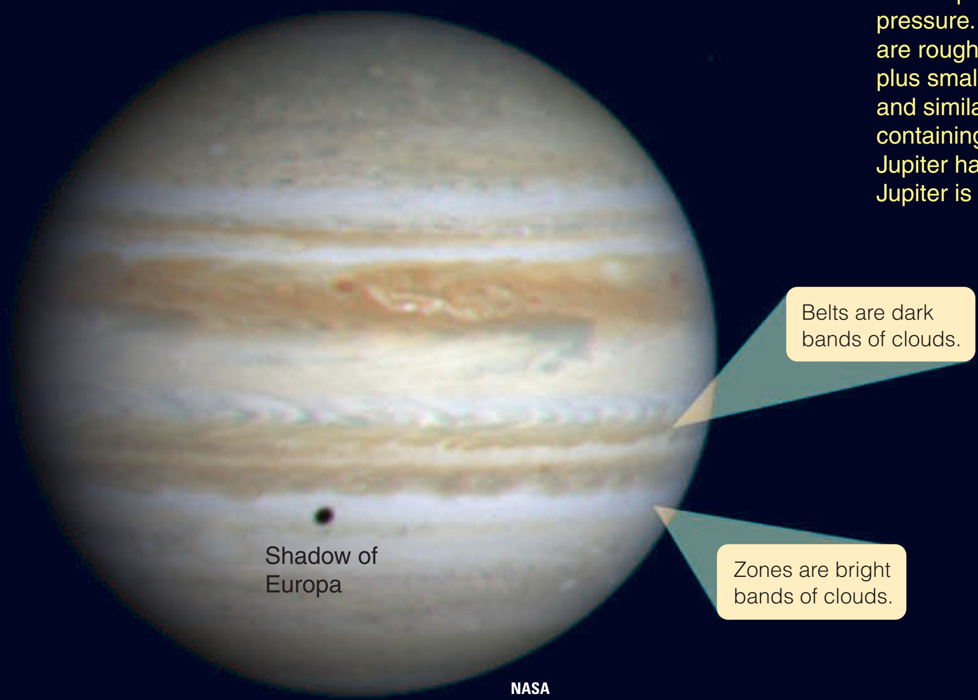


■ **Figure 18-3**

The Galilean moons of Jupiter, from left to right, are Io, Europa, Ganymede, and Callisto. The white circle around Europa shows the size of Earth's moon. (NASA)

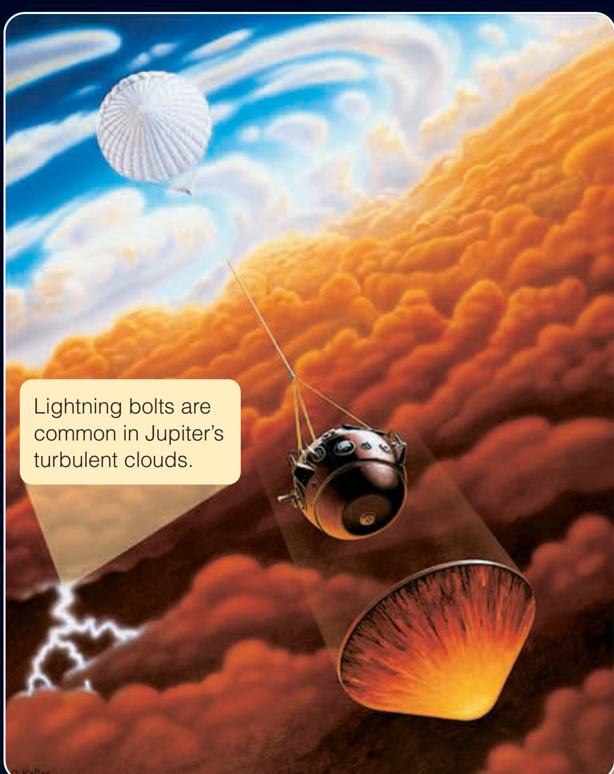
Jupiter's Atmosphere

1 Humans will probably never visit Jupiter's atmosphere. Its cloud layers are deathly cold, and the deeper layers that are warmer have a crushingly high pressure. There is no free oxygen to breathe; the gases are roughly three-quarters hydrogen and a quarter helium, plus small amounts of water vapor, methane, ammonia, and similar molecules. Traces of sulfur and molecules containing sulfur probably make it smell bad. Of course, Jupiter has no surface, so there isn't even a place to stand. Jupiter is a nice planet to look at, but it's not a place to visit.



NASA

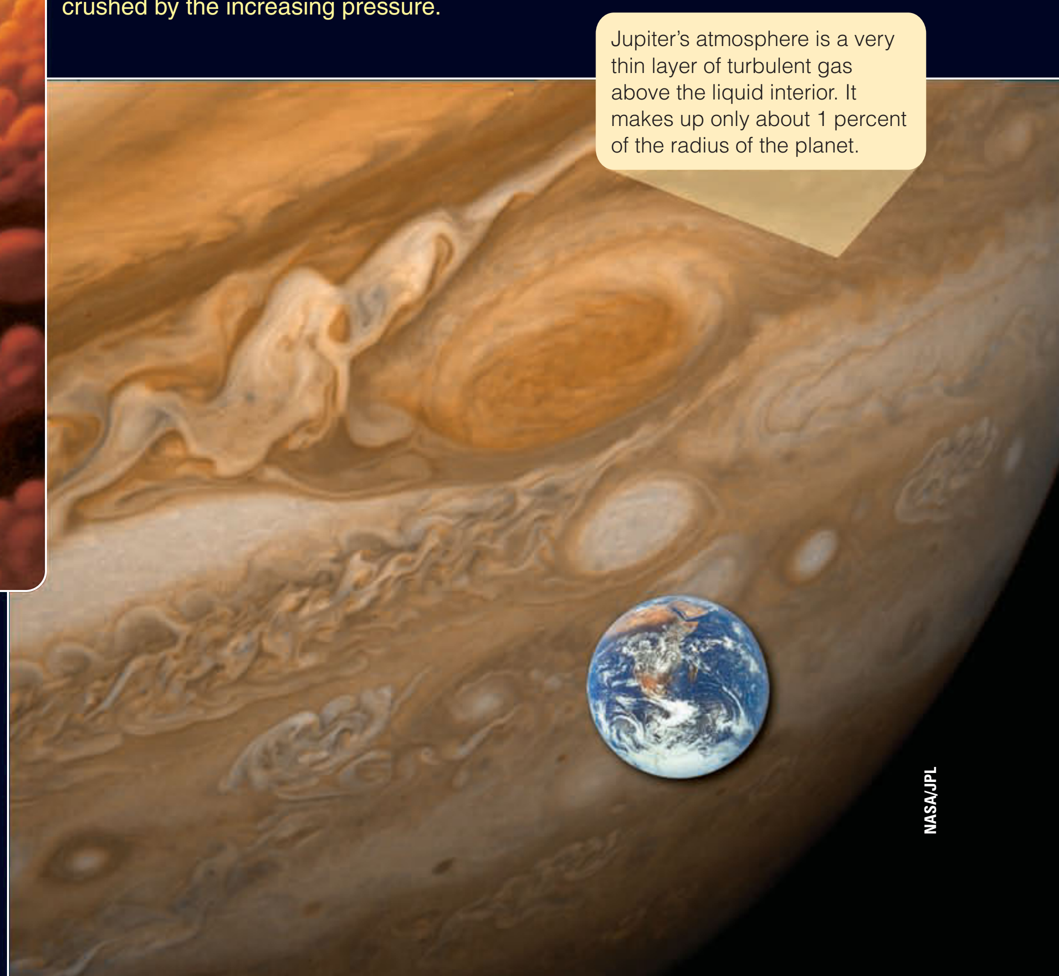
NASA/JPL/Univ. of AZ



Hughes Aircraft Co

1a The only spacecraft to enter Jupiter's atmosphere was the Galileo probe. Released from the Galileo spacecraft, the probe entered Jupiter's atmosphere in December 1995. It parachuted through the upper atmosphere of clear hydrogen, released its heat shield, and then fell through Jupiter's stormy atmosphere until it was crushed by the increasing pressure.

The Great Red Spot at right is a giant circulating storm in one of the southern zones. It has lasted at least 300 years since astronomers first noticed it after the invention of the telescope. Smaller spots are also circulating storms.

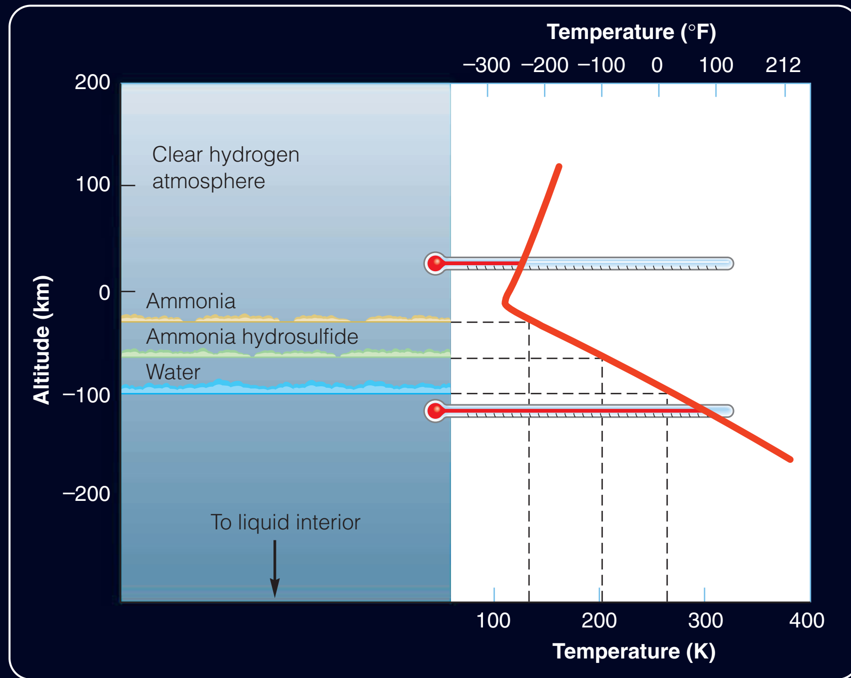


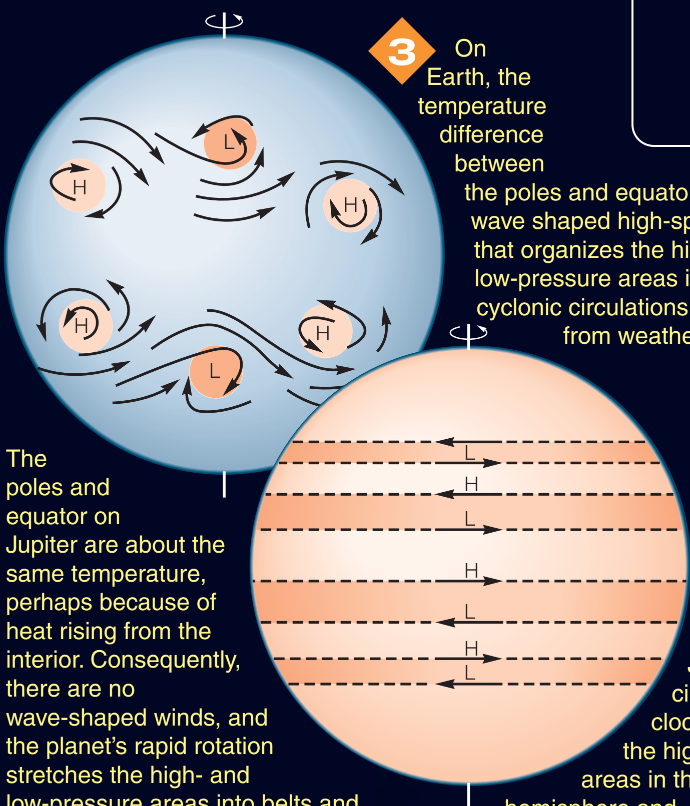
NASA/JPL

2 The visible clouds on Jupiter are composed of ammonia crystals, but models predict that deeper layers of clouds contain ammonia hydrosulfide crystals, and deeper still lies a cloud layer of water droplets. These compounds are normally white, so planetary scientists think the colors arise from small amounts of other molecules formed in reactions powered by lightning or sunlight.

If you could put thermometers in Jupiter's atmosphere at different levels, you would discover that the temperature rises below the uppermost clouds.

Far below the clouds, the temperature and pressure climb so high the gaseous atmosphere merges gradually with the liquid hydrogen interior and there is no surface.



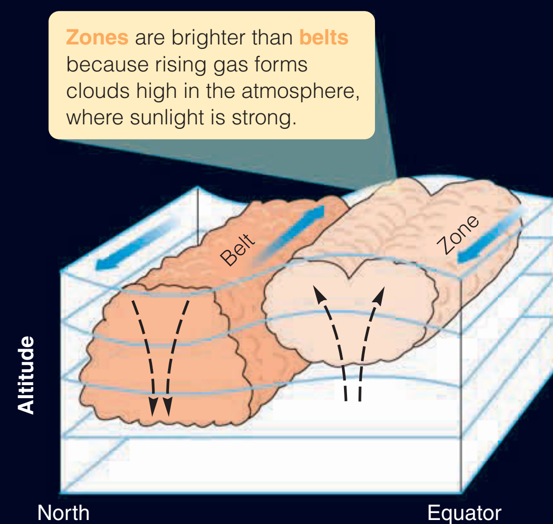


3 On Earth, the temperature difference between the poles and equator drives a wave shaped high-speed wind that organizes the high- and low-pressure areas into cyclonic circulations familiar from weather maps.

The poles and equator on Jupiter are about the same temperature, perhaps because of heat rising from the interior. Consequently, there are no wave-shaped winds, and the planet's rapid rotation stretches the high- and low-pressure areas into belts and zones that circle the planet.

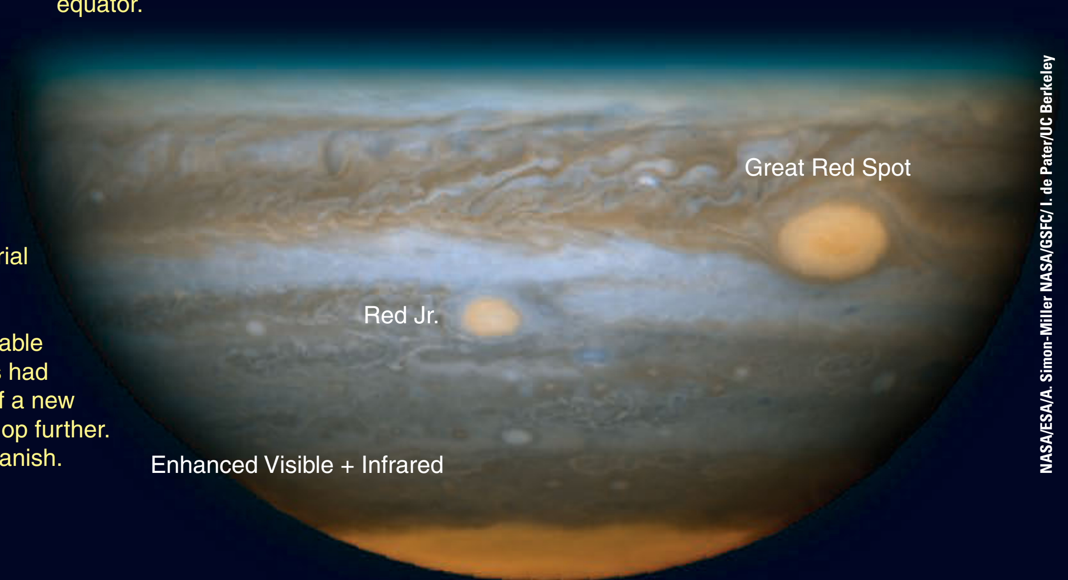
On both Earth and Jupiter, winds circulate clockwise around the high-pressure areas in the northern hemisphere and counter-clockwise south of the equator.

CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see the Active Figure called "Planetary Atmospheres." Notice the temperatures at which the cloud layers form.



4 Three circulating storms visible as white ovals since the 1930s merged in 1998 to form a single white oval. In 2006, the storm intensified and turned red like the Great Red Spot. The reason for the red color is unknown, but it may show that the storm is bringing material up from lower in the atmosphere.

Storms in Jupiter's atmosphere may be stable for decades or centuries, but astronomers had never before witnessed the appearance of a new red spot. It may eventually vanish or develop further. Even the Great Red Spot may someday vanish.



NASA/ESA/A. Simon-Miller NASA/GSFC/ I. de Pater/UC Berkeley

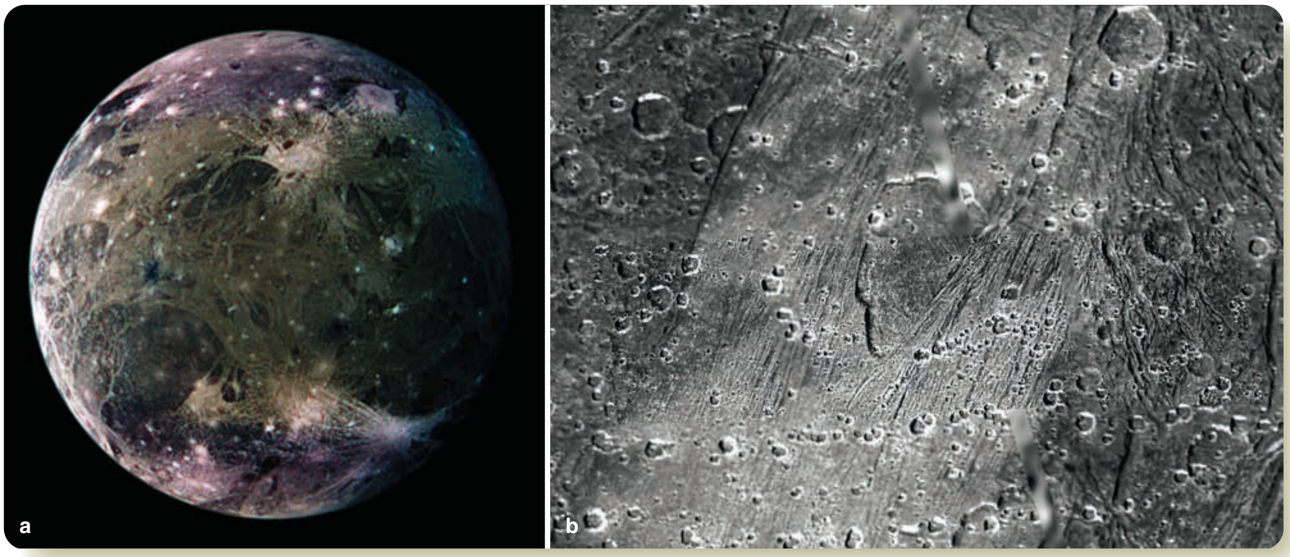


Figure 18-4

(a) This color-enhanced visual-wavelength image of Ganymede shows the frosty poles at top and bottom, the old dark terrain, and the brighter grooved terrain. (b) A band of bright terrain on Ganymede runs from lower left to upper right, and a collapsed area, a possible caldera, lies at the center in this visual-wavelength image. Calderas form where subsurface liquid has drained away, and the bright areas contain other features associated with flooding by water. (NASA)

heats Io's interior. That heat flowing outward causes the volcanism. Europa is not as active as Io, but it also must have a heat source, presumably tidal heating. Ganymede is no longer active, but when it was younger it must have had internal heat to break the crust and produce the grooved terrain.

In fact, the three moons are linked together in **orbital resonances.** Io orbits Jupiter four times while Europa orbits twice and Ganymede orbits once. These resonances keep all three moons' orbits slightly elliptical and drive tidal heating that makes them active now or made them active in the past. Distant Callisto is not caught in this orbital resonance and appears never to have been strongly active.

Jupiter's Ring

Astronomers have known for centuries that Saturn has rings, but Jupiter's ring was not discovered until 1979, when the Voyager 1 spacecraft sent back photos. Less than 1 percent as bright as Saturn's icy rings, Jupiter's ring is very dark and reddish, showing that the ring is rocky rather than icy.

Astronomers can also conclude that the ring particles are mostly microscopic. Photos of the ring show that it is very bright when il-

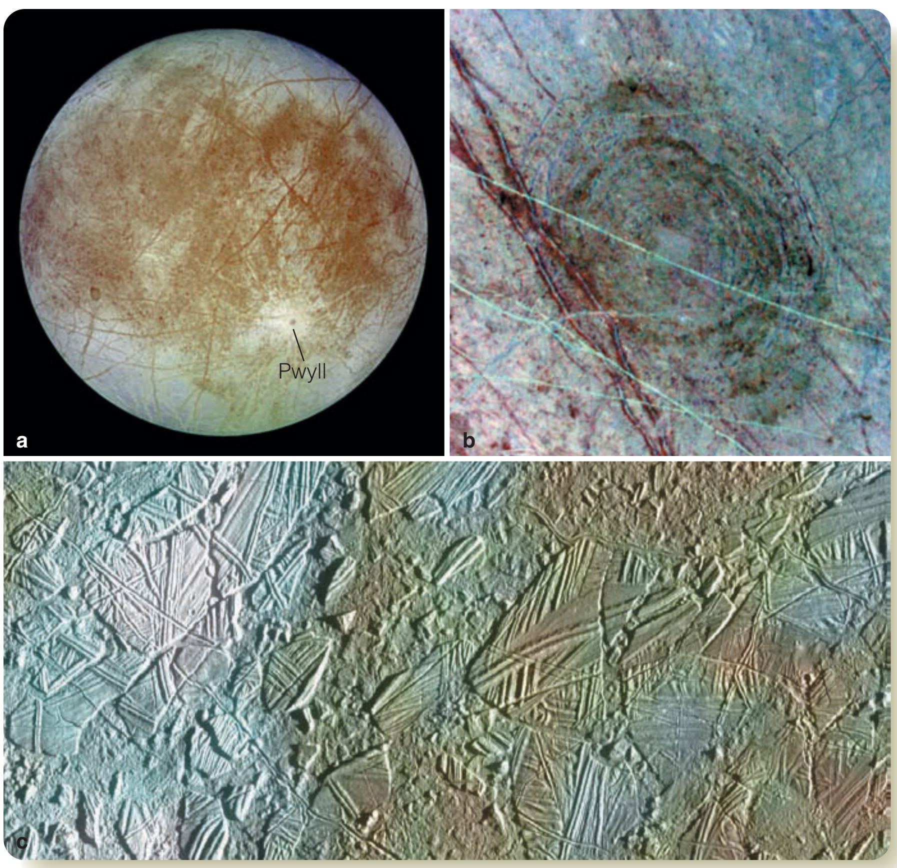


Figure 18-5

(a) The icy surface of Europa is shown here in natural color. Many faults are visible on its surface, but very few craters. The bright crater is Pwyll, a young impact feature. (b) This circular bull's-eye is the remains of a crater 140 km (85 mi) in diameter. Notice the younger cracks and faults that cross the older impact feature. (c) Like icebergs on an arctic ocean, blocks of crust on Europa appear to have floated apart and rotated. The blue icy surface is stained brown by mineral-rich water venting from below the crust. White areas are ejecta from the impact that formed Pwyll crater. (NASA)

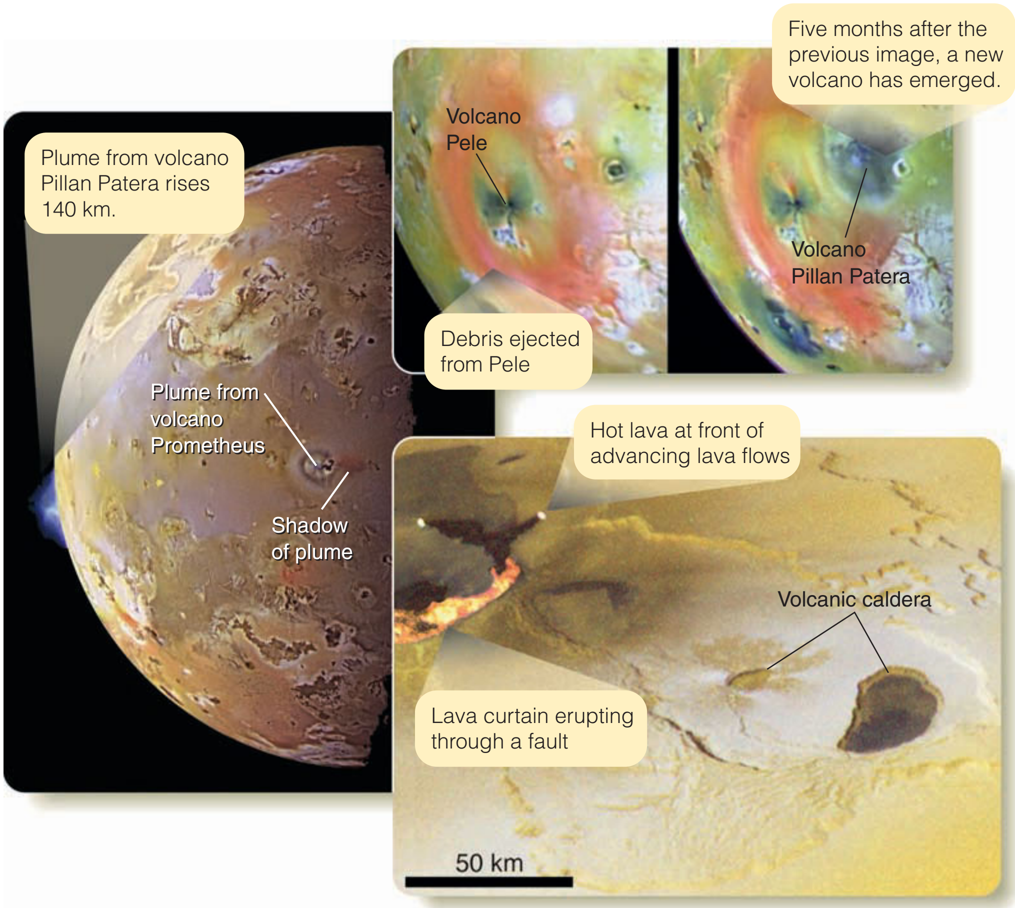


Figure 18-6

These images of volcanic features on Io were produced by combining visual and near-infrared images and digitally enhancing the colors. To human eyes, most of Io would look pale yellow and light orange. (NASA)

Roche limit, tidal forces would overcome the moon's gravity and pull the moon apart. Also, raw material for a moon cannot coalesce inside the Roche limit. The Roche limit is about 2.4 times the planet's radius, depending somewhat on the relative densities of the planet and the orbiting material. Jupiter's rings, as well as those of Saturn, Uranus, and Neptune, lie inside the respective Roche limits for each planet.

Now you can understand Jupiter's dusty ring. If a dust speck gets knocked loose from a larger rock inside the Roche limit, the rock's gravity cannot hold the dust speck. For that same reason, the billions of dust specks in the ring can't pull

luminated from behind (Figure 18-7a)—called **forward scattering.** Large particles do not scatter light forward: A ring filled with basketball-size particles would look dark when illuminated from behind. Forward scattering tells you that Jupiter's ring is made of tiny grains with diameters approximately equal to the wavelengths of light, less than a millionth of a meter, about the size of particles in cigarette smoke.

The ring orbits inside the **Roche limit,** the distance from a planet within which a moon cannot hold itself together by its own gravity. If a moon were to come inside the

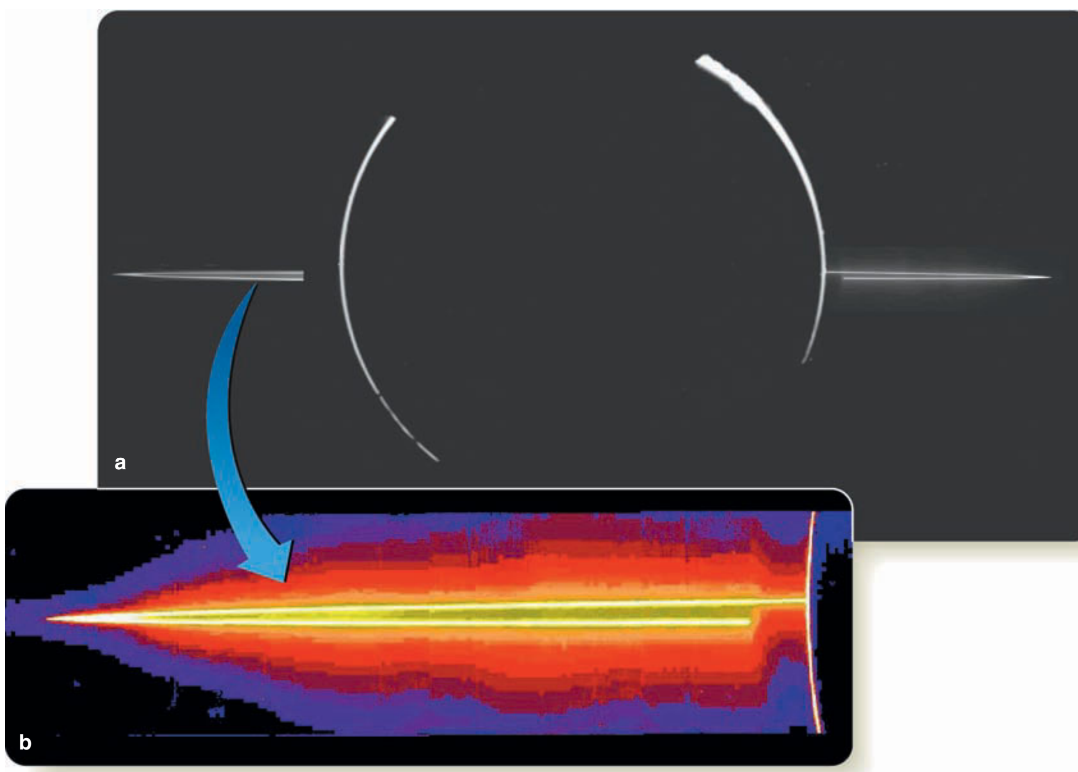


Figure 18-7

(a) The main ring of Jupiter, illuminated from behind, glows brightly in this visual-wavelength image made by the Galileo spacecraft located within Jupiter's shadow. (b) Digital enhancement and false color reveal the halo of ring particles that extends above and below the main ring. The halo is just visible in panel (a). (NASA) **Animated!**

themselves together to make a new moon because of tidal forces inside the Roche limit.

You can be sure that Jupiter's ring particles are not old. The pressure of sunlight and the planet's powerful magnetic field can quickly alter the orbits of the particles. Images show faint ring material extending down toward the cloud tops, evidently dust specks spiraling toward Jupiter. Dust is also lost from the ring as electromagnetic effects force it out of the central plane to form a low-density halo above and below the ring (Figure 18-7b). Another reason the ring particles can't be old is that the intense radiation around Jupiter tends to grind the particles down to nothing in a century or so. The ring you see today therefore can't be material left over in its current situation since the formation of Jupiter. Instead, the ring must be continuously resupplied with new dust. Observations made by the Galileo spacecraft provide evidence that the source of the ring material is meteoroids eroding the small moons Adrastea, Metis, Amalthea, and Thebe that orbit near or within the rings.

The rings around Saturn, Uranus, and Neptune are also known to be short lived, and they also must be resupplied by new material, probably eroded from nearby moons. Aside from supplying the Jovian planets' rings with particles, moons also act to confine the rings, keep them from spreading outward, and alter their shapes. You will explore these processes in detail when you study the rings of the other planets later in this chapter.

A History of Jupiter

Can you put all of the evidence together and tell the story of Jupiter? Creating such a logical argument of evidence and hypotheses is the ultimate goal of planetary astronomy.

Jupiter formed far enough from the sun to incorporate large numbers of icy planetesimals, and it must have grown rapidly. Once it was about 15 times more massive than Earth, it could grow by gravitational collapse (see Chapter 16), capturing gas directly from the solar nebula. Thus, it grew rich in hydrogen and helium from the solar nebula. Its present composition resembles the composition of the sun and the solar nebula. Jupiter's gravity is strong enough to hold onto all its gases, even hydrogen (see Figure 17-15).

The large family of moons may be mostly captured asteroids, and Jupiter may still encounter a wandering asteroid or comet now and then. Some of these are deflected, and some, like comet Shoemaker-Levy 9 that struck Jupiter in 1994, actually fall into the planet (see Chapter 19). Dust blasted off the inner moons by meteoroid impacts settles into the equatorial plane to form Jupiter's ring.

The four Galilean moons are large and seem to have formed like a mini-solar system in a disk of gas and dust around the forming planet. The innermost Galilean moon, Io, is densest, and the densities of the others decrease as you move away from Jupiter, similar to the way the densities of the planets decrease with distance from the sun. Perhaps the inner moons incorporated less ice because they formed closer to the heat of the growing planet. You can recognize that tidal heating also has been important, and the intense heating of the inner moons could have driven off much of their ices. Thus, a combination of two processes may be responsible for the compositions of the Galilean moons.

◄ SCIENTIFIC ARGUMENT ►

Why is Jupiter so large?

You can analyze this question by constructing a logical argument that relates the formation of Jupiter to the solar nebula theory. Jupiter grew so rapidly from icy planetesimals in the outer solar nebula that it eventually became large enough to be able to continue growing by gravitational collapse. By the time the solar nebula cleared away and ended planet building, Jupiter had captured large amounts of hydrogen and helium. The Terrestrial planets are made from a small fraction of the elements present in the solar nebula, but the Jovian planets incorporated abundant hydrogen and helium and were able to become very massive.

Now create a new argument. **How is the activity on Io and Europa powered?**

◄ ►

18-3 Saturn

Saturn is most famous for its beautiful rings, easily visible through the telescopes of modern amateur astronomers. Large Earth-based telescopes have explored the planet for many decades. The first close-up views came when two Voyager probes flew past Saturn in 1980 and 1981. The Cassini spacecraft went into orbit around Saturn in 2004 and began an extended exploration of the planet, its rings, and its moons (■ How Do We Know? 18-2).

Saturn the Planet

As you can see in Figure 18-1, Saturn shows only faint belt–zone circulation, but Voyager, Cassini, and Hubble Space Telescope images confirm that belts and zones are present and that the associated winds are up to three times faster than the winds on Jupiter. Belts and zones on Saturn are less visible than on Jupiter because they occur deeper in Saturn's colder atmosphere, below a layer of methane haze (■ Figure 18-8).

Saturn is less dense than water (it would float!), suggesting that it is, like Jupiter, rich in hydrogen and helium. Saturn is the most oblate of the planets, and this evidence tells you that its interior is mostly liquid with only a small core of heavy elements (■ Celestial Profile 8). Because its internal pressure is lower, Saturn has less liquid metallic hydrogen than Jupiter. Perhaps that is why Saturn's magnetic field is 20 times weaker than Jupiter's. Like Jupiter, Saturn radiates more energy than it receives from the sun, and models predict that it, too, has a very hot interior.

How Do We Know? 18-2

Funding for Basic Research

Who pays for science? Science is an expensive enterprise, and that raises the question of payment. Some science has direct technological applications, but some basic science is of no immediate practical value. Who pays the bill? In "How Do We Know?" 18-1, you saw that technology is of immediate practical use. For that reason, business enterprises fund much of this type of research. Auto manufacturers need inexpensive, durable, quick-drying paint for their cars, and they find it worth the cost to hire chemists to study the way paint dries. Many industries have large research budgets, and some industries, such as pharmaceutical manufacturers, depend exclusively on scientific research to discover and develop new products.

If a field of research has immediate potential to help society, it is likely that government will supply funds. Much of the public health research in the United States is funded by government institutions such as the National Institutes of Health and the National Science Foundation. The practical benefit of finding new ways to prevent disease, for example, is considered well worth the tax dollars.

Basic science, however, has no immediate practical use. That doesn't mean it is useless, but it does mean that the practical-minded stockholders of a company will not approve major investments in such research. Digging up dinosaur bones, for instance, is very poorly funded because no industry can make a profit from the discovery of a new dinosaur. Astronomy is another field of science that has few direct applications, and thus not much astronomical research is funded by industry.

The value of basic research is twofold. Discoveries that have no known practical use today may be critically important years from now. Thus, society needs to continue basic research to protect its own future. But basic research, such as studying Saturn's rings or digging up dinosaur bones, is also of cultural value because it tells us what we are. Each of us benefits in intangible ways from such research, and thus society needs to fund basic research for the same reason it funds art galleries and national parks—to make our lives richer and more fulfilling.

Because there is no immediate financial return from this kind of research, it falls to government institutions and private foundations to pay the bill. The Keck Foundation has built two giant telescopes with no expectation of financial return, and the National Science Foundation has funded thousands of astronomy research projects for the benefit of society. Debates rage as to how much money is enough and how much is too much, but, ultimately, funding basic scientific research is a public responsibility that society must balance against other needs. There isn't anyone else to pick up the tab.

Sending the Cassini spacecraft to Saturn costs each American $0.56 per year over the life of the project. (NASA/STScI)

Saturn's Moons

Saturn has nearly 50 known moons, many of which are small and all of which contain mixtures of ice and rock. Many are probably captured objects.

The largest of Saturn's moons is Titan, a bit larger than the planet Mercury. Its density suggests that it must contain a rocky core under a thick mantle of ices. Titan is so cold that its gas molecules do not travel fast enough to escape. It has an atmosphere composed mostly of nitrogen with traces of argon and methane. Sunlight converts some of the methane into complex carbon-rich molecules that collect into small particles

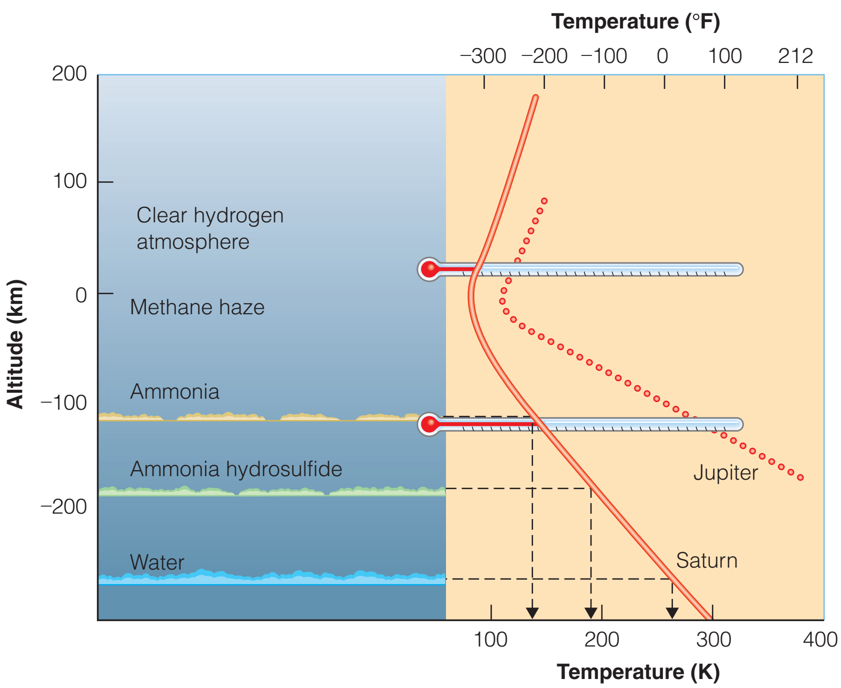


Figure 18-8

Because Saturn is farther from the sun, its atmosphere is colder than Jupiter's (dotted line). The cloud layers on Saturn form at the same temperature as do the cloud layers on Jupiter, but that puts them deeper in Saturn's hazy atmosphere, where they are not as easy to see from the outside as Jupiter's clouds. (See page 385.) **Animated!**

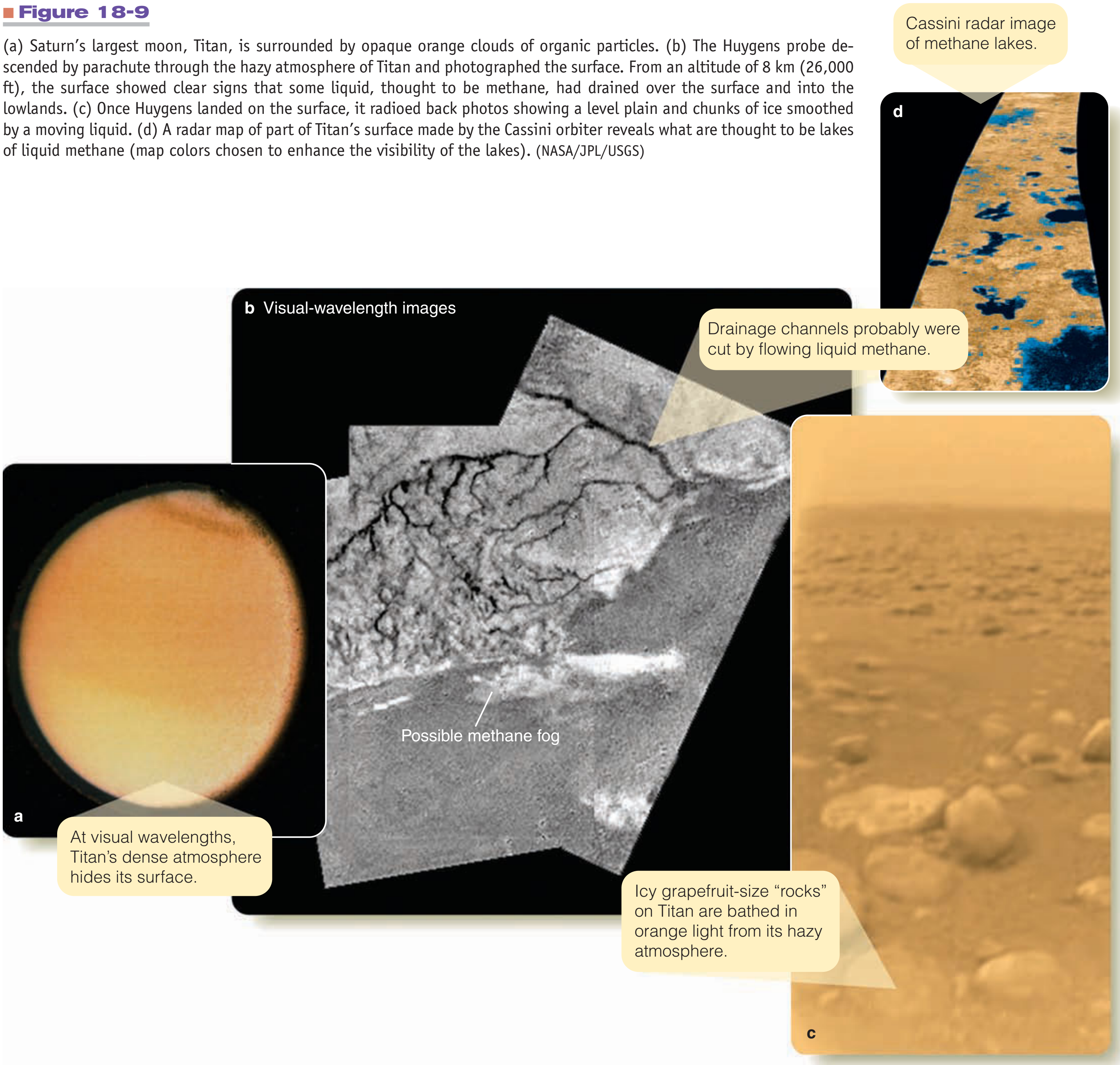
filling the atmosphere with orange smog (■ Figure 18-9a). These particles settle slowly downward to coat the surface with what has been described as dark organic "goo," meaning it is composed of carbon-rich molecules and is probably semiliquid.

Titan's surface is mainly composed of ices of water and methane at −180°C (−290°F). The Cassini spacecraft dropped the Huygens probe into the atmosphere of Titan. The probe photographed dark drainage channels suggesting that liquid methane falls as rain, washes the dark goo off the higher terrain, and drains into the lowlands (Figure 18-9b). Such methane downpours may be rare, however. No direct evidence of liquid methane was detected as the probe descended, but later radar images made by the Cassini orbiter have detected what appear to be lakes presumably containing liquid methane (Figure 18-9d). Infrared images suggest the presence of methane volcanoes that replenish the methane in the atmosphere, so Titan must have some internal heat source to power the activity.

Most of the remaining moons of Saturn are small and icy, have no atmospheres, are heavily cratered and have dark, ancient surfaces. The moon Enceladus, however, shows signs of recent geological activity. Some parts of its surface contain 1000 times fewer craters than other regions, and infrared observations show that its south polar region is unusually warm and venting water and ice geysers (■ Figure 18-10). Evidently, a reservoir of liquid

■ Figure 18-9

(a) Saturn's largest moon, Titan, is surrounded by opaque orange clouds of organic particles. (b) The Huygens probe descended by parachute through the hazy atmosphere of Titan and photographed the surface. From an altitude of 8 km (26,000 ft), the surface showed clear signs that some liquid, thought to be methane, had drained over the surface and into the lowlands. (c) Once Huygens landed on the surface, it radioed back photos showing a level plain and chunks of ice smoothed by a moving liquid. (d) A radar map of part of Titan's surface made by the Cassini orbiter reveals what are thought to be lakes of liquid methane (map colors chosen to enhance the visibility of the lakes). (NASA/JPL/USGS)



water lies just below the surface. At some point in its history, this moon may have been caught in a resonance with another moon and had its interior warmed by tidal heating. Ice particles flying into space from geysers on Enceladus appear to maintain Saturn's large, low-density E ring that extends far beyond the visible rings.

Like nearly all moons in the solar system, Saturn's moons are tidally locked to their planet, rotating to keep the same side facing the planet. The leading side of these moons, the side facing forward in the orbit, is sometimes modified by debris. Iapetus, for example, has a cratered trailing side about as bright as dirty snow, but its leading side is as dark as fresh asphalt. One hypothesis is that the dark material is carbon-rich dust from meteoroid impacts on Phoebe, the next moon outward from Saturn. Iapetus also has a strange equatorial ridge, which may have been produced by rapid rotation when it was young.

Saturn's Rings

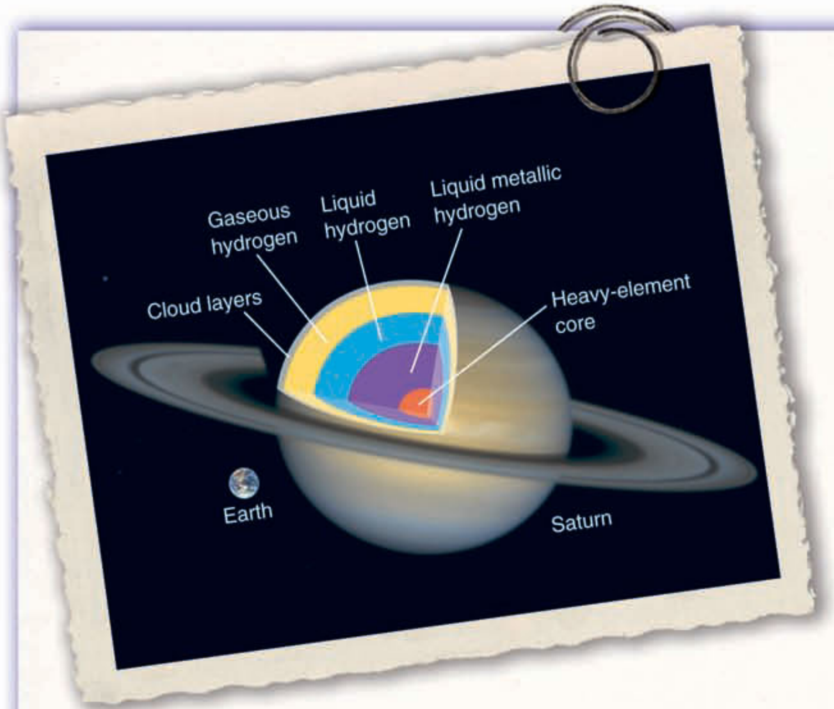
The rings of Saturn are perhaps the most beautiful sight in our solar system. Study ■ **The Ice Rings of Saturn** on pages 392–393 and notice three things:

1. The rings are made up of billions of ice particles, each in its own orbit around the planet. The ring particles you observe now can't be as old as Saturn. The rings must be replenished by impacts on Saturn's icy moons or other processes. The same is true of the rings around the other Jovian planets.
2. The gravitational effects of small moons can confine some rings in narrow strands or keep the edges of rings sharp. Moons can also produce waves in the rings that are visible as tightly wound ringlets.
3. The ring particles are confined in a thin plane spread among small moons and confined by gravitational interactions with larger moons. The rings of Saturn, and the rings of the other Jovian worlds, are created by and controlled by the planet's moons. Without the moons, there would be no rings.

Observations made by the Cassini spacecraft show that the ring particles have compositions resembling that of Saturn's distant icy moon Phoebe. A large impact on Phoebe may be part of the complex history of Saturn's rings.

The History of Saturn

Saturn formed in the outer solar nebula, where ice particles were stable and may have contained more trapped gases. The protoplanet grew rapidly and became massive enough to attract hydrogen and helium by gravitational collapse. The heavier elements sank to the middle to form a small core, and the hydrogen formed a liquid mantle containing liquid metallic hydrogen. The outward flow of heat from the interior drives convection inside the planet that helps produce its magnetic field. Because Saturn is smaller than Jupiter, the internal pressure is less, the planet



Sending the Cassini spacecraft to Saturn costs each American $0.56 per year over the life of the project. (NASA)

Celestial Profile 8: Saturn

Motion:

Average distance from the sun	9.54 AU (1.43×10^9 km)
Eccentricity of orbit	0.0560
Inclination of orbit to ecliptic	2.5°
Orbital period	29.5 y
Period of rotation	10.66 h
Inclination of equator to orbit	26.4°

Characteristics:

Equatorial diameter	1.30×10^5 (9.4 $D_\oplus$)
Mass	5.69×10^{26} kg (95.1 $M_\oplus$)
Average density	0.69 g/cm^3
Gravity at top of clouds	1.2 Earth gravities
Escape velocity	35.6 km/s (3.2 $V_\oplus$)
Temperature at cloud tops	95°K (−290°F)
Albedo	0.61
Oblateness	0.102

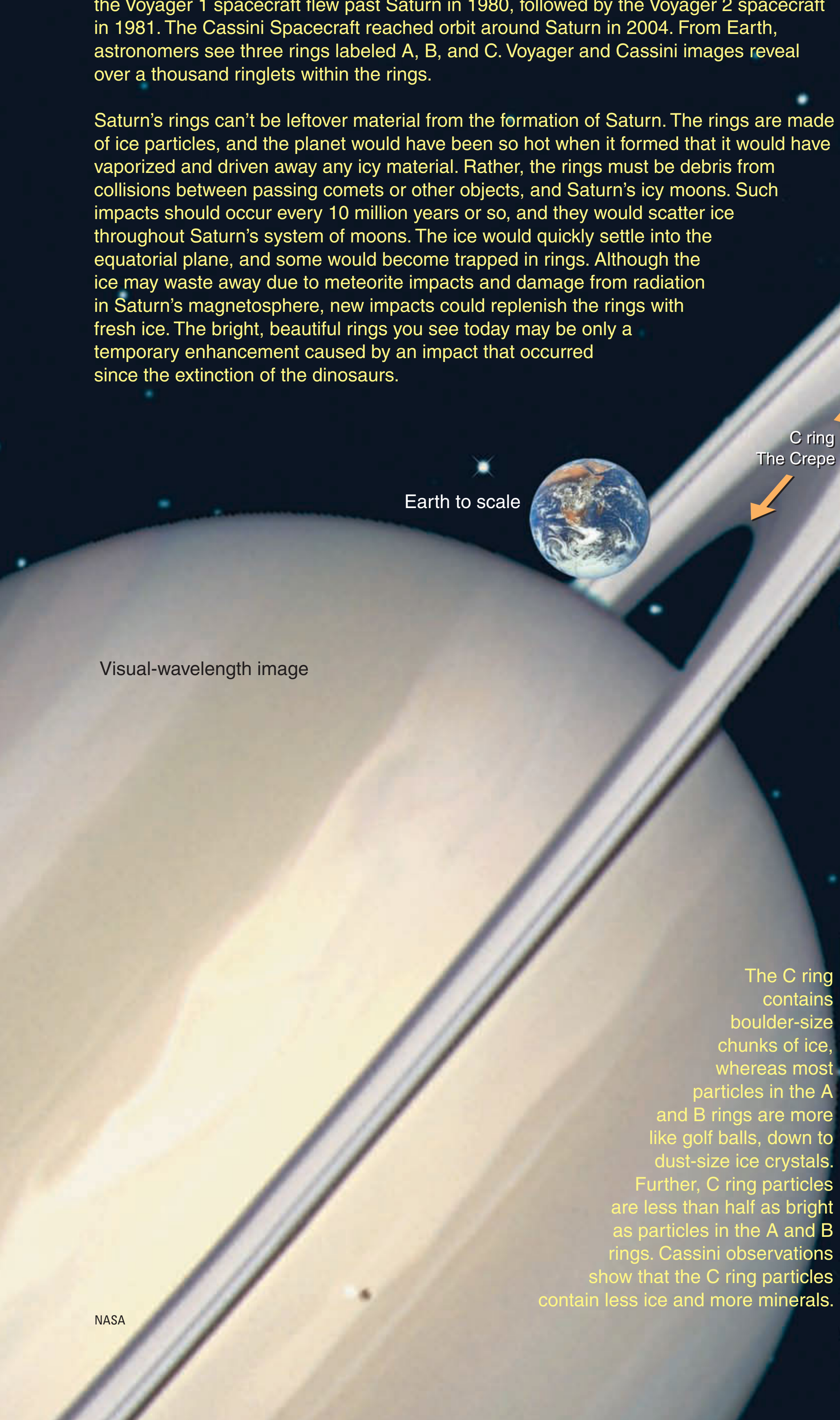
Personality Point:

The Greek god Cronus was forced to flee when his son Zeus took power. Cronus went to Italy where the Romans called him Saturn, protector of the sowing of seed. He was celebrated in a weeklong Saturnalia festival at the time of the winter solstice in late December. Early Christians took over the holiday to celebrate Christmas.

The Ice Rings of Saturn

1 The brilliant rings of Saturn are made up of billions of ice particles ranging from microscopic specks to chunks bigger than a house. Each particle orbits Saturn in its own circular orbit. Much of what astronomers know about the rings was learned when the Voyager 1 spacecraft flew past Saturn in 1980, followed by the Voyager 2 spacecraft in 1981. The Cassini Spacecraft reached orbit around Saturn in 2004. From Earth, astronomers see three rings labeled A, B, and C. Voyager and Cassini images reveal over a thousand ringlets within the rings.

Saturn's rings can't be leftover material from the formation of Saturn. The rings are made of ice particles, and the planet would have been so hot when it formed that it would have vaporized and driven away any icy material. Rather, the rings must be debris from collisions between passing comets or other objects, and Saturn's icy moons. Such impacts should occur every 10 million years or so, and they would scatter ice throughout Saturn's system of moons. The ice would quickly settle into the equatorial plane, and some would become trapped in rings. Although the ice may waste away due to meteorite impacts and damage from radiation in Saturn's magnetosphere, new impacts could replenish the rings with fresh ice. The bright, beautiful rings you see today may be only a temporary enhancement caused by an impact that occurred since the extinction of the dinosaurs.



As in the case of Jupiter's ring, Saturn's rings lie inside the planet's Roché limit where the ring particles cannot pull themselves together to form a moon.

Because it is so dark, the C ring has been called the Crepe ring, referring to the black, semitransparent cloth associated with funerals.

Visual-wavelength image

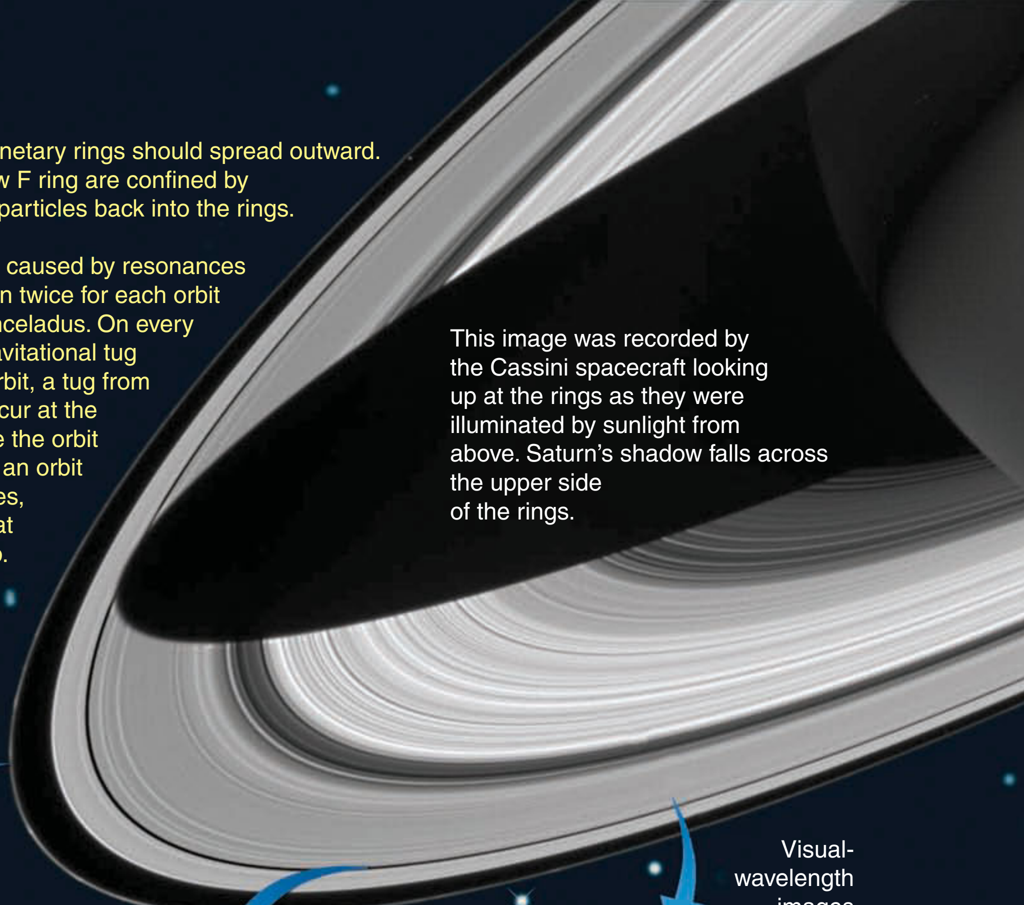
1a An astronaut could swim through the rings. Although the particles orbit Saturn at high velocity, all particles at the same distance from the planet orbit at about the same speed, so they collide gently at low velocities. If you could visit the rings, you could push your way from one icy particle to the next. This artwork is based on a model of particle sizes in the A ring.

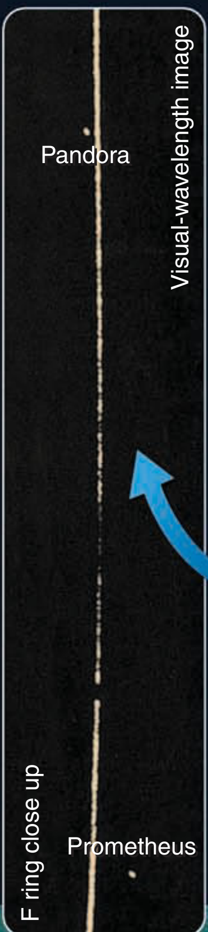
The C ring contains boulder-size chunks of ice, whereas most particles in the A and B rings are more like golf balls, down to dust-size ice crystals. Further, C ring particles are less than half as bright as particles in the A and B rings. Cassini observations show that the C ring particles contain less ice and more minerals.

NASA

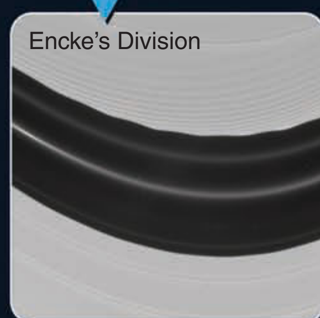
2 Because of collisions among ring particles, planetary rings should spread outward. The sharp outer edge of the A ring and the narrow F ring are confined by **shepherd satellites** that gravitationally usher straying particles back into the rings.

Some gaps in the rings, such as Cassini's Division, are caused by resonances with moons. A particle in Cassini's Division orbits Saturn twice for each orbit of the moon Mimas and three times for each orbit of Enceladus. On every other orbit, the particle feels a gravitational tug from Mimas and, on every third orbit, a tug from Enceladus. These tugs always occur at the same places in the orbit and force the orbit to become slightly elliptical. Such an orbit crosses the orbits of other particles, which results in collisions, and that removes the particle from the gap.



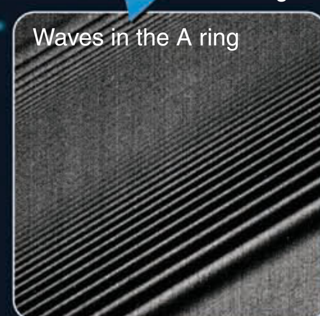


The F ring is clumpy and braided because of two shepherd satellites.

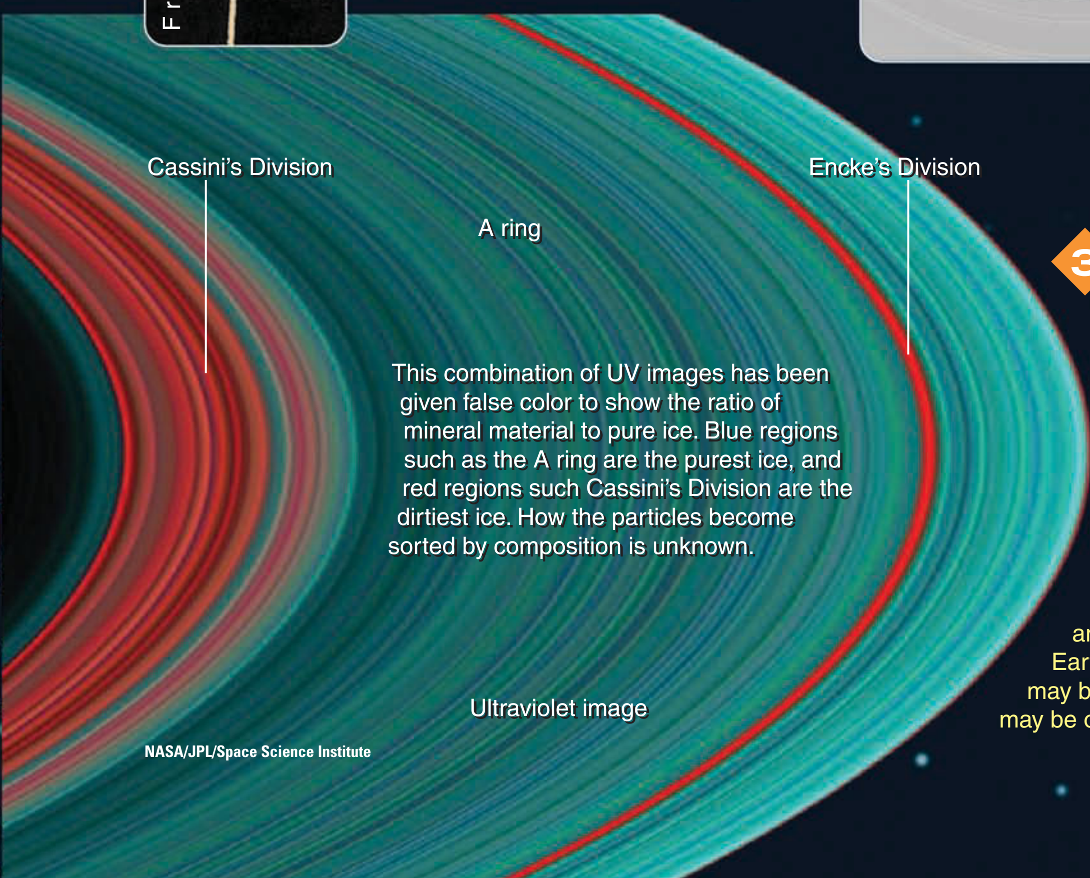


Encke's Division is not empty. Note the ripples at the inner edge. A small moon orbits inside the division.

Visual-wavelength images



Saturn does not have enough moons to produce all of its ringlets by resonances. Many are produced by tightly wound waves, much like the spiral arms found in disk galaxies.



This combination of UV images has been given false color to show the ratio of mineral material to pure ice. Blue regions such as the A ring are the purest ice, and red regions such Cassini's Division are the dirtiest ice. How the particles become sorted by composition is unknown.

NASA/JPL/Space Science Institute

3 How do moons happen to be at just the right places to confine the rings? That puts the cosmic cart before the horse. The ring particles get caught in the most stable orbits among Saturn's innermost moons. The rings push against the inner moons, but those moons are locked in place by resonances with larger, outer moons. Without the moons, the rings would spread and dissipate.

Saturn's rings are a very thin layer of particles and nearly vanish when the rings turn edge-on to Earth. Although ripples in the rings caused by waves may be hundreds of meters high, the sheet of particles may be only about ten meters thick.

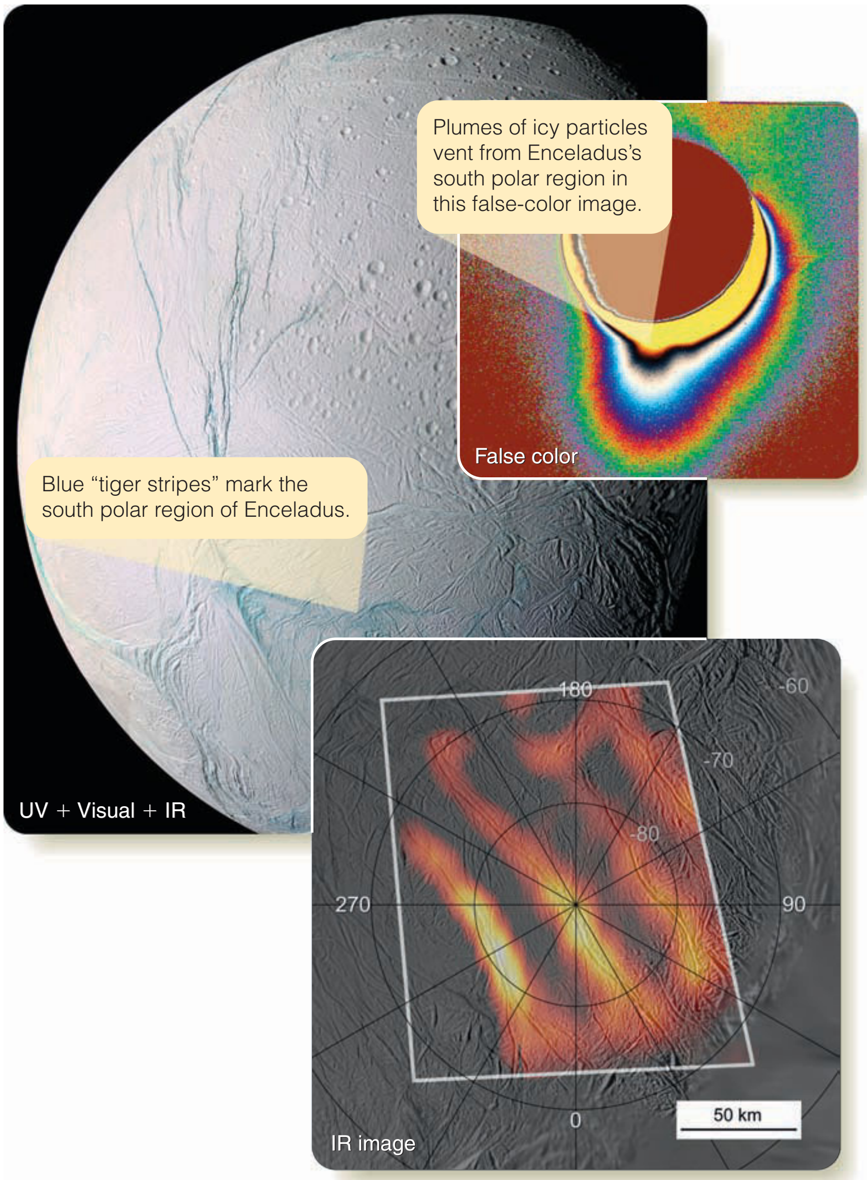


Figure 18-10

Saturn's moon Enceladus is venting water, ice, and organic molecules from geysers near its south pole. A thermal infrared image reveals internal heat leaking to space from the "tiger stripe" cracks where the geysers are located. (NASA/JPL/Space Science Institute)

contains less liquid metallic hyrogen, and its magnetic field is weaker.

The rings can't be *primordial.* That is, they can't be material left over from the formation of the planet. Such ices would have been vaporized and driven away by the heat of the protoplanet. Rather, you can suppose that the rings are debris from the occasional impacts of meteoroids, asteroids, and comets on Saturn's icy moons.

Some of Saturn's moons are probably captured asteroids that wandered too close, but some of the moons probably formed with Saturn. Many have ancient surfaces. The giant moon Titan may have formed with Saturn, or it may be a very large icy planetesimal captured into orbit around Saturn. You will see more evidence for this capture hypothesis for moons when you explore farther from the sun.

◄ SCIENTIFIC ARGUMENT ►

Why do the belts and zones on Saturn look so indistinct?

This argument compares Saturn with Jupiter. In a Jovian planet, the light-colored zones form where rising gas cools and condenses to form icy crystals of ammonia, which are visible as bright clouds. Saturn is twice as far from the sun as Jupiter, so sunlight is weaker and the atmosphere is colder. The gas in Saturn's atmosphere doesn't have to rise as high to reach temperatures cold enough to form clouds. Because the clouds form deeper in the hazy atmosphere, they are not as brightly illuminated by sunlight and look dimmer. Also, a layer of methane haze above the clouds makes the belts and zones look even less distinct.

You have used some simple physics to construct a logical argument that explains the hazy cloud features on Saturn. Now build a new argument. **Why do Saturn's rings have gaps and ringlets?**

◄ ►

18-4 Uranus

Now that you are familiar with the liquid giants in our solar system, you will be able to appreciate how weird the ice giants, Uranus and Neptune, are. Uranus, especially, seems to have forgotten how to behave like a planet.

Uranus the Planet

Uranus is only one-third the diameter of Jupiter and only one-twentieth as massive. Four times farther from the sun than Jupiter, its atmosphere is almost 100°C colder than Jupiter's (■ Celestial Profile 9).

Uranus never grew massive enough to capture large amounts of gas from the nebula as Jupiter and Saturn did, so it has much less hydrogen and helium. Its internal pressure is enough less than Jupiter's that it should not contain any liquid metallic hydrogen. Models of Uranus based in part on its density and oblateness suggest that it has a small core of heavy elements and a deep mantle of partly solid water. Although that material is referred to as ice, it would not be anything like ice on Earth at the temperatures and pressures inside Uranus. The mantle also probably contains rocky material plus dissolved ammonia and methane. Circulation in that electrically conducting mantle may generate the planet's peculiar magnetic field, which is highly inclined to its axis of rotation. Above the mantle lies a deep hydrogen and helium atmosphere.

Uranus rotates on its side, with its equator inclined 98° to its orbit. As a result the winter–summer contrast is extreme, with the sun passing near each of the planet's celestial poles at the solstices, so half of the planet is in perpetual darkness and the other half in perpetual light for the 21 year–long summer and winter seasons (■ Figure 18-11). Compare that with seasons on Earth, discussed in Chapter 3. When Voyager 2 flew past in 1986, the planet's south pole was pointed almost directly at the sun. Uranus's odd rotation may have been produced when Uranus collided with a very large planetesimal late in its formation or by tidal interactions with the other giant planets as Uranus may have migrated outward early in the history of the solar system (see Chapter 16).

Voyager 2 photos show a nearly featureless ball (■ Figure 18-12a). The atmosphere is mostly hydrogen and helium, but traces of methane absorb red light and thus make the atmosphere look green-blue. There is no belt–zone circulation visible in the Voyager photographs, although extreme computer enhancement revealed a few clouds and bands around the south pole. In the decades since Voyager 2 flew past Uranus, spring has come to the northern hemisphere of Uranus and autumn to the southern hemisphere. Images made by the Hubble Space Telescope and new large Earth-based telescopes reveal changing clouds and cloud bands in both hemispheres (Figure 18-12b).

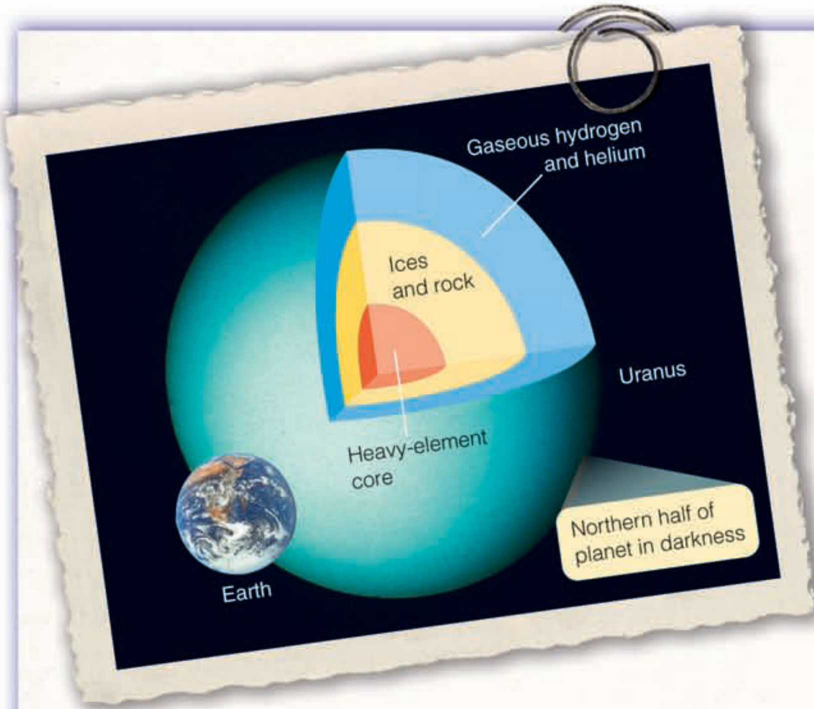
Infrared measurements show that Uranus is radiating about the same amount of energy that it receives from the sun, meaning it has much less heat flowing out of its interior than Jupiter, Saturn, or Neptune. This may account for its limited atmospheric activity. Astronomers are not sure why Uranus differs in this respect from the other Jovian worlds.

Uranus's Moons

Until recently, astronomers could see only five moons orbiting Uranus. Voyager 2 discovered ten more small moons in 1986, and more have been found in images recorded by powerful telescopes on Earth. Note that the IAU has declared that the moons of Uranus are to be named after characters in Shakespeare's plays.

The five major moons of Uranus are smaller than Earth's moon and have old, dark, cratered surfaces. A few have deep cracks, produced, perhaps, when the interior froze and expanded. In some cases, liquid water "lava" appears to have erupted and smoothed regions. Ariel is marked by broad, smooth-floored valleys that may have been cut by flowing ice (■ Figure 18-13a).

Miranda, the innermost moon, is only 14 percent the diameter of Earth's moon, but its surface is marked by grooves called **ovoids** (Figure 18-13b). These may have been caused by internal heat driving convection in the icy mantle. Rising currents of ice have deformed the crust and created the ovoids. By counting craters on the ovoids, astronomers conclude that the entire surface is old and the moon is no longer active. Perhaps it was warmed by tidal heating long ago.



Uranus rotates on its side, and when Voyager 2 flew past in 1986, the planet's south pole was pointed almost directly at the sun. (NASA)

Celestial Profile 9: Uranus

Motion:

Average distance from the sun	19.2 AU (2.87×10^9 km)
Eccentricity of orbit	0.0461
Inclination of orbit to ecliptic	0.8°
Orbital period	84.0 y
Period of rotation	17.23 h
Inclination of equator to orbit	97.9° (retrograde rotation)

Characteristics:

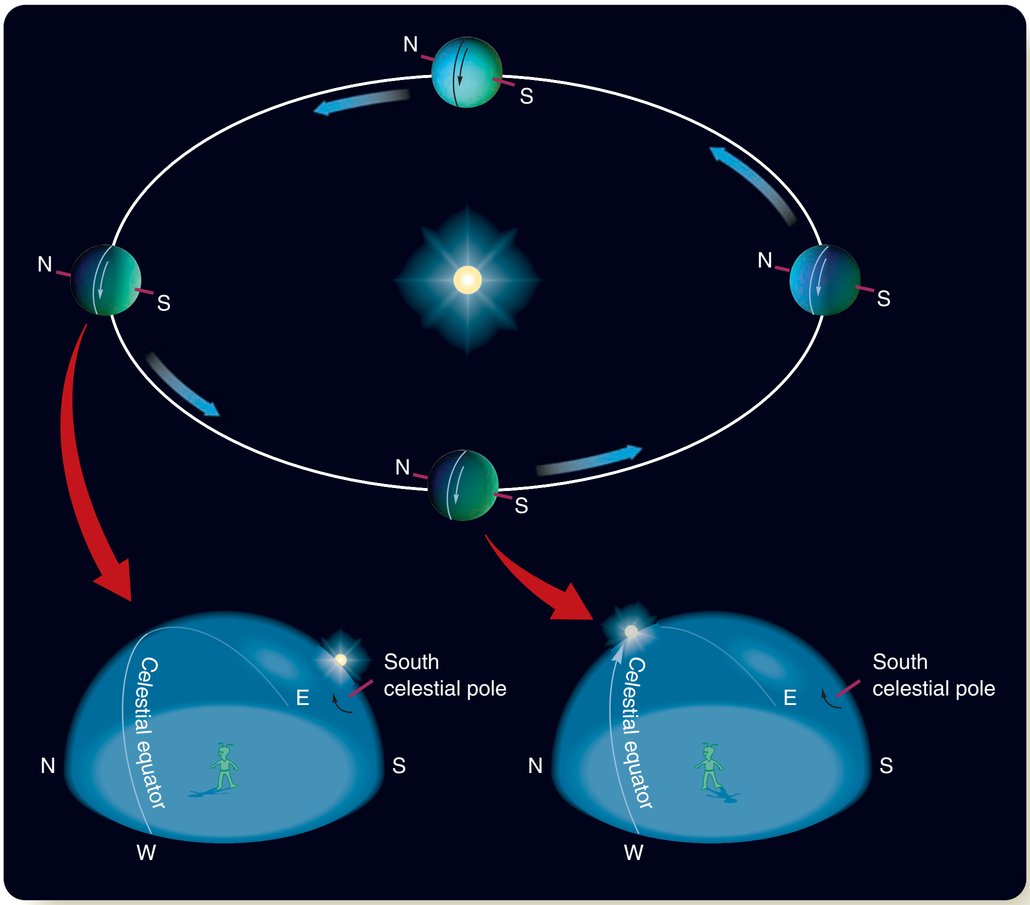
Equatorial diameter	5.11×10^4 km (4.01 $D_\oplus$)
Mass	8.69×10^{25} kg (14.5 $M_\oplus$)
Average density	1.29 g/cm^3
Gravity	0.9 Earth gravity
Escape velocity	22 km/s (2.0 $V_\oplus$)
Temperature above cloud tops	55°K (−360°F)
Albedo	0.35
Oblateness	0.023

Personality Point:

Uranus was discovered in 1781 by William Herschel, a German-born scientist who lived and worked most of his life in England. He named the new planet *Georgium Sidus,* meaning "George's Star" in Latin, after the English King George III. European astronomers, especially the French, refused to accept a planet named after an English king. Instead, they called the planet Herschel. Years later, German astronomer J. E. Bode suggested it be named Uranus after the oldest of the Greek gods.

■ **Figure 18-11**

Uranus rotates on an axis that is tipped 97.9° from the perpendicular to its orbit, so its seasons are extreme. When one of its poles is pointed nearly at the sun (a solstice), an inhabitant of Uranus would see the sun near a celestial pole, never rising or setting. As Uranus orbits the sun, the planet maintains the direction of its axis in space, and thus the sun moves from pole to pole. At the time of an equinox on Uranus, the sun would be on the celestial equator and would rise and set with each rotation of the planet. Compare with similar diagrams for Earth on page 25.

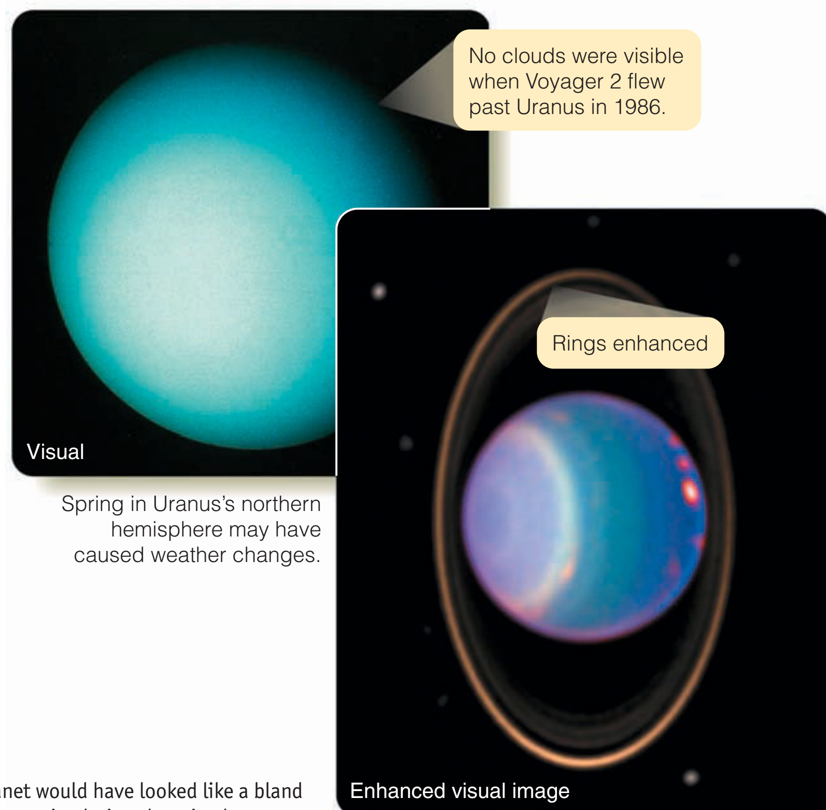


Uranus's Rings

The rings of Uranus are not easily visible from Earth. The first hint that Uranus has rings came from **occultations,** the passage of the planet in front of a star during which the rings momentarily blocked the star's light, observed by astronomers onboard the Kuiper Airborne Observatory in 1977. Most of what astronomers know about these rings comes from the observations by the Voyager 2 spacecraft. Their composition appears to be water ice mixed with methane that has been darkened by exposure to radiation.

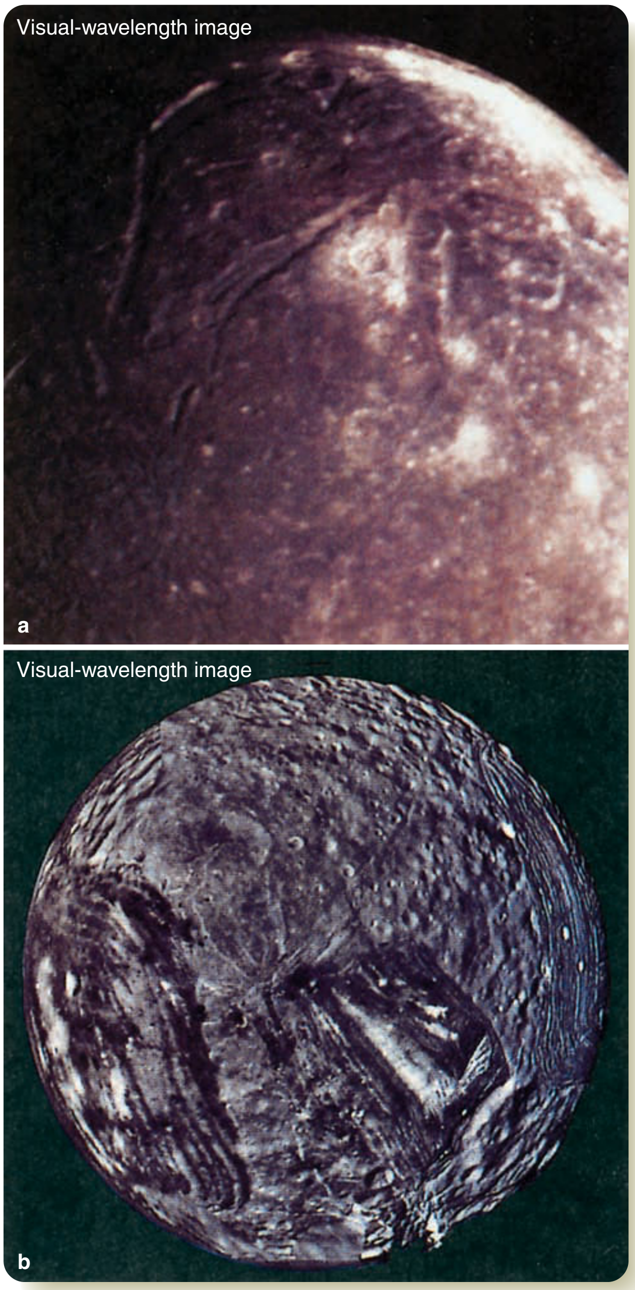
Study ■ **The Rings of Uranus and Neptune** on pages 398–399 and notice three important points:

1. The rings of Uranus were discovered during an occultation when Uranus crossed in front of a star.
2. The rings are dark, contain little dust, and are confined by small moons.
3. Particles orbiting in the rings around Uranus and Neptune cannot survive for long periods, so the rings need to be resupplied with material from impacts on moons, as is also true for the rings around Jupiter and Saturn.



■ **Figure 18-12**

If you had been riding onboard Voyager as it passed Uranus, the planet would have looked like a bland blue-green ball. Contrast-enhanced images reveal traces of belt–zone circulation deep in the atmosphere. (NASA and Erich Karkoschka, University of Arizona)



■ **Figure 18-13**

Evidence of geological activity on two Uranian moons. (a) Ariel has an old cratered surface, but some regions are marked by broad, shallow valleys with few craters. (b) The face of Miranda is marred by ovoids, which are believed to have formed when internal heating caused slow convection in the ice of the moon's mantle. Note the 5 km-high cliff at the lower edge of the moon. (NASA)

When you read about Neptune's rings later in this chapter, you can return to this artwork and see how closely the two ring systems compare.

In 2006, astronomers found two new, very faint rings orbiting far outside the previously known rings of Uranus. The newly discovered satellite Mab appears to be the source of particles for the larger ring, and the smaller of the new rings is confined between the orbits of the moons Portia and Rosalind.

A History of Uranus

Uranus never grew massive enough to capture large amounts of gas from the solar nebula as did Jupiter and Saturn. Uranus is rich in water and ice rather than in hydrogen and helium.

Modern models of the origin of the solar system suggest that Uranus and Neptune formed closer to the sun than their present positions. Interactions with massive Jupiter and Saturn could have gradually moved Uranus and Neptune outward, and tidal effects may have produced Uranus's peculiar rotation. Another theory is that Uranus was struck by a large planetesimal as it was forming and given its highly inclined rotation.

The highly inclined magnetic field of Uranus may be produced by convection in its electrically conducting mantle. With very little heat flowing out of the interior, this convection must be limited.

◄ **SCIENTIFIC ARGUMENT** ►

Why are the rings of Uranus so narrow?

Unlike the rings of Jupiter and Saturn, the rings of Uranus are quite narrow, like hoops of wire. You would expect collisions among ring particles to gradually spread the rings out into thin sheets, so something must be confining the narrow rings. In fact, two small moons have been found orbiting just inside and outside the ϵ ring. If a ring particle drifts away from the ring, the corresponding moon's gravity will nudge it back into the ring, so they are called "shepherd satellites" or "shepherd moons." More shepherd moons, too small to have been detected so far, are thought to control the other rings. Thus, the rings of Uranus resemble Saturn's narrow F ring.

Now expand your argument. **How do moons happen to be in the right place to keep the rings narrow?**

◄ ►

18-5 Neptune

Through a telescope on Earth, Neptune looks like a tiny blue dot with no visible cloud features. In 1989, Voyager 2 flew past and revealed some of Neptune's secrets.

Planet Neptune

Almost exactly the same size as Uranus, Neptune has a similar interior. Model calculations predict that a small core of heavy elements lies within a slushy mantle of water, ices, and rocky materials below a hydrogen-rich atmosphere (■ Celestial Profile 10). Yet, Neptune looks quite different on the outside from Uranus; Neptune is dramatically blue and has active cloud formations. Neptune's dark blue tint is caused by its atmospheric composition of one and a half times more methane than Uranus. Methane absorbs red photons better than blue and scatters blue photons better than red, giving Neptune a blue color and Uranus a green-blue color.

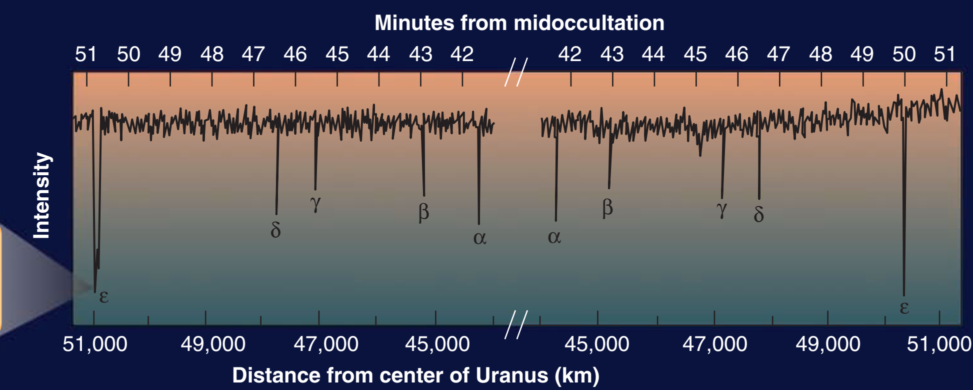
Atmospheric circulation on Neptune is much more dramatic than on Uranus. When Voyager 2 flew by Neptune in 1989, the largest feature was the Great Dark Spot (■ Figure 18-14). Roughly

The Rings of Uranus and Neptune

1 The rings of Uranus were discovered in 1977, when Uranus crossed in front of a star. During this occultation, astronomers saw the star dim a number of times before and again after the planet crossed over the star. The dips in brightness were caused by rings circling Uranus.

More rings were discovered by Voyager 2. The rings are identified in different ways depending on when and how they were discovered.

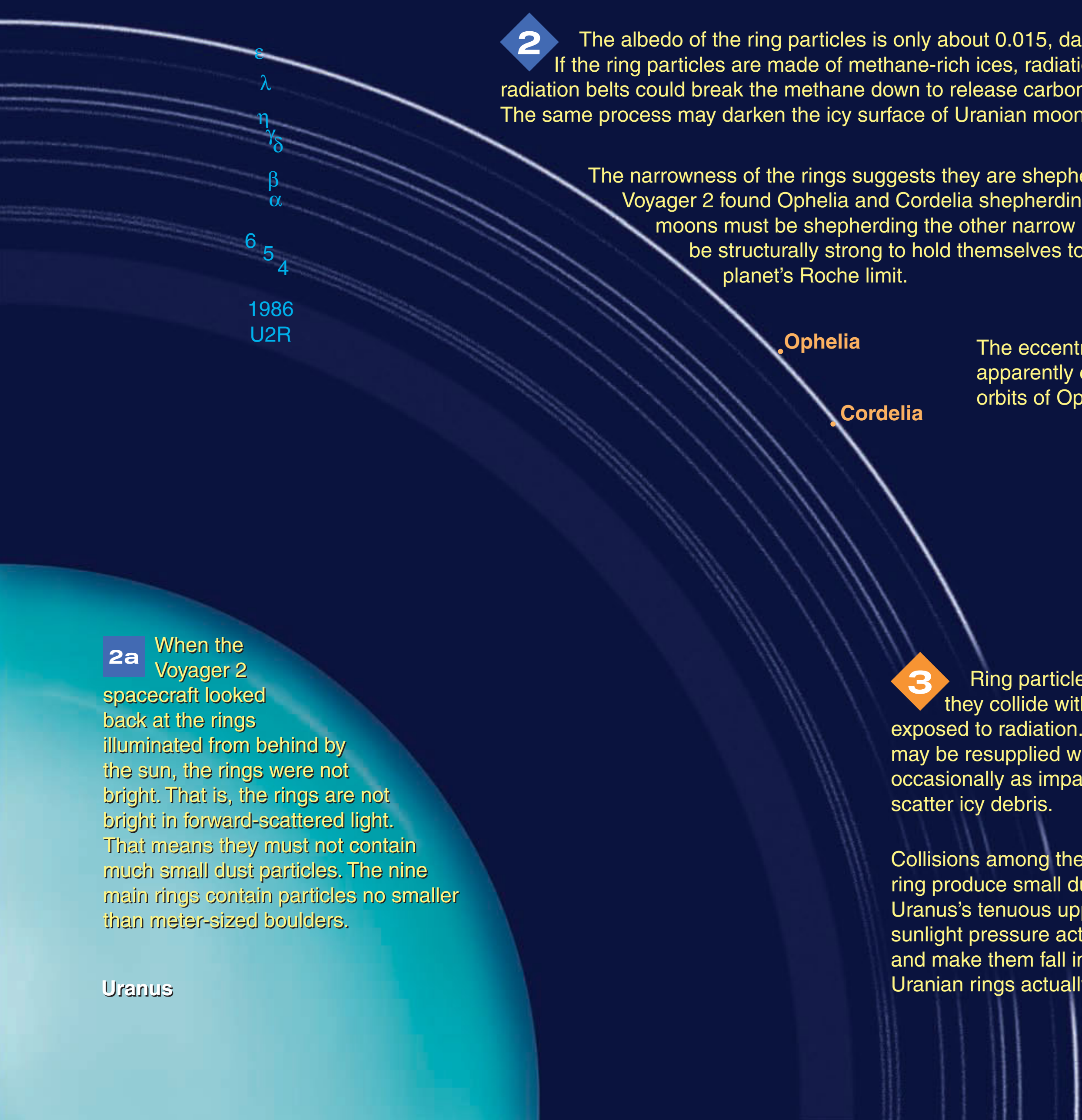
CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see the Active Figure "Uranus's Ring Detection" and animate this diagram.



Notice the eccentricity of the ε ring. It lies at different distances on opposite sides of the planet.

2 The albedo of the ring particles is only about 0.015, darker than lumps of coal. If the ring particles are made of methane-rich ices, radiation from the planet's radiation belts could break the methane down to release carbon and darken the ices. The same process may darken the icy surface of Uranian moons.

The narrowness of the rings suggests they are shepherded by small moons. Voyager 2 found Ophelia and Cordelia shepherding the ε ring. Other small moons must be shepherding the other narrow rings. Such moons must be structurally strong to hold themselves together inside the planet's Roche limit.

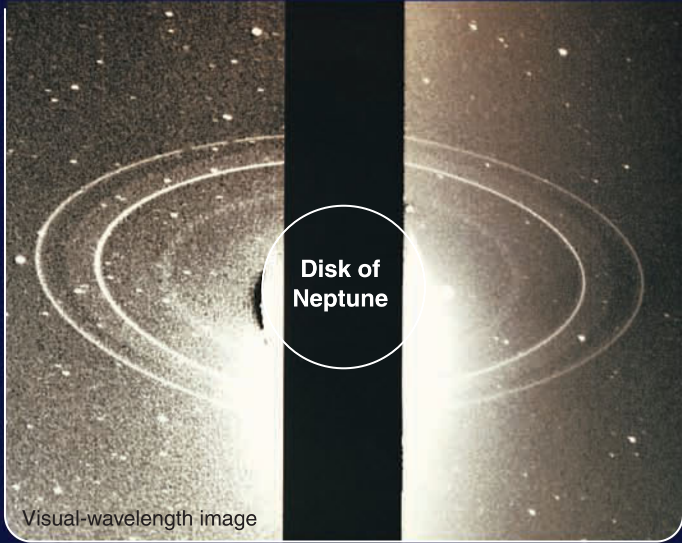


The eccentricity of the ε ring is apparently caused by the eccentric orbits of Ophelia and Cordelia.

2a When the Voyager 2 spacecraft looked back at the rings illuminated from behind by the sun, the rings were not bright. That is, the rings are not bright in forward-scattered light. That means they must not contain much small dust particles. The nine main rings contain particles no smaller than meter-sized boulders.

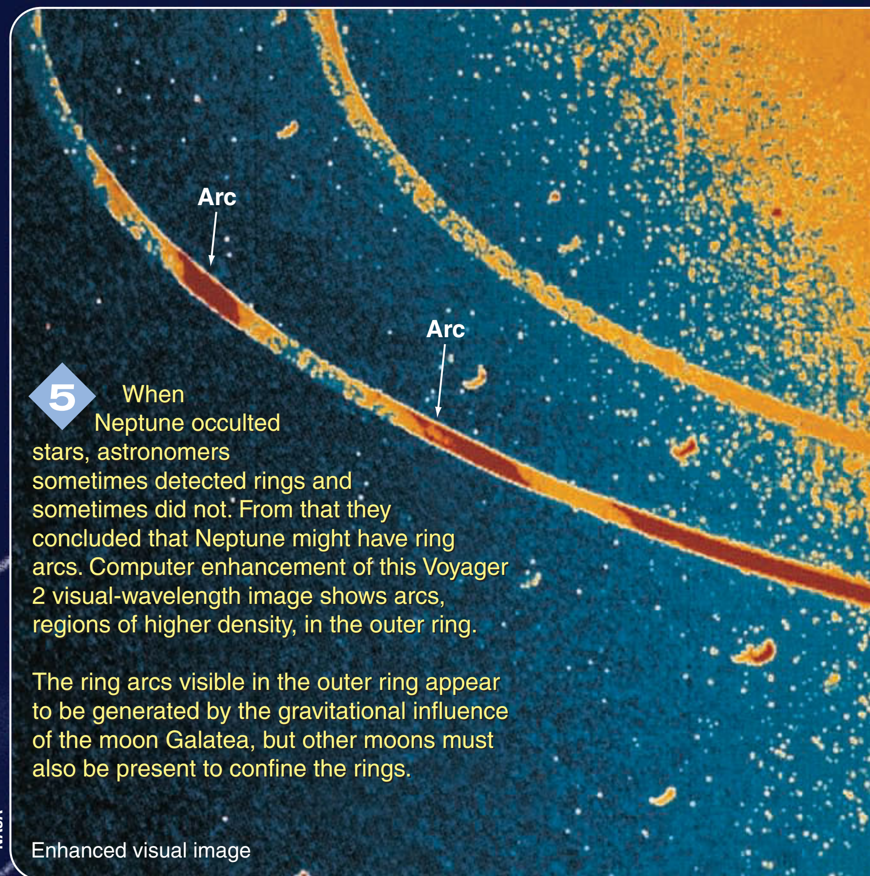
3 Ring particles don't last forever as they collide with each other and are exposed to radiation. The rings of Uranus may be resupplied with fresh particles occasionally as impacts on icy moons scatter icy debris.

Collisions among the large particles in the ring produce small dust grains. Friction with Uranus's tenuous upper atmosphere plus sunlight pressure act to slow the dust grains and make them fall into the planet. The Uranian rings actually contain very little dust.



The rings of Neptune are bright in forward-scattered light, as in the image above, and that indicates that the rings contain significant amounts of dust. The ring particles are as dark as those that circle Uranus, so they probably also contain methane-rich ice darkened by radiation.

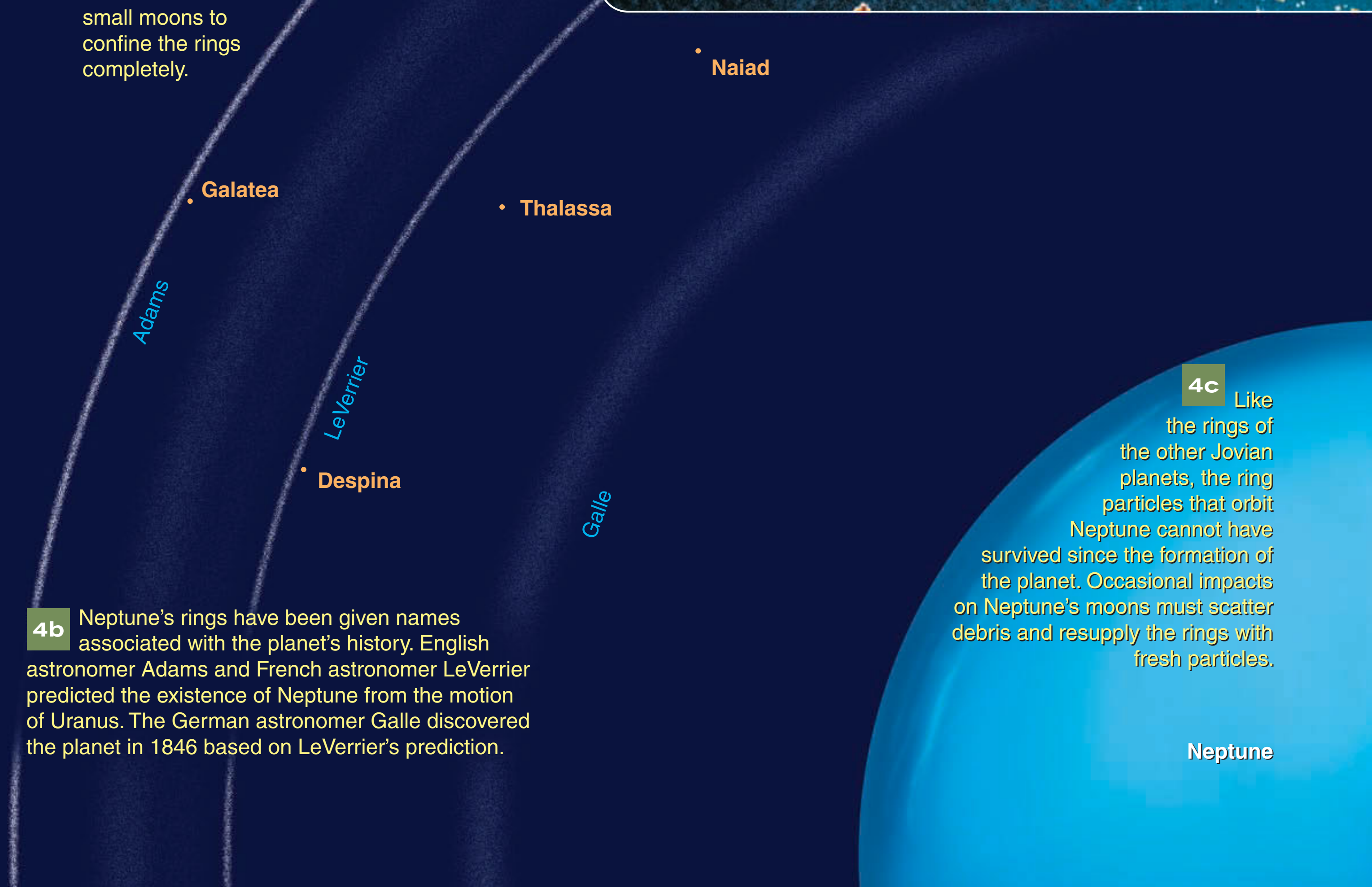
4 The brightness of Neptune is hidden behind the black bar in this Voyager 2 image. Two narrow rings are visible, and a wider, fainter ring lies closer to the planet. More ring material is visible between the two narrow rings.



5 When Neptune occulted stars, astronomers sometimes detected rings and sometimes did not. From that they concluded that Neptune might have ring arcs. Computer enhancement of this Voyager 2 visual-wavelength image shows arcs, regions of higher density, in the outer ring.

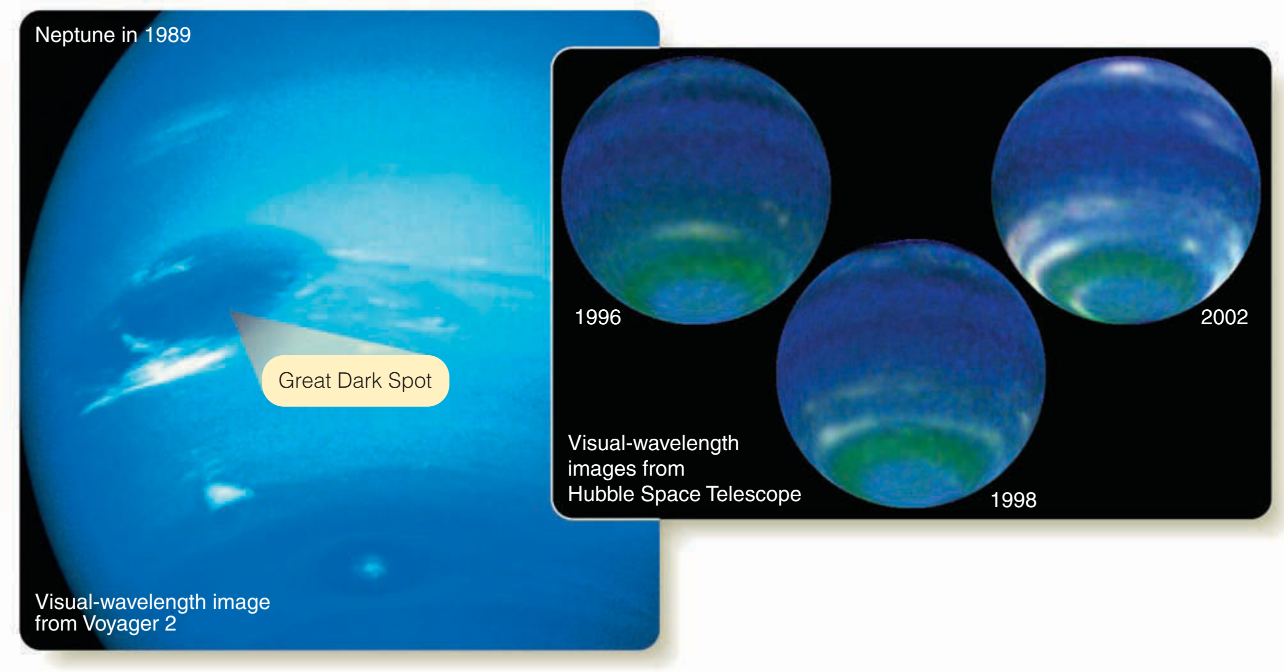
The ring arcs visible in the outer ring appear to be generated by the gravitational influence of the moon Galatea, but other moons must also be present to confine the rings.

4a Neptune's rings lie in the plane of the planet's equator and inside the Roche limit. The narrowness of the rings suggests that shepherd moons must confine them, and a few such moons have been found among the rings. There must be more undiscovered small moons to confine the rings completely.



4b Neptune's rings have been given names associated with the planet's history. English astronomer Adams and French astronomer LeVerrier predicted the existence of Neptune from the motion of Uranus. The German astronomer Galle discovered the planet in 1846 based on LeVerrier's prediction.

4c Like the rings of the other Jovian planets, the ring particles that orbit Neptune cannot have survived since the formation of the planet. Occasional impacts on Neptune's moons must scatter debris and resupply the rings with fresh particles.



■ **Figure 18-14**

Neptune's axis is inclined almost 29 degrees to its orbit. It experiences seasons that each last about 40 years. Since Voyager visited in 1989, spring has come to the southern hemisphere, and the weather has clearly changed, which is surprising because sunlight at Neptune is 900 times dimmer than at Earth. (NASA, L. Sromovsky, and P. Fry, University of Wisconsin–Madison)

the size of Earth, the spot seemed to be an atmospheric circulation pattern much like Jupiter's Great Red Spot. Smaller spots were visible in Neptune's atmosphere, and photos showed they were circulating like hurricanes. More recently, the Hubble Space Telescope has photographed Neptune and found that the Great Dark Spot is gone and new cloud formations have appeared, as shown in Figure 18-14. Evidently, the weather on Neptune is changeable.

The atmospheric activity on Neptune is apparently driven by heat flowing from the interior plus some contribution by dim sunlight 30 AU from the sun. The heat causes convection in the atmosphere, which the rapid rotation of the planet converts into high-speed winds, high-level white clouds of methane ice crystals, and rotating storms visible as spots. Neptune may have more activity than Uranus because it has more heat flowing out of its interior, for reasons that are unclear.

Like Uranus, Neptune has a highly inclined magnetic field that must be linked to circulation in the interior. In both cases, astronomers suspect that ammonia dissolved in the liquid water mantle makes the mantle a good electrical conductor and that convection in the water, coupled with the rotation of the planet, drives the dynamo effect and generates the magnetic field.

Neptune's Moons

Neptune has two moons that were discovered from Earth before Voyager 2 flew past in 1989. Voyager discovered six more very small moons. Since then, a few more small moons have been found by astronomers using Earth-based telescopes.

The two largest moons have peculiar orbits. Nereid, about a tenth the size of Earth's moon, follows a large, elliptical orbit, taking nearly an Earth year to circle Neptune once. Triton, almost 80 percent the size of Earth's moon, orbits Neptune backward. These odd orbits suggest that the system was disturbed long ago in an interaction with some other body, such as a massive planetesimal.

With a temperature of 37 K (−393°F), Triton has an atmosphere of nitrogen and methane about 10^5 times less dense than Earth's (see Figure 17-15). A significant part of Triton is ice, and deposits of nitrogen frost are visible at the southern pole (■ Figure 18-15a), which, at the time Voyager 2 flew past, had been turned toward sunlight for 30 years. The nitrogen frost appears to be vaporizing in the sunlight and is probably refreezing in the darkness at Triton's north pole.

Many features on Triton suggest it has had an active past. It has few craters on its surface, but it does have long faults that appear to have formed when the icy crust broke. Some approximately round basins about 400 km in diameter appear to have been flooded time after time by liquids from the interior (Figure 18-15b). Analysis of dark smudges visible in the southern polar cap (Figure 18-15a) reveals that these are deposits produced when liquid nitrogen in the crust, warmed by the sun, erupts volcanically through vents and spews up to 8 km high into the atmosphere. Methane in the gas is converted by sunlight into dark deposits that fall back, leaving black smudges.

By counting craters on Triton, planetary scientists conclude that the surface has been active as recently as a million years ago

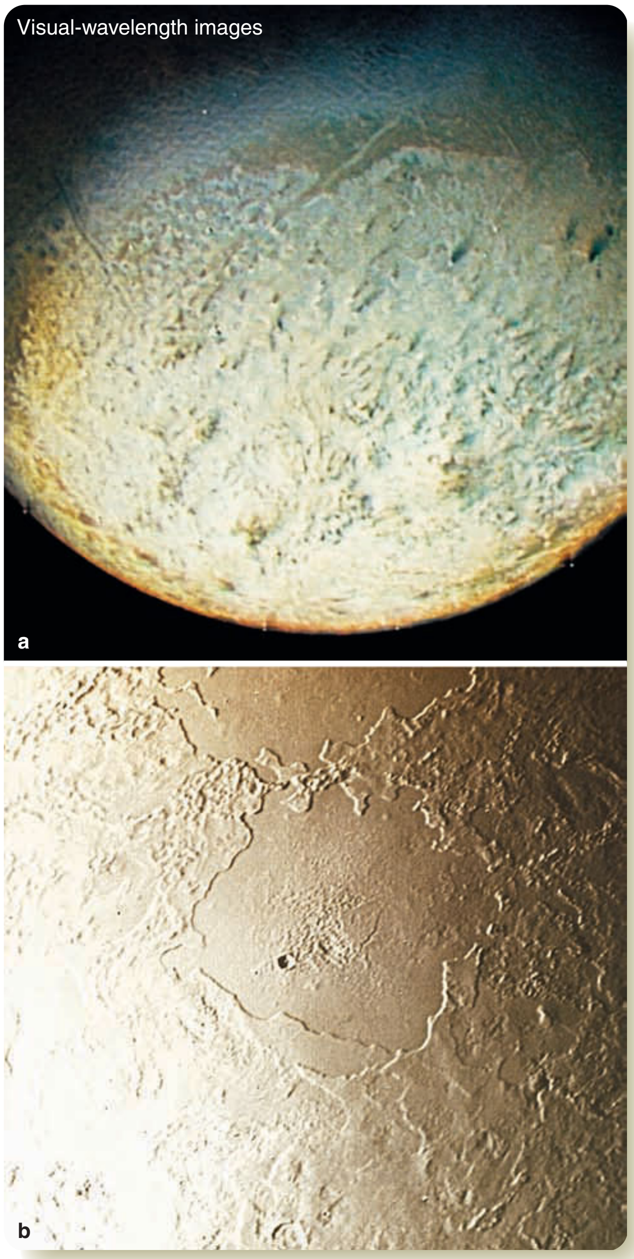
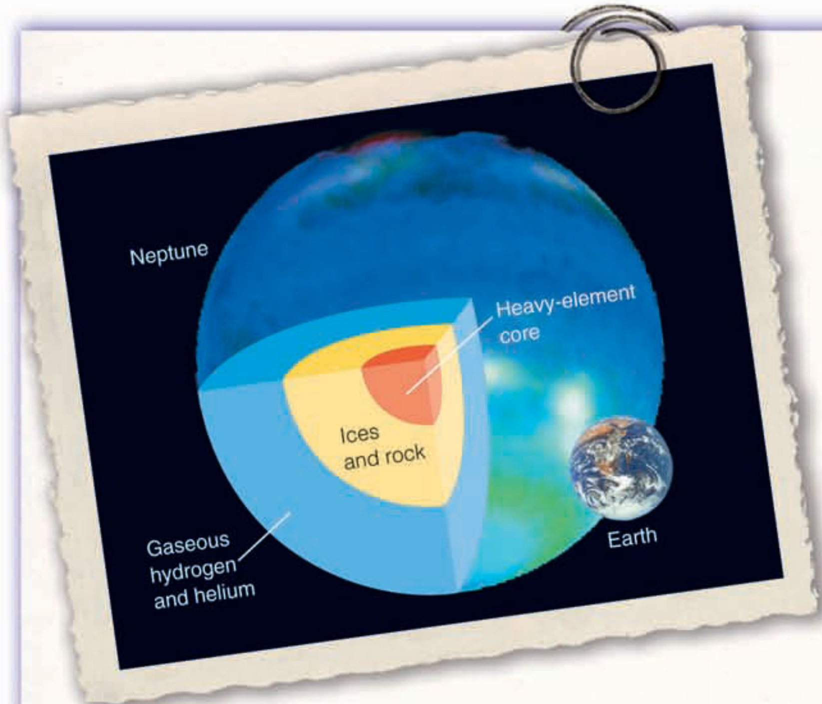


Figure 18-15

Visible-wavelength images of Neptune's moon Triton. (a) Triton's southern polar cap is formed of nitrogen frost. Note dark smudges caused by organic compounds sprayed from nitrogen geysers, and the absence of craters. (b) These round basins on Triton appear to have been repeatedly flooded by liquid from the interior. (NASA)

and may still be active. The energy source for Triton's volcanism could come from radioactive decay. The moon is two-thirds rock, and although such a small world would not be able to generate sufficient radioactive decay to keep molten rock flowing to its surface, frigid Triton may be the site of water–ammonia volcanism: A mixture of water and ammonia could melt at very low temperatures and erupt to resurface parts of the moon.



Neptune was tipped slightly away from the sun when the Hubble Space Telescope recorded this image. The interior is much like that of Uranus, but Neptune has more heat flowing outward. (NASA)

Celestial Profile 10: Neptune

Motion:

Average distance from the sun	30.1 AU (4.50×10^9 km)
Eccentricity of orbit	0.0100
Inclination of orbit to ecliptic	1.8°
Orbital period	164.8 y
Period of rotation	16.05 h
Inclination of equator to orbit	28.8°

Characteristics:

Equatorial diameter	4.95×10^4 km (3.93 $D_\oplus$)
Mass	1.03×10^{26} kg (17.2 $M_\oplus$)
Average density	1.66 g/cm^3
Gravity	1.2 Earth gravities
Escape velocity	25 km/s (2.2 $V_\oplus$)
Temperature above cloud tops	55°K (−360°F)
Albedo	0.35
Oblateness	0.017

Personality Point:

A British and a French astronomer independently calculated the existence and location of Neptune from its gravitational influence on the motion of Uranus. British observers were too slow to act on this information; Neptune was discovered in 1846, and the French astronomer got the credit. Because of its blue color, astronomers named Neptune after the god of the sea.

Neptune's Rings

Neptune's rings are faint and very hard to detect from Earth, but they illustrate some interesting processes of comparative planetology.

Look again at ■ **The Rings of Uranus and Neptune** on pages 398–399 and compare the rings of Neptune with those of Uranus. Notice two additional points:

4 Neptune's rings, named after the astronomers involved in the discovery of the planet, are similar to those of Uranus but contain more small dust particles.

5 Also, Neptune's rings show another way that moons can interact with rings: One of Neptune's moons is producing short arcs in the outermost ring.

Neptune's rings resemble the rings of Uranus, Saturn, and Jupiter in one important way. As you have already learned, these rings can't be primordial. That is, they can't have lasted since the formation of the planets. Planetary rings are constantly being remade.

The History of Neptune

Neptune must have formed much as Uranus did, growing slowly and never becoming massive enough to trap large amounts of hydrogen and helium. It developed a core of heavy elements, a mantle of slushy ices and rock, and a deep hydrogen-rich atmosphere. Neptune's internal heat may be generated partly by radioactive decay in its core and partly by dense material sinking inward. Laboratory experiments show that, at the temperatures and pressures expected deep in the atmospheres of Neptune and Uranus, methane can decompose, and the released carbon might form diamond crystals perhaps as large as pebbles. A continuous flow of diamonds falling into a planet's interior would release energy and could help warm the planet. This process may be the source of some of Neptune's internal heat, but the lack of internal heat in Uranus remains a puzzle. (The possibility of a planet-wide hailstorm of diamonds serves to warn you that other worlds are truly un-Earthly and may harbor things you can hardly imagine.) The heat flowing outward toward Neptune's surface can drive convection, produce a magnetic field, and help create atmospheric circulation.

The moons of Neptune suggest some cataclysmic encounter long ago that put Nereid into a long-period elliptical orbit and Triton into a retrograde orbit. You have seen evidence of major impacts throughout the solar system, so such interactions may have been fairly common. Certainly, impacts on the satellites could provide the debris that is trapped among the smaller moons to form the rings.

◄ SCIENTIFIC ARGUMENT ►

Why is Neptune blue?

In building this argument, you must be careful not to be misled by the words you use. When you look at something, you really turn your eyes toward it and receive light from the object. The light Earth receives from Neptune is sunlight that is scattered from various layers of Neptune and journeys to your eyes. Sunlight entering Neptune's atmosphere must pass through hydrogen gas that contains a small amount of methane, which is a good absorber of longer wavelengths. As a result, red photons are more likely to be absorbed than blue photons, and that makes the light bluer. Furthermore, when the light is scattered in deeper layers, the shorter-wavelength photons are most likely to be scattered, and thus the light that finally emerges from the atmosphere and reaches your telescope is poor in longer wavelengths. It looks blue.

This argument shows how a careful, step-by-step analysis of a natural process can help you better understand how nature works. Now expand your argument. **Why do the clouds on Neptune look white?**

◄ ►

18-6 Pluto—The First Dwarf Planet

Out on the edge of the solar system orbits a family of small, icy worlds. Pluto was the first to be discovered, in 1930, but modern telescopes have found more.

You may have learned in school that there are nine planets in our solar system, but in 2006 the International Astronomical Union voted to remove Pluto from the list of planets and reclassify it as a "dwarf planet." Pluto is a very small, icy world: It isn't Jovian, and it isn't Terrestrial. Its orbit is highly inclined and elliptical enough that Pluto actually comes closer to the sun than Neptune at times. To understand Pluto's status, you must use comparative planetology to analyze Pluto and then compare it with its neighbors.

Pluto is very difficult to observe from Earth. It has only 65 percent the diameter of Earth's moon. In Earth-based telescopes, it never looks like more than a faint point of light and even in Hubble Space Telescope images it shows little detail. Orbiting so far from the sun, Pluto is cold enough to freeze most compounds you think of as gases, and spectroscopic observations have found evidence of nitrogen ice on its surface. Pluto has a thin atmosphere of nitrogen and carbon monoxide with small amounts of methane.

Pluto has three moons. Two, named Nix and Hydra, are quite small, but Charon is relatively large, with half of Pluto's diameter. Charon orbits Pluto with a period of 6.4 days in an orbit highly inclined to the ecliptic (■ Figure 18-16). Pluto and Charon are tidally locked to face each other, so Pluto's rotation is also highly inclined.

Charon's orbit size and period plus Kepler's third law reveal that the mass of the system is only about 0.002 Earth mass. Most of that mass is Pluto, which has about 12 times the mass of Charon. Knowing the diameters and masses of Pluto and Charon allows astronomers to calculate that their densities are both about 2 g/cm^3. This indicates that Pluto and Charon must contain about 35 percent ice and 65 percent rock.

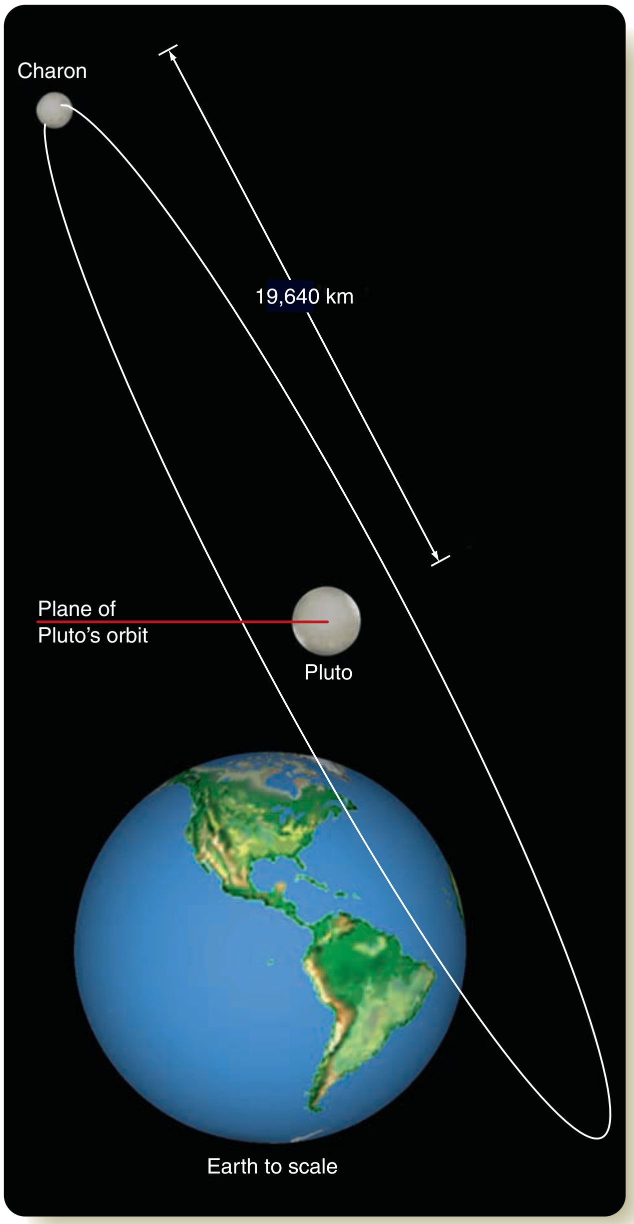


Figure 18-16

Pluto, Pluto's moon Charon, and Charon's orbit are pictured here in scale relative to Earth. Charon's orbit is tipped 118° to the plane of Pluto's orbit.

The best photos by the Hubble Space Telescope reveal almost no surface detail, but you know enough about icy moons to guess that Pluto has craters and probably shows signs of tidal heating caused by interaction with its large moon Charon. The New Horizons spacecraft will fly past Pluto in July 2015, and the images radioed back to Earth will certainly show that Pluto is an interesting world.

What Defines a Planet?

To understand why Pluto is no longer considered a planet, you should recall the Kuiper belt (Chapter 16). Since 1992 astronomers have discovered more than a thousand icy bodies orbiting beyond Neptune. There may be as many as 100 million objects in the Kuiper belt larger than 1 km in diameter. They appear to be icy planetesimals left over from the outer solar nebula. Some of the Kuiper belt objects are quite large, and one, Eris, is 5 percent larger in diameter than Pluto. Three other Kuiper belt objects found so far, Sedna, Quaoar (pronounced *kwah-o-wahr*), and Orcus, are half the size of Pluto or larger. Some of these objects have moons of their own. In that way, they resemble Pluto and its three moons.

A bit of comparative planetology shows that Pluto is not related to the Jovian or Terrestrial planets; it is obviously a member of a newfound family of icy worlds that orbit beyond Neptune. These bodies must have formed at about the same time as the eight classical planets of the solar system, but they did not grow massive enough to clear their orbital zones of remnant planetesimals and consequently remain embedded among a swarm of other objects in the Kuiper belt.

The IAU's criteria for full planet status is that an object must be large enough that its gravity has pulled it into a spherical shape and also large enough to dominate and gravitationally clear its orbital region of most or all other objects over a span of billions of years. Eris and Pluto, the largest objects found so far in the Kuiper belt, and Ceres, the largest object in the asteroid belt, are too small to clear their orbital zones of other objects and therefore do not meet the standard for being called planets. On the other hand all three are large enough to be spherical, so they are the prototypes of a new class of objects defined by the IAU as **dwarf planets.**

Pluto and the Plutinos

No, this section is not about a 1950s rock band. It is about the history of the dwarf planets, and it will take you back 4.6 billion of years to watch the outer planets form.

Over a hundred Kuiper belt objects are known that are caught with Pluto in a 3:2 resonance with Neptune. That is, they orbit the sun twice while Neptune orbits three times. These Kuiper belt objects have been named **plutinos.** The plutinos formed in the outer solar nebula, but how did they get caught in resonances with Neptune? You have learned (Chapter 16) that models of the formation of the planets suggest that Uranus and Neptune may have formed closer to the sun and that sometime later, gravitational interactions among the Jovian planets could have gradually shifted Uranus and Neptune outward. As Neptune migrated outward, its orbital resonances could have swept up small objects like a strange kind of snowplow. The plutinos are caught in the 3:2 resonance, and other Kuiper belt objects are caught in other resonances. The evidence appears to support models that predict that Uranus and Neptune migrated outward.

The migration of the outer planets would have dramatically upset the motion of some of these Kuiper belt objects, and some could have been thrown inward where they could interact with the Jovian planets. Some of those objects may have been captured as moons, and astronomers wonder if moons such as Neptune's Triton could have started life as Kuiper belt objects. Other objects may have hit bodies in the inner solar system and caused the late heavy bombardment episode especially evident on the surface of Earth's moon. The small frozen worlds on the fringes of the solar system may hold clues to the formation of the planets 4.6 billion years ago, and the subsequent history of Earth.

◄ SCIENTIFIC ARGUMENT ►

What evidence indicates that cataclysmic impacts have occurred in our solar system?

To build this argument, you must cite plenty of evidence. The peculiar orbits of Neptune's moons Triton and Nereid, the peculiar rotation of Uranus, and Pluto's large satellite Charon, all hint that impacts and encounters with large planetesimals or comet nuclei have been important in the history of these worlds. Furthermore, the existence of planetary rings suggests that impacts have scattered small particles and replenished the ring systems. Even in the inner solar system, the retrograde rotation of Venus, the high density of Mercury, the smooth northern lowlands of Mars and the formation of Earth's moon are possible consequences of major impacts when the solar system was young.

Now assemble evidence in a new argument. **How does the origin of the plutinos give a clue about one possible source of impacting bodies in the solar system?**

◄ ►

What Are We? Trapped

No one has ever been further from Earth than the moon. We humans have sent robotic spacecraft to visit most of the larger worlds in our solar system, and we have found them to be strange and wonderful places, but no human has ever set foot on any of them. We are trapped on Earth.

We lack the technology to leave Earth. Getting away from Earth's gravitational field is difficult and calls for very large rockets. America built such rockets in the 1960s and early 1970s. They could send astronauts to the moon, but such rockets no longer exist. The best technology today can carry astronauts just a few hundred kilometers above Earth's surface to orbit above the atmosphere. The United States is beginning an ambitious plan to build a new generation of human-piloted rockets meant to carry people back to the moon, and eventually to Mars. Does Earth's civilization have the resources to build spacecraft capable of carrying human explorers to other worlds? We'll have to wait and see.

In the previous chapter, you discovered another reason we Earthlings are trapped on Earth. We have evolved to fit the environment on Earth. None of the planets or moons you explored in this chapter would welcome you. Lack of air, and extreme heat or cold, are obvious problems, but, also, Earthlings have evolved to live with Earth's gravity. Astronauts in space for just a few weeks suffer biomedical problems because they are no longer in Earth's gravity. Living in a colony on Mars or the moon might raise similar problems. Just getting to the outer planets would take decades of space travel; living for years in a colony on one of the Jovian moons under low gravity and exposed to the planet's radiation belts may be beyond the capability of the human body. We may be trapped on Earth not because we lack large enough rockets but because we need Earth's protection.

It seems likely that we need Earth more than it needs us. The human race is changing the world we live on at a startling pace, and some of those changes could make Earth less hospitable to human life. All of your exploring of un-Earthly worlds serves to remind you of the nurturing beauty of our home planet.

Study and Review

Summary

- The Jovian planets—Jupiter, Saturn, Uranus, and Neptune—are large, massive low-density worlds located in the outer solar system.
- Models of planetary interiors can be calculated based on each planet's density and **oblateness (p. 381)**, the fraction by which its equatorial diameter exceeds its polar diameter.
- Jupiter and Saturn are composed mostly of liquid hydrogen, and for this reason they are sometimes referred to as "liquid giants."
- Uranus and Neptune contain abundant water in solid form and are sometimes called "ice giants."
- All the Jovian planets have large systems of satellites and rings that have had complex histories.
- Jupiter is observed to have heat flowing out of it at a high rate, indicating that its interior is very hot.
- Jupiter has a core of heavy elements surrounded by a deep mantle of **liquid metallic hydrogen (p. 381)** in which the planet's large and strong magnetic field is generated.

- The **magnetosphere (p. 382)** around Jupiter traps high-energy particles from the sun to form intense radiation belts.
- Jupiter's atmosphere contains three layers of clouds formed of hydrogen-rich molecules, including water and ammonia.
- Atmospheres of the Jovian planets are marked by **belt–zone circulation (p. 382)** that produces cloud belts parallel to their equators. Zones are high-pressure regions of rising gas, and belts are lower-pressure areas of sinking gas.
- Spots in Jupiter's atmosphere, including the Great Red Spot, are circulating weather patterns.
- The four large **Galilean moons (p. 383)** show signs of geologic activity. **Grooved terrain (p. 383)** on Ganymede, smooth ice and cracks on Europa, and active volcanoes on Io show that **tidal heating (p. 383)** driven by **orbital resonances (p. 386)** has made these moons active.
- Many of Jupiter's moons are small, rocky bodies that could be captured objects. They are too small to retain heat and are not geologically active.
- Jupiter's ring is composed of dark particles that are bright when illuminated from behind, which is called **forward scattering (p. 387).** This shows that the particles are very small, and are probably dust from meteoroid impacts on moons.
- Jupiter's ring, like all of the rings in the solar system, lies inside the planet's **Roche limit (p. 387),** within which tidal stress can destroy a moon or prevent one from forming.
- Saturn is less dense than water and contains a small core and less metallic hydrogen than Jupiter.
- Cloud layers on Saturn occur at the same temperature as those on Jupiter, but, because Saturn is farther from the sun and colder, the cloud layers are deeper in the hydrogen atmosphere below a layer of methane haze and so are less prominent.
- Saturn's moons are icy and mostly heavily cratered.
- Saturn's largest moon, Titan, has a cold, cloudy nitrogen atmosphere. It may have had methane falling as rain on its icy surface and forming lakes and rivers.
- Sunlight entering Titan's atmosphere can convert methane into complex carbon-rich molecules to form haze and particles that settle out to coat the surface with dark organic material.
- Enceladus has a light surface with some uncratered regions. Geysers of water and ice vent from the region around its south pole and provide ice particles to the E ring.
- Saturn's rings are composed of icy particles ranging in size from boulders to dust. The composition and brightness of the ring particles vary from place to place in the rings.
- Grooves in the rings can be produced by orbital resonances, or waves that propagate through the rings, caused by moons near or within the rings.
- Narrow rings and sharp ring edges can be produced by **shepherd satellites (p. 393).**
- The material observed now in the Jovian planets' rings cannot have lasted since the formation of the solar system. The rings must be replenished occasionally with material produced by meteoroids, asteroids, and comets colliding with moons.
- Uranus is much less massive than Jupiter, and its internal pressure cannot produce liquid hydrogen. It has a heavy-element core and a mantle of solid or slushy ice and rocky material below a hydrogen-rich atmosphere.
- Little heat flows out of Uranus, so it cannot be very hot inside.
- Uranus's atmosphere is almost featureless at visual wavelengths with a pale blue color caused by traces of methane, which absorbs red light. Images at selected wavelengths can be enhanced to show traces of belt–zone circulation.
- Uranus rotates on its side, perhaps because of a major impact during its early history or because of tidal interactions with other planets when it was young.
- The larger moons of Uranus are icy and heavily cratered, with signs on some of past geological activity, including **ovoids (p. 395)** on Miranda.
- The rings of Uranus, discovered by stellar **occultations (p. 396)**, are narrow hoops confined by shepherd satellites. The particles appear to be ice with traces of methane darkened by the radiation belt.
- Neptune is an ice giant like Uranus with no liquid hydrogen. Unlike Uranus, Neptune does have heat flowing outward from its interior.
- The atmosphere of Neptune, marked by traces of belt–zone circulation, is rich in hydrogen and colored blue by traces of methane.
- Neptune's satellite system is odd in that distant Nereid follows an elliptical orbit and Triton orbits backward. These may be signs of catastrophic encounters with other objects early in the solar system's history.
- Triton is icy with a thin atmosphere and frosty polar caps. Smooth areas suggest past geological activity, and dark smudges mark the location of active nitrogen geysers.
- Neptune's rings are made of icy particles in narrow hoops and contain arcs produced by the gravitational influence of one or more moons.
- Pluto is a small, icy world with three moons, one of which, Charon, is quite large in relation to Pluto. The moons' orbital plane and Pluto's equator are highly inclined to Pluto's orbit around the sun.
- Pluto, now classified by the **IAU (p. 380)** as a **dwarf planet (p. 403),** is a member of a family of icy bodies in the Kuiper belt orbiting beyond Neptune. At least one of these objects, Eris, is a bit larger than Pluto.
- Some Kuiper belt objects, called **plutinos (p. 403),** follow orbits like Pluto that have an orbital resonance with Neptune.
- Model calculations suggest that Uranus and Neptune formed closer to the sun and migrated outward, pushing millions of icy bodies in orbital resonances farther from the sun to form the Kuiper belt. Pluto may be one of those objects.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. Why is Jupiter so much richer in hydrogen and helium than Earth?
2. How can Jupiter have a liquid interior and not have a definite liquid surface?
3. How does the dynamo effect account for the magnetic fields of Jupiter, Saturn, Uranus, and Neptune?
4. Why are the belts and zones on Saturn less distinct than those on Jupiter?
5. Why do astronomers conclude that none of the Jovian planets' rings can be left over from the formation of the planets?
6. How can a moon produce a gap in a planetary ring system?
7. Explain why the amount of geological activity on Jupiter's moons varies with distance from the planet.
8. What makes Saturn's F ring and the rings of Uranus and Neptune so narrow?
9. Why is the atmospheric activity of Uranus less than that of Saturn and Neptune?
10. Why do astronomers conclude that Saturn's moon Enceladus is geologically active?
11. What are the seasons on Uranus like?
12. Why are Uranus and Neptune respectively blue-green and blue?
13. What evidence is there that Neptune's moon Triton has been geologically active recently?
14. How do astronomers account for the origin of Pluto?
15. What evidence indicates that catastrophic impacts have occurred in the solar system's past?

16. **How Do We Know?** What is the difference between technology and basic science?
17. **How Do We Know?** Why are private foundations and the government, rather than industry, the usual sources of funding for basic research in astronomy.

Discussion Questions

1. Some astronomers argue that Jupiter and Saturn are unusual, while other astronomers argue that all planetary systems should contain one or two such giant planets. What do you think? Support your argument with evidence.
2. Why don't the Terrestrial planets have rings? If you were to search for a ring among the Terrestrial planets, where would you look first?

Problems

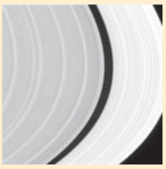
1. What is the maximum angular diameter of Jupiter as seen from Earth? Repeat this calculation for Neptune. (*Hints:* See Celestial Profiles 7 and 10, plus Reasoning with Numbers 3-1.)
2. What is the angular diameter of Jupiter as seen from Callisto? From Amalthea? (*Hint:* See Appendix A, plus Reasoning with Numbers 3-1.)
3. Measure the polar and equatorial diameters of Saturn in the photograph in Celestial Profile 8 and calculate the planet's oblateness.
4. If you observe light reflected from Saturn's rings, you should see a redshift at one edge of the rings and a blueshift at the other edge. If you observe a spectral line and see a difference in wavelength of 0.0560 nm, and the unshifted wavelength (observed in the laboratory) is 500 nm, what is the orbital velocity of particles at the outer edge of the rings? (*Hint:* See Reasoning with Numbers 6-2.)
5. One way to recognize a distant planet is by its motion along its orbit. If Uranus circles the sun in 84 years, how many arc seconds will it move in 24 hours? (*Hint:* Ignore the motion of Earth.)
6. If Uranus's ϵ ring is 50 km wide and the orbital velocity of Uranus is 6.8 km/s, how long should the occultation last that you expect to observe when the ring crosses in front of the star?
7. If Neptune's clouds have a temperature of 55 K, at what wavelength will they radiate the most energy? (*Hint:* See Reasoning with Numbers 6-1.)
8. How long did it take radio commands to travel from Earth to Voyager 2 as it passed Neptune? (*Hints:* See Appendix A, and assume Earth and Neptune were as close as possible during the Voyager 2 encounter.)
9. What is the angular diameter of Pluto as seen from the surface of Charon? (*Hint:* See Figure 18-16.)
10. Use the orbital radius and orbital period of Charon to calculate the mass of the Pluto–Charon system. (*Hints:* Express the orbital radius in meters and the period in seconds. Then, see Reasoning with Numbers 4-1.)

Learning to Look

1. This photo shows a segment of the surface of Jupiter's moon Callisto. Why is the surface mostly dark? Why are some craters dark and some white? What does this image tell you about the history of Callisto?

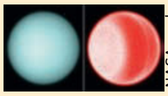
NASA/JPL

2. The Cassini spacecraft recorded this photo of Saturn's A ring and Encke's division. What do you see in this photo that tells you about processes that confine and shape planetary rings?



NASA/JPL/Space Science Institute

3. Two images of Uranus's northern hemisphere show it as it would look to the eye and through a red filter that enhances methane clouds. What do the atmospheric features tell you about circulation on Uranus?



NASA and Heidi Hammel

19 Meteorites, Asteroids, and Comets

Comets can be terrifying to the superstitious, but they are dramatically beautiful and carry clues to the origin of the solar system. This artist's conception shows the Stardust spacecraft about to fly through the dust and gas spewing from the nucleus of Comet Wild 2 in 2004. (NASA/JPL)

Guidepost

In Chapter 16 you began your study of planetary astronomy by pondering evidence about how our solar system formed. In the two chapters that followed you surveyed the planets and found more clues about the origin of the solar system. However, most traces of the planets' earliest histories have been erased by geological activity or other processes. Now you are ready to study smaller, better-preserved objects that can tell you more about the age of planet building.

Compared with planets, the comets and asteroids are unevolved objects, leftover planet construction "bricks." You will find them much as they were when they formed 4.6 billion years ago. The fragments of these objects that reach Earth can give you a close look at these ancient planetesimals. As you explore, you will find answers to four important questions:

Where do meteors and meteorites come from?

What are asteroids?

What are comets?

What happens when an asteroid or comet hits Earth?

As you finish this chapter, you will have acquired real insight into your place in nature. You live on the surface of a planet. There are other planets—are they inhabited too? That is the subject of the next chapter.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

When they shall cry "PEACE, PEACE"
then cometh sudden destruction!
COMET'S CHAOS?—
What Terrible events will the Comet bring?

FROM A PAMPHLET PREDICTING THE END OF THE WORLD BECAUSE OF THE APPEARANCE OF COMET KOHOUTEK IN 1973

YOU ARE NOT afraid of comets, of course; but, not long ago, people viewed them with terror. For example, in 1910, Comet Halley was spectacular. On the night of May 19, 1910, Earth actually passed through the tail of the comet—and millions of people panicked. The spectrographic discovery of cyanide gas in the tails of comets led many to believe that life on Earth would end. Householders in Chicago stuffed rags around doors and windows to keep out the gas, and bottled oxygen was sold out. Con artists in Texas sold comet pills and inhalers to ward off the noxious fumes. An Oklahoma newspaper reported (in what was apparently a hoax) that a religious sect tried to sacrifice a virgin to the comet.

Throughout history, bright comets have been seen as portents of doom. Even the more recent appearance of bright comets has generated predictions of the end of the world. Comet Kohoutek in 1973, Comet Halley in 1986, and Comet Hale–Bopp in 1997 caused concern among the superstitious. A bright comet moving slowly from night to night against the stellar background is so out of the ordinary (■ Figure 19-1) that you should not be surprised that it can generate some instinctive alarm.

In fact, comets are graceful and beautiful visitors to our skies. Astronomers think of comets as one type of messenger from the age of planet building. By studying comets, you can learn something about conditions in the solar nebula from which planets formed. Comets, however, are only the remains of the icy parts of the solar nebula; in contrast, asteroids are rocky debris, also left over from planet building. Because you cannot easily visit comets and asteroids, you can begin by learning about the fragments of those bodies that reach Earth—meteors and meteorites.

■ Figure 19-1

Comet McNaught swept through the inner solar system in 2007 and was a dramatic sight in the southern sky. Seen here from Big Swamp, South Australia, the comet was on its way back into deep space after making its closest approach to the sun ten days earlier. This comet originally came from the outskirts of the solar system on an orbit with a period of about 300,000 years. Gravitational perturbations by the planets during this passage caused its orbit to become hyperbolic, so it will never return but instead is leaving the solar system entirely to journey forever in interstellar space. (© John White Photos)

19-1 Meteoroids, Meteors, and Meteorites

You learned some things about meteorites in Chapter 16 when you studied the age of the solar system. There you saw that the solar system is filled with small particles called meteoroids. Some of them collide with Earth's atmosphere at speeds of 10 to 40 km/s. Friction with the air heats the meteoroids enough so that they glow, and you see them vaporize as streaks across the night sky, called meteors. If a meteoroid is big enough and strong enough, it can survive its plunge through the atmosphere and reach Earth's surface. Once the object strikes Earth's surface, it is called a meteorite. The largest of these objects can blast out craters on Earth's surface, but such impacts are rare. The vast majority of meteorites are too small to form craters. These meteorites fall all over Earth, and their value lies in what they can reveal about the origin of the planets.

Inside Meteorites

Meteorites can be divided into three broad categories. *Iron* meteorites are solid chunks of iron and nickel. *Stony* meteorites are silicate masses that resemble Earth rocks. *Stony-iron* meteorites are mixtures of iron and stone. These types of meteorites are illustrated in ■ Figure 19-2.

Iron meteorites are very dense. They often have dark, rusted surfaces and complicated shapes caused by their passage through the atmosphere. When they are sliced open, polished, and etched with acid, they reveal regular bands called **Widmanstätten patterns** (Figure 19-2). The patterns arise from crystals of nickel-iron alloys that have grown large, indicating that the meteorite cooled from a molten state no faster than a few degrees of temperature per million years. Learning later in this chapter how iron meteorites could have cooled so slowly will be a major step in understanding their history.

Stony meteorites called **chondrites** (pronounced *KON-drites*) have chemical compositions that closely resemble a cooled lump of matter from the sun with the helium and much of the hydrogen removed. Most types of chondrites contain **chondrules,** rounded bits of glassy rock ranging from microscopic to several millimeters across (Figure 19-2). Details of the origin of chondrules are unclear, but they appear to have formed in the young solar system as droplets of molten rock that cooled and hardened rapidly. The presence of chondrule particles inside chondrite meteorites indicates that those chondrites have never been melted since they formed, because melting would have destroyed the chondrules.

Some chondrites contain volatiles and organic (carbon) compounds, and may have formed in the presence of water. **Carbonaceous chondrites** are especially rich in volatile and

■ **Figure 19-2**

The three main types of meteorites—irons, stones, and stony-irons—have distinctive characteristics. (Lab photos courtesy of Russell Kempton, New England Meteoritical)

How Do We Know? 19-1

Enjoying the Natural World

Do you enjoy understanding things? You can admire a meteor shower as Mother Nature's fireworks, but your enjoyment is much greater once you begin to understand what causes meteors and why meteors in a shower follow a pattern. When you know that meteor showers help you understand the origin of Earth, the evening display of shooting stars is even more exciting. Science typically increases your enjoyment of the natural world by revealing the significance of things you might otherwise enjoy only in a casual way.

Everyone likes flowers, for example. But botanists know that flowers have evolved to attract insects and spread pollen. The bright colors signal to insects, and the shapes of the flowers provide little runways so the insect will find it easy to land. Some color patterns even guide the insect in like landing lights at an airport, and many flowers, such as orchids and snapdragons, force the insect to crawl inside in just the right way to exchange pollen and fertilize the flower. Nectar is bug bait. Once you begin to understand what flowers are for, a visit to a garden becomes not only an adventure of color and fragrance but also an adventure in meaning as well.

In addition, your understanding of one natural phenomenon can help you understand and enjoy related phenomena. For example, some flowers attract flies for pollination, and such blossoms smell like rotting meat. A few flowers depend on bats, and they open their blossoms at night. Flowers that depend on hummingbirds have long trumpet-shaped blossoms into which the hummingbirds' beaks just fit.

The more science tells you about nature, the more enjoyable the natural world is. You can enjoy a meteor shower as just a visual treat, but the more you know about it, the more interesting it becomes. The natural world is filled with meaning, and science, as a way of discovering and understanding that meaning, gives you new opportunities to enjoy the world around you.

The beauty of flowers becomes more interesting when you know the reasons for their shapes and colors. (M. Seeds)

organic substances, and also have chondrules. Other types of chondrites, in contrast, are poor in volatiles. The condensing solar nebula should have incorporated water plus other volatiles and organic compounds into solid particles as they formed. If that material had later been heated it would have lost the volatiles, and many of the organic compounds would have been destroyed. The special significance of carbonaceous chondrites meteorites is that they include some of the least modified material in our solar system.

Some stony meteorites contain no chondrules, and they are called **achondrites.** They also lack volatiles and appear to have been subjected to intense heat that melted chondrules and completely drove off the volatiles, leaving behind rock with compositions similar to Earth's lavas.

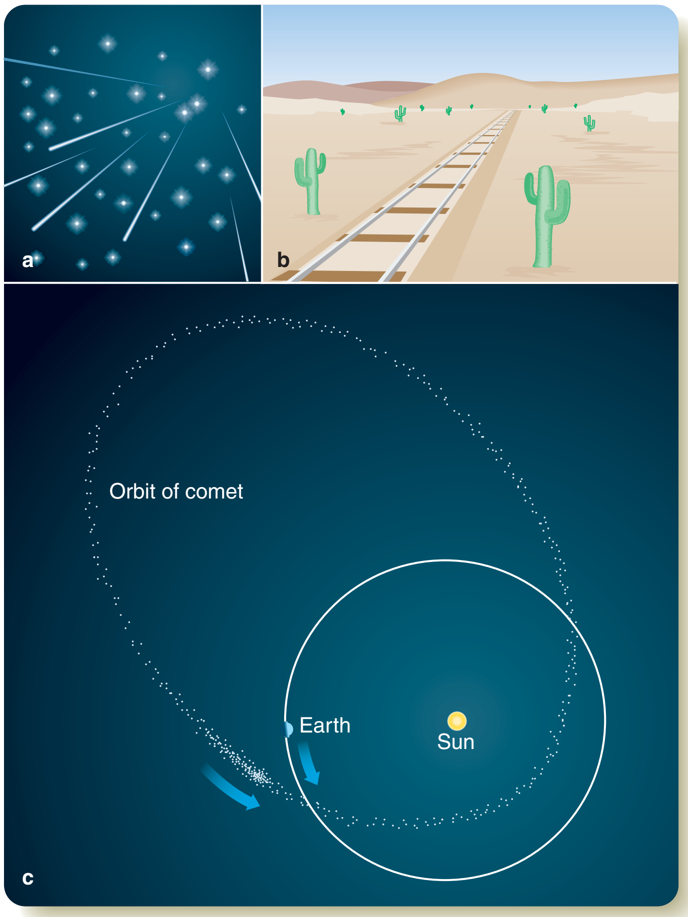
Stony-iron meteorites appear to have formed when a mixture of molten iron and rock cooled and solidified. This type of meteorite also contains no chondrules or volatiles.

The different types of meteorites have had a wide variety of histories. Some chondrites were heated (but not melted) after they formed, others somehow avoided being heated at all. Some achondrites seem like pieces of lava flows, whereas stony-iron and iron meteorites apparently were once deep inside the molten interiors of differentiated objects. Meteorites provide evidence that the early history of the solar system was complex.

The Origins of Meteors and Meteorites

On any clear, moonless night of the year you might see a few meteors per hour, but these usually would not be coming from the same direction and are not related to each other. On some special nights each year you can observe **meteor showers,** displays of meteors that are clearly related in a common origin. For example, the Perseid meteor shower occurs each year in August (Appendix A-12), and during the height of the shower you could see as many as 40 meteors an hour if you stretch out on a lawn chair at a dark site and watch the sky long enough for your eyes to adapt to the darkness (■ How Do We Know? 19-1).

During a meteor shower you will see meteors that are related in that they seem to come from a single spot on the sky. The Perseid shower, for example, appears to come from a spot in the constellation Perseus. These showers are seen when Earth passes near the orbit of a comet (see Section 19-3). The meteors in meteor showers are understood to be produced by dust and debris released from the icy nucleus of the comet. The meteors appear to come from a single place in the sky because they are particles traveling along parallel paths through space. Like railroad tracks extending from a point on the horizon, the meteors appear to approach from one point in space (■ Figure 19-3). The orbits of comets are filled with such debris. The Infrared As-



■ **Figure 19-3**

(a) Meteors in a meteor shower enter Earth's atmosphere along parallel paths, but perspective makes them appear to diverge from a single point in the sky. (b) Similarly, parallel railroad tracks appear to diverge from a point on the horizon. (c) Meteors in a shower are debris left behind as a comet's icy nucleus vaporizes. The rocky and metallic bits of matter spread along the comet's orbit. If Earth passes through such material, you can see a meteor shower.

tronomy Satellite (IRAS) space telescope detected sun-warmed dust glowing at far-infrared wavelengths scattered along the orbits of several comets.

Studies of meteors show that most of the meteors you see on any given night (whether or not there is a shower) are produced by tiny bits of debris from comets. These specks of matter are so small and so weak they vaporize completely in the atmosphere and never reach the ground, but from their motions astronomers can deduce their orbits, and those orbits match the orbits of comets. Thus, nearly all visible meteors come from comets.

It is a **Common Misconception** that a bright meteor disappearing behind a distant hill or line of trees probably landed just a mile or two away. This has triggered hilarious wild goose chases as police, fire companies, and TV crews try to find the impact site. Almost every meteor you see burns up 80 km (50 mi) or more above Earth's surface. Only rarely is an object strong enough to reach Earth's surface and be considered a meteorite, landing as much as 100 miles from where you are standing when you see it.

Evidence suggests that, in contrast to meteors, meteorites are fragments of parent bodies that were large enough to grow hot from radioactive decay or other processes. They then melted and differentiated to form iron-nickel cores and rocky mantles. The molten iron cores would have been well insulated by the thick rocky mantles, so that the iron would have cooled slowly enough to produce big crystals that result in Widmanstätten patterns. Collisions could break up such differentiated bodies and produce different kinds of meteorites (■ Figure 19-4). Iron meteorites appear to be fragments from the parent bodies' iron cores. Some stony meteorites that have been strongly heated appear instead to have come from the mantles or surfaces of such bodies. Stony-iron meteorites apparently come from boundaries between stony mantles and iron cores. In contrast, chondrites are probably fragments of smaller bodies that never melted, and carbonaceous chondrites may be from such unaltered bodies that formed especially far from the sun.

These hypotheses trace the origin of meteorites to planetesimal-like parent bodies, but they leave you with a puzzle. The small meteoroids now flying through the solar system cannot have existed in their present form since the formation of the solar system because they would have been swept up by the planets in a billion years or less. They could not have survived for 4.6 billion years.

In fact, when astronomers study the orbits of meteorites actually seen to fall, the orbits lead back into the asteroid belt. Thus astronomers have good evidence that the meteorites now in museums all over the world must have been produced by asteroid collisions within the last billion years. Nearly all meteors are pieces of comets, but meteorites are pieces broken off asteroids.

◄ SCIENTIFIC ARGUMENT ►

How can you say that meteors come from comets, but meteorites come from asteroids?

First, remember the distinction between meteors and meteorites. A meteor is the streak of light seen in the sky when a particle from space is heated by friction with Earth's atmosphere. A meteorite is a piece of space material that actually reaches the ground.

The distinction between comet and asteroid sources must take into account two very strong effects that prevent you from finding meteorites that originated in comets. First, evidence is that cometary particles are physically weak, and they vaporize in Earth's atmosphere easily. Very few ever reach the ground, and you are unlikely to find them. Second, even if a cometary particle reached the ground, it would be so fragile that it would weather away rapidly, and, again, you would be unlikely to find it. Asteroidal particles, however, are made from rock and metal and are stronger. They are more likely to survive their plunge through the atmosphere and more likely to survive erosion on the ground. Every known meteorite is from the asteroids—not a single meteorite is known to be cometary. In contrast, meteor tracks show that most meteors you see come from comets, and very few are coming from the asteroid belt.

Now build a new argument. **What evidence suggests that meteorites were once part of larger bodies broken up by impacts?**

◄ ►

The Origin of Meteorites

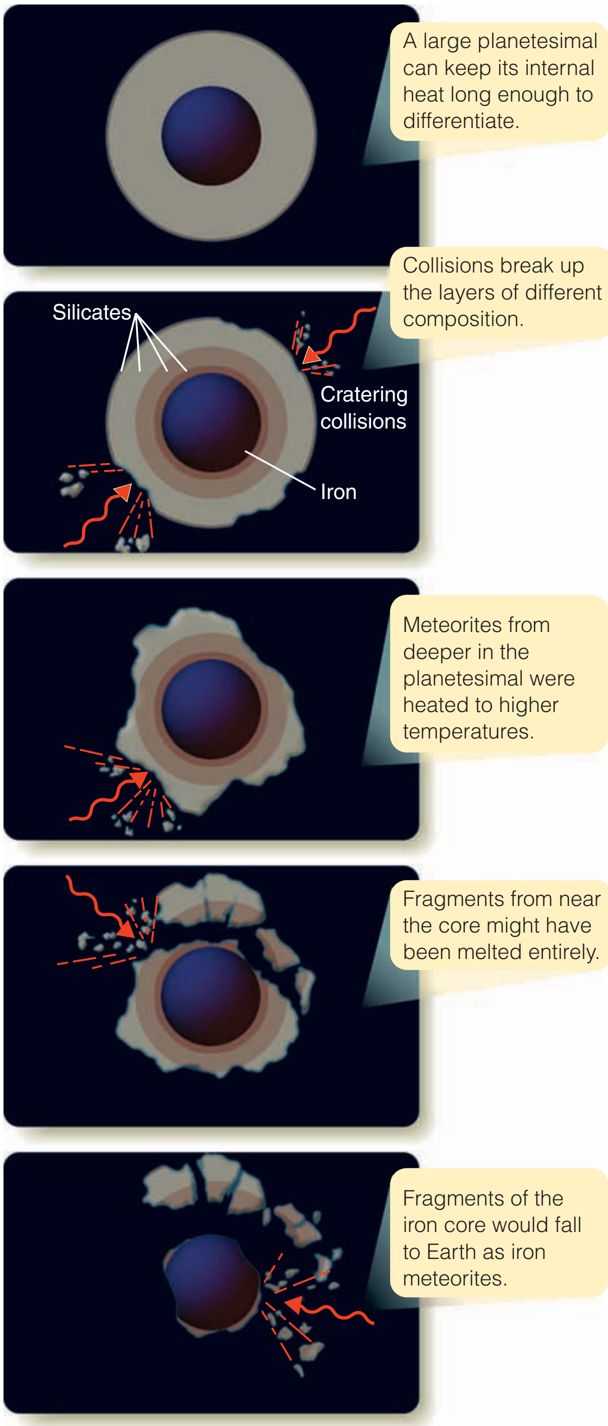


Figure 19-4

Planetesimals that formed when the solar system was forming may have melted and separated into layers of different density and composition. The fragmentation of such a body could produce different types of meteorites. (Adapted from a diagram by Clark Chapman)

19-2 Asteroids

Space pirates lurk in the asteroid belt in science fiction, but astronomers have found that there isn't much in the asteroid belt for a pirate to stand on. Hundreds of thousands of asteroids are known, but most are quite small. Movies and TV have created a **Common Misconception** that flying through an asteroid belt is a hair-raising plunge requiring constant dodging left and right. The asteroid belt between Mars and Jupiter is actually mostly empty space. In fact, if you were standing on an asteroid, it would be many months or years between sightings of other asteroids.

Properties of Asteroids

Asteroids are distant objects too small to study in detail with Earth-based telescopes. Astronomers nevertheless have learned a surprising amount about these little worlds, and spacecraft plus space telescopes have provided a few close-ups.

Study ■ **Observations of Asteroids** on pages 414–415 and notice four important points:

1. Most asteroids are irregular in shape and battered by impact cratering. Many asteroids seem to be rubble piles of broken fragments.
2. Some asteroids are double objects or have small moons in orbit around them. This is further evidence of collisions among the asteroids.
3. A few larger asteroids show signs of geological activity that happened on their surfaces when those asteroids were young.
4. Asteroids can be classified by their albedo, color, and spectra to reveal clues to their compositions. This also allows them to be compared to meteorites in labs on Earth.

Before you continue, you should note that not all asteroids lie in the asteroid belt. A large number of asteroids, perhaps as many as half the number in the main belt, travel in Jupiter's orbit ahead of and behind the planet. These are called **Trojan asteroids** because the largest ones are named after heroes of the Trojan War. Also, a few thousand objects larger than 1 km follow orbits that cross Earth's orbit. A number of searches are under way to locate these **Near-Earth Objects** (**NEOs**). For example, LONEOS (Lowell Observatory Near Earth Object Search) is searching the entire sky once a month, and these searches should be able to locate all of the largest NEOs by 2010. Astronomers are searching for these asteroids to understand asteroids better but also because such asteroids can collide with Earth. Although such collisions occur very rarely, a single impact could cause planetwide devastation. You will learn about such impacts on Earth later in this chapter.

The Origin of the Asteroids

An old hypothesis proposed that asteroids are the remains of a planet that exploded. Planet-shattering death rays may make for exciting science-fiction movies, but in reality planets do not explode. The gravitational field of a planet holds the mass so tightly that completely disrupting the planet would take tremendous energy. In addition, the present-day total mass of the asteroids is only about one-twentieth the mass of the moon, hardly enough to be the remains of a planet.

Astronomers have evidence that the asteroids are the remains of material lying 2 to 4 AU from the sun that was unable to form a planet because of the gravitational influence of Jupiter, the next planet outward. Over the 4.6-billion-year history of the solar system, most of the objects originally in the asteroid belt collided and fragmented and also were perturbed by the gravity of Jupiter and other planets into orbits that collided with planets or caused some asteroids to leave the solar system. The present-day asteroids are a very minor remnant of the original population of planetesimals in that zone, mostly fragmented by collisions with one another.

Rocky **S-type** asteroids are believed to be fragments from the crust and mantle, and **M-types** from the metallic cores, of differentiated asteroids. **C-type** asteroids, which appear to have plentiful carbon compounds, are more common in the outer asteroid belt. It is cooler there, and the condensation sequence (see Chapter 16) predicts that carbonaceous material would form there more easily than in the inner belt.

As you saw in the case of Vesta, a few large asteroids may have been geologically active with lava flowing on their surfaces when they were young. Perhaps they incorporated short-lived radioactive elements such as aluminum-26. Those radioactive elements were produced by a supernova explosion that might also have been the trigger for the formation of the sun and planets while seeding the young solar system with its nucleosynthesis products (see Chapters 9 and 10). Not all large asteroids have been active. Ceres, 900 km in diameter, is almost twice as big as Vesta, but it shows no spectroscopic sign of past activity and evidently has an ice-rich mantle.

Although there are still mysteries to solve, you can understand the compositions of the meteorites. They are fragments of planetesimals, some of which developed molten cores, differentiated, may have had lava flows on their surfaces, and then cooled slowly. The largest asteroids astronomers see today may be nearly unbroken planetesimals, but the rest are fragments produced by 4.6 billion years of collisions.

◄ SCIENTIFIC ARGUMENT ►

What is the evidence that asteroids have been fragmented?

First, your argument might note that the solar nebula theory predicts that planetesimals collided and either stuck together or fragmented. This is suggestive, but it is not evidence. A theory can never be used as evidence to support some other theory or hypothesis. Evidence means observations or the results of experiments, so your argument must cite observations. The spacecraft photographs of asteroids show irregularly shaped little worlds heavily scarred by impact craters. In fact, radar images show what may be pairs of bodies split apart but still in contact, and images of Ida reveal a small satellite, Dactyl. Other asteroids with moons have been found. These double asteroids and asteroids with moons probably reveal the results of fragmenting collisions between asteroids. Furthermore, meteorites appear to have come from the asteroid belt astronomically recently, so fragmentation must be a continuing process there.

Now build an argument to combine what you know of meteorites with your experience with asteroids. **What evidence could you cite to show what the first planetesimals were like?**

◄ ►

19-3 Comets

Of all the fossils left behind by the solar nebula, comets are the most beautiful. Asteroids are dark, rocky worlds, and meteors are flitting specks of fire in Earth's atmosphere, but comets move with the grace and beauty of a great ship at sea (■ Figure 19-5).

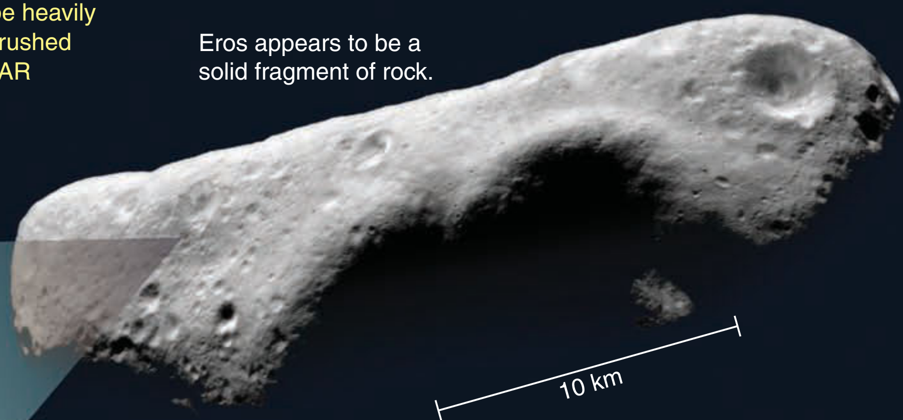
■ **Figure 19-5**

Comet Hale–Bopp was very bright in the sky in 1997. A comet can remain visible in the sky for weeks as it sweeps along its orbit through the inner solar system. (Dean Ketelsen)

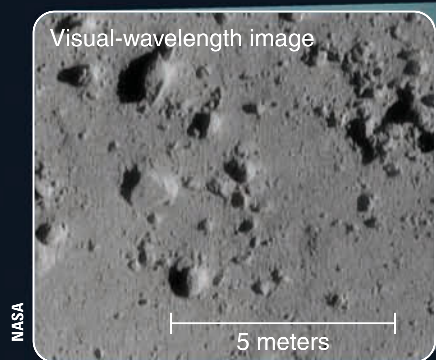
Observations of Asteroids

1 Seen from Earth, asteroids look like faint points of light moving in front of distant stars. Not many years ago they were known mostly for drifting slowly through the field of view and spoiling long time exposures. Some astronomers referred to them as "the vermin of the sky." Spacecraft have now visited asteroids, and the images radioed back to Earth show that the asteroids are mostly small, gray, irregular worlds heavily cratered by impacts.

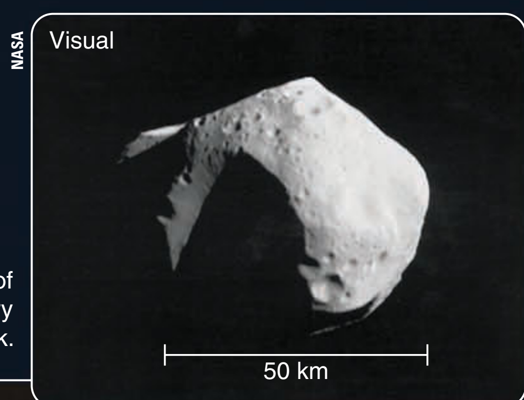
The Near Earth Asteroid Rendezvous (NEAR) spacecraft visited the asteroid Eros in 2000 and found it to be heavily cratered by collisions and covered by a layer of crushed rock ranging from dust to large boulders. The NEAR spacecraft eventually landed on Eros.



Eros appears to be a solid fragment of rock.

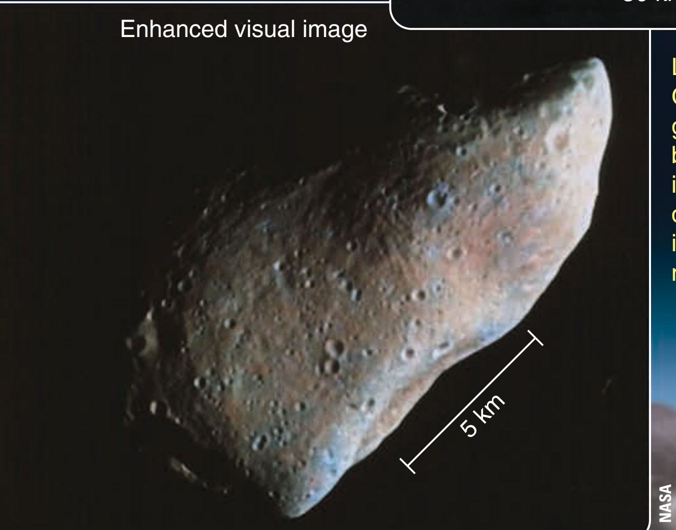


Most asteroids are too small for their gravity to pull them into a spherical shape. Impacts break them into irregularly shaped fragments.



The surface of Mathilde is very dark rock.

1a The mass of an asteroid can be found from its gravitational influence on passing spacecraft. Its volume can be measured using images made from a range of perspectives. The density is mass divided by volume. Mathilde, at left, has such a low density that it cannot be solid rock. Like many asteroids, Mathilde may be a rubble pile of broken fragments with large empty spaces between fragments.



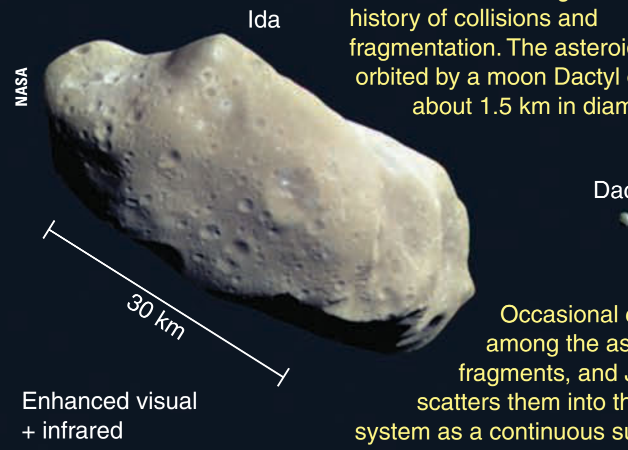
Like most asteroids, Gaspra would look gray to your eyes; but, in this enhanced image at left, color differences probably indicate difference in mineralogy.

If you walked across the surface of an irregularly shaped asteroid such as Eros, you would find gravity very weak; and in many places, it would not be perpendicular to the surface.

NASA

2

Asteroids that pass near Earth can be imaged by radar. The asteroid Toutatis is revealed to be a double object—two objects orbiting close to each other or actually in contact.



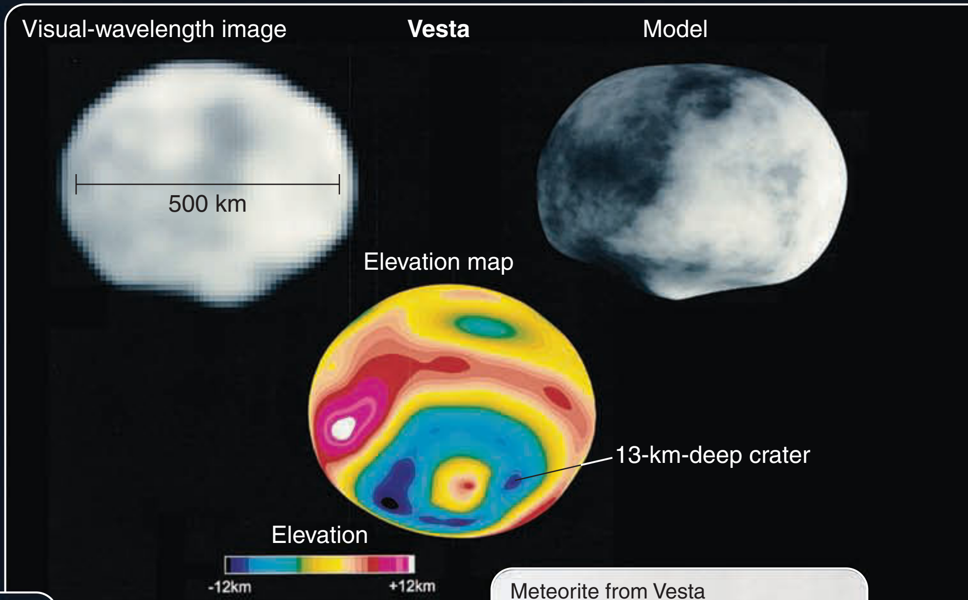
Double asteroids are more common than was once thought, reflecting a history of collisions and fragmentation. The asteroid Ida is orbited by a moon Dactyl only about 1.5 km in diameter.

Occasional collisions among the asteroids release fragments, and Jupiter's gravity scatters them into the inner solar system as a continuous supply of meteorites.

3

The large asteroid Vesta, as shown at right, provides evidence that some asteroids once had geological activity. No spacecraft has visited it, but its spectrum resembles that of solidified lava. Images made by the Hubble Space Telescope allow the creation of a model of its shape. It has a huge crater at its south pole. A family of small asteroids is evidently composed of fragments from Vesta, and a certain class of meteorites, spectroscopically identical to Vesta, are believed to be fragments from the asteroid. The meteorites appear to be solidified basalt.

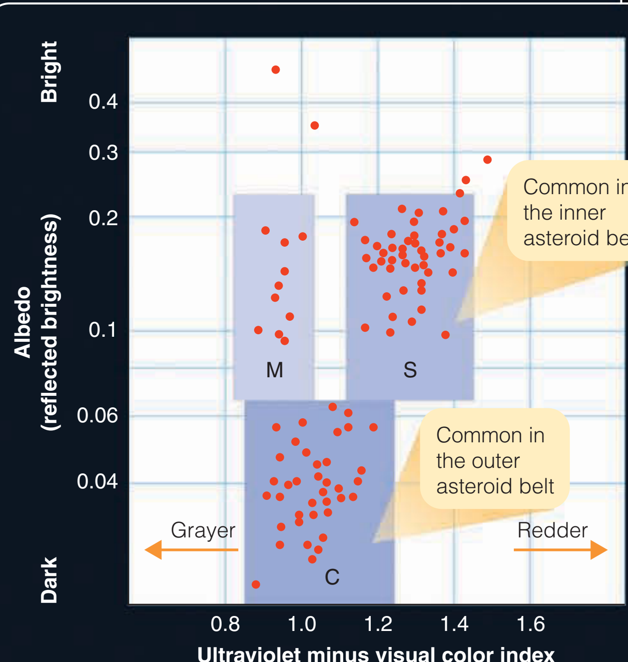
NASA



Courtesy of Russell Kemton, New England Meteoritical

3a

Vesta appears to have had internal heat at some point in its history, perhaps due to the decay of radioactive minerals. Lava flows have covered at least some of its surface.



4

Although asteroids would look gray to your eyes, they can be classified according to their albedos (reflected brightness) and spectroscopic colors. As shown at left, S-types are brighter and tend to be reddish. They are the most common kind of asteroid and appear to be the source of the most common chondrites.

M-type asteroids are not too dark but are also not very red. They may be mostly iron-nickel alloys.

C-type asteroids are as dark as lumps of sooty coal and appear to be carbonaceous.

Properties of Comets

As always, you can begin your study of a new kind of object by summarizing its observational properties. What do comets look like, and how do they behave?

Study ■ **Observations of Comets** on pages 418–419 and notice three important properties of comets and three new terms:

1. Comets have two types of tails shaped by the solar wind and solar radiation, *type I (gas)* and *type II (dust)*. The two types of tails show that the nucleus contains water ice and other frozen compounds plus rocky material, mostly in the form of dust. Comet tails point away from the sun no matter in what direction the comet is moving. Comet heads are called *coma.*

2. Dust in comets not only produces very visible dust tails but spreads throughout the solar system.

3. There is evidence that some comet nuclei are fragile and can break into pieces. An astronomer recently commented that the nuclei of comets can be "as fragile as the meringue in lemon-meringue pie."

Go to **academic.cengage.com/astronomy/seeds** to see Astronomy Exercise "Comets."

The Geology of Comet Nuclei

The nuclei of comets are quite small and cannot be studied in detail with Earth-based telescopes. Nevertheless, astronomers are beginning to understand the geology of these interesting worlds.

Comet nuclei contain ices of water and other volatile compounds such as carbon dioxide, carbon monoxide, methane, ammonia, and so on. These ices are the kinds of compounds that should have condensed from the outer solar nebula, which makes astronomers think that comets are ancient samples of the gases and dust from which the planets formed.

Five spacecraft flew past the nucleus of Comet Halley when it visited the inner solar system in 1985 and 1986. Since then spacecraft have visited the nuclei of Comet Borrelly, Comet Wild 2 (pronounced *Vilt-two*), and Comet Tempel 1. Images show that comet nuclei are irregular in shape and very dark, with jets of gas and dust spewing from active regions on the nuclei (■ Figure 19-6). In general, these nuclei are darker than a lump of coal, which suggests their composition is similar to the carbon-rich carbonaceous chondrite meteorites. Interestingly, samples of dust from the tail of Comet Wild 2 returned to Earth by the *Stardust* probe include chondrule-like particles, another similarity to carbonaceous chondrite meteorites. This is unexpected evidence that when comets formed they somehow included material from very hot, as well as cold, parts of the solar nebula. Comets, like asteroids, show that the early history of the solar system was complex.

From the gravitational influence of a nucleus on a passing spacecraft, astronomers can find the mass and density of the nucleus. Comet nuclei appear to have densities of 0.1 to 0.25 g/cm^3, much less than the density of ice. The shapes and low densities of comet nuclei suggest that they are not solid objects. They have been described as dirty snowballs or icy mudballs, but that is misleading. The evidence leads astronomers to conclude that most comet nuclei are not solid balls of ice but must be fluffy mixtures of ices and dust with significant amounts of empty space. On the other hand, one comet nucleus, that of Comet Wild 2, has cliffs, pinnacles, and other features that show the material has enough strength to stand against the weak gravity of the comet.

Photographs of the comas of comets often show jets springing from the nucleus and being swept back by the pressure of sunlight and by the solar wind to form one or more tails (Figure 19-6). The jets originate from active regions that may be similar to volcanic faults or vents. As the rotation of a comet nucleus carries an active region into sunlight, it begins emitting gas and dust, and as the active region rotates into darkness it shuts down. Although comet nuclei seem to have porous crusts of dark material, the ice and rock are not uniformly mixed through the interior. Breaks in the crust can expose pockets of highly volatile ices and cause sudden bursts of gas production.

The nuclei of comets are only a few kilometers in diameter. Each time they round the sun, they lose many millions of tons of ices, so the nuclei shrink until there is nothing left but dust and rock.

The Origin of Comets

A comet may last only 100 to 1000 orbits around the sun before it has lost all its ices. The comets seen in our skies can't have survived 4.6 billion years since the formation of the solar system, so there must be a continuous supply of new comets. Where do they come from?

Family relationships among the comets provide clues to their origin. Most comets have long, elliptical orbits with periods greater than 200 years. These are known as long-period comets. Their orbits are randomly inclined at all angles to the main plane of the solar system. Approximately equal numbers of long-period comets circle the sun in the retrograde direction as in the prograde direction (the direction in which the planets orbit).

In contrast, about 100 or so of the 600 well-studied comets have orbits with periods less than 200 years. These short-period comets follow orbits that lie within 30° of the plane of the solar system, and most revolve around the sun in the prograde direction.

In the 1950s, Dutch astronomer Jan Oort proposed that the long-period comets are objects that fall into the inner solar system from a spherical cloud of several trillion icy bodies evidently extending from about 10,000 to as much as 100,000 AU from

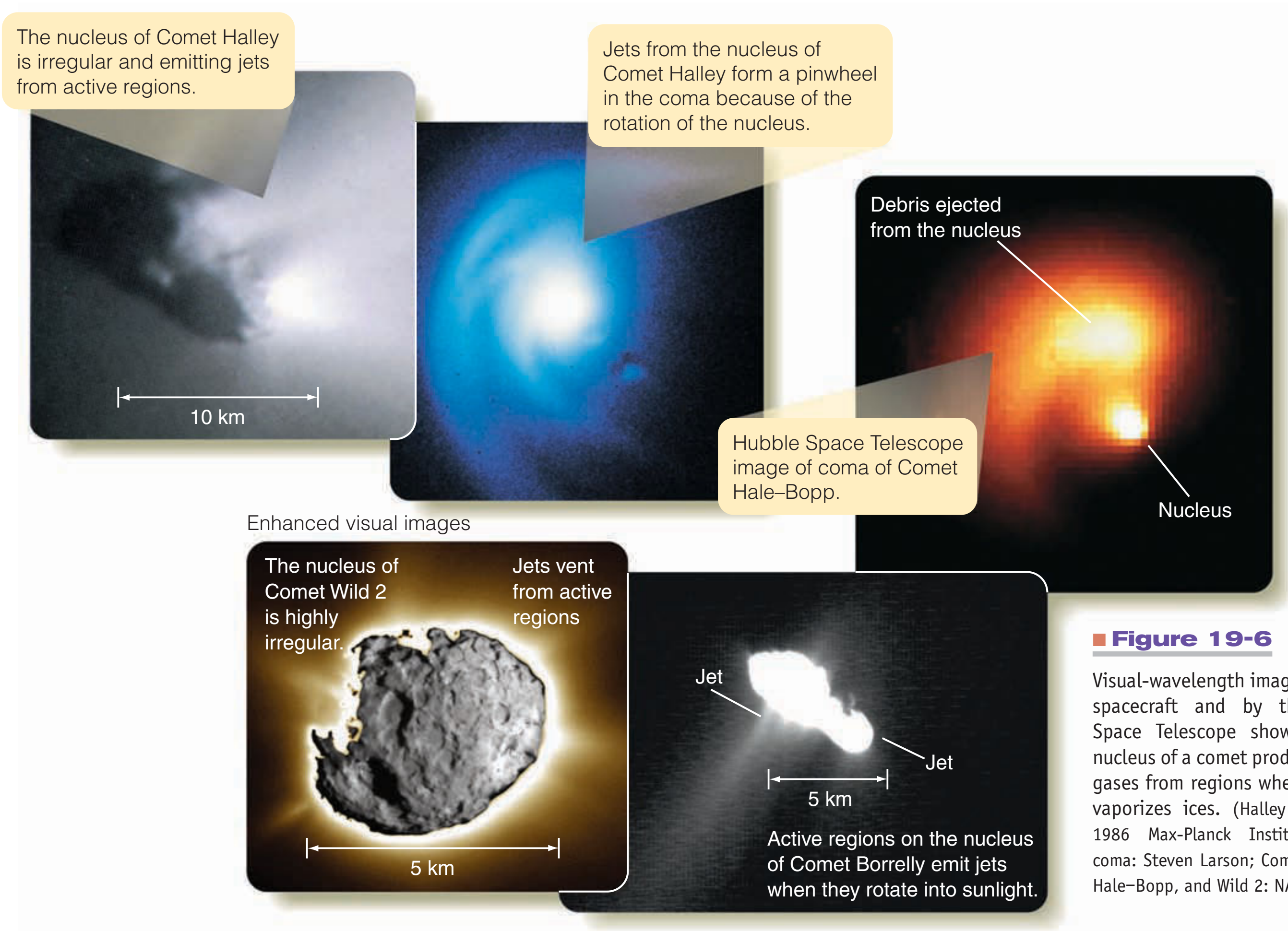


Figure 19-6

Visual-wavelength images made by spacecraft and by the Hubble Space Telescope show how the nucleus of a comet produces jets of gases from regions where sunlight vaporizes ices. (Halley nucleus: © 1986 Max-Planck Institute; Halley coma: Steven Larson; Comets Borrelly, Hale–Bopp, and Wild 2: NASA)

the sun, later named the **Oort cloud.** Far from the sun, they are very cold, lack comas and tails, and are invisible. The gravitational influence of occasional passing stars could perturb a few of these objects toward the inner solar system, where the heat of the sun warms their ices and transforms them into comets. Because the Oort cloud is spherical, these long-period comets arrive in our neighborhood from random directions.

Some of the short-period comets, including Comet Halley, whose orbit is retrograde, appear to have originated in the Oort cloud and probably had their original long elliptical orbits altered by a close encounter with Jupiter or another planet. Many of the short-period comets, however, cannot have begun in the Oort cloud. Detailed calculations show that interactions with a planet can't easily redirect objects from the Oort cloud into the orbits that those comets occupy. Rather, those comets must have originated in the Kuiper belt (see Chapters 16 and 18). The Kuiper belt objects that have been found are much bigger than ordinary comet nuclei, but astronomers estimate the Kuiper belt holds as many as 100 million smaller objects. When a Kuiper belt object is perturbed into the inner solar system, it can interact with planets and have its orbit shortened into that of a short-period comet.

The Kuiper belt objects evidently formed as icy planetesimals in the outer solar nebula not much farther from the sun than the outer planets. The objects in the Oort cloud lie much farther from the sun, but they can't have formed out there. The solar nebula would have had too low a density at such great distances. Also, you would expect objects that formed from the nebula to be confined to a disk and not be distributed spherically. Astronomers think the evidence points toward the region among the Jovian planets as the original source of objects now in the Oort cloud. As the Jovian planets grew more massive, they swept up some of the local planetesimals but ejected others to form the Oort cloud.

◄ SCIENTIFIC ARGUMENT ►

How do comets help explain the formation of the planets?

This argument must refer to the solar nebula hypothesis. The planetesimals that formed in the inner solar nebula were warm and could not incorporate much ice. The asteroids may be the last remains of such rocky bodies. In the outer solar nebula, the planetesimals contained large amounts of ices. Many were destroyed when they accreted together to make the Jovian planets, but some survived. The icy bodies of the Oort cloud and the Kuiper belt may be the solar system's last surviving icy planetesimals. When those icy objects have their orbits perturbed by the gravity of the planets or

Observations of Comets

1 A **type I** or **gas tail** is produced by ionized gas carried away from the nucleus by the solar wind. The spectrum of a gas tail is an emission spectrum. The atoms are ionized by the ultraviolet light in sunlight. The wisps and kinks in gas tails are produced by the magnetic field embedded in the solar wind.

Spectra of gas tails reveal atoms and ions such as H_2O, CO_2, CO, H, OH, O, S, C, and so on. These are released by the vaporizing ices or produced by the breakdown of those molecules. Some gases, such as hydrogen cyanide (HCN), must be formed by chemical reactions.

1a A **type II** or **dust tail** is produced by dust that was contained in the the vaporizing ices of the nucleus. The dust is pushed gently outward by the pressure of sunlight, and it reflects an absorption spectrum, the spectrum of sunlight. The dust is not affected by the magnetic field of the solar wind, so dust tails are more uniform than gas tails. Dust tails are often curved because the dust particles follow their individual orbits around the sun once they leave the nucleus. Because of the forces acting on them, both gas and dust tails extend away from the sun.

1b The nucleus of a comet (not visible here) is a small, fragile lump of porous rock containing ices of water, carbon dioxide, ammonia, and so on. Comet nuclei can be 10 to 100 km in diameter.

The **coma** of a comet is the cloud of gas and dust that surrounds the nucleus. It can be over 1,000,000 km in diameter, bigger than the sun.

1c Comet Mrkos in 1957 shows how the gas tail can change from night to night due to changes in the magnetic field in the solar wind.

Visual-wavelength images

CENGAGENOW™

Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see Active Figure "Build a Comet." See how energy from the sun shapes a comet.

2 As the ices in a comet nucleus vaporize, they release dust particles that not only form the dust tail, but also spread throughout the solar system.

The Deep Impact spacecraft released an instrumented probe into the path of Comet Temple 1. When the comet slammed into the probe at 10.2 km/s as shown at right huge amounts of gas and dust were released. From the results, scientists conclude that the nucleus of the comet is rich in dust finer than the particles of talcum powder. The nucleus is marked by craters, but it is not solid rock. It is about the density of fresh fallen snow.

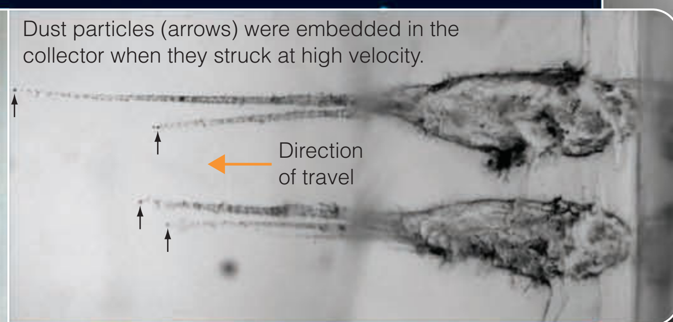

NASA/JPL-Caltech/UMD

Only seconds before impact, craters are visible on the dark surface.

Images of Comet Temple 1 from the flyby probe 13 seconds after the impact probe hit. Gas and dust are thrown out of the impact crater.


A microscopic mineral crystal from Comet Wild 2

NASA



Dust particles (arrows) were embedded in the collector when they struck at high velocity.

JPL/ NASA

2a The Stardust spacecraft flew past the nucleus of Comet Wild 2 and collected dust particles (as shown above) in an exposed target that was later parachuted back to Earth. The dust particles hit the collector at high velocity and became embedded, but they can be extracted for study.

Some of the collected dust is made of high temperature minerals that could only have formed near the sun. This suggests that material from the inner solar nebula was mixed outward and became part of the forming comets in the outer solar system. Other minerals found include olivine, a very common mineral but not one that scientists expected to find in a comet.

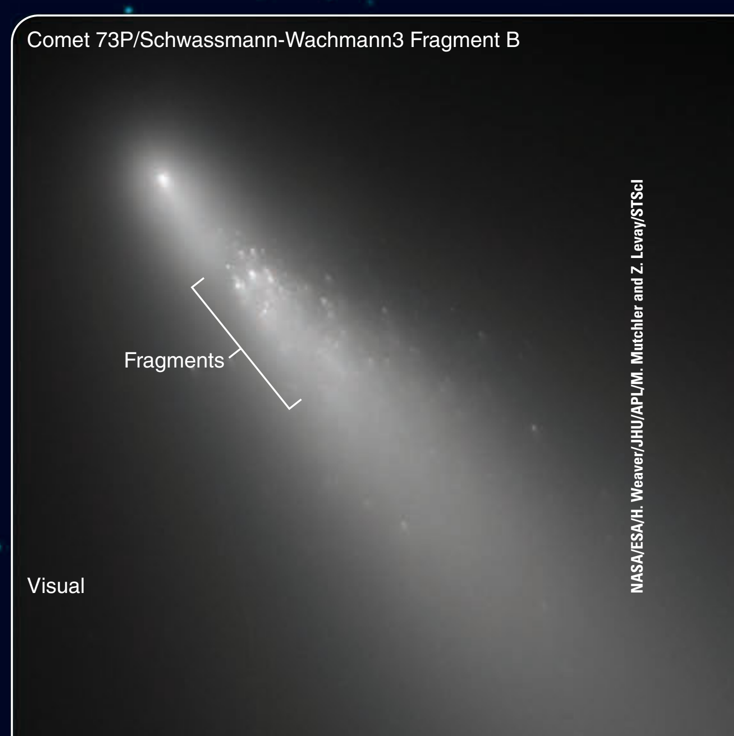

NASA

This dust particle was collected by spacecraft above Earth's atmosphere. It is almost certainly from a comet.

3 The nuclei of comets are not strong and can break up. In 2006, Comet Schwassmann-Wachmann 3 broke into a number of fragments which themselves fragmented. Fragment B is shown at the right breaking into smaller pieces. The gas and dust released by the break up made the comet fragments bright in the night sky and some were visible with binoculars. As its ices vaporize and its dust spreads, the comet may totally disintegrate and leave nothing but a stream of debris along its previous orbit.

Comets most often break up as they pass close to the sun or close to a massive planet like Jupiter. Comet LINEAR broke up in 2000 as it passed by the sun. The comet that hit Jupiter in 1994 was first ripped to pieces by tidal stresses from Jupiter's gravity. Comets can also fragment far from planets, perhaps because of the vaporization of critical areas of ice.

Comet 73P/Schwassmann-Wachmann3 Fragment B



NASA/ESA/H. Weaver/JHU/APL/M. Mutchler and Z. Levay/STScI passing stars, some are redirected into the inner solar system where you see them as comets. The gases released by comets indicate that they are rich in volatile materials such as water and carbon dioxide. These are the ices you would expect to find in the icy planetesimals. Furthermore, comets are rich in dust with rocklike chemical composition, and the planetesimals must have included large amounts of such dust frozen into the ices when they formed. Thus, the nuclei of comets seem to be frozen samples of the ancient solar nebula.

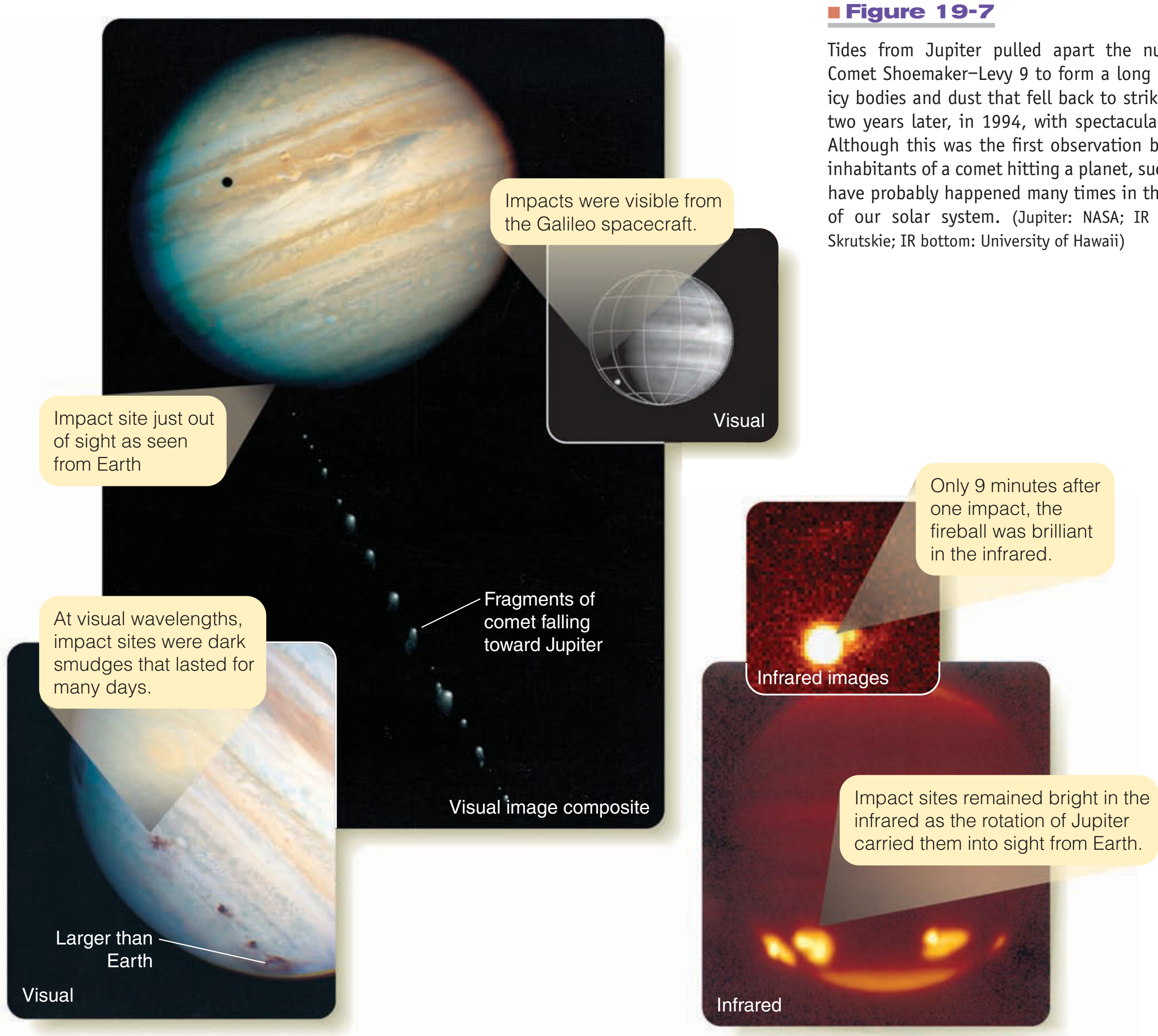
Nearly all of the mass of a comet is in the nucleus, but the light you see comes from the coma and the tail. Build a new argument to discuss observations. **What do spectra of comets tell you about the process that converts dirty ice into a comet?**

◀ ▶

19-4 Impacts on Earth

For centuries, superstitious people have associated comets with doom, which seems silly. Comets are graceful visitors from the icy fringes of the solar system. But comets and asteroids do hit planets now and then, so you might wonder just how dangerous such impacts would be.

Astronomers have good reason to believe that comets, very large meteorites, and even asteroids can hit planets. Earthlings watched in awe in 1994 as fragments from the nucleus of comet Shoemaker–Levy 9 slammed into Jupiter and produced impacts equaling millions of megatons of TNT (■ Figure 19-7). Also,



■ Figure 19-7

Tides from Jupiter pulled apart the nucleus of Comet Shoemaker–Levy 9 to form a long strand of icy bodies and dust that fell back to strike Jupiter two years later, in 1994, with spectacular results. Although this was the first observation by Earth's inhabitants of a comet hitting a planet, such events have probably happened many times in the history of our solar system. (Jupiter: NASA; IR top: Mike Skrutskie; IR bottom: University of Hawaii) astronomers have found chains of craters on solar system objects that seem to have been formed by fragmented comets (■ Figure 19-8).

Meteorites hit Earth every day, and occasionally a large one can form a crater (■ Figure 19-9). Earth is marked by about 150 known meteorite craters. No one has ever seen an asteroid hit a planet, but there are some very big craters in the solar system that show what can happen. A large impact on Earth could have devastating consequences.

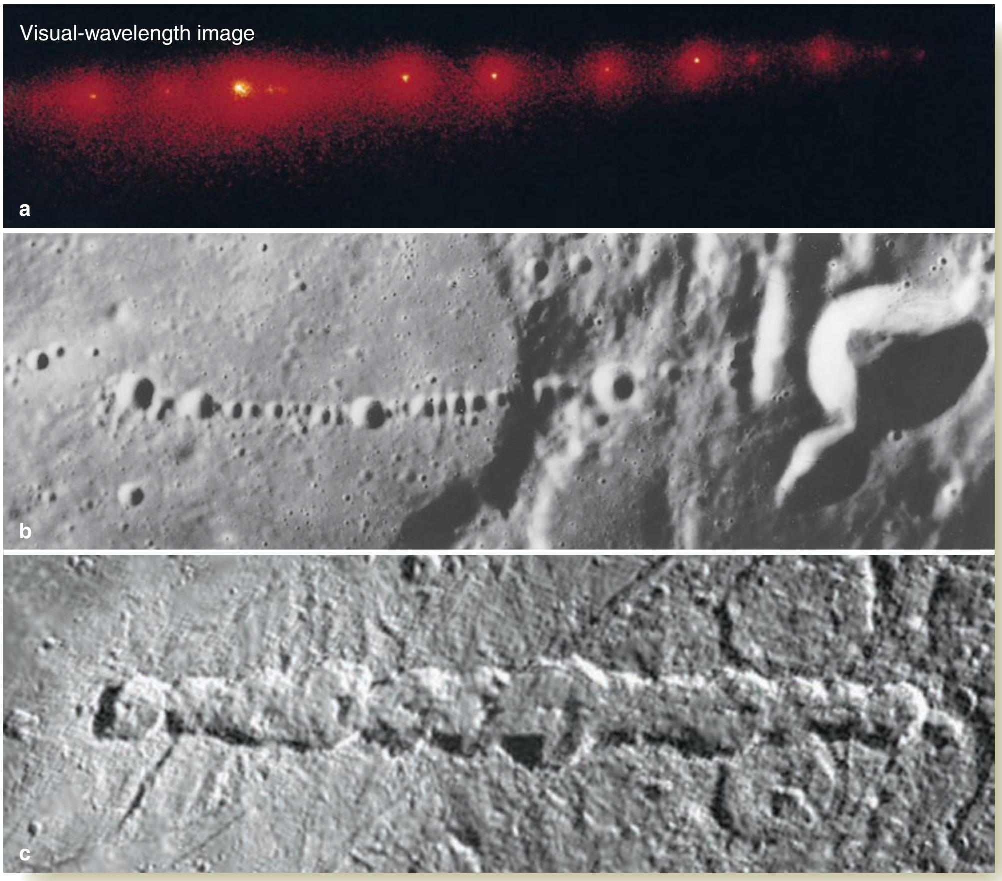
Sixty-five million years ago, at the end of the Cretaceous period, over 75 percent of the species on Earth, including the dinosaurs, went extinct. Scientists have found a thin layer of clay all over the world that was laid down at that time, and it is rich in the element iridium—common in meteorites but rare in Earth's crust. This suggests that an impact occurred that was large enough to have altered Earth's climate and caused the worldwide extinction.

Mathematical models combined with observations create a plausible scenario of a major impact on Earth. Of course, creatures living near the site of the impact would die in the initial shock, but then things would get bad elsewhere. An impact at sea would create tsunamis (tidal waves) many hundreds of meters high that would sweep around the world, devastating regions far inland from coasts. On land or sea, a major impact would eject huge amounts of pulverized rock high above the atmosphere. As this material fell back, Earth's atmosphere would be turned into a glowing oven of red-hot meteorites streaming through the air, and the heat would trigger massive forest fires around the world. Soot from such fires has been found in the layers of clay laid down at the end of the Cretaceous period. Once the firestorms cooled, the remaining dust in the atmosphere would block sunlight and produce deep darkness for a year or more, killing off most plant life. At the same time, if the impact site was at or near limestone deposits, large amounts of carbon dioxide could be released into the atmosphere and produce intense acid rain.

Geologists have located a crater at least 180 km (110 mi) in diameter centered near the village of **Chicxulub** (pronounced *cheek-shoe-lube*) in the northern Yucatán region of Mexico (■ Figure 19-10). Although the crater is now completely covered by sediments, mineral samples show that it contains shocked quartz typical of impact sites and that it is exactly the right age. The impact of an object 10 to 14 km in diameter formed the crater about 65 million years ago, just when the dinosaurs and many other species died out. Most Earth scientists now conclude that this is the scar of the impact that ended the Cretaceous period.

There are a number of major extinctions in the fossil record, and at least some of these were probably caused by large impacts. Large asteroid impacts on Earth happen many millions of years apart, but they continue to happen. In mid-March 1998, newspaper headlines announced, "Mile-Wide Asteroid to Hit Earth in October 2028." The news media did not emphasize the uncertainty in the asteroid's orbit. Within days, astronomers found more images of the asteroid on old photographic plates, recalculated the orbit adding the new data, and concluded that the asteroid, known as $1997XF_{11}$, would miss Earth by 600,000 miles. There will be no impact by that asteroid in 2028, but there are plenty more asteroids in Earth-crossing orbits that haven't been discovered. It is just a matter of time.

Go to **academic.cengage.com/astronomy/seeds** to see the Astronomy Exercise "Cratering."

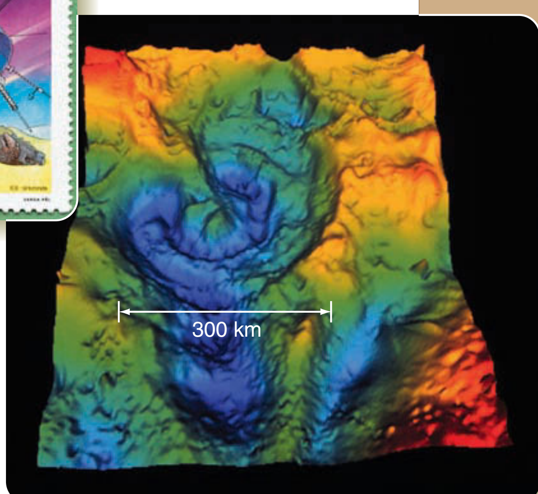


■ Figure 19-8

(a) Fragments of Comet Shoemaker–Levy 9 headed toward impact on Jupiter. (b) A 40-km-long crater chain on Earth's moon and (c) a 140-km-long crater chain on Jupiter's moon Callisto were apparently formed by the impact of fragmented comet nuclei similar to Comet Shoemaker–Levy 9. (NASA)

■ **Figure 19-9**

(a) The Barringer Meteorite Crater (near Flagstaff, Arizona) is nearly a mile in diameter and was formed about 50,000 years ago by the impact of an iron meteorite, estimated to have been roughly 90 meters in diameter. It hit with energy equivalent to that of a 3-megaton hydrogen bomb. Notice the raised and deformed rock layers all around the crater. The brick museum building visible on the far rim at right provides some idea of scale. (M. A. Seeds) (b) Like all larger-impact features, the Barringer Meteorite Crater has a raised rim and scattered ejecta. (USGS) **Animated!**



■ **Figure 19-10**

The theory that the impact of one or more comets altered Earth's climate and drove dinosaurs to extinction has become so popular it appeared on this Hungarian stamp. The spacecraft shown (ICE) flew through the tail of a comet in 1985. Note the dead dinosaurs in the background. The giant impact scar buried in Earth's crust near the village of Chicxulub in the northern Yucatán peninsula was formed about 65 million years ago by the impact of a large asteroid or comet. This gravity map shows the extent of the crater hidden below limestone deposited long after the impact. (Virgil L. Sharpton, University of Alaska, Fairbanks)

What Are We? Sitting Ducks

Human civilization is spread out over Earth's surface and exposed to anything that falls out of the sky. Meteorites, asteroids, and comets bombard Earth, producing impacts that vary from dust settling gently on rooftops to disasters capable of destroying all life. In this case, the scientific evidence is conclusive and highly unwelcome.

Statistically we are quite safe. The chance that a major impact will occur during your lifetime is so small it is hard to estimate. But the consequences of such an impact are so severe that humanity should be preparing. One way to prepare is to find those objects that could hit us, map their orbits, and identify any that are dangerous.

What we do next isn't clear. Blowing up a dangerous asteroid in space might make a good movie, but converting one big projectile into a thousand small ones might not be very smart. Changing an asteroid's orbit could be difficult without a few decades' advance warning. Unlikely or not, large impacts demand consideration and preparation.

Throughout the universe there may be two kinds of inhabited worlds. On one type of world, intelligent creatures have developed ways to prevent asteroid and comet impacts from altering their climates and destroying their civilizations. But on other worlds, including Earth, intelligent races have not yet found ways to protect themselves. Some of those civilizations survive. Some don't.

Study and Review

Summary

- The term meteoroid refers to small solid particles orbiting in the solar system. The term meteor refers to a visible streak of light from a meteoroid heated and glowing as it enters Earth's atmosphere. The term meteorite refers to space material that has reached Earth's surface.
- Iron meteorites are mostly iron and nickel; when sliced open, polished, and etched, they show **Widmanstätten patterns (p. 409).** These reveal that the metal cooled from a molten state very slowly.
- Stony meteorites included **chondrites (p. 409),** which contain small, glassy particles called **chondrules (p. 409),** believed to be very ancient droplets of molten material formed in the solar nebula.
- Stony meteorites that are rich in volatiles and carbon are called **carbonaceous chondrites (p. 409).** They are among the least modified meteorites.
- An **achondrite (p. 410)** is a stony meteorite that contains no chondrules and no volatiles. Achondrites appear to have been melted after they formed and, in some cases, resemble solidified lavas.
- Many meteorites appear to have formed as part of larger bodies that melted, differentiated, and cooled very slowly. Later these bodies were broken up, and fragments from the core became iron meteorites, while fragments from the outer layers became stony meteorites.
- The evidence, including the orbits of meteorites seen to fall, suggests that meteorites are fragments of asteroids.
- The vast majority of meteors (visible streaks of light from particles heated by passage through the atmosphere), including meteors in **meteor showers (p. 410),** appear to be low-density, fragile bits of debris from comets.
- Asteroids are irregular in shape and heavily cratered from collisions. Their surfaces are covered by gray, pulverized rock, and some asteroids have such low densities they must be fragmented rubble piles.
- Most asteroids lie in a belt between Mars and Jupiter, although the **Trojan asteroids (p. 412)** share Jupiter's orbit and others have orbits that cross into the inner solar system. If they pass near Earth, they are called **Near-Earth Objects** (**NEOs**) **(p. 412).**
- **C-type (p. 413)** asteroids are more common in the outer asteroid belt where the solar nebula was cooler. They are darker and may be carbonaceous.
- **S-type (p. 413)** asteroids are the most common and may be the source of the most common kind of meteorites, the chondrites. S-type asteroids are more frequently found in the inner belt.
- Carbonaceous chondrites appear to have formed further from the sun.

- **M-type (p. 413)** asteroids appear to have nickel-iron compositions and may be the cores of differentiated asteroids shattered by collisions.
- The asteroids formed as rocky planetesimals between Mars and Jupiter, but Jupiter prevented them from forming a planet. Collisions have fragmented all but the largest of the asteroids. Most of the material inferred to have been originally in the asteroid belt has been gravitationally perturbed and swept up by the planets or tossed out of the solar system.
- A comet is produced by a lump of ices and rock about 1 to 10 km in diameter, referred to as the comet nucleus. In long, elliptical orbits, the icy nucleus stays frozen until it nears the sun. Then, some of the ices vaporize and release dust and gas that is blown away to form a tail.
- A **type I (gas) (p. 418)** comet tail is ionized gas carried away by the solar wind. A **type II (dust) (p. 418)** tail is solid debris released from the nucleus and blown outward by the pressure of sunlight. A comet's tail always points away from the sun, no matter in what direction the comet is moving.
- The **coma (p. 418),** or head, of a comet can be up to a million kilometers in diameter.
- Spacecraft flying past comets have revealed that they have very dark, rocky crusts and that jets of vapor and dust issue from active regions on the sunlit side.
- The low density of comet nuclei shows that they are irregular mixtures of ices and silicates, probably containing large voids. At least one comet nucleus has surface features showing the material has a surprising amount of strength.
- Comets are believed to have formed as icy planetesimals in the outer solar system, and some were ejected to form the **Oort cloud (p. 417).** Comets perturbed inward from the Oort cloud become long-period comets.
- Other icy bodies formed in the outer solar system and now make up the Kuiper belt beyond Neptune. Objects from the Kuiper belt that are perturbed into the inner solar system can become short-period comets.
- A major impact on Earth can trigger extinctions because of global forest fires caused by heated material falling back into the atmosphere, tsunamis inundating coastal regions around the world, acid rain resulting from large amounts of carbon dioxide released into the atmosphere, and climate change caused by the atmosphere filling with dust, plunging the entire Earth into darkness for years.
- An impact at **Chicxulub (p. 421)** in Mexico's Yucatán region 65 million years ago appears to have triggered the extinction of 75 percent of the species then on Earth, including the dinosaurs.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. What do Widmanstätten patterns indicate about the history of iron meteorites?
2. What do chondrules tell you about the history of chondrites?
3. Why are there no chondrules in achondritic meteorites?
4. Why do astronomers refer to carbonaceous chondrites as unmodified or "primitive" material?
5. How do observations of meteor showers reveal one of the sources of meteoroids?
6. How can most meteors be cometary if all meteorites are asteroidal?
7. Why do astronomers think the asteroids were never part of a full-sized planet?
8. What evidence indicates that the asteroids are mostly fragments of larger bodies?
9. What evidence indicates that some asteroids have differentiated?
10. What evidence indicates that some asteroids have had geologically active surfaces?
11. How is the composition of meteorites related to the formation and evolution of asteroids?
12. What is the difference between a type I comet tail versus a type II tail?
13. What evidence indicates that cometary nuclei are rich in ices?
14. Why do short-period comets tend to have orbits near the plane of the solar system?
15. What are the hypotheses for how the bodies in the Kuiper belt and the Oort cloud formed?
16. **How Do We Know?** How can scientific understanding increase your enjoyment of the natural world?

Discussion Questions

1. It has been suggested that humans may someday mine the asteroids for materials to build and supply space colonies. What kinds of materials could Earthlings get from asteroids? (*Hint:* What are S-, M-, and C-type asteroids made of?)
2. If cometary nuclei were heated during the formation of the solar system by internal radioactive decay rather than by solar radiation, how would comets differ from what is observed?
3. Do you think the government should spend money to find near-Earth asteroids? How serious is the risk?

Problems

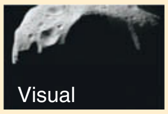
1. Large meteorites are hardly slowed by Earth's atmosphere. Assuming the atmosphere is 100 km thick and that a large meteorite falls perpendicular to the surface, how long does it take to reach the ground? (*Hint:* About how fast do meteoroids travel?)
2. If a single asteroid 1 km in diameter were fragmented into meteoroids 1 m in diameter, how many would it yield? (*Hint:* The volume of a sphere $= \frac{4}{3}\pi r^3$.)
3. What is the orbital period of a typical asteroid? (*Hint:* Use Kepler's third law. See Table 4-1.)
4. If a trillion (10^{12}) asteroids, each 1 km in diameter, were assembled into one body, how large would it be? (*Hint:* The volume of a sphere $= \frac{4}{3}\pi r^3$.) Compare that to the size of Earth.
5. What is the maximum angular diameter of the largest asteroid, Ceres, as seen from Earth? Could Earth-based telescopes detect surface features? Could the Hubble Space Telescope? (*Hints:* See Reasoning with Numbers 3-1. The angular resolution of Earth-based telescopes is about 1 arcsec, and of Hubble about 0.1 arcsec. Ceres's average distance from the sun is 2.8 AU.)
6. If the velocity of the solar wind is about 400 km/s and the visible tail of a comet is 1×10^8 km long, how long does it take a solar wind atom to travel from the nucleus to the end of the visible tail?
7. If you saw Comet Halley when it was 0.7 AU from Earth and it had a visible tail 5° long, how long was the tail in kilometers? Suppose that the tail was not perpendicular to your line of sight. Is your first answer too large or too small? (*Hint:* See Reasoning with Numbers 3-1.)
8. What is the orbital period of a comet nucleus in the Oort cloud? What is its orbital velocity? (*Hints:* Use Kepler's third law. The circumference of a circular orbit $= 2\pi r$.)
9. The mass of an average comet's nucleus is about 10^{12} kg. If the Oort cloud contains 2×10^{11} comet nuclei, what is the mass of the cloud in Earth masses? Compare that with Jupiter's mass. (*Hint:* See Appendix A.)

Learning to Look

1. What do you see in the image to the right that tells you the size of planetesimals when the solar system was forming?

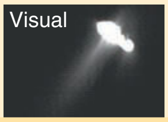
Russell Kempton, New England Meteoritical

2. Discuss the surface of the asteroid Mathilde, pictured to the right. What do you see that tells you something about the history of the asteroids?



NASA

3. What do you see in this image of the nucleus of Comet Borrelly that tells you how comets produce their comas and tails?



NASA

20 Life on Other Worlds

Every life form on Earth has evolved to survive in some ecological niche. The Wekiu bug lives with the astronomers at an altitude of 13,800 feet atop the Hawaiian volcano Mauna Kea. It inhabits spaces between the icy cinders and eats insects carried up by ocean breezes.
(Kris Koenig/Coast Learning Systems)

Guidepost

This chapter is either unnecessary or vital. If you believe that astronomy is the study of the physical universe above the clouds, then you are done; the previous 19 chapters completed your study of astronomy. But, if you believe that astronomy is the study not only of the physical universe but also of your role as a living being in the evolution of the universe, then everything you have learned so far from this book has been preparation for this final chapter.

As you read this chapter, you will ask four important, related questions:

What is life?

How did life originate on Earth?

Could life begin on other worlds?

Can humans on Earth communicate with civilizations on other worlds?

You won't get more than the beginnings of answers to those questions here, but often in science asking a question is more important than getting an immediate answer.

You have explored the universe from the phases of the moon to the big bang, from the origin of Earth to the death of the sun. Astronomy is meaningful, not just because it is about the universe, but because it is also about you. Now that you know astronomy, you see yourself and your world in a different way. Astronomy has changed you.

Animated! **This bar denotes active figures that may be found at academic.cengage.com/astronomy/seeds.**

Did I solicit thee from darkness to promote me?

ADAM, TO GOD,
IN JOHN MILTON'S *PARADISE LOST*

As a living thing, you have been promoted from darkness. The atoms of carbon, oxygen, and other heavy elements that are necessary components of your body did not exist at the beginning of the universe but were created by successive generations of stars.

The elements from which you are made are common everywhere in the observable universe, so it is possible that life began on other worlds and evolved to intelligence. If so, perhaps those other civilizations will be detected from Earth. Future astronomers may discover distant alien species completely different from any life on Earth.

Your goal in this chapter is to try to understand truly intriguing puzzles—the origin and evolution of life on Earth, and whether there is life on other worlds (■ How Do We Know? 20-1)

20-1 The Nature of Life

What is life? Philosophers have struggled with that question for thousands of years, and it is not possible to answer it in a single chapter. An attempt at a general definition of what living things do, distinguishing them from nonliving things, might be the following: Life is a *process* by which an organism extracts energy from the surroundings, maintains itself, and modifies the surroundings to promote its own survival and reproduction.

One very important observation is that all living things on Earth, no matter how apparently different, share certain characteristics in how they perform the process of life.

The Physical Basis of Life

The physical basis of life on Earth is the element carbon (■ Figure 20-1). Because of the way carbon atoms bond to each other and to other atoms, they can join into long, complex, stable chains that are capable, among many other feats, of storing and

How Do We Know? 20-1

The Nature of Scientific Explanation

Must science and religion be in conflict? Science is a way of understanding the world around you, and at the heart of that understanding are explanations that science gives you for natural phenomena. Whether you call these explanations stories, histories, hypotheses, or theories, they are attempts to describe how nature works based on evidence and intellectual honesty. While you may take these explanations as factual truth, you should understand that they are not the only explanations that describe the world.

A separate class of explanations involves religion. The Old Testament description of the creation of the world does not fit with scientific observations, but it is a way of understanding the world nonetheless. Religious explanations are based partly on faith rather than on strict rules of logic and evidence, and it is wrong to demand that they follow the same rules as scientific explanations. In the same way, it is wrong to demand that scientific explanations take into account religious beliefs. The so-called conflict between science and religion arises when people fail to recognize that science and religion are different ways of knowing about the world.

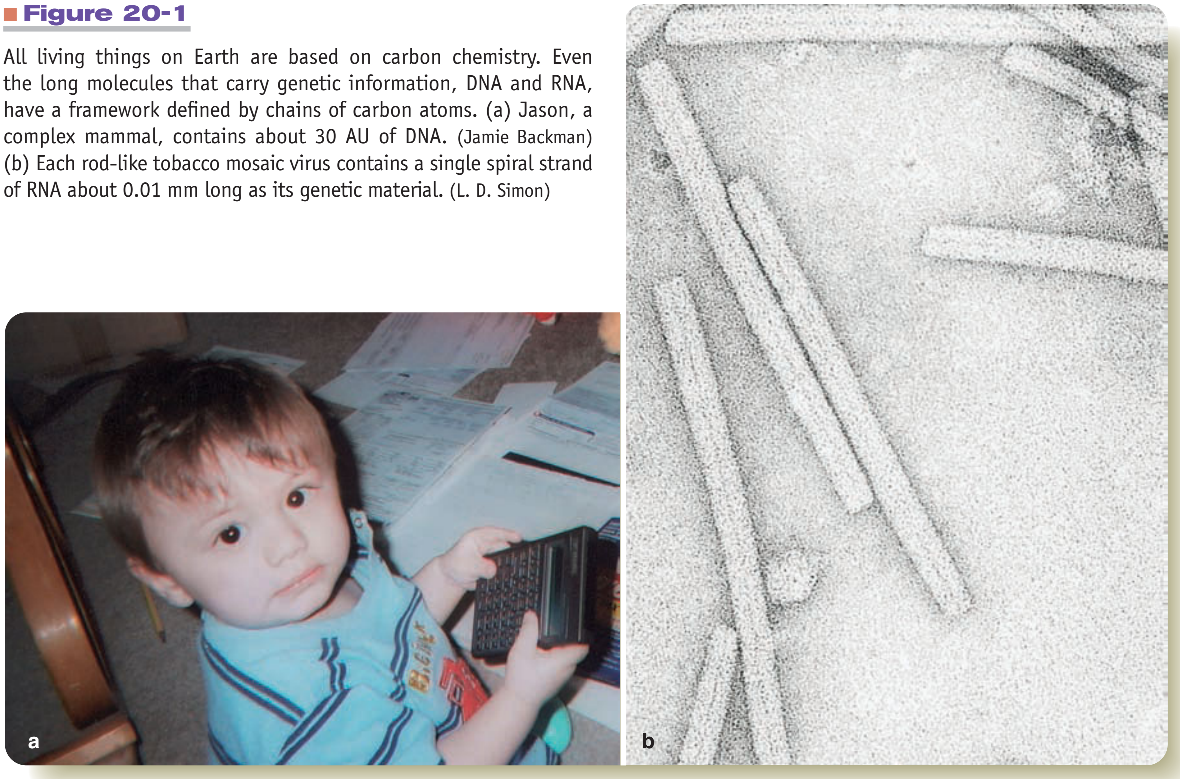
Scientific explanations are very convincing because science has been very successful at producing technological innovations that have changed the world you live in. From new vaccines to digital music players to telescopes that can observe the most distant galaxies, the products of the scientific process are all around you. Scientific explanations have provided tremendous insights into the workings of nature, and it is easy to forget that there are other explanations.

Science and religion both have a lot to offer by their differing ways of explaining the world, but the two ways follow separate rules and cannot be judged by each other's standards. The trial of Galileo can be understood as a conflict between these two ways of knowing.

Galileo's telescope gave him a new way to know about the world.

Figure 20-1

All living things on Earth are based on carbon chemistry. Even the long molecules that carry genetic information, DNA and RNA, have a framework defined by chains of carbon atoms. (a) Jason, a complex mammal, contains about 30 AU of DNA. (Jamie Backman) (b) Each rod-like tobacco mosaic virus contains a single spiral strand of RNA about 0.01 mm long as its genetic material. (L. D. Simon)



transmitting information. A large amount of information is necessary to control the activities, and maintain the forms, of living things.

It is possible that life on other worlds could use silicon instead of carbon. Silicon is right below carbon in the periodic chart (see Appendix Table A-16), which means that it shares many of carbon's chemical properties. But life based on silicon seems unlikely because silicon chains are harder to assemble and disassemble than their carbon counterparts and can't be as lengthy. Science fiction has proposed even stranger life forms based on, for example, electromagnetic fields and ionized gas, and none of these possibilities can be ruled out. These hypothetical life forms make for fascinating speculation, but they can't be studied as systematically as life on Earth. This chapter is concerned with the origin and evolution of life as it is on Earth, based on carbon, not because of lack of imagination, but because it is the only form of life about which we know anything.

Even carbon-based life has its mysteries. What makes a lump of carbon-based molecules a living thing? An important part of the answer lies in the transmission of information from one molecule to another.

Information Storage and Duplication

Most actions performed by living cells are carried out by molecules built within the cells. Cells must store recipes for making all these molecules as well as how and when to use them, and then somehow pass the recipes on to their offspring.

Study **DNA: The Code of Life** on pages 428–429 and notice three important points and seven new terms:

1. The chemical recipes of life are stored as templates on *DNA (deoxyribonucleic acid)* molecules. DNA templates guide specific chemical reactions within the cell.
2. The instructions stored in DNA are genetic information handed down to offspring.
3. The DNA molecule reproduces itself when a cell divides so that each new cell contains a copy of the original information.

To produce viable offspring, a cell must be able to make copies of its DNA. Surprisingly, it is important for the continued existence of all life that the copying process includes mistakes.

Modifying the Information

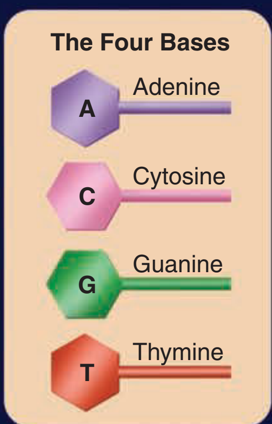
Earth's environment changes continuously. To survive, species must change as their food supply, climate, or home terrain changes. If the information stored in DNA could not change, then life would go extinct quickly. The process by which life adjusts itself to changing environments is called **biological evolution.**

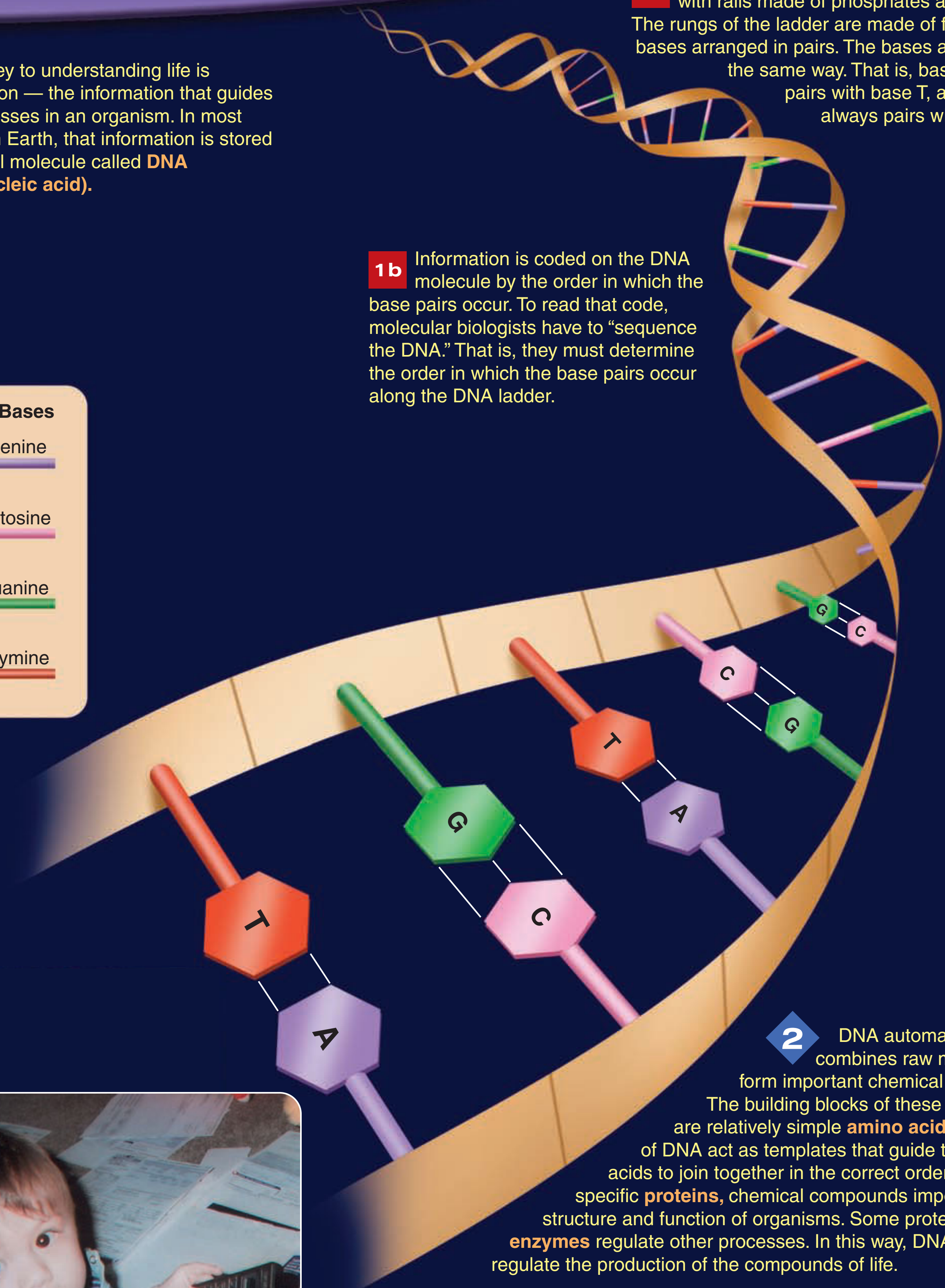
When an organism reproduces, its offspring receive a copy of its DNA. Sometimes external effects like radiation alter the DNA during the parent organism's lifetime, and sometimes mistakes are made in the copying process, so that occasionally the

DNA: The Code of Life

1 The key to understanding life is information — the information that guides all of the processes in an organism. In most living things on Earth, that information is stored on a long spiral molecule called **DNA (deoxyribonucleic acid).**

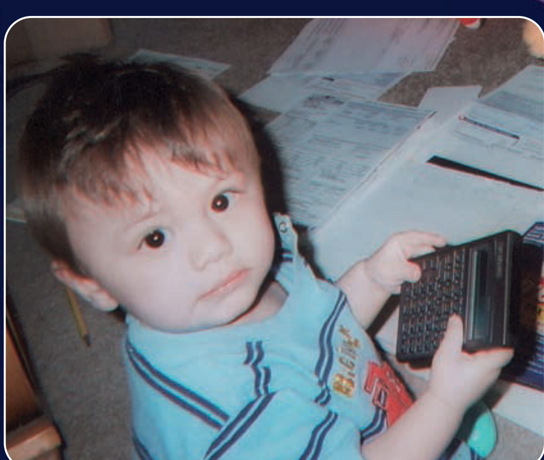
1a The DNA molecule looks like a spiral ladder with rails made of phosphates and sugars. The rungs of the ladder are made of four chemical bases arranged in pairs. The bases always pair the same way. That is, base A always pairs with base T, and base G always pairs with base C.





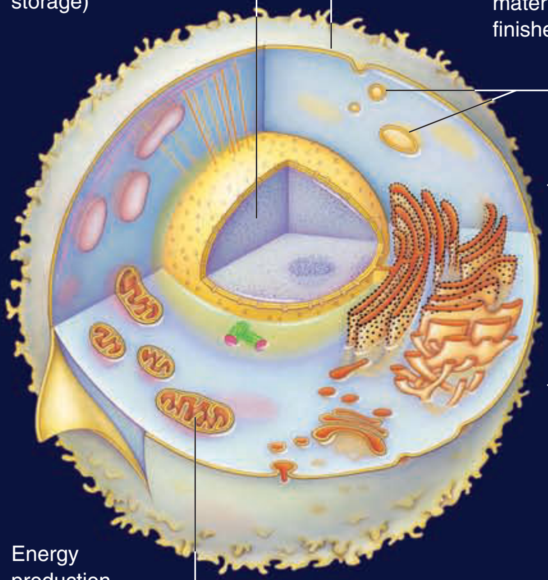
1b Information is coded on the DNA molecule by the order in which the base pairs occur. To read that code, molecular biologists have to “sequence the DNA.” That is, they must determine the order in which the base pairs occur along the DNA ladder.

2 DNA automatically combines raw materials to form important chemical compounds. The building blocks of these compounds are relatively simple **amino acids.** Segments of DNA act as templates that guide the amino acids to join together in the correct order to build specific **proteins,** chemical compounds important to the structure and function of organisms. Some proteins called **enzymes** regulate other processes. In this way, DNA recipes regulate the production of the compounds of life.



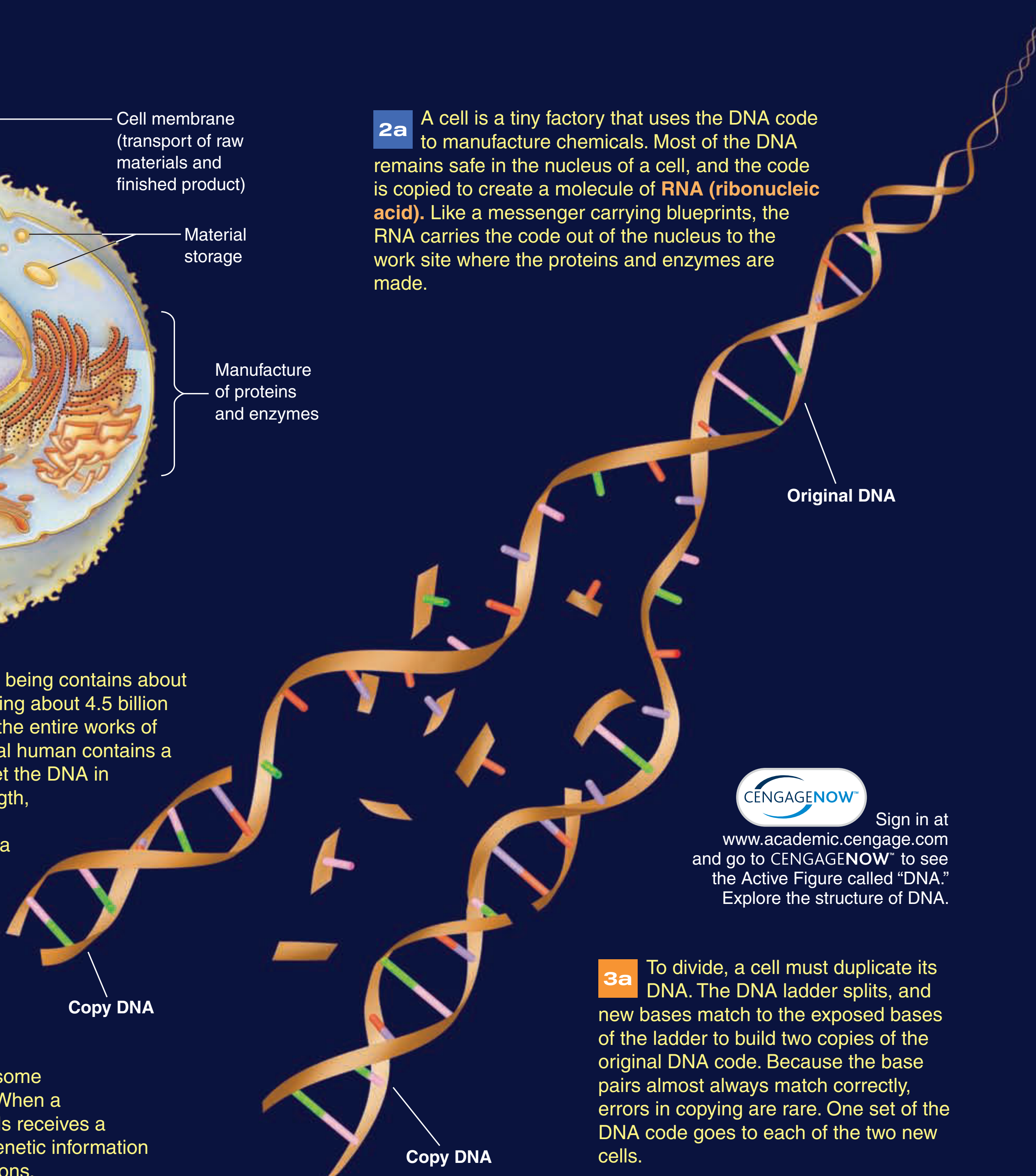
Jamie Backman

The traits you inherit from your parents, the chemical processes that animate you, and the structure of your body are all encoded in your DNA. When people say “you have your mother’s eyes,” they are talking about DNA codes.



2a A cell is a tiny factory that uses the DNA code to manufacture chemicals. Most of the DNA remains safe in the nucleus of a cell, and the code is copied to create a molecule of **RNA (ribonucleic acid).** Like a messenger carrying blueprints, the RNA carries the code out of the nucleus to the work site where the proteins and enzymes are made.

2b A single cell from a human being contains about 1.5 meters of DNA containing about 4.5 billion base pairs — enough to record the entire works of Shakespeare 200 times. A typical human contains a total of about 600 AU of DNA. Yet the DNA in each cell, only 1.5 meters in length, contains all of the information to create a new human. A clone is a new creature created from the DNA code found in a single cell.

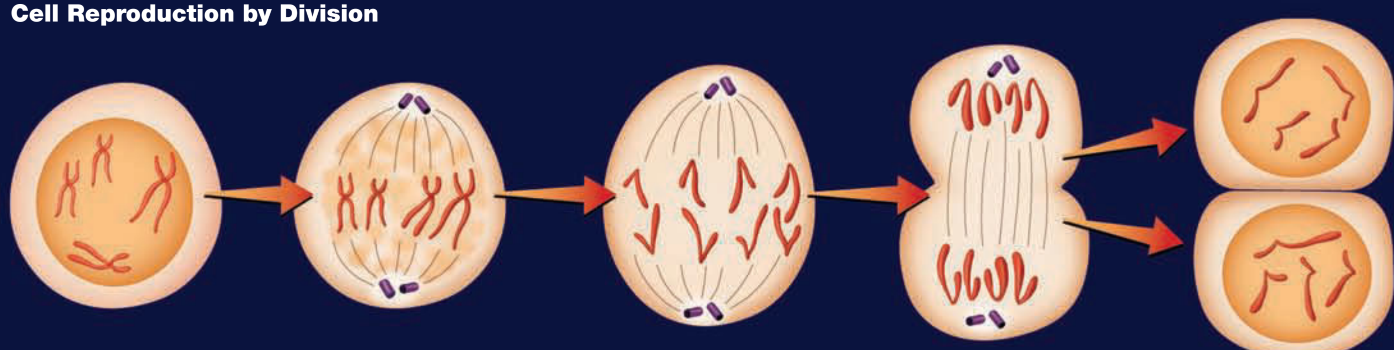


CENGAGENOW™ Sign in at www.academic.cengage.com and go to CENGAGENOW™ to see the Active Figure called "DNA." Explore the structure of DNA.

3 DNA, coiled into a tight spiral, makes up the **chromosomes** that are the genetic material in a cell. A **gene** is a segment of a chromosome that controls a certain function. When a cell divides, each of the new cells receives a copy of the chromosomes, as genetic information is handed down to new generations.

3a To divide, a cell must duplicate its DNA. The DNA ladder splits, and new bases match to the exposed bases of the ladder to build two copies of the original DNA code. Because the base pairs almost always match correctly, errors in copying are rare. One set of the DNA code goes to each of the two new cells.

Cell Reproduction by Division



As a cell begins to divide, its DNA duplicates itself.

The duplicated chromosomes move to the middle.

The two sets of chromosomes separate, and . . .

the cell divides to produce . . .

two cells, each containing a full set of the DNA code.

copy is slightly different from the original. Offspring born with random alterations to their DNA are called **mutants.** Most mutations make no difference, but some mutations are fatal, killing the afflicted organisms before they can reproduce. In rare but vitally important cases, a mutation can actually help an organism survive.

These changes produce variation among the members of a species. All of the squirrels in the park may look the same, but they carry a range of genetic variation. Some may have slightly longer tails or faster-growing claws. These variations make almost no difference until the environment changes. For example, if the environment becomes colder, a squirrel with a heavier coat of fur will, on average, survive longer and produce more offspring than its normal contemporaries. Likewise, the offspring that inherit this beneficial variation will also live longer and have more offspring of their own. These differing rates of survival and reproduction are examples of **natural selection.** Over time, the beneficial variation increases in frequency, and a species can evolve until the entire population shares the trait. In this way, natural selection adapts species to their changing environments by selecting, from the huge array of random variations, those that would most benefit the survival of the species.

It is a **Common Misconception** that evolution is random, but that is not true. The underlying variation within species is random, but natural selection is not random because progressive changes in a species are directed by changes in the environment.

◄ SCIENTIFIC ARGUMENT ►

Why is it important that errors occur in copying DNA?

Sometimes the most valuable scientific arguments are those that challenge what seems like common sense. It seems obvious that mistakes shouldn't be made in copying DNA, but in fact variation is necessary for long-term survival of a species. For example, the DNA in a starfish contains all the information the starfish needs to grow, develop, survive, and reproduce. The information must be passed on to the starfish's offspring for them to survive. But that information must change if the environment changes. A change in the ocean's temperature may kill the specific shellfish that the starfish eat. If none of the starfish are able to digest another kind of food—if all the starfish have exactly the same DNA—they will all die. But if a few starfish are born with the ability to make enzymes that can digest a different kind of shellfish, the species may be able to carry on.

Variations in DNA are caused both by external factors such as radiation and by occasional mistakes in the copying process. The survival of life depends on this delicate balance between mostly reliable reproduction and the introduction of small variations in DNA. Now build a new argument. **How does DNA make copies of itself?**

◄ ►

20-2 The Origin of Life

It is obvious that the 4.5 billion chemical bases that make up human DNA did not just come together in the right order by chance. The key to understanding the origin of life lies in the processes of evolution. The complex interplay of environmental factors with the DNA of generation after generation of organisms drove some life forms to become more sophisticated over time, until they became the unique and specialized creatures on Earth today.

This means that life on Earth could have begun very simply, even as simple as carbon-chain molecules able to copy themselves. Of course, this is a hypothesis for which you can seek evidence. What evidence exists regarding the origin of life on Earth?

Life on Earth

The oldest fossils are the remains of sea creatures, and this indicates that life began in the sea. Identifying the oldest fossils is not easy, however. Rocks from western Australia that are at least 3.4 billion years old contain features that biologists identify as **stromatolites,** fossilized remains of colonies of single-celled organisms (■ Figure 20-2). Fossils this old are difficult to recognize because the earliest living things did not contain easily preserved hard parts like bones or shells and because the individual organisms were microscopic. The evidence, though scarce, indicates that simple organisms lived in Earth's oceans at least 3.4 billion years ago, less than 1.2 billion years after Earth formed. Stromatolite colonies of microorganisms are more complex than individual cells, so you can imagine there were probably earlier, simpler organisms. Where did those first simplest organisms come from?

An important experiment performed by Stanley Miller and Harold Urey in 1952 sought to recreate the conditions in which life on Earth began. The **Miller experiment** consisted of a sterile, sealed glass container holding water, hydrogen, ammonia, and methane. An electric arc inside the apparatus created sparks to simulate the effects of lightning in Earth's early atmosphere (■ Figure 20-3).

■ **Figure 20-2**

A fossil stromatolite from western Australia that is more than 3 billion years old, constituting some of the oldest evidence of life on Earth. Stromatolites are formed, layer upon layer, by mats of bacteria living in shallow water where they are covered repeatedly by sediments. (Chip Clark, National Museum of Natural History)

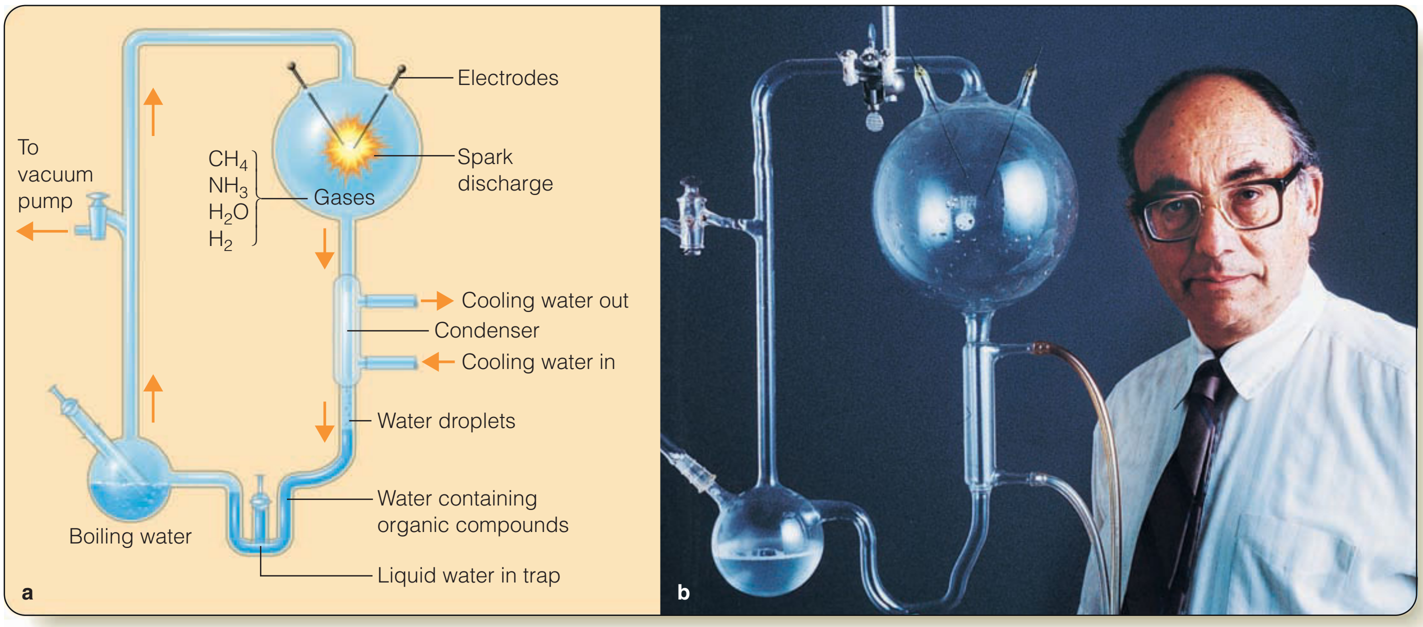


Figure 20-3

(a) The Miller experiment circulated gases through water in the presence of an electric arc. This simulation of primitive conditions on Earth produced many complex organic molecules, including amino acids, the building blocks of proteins. (b) Stanley Miller with a Miller apparatus. (Stanley Miller)

Miller and Urey let the experiment run for a week and then analyzed the material inside. They found that the interaction between the electric arc and the simulated atmosphere had produced many organic molecules from the raw material of the experiment, including such important building blocks of life as amino acids. When the experiment was run again using different energy sources such as hot silica to represent molten lava spilling into the ocean, similar molecules were produced. Even a source of ultraviolet radiation representing the small amount of UV in sunlight was sufficient to produce complex organic molecules.

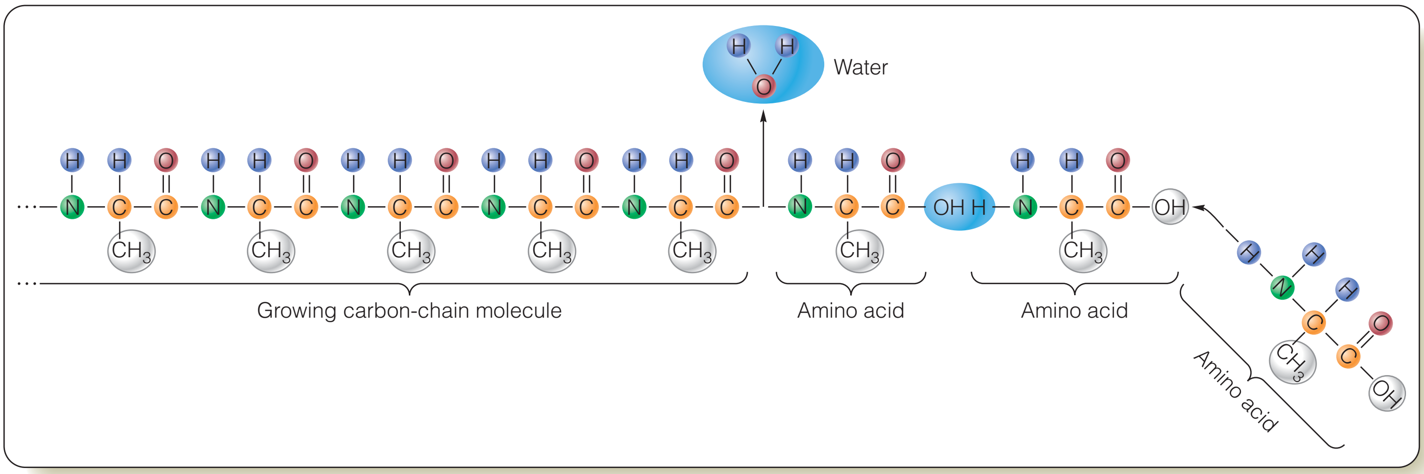
Scientists are professionally skeptical about scientific findings (see "How Do We Know" 16-3), and they re-evaluated the Miller-Urey experiment in light of new information. According to updated models of the formation of the solar system and Earth (Chapters 16 and 17), Earth's early atmosphere probably consisted mostly of carbon dioxide, nitrogen, and water vapor instead of the mix of hydrogen, ammonia, methane, and water vapor assumed by Miller and Urey. When gases corresponding to the newer understanding of the early Earth atmosphere are processed in a Miller apparatus, lesser but still significant numbers of organic molecules are created.

The Miller experiment is important because it shows that complex organic molecules form naturally in a wide variety of circumstances. Lightning, sunlight, and hot lava are just some of the energy sources that can naturally rearrange simple common molecules into the complex molecules that make life possible. If you could travel back in time, you would expect to find Earth's first oceans filled with a rich mixture of organic compounds called the **primordial soup.**

Many of these organic compounds would have been able to link up to form larger molecules. Amino acids, for example, can link together to form proteins by joining ends and releasing a water molecule (■ Figure 20-4). That reaction, however, does not proceed easily in a water solution. Scientists hypothesize that this step may have been more likely to happen in sun-warmed tidal pools or on shorelines where organic molecules from the primordial soup could have been concentrated by evaporation. The production of large organic molecules may have been aided in such semi-dry environments by clay crystals acting as templates to hold the organic subunits close together.

These complex organic molecules were still not living things. Even though some proteins may have contained hundreds of amino acids, they did not reproduce but rather linked and broke apart at random. Because some molecules are more stable than others, and some bond together more easily than others, scientists hypothesize that a process of **chemical evolution** eventually concentrated the various smaller molecules into the most stable larger forms. Eventually, according to the hypothesis, somewhere in the oceans, after sufficient time, a molecule formed that could copy itself. At that point, the chemical evolution of molecules became the biological evolution of living things.

An alternate theory for the origin of life holds that reproducing molecules may have arrived here from space. Radio astronomers have found a wide variety of organic molecules in the interstellar medium, and similar compounds have been found inside meteorites (■ Figure 20-5). The Miller experiment showed how easy it is to create complex organic molecules from simpler compounds, so it is not surprising to find them in space. Although



■ **Figure 20-4**

Amino acids can link together via the release of a water molecule to form long carbon-chain molecules. The amino acid in this hypothetical example is alanine, one of the simplest.

■ **Figure 20-5**

A piece of the Murchison meteorite, a carbonaceous chondrite (see Chapter 19) that fell in 1969 near Murchison, Australia. Analysis of the interior of the meteorite revealed the presence of amino acids. Whether the first building blocks of life originated in space is unknown, but the amino acids found in meteorites illustrate how commonly amino acids and other complex organic molecules occur in the universe, even in the absence of living things. (Chip Clark, National Museum of Natural History)

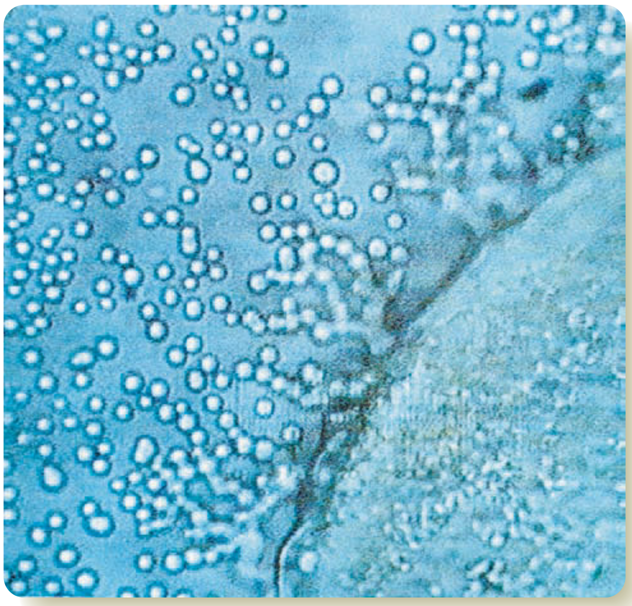
speculation is fun, the hypothesis that life arrived on Earth from space is presently more difficult to test than the hypothesis that life on Earth originated on Earth.

Whether the first reproducing molecules formed here on Earth or in space, the important thing is that they could have formed by natural processes. Scientists know enough about those processes to feel confident about them, even though some of the steps remain unknown.

The details of the evolution of the first cells are unknown, but the first reproducing molecule to surround itself with a protective membrane would have gained an important survival advantage. Experiments have shown that microscopic spheres the size of cells containing organic molecules form relatively easily in water (■ Figure 20-6), so the evolution of cell membranes is not surprising.

The first cells must have been simple single-celled organisms much like modern bacteria. As you learned earlier, these kinds of cells are preserved in stromatolites (Figure 20-2), mineral formations produced by layers of photosynthetic bacteria and shallow ocean sediments. Stromatolites are in rocks with radioactive ages of 3.4 billion years, and they also still form in some places today. Stromatolites and other photosynthetic organisms would have begun adding oxygen, a product of photosynthesis, to Earth's early atmosphere (■ Figure 20-7). An oxygen abundance of only 0.1 percent would have created an ozone screen, protecting organisms from the sun's ultraviolet radiation and later allowing life to colonize the land.

Over the course of eons, the natural processes of evolution gave rise to stunningly complex **multicellular** life forms with



■ **Figure 20-6**

Single amino acids can be assembled into long proteinlike molecules. When such material cools in water, it can form microspheres, microscopic globules with double-layered boundaries similar to cell membranes. Microspheres may have been an intermediate stage in the evolution of life between complex but nonliving molecules and living cells holding molecules reproducing genetic information. (Sidney Fox and Randall Grubbs)

their own widely differing ways of life. It is a **Common Misconception** to imagine that life is too complex to have evolved from such simple beginnings. It is possible because small variations can accumulate, although that accumulation requires huge amounts of time.

Geologic Time

Life has existed on Earth for at least 3.4 billion years, but there is little evidence of anything more than simple organisms until about 540 million years ago, when life suddenly branched into a wide variety of complex forms like the trilobites (■ Figure 20-8). This sudden increase in complexity is known as the **Cambrian explosion,** and marks the beginning of the Cambrian period.

If you represented the entire history of Earth on a scale diagram, the Cambrian explosion would be near the top of the column, as shown at the left of ■ Figure 20-9. The emergence of most animals familiar to you today, including fishes, amphibians, reptiles, birds, and mammals, would be crammed into the topmost part of the chart, above the Cambrian explosion.

If you magnify that portion of the diagram, as shown on the right side of Figure 20-9, you can get a better idea of when these events occurred in the history of life. Humanoid creatures have walked on Earth for about 4 million years. This is a long time by the standard of a human lifetime, but it makes only a narrow red

■ **Figure 20-7**

Artist's conception of a scene on the young Earth, 3 billion years ago, with stromatolite bacterial mats growing near the shores of an ocean. (National Museum of Natural History, Peter Sawyer, © 2009 Smithsonian Institution)

Animated!

Figure 20-8

Trilobites made their first appearance in the Cambrian oceans. The smallest were almost microscopic, and the largest were bigger than dinner plates. This example, about the size of a human hand, lived 400 million years ago in an ocean floor that is now a limestone deposit in Pennsylvania. (Franklin & Marshall College/Grundy Observatory/Michael Seeds)

line at the top of the diagram. All of recorded history would be a microscopically thin line at the very top of the column.

To understand just how thin that line is, imagine that the entire 4.6-billion-year history of the Earth has been compressed onto a yearlong video and that you began watching this video on January 1. You would not see any signs of life until March or early April, and the slow evolution of the first simple forms would take the next six or seven months. Suddenly, in mid-November, you would see the trilobites and other complex organisms of the Cambrian explosion.

You would see no life of any kind on land until November 28, but once life appeared it would diversify quickly, and by December 12 you would see dinosaurs walking the continents. By the day after Christmas they would be gone, and mammals and birds would be on the rise.

If you watched closely, you might see the first humanoid forms by late afternoon on New Year's Eve, and by late evening you could see humans making the first stone tools. The Stone Age would last until 11:59 PM, after which the first towns, and then cities would appear. Suddenly things would begin to happen at lighting speed. Babylon would flourish, the Pyramids would rise, and Troy would fall. The Christian era would begin 14 seconds before the New Year. Rome would fall, the Middle Ages and the Renaissance would flicker past. The American and French revolutions would occur one and a half seconds before the end of the video.

By imagining the history of Earth as a yearlong video, you have gained some perspective on the rise of life. Tremendous amounts of time were needed for the first simple living things to evolve in the oceans; but, as life became more complex, new forms arose more and more quickly as the hardest problems—how to reproduce, how to take energy efficiently from the environment, how to move around—were "solved" by the process of biological evolution. The easier problems, like what to eat, where to live, and how raise young, were managed in different ways by different organisms, leading to the diversity that is seen today.

Even intelligence—that which appears to set humans apart from other animals—may be a unique solution to an evolutionary problem posed to humanity's ancient ancestors. A smart animal is better able to escape predators, outwit its prey, and feed and shelter itself and its offspring, so under certain conditions evolution is likely to select for intelligence. Could intelligent life arise on other worlds? To try to answer this question, you can estimate the chances of any type of life arising on other worlds, then assess the likelihood of that life developing intelligence.

Life in Our Solar System

Could there be carbon-based life elsewhere in our solar system? Liquid water seems to be a requirement of carbon-based life, necessary both as the medium for vital chemical reactions and to transport nutrients and wastes. It is not surprising that life developed in Earth's oceans and stayed there for billions of years before it was able to colonize the land.

Scientists are in agreement that any world harboring living things must have significant quantities of some type of liquid. Water is a cosmically abundant substance with properties such as high heat capacity that set it apart from other common molecules that are liquid at the temperatures of planetary surfaces. Maybe there are other liquids that can support the processes of life on other planets, but water, aside from being common in the universe, has characteristics that cause Earth scientists to regard it as special, and not just because they are made of water themselves.

Many worlds in the solar system can be eliminated immediately as hosts for life because liquid water is not possible there. The moon and Mercury are airless, and water would boil away into space immediately. Venus has traces of water vapor in its atmo-

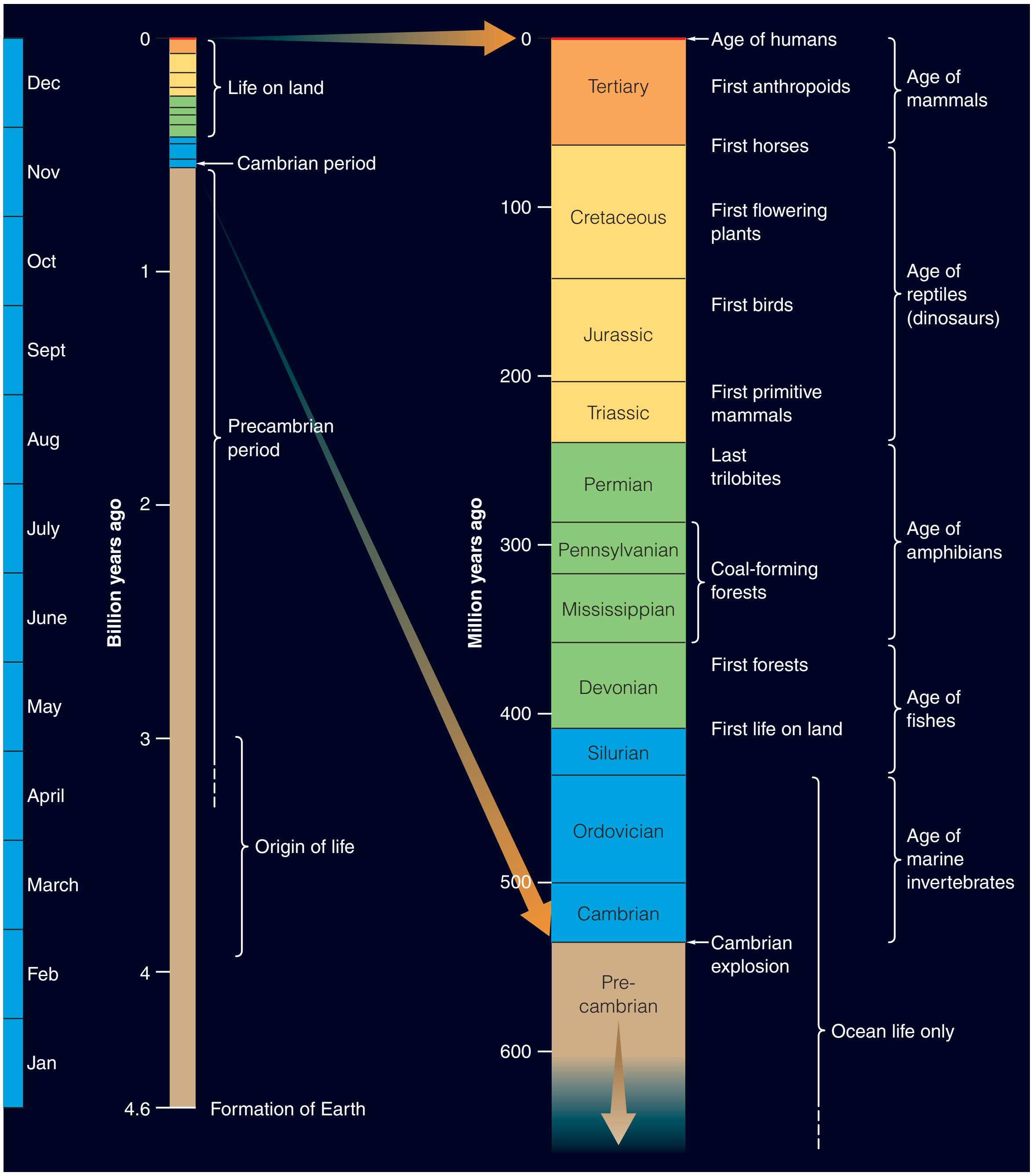


Figure 20-9

Complex life has developed on Earth only recently. If the entire history of Earth were represented in a time line (left), you would have to magnify the end of the line to see details such as life leaving the oceans and dinosaurs appearing. The age of humans would still be only a thin line at the top of your diagram. If the history of Earth were a yearlong videotape, humans would not appear until the last hours of December 31.

sphere, but it is too hot for liquid water to survive on the surface. The Jovian planets have deep atmospheres, and at a certain level it is likely that water condenses into liquid droplets. However, it seems unlikely that life could have originated there. The Jovian planets do not have solid surfaces (see Chapter 18), so isolated water droplets cannot mingle to mimic the rich primordial oceans of Earth, where organic molecules grew and interacted. Additionally, powerful downdraft currents in the gas giants' atmospheres would quickly carry any reproducing molecules that did form there into inhospitably hot regions of the atmosphere.

As you learned in Chapter 18, at least one of the Jovian satellites could potentially support life. Jupiter's moon Europa appears to have a liquid-water ocean below its icy crust, and minerals dissolved in the water could provide a rich source of raw material for chemical evolution. Europa's ocean is kept warm and liquid now by tidal heating. There also may be liquid water layers under the surfaces of Ganymede and Callisto. That can change as the orbits of the moons interact; Europa, Ganymede, and Callisto may have been frozen solid at other times in their histories.

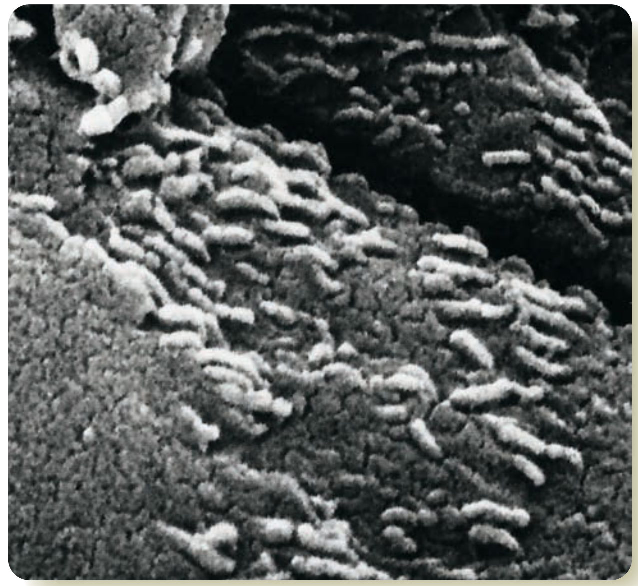
Saturn's moon Titan is rich in organic molecules. Chapter 18 described how sunlight converts the methane in Titan's atmosphere into organic smog particles that settle to the surface. The chemistry of life that could have evolved from those molecules and survived in Titan's lakes of methane is unknown. It is fascinating to consider possibilities, but Titan's extremely low temperature of −180°C (−290°F) would make chemical reactions slow to the point where life processes seem unlikely.

Water containing organic molecules has been observed venting from the south polar region of Saturn's moon Enceladus. (See Chapter 18.) It is possible that life could exist in that water under Enceladus's crust, but the moon is very small and its tidal heating may operate only occasionally. Enceladus may not have had plentiful liquid water for the extended time necessary for the rise of life.

Mars is the most likely place for life to exist in the solar system because, as you learned in Chapter 17, there is a great deal of evidence that liquid water once flowed on its surface. Even so, results from searches for signs of life on Mars are not encouraging. The robotic spacecraft Viking 1 and Viking 2 landed on Mars in 1976 and tested soil samples for living organisms. Some of the tests had puzzling semi-positive results that scientists hypothesize were caused by nonbiological chemical reactions in the soil. No evidence clearly indicates the presence of life or even of organic molecules in the Martian soil.

There was a splash of news stories in the 1990s regarding supposed chemical and physical traces of life on Mars discovered inside a Martian meteorite found in Antarctica (■ Figure 20-10). Scientists were excited by the announcement, but they employed professional skepticism and immediately began testing the evidence. Their results suggest that the unusual chemical signatures in the rock may have formed by processes that did not involve life. Tiny features in the rock that were originally thought to be fossils of ancient Martian microorganisms could possibly be nonbiological mineral formations instead. This is the only direct evidence yet found regarding potential life on Mars, but it remains highly controversial. Conclusive evidence of life on Mars may have to wait until a geologist from Earth can scramble down dry Martian streambeds and crack open rocks looking for fossils.

There is no strong evidence for the existence of life elsewhere in the solar system. Now your search will take you to distant planetary systems.



■ Figure 20-10

Meteorite ALH84001 is one of a dozen meteorites known to have originated on Mars. A research group claimed that the meteorite contained chemical and physical traces of ancient life on Mars, including what appear to be fossils of microscopic organisms. That evidence has not been confirmed, and the claim continues to be tested and debated. (NASA)

Life in Other Planetary Systems

Could life exist in other planetary systems? You already know that there are many different kinds of stars and that many of these stars have planetary systems. As a first step toward answering this question, you can try to identify the kinds of stars that seem most likely to have stable planetary systems where life could evolve.

If a planet is to be a suitable home for living things, it must be in a stable orbit around its sun. That is easy in a planetary system like our own, but planet orbits in binary star systems would be unstable unless the component stars are very close together or very far apart. Astronomers can calculate that, in binary systems with stars separated by intermediate distances of a few AU, the planets should eventually be swallowed up by one of the stars or ejected from the system. Half the stars in the galaxy are members of binary systems, and many of them are unlikely to support life on planets.

Moreover, just because a star is single does not necessarily make it a good candidate for sustaining life. Earth required perhaps as much as 1 billion years to produce the first cells and 4.6 billion years for intelligence to emerge. Massive stars that live only a few million years do not meet this criterion. If the history of life on Earth is representative, then stars more massive and luminous than about spectral type F5 are too short lived for complex life to develop. Main-sequence stars of types G and K, and possibly some of the faint M stars, are the best candidates.

The temperature of a planet is also important, and that depends on the type of star it orbits and its distance from the star. Astronomers have defined a **habitable zone** around a star as a region within which planets have temperatures that permit the existence of liquid water. The sun's habitable zone extends from around the orbit of Venus to the orbit of Mars, with Earth right in the middle. A low-luminosity star has a small habitable zone, and a high-luminosity star has a large one.

Scientists on Earth are finding life in places previously judged inhospitable, such as the bottoms of ice-covered lakes in Antarctica and far underground inside solid rock. Life has also been found in boiling hot springs with highly acidic water. As a result, it is difficult to pin down a range of environments and be sure that life cannot exist outside those conditions. You should also note that three of the environments listed as possible havens for life, Europa, Titan, and Enceladus, are in the outer solar system, far outside the sun's conventionally defined habitable zone. Stable planets inside the habitable zones of long-lived stars are the places where life seems most likely, but, given the tenacity and resilience of Earth's life forms, there might be other, seemingly inhospitable, places in the universe where life exists.

◄ SCIENTIFIC ARGUMENT ►

What evidence indicates that life is possible on other worlds?

A good scientific argument involves careful analysis of evidence. Fossils on Earth show that life originated in the oceans at least 3.4 billion years ago, and biologists have proposed likely chemical processes that could have eventually yielded reproducing molecules inside membranes, the first simple life forms. Fossils show that life developed slowly at first. The pace of evolution quickened about half a billion years ago, when life took on complex forms. Later, when life emerged onto the land, it evolved rapidly into diverse forms. Intelligence is a relatively recent development: It is only a few million years old.

If this evolutionary process occurred on Earth, it seems reasonable that it could have occurred on other worlds as well. Earth-like worlds could be plentiful in the universe. Life may begin and eventually evolve to intelligence on any world where conditions are right. Now make a related argument. **What are the conditions you should expect on other worlds that host life?**

◄ ►

20-3 Communication with Distant Civilizations

Visiting extrasolar planets is, for now, impossible. Nevertheless, if other civilizations exist, it is possible humans can communicate with them. Nature puts restrictions on the pace of such conversations, but the main problem lies in the unknown life expectancy of civilizations.

Travel Between the Stars

The distances between stars are almost beyond comprehension. The space shuttle would take about 150,000 years to reach the nearest star. The obvious way to overcome these huge distances is with tremendously fast spaceships, but even the closest stars are many light-years away.

Nothing can exceed the speed of light, and accelerating a spaceship close to the speed of light takes huge amounts of energy. Even if you travel more slowly, your rocket would still require massive amounts of fuel. If you were piloting a spaceship with the mass of a yacht to the nearest star, 4 light-years away, and you wanted to travel at half the speed of light so as to arrive in 8 years, the trip would require 400 times as much energy as the entire United States consumes in a year.

These limitations not only make it difficult for humans to leave the solar system, but they would also make it difficult for aliens to visit Earth. Reputable scientists have studied "unidentified flying objects" (UFOs) and have never found any evidence that Earth is being visited or has ever been visited by aliens (■ How Do We Know? 20-2). Humans are unlikely ever to meet aliens face to face. However, communication by radio across interstellar distances takes relatively little energy.

Radio Communication

Nature puts restrictions on travel through space, and it also restricts astronomers' ability to communicate with distant civilizations by radio. One restriction is based on simple physics. Radio signals are electromagnetic waves and travel at the speed of light. Due to the distances between the stars, the speed of radio waves would severely limit astronomers' ability to carry on normal conversations with distant civilizations. Decades could elapse between asking a question and getting an answer.

So, rather than try to begin a conversation, one group of astronomers decided in 1974 to broadcast a message of greeting toward the globular cluster M13, 26,000 light-years away, using the Arecibo radio telescope (see Figure 5-20). When the signal arrives 26,000 years in the future, alien astronomers may be able to understand it because the message is **anticoded,** meaning that it is intended to be decoded by beings about whom we know nothing except that they build radio telescopes. The message is a string of 1679 pulses and gaps. Pulses represent 1s, and gaps represent 0s. The string can be arranged in only two possible ways: as 23 rows of 73 or as 73 rows of 23. The second arrangement forms a picture containing information about life on Earth (■ Figure 20-11).

Although the 1974 Arecibo beacon was the only powerful signal sent purposely from Earth to other solar systems, Earth is sending out many other signals. Short-wave radio signals, such as TV and FM, have been leaking into space for the last 50 years or so. Any civilization within 50 light-years could already have detected Earth's civilization. That works both ways: Alien signals, whether intentional messages of friendship or the blather of their equivalent to daytime TV, could be arriving at Earth now. Astronomers all over the world are pointing radio telescopes at the most likely stars and listening for alien civilizations.

Which channels should astronomers monitor? Wavelengths longer than 30 cm would get lost in the background noise of the Milky Way Galaxy, while wavelengths shorter than about 1 cm are mostly absorbed in Earth's atmosphere. Between those wavelengths is a radio window that is open for communication. Even this restricted window contains millions of possible radio-

How Do We Know? | 20-2

UFOs and Space Aliens

Has Earth been visited by aliens? If you conclude that there is likely to be life on other worlds, then you might be tempted to use UFO sightings as evidence to test your hypothesis. Scientists don't do this for two reasons.

First, the reputation of UFO sightings and alien encounters does not inspire confidence that these data are reliable. Most people hear of such events in grocery store tabloids, daytime talk shows, or sensational "specials" on viewer-hungry cable networks. You should take note of the low reputation of the media that report UFOs and space aliens. Most of these reports, like the reports that Elvis is alive and well, are simply made up for the sake of sensation, and you cannot use them as reliable evidence.

Second, the few UFO sightings that are not made up do not survive careful examination. Most are mistakes and unintentional misinterpretations, committed by honest people, of natural events or human-made objects. Over many decades, experts have studied these incidents and found none that are convincing. In short, despite false claims to the contrary on TV shows, there is no dependable evidence that Earth has ever been visited by aliens.

In a way, that's too bad. A confirmed visit by intelligent creatures from beyond our solar system would answer many questions. It would be exciting, enlightening, and, like any real adventure, a bit scary. But scientists must professionally pay attention to what is supported by evidence rather than what might be thrilling. There is not yet any direct evidence of life on other worlds.

Flying saucers from space are fun to think about, but there is no evidence that they are real.

frequency bands and is too wide to monitor easily, but astronomers may have thought of a way to narrow the search. Within this window lie the 21-cm spectral line of neutral hydrogen and the 18-cm line of OH (■ Figure 20-12). The interval between those lines has low background interference and is named the **water hole** because H plus OH yields water. Any civilizations sophisticated enough to do radio astronomy research must know of these lines and might appreciate their significance in the same way as do Earthlings.

A number of searches for extraterrestrial radio signals have been made, and some are now under way. This field of study is known as **SETI, Search for Extra-Terrestrial Intelligence,** and it has generated heated debate among astronomers, philosophers, theologians, and politicians. Congress funded a NASA search for a short time but ended support in the early 1990s because it feared public reaction. In fact, the annual cost of a major search is only about as much as a single Air Force attack helicopter, but much of the reluctance to fund searches probably stems from issues other than cost. One point to keep in mind is that the discovery of real alien intelligence would cause a huge change in humans' worldview, akin to Galileo's discovery that the moons of Jupiter do not go around the Earth. Some turmoil would likely result.

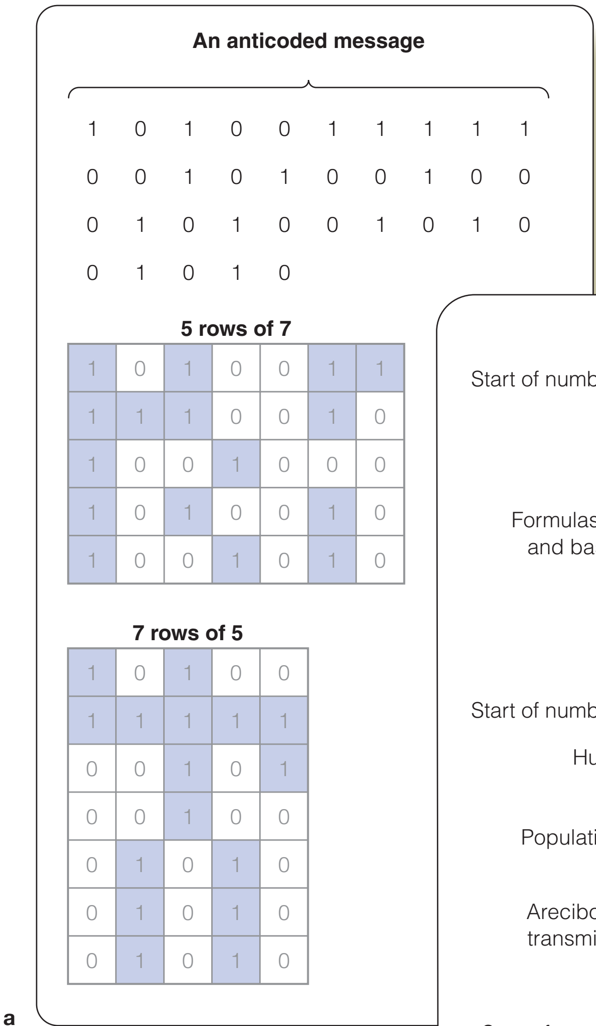
In spite of the controversy, the search continues. The NASA SETI project canceled by Congress was renamed Project Phoenix and completed using private funds. The SETI Institute, founded in 1984, managed Project Phoenix plus several other important searches and is currently building a new radio telescope array in northern California, collaborating with the University of California, Berkeley, and partly funded by Paul Allen of Microsoft.

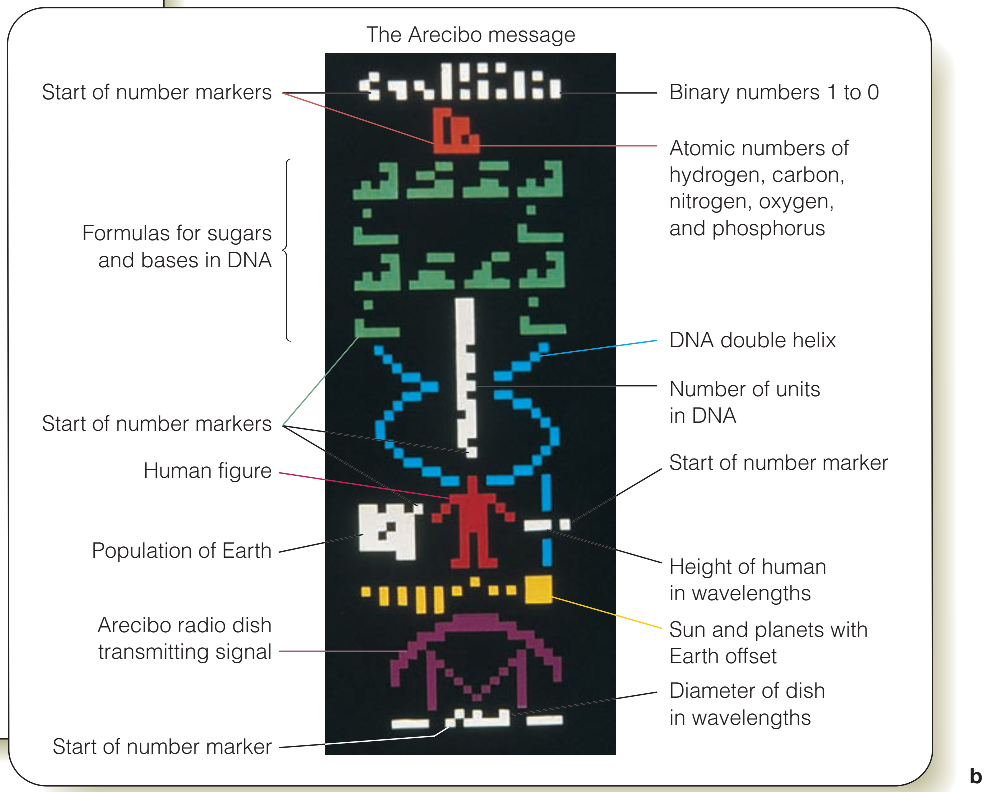
There is even a way for you to help with searches. The Berkeley SETI team (*Note:* they are separate from the SETI Institute), with the support of the Planetary Society, has recruited about 4 million owners of personal computers that are connected to the Internet. You can download a screen saver that searches data files from the Arecibo radio telescope for signals whenever you are not using the computer. For information, locate the seti@home project at http://setiathome.ssl.berkeley.edu/.

The search continues, but radio astronomers struggle to hear anything against the worsening babble of radio noise from human civilization. Wider and wider sections of the electromagnetic spectrum are being used for Earthly communication, and this, combined with stray electromagnetic radiation from electronic devices including everything from computers to refrigerators, makes hearing faint radio signals difficult. It would be ironic if humans fail to detect faint signals from another world because our own world has become too noisy. Ultimately, the chance of success depends on the number of inhabited worlds in the galaxy.

How Many Inhabited Worlds?

Given enough time, the searches will find other worlds with civilizations, assuming that there are at least a few out there. If intelligence is common, scientists should find signals relatively





■ **Figure 20-11**

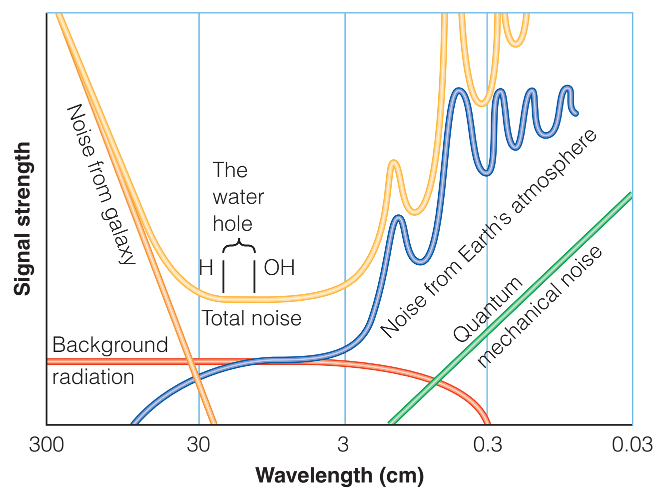
(a) An anticoded message is designed for easy decoding. Here a string of 35 radio pulses, represented as 1s and 0s, can be arranged in only two ways, as 5 rows of 7 or 7 rows of 5. The second way produces a friendly message. (b) The Arecibo message describes life on Earth (color added for clarity). Binary numbers give the height of the human figure (1110) and the diameter of the telescope dish (100101111110) in units of the wavelength of the signal, 12.3 cm. (NASA) **Animated!**

soon—within the next few decades—but if intelligence is rare, it may take much longer.

Simple arithmetic can give you an estimate of the number of technological civilizations in the Milky Way Galaxy with which you might communicate, N_c. The formula proposed for discussions about N_c is named the **Drake equation** after the radio astronomer Frank Drake, a pioneer in the search for extraterrestrial intelligence. The version of the Drake equation presented here is modified slightly from its original form:

$$N_c = N_* \cdot f_P \cdot n_{HZ} \cdot f_L \cdot f_I \cdot f_S$$

N_* is the number of stars in our galaxy, and f_P represents the fraction of stars that have planets. If all single stars have planets, f_P is about 0.5. The factor n_{HZ} is the average number of planets in each planetary system suitably located in the habitable zone—meaning for the sake of the present discussion, the number of planets per planetary system possessing liquid water. Europa and Enceladus in our solar system show that liquid water can exist due to tidal heating outside the conventional habitable zone that in our system contains Earth's orbit. Thus, n_{HZ} may be



■ **Figure 20-12**

Radio noise from various astronomical sources makes it difficult to detect distant signals at wavelengths longer than 30 cm or shorter than 1 cm. In this range, wavelengths of radio emission lines from H atoms and from OH molecules mark a small wavelength range named the water hole that may be a likely channel for communication.

larger than had been originally thought. The factor f_L is the fraction of suitable planets on which life begins, and f_I is the fraction of those planets where life evolves to intelligence.

The six factors on the right-hand side of the Drake equation can be roughly estimated, with decreasing certainty as you proceed from left to right, The final factor is extremely uncertain. That factor f_S is the fraction of a star's life during which an intelligent species is communicative. If a society survives at a technological level for only 100 years, the chances of communicating with it are small. On the other hand, a society that stabilizes and remains technological for a long time is much more likely to be detected. For a star with a life span of 10 billion years, f_S might conceivably range from 10^{-8} for extremely short-lived societies to 10^{-4} for societies that survive for a million years. ■ Table 20-1 summarizes what many scientists consider a reasonable range of values for f_S and the other factors.

If the optimistic estimates are true, there could be a communicative civilization within a few tens of light-years from Earth. On the other hand, if the pessimistic estimates are true, Earth may be the only planet that is capable of communication within thousands of the nearest galaxies.

◄ SCIENTIFIC ARGUMENT ►

***Why does the number of civilizations that could be detected* depend on how long civilizations survive at a technological level?**

This scientific argument depends on the timing of events. If you turned a radio telescope to the sky and scanned millions of frequency bands for many stars, you would be taking a snapshot of the universe at a particular time. Broadcasts from other civilizations must be arriving at that same time if they are to be detected. If most civilizations survive for a long time, there is a much greater chance that you will detect one of them in your snapshot than if civilizations tend to disappear quickly due to nuclear war or environmental collapse. If most civilizations last only a short time, there may be none capable of transmitting during the short interval when Earthlings are capable of building radio telescopes to listen for them.

The speed at which astronomers can search for signals is limited because computers must search many frequency intervals, but not all frequencies inside Earth's radio window are subject to intensive search. Build a new argument to explain: **Why is the water hole an especially good frequency band in which to listen?**

◄ ►

■ **Table 20-1** | **The Number of Technological Civilizations per Galaxy**

Estimates	Variables	Pessimistic	Optimistic
N_*	Number of stars per galaxy	2×10^{11}	2×10^{11}
f_P	Fraction of stars with planets	0.1	0.5
n_{HZ}	Number of planets per star that lie in habitable zone for longer than 4 billion years	0.01	1
f_L	Fraction of suitable planets on which life begins	0.01	1
f_I	Fraction of planets with life where life forms evolve to intelligence	0.01	1
f_S	Fraction of star's existence during which a technological society survives	10^{-8}	10^{-4}
N_c	Number of communicative civilizations per galaxy	2×10^{-4}	1×10^{7}

What Are We? Matter and Spirit

There are over 4000 religions around the world, and nearly all hold that humans have a dual nature. We are physical objects made of atoms, but we are also spiritual beings. Science is unable to examine the spiritual side of existence, but it can tell us about our physical nature.

The matter you are made of appeared in the big bang and was cooked into a wide range of elements inside stars. Your atoms may have been inside at least two or three generations of stars. Eventually, your atoms became part of a nebula that contracted to form our sun and the planets of the solar system

Your atoms have been part of Earth for the last 4.6 billion years. They have been recycled many times through dinosaurs, stromatolites, fish, bacteria, grass, birds, worms, and other living things. You are using your atoms now, but when you are done with them, they will go back to Earth and be used again and again.

When the sun swells into a red giant star and dies in a few billion years, Earth's atmosphere and oceans will be driven away, and at least the outer few kilometers of Earth's crust will be vaporized and blown outward to become part of the nebula around the white-dwarf remains of the sun. Your atoms are destined to return to the interstellar medium and will become part of future generations of stars and planets.

The message of astronomy is that humans are not just observers: We are participants in the universe. Among all of the galaxies, stars, planets, planetesimals, and bits of matter, humans are objects that can think, and that means we can understand what we are.

Is the human race the only thinking species? If so, we bear the sole responsibility to understand and admire the universe. The detection of signals from another civilization would demonstrate that we are not alone, and such communication would end the self-centered isolation of humanity and stimulate a reevaluation of the meaning of human existence. We may never realize our full potential as humans until we communicate with nonhuman intelligent life.

Study and Review

Summary

- The process of life extracts energy from the surroundings, maintains an organism, and modifies the surroundings to promote the organism's survival.
- Living things have a physical basis—the arrangement of matter and energy that makes life possible. Life on Earth is based on carbon chemistry.
- Living things must also have a controlling unit of information that can be passed to successive generations.
- Genetic information for life on Earth is stored in long carbon-chain molecules such as **DNA (deoxyribonucleic acid) (p. 428).**
- The DNA molecule stores information in the form of chemical bases linked together like the rungs of a ladder. Copied by the **RNA (ribonucleic acid) (p. 429)** molecule, the patterns of bases act as recipes for connecting together **amino acid (p. 428)** subunits to construct **proteins (p. 428),** including **enzymes (p. 428),** that are respectively the main structural and control components of the life process.
- The unit of heredity is a **gene (p. 429),** a piece or several pieces of DNA that in most cases specifies the construction of just one particular type of protein molecule. Genes are connected together in structures called **chromosomes (p. 429),** which are essentially single long DNA molecules.
- When a cell divides, the chromosomes split lengthwise and duplicate themselves so that each of the new cells can receive a copy of the genetic information.
- **Biological evolution (p. 427)** begins when molecules begin reproducing.
- Errors in duplication or damage to the DNA molecule can produce **mutants (p. 430),** organisms that contain new DNA information and therefore have new properties. Variation in genetic codes can become widespread among individuals in a species. **Natural selection (p. 430)** determines which of these variations are best suited to survive, and the species evolves to fit its environment.
- The oldest definitely identified fossils on Earth, structures called **stromatolites (p. 430)** that are composed of stacks of bacterial mats and sediments, are at least 3.4 billion years old. Those fossils provide evidence that life began in the oceans.
- Fossil evidence indicates that life began on Earth as simple single-celled organisms like bacteria and much later evolved into more complex, **multicellular (p. 432)** creatures.
- The **Miller experiment (p. 430)** shows that the chemical building blocks of life form naturally under a wide range of circumstances.
- Biologists hypothesize that **chemical evolution (p. 431)** concentrated simple molecules into a diversity of larger stable organic molecules dissolved in the young Earth's oceans, but those molecules did not reproduce copies of themselves. The hypothetical organic-rich water is sometimes referred to as the **primordial soup (p. 431).**
- Life forms did not become large and complex until about 0.5 billion years ago, during what is called the **Cambrian explosion (p. 433).**
- Life emerged from the oceans only about 0.4 billion years ago, and human intelligence developed over the last 4 million years.
- Life as it is known on Earth requires liquid water and thus a specific range of temperatures.
- No other planet in our solar system appears to harbor life at present. Most are too hot or too cold, although life might have begun on Mars before it became too cold and dry.
- Liquid water exists, and therefore Earth-like life is at least possible, under the surfaces of Jupiter's moon Europa and Saturn's moon Enceladus. Saturn's moon Titan has abundant organic compounds but does not have liquid water.
- Because the origin of life and its evolution into intelligent creatures took so long on Earth, scientists do not consider short-lived middle- and upper-main-sequence stars as likely homes for life.
- Main-sequence G and K stars are thought to be likely candidates to host planets with life. Scientists are not sure whether the fainter M stars are also good candidates or not.
- The **habitable zone (p. 436)** around a star, within which planets can have liquid water on their surfaces, may be larger than scientists had expected, given the wide variety of living things now found in extreme environments on Earth and the possibility of tidal heating of moons orbiting large planets.
- Because of distance, speed, and fuel, travel between the stars seems almost impossible for humans, or for aliens who might visit Earth.
- Radio communication between planetary systems may be possible, but a real conversation would be difficult because of very long travel times for radio signals.
- Broadcasting a radio beacon of pulses would distinguish the signal from naturally occurring radio emission and identify the source as a technological civilization. The signal can be **anticoded (p. 437)** in the hope it would be easy for another civilization to decode.
- One good part of the radio spectrum for communication is called the **water hole (p. 438),** the wavelength range from the 21-cm spectral line of hydrogen to the 18-cm line of OH. Even so, millions of radio wavelengths need to be tested to fully survey the water hole for a given target star.
- Sophisticated searches are now underway to detect radio transmissions from civilizations on other worlds, but such **SETI (Search for Extra-Terrestrial Intelligence) (p. 438)** programs are hampered by limited computer power and radio noise pollution from human civilization.
- The number of civilizations in our galaxy that are at a technological level and able to communicate while humans are listening can be estimated by the **Drake equation (p. 439);** this number is limited primarily by the lifetimes of their and our civilizations.

Review Questions

To assess your understanding of this chapter's topics with additional quizzing and animations, go to **academic.cengage.com/astronomy/seeds**

1. If life is based on information, what is that information?
2. What would happen to a life form if the information handed down to offspring was always the same? How would that endanger the future of the life form?
3. How does the DNA molecule produce a copy of itself?
4. Give an example of natural selection acting on new DNA patterns to select the most advantageous characteristics.
5. What evidence do scientists have that life on Earth began in the sea?
6. Why do scientists think that liquid water is necessary for the origin of life?
7. What is the difference between chemical evolution and biological evolution?
8. What is the significance of the Miller experiment?
9. How does intelligence make a creature more likely to survive?
10. Why are upper-main-sequence (high-luminosity) stars unlikely sites for intelligent civilizations?
11. Why is it reasonable to suspect that travel between stars is nearly impossible?
12. How does the stability of technological civilizations affect the probability that Earth can communicate with them?

13. What is the water hole, and why would it be a good "place" to look for other civilizations?
14. **How Do We Know?** How do science and religion have complimentary explanations of the world?
15. **How Do We Know?** Why are scientists sure Earth has never been visited by aliens?

Discussion Questions

1. Do you expect hypothetical alien recipients of the Arecibo message will be able to decode it? Why or why not?
2. What do you think it would mean if decades of careful searches for radio signals for extraterrestrial intelligence turn up nothing?

Problems

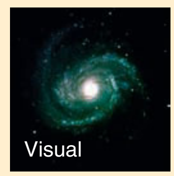
1. A single human cell encloses about 1.5 m of DNA, containing 4.5 billion base pairs. What is the spacing between these base pairs in nanometers? That is, how far apart are the rungs on the DNA ladder?
2. If you represent the history of the Earth by a line 1 m long, how long a segment would represent the 400 million years since life moved onto the land? How long a segment would represent the 4-million-year history of human life?
3. If a human generation, the average time from birth to childbearing, has been 20 years long, how many generations have passed in the last 1 million years?
4. If a star must remain on the main sequence for at least 5 billion years for life to evolve to intelligence, what is the most massive a star can be and still possibly harbor intelligent life on one of its planets? (*Hint:* See Reasoning with Numbers 9-1.)
5. If there are about 1.4×10^{-4} stars like the sun per cubic light-year, how many lie within 100 light-years of Earth? (*Hint:* The volume of a sphere is $\frac{4}{3}\pi r^3$.)
6. Mathematician Karl Gauss suggested planting forests and fields in a gigantic geometric proof to signal to possible Martians that intelligent life exists on Earth. If Martians had telescopes that could resolve details no smaller than 1 arc second, how large would the smallest element of Gauss's signal have to be for it to be visible at Mars's closest approach to Earth? (*Hint:* See Reasoning with Numbers 3-1 and Appendix A.)
7. If you detected radio signals with an average wavelength of 20 cm and suspected that they came from a civilization on a distant planet, roughly how much of a change in wavelength should you expect to see because of the orbital motion of the distant planet? (*Hint:* See Reasoning with Numbers 6-2.)
8. Calculate the number of communicative civilizations per galaxy using your own estimates of the factors in Table 20-1.

Learning to Look

1. The star cluster shown in the image to the right, contains cool red giants and main-sequence stars from hot blue stars all the way down to red dwarfs. Discuss the likelihood that planets orbiting any of these stars might be home to life. (*Hint:* Estimate the age of the cluster.)


NASA

2. If you could search for life in the galaxy shown in the image to the right, would you look among disk stars or halo stars? Discuss the factors that influence your decision.



ESO

Afterword

The aggregate of all our joys and sufferings, thousands of confident religions, ideologies and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilizations, every king and peasant, every young couple in love, every hopeful child, every mother and father, every inventor and explorer, every teacher of morals, every corrupt politician, every superstar, every supreme leader, every saint and sinner in the history of our species, lived there on a mote of dust, suspended in a sunbeam.

CARL SAGAN (1934–1996)

Earth photographed by Voyager 1 from the edge of the solar system. NASA

Our journey together is over, but before we part company, let's ponder one final time the primary theme of this book—humanity's place in the physical universe. Astronomy gives us some comprehension of the workings of stars, galaxies, and planets, but its greatest value lies in what it teaches us about ourselves. Now that you have surveyed astronomical knowledge, you can better understand your own position in nature.

To some, the word *nature* conjures up visions of furry rabbits hopping about in a forest glade. To others, nature is the blue-green ocean depths, and still others think of nature as windswept mountaintops. As diverse as these images are, they are all Earth-bound. Having studied astronomy, you can see nature as a beautiful mechanism composed of matter and energy, interacting according to simple rules, forming galaxies, stars, planets, mountaintops, ocean depths, forest glades, and people.

Perhaps the most important astronomical lesson is that humanity is a small but important part of the universe. Most of the universe is probably lifeless. The vast reaches between the galaxies appear to be empty of all but the thinnest gas, and stars are much too hot to preserve the chemical bonds that seem necessary for life to survive and develop. It seems that only on the surfaces of a few planets, where temperatures are moderate, can atoms link together in special ways to form living matter.

If life is special, then intelligence is precious. The universe must contain many planets devoid of life, planets where sunlight has shined unfelt for billions of years. There may also exist planets on which life has developed but has not become complex, planets where the wind stirs wide plains of grass and rustles through dark forests. On some planets, creatures resembling Earth's insects, fish, birds, and animals may watch the passing days only dimly aware of their own existence. It is intelligence, human or alien, that gives meaning to the landscape.

Science is the process by which Earth's intelligence has tried to understand the physical universe. Science is not the invention of new devices or processes. It does not create home computers, cure the mumps, or manufacture plastic spoons—those are engineering and technology, the adaptation of scientific understanding for practical purposes. Science is the understanding of nature, and astronomy is that understanding on the grandest scale. Astronomy is the science by which the universe, through its intelligent lumps of matter, tries to understand its own existence.

As the primary intelligent species on this planet, we are the custodians of a priceless gift—a planet filled with living things. This is especially true if life is rare in the universe. In fact, if Earth is the only inhabited planet, our responsibility is overwhelming. We are the only creatures who can take action to preserve the existence of life on Earth, and, ironically, our own actions are the most serious hazards.

The future of humanity is not secure. We are trapped on a tiny planet with limited resources and a population growing faster than our ability to produce food. We have already driven some creatures to extinction and now threaten others. We are changing the climate of our planet in ways we do not fully understand. Even if we reshape our civilization to preserve our world, the sun's evolution will eventually destroy Earth.

This may be a sad prospect, but a few factors are comforting. First, everything in the universe is temporary. Stars die, galaxies die; perhaps the entire universe will someday end. Our distant future is limited, and this assures us that we are a part of a much larger whole. Second, we have a few billion years to prepare, and a billion years is a very long time. Only a few million years ago, our ancestors were starting to walk upright and communicate. A billion years ago, our ancestors were microscopic organisms living in the oceans. To suppose that a billion years hence there will be beings resembling today's humans, or that humans will still be the dominant intelligence on Earth, or that humans will even exist, are ultimately conceits.

Our responsibility is not to save our race for all eternity but to behave as dependable custodians of our planet, preserving it, admiring it, and trying to understand it. That calls for drastic changes in our behavior toward other living things and a revolution in our attitude toward our planet's resources. Whether we can change our ways is debatable—humanity is far from perfect in its understanding, abilities, or intentions. However, you must not imagine that we, and our civilization, are less than precious. We have the gift of intelligence, and that is the finest thing this planet has ever produced.

Appendix A
Units and Astronomical Data

Introduction

The metric system is used worldwide as the system of units, not only in science but also in engineering, business, sports, and daily life. Developed in 18th-century France, the metric system has gained acceptance in almost every country in the world because it simplifies computations.

A system of units is based on the three fundamental units for length, mass, and time. Other quantities, such as density and force, are derived from these fundamental units. In the English (or British) system of units (commonly used only in the United States, Tonga, and Southern Yemen, but, ironically, not in Great Britain) the fundamental unit of length is the foot, composed of 12 inches. The metric system is based on the decimal system of numbers, and the fundamental unit of length is the meter, composed of 100 centimeters.

Because the metric system is a decimal system, it is easy to express quantities in larger or smaller units as is convenient. You can give distances in centimeters, meters, kilometers, and so on. The prefixes specify the relation of the unit to the meter. Just as a cent is 1/100 of a dollar, so a centimeter is 1/100 of a meter. A kilometer is 1000 m, and a kilogram is 1000 g. The meanings of the commonly used prefixes are given in ■ Table A-1.

The SI Units

Any system of units based on the decimal system would be easy to use, but by international agreement, the preferred set of units, known as the *Système International d'Unités* (SI units) is based on the meter, kilogram, and second. These three fundamental units define the rest of the units, as given in ■ Table A-2.

The SI unit of force is the newton (N), named after Isaac Newton. It is the force needed to accelerate a 1 kg mass by 1 m/s^2, or the force roughly equivalent to the weight of an apple at Earth's surface. The SI unit of energy is the joule (J), the energy produced by a force of 1 N acting through a distance of 1 m. A joule is roughly the energy in the impact of an apple falling off a table.

Exceptions

Units can help you in two ways. They make it possible to make calculations, and they can help you to conceive of certain quantities. For calculations, the metric system is far superior, and it is used for calculations throughout this book.

Americans commonly use the English system of units, so for conceptual purposes this book also expresses quantities in English units. Instead of saying the average person would weigh 133 N on the moon, it might be more helpful to some readers for that weight to be expressed as 30 lb. Consequently, this text commonly gives quantities in metric form followed by the English form in parentheses: The radius of the moon is 1738 km (1080 mi).

In SI units, density should be expressed as kilograms per cubic meter, but no human hand can enclose a cubic meter, so that unit does not help you grasp the significance of a given density. This book refers to density in grams per cubic centimeter. A gram is roughly the mass of a paperclip, and a cubic centimeter is the size of a small sugar cube, so you can easily conceive of a density of 1 g/cm^3, roughly the density of water. This is not a

■ **Table A-1** | **Metric Prefixes**

Prefix	Symbol	Factor
Mega	M	10^6
Kilo	k	10^3
Centi	c	10^{-2}
Milli	m	10^{-3}
Micro	μ	10^{-6}
Nano	n	10^{-9}

■ **Table A-2** | **SI *(Système International)* Metric Units**

Quantity	SI Unit
Length	Meter (m)
Mass	Kilogram (kg)
Time	Second (s)
Force	Newton (N)
Energy	Joule (J)

bothersome departure from SI units because you will not have to make complex calculations using density.

Conversions

To convert from one metric unit to another (from meters to kilometers, for example), you have only to look at the prefix. However, converting from metric to English or English to metric is more complicated. The conversion factors are given in ■ Table A-3.

Example: The radius of the moon is 1738 km. What is this in miles? Table A-3 indicates that 1.000 mile equals 1.609 km, so

$$1738 \text{ km} \times \frac{1.000 \text{ mi}}{1.609 \text{ km}} = 1080 \text{ mi}$$

Temperature Scales

In astronomy, as in most other sciences, temperatures are expressed on the Kelvin scale, although the centigrade (or Celsius) scale is also used. The Fahrenheit scale commonly used in the United States is not used in scientific work.

Temperatures on the Kelvin scale are measured from absolute zero, the temperature of an object that contains no extractable heat. In practice, no object can be as cold as absolute zero, although laboratory apparatuses have reached temperatures lower than 10^{-6} K. The Kelvin scale is named after the Scottish mathematical physicist William Thomson, Lord Kelvin (1824–1907).

The centigrade scale refers temperatures to the freezing point of water (0°C) and to the boiling point of water (100°C). One degree Centigrade is $^1/_{100}$, the temperature difference between the freezing and boiling points of water, thus the prefix *centi.* The centigrade scale is also called the Celsius scale after its inventor, the Swedish astronomer Anders Celsius (1701–1744).

The Fahrenheit scale fixes the freezing point of water at 32°F and the boiling point at 212°F. Named after the German physicist Gabriel Daniel Fahrenheit (1686–1736), who made the first successful mercury thermometer in 1720, the Fahrenheit scale is used routinely only in the United States.

It is easy to convert temperatures from one scale to another using the information given in ■ Table A-4.

Powers of 10 Notation

Powers of 10 make writing very large numbers much simpler. For example, the nearest star is about 43,000,000,000,000 km from the sun. Writing this number as 4.3 x 10^{13} km is much easier.

Very small numbers can also be written with powers of 10. For example, the wavelength of visible light is about 0.0000005 m. In powers of 10 this becomes 5×10^{-7} m.

The powers of 10 used in this notation appear below. The exponent tells you how to move the decimal point. If the exponent is positive, move the decimal point to the right. If the exponent is negative, move the decimal point to the left. For example, 2.0000×10^3 equals 2000, and 2×10^{-3} equals 0.002.

$$10^5 = 100{,}000$$
$$10^4 = 10{,}000$$
$$10^3 = 1{,}000$$
$$10^2 = 100$$
$$10^1 = 10$$
$$10^0 = 1$$
$$10^{-1} = 0.1$$
$$10^{-2} = 0.01$$
$$10^{-3} = 0.001$$
$$10^{-4} = 0.0001$$

If you use scientific notation in calculations, be sure you correctly enter the numbers into your calculator. Not all calculators accept scientific notation, but those that can have a key labeled EXP, EEX, or perhaps EE that allows you to enter the ex-

■ **Table A-3 | Conversion Factors Between British and Metric Units**

1 inch = 2.54 centimeters	1 centimeter = 0.394 inch
1 foot = 0.3048 meter	1 meter = 39.36 inches = 3.28 feet
1 mile = 1.609 kilometers	1 kilometer = 0.6214 mile
1 slug = 14.59 kilograms	1 kilogram = 0.0685 slug
1 pound = 4.448 Newtons	1 Newton = 0.2248 pound
1 foot-pound = 1.356 Joules	1 Joule = 0.7376 foot-pound
1 horsepower = 745.7 Joules/s	1 Joule/s = 0.001341 horsepower
	1 Joule/s = 1 Watt

■ **Table A-4 | Temperature Scales and Conversion Formulas**

	Kelvin (K)	Centigrade (°C)	Fahrenheit (°F)
Absolute zero	0 K	−273°C	−459°F
Freezing point of water	273 K	0°C	32°F
Boiling point of water	373 K	100°C	212°F

Conversions:

$K = °C + 273$

$°C = {}^5/_9\,(°F - 32)$

$°F = {}^9/_5\,(°C) + 32$

ponent of ten. To enter a number such as 3×10^8, press the keys 3 EXP 8. To enter a number with a negative exponent, you must use the change-sign key, usually labeled +/− or CHS. To enter the number 5.2×10^{-3}, press the keys 5.2 EXP +/− 3. Try a few examples.

To read a number in scientific notation from a calculator you must read the exponent separately. The number 3.1×10^{25} may appear in a calculator display as 3.1 25 or on some calculators as 3.1 10^{25}. Examine your calculator to determine how such numbers are displayed.

Astronomy Units and Constants

Astronomy, and science in general, is a way of learning about nature and understanding the universe. To test hypotheses about how nature works, scientists use observations of nature. The tables that follow contain some of the basic observations that support science's best understanding of the astronomical universe. Of course, these data are expressed in the form of numbers, not because science reduces all understanding to mere numbers, but because the struggle to understand nature is so demanding that science must use every valid means available. Quantitative thinking—reasoning mathematically—is one of the most powerful techniques ever invented by the human brain. Thus, these tables are not nature reduced to mere numbers but numbers supporting humanity's growing understanding of the natural world around us.

Table A-5 | Astronomical Constants

Velocity of light (c)	$= 3.00 \times 10^8$ m/s
Gravitational constant (G)	$= 6.67 \times 10^{-11}$ m^3/s^{2}kg
Mass of H atom	$= 1.67 \times 10^{-27}$ kg
Mass of Earth ($M_{\oplus}$)	$= 5.98 \times 10^{24}$ kg
Earth equatorial radius ($R_{\oplus}$)	$= 6.38 \times 10^3$ km
Mass of sun ($M_{\odot}$)	$= 1.99 \times 10^{30}$ kg
Radius of sun ($R_{\odot}$)	$= 6.96 \times 10^8$ m
Solar luminosity ($L_{\odot}$)	$= 3.83 \times 10^{26}$ J/s
Mass of moon	$= 7.35 \times 10^{22}$ kg
Radius of moon	$= 1.74 \times 10^3$ km

Table A-6 | Units Used in Astronomy

1 Angstrom (Å)	$= 10^{-8}$ cm
	$= 10^{-10}$ m
	= 10 nm
1 astronomical unit (AU)	$= 1.50 \times 10^{11}$ m
	$= 93.0 \times 10^6$ mi
1 light-year (ly)	$= 6.32 \times 10^4$ AU
	$= 9.46 \times 10^{15}$ m
	$= 5.88 \times 10^{12}$ mi
1 parsec (pc)	$= 2.06 \times 10^5$ AU
	$= 3.09 \times 10^{16}$ m
	= 3.26 ly
1 kiloparsec (kpc)	= 1000 pc
1 megaparsec (Mpc)	= 1,000,000 pc

Table A-7 | Properties of Main-Sequence Stars

Spectral Type	Absolute Visual Magnitude (M_v)	Luminosity*	Temp. (K)	λ_{max} (nm)	Mass*	Radius*	Average Density (g/cm^3)
O5	−5.8	501,000	40,000	72.4	40	17.8	0.01
B0	−4.1	20,000	28,000	100	18	7.4	0.1
B5	−1.1	790	15,000	190	6.4	3.8	0.2
A0	+0.7	79	9900	290	3.2	2.5	0.3
A5	+2.0	20	8500	340	2.1	1.7	0.6
F0	+2.6	6.3	7400	390	1.7	1.4	1.0
F5	+3.4	2.5	6600	440	1.3	1.2	1.1
G0	+4.4	1.3	6000	480	1.1	1.0	1.4
G5	+5.1	0.8	5500	520	0.9	0.9	1.6
K0	+5.9	0.4	4900	590	0.8	0.8	1.8
K5	+7.3	0.2	4100	700	0.7	0.7	2.4
M0	+9.0	0.1	3500	830	0.5	0.6	2.5
M5	+11.8	0.01	2800	1000	0.2	0.3	10.0
M8	+16	0.001	2400	1200	0.1	0.1	63

*Luminosity, mass, and radius are given in terms of the sun's luminosity, mass, and radius.

Table A-8 | The 15 Brightest Stars

Star	Name	Apparent Visual Magnitude (m_v)	Spectral Type	Absolute Visual Magnitude (M_v)	Distance (ly)
α CMa A	Sirius	−1.47	A1	1.4	8.7
α Car	Canopus	−0.72	F0	−3.1	98
α Cen	Rigil Kentaurus	−0.01	G2	4.4	4.3
α Boo	Arcturus	−0.06	K2	−0.3	36
α Lyr	Vega	0.04	A0	0.5	26.5
α Aur	Capella	0.05	G8	−0.6	45
β Ori A	Rigel	0.14	B8	−7.1	900
α CMi A	Procyon	0.37	F5	2.7	11.3
α Ori	Betelgeuse	0.41	M2	−5.6	520
α Eri	Achernar	0.51	B3	−2.3	118
β Cen AB	Hadar	0.63	B1	−5.2	490
α Aql	Altair	0.77	A7	2.2	16.5
α Tau A	Aldebaran	0.86	K5	−0.7	68
α Cru	Acrux	0.90	B2	3.5	260
α Vir	Spica	0.91	B1	−3.3	220

Table A-9 | The 15 Nearest Stars

Name	Absolute Magnitude (M_v)	Distance (ly)	Spectral Type	Apparent Visual Magnitude (m_v)
Sun	4.83		G2	−26.8
Proxima Cen	15.45	4.28	M5	11.05
α Cen A	4.38	4.3	G2	0.1
α Cen B	5.76	4.3	K5	1.5
Barnard's Star	13.21	5.9	M5	9.5
Wolf 359	16.80	7.6	M6	13.5
Lalande 21185	10.42	8.1	M2	7.5
Sirius A	1.41	8.6	A1	−1.5
Sirius B	11.54	8.6	white dwarf	7.2
Luyten 726-8A	15.27	8.9	M5	12.5
Luyten 726-8B (UV Cet)	15.8	8.9	M6	13.0
Ross 154	13.3	9.4	M5	10.6
Ross 248	14.8	10.3	M6	12.2
ϵ Eri	6.13	10.7	K2	3.7
Luyten 789-6	14.6	10.8	M7	12.2

Table A-10 | Properties of the Planets

PHYSICAL PROPERTIES (Earth = ⊕)

Planet	Equatorial Radius (km)	Equatorial Radius (⊕ = 1)	Mass (⊕ = 1)	Average Density (g/cm^3)	Surface Gravity (⊕ = 1)	Escape Velocity (km/s)	Sidereal Period of Rotation	Inclination of Equator to Orbit
Mercury	2439	0.38	0.056	5.44	0.38	4.3	58.65d	0°
Venus	6052	0.95	0.815	5.24	0.90	10.3	243.02d	177°
Earth	6378	1.00	1.000	5.50	1.00	11.2	23.93h	23°.5
Mars	3396	0.53	0.108	3.94	0.38	5.0	24.62h	25°.3
Jupiter	71,494	11.20	317.8	1.34	2.54	61	9.92h	3°.1
Saturn	60,330	9.42	95.2	0.69	1.16	35.6	10.23h	26°.4
Uranus	25,559	4.01	14.5	1.19	0.92	22	17.23h	97°.9
Neptune	24,750	3.93	17.2	1.66	1.19	25	16.05h	28°.8

ORBITAL PROPERTIES

Planet	Semimajor Axis (*a*) (AU)	Semimajor Axis (*a*) (10^6 km)	Orbital Period (*P*) (y)	Orbital Period (*P*) (days)	Average Orbital Velocity (km/s)	Orbital Eccentricity	Inclination to Ecliptic
Mercury	0.39	57.9	0.24	87.97	47.89	0.206	7°.0
Venus	0.72	108.2	0.62	224.68	35.03	0.007	3°.4
Earth	1.00	149.6	1.00	365.26	29.79	0.017	0° *
Mars	1.52	227.9	1.88	686.95	24.13	0.093	1°.9
Jupiter	5.20	778.3	11.87	4334.3	13.06	0.048	1°.3
Saturn	9.54	1427.0	29.46	10,760	9.64	0.056	2°.5
Uranus	19.18	2869.0	84.01	30,685	6.81	0.046	0°.8
Neptune	30.06	4497.1	164.79	60,189	5.43	0.010	1°.8

* By definition.

Table A-11 | Principal Satellites of the Solar System

Planet	Satellite	Radius (km)	Distance from Planet (10^3 km)	Orbital Period (days)	Orbital Eccentricity	Orbital Inclination
Earth	Moon	1738	384.4	27.32	0.055	5°.1
Mars	Phobos	14×12×10	9.4	0.32	0.018	1°.0
	Deimos	8×6×5	23.5	1.26	0.002	2°.8
Jupiter	Amalthea	135×100×78	182	0.50	0.003	0°.4
	Io	1820	422	1.77	0.000	0°.3
	Europa	1565	671	3.55	0.000	0°.5
	Ganymede	2640	1071	7.16	0.002	0°.2
	Callisto	2420	1884	16.69	0.008	0°.2
	Himalia	~85	11,470	250.6	0.158	27°.6
Saturn	Janus	110×80×100	151.5	0.70	0.007	0°.1
	Mimas	196	185.5	0.94	0.020	1°.5
	Enceladus	250	238.0	1.37	0.004	0°.0
	Tethys	530	294.7	1.89	0.000	1°.1
	Dione	560	377	2.74	0.002	0°.0
	Rhea	765	527	4.52	0.001	0°.4
	Titan	2575	1222	15.94	0.029	0°.3
	Hyperion	205×130×110	1484	21.28	0.104	~0°.5
	Iapetus	720	3562	79.33	0.028	14°.7
	Phoebe	110	12,930	550.4	0.163	150°
Uranus	Miranda	242	129.9	1.41	0.017	3°.4
	Ariel	580	190.9	2.52	0.003	0°
	Umbriel	595	266.0	4.14	0.003	0°
	Titania	805	436.3	8.71	0.002	0°
	Oberon	775	583.4	13.46	0.001	0°
Neptune	Proteus	205	117.6	1.12	~0	~0°
	Triton	1352	354.59	5.88	0.00	160°
	Nereid	170	5588.6	360.12	0.76	27.7°

Table A-12 | Meteor Showers

Shower	Dates	Hourly Rate	Radiant of shower		Associated Comet
			R.A.	Dec.	
Quadrantids	Jan. 2–4	30	15h24m	50°	
Lyrids	April 20–22	8	18h4m	33°	1861 I
η Aquarids	May 2–7	10	22h24m	0°	Halley?
δ Aquarids	July 26–31	15	22h36m	−10°	
Perseids	Aug. 10–14	40	3h4m	58°	1982 III
Orionids	Oct. 18–23	15	6h20m	15°	Halley?
Taurids	Nov. 1–7	8	3h40m	17°	Encke
Leonids	Nov. 14–19	6	10h12m	22°	1866 I Temp
Geminids	Dec. 10–13	50	7h28m	32°	

Table A-13 | Greatest Elongations of Mercury*

Evening Sky	Morning Sky
Jan. 4, 2009	Feb. 13, 2009
Apr. 26, 2009**	June 13, 2009
Aug. 24, 2009	Oct. 6, 2009**
Dec. 18, 2009	Jan. 27, 2010
Apr. 8, 2010**	May 26, 2010
Aug. 7, 2010	Sept. 19, 2010**
Dec. 1, 2010	Jan. 9, 2011
March 23, 2011**	May 7, 2011
July 20, 2011	Sept. 3, 2011**
Nov. 14, 2011	Dec. 23, 2011
March 5, 2012**	April 18, 2012
July 1, 2012	Aug. 16, 2012
Oct. 26, 2012	Dec. 4, 2012

*Elongation is the angular distance from the sun to a planet.

**Most favorable elongations.

Table A-14 | Greatest Elongations of Venus

Evening Sky	Morning Sky
Jan. 14, 2009	June 5, 2009
Aug. 20, 2010	Jan. 8, 2011
March 27, 2012	Aug. 15, 2012
Nov. 1, 2013	March 22, 2014
June 6, 2015	Oct. 26, 2015
Jan. 12, 2017	June 3, 2017
Aug. 17, 2018	Jan. 6, 2019

Table A-15 | The Greek Alphabet

A, α	alpha	H, η	eta	N, ν	nu	T, τ	tau
B, β	beta	Θ, θ	theta	Ξ, ξ	xi	Υ, υ	upsilon
Γ, γ	gamma	I, ι	iota	O, o	omicron	Φ, ϕ	phi
Δ, δ	delta	K, κ	kappa	Π, π	pi	X, χ	chi
E, ϵ	epsilon	Λ, λ	lambda	P, ρ	rho	Ψ, ψ	psi
Z, ζ	zeta	M, μ	mu	Σ, σ	sigma	Ω, ω	omega

■ Table A-16 | Periodic Table of the Elements

Atomic number → 11; Symbol → Na; Atomic mass → 22.99

Atomic masses are based on carbon-12. Numbers in parentheses are mass numbers of most stable or best-known isotopes of radioactive elements.

Group / Period; Transition Elements (groups 3–12); Noble Gases (18)

Period	IA(1)	IIA(2)	IIIB(3)	IVB(4)	VB(5)	VIB(6)	VIIB(7)	VIII (8)	VIII (9)	VIII (10)	IB(11)	IIB(12)	IIIA(13)	IVA(14)	VA(15)	VIA(16)	VIIA(17)	Noble Gases (18)
1	1 H 1.008																	2 He 4.003
2	3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
3	11 Na 22.99	12 Mg 24.31											13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.06	17 Cl 35.45	18 Ar 39.95
4	19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.90	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.93	28 Ni 58.7	29 Cu 63.55	30 Zn 65.38	31 Ga 69.72	32 Ge 72.59	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
5	37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc 98.91	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.6	53 I 126.9	54 Xe 131.3
6	55 Cs 132.9	56 Ba 137.3	57* La 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.9	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (210)	85 At (210)	86 Rn (222)
7	87 Fr (223)	88 Ra 226.0	89** Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (262)	108 Hs (265)	109 Mt (266)	110 Ds (269)	111 Uuu (272)	112 Uub (277)	113 Uub (284)	114 Uuq (285)	115 Uub (288)	116 Uuh (289)		

Inner Transition Elements

* Lanthanide Series 6	58 Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (145)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.0	71 Lu 175.0
** Actinide Series 7	90 Th 232.0	91 Pa 231.0	92 U 238.0	93 Np 237.0	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (260)

The Elements and Their Symbols

Actinium	Ac	Cesium	Cs	Hafnium	Hf	Mercury	Hg	Protactinium	Pa	Tellurium	Te
Aluminum	Al	Chlorine	Cl	Hassium	Hs	Molybdenum	Mo	Radium	Ra	Terbium	Tb
Americium	Am	Chromium	Cr	Helium	He	Neodymium	Nd	Radon	Rn	Thallium	Tl
Antimony	Sb	Cobalt	Co	Holmium	Ho	Neon	Ne	Rhenium	Re	Thorium	Th
Argon	Ar	Copper	Cu	Hydrogen	H	Neptunium	Np	Rhodium	Rh	Thulium	Tm
Arsenic	As	Curium	Cm	Indium	In	Nickel	Ni	Rubidium	Rb	Tin	Sn
Astatine	At	Darmstadtium	Ds	Iodine	I	Niobium	Nb	Ruthenium	Ru	Titanium	Ti
Barium	Ba	Dubnium	Db	Iridium	Ir	Nitrogen	N	Rutherfordium	Rf	Tungsten	W
Berkelium	Bk	Dysprosium	Dy	Iron	Fe	Nobelium	No	Samarium	Sm	Uranium	U
Beryllium	Be	Einsteinium	Es	Krypton	Kr	Osmium	Os	Scandium	Sc	Vanadium	V
Bismuth	Bi	Erbium	Er	Lanthanum	La	Oxygen	O	Seaborgium	Sg	Xenon	Xe
Bohrium	Bh	Europium	Eu	Lawrencium	Lr	Palladium	Pd	Selenium	Se	Ytterbium	Yb
Boron	B	Fermium	Fm	Lead	Pb	Phosphorous	P	Silicon	Si	Yttrium	Y
Bromine	Br	Fluorine	F	Lithium	Li	Platinum	Pt	Silver	Ag	Zinc	Zn
Cadmium	Cd	Francium	Fr	Lutetium	Lu	Plutonium	Pu	Sodium	Na	Zirconium	Zr
Calcium	Ca	Gadolinium	Gd	Magnesium	Mg	Polonium	Po	Strontium	Sr		
Californium	Cf	Gallium	Ga	Manganese	Mn	Potassium	K	Sulfur	S		
Carbon	C	Germanium	Ge	Meitnerium	Mt	Praseodymium	Pr	Tantalum	Ta		
Cerium	Ce	Gold	Au	Mendelevium	Md	Promethium	Pm	Technetium	Tc		

Appendix B
Observing the Sky

Observing the sky with the naked eye is as important to modern astronomy as picking up pretty pebbles is to modern geology. The sky is inspiring—a natural wonder unimaginably bigger than the Grand Canyon, the Rocky Mountains, or any other site that tourists visit every year. To neglect the beauty of the sky is equivalent to geologists neglecting the beauty of the minerals they study. This supplement is meant to act as a tourist's guide to the sky. You analyzed the universe in the textbook's chapters; here you can admire it.

The brighter stars in the sky are visible even from the centers of cities with their air and light pollution. But in the countryside, only a few miles beyond the cities, the night sky is a velvety blackness strewn with thousands of glittering stars. From a wilderness location, far from the city's glare, and especially from high mountains, the night sky is spectacular.

Using Star Charts

The constellations are a fascinating cultural heritage of our planet, but they are sometimes a bit difficult to learn because of Earth's motion. The constellations above the horizon change with the time of night and the seasons.

Because Earth rotates eastward, the sky appears to rotate westward around Earth. A constellation visible overhead soon after sunset will appear to move westward, and in a few hours it will disappear below the horizon. Other constellations will rise in the east, so the sky changes gradually through the night.

In addition, Earth's orbital motion makes the sun appear to move eastward among the stars. Each day the sun moves about twice its own diameter, or about one degree, eastward along the ecliptic; consequently, each night at sunset, the constellations are shifted about one degree farther toward the west.

Orion, for instance, is visible in the evening sky in January; but, as the days pass, the sun moves closer to Orion. By March, Orion is difficult to see in the western sky soon after sunset. By June, the sun is so close to Orion it sets with the sun and is invisible. Not until late July is the sun far enough past Orion for the constellation to become visible rising in the eastern sky just before dawn.

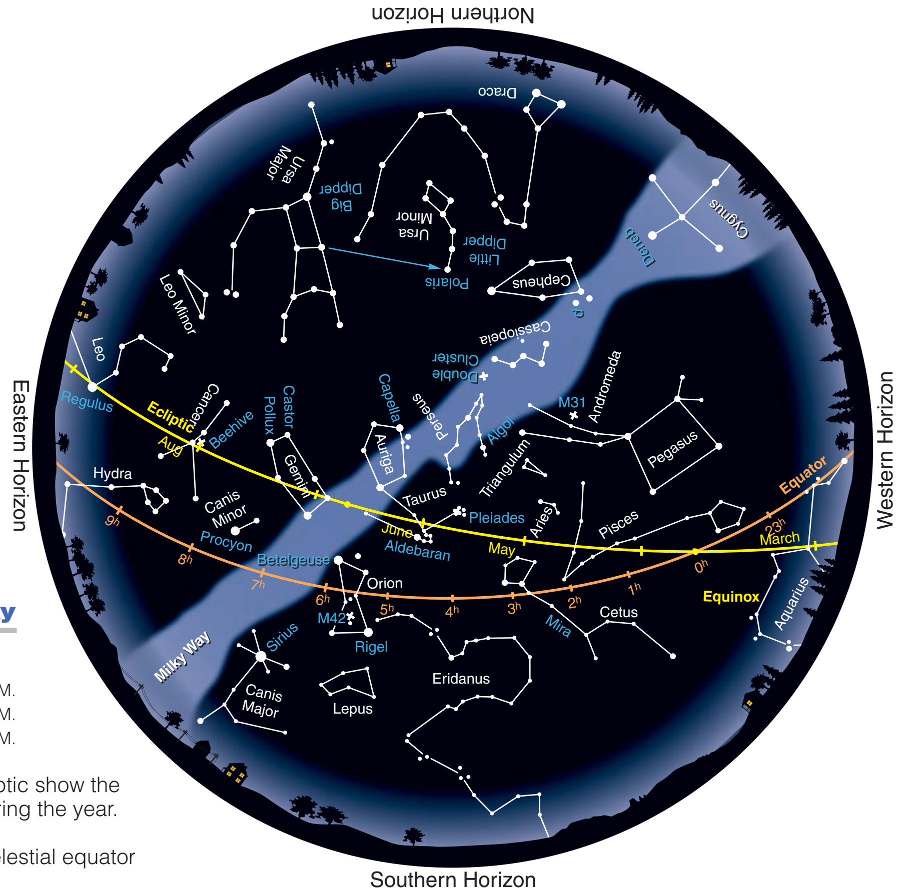
Because of the rotation and orbital motion of Earth, you need more than one star chart to map the sky. Which chart you select depends on the month and the time of night.

Two sets of charts are included for two typical locations on Earth. The Northern Hemisphere star charts show the sky as seen from a northern latitude typical for the United States and Central Europe. The Southern Hemisphere star charts are appropriate for readers in Earth's Southern Hemisphere, including Australia, southern South America, and southern Africa.

To use the charts, select the appropriate chart and hold it overhead as shown in ■ Figure B-1. If you face south, turn the chart until the words *southern horizon* are at the bottom of the chart. If you face other directions, turn the chart appropriately.

■ **Figure B-1**

To use the star charts in this book, select the appropriate chart for the season and time. Hold it overhead and turn it until the direction at the bottom of the chart is the same as the direction you are facing.



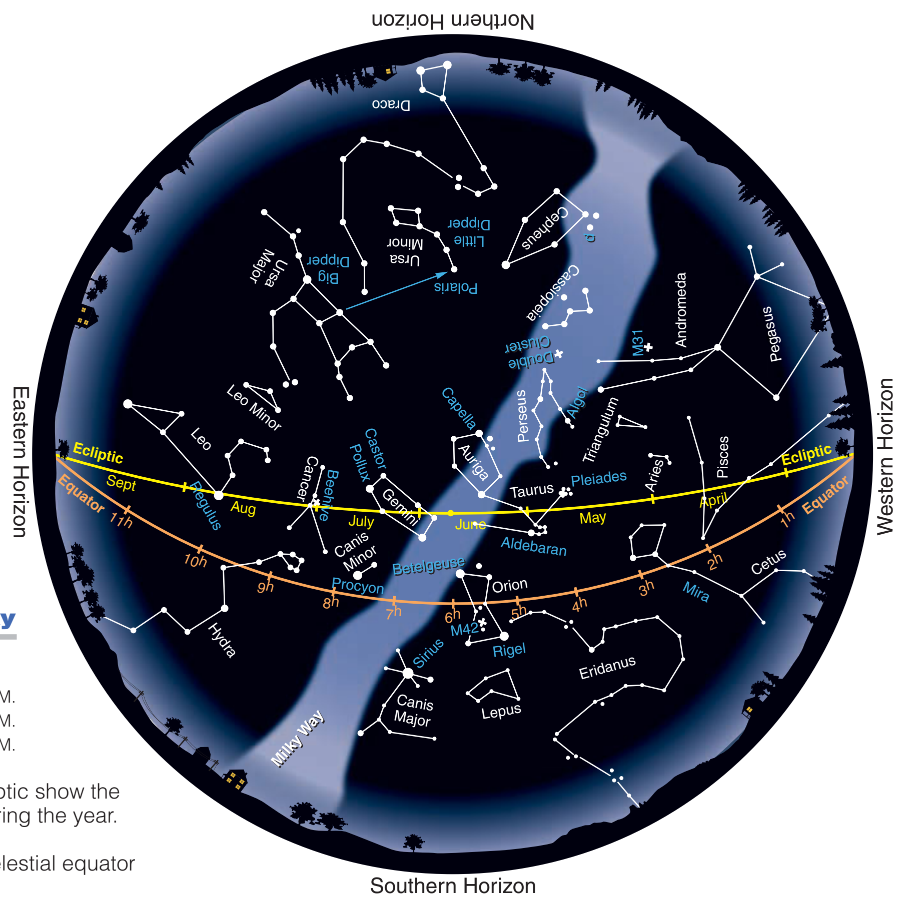
Northern Hemisphere Sky

JANUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



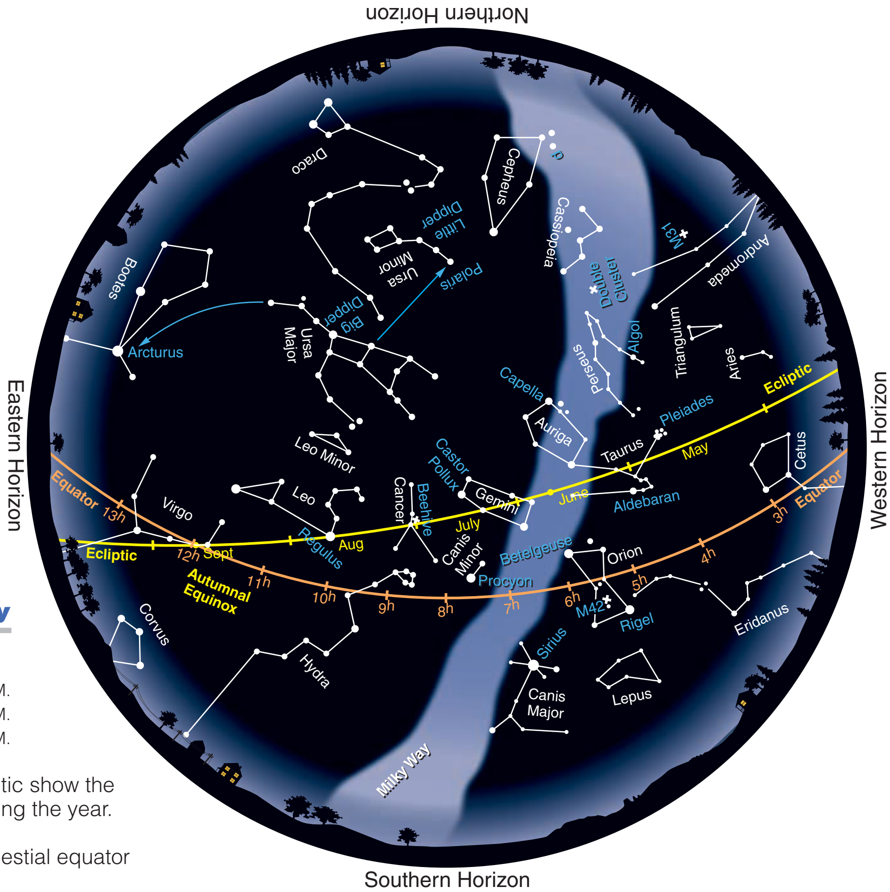
Northern Hemisphere Sky

FEBRUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



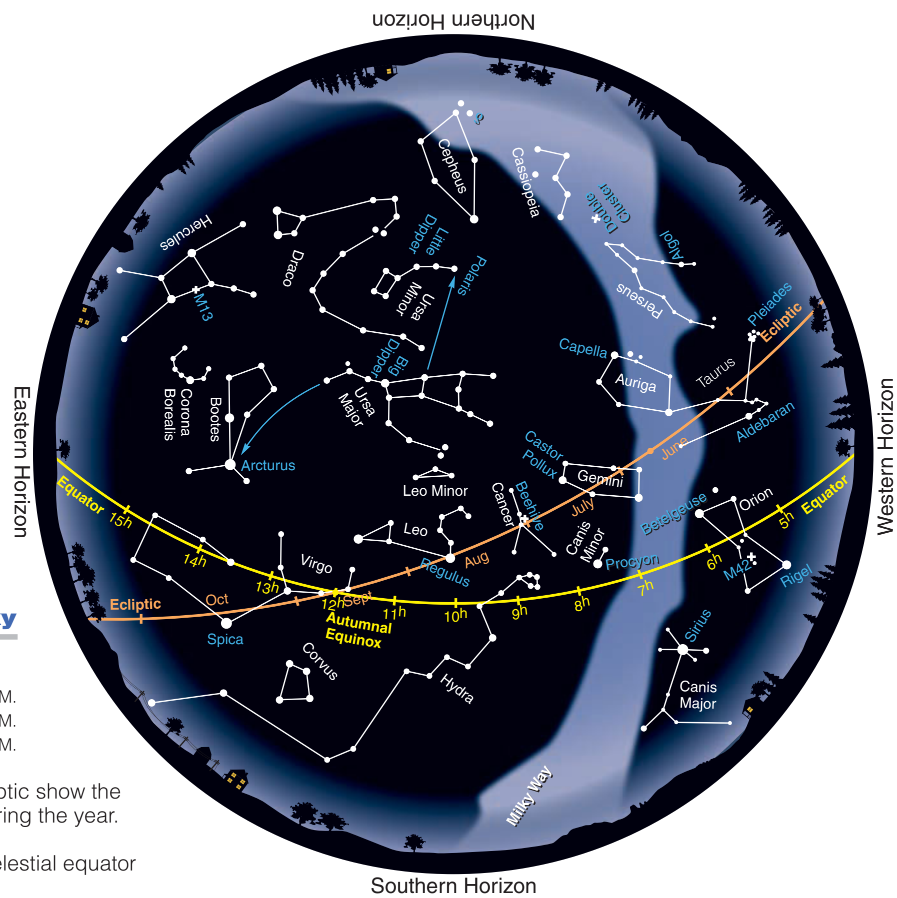
Northern Hemisphere Sky

MARCH

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



Northern Hemisphere Sky

APRIL

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

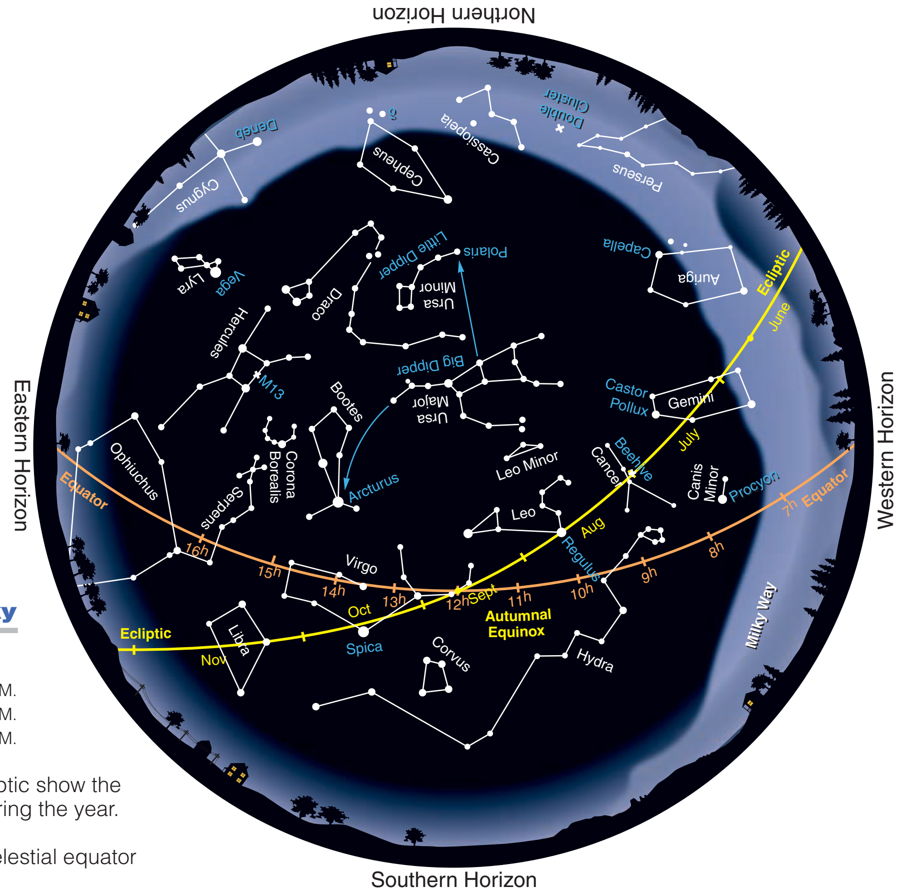
Northern Hemisphere Sky

MAY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



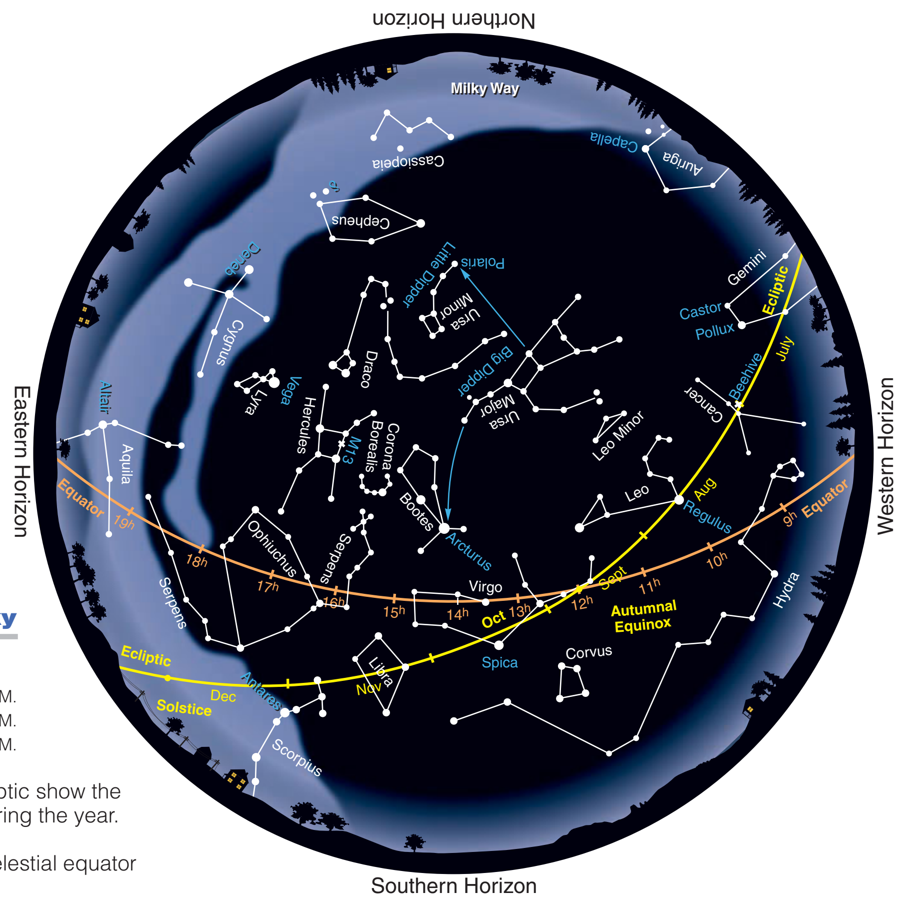
Northern Hemisphere Sky

JUNE

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



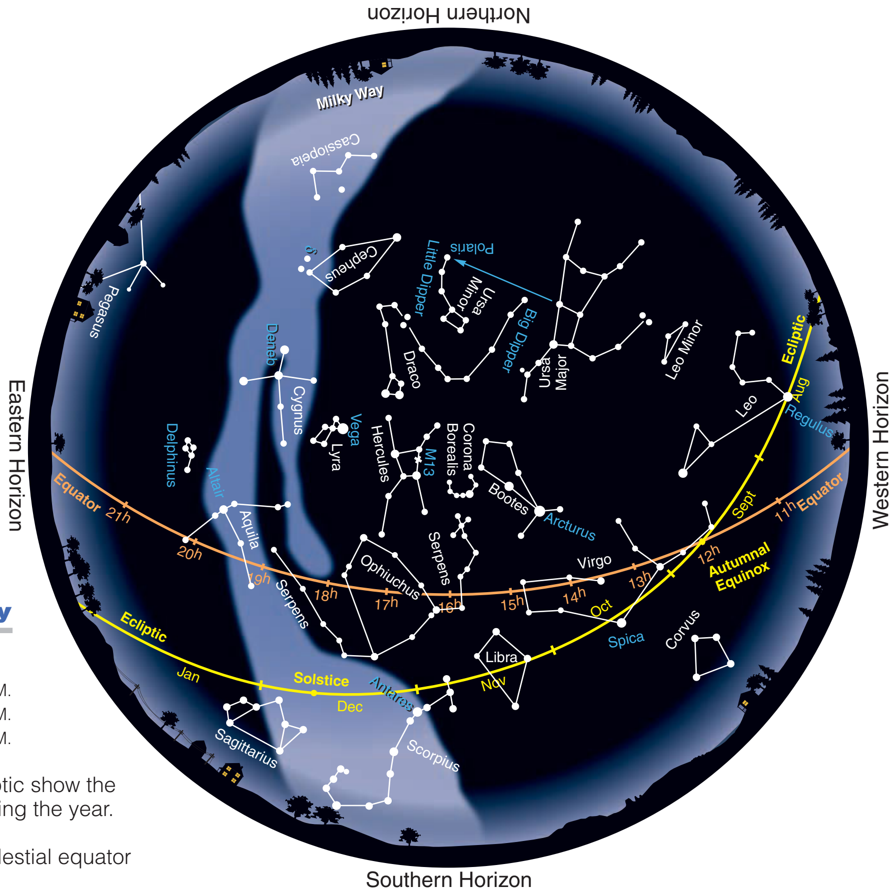
Northern Hemisphere Sky

JULY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



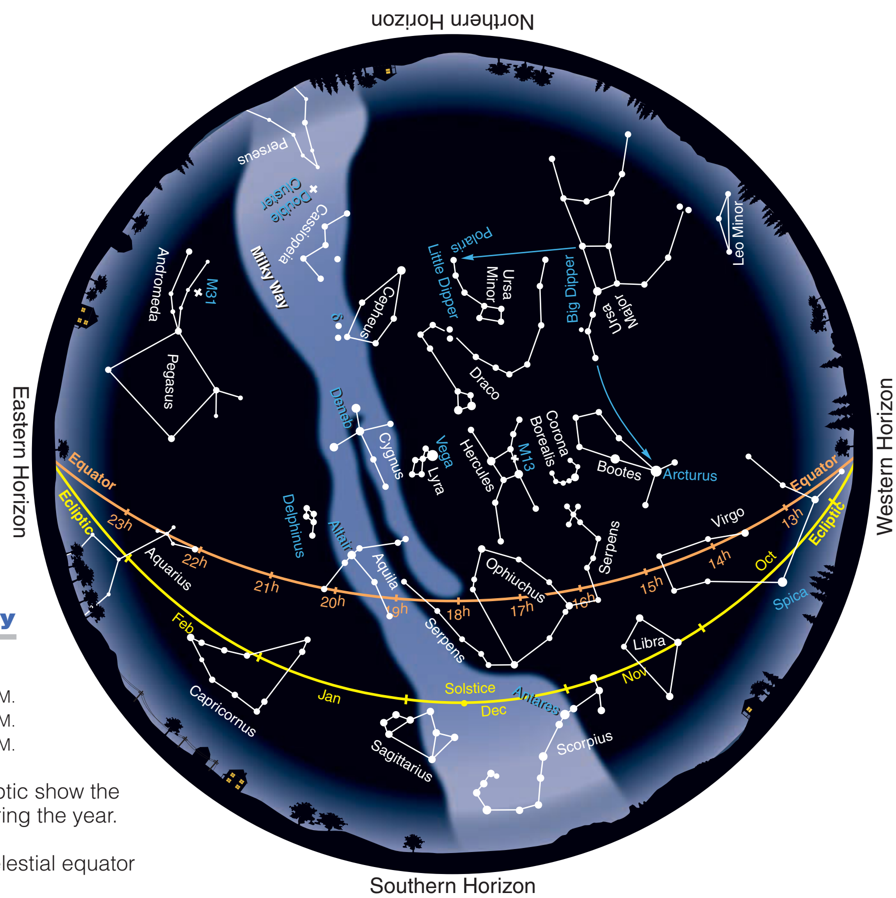
Northern Hemisphere Sky

AUGUST

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



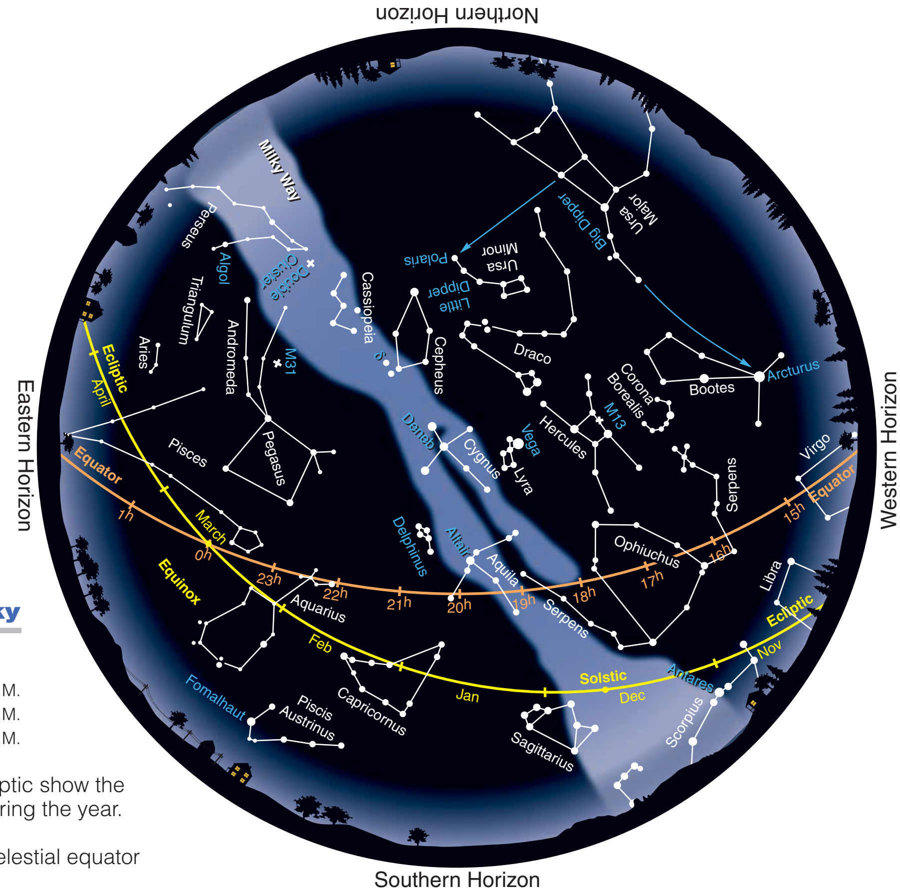
Northern Hemisphere Sky

SEPTEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



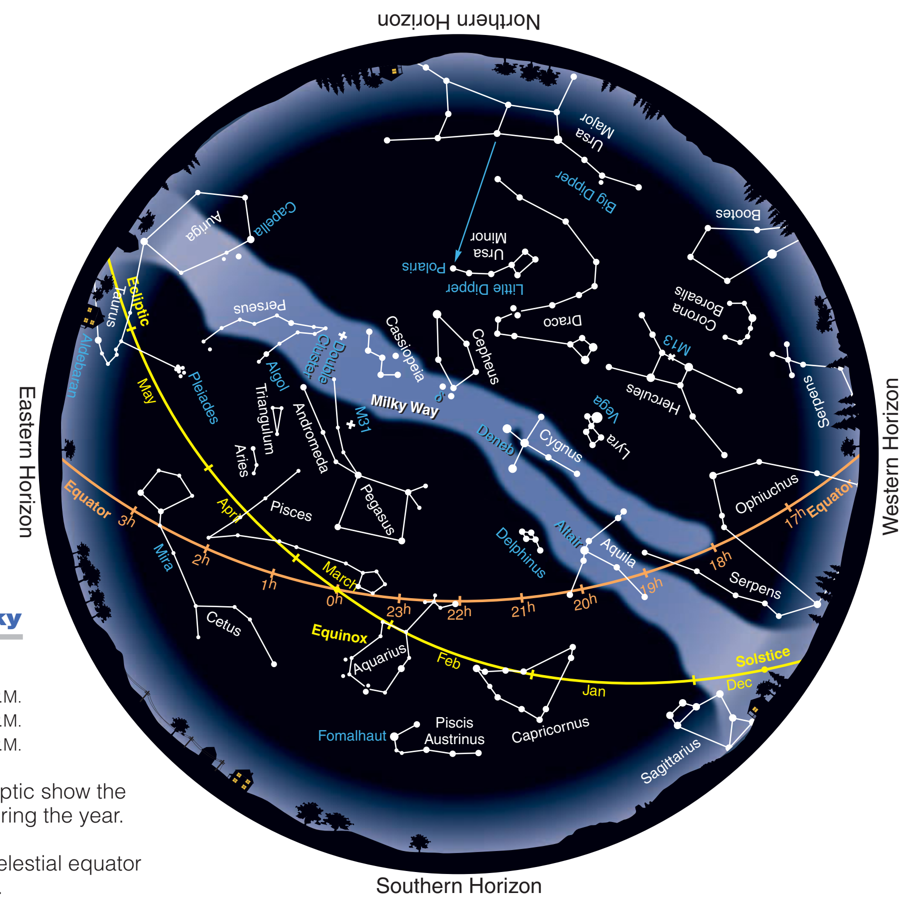
Northern Hemisphere Sky

OCTOBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



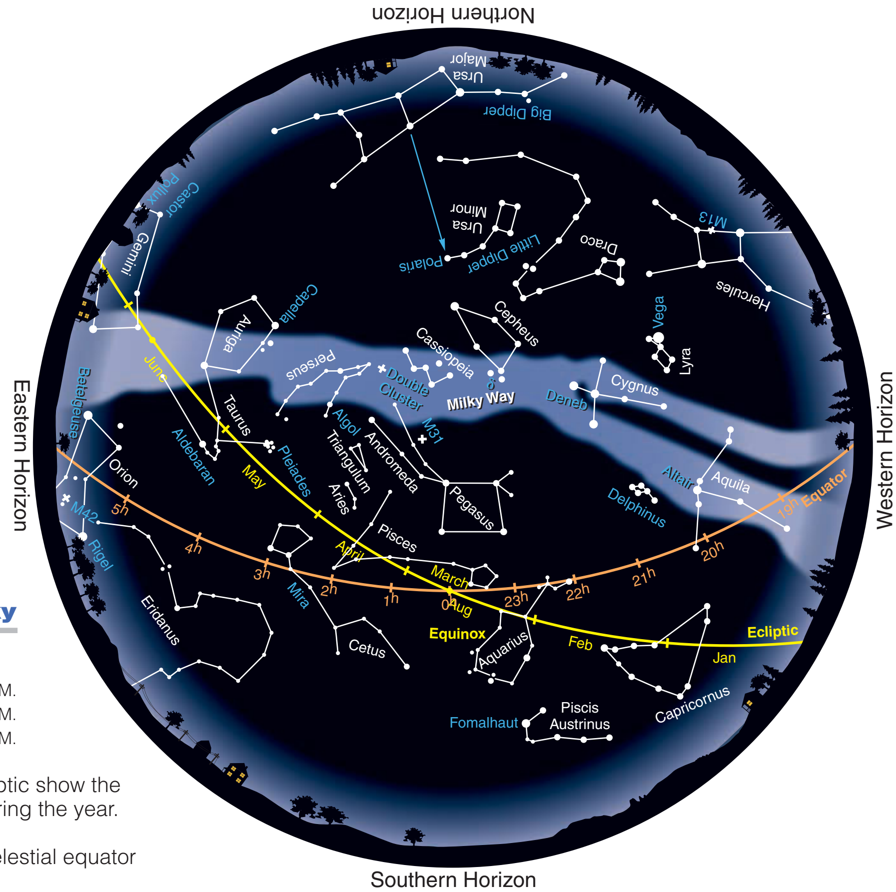
Northern Hemisphere Sky

NOVEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



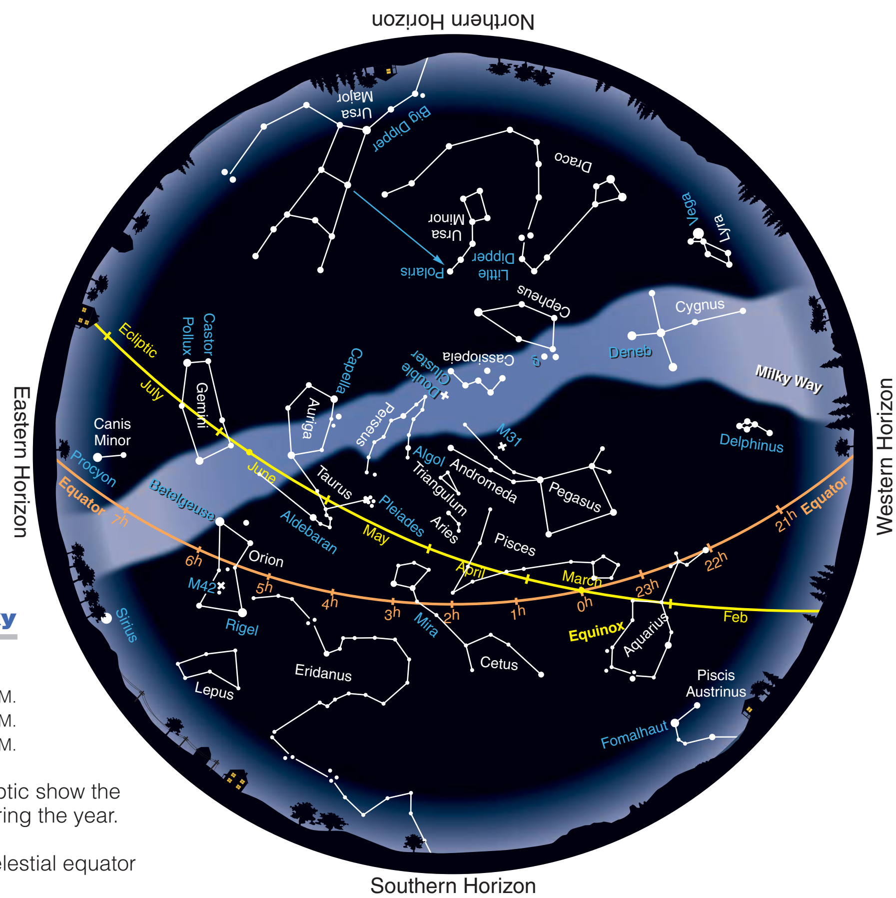
Northern Hemisphere Sky

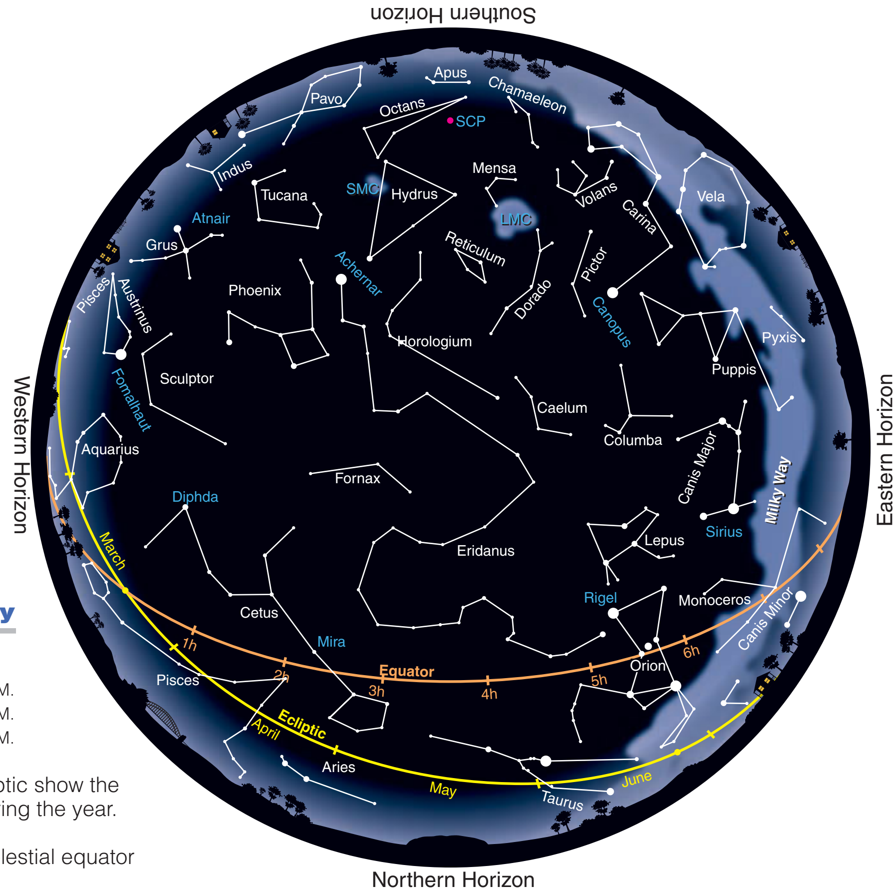
DECEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.





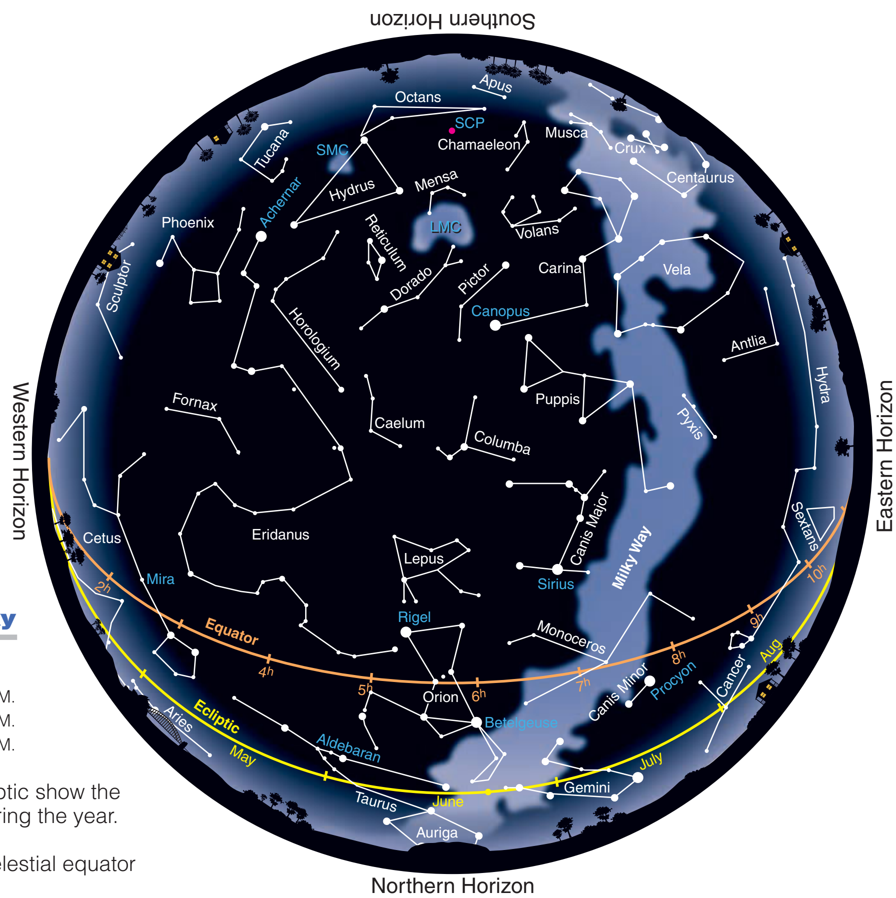
Southern Hemisphere Sky

JANUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



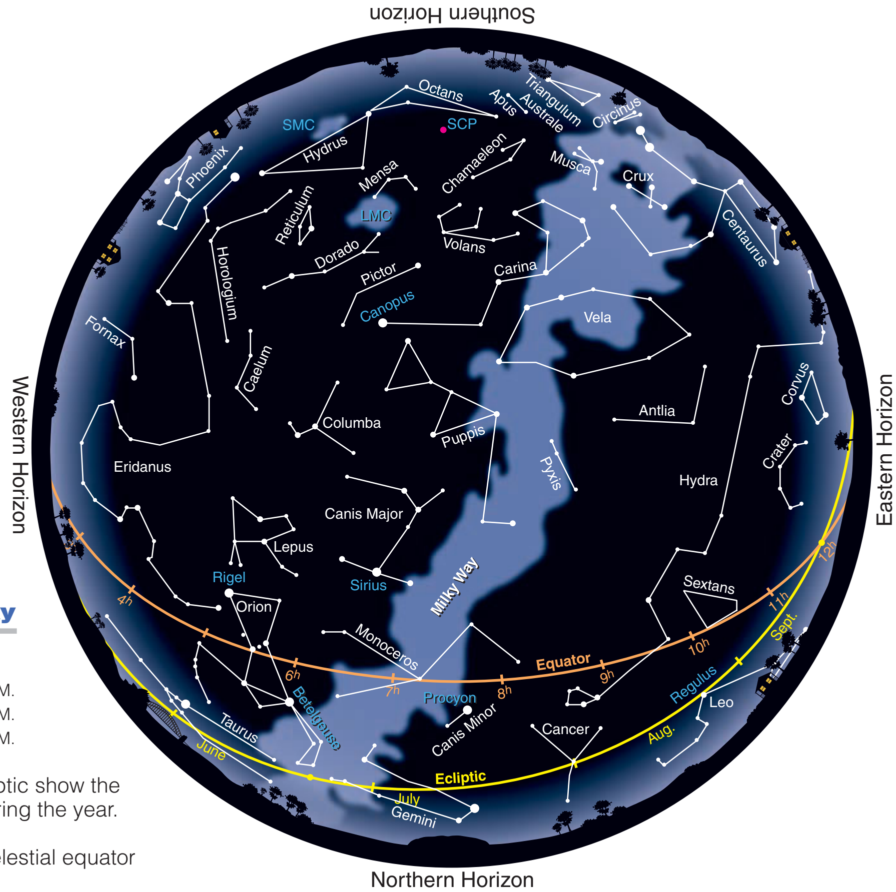
Southern Hemisphere Sky

FEBRUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



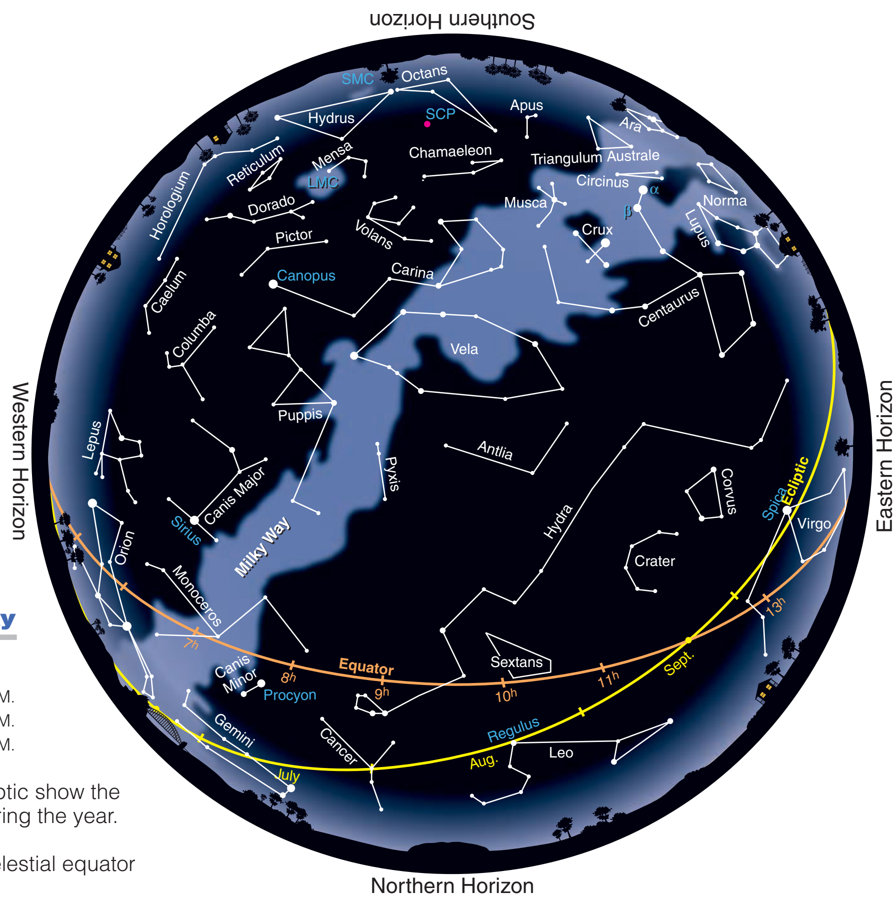
Southern Hemisphere Sky

MARCH

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



Southern Hemisphere Sky

APRIL

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

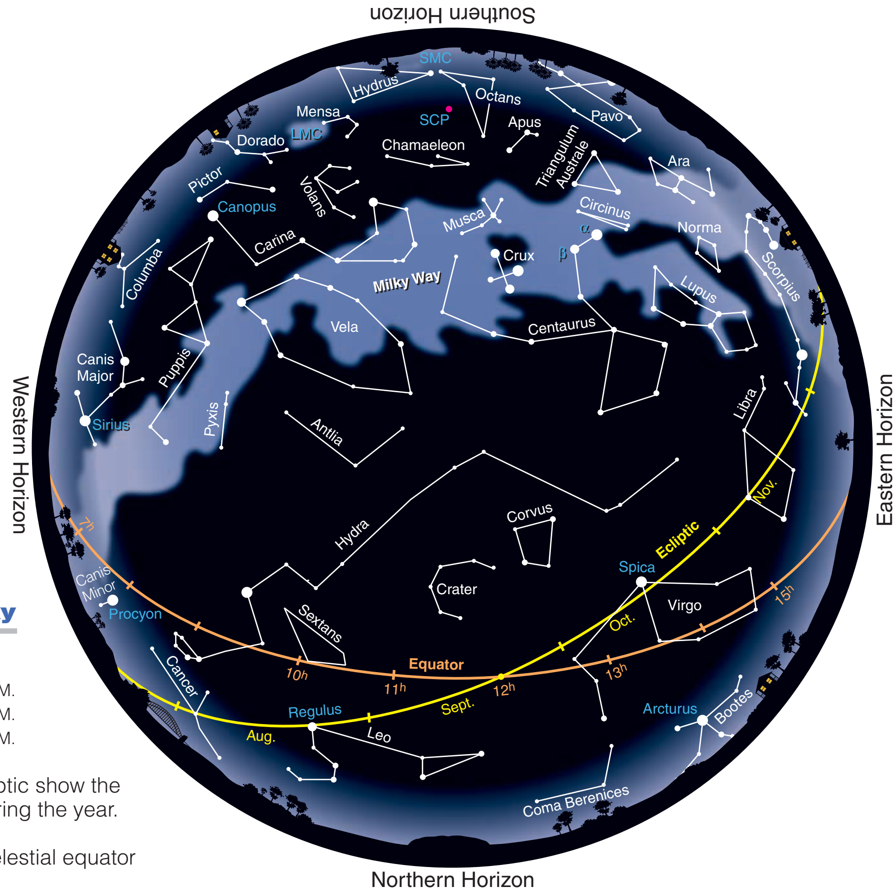
Southern Hemisphere Sky

MAY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



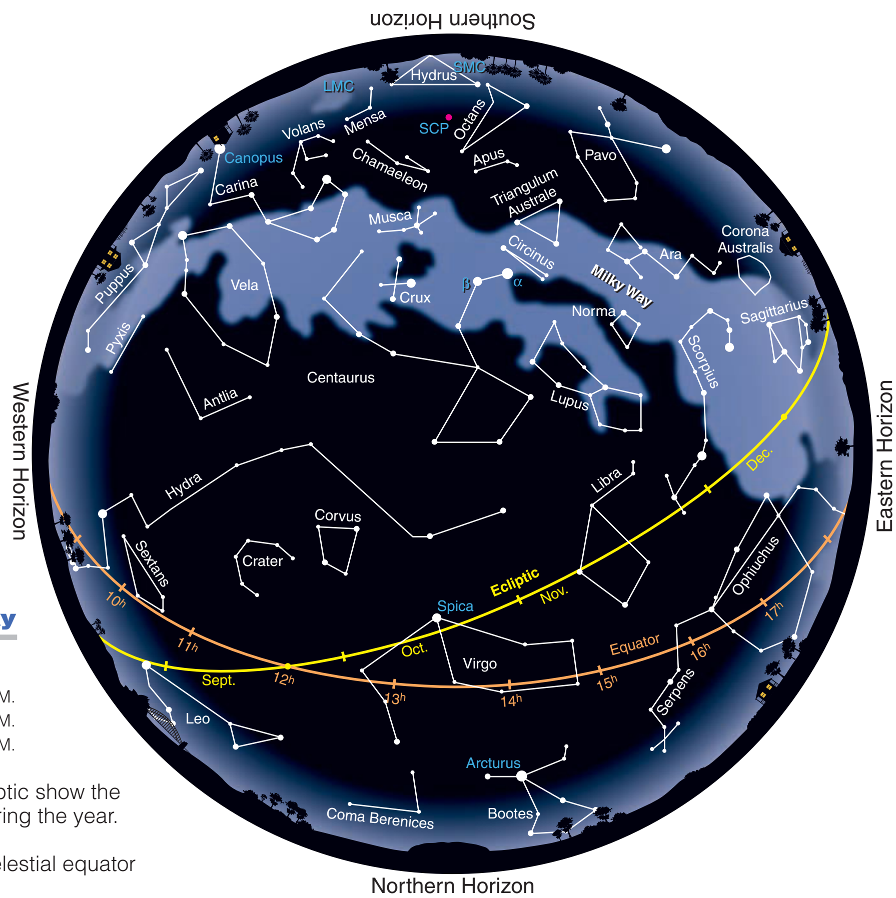
Southern Hemisphere Sky

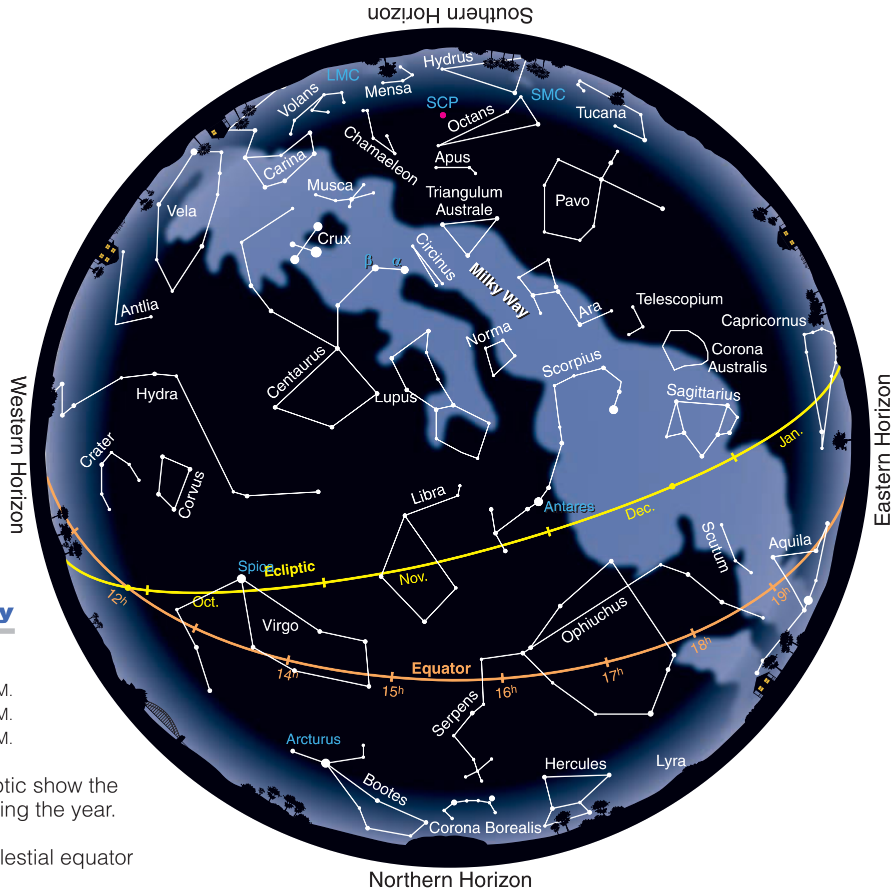
JUNE

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.





Southern Hemisphere Sky

JULY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

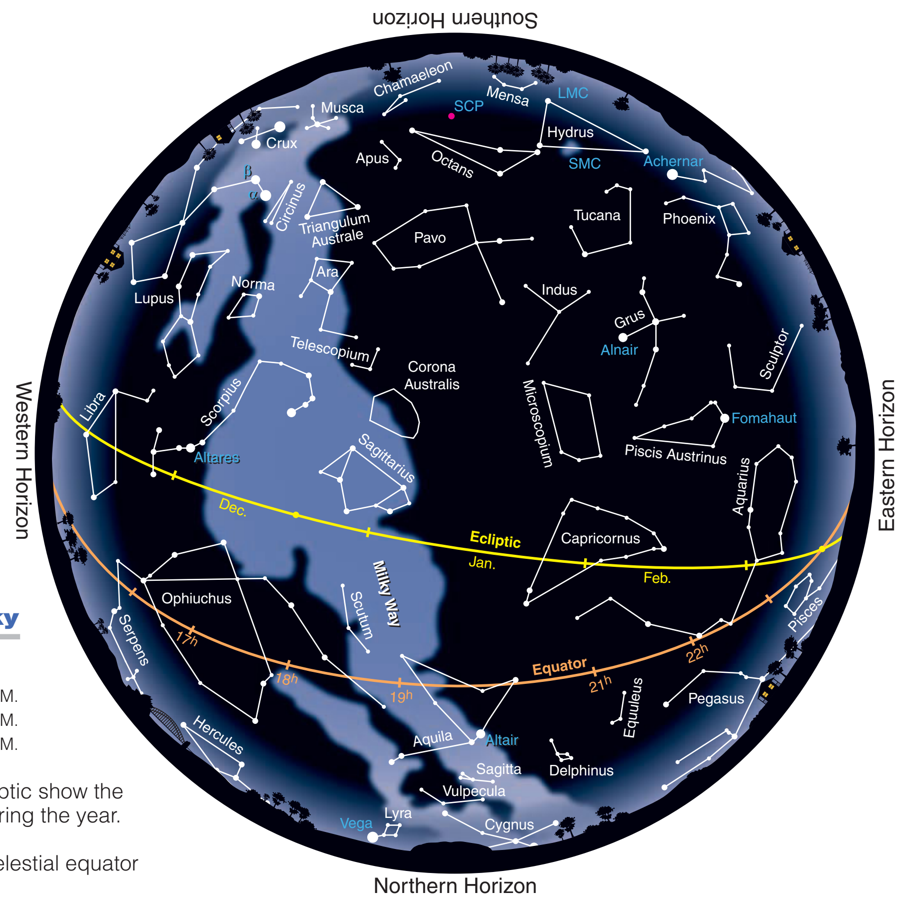
Southern Hemisphere Sky

AUGUST

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



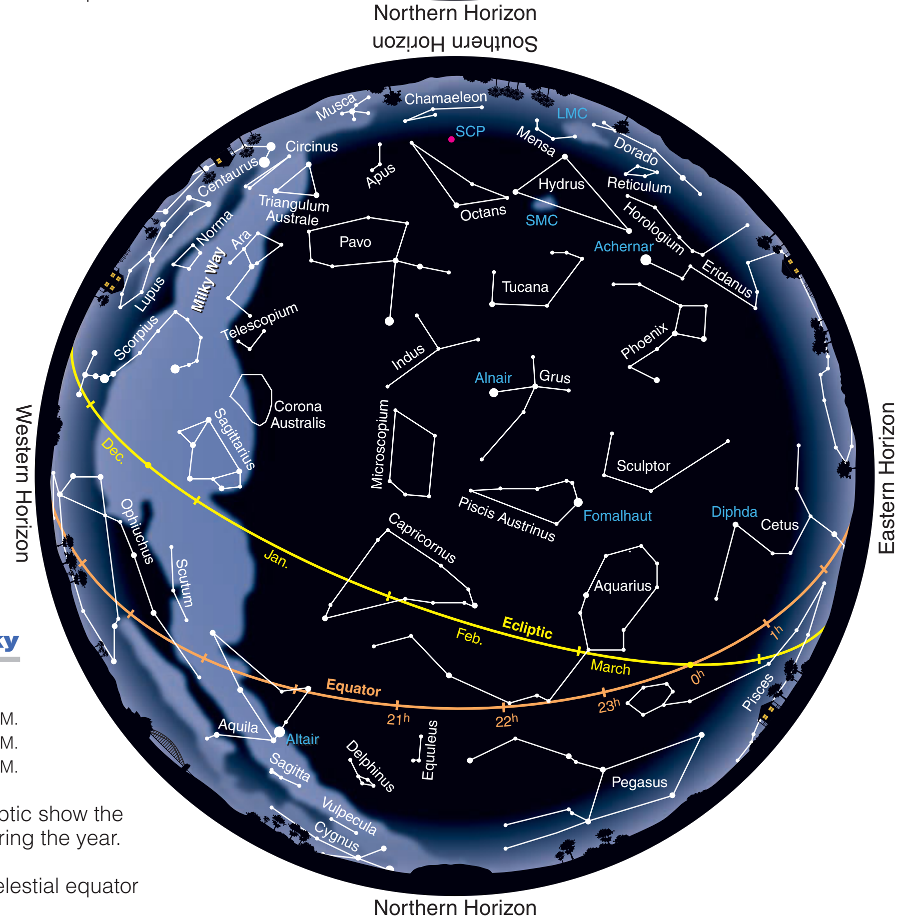
Southern Hemisphere Sky

SEPTEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



Southern Hemisphere Sky

OCTOBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

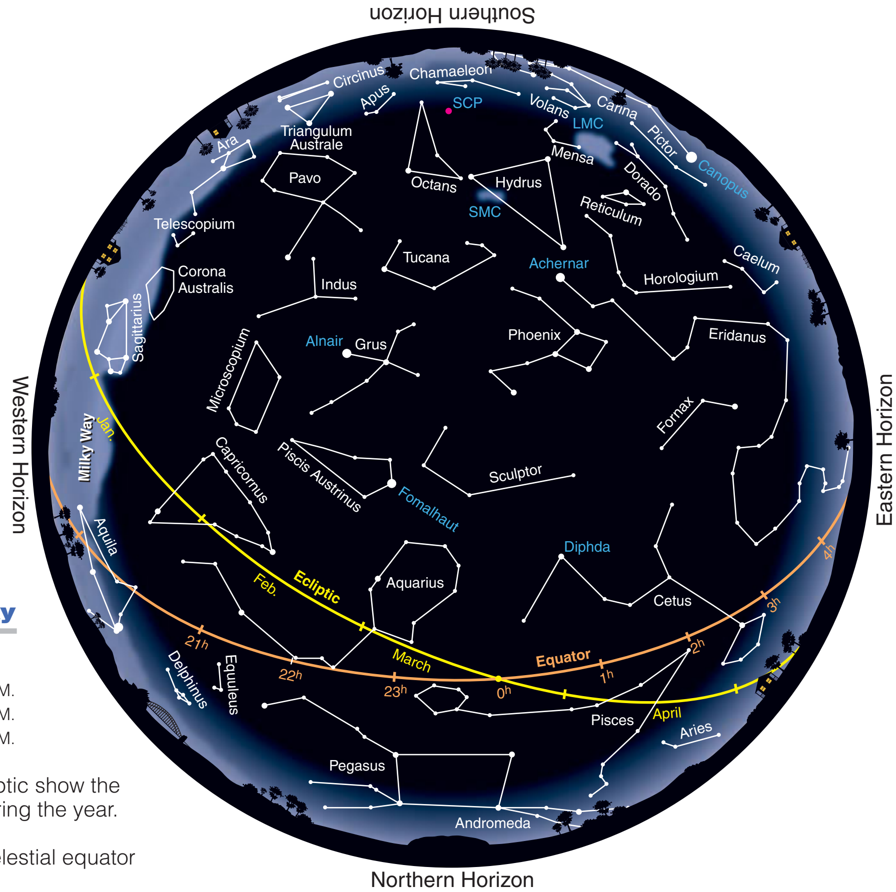
Southern Hemisphere Sky

NOVEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



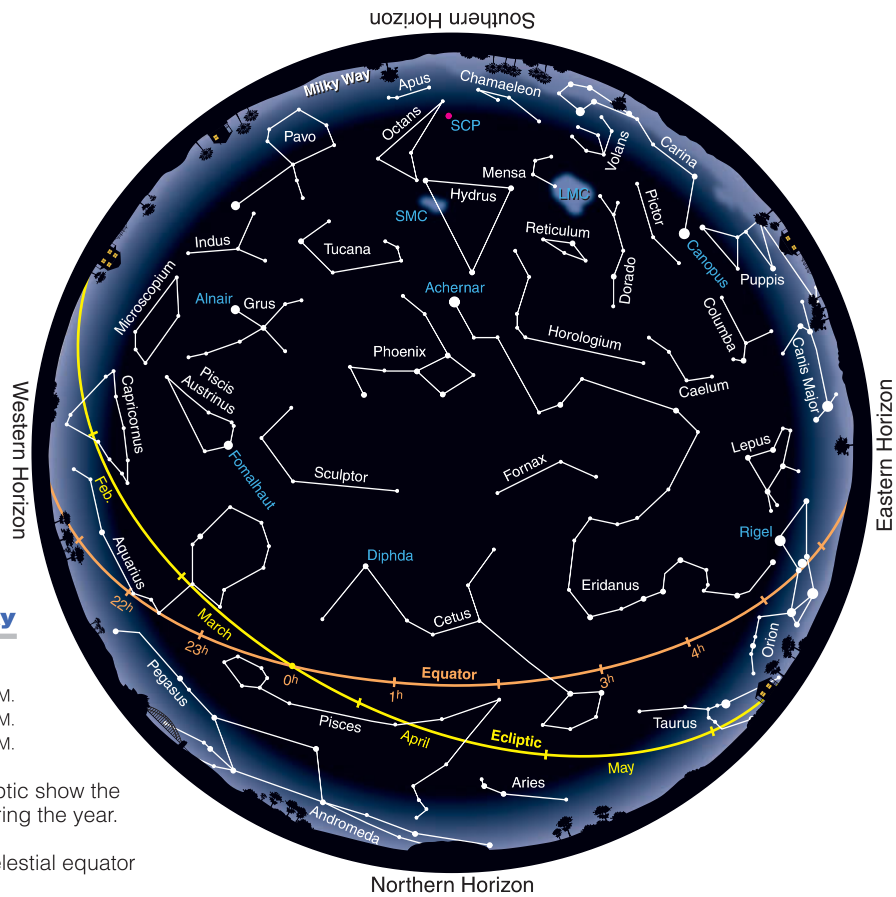
Southern Hemisphere Sky

DECEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.



Glossary

Pronunciation Guide

ā pay	ē be	ī pie	ō so	ŭ cut
ă hat	ĕ pet	ĭ pit	ŏ pot	
ä father		î pier	ô paw	
			o͞o food	

Numbers in parentheses refer to the page where the term is first discussed in the text.

absolute visual magnitude *(M_v)* (138) Intrinsic brightness of a star. The apparent visual magnitude the star would have if it were 10 pc away.

absolute zero (99) The theoretical lowest possible temperature at which a material contains no extractable heat energy. Zero on the Kelvin temperature scale.

absorption line (102) A dark line in a spectrum caused by the absence of photons absorbed by atoms or molecules.

absorption spectrum (dark-line spectrum) (102) A spectrum that contains absorption lines.

accretion (331) The sticking together of solid particles to produce a larger particle.

accretion disk (198) The whirling disk of gas that forms around a compact object such as a white dwarf, neutron star, or black hole as matter is drawn in.

achondrite (ā·kŏn´drīt) (410) Stony meteorite containing no chondrules or volatiles.

achromatic lens (73) A telescope lens composed of two lenses ground from different kinds of glass and designed to bring two selected colors to the same focus and correct for chromatic aberration.

active galactic nuclei (AGN) (283) The centers of active galaxies that are emitting large amounts of excess energy. *See also* active galaxy.

active galaxy (283) A galaxy whose center emits large amounts of excess energy, often in the form of radio emission. Active galaxies are suspected of having massive black holes in their centers into which matter is flowing.

active optics (81) Thin telescope mirrors that are controlled by computers to maintain proper shape as the telescope moves.

active region (127) A magnetic region on the solar surface that includes sunspots, prominences, flares, and the like.

adaptive optics (82) A computer-controlled optical system in an astronomical telescope used to partially correct for seeing.

albedo (352) The ratio of the light reflected from an object to the light that hits the object. Albedo equals 0 for perfectly black and 1 for perfectly white.

alt-azimuth mounting (81) A telescope mounting that allows the telescope to move in altitude (perpendicular to the horizon) and in azimuth (parallel to the horizon). *See also* equatorial mounting.

amino acid (ŭ·mē´nō) (428) Carbon-chain molecule that is the building block of protein.

Angstrom (Å) (ăng´strŭm) (71) A unit of distance commonly used to measure the wavelength of light. $1 \text{ Å} = 10^{-10}$ m.

angular diameter (17) The angle formed by lines extending from the observer to opposite sides of an object.

angular distance (17) The angle formed by lines extending from the observer to two locations.

angular momentum (198) A measure of the tendency of a rotating body to continue rotating. Mathematically, the product of mass, velocity, and radius.

annular eclipse (36) A solar eclipse in which the solar photosphere appears around the edge of the moon in a bright ring, or annulus. The corona, chromosphere, and prominences cannot be seen.

anorthosite (ăn·ôr´thŭ·sīt) (352) Rock of aluminum and calcium silicates found in the lunar highlands.

anticoded (437) Refers to a message that is meant to be understood by a recipient about whom little or nothing is known. Opposite of coded.

antimatter (303) Matter composed of antiparticles, which upon colliding with a matching particle of normal matter annihilate and convert the mass of both particles into energy. The antiproton is the antiparticle of the proton, and the positron is the antiparticle of the electron.

aphelion (ŭ·fē´le·ŭn) (25) The orbital point of greatest distance from the sun.

apogee (36) The point farthest from Earth in the orbit of a body circling Earth.

apparent visual magnitude (m_v) (13) The brightness of a star as seen by human eyes on Earth.

association (164) Group of widely scattered stars (10 to 100) moving together through space. Not gravitationally bound into clusters.

asterism (11) A named grouping of stars that is not one of the recognized constellations. Examples are the Big Dipper and the Pleiades.

asteroid (324) Small, rocky world. Most asteroids lie between Mars and Jupiter in the asteroid belt.

astronomical unit (AU) (3) Average distance from Earth to the sun; 1.5×10^8 km, or 93×10^6 mi.

atmospheric window (72) Wavelength region in which our atmosphere is transparent—at visual, infrared, and radio wavelengths.

aurora (ô·rôr´ŭ) (131) The glowing light display that results when a planet's magnetic field guides charged particles toward the north and south magnetic poles, where they strike the upper atmosphere and excite atoms to emit photons.

autumnal equinox (24) The point on the celestial sphere where the sun crosses the celestial equator going southward. Also, the time when the sun reaches this point and autumn begins in the Northern Hemisphere—about September 22.

Babcock model (125) A model of the sun's magnetic cycle in which the differential rotation of the sun winds up and tangles the solar magnetic field in a 22-year cycle. This is thought to be responsible for the 11-year sunspot cycle.

Balmer series (103) A series of spectral lines produced by hydrogen in the near-ultraviolet and visible parts of the spectrum. The three longest-wavelength Balmer lines are visible to the human eye.

barred spiral galaxy (262) A spiral galaxy with an elongated nucleus resembling a bar from which the arms originate.

basalt (350) Dark igneous rock, characteristic of solidified lava.

belt–zone circulation (382) The atmospheric circulation typical of Jovian planets in which dark belts and bright zones encircle the planet parallel to its equator.

big bang (300) The high-density, high-temperature state from which the expanding universe of galaxies began.

big rip (313) The fate of the universe if dark energy increases with time and galaxies, stars, and even atoms are eventually ripped apart by the accelerating expansion of the universe.

binary stars (145) Pairs of stars that orbit around their common center of mass.

binding energy (96) The energy needed to pull an electron away from its atom.

biological evolution (427) The process by which living things adapt to changing environments via changes in genetic makeup.

bipolar flow (169) Jets of gas flowing away from a central object in opposite directions. A term usually applied to protostars.

birth line (166) In the H–R diagram, the line above the main sequence where protostars first become visible.

blackbody radiation (99) Radiation emitted by a hypothetical perfect radiator. The spectrum is continuous, and the wavelength of maximum emission depends on the body's temperature.

black dwarf (196) The end state of a white dwarf that has cooled to low temperature.

black hole (223) A mass that has collapsed to such a small volume that its gravity prevents the escape of all radiation. Also, the volume of space from which radiation may not escape.

blueshift (108) A Doppler shift toward shorter wavelengths caused by a velocity of approach.

Bok globule (168) Small, dark cloud only about 1 ly in diameter that contains 10 to 1000 solar masses of gas and dust, believed to be related to star formation.

breccia (brĕch´ē·ŭ) (352) Rock composed of fragments of earlier rocks bonded together.

bright-line spectrum *See* emission spectrum.

brown dwarf (179) A star whose mass is too low to ignite nuclear fusion and that is heated by contraction.

calibration (237) The establishment of the relationship between a parameter that is easily determined and a parameter that is more difficult to determine. For example, the periods of Cepheid variables have been calibrated to reveal absolute magnitudes, which can then be used to find distance. Thus astronomers say Cepheids have been calibrated as distance indicators.

Cambrian explosion (kăm´brē·ŭn) (433) The sudden evolution of many complex creatures about 530 million years ago at the beginning of the Cambrian period.

carbonaceous chondrite (kär·bŭ·nā´·shŭs kŏn´-drīt) (409) Stony meteorite that contains both chondrules and volatiles. These chondrites may be the least-altered remains of the solar nebula still present in the solar system.

carbon–nitrogen–oxygen (CNO) cycle (173) A series of nuclear reactions that use carbon as a catalyst to combine four hydrogen atoms to make one helium atom plus energy, effective in stars more massive than the sun.

Cassegrain focus (kăs´ŭ·grān) (80) The optical design in which the secondary mirror reflects light back down the tube through a hole in the center of the objective mirror.

CCD *See* charge-coupled device.

celestial equator (16) The imaginary line around the sky directly above Earth's equator.

celestial pole (north or south) (16) One of the two points on the celestial sphere directly above Earth's poles.

celestial sphere (16) An imaginary sphere of very large radius surrounding Earth to which the planets, stars, sun, and moon seem to be attached.

center of mass (63) The balance point of a body or system of masses. The point about which a body or system of masses rotates in the absence of external forces.

Cepheid variable star (sĕ·fē´ĭd) (236) Variable star with a period of 60 days whose period of variation is related to luminosity.

Chandrasekhar limit (shăn´drä·sā´·kär) (196) The maximum mass of a white dwarf, about 1.4 solar masses. A white dwarf of greater mass cannot support itself and will collapse.

charge-coupled device (CCD) (84) An electronic device consisting of a large array of light-sensitive elements used to record very faint images.

chemical evolution (431) The chemical process that led to the growth of complex molecules on primitive Earth. This did not involve the reproduction of molecules.

Chicxulub (421) A barely discernible crater on the coast of Mexico's Yucatan peninsula thought to be the site of the impact that caused the K-T event, the end of the Cretaceous period and the extinction of the dinosaurs. Named after a small town located near the center of the crater.

chondrite (kŏn´drīt) (409) A stony meteorite that contains chondrules.

chondrule (kŏn´dro͞ol) (409) Round, glassy body, found in some stony meteorites, believed to have solidified very quickly from molten drops of silicate material.

chromatic aberration (krō·măt´ĭk) (73) A distortion found in refracting telescopes because lenses focus different colors at slightly different distances. Images are consequently surrounded by color fringes.

chromosome (429) A body within a living cell that contains genetic information responsible for the determination and transmission of hereditary traits.

chromosphere (krō´mŭ·sfîr) (35) Bright gases just above the photosphere of the sun.

circular velocity (62) The velocity an object needs to stay in orbit around another object.

circumpolar constellation (north or south) (17) A constellation so close to one of the celestial poles that it never sets or never rises as seen from a particular latitude.

closed orbit (63) Either a circular orbit or an elliptical orbit that returns to the same starting point over and over.

closed universe (307) A model universe in which the average density is great enough to stop the expansion and make the universe contract.

cluster method (269) The method of determining the masses of galaxies based on the motions of galaxies in a cluster.

CNO cycle *See* carbon–nitrogen–oxygen cycle.

cold dark matter (309) Mass in the universe, as yet undetected except for its gravitational influence, which is made up of slow-moving particles.

coma (418) The glowing head of a comet.

comet (324) One of the small, icy bodies that orbit the sun and produce tails of gas and dust when they approach the sun.

comparative planetology (344) The study of planets in relation to one another.

comparison spectrum (86) A spectrum of known spectral lines used to identify unknown spectral lines in an object's spectrum.

composite volcano (364) A volcano formed by successive lava and ash flows. This type of volcano has steep sides and, on Earth, is found along a subduction zone.

condensation (331) The growth of a particle by addition of material from surrounding gas, atom by atom.

condensation sequence (331) The sequence in which different materials condense from the solar nebula with decreasing temperature outward from the sun.

conservation of energy (175) One of the basic laws of stellar structure: The amount of energy flowing out of the top of a shell must equal the amount coming in at the bottom plus whatever energy is generated within the shell.

conservation of mass (175) One of the basic laws of stellar structure: The total mass of the star must equal the sum of the masses of the shells, and the mass must be distributed smoothly through the star.

constellation (11) One of the stellar patterns identified by name, usually of mythological gods, people, animals, or objects. Also, the region of the sky containing that star pattern.

continuous spectrum (102) A spectrum in which there are no absorption or emission lines.

convection (115) Circulation in a fluid driven by heat. Hot material rises and cool material sinks.

convective zone (121) The region inside a star where energy is carried outward as rising hot gas and sinking cool gas.

corona (35) On the sun, the faint outer atmosphere composed of low-density, high-temperature gas.

coronae (363) On Venus, large, round geological faults in the crust caused by the intrusion of magma below the crust.

coronagraph (116) A telescope designed to photograph the inner corona of the sun.

coronal hole (131) An area of the solar surface that is dark at X-ray wavelengths, thought to be associated with divergent magnetic fields and the source of the solar wind.

coronal mass ejection (CME) (131) Matter ejected from the sun's corona in powerful surges guided by magnetic fields.

cosmic microwave background radiation (301) Radiation from the hot clouds of the big bang explosion. The large red shift makes it appear to come from a body whose temperature is only 2.7 K.

cosmic ray (91) A subatomic particle traveling at tremendous velocity that strikes the Earth's atmosphere from space.

cosmological constant (312) A constant in Einstein's equations of space and time that represents a force of repulsion.

cosmological principle (306) The assumption that any observer in any galaxy sees the same general features of the universe.

cosmology (296) The study of the nature, origin, and evolution of the universe.

Coulomb barrier (ko͞o·lôm) (121) The electrostatic force of repulsion between bodies of like charge, commonly applied to atomic nuclei.

Coulomb force (96) The electrostatic force of repulsion or attraction between charged bodies.

critical density (307) The average density of the universe needed to make its curvature flat.

C-type (413) The type of asteroid with reflectance spectra indicating abundant carbon compounds, found mostly in the outer belt.

dark age (304) The period of time after the glow of the big bang faded into the infrared and before the birth of the first stars, during which the universe expanded in darkness.

dark energy (313) The energy believed to fill empty spaces and drive the acceleration of the expanding universe.

dark halo (242) The low-density extension of the halo of our galaxy, believed to be composed of dark matter.

dark-line spectrum *See* absorption spectrum.

dark matter (243) Nonluminous matter that is detected only by its gravitational influence.

dark nebula (161) A cloud of gas and dust seen silhouetted against a brighter nebula.

debris disk (337) A disk of dust around some stars found by infrared observations. The dust is debris from collisions among asteroids, comets, and Kuiper belt objects.

deferent (dĕf´ŭr·ŭnt) (45) In the Ptolemaic theory, the large circle around Earth along which the center of the epicycle was thought to move.

degenerate matter (187) Extremely high-density matter in which pressure no longer depends on temperature due to quantum mechanical effects.

density wave theory (252) Theory proposed to account for spiral arms as compressions of the interstellar medium in the disk of the galaxy.

deoxyribonucleic acid (428) DNA: The long carbon-chain molecule that records information to govern the biological activity of the organism. DNA carries the genetic data passed to offspring.

deuterium (120) An isotope of hydrogen in which the nucleus contains a proton and a neutron.

diamond-ring effect (36) During a total solar eclipse, the momentary appearance of a spot of photosphere at the edge of the moon, producing a brilliant glare set in the silvery ring of the corona.

differential rotation (125) The rotation of a body in which different parts of the body have different periods of rotation. This occurs in the sun, the Jovian planets, and the disk of the galaxy.

differentiation (332) The separation of planetary material according to density.

diffraction fringe (74) Blurred fringe surrounding any image, caused by the wave properties of light. Because of this, no image detail smaller than the fringe can be seen.

disk component (240) All material confined to the plane of the galaxy.

distance indicator (265) Object whose luminosity or diameter is known, used to find the distance to a star cluster or galaxy.

distance scale (265) The combined calibration of distance indicators used by astronomers to find the distances to remote galaxies.

DNA *See* deoxyribonucleic acid.

Doppler effect (107) The change in the wavelength of radiation due to relative radial motion of source and observer.

double-exhaust model (286) The theory that double radio lobes are produced by pairs of jets emitted in opposite directions from the centers of active galaxies.

double-lobed radio source (285) A galaxy that emits radio energy from two regions (lobes) located on opposite sides of the galaxy.

Drake equation (439) The equation that estimates the total number of communicative civilizations in our galaxy.

dust tail (418) A comet tail composed of dust released from the nucleus and pushed away by the pressure of sunlight, also known as a type II tail.

dwarf planet (403) A body that orbits the sun, is not a satellite of a planet, is massive enough to pull itself into a spherical shape but not massive enough to clear out other bodies in and near its orbit. For example, Pluto, Eris, and Ceres.

dynamo effect (124) The process by which a rotating, convecting body of conducting matter, such as Earth's core, can generate a magnetic field.

east point (16) One of the four cardinal directions, the point on the horizon directly east.

eccentricity, *e* (53) A number between 1 and 0 that describes the shape of an ellipse, the distance from one focus to the center of the ellipse divided by the semimajor axis.

eclipsing binary system (149) A binary star system in which the stars eclipse each other.

ecliptic (22) The apparent path of the sun around the sky.

ejecta (354) Pulverized rock scattered by meteorite impacts on a planetary surface.

electromagnetic radiation (70) Changing electric and magnetic fields that travel through space and transfer energy from one place to another; examples are light or radio waves.

electron (95) Low-mass atomic particle carrying a negative charge.

ellipse (53) A closed curve around two points, called the foci, such that the total distance from one focus to the curve and back to the other focus remains constant.

elliptical galaxy (262) A galaxy that is round or elliptical in outline and contains little gas and dust, no disk or spiral arms, and few hot, bright stars.

emission line (102) A bright line in a spectrum caused by the emission of photons from atoms.

emission nebula (160) A cloud of glowing gas excited by ultraviolet radiation from hot stars.

emission spectrum (bright-line spectrum) (102) A spectrum containing emission lines.

energy level (97) One of a number of states an electron may occupy in an atom, depending on its binding energy.

energy transport (176) Flow of energy from hot regions to cooler regions by one of three methods: conduction, convection, or radiation.

enzyme (428) Special protein that controls processes in an organism.

epicycle (ĕp´ŭ·sī·kŭl) (45) The small circle followed by a planet in the Ptolemaic theory. The center of the epicycle follows a larger circle (the deferent) around Earth.

equant (ē´kwŭnt) (45) In the Ptolemaic theory, the point off-center in the deferent from which the center of the epicycle appears to move uniformly.

equatorial mounting (80) A telescope mounting that allows motion parallel to and perpendicular to the celestial equator.

escape velocity (63) The initial velocity an object needs to escape from the surface of a celestial body.

evening star (26) Any planet visible in the sky just after sunset.

event horizon (223) The boundary of the region of a black hole from which no radiation may escape. No event that occurs within the event horizon is visible to a distant observer.

evolutionary track (165) The path a star follows in the H–R diagram as it gradually changes its surface temperature and luminosity.

excited atom (98) An atom in which an electron has moved from a lower to a higher energy level.

expanding universe (299) The observed property of the universe that the average distance between galaxies is increasing with time.

extrasolar planet (338) A planet orbiting a star other than the sun.

eyepiece (72) A short-focal-length lens used to enlarge the image in a telescope. The lens nearest the eye.

false-color image (84) A representation of graphical data with added or enhanced color to reveal detail.

field of view (2) The area visible in an image. Usually given as the diameter of the region.

filament (130, 314) A solar prominence, seen from above, silhouetted against the bright photosphere.

filtergram (115) A photograph (usually of the sun) taken in the light of a specific region of the spectrum — for example, an H_α filtergram.

flare (131) A violent eruption on the sun's surface.

flatness problem (310) In cosmology, the peculiar circumstance that the early universe must have contained almost exactly the right amount of matter to make space-time flat.

flat universe (307) A model of the universe in which space-time is not curved.

flocculent (253) Describes a galaxy whose spiral arms have a woolly or fluffy appearance.

flux (14, 137) A measure of the flow of energy through a surface. Usually applied to light.

focal length (72) The focal length of a lens is the distance from the lens to the point where it focuses parallel rays of light.

folded mountain range (350) A long range of mountains formed by the compression of a planet's crust.

forward scattering (387) The optical property of finely divided particles to preferentially direct light in the original direction of the light's travel.

frequency (72) The number of times a given event occurs in a given time; for a wave, the number of cycles that pass the observer in 1 second.

galactic cannibalism (276) The theory that large galaxies absorb smaller galaxies.

galactic corona (242) The extended, spherical distribution of low-luminosity matter believed to surround the Milky Way and other galaxies.

galaxy (5) A large system of stars, star clusters, gas, dust, and nebulae orbiting a common center of mass.

Galilean moons (găl·ŭ·lē´ŭn) (383) The four largest satellites of Jupiter, named after their discoverer, Galileo.

gamma-ray burst (228) A sudden, powerful burst of gamma rays.

gas tail (418) A comet tail composed of ionized gas atoms released from the nucleus and carried outward by the solar wind, also called a type I comet tail.

gene (429) A unit of DNA — or sometimes RNA — information responsible for controlling an inherited physiological trait.

geocentric universe (44) A model universe with Earth at the center, such as the Ptolemaic universe.

geosynchronous satellite (62) A satellite that orbits eastward around Earth with a period of 24 hours and remains above the same spot on Earth's surface.

giant (141) Large, cool, highly luminous star in the upper right of the H–R diagram, typically 10 to 100 times the diameter of the sun.

global warming (349) The gradual increase in the surface temperature of Earth caused by human modifications to Earth's atmosphere.

globular cluster (192) A star cluster containing 100,000 to 1 million stars in a sphere about 75 ly in diameter, generally old, metal-poor, and found in the spherical component of a galaxy.

grand unified theories (GUTs) (311) Theories that attempt to unify (describe in a similar way) the electromagnetic, weak, and strong forces of nature.

granulation (114) The fine structure of bright grains covering the sun's surface.

grating (84) A piece of material in which numerous microscopic parallel lines are scribed. Light encountering a grating is dispersed to form a spectrum.

gravitational collapse (332) The process by which a forming body such as a planet gravitationally captures gas from its surroundings.

gravitational lensing (271) The process by which the gravitational field of a massive object focuses the light from a distant object to produce multiple images of the distant object or to make the distant object look brighter.

gravitational radiation (217) Disturbances in a gravitational field traveling at the velocity of light and carrying energy away from an object with a rapidly changing mass distribution.

gravitational redshift (224) The lengthening of the wavelength of a photon due to its escape from a gravitational field.

greenhouse effect (349) The process by which carbon dioxide in an atmosphere traps heat and raises the temperature of a planetary surface.

grooved terrain (383) Regions of the surface of Ganymede consisting of parallel grooves, believed to have been formed by repeated fracture of the icy crust.

ground state (97) The lowest permitted electron energy level in an atom.

GUTs *See* grand unified theories.

habitable zone (436) The region around a star within which a planet the size of Earth with a comparably thick atmosphere would have an average surface temperature between 0°C and 100°C, allowing liquid water to exist.

half-life (325) The time required for half of the atoms in a radioactive sample to decay.

halo (241) The spherical region of a spiral galaxy, containing a thin scattering of stars, star clusters, and small amounts of gas.

heat (99) Energy stored in a material as agitation among its particles.

heat of formation (332) In planetology, the heat released by infalling matter during the formation of a planetary body.

heavy bombardment (335) The intense cratering during the first 0.5 billion years in the history of the solar system.

heliocentric universe (46) A model of the universe with the sun at the center, such as the Copernican universe.

helioseismology (117) The study of the interior of the sun by the analysis of its modes of vibration.

helium flash (188) The explosive ignition of helium fusion that takes place in some giant stars.

Herbig–Haro object (169) A small nebula that varies irregularly in brightness, believed to be associated with star formation.

Hertzsprung–Russell (H–R) diagram (hĕrt´ sprŭng·rŭs·ŭl) (140) A plot of the intrinsic brightness versus the surface temperature of stars. It separates the effects of temperature and surface area on stellar luminosity and is commonly plotted as absolute magnitude versus spectral type but also as luminosity versus surface temperature or color.

homogeneity (306) The assumption that, on the large scale, matter is uniformly spread through the universe.

horizon (16) The circular boundary between the sky and Earth.

horizon problem (310) In cosmology, the circumstance that the primordial background radiation seems much more isotropic than can be explained by the standard big bang theory.

horizontal branch (193) In the H–R diagram, stars fusing helium in a shell and evolving back toward the red giant region.

hot dark matter (309) Dark matter made up of particles, such as neutrinos, traveling at or nearly at the speed of light.

hot spot (286) In geology, a place on Earth's crust where volcanism is caused by a rising convection cell in the mantle below. In radio astronomy, a bright spot in a radio lobe.

H–R diagram *See* Hertzsprung–Russell diagram.

HII region (160) A region of ionized hydrogen around a hot star.

Hubble constant *(H)* (268) A measure of the rate of expansion of the universe, the average value of velocity of recession divided by distance, presently believed to be about 70 km/s/Mpc.

Hubble law (268) The linear relation between the distances to galaxies and their velocity of recession.

Hubble time (300) One divided by the Hubble constant. The Hubble time is the age of the universe if it has expanded since the big bang at a constant rate.

hydrostatic equilibrium (175) The balance between the weight of the material pressing downward on a layer in a star and the pressure in that layer.

hypernova (229) Produced when a very massive star collapses into a black hole. Thought to be a possible source of gamma-ray bursts.

hypothesis (53) A conjecture, subject to further tests, that accounts for a set of facts.

IAU (380) International Astronomical Union, the international society of astronomers that is responsible for, among other things, deciding rules for the names of astronomical objects and their permanent surface features.

ice line (331) Boundary beyond which water vapor could freeze to form ice.

inflationary universe (310) A version of the big bang theory, derived from grand unified theories, that includes a rapid expansion when the universe was very young.

infrared radiation (72) Electromagnetic radiation with wavelengths intermediate between visible light and radio waves.

inner Lagrangian point (198) The point of gravitational equilibrium between two orbiting stars through which matter can flow from one star to the other.

instability strip (236) The region of the H–R diagram in which stars are unstable to pulsation. A star passing through this strip becomes a variable star.

interferometry (83) The observing technique in which separated telescopes combine to produce a virtual telescope with the resolution of a much larger-diameter telescope.

interstellar dust (159) Microscopic solid grains in the interstellar medium.

interstellar medium (159) The gas and dust distributed between the stars.

interstellar reddening (159) The process in which dust scatters blue light out of starlight and makes the stars look redder.

inverse square relation (60) A rule that the strength of an effect (such as gravity) decreases in proportion as the distance squared increases.

ion (96) An atom that has lost or gained one or more electrons.

ionization (96) The process in which atoms lose or gain electrons.

irregular galaxy (263) A galaxy with a chaotic appearance, large clouds of gas and dust, and both population I and II stars, but without spiral arms.

isotopes (96) Atoms that have the same number of protons but a different number of neutrons.

isotropy (ī·sŏt´rŭ·pē) (306) The assumption that, in its general properties, the universe looks the same in every direction.

joule (J) (jo͞ol) (100) A unit of energy equivalent to a force of 1 newton acting over a distance of 1 m. One joule per second equals 1 watt of power.

Jovian planet (326) Jupiter-like planet with a large diameter and low density.

Kelvin temperature scale (99) A temperature scale using Celsius degrees and based on zero at absolute zero.

kiloparsec (kpc) (240) A unit of distance equal to 1000 pc or 3260 ly.

Kirchhoff's laws (102) A set of laws that describe the absorption and emission of light by matter.

Kuiper belt (324) The collection of icy planetesimals believed to orbit in a region from just beyond Neptune out to 100 AU or more.

Lagrangian points (198) Points of gravitational stability in the orbital plane of a binary star system or of a planet and its moon.

large-impact hypothesis (353) The theory that the moon formed from debris ejected during a collision between Earth and a large planetesimal.

large-scale structure (314) The distribution of clusters and superclusters of galaxies in filaments and walls enclosing voids.

L dwarf (106) A main-sequence star cooler than an M star.

light curve (149) A graph of brightness versus time commonly used in analyzing variable stars and eclipsing binaries.

light-gathering power (74) The ability of a telescope to collect light, proportional to the area of the telescope's objective lens or mirror.

lighthouse model (213) The explanation of a pulsar as a spinning neutron star sweeping beams of electromagnetic radiation around the sky.

light pollution (76) The illumination of the night sky by waste light from cities and outdoor lighting, which prevents the observation of faint objects.

light-year (ly) (4) The unit of distance light travels in 1 year.

liquid metallic hydrogen (381) A form of liquid hydrogen that is a good electrical conductor, found in the interiors of Jupiter and Saturn.

lobate scarp (lō´bāt·skärp) (359) A curved cliff such as those found on Mercury.

Local Group (273) The small cluster of a few dozen galaxies that contains our Milky Way Galaxy.

look-back time (268) The amount by which you look into the past when you look at a distant galaxy, a time equal to the distance to the galaxy in light-years.

luminosity *(L)* (138) The total amount of energy a star radiates in 1 second.

luminosity class (143) A category of stars of similar luminosity, determined by the widths of lines in their spectra.

lunar eclipse (30) The darkening of the moon when it moves through Earth's shadow.

Lyman series (lī´mĕn) (103) Spectral lines in the ultraviolet spectrum of hydrogen produced by transitions whose lowest energy level is the ground state.

magma ocean (356) A deep layer of molten rock covering all or most of the surface of a planet.

magnetar (229) A class of neutron star having a very strong magnetic field.

magnetic carpet (116) The network of small magnetic loops that covers the solar surface.

magnetosphere (măg·nē´tō·sfîr) (382) The volume of space around a planet within which the motion of charged particles is dominated by the planetary magnetic field rather than the solar wind.

magnifying power (75) The ability of a telescope to make an image larger.

magnitude scale (13) The astronomical brightness scale. The larger the number, the fainter the star.

main sequence (141) The region of the H–R diagram running from upper left to lower right, which includes roughly 90 percent of all stars.

mare (mä´rā) (plural: **maria**) (352) One of the lunar lowlands filled by successive flows of dark lava, from the Latin for "sea."

mass (60) A measure of the amount of matter making up an object.

mass–luminosity relation (152) The more massive a star is, the more luminous it is.

Maunder butterfly diagram (môn´dŭr) (126) A graph showing the latitude of sunspots versus time, first plotted by W. W. Maunder in 1904.

Maunder minimum (môn´dŭr) (127) A period between 1645 and 1715 of less numerous sunspots and other solar activity.

megaparsec (Mpc) (265) A unit of distance equal to 1,000,000 pc.

metals (244) In astronomical usage, all atoms heavier than helium.

meteor (324) A small bit of matter heated by friction to incandescent vapor as it falls through Earth's atmosphere.

meteorite (325) A meteor that survives its passage through the atmosphere and strikes the ground.

meteoroid (325) A meteor in space before it enters Earth's atmosphere.

meteor shower (410) A multitude of meteors that appear to come from the same region of the sky, believed to be caused by comet debris.

midocean rise (350) One of the undersea mountain ranges that push up from the seafloor in the center of the oceans.

Milankovitch hypothesis (27) Suggestion that Earth's climate is determined by slow periodic changes in the shape of its orbit, the angle of its axis, and precession.

Milky Way (5) The hazy band of light that circles our sky, produced by the glow of our galaxy.

Milky Way Galaxy (5) The spiral galaxy containing our sun, visible in the night sky as the Milky Way.

Miller experiment (430) An experiment that attempted to reproduce the conditions under which life began on Earth and manufactured amino acids and other organic compounds.

millisecond pulsar (219) A pulsar with a pulse period of only a few milliseconds.

minute of arc (17) An angular measure, one-sixtieth of a degree.

model *See* scientific model.

molecular cloud (163) A dense interstellar gas cloud in which atoms are able to link together to form molecules such as H_2 and CO.

molecule (96) Two or more atoms bonded together.

morning star (26) Any planet visible in the sky just before sunrise.

M-type (413) The type of asteroid with reflectance spectra indicating metallic composition, possibly fragments of the cores of larger, differentiated parent bodies.

multicellular (432) An organism consisting of more than one cell.

multiringed basin (355) Large impact feature (crater) containing two or more concentric rims formed by fracturing of the planetary crust.

mutant (430) Offspring born with altered DNA.

nadir (16) The point on the celestial sphere directly below the observer; the opposite of the zenith.

nanometer (nm) (71) A unit of distance equaling one-billionth of a meter (10^{-9} m).

natural law (53) A theory that is almost universally accepted as true.

natural selection (430) The process by which the best traits are passed on, allowing the most able to survive.

neap tide (65) Ocean tide of low amplitude occurring at first- and third-quarter moon.

Near-Earth Objects (NEOs) (412) Asteroids (or, possibly, extinct comet nuclei) with orbits that cross Earth's orbit that can occasionally make close approaches to our planet.

nebula (159) A glowing cloud of gas or a cloud of dust reflecting the light of nearby stars.

neutrino (120) A neutral, massless atomic particle that travels at or nearly at the speed of light.

neutron (95) An atomic particle with no charge and about the same mass as a proton.

neutron star (211) A small, highly dense star, with radius about 10 km, composed almost entirely of tightly packed neutrons.

Newtonian focus (80) The optical design in which a diagonal mirror reflects light out the side of the telescope tube for easier access.

node (36) The point where an object's orbit passes through the plane of Earth's orbit.

nonbaryonic matter (308) Proposed dark matter made up of particles other than protons and neutrons (baryons).

north celestial pole (16) The point on the celestial sphere directly above Earth's North Pole.

north point (16) One of the four cardinal directions, the point on the horizon directly north.

nova (185) From the Latin, meaning "new" a sudden brightening of a star making it appear as a new star in the sky, believed to be associated with eruptions on white dwarfs in binary systems.

nuclear bulge (241) The spherical cloud of stars that lies at the center of spiral galaxies.

nuclear fission (119) Reactions that break the nuclei of atoms into fragments.

nuclear fusion (119) Reactions that join the nuclei of atoms to form more massive nuclei.

nucleus (of an atom) (95) The central core of an atom containing protons and neutrons that carries a net positive charge.

objective lens (72) In a refracting telescope, the long-focal-length lens that forms an image of the object viewed. The lens closest to the object.

objective mirror (72) In a reflecting telescope, the principal mirror (reflecting surface) that forms an image of the object viewed.

oblateness (381) The flattening of a spherical body, usually caused by rotation.

observable universe (299) The part of the universe that you can see from your location in space and in time.

occultation (396) The passage of a larger body in front of a smaller body.

Olbers's paradox (ôl´bŭrs) (297) The conflict between observation and theory about why the night sky should or should not be dark.

Oort cloud (ōrt) (417) The hypothetical source of comets, a swarm of icy bodies believed to lie in a spherical shell extending to 100,000 AU from the sun.

opacity (176) The resistance of a gas to the passage of radiation.

open cluster (192) A cluster of 100 to 1000 stars with an open, transparent appearance, usually relatively young and located in the disk of the galaxy. The stars are not tightly grouped.

open orbit (63) An orbit that carries an object away, never to return to its starting point.

open universe (307) A model of the universe in which the average density is less than the critical density needed to halt the expansion.

orbital resonances (386) Orbital periods of objects related by simple arithmetic relationships. For example, Io's orbital period around Jupiter is exactly $^1/_2$ of Europa's, and exactly $^1/_4$ of Ganymede's; as a result, the objects strongly influence each other gravitationally.

outflow channel (369) Geological feature on Mars produced by the rapid motion of floodwaters.

outgassing (333) The release of gases from a planet's interior.

ovoid (395) The oval features found on Miranda, a satellite of Uranus.

paradigm (49) A commonly accepted set of scientific ideas and assumptions.

parallax (44) The apparent change in position of an object due to a change in the location of the observer. Astronomical parallax is measured in seconds of arc.

parsec (pc) (pär´sĕk) (136) The distance to a hypothetical star whose parallax is 1 second of arc. 1 pc = 206,265 AU = 3.26 ly.

Paschen series (pä´shŭn) (103) Spectral lines in the infrared spectrum of hydrogen produced by transitions whose lowest energy level is the third.

penumbra (pĭ·nŭm´brŭ) (30) The portion of a shadow that is only partially shaded.

perigee (36) The point closest to Earth in the orbit of a body circling Earth.

perihelion (pĕr·ŭ·hē´lē·ŭn) (25) The orbital point of closest approach to the sun.

period–luminosity relation (237) The relation between period of pulsation and intrinsic brightness among Cepheid variable stars.

permafrost (368) Permanently frozen soil.

permitted orbit (96) One of the energy levels that an electron may occupy in an atom.

photon (71) A quantum of electromagnetic energy that carries an amount of energy that depends inversely on its wavelength.

photosphere (35) The bright visible surface of the sun.

planet (3) A small, nonluminous body formed by accretion in a disk around a protostar.

planetary nebula (191) An expanding shell of gas ejected from a star during the latter stages of its evolution.

planetesimal (plăn·ŭ·tĕs´ŭ·mŭl) (331) One of the small bodies that formed from the solar nebula and eventually grew into protoplanets.

plastic (348) A material with the properties of a solid but capable of flowing under pressure.

plate tectonics (350) The constant destruction and renewal of Earth's surface by the motion of sections of crust.

Plutino (403) One of the icy Kuiper belt objects that, like Pluto, is caught in a 3:2 orbital resonance with Neptune.

polar axis (81) In an equatorial telescope mounting, the axis that is parallel to Earth's axis.

poor galaxy cluster (273) An irregularly shaped cluster that contains fewer than 1000 galaxies, many of which are spiral, and no giant ellipticals.

population I (244) Stars rich in atoms heavier than helium, nearly always relatively young stars found in the disk of the galaxy.

population II (244) Stars poor in atoms heavier than helium, nearly always relatively old stars found in the halo, globular clusters, or the nuclear bulge.

positron (120) The antiparticle of the electron.

precession (18) The slow change in the direction of Earth's axis of rotation. One cycle takes nearly 26,000 years.

primary atmosphere (348) The original atmosphere of a planet, before removal or alteration and replacement by a later, secondary atmosphere.

primary lens (72) In a refracting telescope, the largest lens.

primary mirror (72) In a reflecting telescope, the largest mirror.

prime focus (80) The point at which the objective mirror forms an image in a reflecting telescope.

primeval atmosphere (375) Earth's first air.

primordial soup (431) The rich solution of organic molecules in Earth's first oceans.

prominence (36, 130) Eruption on the solar surface. Visible during total solar eclipses.

proper motion (237) The rate at which a star moves across the sky, measured in seconds of arc per year.

protein (428) Complex molecule composed of amino acid units.

proton (95) A positively charged atomic particle contained in the nucleus of an atom. The nucleus of a hydrogen atom.

proton–proton chain (120) A series of three nuclear reactions that builds a helium atom by adding together protons. The main energy source in the sun.

protoplanet (331) Massive object, destined to become a planet, resulting from the coalescence of planetesimals in the solar nebula.

protostar (165) A collapsing cloud of gas and dust destined to become a star.

pseudoscience (26) A subject that claims to obey the rules of scientivic reasoning but does not. Examples include astrology, crystal power, and pyramid power.

pulsar (212) A source of short, precisely timed radio bursts, believed to be spinning neutron stars.

pulsar wind (216) The breeze of high-energy particles flowing away from a spinning neutron star.

***P* wave** (346) A pressure wave. A type of seismic wave produced in Earth by the compression of the material.

quantum mechanics (96) The study of the behavior of atoms and atomic particles.

quasar (quasi-stellar object, or QSO) (kwā´zär) (285) Small, powerful source of energy believed to be the active core of a very distant galaxy.

quintessence (313) The postulated energy that fills empty space and drives the acceleration of the universe.

radial velocity (V_r) (109) That component of an object's velocity directed away from or toward Earth.

radiation pressure (335) The force exerted on the surface of a body by its absorption of light. Small particles floating in the solar system can be blown outward by the pressure of the sunlight.

radiative zone (121) The region inside a star where energy is carried outward as photons.

radio galaxy (283) A galaxy that is a strong source of radio signals.

radio interferometer (87) Two or more radio telescopes that combine their signals to achieve the resolving power of a larger telescope.

rays (354) Ejecta from meteorite impacts forming white streamers radiating from some lunar craters.

recombination (304) The stage in the early history of the universe within 300,000 years of the big bang when the gas became transparent to radiation.

reconnection (131) On the sun, the merging of magnetic fields to release energy in the form of flares.

red dwarf (141) A faint, cool, low-mass, main-sequence star.

redshift (108) A Doppler shift toward longer wavelengths caused by a velocity of recession.

reflecting telescope (72) A telescope that uses a concave mirror to focus light into an image.

reflection nebula (160) A nebula produced by starlight reflecting off dust particles in the interstellar medium.

refracting telescope (72) A telescope that forms images by bending (refracting) light with a lens.

reionization (304) The stage in the early history of the universe when ultraviolet photons from the first stars ionized the gas filling space.

resolving power (74) The ability of a telescope to reveal fine detail. Depends on the diameter of the telescope objective.

retrograde motion (44) The apparent backward (westward) motion of planets as seen against the background of stars.

revolution (22) Orbital motion about a point located outside the orbiting body. *See also* rotation.

ribonucleic acid (RNA) (429) Long carbon-chain molecules that use the information stored in DNA to manufacture complex molecules necessary to the organism.

rich galaxy cluster (273) A cluster containing over 1000 galaxies, mostly elliptical, scattered over a volume about 3 Mpc in diameter.

rift valley (350) A long, straight, deep valley produced by the separation of crustal plates.

ring galaxy (277) A galaxy that resembles a ring around a bright nucleus, believed to be the result of a head-on collision of two galaxies.

RNA *See* ribonucleic acid.

Roche limit (rōsh) (387) The minimum distance between a planet and a satellite that holds itself together by its own gravity. If a satellite's orbit brings it within its planet's Roche limit, tidal forces will pull the satellite apart.

Roche lobe (198) The volume of space a star controls gravitationally within a binary system.

Roche surface (198) The dumbbell-shaped surface that encloses the Roche lobes around a close binary star.

rotation (22) Motion around an axis passing through the rotating body. *See also* revolution.

rotation curve (242, 269) A graph of orbital velocity versus radius in the disk of a galaxy.

rotation curve method (269) A method of determining a galaxy's mass by observing the orbital velocities and radius of the orbits of stars in the galaxy.

RR Lyrae variable star (är·är·lī´rē) (236) Variable star with period of from 12 to 24 hours, common in some globular clusters.

Sagittarius A* (248) The powerful radio source located at the core of the Milky Way Galaxy.

Saros cycle (sĕ´rōs) (39) An 18-year, 11-⅓ day period after which the pattern of lunar and solar eclipses repeats.

Schmidt-Cassegrain focus (80) The optical design that uses a thin corrector plate at the entrance to the telescope tube. A popular design for small telescopes.

Schwarzschild radius (R_S) (schwôrts´ shēld) (223) The radius of the event horizon around a black hole.

scientific argument (28) An honest, logical discussion of observations and theories intended to reach a valid conclusion.

scientific method (5) The reasoning style by which scientists test theories against evidence to understand how nature works.

scientific model (15) A tentative description of a phenomenon for use as an aid to understanding.

scientific notation (3) The system of recording very large or very small numbers by using powers of 10.

secondary atmosphere (348) The gases outgassed from a planet's interior; rich in carbon dioxide.

secondary crater (354) Impact crater formed by debris ejected from a larger impact.

secondary mirror (80) In a reflecting telescope, a mirror that directs the light from the primary mirror to a focal position.

second of arc (17) An angular measure, one-sixtieth of a minute of arc.

seeing (75) Atmospheric conditions on a given night. When the atmosphere is unsteady, producing blurred images, the seeing is said to be poor.

self-sustaining star formation (254) The process by which the birth of stars compresses the surrounding gas clouds and triggers the formation of more stars, proposed to explain spiral arms.

semimajor axis, *a* (53) Half of the longest diameter of an ellipse.

SETI (438) The Search for Extra-Terrestrial Intelligence.

Seyfert galaxy (sē´fûrt) (283) An otherwise normal spiral galaxy with an unusually bright, small core that fluctuates in brightness, believed to indicate that the core is erupting.

Shapley–Curtis debate (240) The 1920 debate between Harlow Shapley and Heber Curtis on the nature of spiral nebulae. Curtis argued that they are other galaxies, and Shapley argued that they are internal to our own galaxy.

shepherd satellite (393) A satellite that, by its gravitational field, confines particles to a planetary ring.

shield volcano (364) Wide, low-profile volcanic cone produced by highly liquid lava.

shock wave (164) A sudden change in pressure that travels as an intense sound wave.

sidereal drive (sī·dîr´ē·ŭl) (81) The motor and gears on a telescope that turn it westward to keep it pointed at a star.

sidereal period (33) The time a celestial body takes to turn once on its axis or revolve once around its orbit relative to the stars.

singularity (222) The object of zero radius into which the matter in a black hole is believed to fall.

solar eclipse (34) The event that occurs when the moon passes directly between Earth and the sun, blocking your view of the sun.

solar nebula theory (321) The theory that the planets formed from the same cloud of gas and dust that formed the sun.

solar system (3) The sun and its planets, asteroids, comets, and so on.

solar wind (117) Rapidly moving atoms and ions that escape from the solar corona and blow outward through the solar system.

south celestial pole (16) The point on the celestial sphere directly above Earth's South Pole.

south point (16) One of the four cardinal directions, the point on the horizon directly south.

spectral class or type (104) A star's position in the temperature classification system O, B, A, F, G, K, M, based on the appearance of the star's spectrum.

spectral line (84) A line in a spectrum at a specific wavelength produced by the absorption or emission of light by certain atoms.

spectral sequence (104) The arrangement of spectral classes (O, B, A, F, G, K, M) ranging from hot to cool.

spectrograph (84) A device that separates light by wavelengths to produce a spectrum.

spectroscopic binary system (147) A star system in which the stars are too close together to be visible separately. We see a single point of light. Only in a spectrum can the two stars be detected.

spectroscopic parallax (144) The method of determining a star's distance by comparing its apparent magnitude with its absolute magnitude as estimated from its spectrum.

spectrum (71) An arrangement of electromagnetic radiation in order of wavelength or frequency.

spherical component (241) The part of the galaxy including all matter in a spherical distribution around the center (the halo and nuclear bulge).

spicule (spĭk′yo͞ol) (115) A small, flamelike projection in the chromosphere of the sun.

spiral arm (5, 241) Long spiral pattern in a disk galaxy composed of bright stars, star clusters, gas, and dust. Spiral arms extend from the center to the edge of the disk of spiral galaxies.

spiral galaxy (262) A galaxy with an obvious disk component containing gas; dust; hot, bright stars; and spiral arms.

spiral tracer (248) Object used to map a galaxy's spiral arms—for example, O and B associations, open clusters, clouds of ionized hydrogen, and some types of variable stars.

spring tide (65) Ocean tide of high amplitude that occurs at full and new moon.

standard candle (265) Object of known brightness that astronomers use to find distance—for example, Cepheid variable stars and supernovae.

star (3) A globe of gas held together by its own gravity and supported by the internal pressure of its hot gases, which generate energy by nuclear fusion.

star cluster (164) A group of stars that formed together and orbit a common center of mass.

starburst galaxy (275) A galaxy undergoing a rapid burst of star formation.

star-formation pillar (172) The column of gas produced when a dense core of gas protects the nebula behind it from the energy of a nearby hot star that is evaporating and driving away a star-forming nebula.

static (297) Unchanging, not evolving; for example, a static universe, which would have the same overall properties as a function of time, and not be expanding or contracting.

stellar model (177) A table of numbers representing the conditions in various layers within a star.

stellar parallax *(p)* (136) A measure of stellar distance. *See also* parallax.

stromatolite (430) A layered fossil formation caused by ancient mats of algae or bacteria, which built up mineral deposits season after season.

strong force (119) One of the four forces of nature, the strong force binds protons and neutrons together in atomic nuclei.

S-type (413) The type of asteroid with reflectance spectra indicating rocky compositions, possibly fragments of the mantles and crusts of larger, differentiated parent bodies.

subduction zone (350) A region of a planetary crust where one tectonic plate slides under another.

summer solstice (24) The point on the celestial sphere where the sun is at its most northerly point. Also, the time when the sun passes this point, about June 22, and summer begins in the Northern Hemisphere.

sunspot (113) Relatively dark spot on the sun that contains intense magnetic fields.

supercluster (314) A cluster of galaxy clusters.

supergiant (141) Exceptionally luminous star whose diameter is 10 to 1000 times that of the sun.

supergranule (115) Very large convective features in the sun's surface.

supernova (185) The explosion of a star in which it increases its brightness by a factor of about a million.

supernova (type Ia) (203) The explosion of a star, thought to be caused by the transfer of matter to a white dwarf.

supernova remnant (205) The expanding shell of gas marking the site of a supernova explosion.

***S* wave** (346) A shear wave. A type of seismic wave produced in Earth by the lateral motion of the material.

synchrotron radiation (sĭn′krŭ·trŏn) (204) Radiation emitted when high-speed electrons move through a magnetic field.

synodic period (sĭ·nŏd′ĭk) (33) The time a solar system body takes to orbit the sun once and return to the same orbital relationship with Earth, that is, orbital period referenced to Earth.

T dwarf (106) A very cool, low-mass star or brown dwarf located below the L stars on the main sequence.

temperature (99) A measure of the agitation among the atoms and molecules of a material, the intensity of heat.

terminator (352) The dividing line between daylight and darkness on a planet or moon.

terrestrial planet (326) An Earthlike planet—small, dense, rocky.

theory (53) A system of assumptions and principles applicable to a wide range of phenomena that has been repeatedly verified.

thermal energy (99) The energy stored in an object as agitation among its atoms and molecules.

tidal heating (383) The heating of a planet or satellite because of friction caused by tides.

tidal tail (276) A long streamer of stars, gas, and dust torn from a galaxy during its close interaction with another passing galaxy.

time dilation (224) The slowing of moving clocks or clocks in strong gravitational fields.

totality (31) The period during a solar eclipse when the sun's photosphere is completely hidden by the moon, or the period during a lunar eclipse when the moon is completely inside the umbra of Earth's shadow.

transition (103) The movement of an electron from one atomic energy level to another.

transits (339) Events in which a small object passes in front of a larger object from an observer's point of view; for example, when an extrasolar planet moving in its orbit passes in front of its parent star, blocking a small amount of the star's light.

triple-alpha process (174) The nuclear fusion process that combines three helium nuclei (alpha particles) to make one carbon nucleus.

Trojan asteroids (412) Asteroids sharing Jupiter's orbit, clustered around points 60 degrees ahead and behind the planet, named after characters in the Trojan War.

T Tauri star (tôrē) (168) A young star surrounded by gas and dust, believed to be contracting toward the main sequence.

turnoff point (192) The point in an H–R diagram at which a cluster's stars turn off of the main sequence and move toward the red-giant region, revealing the approximate age of the cluster.

type I comet tail *See* gas tail.

type I supernova (203) A supernova explosion caused by the collapse of a white dwarf.

type II comet tail *See* dust tail.

type II supernova (203) A supernova explosion caused by the collapse of a massive star.

ultraviolet radiation (72) Electromagnetic radiation with wavelengths shorter than visible light but longer than X rays.

umbra (ŭm′brŭ) (30) The region of a shadow that is totally shaded.

uncompressed density (330) The density a planet would have if its gravity did not compress it.

unified model (289) An attempt to explain the different types of active galactic nuclei using a single model viewed from different directions.

uniform circular motion (44) The classical belief that the perfect heavens could move only by the combination of uniform motion along circular orbits.

valley network (372) A system of dry drainage channels on Mars that resembles the beds of rivers and tributary streams on Earth.

variable star (234) A star whose brightness changes periodically.

velocity dispersion method (270) A method of finding a galaxy's mass by observing the range of velocities within the galaxy.

vernal equinox (24) The place on the celestial sphere where the sun crosses the celestial equator moving northward. Also, the time of year when the sun crosses this point, about March 21, and spring begins in the Northern Hemisphere.

vesicular basalt (vŭ·sĭk′yŭ·lŭr) (352) A porous rock formed by solidified lava with trapped bubbles.

visual binary system (147) A binary star system in which the two stars are separately visible in the telescope.

voids (314) Extremely large structural features in the universe, regions tens or hundreds of megaparsecs in size, containing few or no visible galaxies.

water hole (438) The interval of the radio spectrum between the 21-cm hydrogen radiation and the 18-cm OH radiation; these are likely wavelengths to use in the search for extraterrestrial life.

wavelength (70) The distance between successive peaks or troughs of a wave, usually represented by *l*.

wavelength of maximum intensity (λ_{max}) (99) The wavelength at which a perfect radiator emits the maximum amount of energy. Depends only on the object's temperature.

weak force (119) One of the four forces of nature, the weak force is responsible for some forms of radioactive decay.

west point (16) One of the four cardinal directions, the point on the horizon directly west.

white dwarf (141) Dying star at the lower left of the H–R diagram that has collapsed to the size of Earth and is slowly cooling off.

Widmanstätten pattern (wĭd´mŭn·stä·tŭn) (409) Bands in iron meteorites due to large crystals of nickel-iron alloys.

WIMPs (309) Weakly Interacting Massive Particles, a possible component of dark matter, consisting of one or more hypothetical types of subatomic particles that have mass, thus exerting gravitational force, but otherwise interact weakly, or not at all, with ordinary matter.

winter solstice (24) The point on the celestial sphere where the sun is farthest south. Also, the time of year when the sun passes this point, about December 22, and winter begins in the Northern Hemisphere.

X-ray burster (219) An object that produces repeated bursts of X rays.

Zeeman effect (127) The splitting of spectral lines into multiple components when the atoms are in a magnetic field.

zenith (16) The point on the sky directly above the observer.

zero-age main sequence (ZAMS) (180) The location in the H–R diagram where stars first reach stability as hydrogen-burning stars.

zodiac (26) A band centered on the ecliptic and encircling the sky.

Answers to Even Problems

Chapter 1
2. 3475 km; **4.** 1.08×10^8 km; **6.** about 1.2 s; **8.** about 75,000 yr; **10.** about 25

Chapter 2
2. 2800; **4.** The sun is 400,000 times brighter than the full moon.

Chapter 3
2. (a) new; (b) full; (c) first quarter; (d) third quarter; **4.** 14 days; **6.** 690 arc seconds or 11 arc minutes or about one-fifth of a degree; **8.** The eclipse would again be visible in Canada on 12 August 2026 [July 10, 1972 + 3 × (6585 $^1/_3$) days)].

Chapter 4
2. Ratio of maximum to minimum distance = 6.1:1. In the Ptolemaic model, Venus is always between Earth and the sun. In the diagram of the Ptolemaic universe, the same ratio is about 1.5:1; **4.** 32 years; **6.** 30.06 AU; **8.** The arrows are about the correct length.

Chapter 5
2. 3 m; **4.** Each of the Keck telescopes has 1.6 million times the lightgathering power of the human eye; **6.** No, the resolution of Galileo's telescopes was about 5.8 seconds of arc; **8.** 13 mm; **10.** 45 cm

Chapter 6
2. 150 nm; **4.** $2^4 = 16$; **6.** 250 nm; **8.** B, F, M, K; **10.** 0.6 nm

Chapter 7
2. 730 km; **4.** 9.0×10^{16} J; **6.** 0.2 kg; **8.** 3.6 times brighter; **10.** 400,000 yr

Chapter 8
2. 28 m
4.

m	M_V	*d* (pc)	*p* (seconds of arc)
7	7	10	0.1
11	1	1000	0.001
1	−2	40	0.025
4	2	25	0.040

6. about B8; **8.** Using Figure 8-10, M_v is estimated to be –7, so $d = 400$ pc. There is some uncertainty in this number in determining the absolute magnitude of the B0V star from the figure; **10.** a, c, c, c, d; **12.** 4.23×10^{11} m, 6.7×10^{10} m, 2.4×10^{31} kg (12 solar masses); **14.** A 4-solarmass main-sequence star has a luminosity of approximately 128 solar luminosities, a 9-solar-mass main-sequence star has a luminosity of approximately 2200 solar luminosities, and a 7-solar-mass main-sequence star has a luminosity of 910 solar luminosities.

Chapter 9
2. 20,000 AU or 0.097 pc; **4.** 100 K; **6.** In both reactions, the net result is that four ^{1}H nuclei are converted into one ^{4}He nucleus and energy. In the case of the proton–proton chain, six ^{1}H nuclei are put into the reaction, and the reaction returns two ^{1}H nuclei, one 4He nucleus, and energy (6 ^{1}H = 2 ^{1}H + 1 ^{4}He + energy, which is equivalent to 4 ^{1}H = 1 ^{4}He + energy). The CNO cycle inputs one ^{12}C nucleus and four ^{1}H nuclei and returns one ^{12}C nucleus, one ^{4}He nucleus, and energy (1 ^{12}C + 4 ^{1}H = 1 ^{12}C + ^{4}He + energy, which is equivalent to 4 ^{1}H = 1 ^{4}He + energy); **8.** 2000 nm = 2 micrometers; **10.** Estimating the absolute magnitude of an O6V star to be –5.6 implies a distance of 1600 pc. The answer is very dependent on the estimate for the absolute magnitude.

Chapter 10
2. about 310 million years; **4.** 1.8 ly; **6.** 32,000 years; **8.** 1.3×10^6 times; **10.** 1200 km/s

Chapter 11
2. 0.00096 AU if the total mass is 2 solar masses; **4.** 3 m, 1.1×10^{-25} m; **6.** 8.1 minutes; **8.** 3.0×10^8 m/s; photons

Chapter 12
2. 1.6%; **4.** 5000 pc; **6.** 25 pc; **8.** 20 times; **10.** 1500 K

Chapter 13
2. 2.6×10^6 pc (2.6 Mpc); **4.** 93 km/s; **6.** 28.6 Mpc; **8.** 165 million years

Chapter 14
2. 7.6 million years; **4.** 0.024 pc = 4900 AU; **6.** –28.5; **8.** 640 Mpc

Chapter 15
2. 57 km/s/Mpc; 17.5 billion years; 11.7 billion years; **4.** 1.6×10^{-27} kg/m^3 ; **6.** 76 km/s/Mpc

Chapter 16
2. It will look $(206{,}265)^2 = 4.3 \times 10^{10}$ times fainter, apparent magnitude about +23; **4.** about 2.3 times the half-life, or 3.0 billion years; **6.** large amounts of methane and water ices, mixed with substances condensing at higher temperatures; **8.** about 1300 impacts per 100 square meters per hour

Chapter 17
2. 16 times; 64 times; **4.** 0.10 arcseconds; **6.** 9.998 cm

Chapter 18
2. From Callisto, Jupiter's diameter would appear to be 4.3 degrees, and from Amalthea it would appear to be 45.0 degrees; **4.** 16.8 km/s; **6.** 7.3 s; **8.** 4.0 hours; **10.** 1.5×10^{22} kg

Chapter 19
2. 10^9; **4.** 10^4 km diameter; **6.** 2.9 days; **8.** 10^6 years at 10^4 AU; 0.30 km/s

Chapter 20
2. 8.7 cm; 0.87 mm; **4.** about 1.4 solar masses; **6.** 380 km; **8.** Answers will vary greatly. Using intermediate values between the two columns of the table, such as 2×10^{11}, 0.1, 0.1, 0.1, 0.1, 10 ^{6}, yields an answer of 20.

Index

Bold page numbers indicate key terms.

A

aberration, chromatic, **73**
absolute visual magnitude (M_V), **138,** 139
absolute zero, **99**
absorption lines, **101, 102**
absorption spectra, **101, 102,** 114
accretion, **331**
accretion disks, **198**–199
 and energy jets, 227–228
 and supermassive black holes, 289, 290, 291
achondrites, **410**
achromatic lenses, **73**
active galactic nuclei (AGN), **283**
active galaxies, **283**–285, 288–289
active optics, **81**
adaptive optics, **82**–83
Adhara, 144
albedo, **352**
Aldebaran, 12, 144
 locating, 13
Algol, 150–151
aliens, space, 438
Allen Telescope Array, 87
Almagest, 43, 45
Alpha Centauri, 4
 locating, 13
 spectral classification of, 106
Alpha Lyrae, 13
Alpher, Ralph, 301
Alpheratz, 11
Altair, 144
alt-azimuth mountings, **79, 81**
Andromeda, 11
Andromeda Galaxy, 234, 276
angstrom (Å), **71**
angular diameter, **14, 17,** 34
 of moon, 36, 37
angular distance, **14**
angular momentum, **198**
annular eclipse, **36**
anorthosite, **352**
antimatter, **303**
aphelion, **23,** 24–**25**
 and Earth's climate, 29
apogee, **36,** 63
apparent visual magnitudes (m_v), **13**–14
Aquarius, 22
Ariel, 395
Aristotelian universe, 43–46
Aristotle, 43
association, **164**
asterisms, **11**
asteroids, **324,** 335, 407, 412–413
 impacts on Earth, 421
 observations of, 414–415
 origin of, 413
 properties of, 412
astrology, 26
astronomer's triangulation method of measurement, 135–136
Astronomia Nova (New Astronomy), 52
astronomical unit (AU), **3**
astronomy
 classical, 43–46
 imaging systems in, 84
 from space, 89–91
astronomy constants, 447
astronomy units, 447
astrophysics, high-energy, 89–90
atmosphere(s)
 Earth's, 348–349, 352
 of Jovian planets, 380
 of Jupiter, 382–383, 384–385
 of Mars, 367–368
 solar, 113–119
 of Terrestrial planets, 345
 of Venus, 361–362
atmospheric windows, **72**
atom(s), 95–97
 in body, 321
 electron shells and, 96–97
 excitation of, 97–98
 ionized, **96**
 kinds of, 95–96
 model, 97
auroras, **128**
autumnal equinox, **23, 24**–25

B

Babcock model, **125**
background radiation, **301**–302, 315
Balmer absorption lines, 101, 104, 105, 106
Balmer series, **101**
Balmer thermometer, 101, 104
barred spiral galaxies, **262**
Barringer Meteorite Crater, 422
baryonic matter, 314
baryons, **308**
Bell, Jocelyn, 212
belt-zone circulation, **382**–383
Beta Pictoris, 337
Betelgeuse, 12, 13, 98
 locating, 13
 parallax of, 144
 spectral classification of, 106
big bang theory, **6, 300**–305
Big Dipper, 11, 12
big rip, **313,** 314
binary pulsars, 217–219
binary star system(s)
 and black holes, 226–227
 eclipsing, 149–**151**
 energy jets from, 227–228
 evolution of, 197–200
 mass of, 145, 146
 spectroscopic, **147**–149
 visual, **147**
binding energy, **96**
biological evolution, **427**
bipolar flows, **167**
birth line, **166**
black body radiation, **99,** 100
black dwarfs, **196**
black holes, 222–227
 candidates for, 226
 definition of, **223**
 energy jets from, 227–228
 escape velocity and, 222
 in galaxies, 285
 gamma ray bursts from, 228–230
 leaping into, 224
 and Milky Way, 248, 251
 Schwarzschild, 222–224
 search for, 224–227
 supermassive, 270–271. *See also* supermassive black holes
blueshift, **108**
Bok globules, **167, 168**
Brahe, Tycho, 50, 51
breccias, **352**
bright-line spectra, **101, 102**
brightness of stars, 13–14, 136–138
brown dwarfs, **179**

C

calibration, **237,** 238
Callisto, 383, 436
Caloris Basin, 359
Cambrian explosion, **433,** 434
Canis Majoris, 12
cannibal galaxy, Milky Way as, 276
capture hypothesis, **353**
carbonaceous chondrites, **409**–410
carbon dioxide
 in Earth's atmosphere, 348–349
 in Venus, 363, 366
Cassegrain focus, **79, 80**
Cassiopeia A, 205

catastrophic hypotheses, 321, 322
Venus in, 334
celestial equator, 16, 17
celestial sphere, 14–15, 18
center of mass, **61**
Cepheid variable stars, 234, 236–237
as distance indicators, 265, 266
kinds of, 244
Ceres, 403, 413
Chandrasekhar limit, **196,** 197
Chandra X-Ray Observatory, 90
charge-coupled devices (CCDs), **84**
Charon, 402
chemical evolution, **431**
Chicxulub, **421**
chondrites, **409,** 411
chondrules, **409**
Chort, 13
chromatic aberration, **73**
chromosphere, **35,** 36
activity of, 128–129
characteristics of, 115–116
circular velocity, **61, 62**
circumpolar constellations, **15, 17**
climate
astronomical influences on, 26–29
on Mars, 373
closed orbits, **61, 63**
closed universe models, **307**
cluster method, **269,** 270
clusters, 6. *See also* globular clusters; open clusters
galaxy, 273
CNA (carbon-nitrogen-oxygen) cycle, **171**
cold dark matter, **309**
Columbus, Christopher, 46
coma, of comet, **418**
Coma cluster, 273
comet(s), **324,** 328, 407, 408, 413, 414–420
characteristics of, 335
geology of nuclei, 416
and meteors, 411
observations of, 418–419
origin of, 416–417, 420
properties of, 416
Comet Halley, 408
Comet Wild 2, 416
communication
with distant civilizations, 437–440
radio, 437–438
comparison spectrum, **86**
composite volcanoes, **362, 364**
Compton Gamma Ray Observatory, 90, 228
condensation, **331**
condensation hypothesis, **353**
condensation sequence, **331**
conduction, 176
conservation of energy, law of, **175**
conservation of mass, law of, **175**
constellations, **11**
continental drift, **351**
continuous spectra, **101, 102**
convection, **115,** 176
convective zone, **121**
conversions, unit, 446
Copernican model, 46
Galileo and, 55, 56–58
Copernicus, Nicolaus, 46–**49**
core, 345
corona, **35,** 36
activity of, 128–129
solar, 116–117
coronae, **363**
coronagraphs, **116**
coronal holes, **128, 131**
coronal mass ejections (CMEs), **128, 131**
cosmic expansion, 299–300
cosmic jets, 286–287
cosmic microwave background radiation, **301**–302
cosmic rays, 90–**91**
cosmological constant, **312**
cosmological principle, **306**
cosmologists, **297**
cosmology, **296.** *See also* universe
basic concepts in, 296–297, 299–300
big bang theory in, 300–305
curvature of universe in, 315–317
dark energy in, 312–313
inflation and acceleration in, 310–313
large-scale structure in, 314–315
matter in, 307, 308–309
space and time in, 306–308
Coulomb barrier, **121**
Coulomb force, **96**
Crab Nebula, 204, 205, 213
pulsar, 216
cratering, **346,** 347
craters, 326
formation of, 356
of moon, 352, 354–355
Venus, 363, 366
critical density, **307,** 308, 310
crust, 345
of Earth, 348, 350–351
of Mars, 374
C-type asteroids, **413**
Curtis, Heber, 240
Cygnus X-1, 225–226

D

dark age, **304**
dark energy, 312–**313**
dark halo, **242**
dark-line spectra, **101, 102**
dark matter, **243**
clumping of, 314
in cosmology, 308–310
in galaxies, 271
gravitational lensing and, 271–273
dark nebulae, **159, 161**
dark side of the moon, 32
Davis neutrino experiment, 122, 123
debris disks, 336–338, **337**
deferent, **43, 45**
degenerate matter, **187**
Deimos, 375
density
of planets, 330–331
in star survey, 152–153
of universe, 307, 308, 310
density wave theory, **252**–254
De Revolutionibus Orbium Coelestium, 46–49
De Stella Nova (The New Star), 50
deuterium, **120,** 308
Devil's Hole, 28
Dialogo Dei Due Massimi Sistemi (Dialogue Concerning the Two Chief World Systems), 56, 57–58
diameter
angular, **14, 17,** 34
of galaxies, 268–269
linear, 34
of stars, 139–145
diamond-ring effect, **36**
Dicke, Robert, 301
differential rotation, **125**
differentiation, **332,** 333
in planetary evolution, **345,** 347
diffraction fringe, **74**
disk component, **240,** 241
disks
gas and dust, 336–338
of Milky Way, 242
distance indicators, **265**
distance scale, **265**
DNA, 427, 428–429, 430
Doppler effect, **107**–109
double-exhaust model, **285, 286**
double-lobed radio galaxies, **285**
double pulsar, 218
Drake, Frank, 439
Drake equation, **439,** 440
dust disks, 336–338

dust tail, **418**
dwarf ellipticals, 261, 278
dwarf irregulars, 261
dwarf planet, Pluto as, 402–404
dynamo effect, **124**

E

Earth, 345–352
age of, 325
atmosphere of, 348–349, 352
carbon dioxide of, 363
characteristics of, 347
climatic influences on, 26–29
composition of, 345
in Copernican system, 47, 49
crust of, 348, 350–351
density of, 330
development of, 345–346
distance from sun, 3, 113
end of, 200
impacts on, 420–423
interior of, 346–348
motion of, 347
orbit of, 27
rotational axis of, 18, 19, 24
size of, 3
surface of, 362
volcanoes on, **362**
east point, **16**
eccentricity (*e*), **53**
eclipses
lunar, **30**–31, **34**
predicting, 36, 39
solar, **34**–36
total, 31
eclipsing binary system, 149–**151**
ecliptic, **22,** 23
Einstein's theory of relativity, 306–307
electromagnetic force, 310–311
electromagnetic radiation, **70,** 99
electrons, **95,** 96, 98
electron shells, 96–97
elements, periodic table of, 452
ellipse, 52, **53**
elliptical galaxies, **262**
ellipticals, 261
emission lines, **101**
emission nebulae, **159, 160**
emission spectra, **101, 102**
Enceladus, 390–391, 436
energy
in cosmology, 306–310
nuclear binding, 119–120
in stellar structure, 174–175
energy flow, in planetary study, 345
energy jets
from compact objects, 227–228
and supermassive black holes, 289, 290
energy transport
in stellar structure, **176**
from sun, 121
epicycles, **43, 45,** 47, 48
Epsilon Eridani, 337
equants, **43, 45,** 47
equatorial mounting, **79, 81**
equinox, **24**
Eris, 403
escape velocity, **61, 63,** 222
Euler Crater, 354
Europa, 383, 435
evening star, **26**
event horizon, **223**
evolution, 430
of cells, 432–433
evolutionary hypotheses, 321, 322
Venus in, 334
evolutionary trace, **165**
excited atom, **98**
expanding universe, **299**
extrasolar planet, **338**
extraterrestrial radio signals, 438
eyepiece, **72**

F

false-color images, **84**
Favorite Stars, 13
field of view, **2**
filaments, **128, 130, 314**
filtergram, **115**
fission hypothesis, **353**
flares, **128**
flatness problem, **310,** 317
flat universe models, **307**
flavors, neutrino, **122,** 123
flocculent, **253**
flooding, **346,** 347
floppy mirrors, 81
flux, **14, 137**
focal length, **72**
foci, **53**
force(s), 310–311
Coulomb, **96**
in law of gravity, 60
forward scattering, **387**
fossils, 430
frequency, **72**
fusion, 173–174

G

galactic cannibalism, **273, 274**
galactic corona, **242**
galaxies, **5,** 259–281
active, 283–285, 288–289
clusters of, 273
colliding, 273, 274–275, 276
color of, 261
definition of, 6
distances to, 265
evolution of, 273–275, 278–279
family of, 260–264
farthest, 278–279
measuring properties of, 265–273
Milky Way. *See* Milky Way Galaxy
number of, 261, 264
shapes of, 260–261, 262–263
Galilean moons, **383,** 388
Galilei, Galileo, 55–58
gamma-ray bursts, **228**–230
gamma rays, 120
Gamow, George, 301
Ganymede, 383, 386, 436
gas, disks of, 336–338
gas tail, **418**
geocentric universe, **43, 44**
geologic time, 433–434
geosynchronous satellites, **61, 62**
giant stars, **141,** 142, 185–190
degenerate matter and, 186–187
expansion into, 185–186
helium fusion in, 187–189
and mass-luminosity relation, 152
star clusters and, 189–190
gibbous moon, 30
GLAST (Gamma-Ray Large Area Space Telescope), 90
global warming, **349**
globular clusters, **190, 192**
age of, 243
as distance indicators, 265
in Milky Way, 237
grand unified theories (GUTs), **311**
granulation, **114,** 115
grating, **84**
gravitation, 60
gravitational collapse, **332**
gravitational field, neutron star, 218–219
gravitational lensing, **271**
and dark matter, 271–273
in galaxy searching, 278
gravitational radiation, **217**
gravitational redshift, **224**
gravity, 59, 60
and inflationary universe, 310–311
Newton's law of, 60
Great Dark Spot of Neptune, 397, 400
Great Nebula of Orion. *See* Orion nebula
Great Square of Pegasus, 11, 12

greenhouse effect, **349**
and Venus, 363, 366
grindstone model of the universe, 234, 235
grooved terrain, **383**
ground state, **97**

H

habitable zone, **436,** 437
half-life, **325**
halo, **241,** 242
halo population stars, 244
Harmonice Mundi (The Harmony of the World), 52
heat, **99**
heat of formation, **332**
heavy bombardment, **335**
heavy-element fusion, 173–174
heliocentric universe, **46,** 47
helioseismology, **117,** 118
helium
in big bang theory, 303, 308
nucleus, 120
helium flash, **188**
helium fusion
in giant stars, 187–189
process, 173–174
Herbig-Haro objects, **167, 169**
Hercules X-1, 218, 219
Herman, Robert, 301
Herschel, Caroline, 234
Herschel, William, 234
Hertzsprung-Russell (H-R) diagram, 140–**143,** 192–193
high-energy astrophysics, 89–90
Hipparchus, 13
homogeneity, **306**
horizon, **310**
horizon problem, **310**
horizontal branch, **190**
horizontal branch stars, **193**
hot dark matter, **309**
hot spots, **285**
Hubble, Edwin, 240
Hubble constant (*H*), **268,** 300, 313
Hubble law, **268,** 269
in cosmic expansion, 299
Hubble space telescope, 89, 90
Hubble time, **300**
in universe acceleration, 313
Hulse, Russell, 217
Hydra, 402
Hydra cluster, 299
hydrogen, in big bang theory, 303
hydrogen fusion, 120–121
in giant star, 185–186
hydrostatic equilibrium, **175**–176
hypernova, **229,** 230
hypothesis, **53**

I

Iapetus, 391
ice ages, 26
ice line, **331**
inflationary universe, **310**
inflation of universe, 310–311
inflation theory, 317
Infrared Astronomical Satellite (IRAS), 89
infrared radiation, **72**
infrared telescopes, 77, 78
instability strip, **236**
INTEGRAL satellite, 90
intensity ratio, 15
interferometers, radio, **87**
interferometry, **83**–84
International Astronomical Union (IAU), **380**
interstellar absorption lines, **162**
interstellar dust, **159**
interstellar medium, 159, 162–164
interstellar reddening, **159**
intrinsic brightness of stars, 136–138
inverse square relation, **60**
Io, 383, 386, 388
ion, **96**
ionized atom, **96**
iron meteorites, **409,** 411
irregular galaxies, **262, 263**
Ishtar Terra, 362
isotopes, **96,** 325
in radioactive dating, 325
isotropy, **306**

J

James Webb Space Telescope, 89, 90
joule (J), **100**
Jovian planets, **324,** 326–327, 378–406
formation of, 333–334
growth of, 333
Jupiter, 381–388. *See also* Jupiter
and life, 435
Neptune, 397–402. *See also* Neptune
Pluto, 402–404. *See also* Pluto
Saturn, 388–394. *See also* Saturn
travel guide to, 379–381
Uranus, 394–397. *See also* Uranus
Jupiter, 381–388
atmosphere of, 382–383, 384–385
characteristics of, 381
composition of, 330
formation of, 334
history of, 388
interior of, 381–382
moons of, 55, 56, 383, 386, 435, 436
motion of, 381
rings of, 386–388
size of, 380

K

Kelvin temperature scale, **99**
Kepler, Johannes, 51–52
Keplerian motion, 242
Kepler's supernova, 52
Kepler's third law of planetary motion, 146
Kepler's three fundamental rules of planetary motion, 52–54
kiloparsec, **240**
Kirchhoff's laws, **101, 102**
second, 115
third, 114
Kuiper belt, **324**
and comets, 417
as debris disk, 337
and Pluto, 403, 404
Kuiper belt objects, 335

L

Lagrangian points, **198,** 200
Lakshmi Planum, 362
large-impact hypothesis, **353**
large-scale structure, **314**
L dwarfs, **106**
Leavitt, Henrietta, 236, 237
life
communication with distant civilizations and, 437–440
definition of, 426
nature of, 426–427, 430
origin of, 430–437
in other planetary systems, 436–437
on other worlds, 425–442
physical basis of, 426–427
in solar system, 434–436
light
interaction with matter, 97–101
spectrum of, 71–72
as wave and particle, 70–71
light curve, **149,** 150
light-gathering power, **74,** 75, 76, 77
lighthouse model, **213,** 214–215
light pollution, **76**
light-year (ly), **4**
linear diameter, 34
liquid metallic hydrogen, **381**
lithium, 308
Little Green Men (LGM), 212
lobate scarps, **359,** 361
Local Group cluster, **273**
look-back time, **268**

lower-main-sequence stars, 190–191, 196–197
luminosity
of galaxies, 268–269
of stars, 138–**139**
in star survey, 152–153
luminosity classes, **143**–144
lunar dust, 353
lunar eclipses, **30**–31, **34**
Lyman series, **101**

M

Mab, 397
magnetars, **229,** 230
magnetic carpet, **116**
magnetic field, sun, 124–125, 128
magnetosphere, **382**
magnification, telescope, 77
magnifying power, **75**
magnitudes, 15
magnitude scale, **13**
main sequence, **141**
main-sequence stars, 178–181
properties of, 448
man in the moon, 29
mantle, 345
Mare Imbrium (Sea of Rains), 356, 358
Mare Orientale, 359
maria, **352,** 357
Mars, 367–375
age of, 328
atmosphere of, 367–368
characteristics of, 369
density of, 330
evidence of water in, 369, 372–374
history of, 374–375
life on, 436
moons of, 375
motion of, 369
surface exploration of, 368–369
volcanoes, 363, 369
mass, **60**
of galaxies, 269–270
in law of gravity, 60
of Milky Way galaxy, 242–243
in star survey, 152–153
in stellar structure, 174–175
mass-luminosity relation, **152, 153**
in stars, 178–179
mass transfer and binary stars, **198**
mathematical models, 177
matter
baryonic, 314
in cosmology, 306–310
dark. *See* dark matter
degenerate, **187**
interaction with light, 97–101
nonbaryonic, **308**
Mauna Loa, 369
Maunder butterfly diagram, **124,** 125, **126**
Maunder minimum, **124, 127**
Maxwell, James Clark, 311
Maxwell Montes, 362
medium-mass stars, 190–191
megaparsec (Mpc), **265**
Mendel, Gregor, 7
Mercury, 358–361
distance to sun, 4
characteristics of, 359
density of, 330
greatest elongations of, 451
history of, 359, 361
motion of, 359
orbit of, 323
size of, 3
spacecraft visiting, 358–359
message, anticoded, **437**
metals, **244**
meteor(s), **324**–325, 409
definition of, 409
origins of, 410–411
meteorite(s), **325,** 407
and age of solar system, 329
definition of, 409
impact on Earth, 421
and moon's crater, 352, 353, 354, 355
origins of, 410–411
types of, 409–410
meteoroids, **325,** 409
meteor showers, **410,** 450
Milankovitch, Milutin, 27
Milankovitch hypothesis, **27,** 28
Milky Way Galaxy, **5,** 233–257
age of, 243–244
analysis of, 240–242
Cepheid variable stars in, 234, 236–237
discovery of, 234, 236–237
history of, 246–247
mass of, 242–243
nucleus of, 248
origin of, 243–247
size of, 237, 239–240
spiral arms and, 248–249, 252–254
star formation in, 254
Miller, Stanley, 430, 431
Miller experiment, **430,** 431
millisecond pulsars, **219**
minute of arc, **17**
Miranda, 395
mirrors, 72, 73, 81
molecular clouds, **163**
molecule, **96**
moon(s), 335
of Jupiter, 383, 386, 388
of Mars, 375
of Neptune, 400–401
of Saturn, 389–391
of Uranus, 395
moon of Earth, 352–357
characteristics of, 353
common misconceptions about, 30
craters in, 354–355
cycles of, 29–39
eclipses of, 34–36, 38, 39
geology of, 352–353
gravity of, 61
history of, 353, 356–357
lunar eclipses of, 30–31, 34
motion of, 29, 353
orbit of and nodes, 36, 39
origin of, 353
phases of, 29–30, 32–33
moon-phase dial, 30
moon rocks, ages of, 328, 356
morning star, **26**
motion, Newton's laws of, 58–60
M-type asteroids, **413**
multicellular life forms, **432**–433
mutants, **430**
Mysterium Cosmographicum, 52

N

nanometer (nm), **71**
natural law, **53**
natural selection, **430**
neap tides, **65**
Near-Earth Objects (NEOs), **412**
nebulae, **159, 191,** 196. *See also specific types*
formation of, 194–195
Neptune, 397–402
characteristics of, 401
formation of, 334
history of, 402
moons of, 400–401
motion of, 401
rings of, 396, 397, 398–399, 402
size of, 380
Nereid, 400, 402
neutrinos, **120**
and dark matter, 309
solar, counting, 121–123
neutrons, **95,** 96
neutron stars, **211**–222
binary pulsars and, 217–219
characteristics of, 211, 212
discovery of, 211
energy jets from, 227–228

neutron stars *(continued)*
fastest pulsars and, 219–220
gamma ray bursts from, 228–230
mass of, 211
model pulsar and, 213, 214–215
pulsar discovery and, 212–213
pulsar evolution and, 213, 216–217
pulsar planets and, 220–221
theoretical prediction of, 211–212
new moon, 32
Newton, Isaac, 58–60
Newtonian focus, **79, 80**
Newtonian universe, 65–66
Newton's law of motion, 58–60
Newton's version of third law, 146
NGC 1275, 285
Nix, 402
nodes, of moon's orbit, **36**
nonbaryonic matter, **308,** 309, 310
north celestial pole, **16,** 19
north point, **16**
North Star, 106
novae, **185,** 199–200
nova explosions, 200
nuclear binding energy, 119–120
nuclear bulge, **241**
and supermassive black holes, 291
nuclear fission reactions, **119**
nuclear fusion, **119**
in massive stars, 201–202
in sun, 119–123
and supernova explosions, 202–203
nuclear reactions
in big bang theory, 303, 308
in stars, 174
nuclei, comets, 416
nucleus, **95**
Milky Way galaxy, 248

O

objective lens, **72**
objective mirror, **72**
oblateness, **381**
observable universe, **299**
occultations, **396**
Olbers's paradox, **297**
Olympus Mons, 369
Oort, Jan, 416
Oort cloud, **417**
opacity, of gas, **176**
open clusters, **190**
age of, 243
open cluster star, **192**
open orbits, **61, 63**
open universe models, **307**
optical telescopes, 72–74
optics, adaptive, **82**–83
orbit, electron, 96, 97
orbital motion, Isaac Newton and, 58–66
orbital resonances, **386**
orbits. *See specific planets*
Orion, 12, 13, 98
spectral classification of, 106
Orion nebula, 170–171, 172–173
outflow channels, **369,** 372
outgassing, **333**
ovoids, **395**
oxygen, in Earth's atmosphere, 349, 352
ozone (O_3), 352

P

paradigm, **49**
parallax, **43, 44, 135,** 136
and distance, 137
measuring, 136
spectroscopic, **144**
and Tycho's supernova, 50
parked car effect, 349
parsec (pc), **136**
Paschen series, **101, 103**
Pauli exclusion principle, 187
penumbra, Earth's, **30, 31,** 34, 35
Penzias, Arno, 301
perigee, **36,** 63
perihelion, **23,** 24–25
and Earth's climate, 29
periodic table of elements, 452
period-luminosity relation, **237**
permafrost, **368**
permitted orbits, **96**
Perseid meteor shower, 410
Perseus galaxy cluster, 285
Phobos, 375
Phoebe, 391
photographic plates, 84
photons, **71**
and atom excitation, 98
energies of, 72
photosphere, **35,** 36, 98, 99
components of, 113–115
vibrations beneath, 117–119
Planck's constant, 71
planet(s). *See also specific planets*
in comparison, 344
composition of, 330
density of, 330–331
detecting, 339, 340
developmental stages of, 345–346, 347
extrasolar, **338**
formation of, 333–334
growth of, 333
Jovian. *See* Jovian planets
motion of, 23, 26
orbiting other stars, 336–340
outer, 380. *See also* Jovian planets
properties of, 449
pulsar, 220–221
retrograde motion of, **43, 44**
size of, **3**
types of, 324, 326–327
planetary motion, 49–55
planetary nebulae, **191,** 194–195, 196
planetesimals, **331**–332, 333
characteristics of, 335
comets as, 417
planet transit, **339**
plastic, **348**
plate tectonics, **350,** 351
Venus and, 366, 367
Plato, 43
Plum Crater, 355
plutinos, 403–404
Pluto, 402–404
polar axis, **79, 81**
Polaris, 15
in binary system, 147
locating, 13
spectral classification of, 106
poor galaxy clusters, **273**
population I stars, **244**
population II stars, **244**
positron, **120**
potassium, 325
powers of telescope, 74–77
powers of ten notation, 446
precession, **18,** 19
and Earth's climate, 29
pressure-temperature thermostat, 174
primary atmosphere, Earth's, **348**
primary lens, **72**
primary mirror, **72**
prime focus, **79, 80**
primordial rings, **394**
primordial soup, **431**
prism, 84
Project Phoenix, 438
prominences, **36, 128, 130**
proper motions, stars, **237**
proton-proton chain, **120**
protons, **95,** 96
protoplanets, **332**–333
protostar(s), **165,** 164–166
pseudosciences, **26**
PSR1257+12, 221
Ptolemaic model, 45
Copernican model and, 46, 49
Venus in, 56

Ptolemy, Claudius, 13, 43
pulsar(s), **212**
  binary, 217–219
  discovery of, 212–213
  evolution of, 213, 216–217
  fastest, 219–220
  model of, 213, 214–215
pulsar planets, 220–221
pulsar PSR1913+16, 217
pulsar PSRJ1740-5340, 220
pulsar wind, **216**
P waves, **346**

Q

quantum mechanics, **96,** 97
  and giant stars, 187
quasars, 285, 288–289, 292
quintessence, **313**

R

radial velocity, **109**
radiation
  background, **301**–302, 315
  definition of, 70
  energy transport by, 176
  from heated object, 98–101
  infrared, 77, 78
radiation pressure, **335**
radiative zone, **121**
radioactive dating, 325, 329
radio communication, 437–438
radio galaxies, **283**
radio interferometer, **87**
radio lobes, 286–287
radio maps, of spiral arms, 249, 252
radio telescopes, 86–89
radio waves, 72
recombination, **304**
reconnection, **128**
red dwarfs, **141,** 190
redshift(s), **108**
  in background radiation, 302
  in cosmic expansion, 299
  cosmological, 306–307
  quasar, 288, 292
reflecting telescopes, **72,** 73
reflection nebulae, **159, 160**
refracting telescopes, **72**–73
reionization era, **304**
resolution limit, 76
resolving power, **74,** 75, 77
retrograde motion, **43, 44**
  in Copernican system, 47, 49
revolution
  definition of, **22**
  of planets, 323–324
Revolutions of the Celestial Spheres, 46–49
rich galaxy clusters, **273**
Rigel, 98
  locating, 13
  luminosity of, 144
  spectral classification of, 106
ring galaxies, **273**
rings, 335, 380
  of Jupiter, 386–388
  of Neptune, 396, 397, 398–399, 402
  of Saturn, 391, 392–393
  of Uranus, 396–397, 398–399
Roche limit, **387,** 388
Roche lobes, **198**
Roche surface, **198**
rocky core, **382**
rotation
  definition of, **22**
  of planets, 323–324
rotation curve, **242, 269**
rotation curve method, **269**
RR Lyrae variable star, **236**
Rudolphine Tables, 51, 52, 54–55

S

Sagittarius, 22
  as center of Milky Way, 248, 250–251
Sagittarius Dwarf, 273
Saros cycle, **39**
satellite systems, 380–381
Saturn, 388–394
  characteristics of, 391
  formation of, 334
  history of, 391, 394
  moons of, 389–391, 436
  motion of, 391
  rings of, 391, 392–393
  size of, 380
Schmidt, Maarten, 285
Schmidt-Cassegrain focus, **79, 80**
Schwarzschild, Karl, 223
Schwarzschild black holes, 222–224
Schwarzschild radius (R_s), **223**
science, **7**
  and religion, 426
scientific argument, **28,** 29
scientific method, **7**
scientific models, **15,** 18
scientific notation, **3**
seasons, cycle of, 23
secondary atmosphere, Earth's, **348**
secondary mirror, **79**
second of arc, **17**
seeing, **75**
seismic waves, 346–347
self-sustaining star formation, **254**
semimajor axis (*a*), **53**
SETI (Search for Extra-Terrestrial Intelligence), **438**
Seyfert, Carl K., 283
Seyfert galaxies, **283**–284
Shapley, Harlow, 237, 239
Shapley-Curtis debate, **240**
shield volcanoes, **362, 364**
shock wave, **164**
Shoemaker-Levy 9, 420
shooting stars, 325
sidereal drive, **79**
sidereal period, **30, 33**
Sidereus Nuncius (The Sidereal Messenger), 55
singularity, **222**
Sirius, 12, 13, 144
  in binary system, 147
  locating, 13
  spectral classification of, 106
  white dwarf of, 196
slow surface evolution, **346,** 347
small-angle formula, **34**
small galaxies, 261
SN1987A, 204, 205
solar activity, 123–131
solar eclipses, **34**–36
solar flares, **131**
solar nebula
  chemical composition of, 329–330, 331
  clearing of, 335
solar nebula theory, **321**–322, **325,** 340
  planet formation in, 334
solar system, **3,** 320–342
  age of, 325, 328–329
  definition of, 6
  life in, 434–436
  orbiting of stars and, 336–340
  origins of, 320, 321–322
  outer. *See* Jovian planets
  planet formation in, 329–336
  principal satellites of, 450
  survey of, 323–325, 328–329
solar wind, **117,** 335
solstice, **24**
sound, 70, 107
south celestial pole, **16**
south point, **16**
space, astronomy from, 89–91
space debris, 324–325
space-time concept, 306–307
spectral classes, **104**
spectral classification, 104–106
spectral lines, **84, 86,** 101
spectral sequence, **104**
spectral types, **104**

spectrograph, **84,** 86
spectroscopic binary system, **147**–149
spectroscopic parallax, **144**
spectrum, **71**
of sun's corona, 116
speed of light, 71
spherical component, **241**
Spica, 13
spicules, **115,** 116
spiral arms, **5, 241**
of Milky Way, 248–249, 252–254, 255
spiral galaxies, 261, **262,** 277, 278
spiral tracers, **248,** 249
Spitzer Space Telescope, 89
spring tides, **65**
Square Kilometre Array, 87
SS433, 227, 228
standard candles, **265**
star(s), 95, 134–157, 158–183, 184–209
absolute visual magnitude of, 138
brightness of, 13–14, 448
Cepheid variable, 234, 236–237
deaths of, 184, 190–191, 196–197, 201–206
diameters of, 139–145
evolution of binary systems of, 197–200
formation of, 158, 163–167, 170, 172–173
fusion in, 173–174
giant, 185–190
interstellar medium of, 159, 162–163
intrinsic brightness of, 136–138
life expectancies of, 179, 180, 181
lower main sequence death(s), 190–191, 196–197
luminosity of, 138–**139,** 140
main-sequence, 178–181
masses of, 145–151
massive, deaths of, 201–206
measuring distances to, 135–136
names of, 11–13
nearest, 449
nebulae and, 160–161, 168–169, 172–173
neutron. *See* neutron stars
nuclear reactions in, 174
Orion nebula and, 170–171
planets orbiting, 336–340
shooting, 325
spectral classification of, 104–106
stellar structure in, 174–178
survey of, 151–156
temperature of, 139–140
travel between, 437
twinkling, 75
starburst galaxies, **277**
star charts, 13
using, 453–465
star cluster(s), **164**
age of, 243, 244
evolution of, 189–190
H-R diagrams of, 192–193
in Milky Way, 237, 239
star 51 Pegasi, 338
star formation
in Orion nebula, 170–171, 172–173
in spiral arms, 254
starlight, factors in, 98–99
star trails, 16
static universe, **297**
Stefan-Boltzmann law, 100
stellar models, 176–178, **177**
stellar parallax (*p*), **136**
stellar populations, in Milky Way, 244
stellar spectra, 101–109
Balmer thermometer in, 101, 104
chemical composition in, 106–107
classification of, 104–106
Doppler effect in, 107–109
formation of, 101
stellar structure, 174–176
energy transport in, 176
hydrostatic equilibrium, 175–176
mass and energy in, 174–175
models of, 176–178
stony-iron meteorites, **409, 410,** 411
stony meteorites, **409,** 410, 411
stromatolites, **430,** 432
strong force, **119**
and inflationary universe, 310–311
S-type asteroids, **413**
subduction zone, **350**
summer solstice, **23, 24**–25
sun, 112–113
activity of, 123–131
age of, 329
annual motion of, 22–23
atmosphere of, 113–119
characteristics of, 113
cycle of, 22–26
distance of Earth to, 3
distance of Mercury to, 4
distance of Venus to, 4
energy transport in, 121
magnetic cycle of, 124–125, 128
as mass-sequence star, 180
mathematical models of, 200
nuclear fusion in, 119–123
observing, 123–124
rotation of, 323
spectrum of, 114
star luminosity vs., 138
temperature of, 113
tidal effect of, 64–65
total and annular eclipses of, 37
sunspots, **113,** 124
cycle of, 125, 126–127
Galileo and, 56
magnetic field and, 125
superclusters, 6, **314**
supergiant stars, **115,** 142, 144
luminosity of, 144
and mass-luminosity relation, 152
supergranules, **115**
supermassive black holes, 289–292
disks and jets in, 289
in galaxies, 270–271, 285
origins of, 290–291
supernovae, **185**
in distant galaxies, 265
observations of, 204–206
pulsars in, 216, 217
types of, **203**
in universe acceleration, 313
supernova explosions, 164, 202–203
supernova remnant, **205**
surveyor's triangulation method of measurement, 135
S waves, **346**–347
synchrotron radiation, **204**
synodic period, **30, 33**
Système International d'Unités (SI units), 445–446

T

Tauri stars, **166, 168**
Taylor, Joseph, 217
T dwarfs, **106**
telescope(s), 70
adaptive optics and, 82–83
above atmosphere, 89–91
Galileo's observations through, 55–56
giant, 79
infrared, 77, 78
interferometry in, **83**–84
new-generation, 78–82
optical, 72–74
powers of, 74–77
radio, 86
selection of, 78
temperature, **99**
in radiation, 99
stars, Balmer lines and, 101
of sun, 113
temperature scales, 446
terminator, **352**
Terrestrial planets, **324,** 326–327, 343–377
composition of, 345

developmental stages of, 345–346 347
Earth, 345. *See also* Earth
formation of, 334
growth of, 333
Mars, 367–369, 372–375
Mercury, 358–361
moon and, 352–353, 356–357. *See also* moon of Earth
travel guide to, 344–345
Venus, 361–363, 366–367
Terzan 5ad, 219
theory, **53**
thermal energy, **99**
Theta Leonis, 13
tidal heating, **383,** 386, 388
tidal tails, **273, 274,** 290
tides, 61, 64–65
time, geologic, 433–434
time dilation, **224**
Titan, 389–390, 394, 436
total eclipse, 31
totality, **31,** 35, 36
transistion, **101**
transit, planet, **339**
Trapezium, 170
travel, between stars, 437
triple-alpha process, **174**
Triton, 400–401, 402
Trojan asteroids, **412**
turnoff point, **190**
Two-Degree-Field Redshift Survey, 314–315
Tycho Crater, 354
Tycho's model, 50, 51
Tycho's supernova, 50
type I (gas) tail, **418**
type II (dust) tail, **418**

U

UFOs, 438
ultraviolet waves, **72**
umbra, Earth's, **30, 31,** 34, 35
uncompressed densities, **330**
unified model of active galaxy cores, **289**–290
uniform circular motion, **43, 44**
in Copernican system, 47, 48
units, 445–452
universe
acceleration of, 311–312, 313
age of, 300
cosmic expansion of, 299–300
curvature of, 315–317
definition of, 6
edge-center problem in, 297
expanding, **306**
introduction to, 296–300
principles of. *See* cosmology
universe models, 307–308
Uranus, 394–397
characteristics of, 395
formation of, 334
history of, 397
moons of, 395
motion of, 395
rings of, 396–397
rotation of, 323, 324, 334
size of, 380
Urey, Harold, 430, 431
Ursa Major, 11

V

Valles Marineris, 369
valley networks, **372**
variable stars, **234**–235
Cepheid, 234, **235,** 236
Vega, 13, 144
debris disk of, 337
locating, 13
spectral classification of, 106
velocity, escape, 222
velocity dispersion method, **270**
Vel pulsar, 216
Venus, 361–367
atmosphere of, 361–362
characteristics of, 363
density of, 330
distance to sun, 4
Galileo's observations of, 56
greatest elongations of, 451
history of, 363, 366
motion of, 363
in Ptolemaic model, 56
rotation of, 323, 324, 334
size of, 3
suface of, 362–363
volcanoes, 363–364
vernal equinox, **23, 24**–25
Very Large Array (VLA), 87
Very Long Baseline Array (VLBA), 87
vesicular basalt, **352**
Vesta, 413, 415
Virgo cluster, 299
visual binary systems, **147**
voids, **314**
volcanoes, **362**
on Mars, 369, 374, 375
on Venus, 362, 365

W

water, on Mars, 368, 369, 372–375
water hole, **438**
wavelength, **70**
wavelength of maximum intensity (λ_{max}), **99**
weak force, **119**
and inflationary universe, 310–311
west point, **16**
white dwarfs, **141,** 142, 144
characteristics of, 196–197
masses of, 153
novae and, 199
vs. pulsar, 212–213
and supernovas, 203
Widmanstätten patterns, **409,** 411
Wien's law, 100
Wilson, Robert, 301
WIMPs (weakly interacting massive particles), **309**
winter solstice, **23, 24**–25

X

X-ray binaries, 225, 226, 227
X-ray bursters, **219**
X-rays, 72
XTEJ1739-285 pulsar, 219
XTEJ1751-305, 220

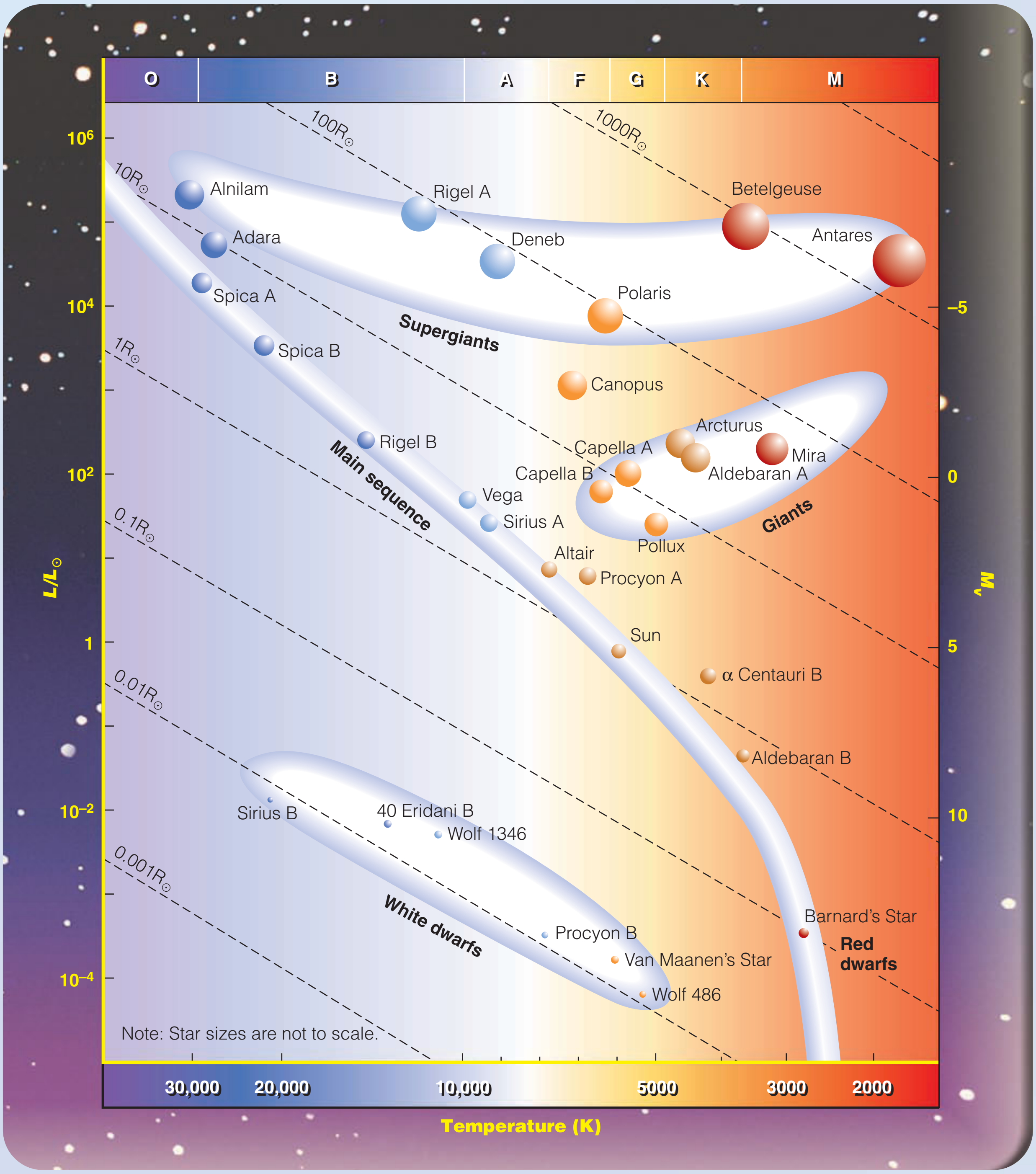
Y

year, definition of, 22

Z

Zeeman effect, **124, 127**
zenith, **16**
zero-age main sequence (ZAMS), **180**
zodiac, **26**

Flash Reference:
H–R Diagram



The H–R diagram is the key to understanding stars, their birth, their long lives, and their eventual deaths. Luminosity ($L/L_\odot$) refers to the total amount of energy that a star emits in terms of the sun's luminosity, and the temperature refers to the temperature of its surface. Together, the temperature and luminosity of a star locate it on the H–R diagram and tell astronomers its radius, its family relationships with other stars, and a great deal about its history and fate.

Flash Reference:
Comparative Planetology

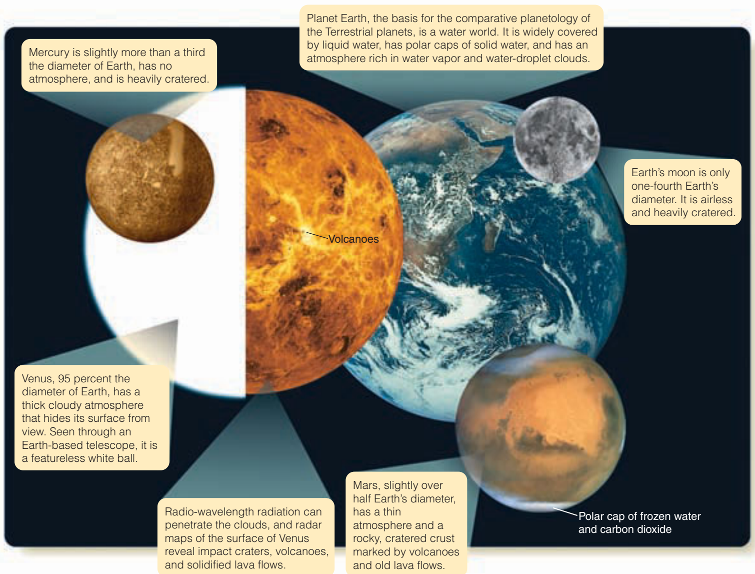
The terrestrial or Earthlike planets lie very close to the sun, and their orbits are hardly visible in a diagram that includes the outer planets.

Mercury, Venus, Earth and its moon, and Mars are small worlds made of rock and metal with little or no atmospheric gases.

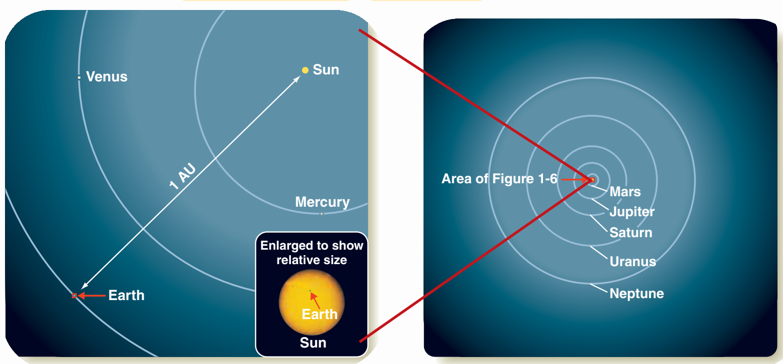
The outer worlds of our solar system orbit far from the sun. Jupiter, Saturn, Uranus, and Neptune are Jovian or Jupiter-like planets much bigger than Earth. They contain large amounts of low-density gases.

Pluto is one of a number of small, icy worlds orbiting beyond Neptune. Astronomers have concluded that Pluto is not really a planet and now refer to it as a dwarf planet.

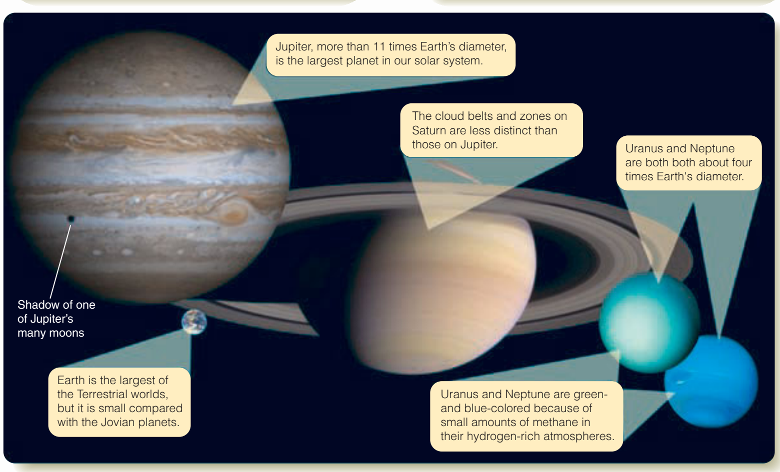
The Terrestrial Worlds



Planetary Orbits



The Outer Worlds

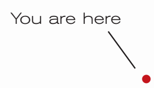


• *Horizons* readers: See page 344 for the terrestrial planets. See pages 3 and 4 for the two orbital diagrams. See page 380 for the outer worlds. *Astronomy* readers: See page 132 for the terrestrial planets. See pages 3 and 4 for the two orbital diagrams. See page 168 for the outer worlds.

Flash Reference:
Arrows

This book is designed to use arrows to alert you to important concepts in diagrams and graphs. Some arrows point things out, but others represent motion, force, or even the flow of light. Look at arrows in the book carefully and use this Flash Reference card to catch all of the arrow clues.

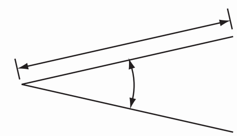
Point at things:



Force:

Process flow:

Measurement:

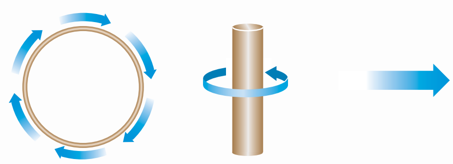


Direction:

Radio waves, infrared, photons:

Motion:

Rotation 2-D *Rotation 3-D* *Linear*



Light flow:

Updated arrow style