Astronomy

Rachel Mastrapa

De Anza College

Physics

CC BY 4.0



Astronomy 2e Openstax Dr. Mastrapa

This book is a remix of Openstax Astronomy, created for ASTR 4 at De Anza College.

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden borne from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and were supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

The LibreTexts name, logo, and book covers are trademarked by LibreTexts, Inc. (a 501(c)(3) not-for-profit organization) and protected against unauthorized use.

This text was compiled on 10/20/2026

TABLE OF CONTENTS

Licensing

1: Introduction to Astronomy

1.1: Introduction

1.2: The Nature of Astronomy

1.3: The Nature of Science

1.4: The Laws of Nature

1.5: Numbers in Astronomy

1.6: Consequences of Light Travel Time

2: The Universe

2.1: A Tour of the Universe

2.2: The Universe on the Large Scale

2.3: The Universe of the Very Small

2.4: A Conclusion and a Beginning

3: Ancient Astronomy

3.1: The Sky Above

3.2: Ancient Astronomy

3.3: Key Terms

4: Modern Astronomy

4.1: The Birth of Modern Astronomy

4.2: The Laws of Planetary Motion

5: Physics and Gravity

5.1: Newton's Great Synthesis

5.2: Newton's Universal Law of Gravitation

5.3: Orbits in the Solar System

5.4: Motions of Satellites and Spacecraft

5.5: Gravity with More Than Two Bodies

5.6: Key Terms

6: Time

6.1: Earth and Sky

6.2: The Seasons

6.3: Keeping Time

6.4: The Calendar

7: The Moon

7.1: Phases and Motions of the Moon

8: The Moon 2

8.1: Ocean Tides and the Moon

8.2: Eclipses of the Sun and Moon

8.3: Key Terms

9: Light

9.1: The Behavior of Light

9.2: The Electromagnetic Spectrum

9.3: Spectroscopy in Astronomy

9.4: The Structure of the Atom

9.5: Formation of Spectral Lines

9.6: The Doppler Effect

9.7: Key Terms

10: Observing

10.1: Telescopes

10.2: Telescopes Today

10.3: Visible-Light Detectors and Instruments

10.4: Radio Telescopes

10.5: Observations outside Earth's Atmosphere

10.6: The Future of Large Telescopes

10.7: Key Terms

11: Solar System

11.1: Overview of Our Planetary System

11.2: Compostion and Structure of Planets

11.3: Dating Planetary Surfaces

11.4: Origin of the Solar System

11.5: Summary

12: Earth

12.1: The Global Perspective

12.2: Earth's Crust

12.3: Earth's Atmosphere

12.4: Life, Chemical Evolution, and Climate Change

12.5: Cosmic Influences on the Evolution of Earth

12.6: Key Terms

13: The Moon and Mercury

13.1: General Properties of the Moon

13.2: The Lunar Surface

13.3: Impact Craters

13.4: The Origin of the Moon

13.5: Mercury

13.6: Key Terms

14: Venus

14.1: The Nearest Planets - An Overview

14.2: The Geology of Venus

14.3: The Massive Atmosphere of Venus

15: Mars

15.1: The Geology of Mars

15.2: Water and Life on Mars

15.3: Divergent Planetary Evolution

15.4: Key Terms

16: Giant Planets

16.1: Exploring the Outer Planets

16.2: The Giant Planets

16.3: Atmosphere of the Giant Planets

16.4: Key Terms

17: Moons

17.1: Ring and Moon Systems Introduced

17.2: The Gailean Moons of Jupiter

17.3: Titan and Triton

18: Asteroids and Comets

18.1: Asteroids

18.2: Asteroids and Planetary Defense

18.3: The "Long-Haired" Comets

18.4: Key Terms

19: The Sun

19.1: The Structure and Composition of the Sun

19.2: The Solar Cycle

19.3: Solar Activity above the Photosphere

19.4: Mass, Energy, and the Theory of Relativity

19.5: The Solar Interior - Theory

19.6: The Solar Interior - Observations

19.7: Key Terms

20: The Stars

20.1: The Brightness of Stars

20.2: Colors of Stars

20.3: The Spectra of Stars (and Brown Dwarfs)

20.4: Measuring Stellar Masses

20.5: The H-R Diagram

21: Exoplanets

21.1: Planets beyond the Solar System- Search and Discovery

21.2: Exoplanets Everywhere - What we are Learning

Index

Glossary

Detailed Licensing

Detailed Licensing

This page titled Astronomy (Mastrapa) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

This page titled Licensing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

1: INTRODUCTION TO ASTRONOMY

1.1: Introduction
1.2: The Nature of Astronomy
1.3: The Nature of Science
1.4: The Laws of Nature
1.5: Numbers in Astronomy
1.6: Consequences of Light Travel Time

This page titled 1: Introduction to Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

1.1: Introduction



Figure 1.1 : Distant Galaxies. These two interacting islands of stars (galaxies) are so far away that their light takes hundreds of millions of years to reach us on Earth (photographed with the Hubble Space Telescope). (credit: modification of work by NASA, ESA, the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration, and K. Noll (STScI))

We invite you to come along on a series of voyages to explore the universe as astronomers understand it today. Beyond Earth are vast and magnificent realms full of objects that have no counterpart on our home planet. Nevertheless, we hope to show you that the evolution of the universe has been directly responsible for your presence on Earth today.

Along your journey, you will encounter:

- a canyon system so large that, on Earth, it would stretch from Los Angeles to Washington, DC (Figure 1.2).



Figure 1.2 : Mars Mosaic. This image of Mars is centered on the Valles Marineris (Mariner Valley) complex of canyons, which is as long as the United States is wide. (credit: modification of work by NASA)

- a crater and other evidence on Earth that tell us that the dinosaurs (and many other creatures) died because of a cosmic collision.

- a tiny moon whose gravity is so weak that one good throw from its surface could put a baseball into orbit.
- a collapsed star so dense that to duplicate its interior we would have to squeeze every human being on Earth into a single raindrop.
- exploding stars whose violent end could wipe clean all of the life-forms on a planet orbiting a neighboring star (Figure 1.3).
- a "cannibal galaxy" that has already consumed a number of its smaller galaxy neighbors and is not yet finished finding new victims.
- a radio echo that is the faint but unmistakable signal of the creation event for our universe.

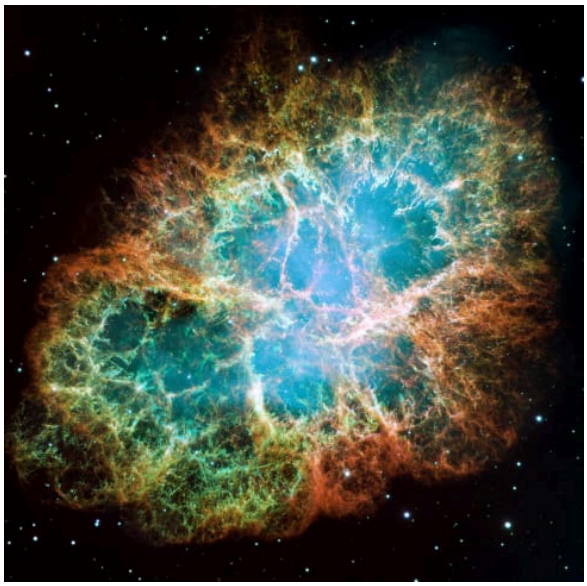


Figure 1.3 : Stellar Corpse. We observe the remains of a star that was seen to explode in our skies in 1054 (and was, briefly, bright enough to be visible during the daytime). Today, the remnant is called the Crab Nebula and its central region is seen here. Such exploding stars are crucial to the development of life in the universe. (credit: NASA, ESA, J. Hester (Arizona State University))

Such discoveries are what make astronomy such an exciting field for scientists and many others—but you will explore much more than just the objects in our universe and the latest discoveries about them. We will pay equal attention to the *process* by which we have come to understand the realms beyond Earth and the tools we use to increase that understanding.

We gather information about the cosmos from the messages the universe sends our way. Because the stars are the fundamental building blocks of the universe, decoding the message of starlight has been a central challenge and triumph of modern astronomy. By the time you have finished reading this text, you will know a bit about how to read that message and how to understand what it is telling us.

This page titled 1.1: Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

- 1.1: Introduction by OpenStax is licensed CC BY 4.0.

1.2: The Nature of Astronomy

Astronomy is defined as the study of the objects that lie beyond our planet Earth and the processes by which these objects interact with one another. We will see, though, that it is much more. It is also humanity's attempt to organize what we learn into a clear history of the universe, from the instant of its birth in the Big Bang to the present moment. Throughout this book, we emphasize that science is a *progress report*—one that changes constantly as new techniques and instruments allow us to probe the universe more deeply.

In considering the history of the universe, we will see again and again that the cosmos *evolves*; it changes in profound ways over long periods of time. For example, the universe made the carbon, the calcium, and the oxygen necessary to construct something as interesting and complicated as you. Today, many billions of years later, the universe has evolved into a more hospitable place for life. Tracing the evolutionary processes that continue to shape the universe is one of the most important (and satisfying) parts of modern astronomy.

This page titled 1.2: The Nature of Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.2: The Nature of Astronomy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

1.3: The Nature of Science

The ultimate judge in science is always what nature itself reveals based on observations, experiments, models, and testing. Science is not merely a body of knowledge, but a *method* by which we attempt to understand nature and how it behaves. This method begins with many observations over a period of time. From the trends found through observations, scientists can *model* the particular phenomena we want to understand. Such models are always approximations of nature, subject to further testing.

As a concrete astronomical example, ancient astronomers constructed a model (partly from observations and partly from philosophical beliefs) that Earth was the center of the universe and everything moved around it in circular orbits. At first, our available observations of the Sun, Moon, and planets did fit this model; however, after further observations, the model had to be updated by adding circle after circle to represent the movements of the planets around Earth at the center. As the centuries passed and improved instruments were developed for keeping track of objects in the sky, the old model (even with a huge number of circles) could no longer explain all the observed facts. As we will see in the chapter on Observing the Sky: The Birth of Astronomy, a new model, with the Sun at the center, fit the experimental evidence better. After a period of philosophical struggle, it became accepted as our view of the universe.

When they are first proposed, new models or ideas are sometimes called *hypotheses*. You may think there can be no new hypotheses in a science such as astronomy—that everything important has already been learned. Nothing could be further from the truth. Throughout this textbook you will find discussions of recent, and occasionally still controversial, hypotheses in astronomy. For example, the significance that the huge chunks of rock and ice that hit Earth have for life on Earth itself is still debated. And while the evidence is strong that vast quantities of invisible “dark energy” make up the bulk of the universe, scientists have no convincing explanation for what the dark energy actually is. Resolving these issues will require difficult observations done at the forefront of our technology, and all such hypotheses need further testing before we incorporate them fully into our standard astronomical models.

This last point is crucial: a hypothesis must be a proposed explanation that can be *tested*. The most straightforward approach to such testing in science is to perform an experiment. If the experiment is conducted properly, its results either will agree with the predictions of the hypothesis or they will contradict it. If the experimental result is truly inconsistent with the hypothesis, a scientist must discard the hypothesis and try to develop an alternative. If the experimental result agrees with predictions, this does not necessarily prove that the hypothesis is absolutely correct; perhaps later experiments will contradict crucial parts of the hypothesis. But, the more experiments that agree with the hypothesis, the more likely we are to accept the hypothesis as a useful description of nature.

One way to think about this is to consider a scientist who was born and lives on an island where only black sheep live. Day after day the scientist encounters black sheep only, so he or she hypothesizes that all sheep are black. Although every observed sheep adds confidence to the hypothesis, the scientist only has to visit the mainland and observe one white sheep to prove the hypothesis wrong.

When you read about experiments, you probably have a mental picture of a scientist in a laboratory conducting tests or taking careful measurements. This is certainly the case for a biologist or a chemist, but what can astronomers do when our laboratory is the universe? It’s impossible to put a group of stars into a test tube or to order another comet from a scientific supply company.

As a result, astronomy is sometimes called an *observational* science; we often make our tests by observing many samples of the kind of object we want to study and noting carefully how different samples vary. New instruments and technology can let us look at astronomical objects from new perspectives and in greater detail. Our hypotheses are then judged in the light of this new information, and they pass or fail in the same way we would evaluate the result of a laboratory experiment.

Much of astronomy is also a *historical* science—meaning that what we observe has already happened in the universe and we can do nothing to change it. In the same way, a geologist cannot alter what has happened to our planet, and a paleontologist cannot bring an ancient animal back to life. While this can make astronomy challenging, it also gives us fascinating opportunities to discover the secrets of our cosmic past.

You might compare an astronomer to a detective trying to solve a crime that occurred before the detective arrived at the scene. There is lots of evidence, but both the detective and the scientist must sift through and organize the evidence to test various hypotheses about what actually happened. And there is another way in which the scientist is like a detective: they both must prove their case. The detective must convince the district attorney, the judge, and perhaps ultimately the jury that his hypothesis

is correct. Similarly, the scientist must convince colleagues, editors of journals, and ultimately a broad cross-section of other scientists that her hypothesis is provisionally correct. In both cases, one can only ask for evidence “beyond a reasonable doubt.” And sometimes new evidence will force both the detective and the scientist to revise their last hypothesis.

This self-correcting aspect of science sets it off from most human activities. Scientists spend a great deal of time questioning and challenging one another, which is why applications for project funding—as well as reports for publication in academic journals—go through an extensive process of *peer review*, which is a careful examination by other scientists in the same field. In science (after formal education and training), everyone is encouraged to improve upon experiments and to challenge any and all hypotheses. New scientists know that one of the best ways to advance their careers is to find a weakness in our current understanding of something and to correct it with a new or modified hypothesis.

This is one of the reasons science has made such dramatic progress. An undergraduate science major today knows more about science and math than did Sir Isaac Newton, one of the most renowned scientists who ever lived. Even in this introductory astronomy course, you will learn about objects and processes that no one a few generations ago even dreamed existed.

This page titled 1.3: The Nature of Science is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.3: The Nature of Science by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

1.4: The Laws of Nature

Over centuries scientists have extracted various *scientific laws* from countless observations, hypotheses, and experiments. These scientific laws are, in a sense, the “rules” of the game that nature plays. One remarkable discovery about nature—one that underlies everything you will read about in this text—is that the same laws apply everywhere in the universe. The rules that determine the motion of stars so far away that your eye cannot see them are the same laws that determine the arc of a baseball after a batter has hit it out of the park.

Note that without the existence of such universal laws, we could not make much headway in astronomy. If each pocket of the universe had different rules, we would have little chance of interpreting what happened in other “neighborhoods.” But, the consistency of the laws of nature gives us enormous power to understand distant objects without traveling to them and learning the local laws. In the same way, if every region of a country had completely different laws, it would be very difficult to carry out commerce or even to understand the behavior of people in those different regions. A consistent set of laws, though, allows us to apply what we learn or practice in one state to any other state.

This is not to say that our current scientific models and laws cannot change. New experiments and observations can lead to new, more sophisticated models—models that can include new phenomena and laws about their behavior. The general theory of relativity proposed by Albert Einstein is a perfect example of such a transformation that took place about a century ago; it led us to predict, and eventually to observe, a strange new class of objects that astronomers call *black holes*. Only the patient process of observing nature ever more carefully and precisely can demonstrate the validity of such new scientific models.

One important problem in describing scientific models has to do with the limitations of language. When we try to describe complex phenomena in everyday terms, the words themselves may not be adequate to do the job. For example, you may have heard the structure of the atom likened to a miniature solar system. While some aspects of our modern model of the atom do remind us of planetary orbits, many other of its aspects are fundamentally different.

This problem is the reason scientists often prefer to describe their models using equations rather than words. In this book, which is designed to introduce the field of astronomy, we use mainly words to discuss what scientists have learned. We avoid complex math, but if this course piques your interest and you go on in science, more and more of your studies will involve the precise language of mathematics.

This page titled 1.4: The Laws of Nature is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.4: The Laws of Nature by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

1.5: Numbers in Astronomy

In astronomy we deal with distances on a scale you may never have thought about before, with numbers larger than any you may have encountered. We adopt two approaches that make dealing with astronomical numbers a little bit easier. First, we use a system for writing large and small numbers called *scientific notation* (or sometimes *powers-of-ten notation*). This system is very appealing because it eliminates the many zeros that can seem overwhelming to the reader. In scientific notation, if you want to write a number such as 500,000,000, you express it as 5×10^8. The second way we try to keep numbers simple is to use a consistent set of units—the metric International System of Units, or SI (from the French *Système International d'Unités*.)

Link to Learning

Watch this brief PBS animation that explains how scientific notation works and why it's useful.

A common unit astronomers use to describe distances in the universe is a light-year, which is the distance light travels during one year. Because light always travels at the same speed, and because its speed turns out to be the fastest possible speed in the universe, it makes a good standard for keeping track of distances. You might be confused because a "light-year" seems to imply that we are measuring time, but this mix-up of time and distance is common in everyday life as well. For example, when your friend asks where the movie theater is located, you might say "about 20 minutes from downtown."

So, how many kilometers are there in a light-year? Light travels at the amazing pace of $3\text{x}10^8$ m/s



Figure 1.4 : Orion Nebula. This beautiful cloud of cosmic raw material (gas and dust from which new stars and planets are being made) called the Orion Nebula is about 1400 light-years away. That's a distance of roughly 1.34×10^{16} kilometers—a pretty big number. The gas and dust in this region are illuminated by the intense light from a few extremely energetic adolescent stars. (credit: NASA, ESA, M. Robberto (Space Telescope Science Institute/ESA) and the Hubble Space Telescope Orion Treasury Project Team)

Example 1.1

Scientific Notation

In 2015, the richest human being on our planet had a net worth of $79.2 billion. Some might say this is an astronomical sum of money. Express this amount in scientific notation.

Solution

$79.2 billion can be written $79,200,000,000. Expressed in scientific notation it becomes $\$7.92 \times 10^{10}$

Example 1.2

Getting Familiar with a Light-Year

How many kilometers are there in a light-year?

Solution

Light travels 3×10^5 km in 1 s. So, let's calculate how far it goes in a year:

- There are 60 $\left(6 \times 10^1\right)$ s in 1 min, and 6×10^1 min in 1 h.
- Multiply these together and you find that there are 3.6×10^3 s/h.
- Thus, light covers $3 \times 10^5 \ \mathrm{km/s} \times 3.6 \times 10^3 \ \mathrm{s/h} = 1.08 \times 10^9 \ \mathrm{km/h}.$
- There are 24 or 2.4×10^1 h in a day, and 365.25 $\left(3.65 \times 10^2\right)$ days in 1 y.
- The product of these two numbers is 8.77×10^3 h/y.
- Multiplying this by 1.08×10^9 km/h gives 9.46×10^{12} km/light-year.

That's almost 10,000,000,000,000 km that light covers in a year. To help you imagine how long this distance is, we'll mention that a string 1 light-year long could fit around the circumference of Earth 236 million times.

This page titled 1.5: Numbers in Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

1.6: Consequences of Light Travel Time

There is another reason the speed of light is such a natural unit of distance for astronomers. Information about the universe comes to us almost exclusively through various forms of light, and all such light travels at the speed of light—that is, 1 light-year every year. This sets a limit on how quickly we can learn about events in the universe. If a star is 100 light-years away, the light we see from it tonight left that star 100 years ago and is just now arriving in our neighborhood. The soonest we can learn about any changes in that star is 100 years after the fact. For a star 500 light-years away, the light we detect tonight left 500 years ago and is carrying 500-year-old news.

Because many of us are accustomed to instant news from the Internet, some might find this frustrating.

"You mean, when I see that star up there," you ask, "I won't know what's actually happening there for another 500 years?"

But this isn't the most helpful way to think about the situation. For astronomers, *now* is when the light reaches us here on Earth. There is no way for us to know anything about that star (or other object) until its light reaches us.

But what at first may seem a great frustration is actually a tremendous benefit in disguise. If astronomers really want to piece together what has happened in the universe since its beginning, they must find evidence about each epoch (or period of time) of the past. Where can we find evidence today about cosmic events that occurred billions of years ago?

The delay in the arrival of light provides an answer to this question. The farther out in space we look, the longer the light has taken to get here, and the longer ago it left its place of origin. By looking billions of light-years out into space, astronomers are actually seeing billions of years into the past. In this way, we can reconstruct the history of the cosmos and get a sense of how it has evolved over time.

This is one reason why astronomers strive to build telescopes that can collect more and more of the faint light in the universe. The more light we collect, the fainter the objects we can observe. On average, fainter objects are farther away and can, therefore, tell us about periods of time even deeper in the past. Instruments such as the Hubble Space Telescope (Figure 1.5) and the Very Large Telescope in Chile (which you will learn about in the chapter on Astronomical Instruments), are giving astronomers views of deep space and deep time better than any we have had before.



Figure 1.5 : Telescope in Orbit. The Hubble Space Telescope, shown here in orbit around Earth, is one of many astronomical instruments in space. (credit: modification of work by European Space Agency)

This page titled 1.6: Consequences of Light Travel Time is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.6: Consequences of Light Travel Time by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

2: THE UNIVERSE

2.1: A Tour of the Universe
2.2: The Universe on the Large Scale
2.3: The Universe of the Very Small
2.4: A Conclusion and a Beginning

This page titled 2: The Universe is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

2.1: A Tour of the Universe

We can now take a brief introductory tour of the universe as astronomers understand it today to get acquainted with the types of objects and distances you will encounter throughout the text. We begin at home with Earth, a nearly spherical planet about 13,000 kilometers in diameter (Figure 1.6). A space traveler entering our planetary system would easily distinguish Earth from the other planets in our solar system by the large amount of liquid water that covers some two thirds of its crust. If the traveler had equipment to receive radio or television signals, or came close enough to see the lights of our cities at night, she would soon find signs that this watery planet has sentient life.

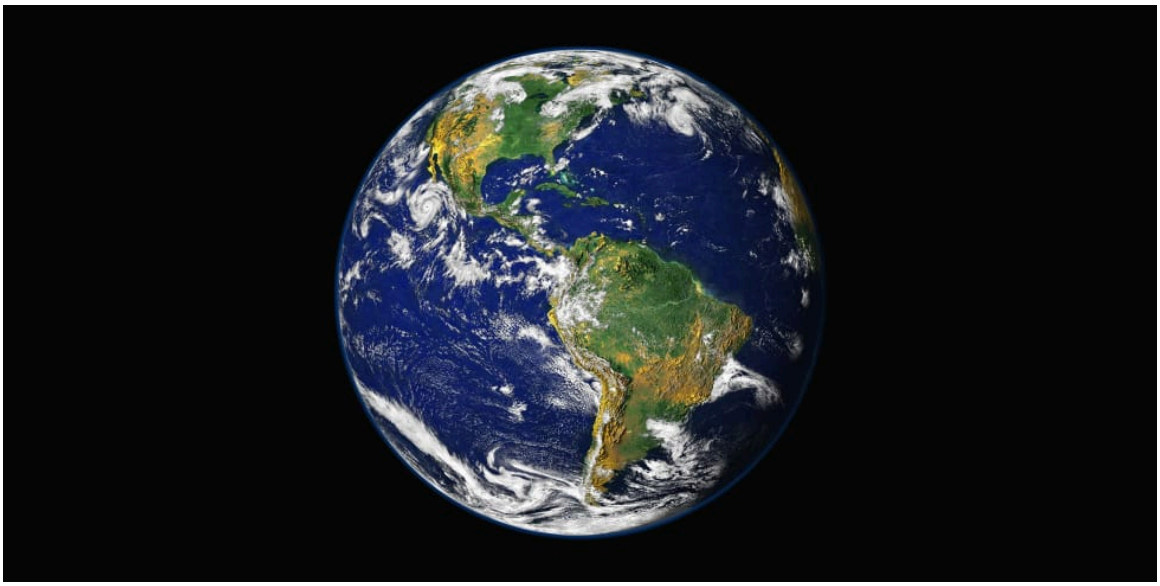


Figure 1.6 : Humanity's Home Base. This image shows the Western hemisphere as viewed from space 35,400 kilometers (about 22,000 miles) above Earth. Data about the land surface from one satellite was combined with another satellite's data about the clouds to create the image. (credit: modification of work by R. Stockli, A. Nelson, F. Hasler, NASA/ GSFC/ NOAA/ USGS)

Our nearest astronomical neighbor is Earth's satellite, commonly called the *Moon*. Figure 1.7 shows Earth and the Moon drawn to scale on the same diagram. Notice how small we have to make these bodies to fit them on the page with the right scale. The Moon's distance from Earth is about 30 times Earth's diameter, or approximately 384,000 kilometers, and it takes about a month for the Moon to revolve around Earth. The Moon's diameter is 3476 kilometers, about one fourth the size of Earth.



Figure 1.7 : Earth and Moon, Drawn to Scale. This image shows Earth and the Moon shown to scale for both size and distance. (credit: modification of work by NASA)

Light (or radio waves) takes 1.3 seconds to travel between Earth and the Moon. If you've seen videos of the Apollo flights to the Moon, you may recall that there was a delay of about 3 seconds between the time Mission Control asked a question and the time the astronauts responded. This was not because the astronauts were thinking slowly, but rather because it took the radio waves almost 3 seconds to make the round trip.

Earth revolves around our star, the Sun, which is about 150 million kilometers away—approximately 400 times as far away from us as the Moon. We call the average Earth–Sun distance an *astronomical unit* (AU) because, in the early days of astronomy, it was the most important measuring standard. Light takes slightly more than 8 minutes to travel 1 astronomical unit, which means the latest news we receive from the Sun is always 8 minutes old. The diameter of the Sun is about 1.5 million kilometers; Earth could fit comfortably inside one of the minor eruptions that occurs on the surface of our star. If the Sun were reduced to the size of a basketball, Earth would be a small apple seed about 30 meters from the ball.

It takes Earth 1 year (3×10^7 seconds) to go around the Sun at our distance; to make it around, we must travel at approximately 110,000 kilometers per hour. (If you, like many students, still prefer miles to kilometers, you might find the following trick helpful. To convert kilometers to miles, just multiply kilometers by 0.6. Thus, 110,000 kilometers per hour becomes 66,000 miles per hour.) Because gravity holds us firmly to Earth and there is no resistance to Earth's motion in the vacuum of space, we participate in this extremely fast-moving trip without being aware of it day to day.

Earth is only one of eight planets that revolve around the Sun. These planets, along with their moons and swarms of smaller bodies such as dwarf planets, make up the solar system (Figure 1.8). A planet is defined as a body of significant size that orbits a star and does not produce its own light. (If a large body consistently produces its own light, it is then called a *star*.) Later in the book this definition will be modified a bit, but it is perfectly fine for now as you begin your voyage.

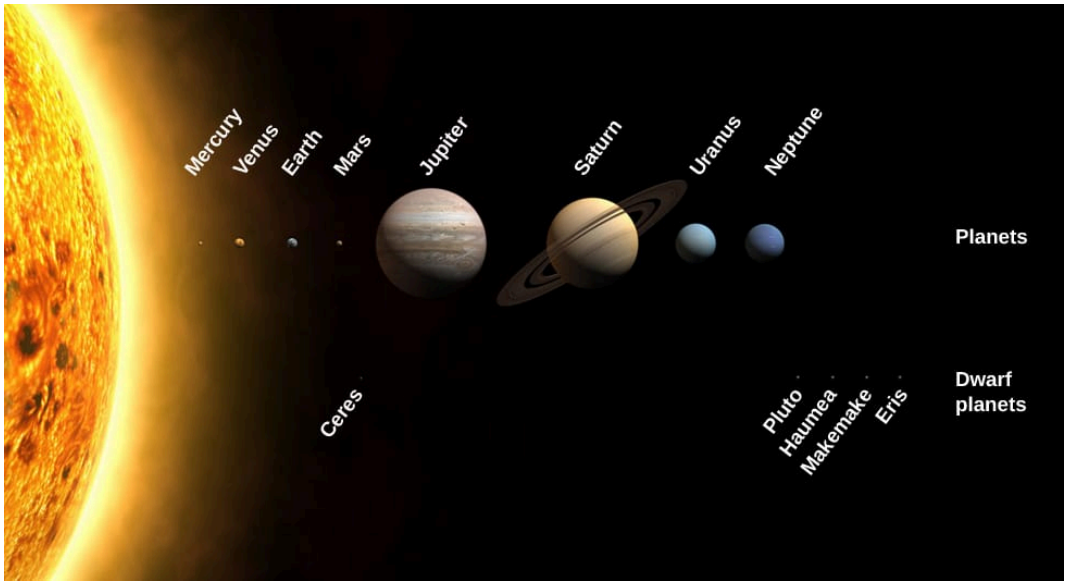


Figure 1.8 : Our Solar Family. The Sun, the planets, and some dwarf planets are shown with their sizes drawn to scale. The orbits of the planets are much more widely separated than shown in this drawing. Notice the size of Earth compared to the giant planets. (credit: modification of work by NASA)

We are able to see the nearby planets in our skies only because they reflect the light of our local star, the Sun. If the planets were much farther away, the tiny amount of light they reflect would usually not be visible to us. The planets we have so far discovered orbiting other stars were found from the pull their gravity exerts on their parent stars, or from the light they block from their stars when they pass in front of them. We can't see most of these planets directly, although a few are now being imaged directly.

The Sun is our local star, and all the other stars are also enormous balls of glowing gas that generate vast amounts of energy by nuclear reactions deep within. We will discuss the processes that cause stars to shine in more detail later in the book. The other stars look faint only because they are so very far away. If we continue our basketball analogy, Proxima Centauri, the nearest star beyond the Sun, which is 4.3 light-years away, would be almost 7000 kilometers from the basketball.

When you look up at a star-filled sky on a clear night, all the stars visible to the unaided eye are part of a single collection of stars we call the *Milky Way Galaxy*, or simply the *Galaxy*. (When referring to the Milky Way, we capitalize *Galaxy*; when talking about other galaxies of stars, we use lowercase *galaxy*.) The Sun is one of hundreds of billions of stars that make up the Galaxy; its extent, as we will see, staggers the human imagination. Within a sphere 10 light-years in radius centered on the Sun, we find roughly ten stars. Within a sphere 100 light-years in radius, there are roughly 10,000 (10^4) stars—far too many to count or name—but we have still traversed only a tiny part of the Milky Way Galaxy. Within a 1000-light-year sphere, we find some ten million (10^7) stars; within a sphere of 100,000 light-years, we finally encompass the entire Milky Way Galaxy.

Our Galaxy looks like a giant disk with a small ball in the middle. If we could move outside our Galaxy and look down on the disk of the Milky Way from above, it would probably resemble the galaxy in Figure 1.9, with its spiral structure outlined by the blue light of hot adolescent stars.



Figure 1.9 : Spiral Galaxy. This galaxy of billions of stars, called by its catalog number NGC 1073, is thought to be similar to our own Milky Way Galaxy. Here we see the giant wheel-shaped system with a bar of stars across its middle. (credit: NASA, ESA)

The Sun is somewhat less than 30,000 light-years from the center of the Galaxy, in a location with nothing much to distinguish it. From our position inside the Milky Way Galaxy, we cannot see through to its far rim (at least not with ordinary light) because the space between the stars is not completely empty. It contains a sparse distribution of gas (mostly the simplest element, hydrogen) intermixed with tiny solid particles that we call *interstellar dust*. This gas and dust collect into enormous clouds in many places in the Galaxy, becoming the raw material for future generations of stars. Figure 1.10 shows an image of the disk of the Galaxy as seen from our vantage point.



Figure 1.10 : Milky Way Galaxy. Because we are inside the Milky Way Galaxy, we see its disk in cross-section flung across the sky like a great milky white avenue of stars with dark “rifts” of dust. In this dramatic image, part of it is seen above Trona Pinnacles in the California desert. (credit: Ian Norman)

Typically, the interstellar material is so extremely sparse that the space between stars is a much better vacuum than anything we can produce in terrestrial laboratories. Yet, the dust in space, building up over thousands of light-years, can block the light of more distant stars. Like the distant buildings that disappear from our view on a smoggy day in Los Angeles, the more distant regions of the Milky Way cannot be seen behind the layers of interstellar smog. Luckily, astronomers have found that stars and raw material shine with various forms of light, some of which do penetrate the smog, and so we have been able to develop a pretty good map of the Galaxy.

Recent observations, however, have also revealed a rather surprising and disturbing fact. There appears to be more—much more—to the Galaxy than meets the eye (or the telescope). From various investigations, we have evidence that much of our Galaxy is made of material we cannot currently observe directly with our instruments. We therefore call this component of the Galaxy *dark matter*. We know the dark matter is there by the pull its gravity exerts on the stars and raw material we can observe, but what this dark matter is made of and how much of it exists remain a mystery. Furthermore, this dark matter is not confined to our Galaxy; it appears to be an important part of other star groupings as well.

By the way, not all stars live by themselves, as the Sun does. Many are born in double or triple systems with two, three, or more stars revolving about each other. Because the stars influence each other in such close systems, multiple stars allow us to measure characteristics that we cannot discern from observing single stars. In a number of places, enough stars have formed together that we recognized them as star clusters (Figure 1.11). Some of the largest of the star clusters that astronomers have cataloged contain hundreds of thousands of stars and take up volumes of space hundreds of light-years across.



Figure 1.11 : Star Cluster. This large star cluster is known by its catalog number, M9. It contains some 250,000 stars and is seen more clearly from space using the Hubble Space Telescope. It is located roughly 25,000 light-years away. (credit: NASA, ESA)

You may hear stars referred to as “eternal,” but in fact no star can last forever. Since the “business” of stars is making energy, and energy production requires some sort of fuel to be used up, eventually all stars run out of fuel. This news should not cause you to panic, though, because our Sun still has at least 5 or 6 billion years to go. Ultimately, the Sun and all stars will die, and it is in their death throes that some of the most intriguing and important processes of the universe are revealed. For example, we now know that many of the atoms in our bodies were once inside stars. These stars exploded at the ends of their lives, recycling their material back into the reservoir of the Galaxy. In this sense, all of us are literally made of recycled “star dust.”

This page titled 2.1: A Tour of the Universe is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.7: A Tour of the Universe by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

2.2: The Universe on the Large Scale

In a very rough sense, you could think of the solar system as your house or apartment and the Galaxy as your town, made up of many houses and buildings. In the twentieth century, astronomers were able to show that, just as our world is made up of many, many towns, so the universe is made up of enormous numbers of galaxies. (We define the universe to be everything that exists that is accessible to our observations.) Galaxies stretch as far into space as our telescopes can see, many billions of them within the reach of modern instruments. When they were first discovered, some astronomers called galaxies *island universes*, and the term is aptly descriptive; galaxies do look like islands of stars in the vast, dark seas of intergalactic space.

The nearest galaxy, discovered in 1993, is a small one that lies 70,000 light-years from the Sun in the direction of the constellation Sagittarius, where the smog in our own Galaxy makes it especially difficult to discern. (A constellation, we should note, is one of the 88 sections into which astronomers divide the sky, each named after a prominent star pattern within it.) Beyond this Sagittarius dwarf galaxy lie two other small galaxies, about 160,000 light-years away. First recorded by Magellan's crew as he sailed around the world, these are called the *Magellanic Clouds* (Figure 1.12). All three of these small galaxies are satellites of the Milky Way Galaxy, interacting with it through the force of gravity. Ultimately, all three may even be swallowed by our much larger Galaxy, as other small galaxies have been over the course of cosmic time.



Figure 1.12 : Neighbor Galaxies. This image shows both the Large Magellanic Cloud and the Small Magellanic Cloud above the telescopes of the Atacama Large Millimeter/Submillimeter Array (ALMA) in the Atacama Desert of northern Chile. (credit: ESO, C. Malin)

The nearest large galaxy is a spiral quite similar to our own, located in the constellation of Andromeda, and is thus called the Andromeda galaxy; it is also known by one of its catalog numbers, M31 (Figure 1.13). M31 is a little more than 2 million light-years away and, along with the Milky Way, is part of a small cluster of more than 50 galaxies referred to as the *Local Group*.

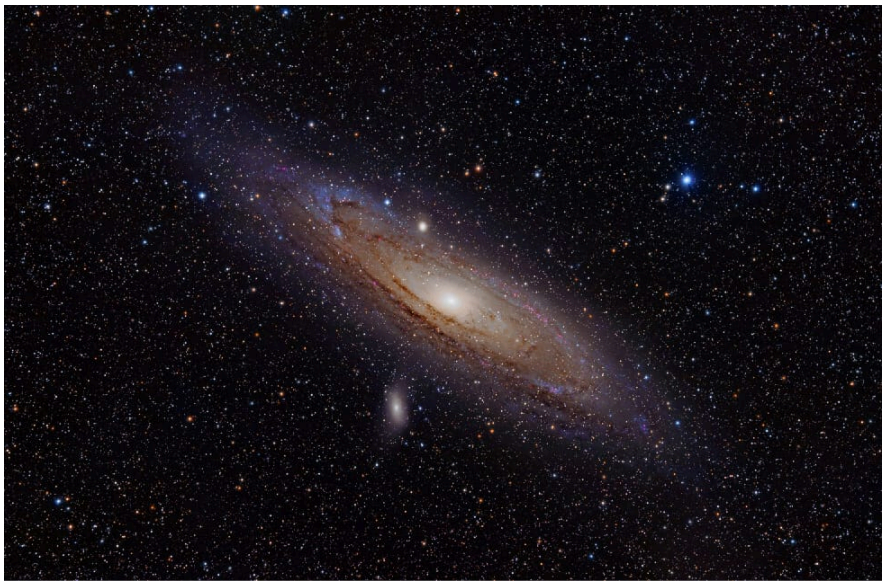


Figure 1.13 : Closest Spiral Galaxy. The Andromeda galaxy (M31) is a spiral-shaped collection of stars similar to our own Milky Way. (credit: Adam Evans)

At distances of 10 to 15 million light-years, we find other small galaxy groups, and then at about 50 million light-years there are more impressive systems with thousands of member galaxies. We have discovered that galaxies occur mostly in clusters, both large and small (Figure 1.14).



Figure 1.14 : Fornax Cluster of Galaxies. In this image, you can see part of a cluster of galaxies located about 60 million light-years away in the constellation of Fornax. All the objects that are not pinpoints of light in the picture are galaxies of billions of stars. (credit: ESO, J. Emerson, VISTA. Acknowledgment: Cambridge Astronomical Survey Unit)

Some of the clusters themselves form into larger groups called *superclusters*. The Local Group is part of a supercluster of galaxies, called the Virgo Supercluster, which stretches over a diameter of 110 million light-years. We are just beginning to explore the structure of the universe at these enormous scales and are already encountering some unexpected findings.

At even greater distances, where many ordinary galaxies are too dim to see, we find *quasars*. These are brilliant centers of galaxies, glowing with the light of an extraordinarily energetic process. The enormous energy of the quasars is produced by gas that is heated to a temperature of millions of degrees as it falls toward a massive black hole and swirls around it. The brilliance of quasars makes them the most distant beacons we can see in the dark oceans of space. They allow us to probe the universe 10 billion light-years away or more, and thus 10 billion years or more in the past.

With quasars we can see way back close to the Big Bang explosion that marks the beginning of time. Beyond the quasars and the most distant visible galaxies, we have detected the feeble glow of the explosion itself, filling the universe and thus coming to us from all directions in space. The discovery of this “afterglow of creation” is considered to be one of the most significant events in twentieth-century science, and we are still exploring the many things it has to tell us about the earliest times of the universe.

Measurements of the properties of galaxies and quasars in remote locations require large telescopes, sophisticated light-amplifying devices, and painstaking labor. Every clear night, at observatories around the world, astronomers and students are at work on such mysteries as the birth of new stars and the large-scale structure of the universe, fitting their results into the tapestry of our understanding.

This page titled 2.2: The Universe on the Large Scale is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.8: The Universe on the Large Scale by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

2.3: The Universe of the Very Small

The foregoing discussion has likely impressed on you that the universe is extraordinarily large and extraordinarily empty. On average, it is 10,000 times more empty than our Galaxy. Yet, as we have seen, even the Galaxy is mostly empty space. The air we breathe has about 10^{19} atoms in each cubic centimeter—and we usually think of air as empty space. In the interstellar gas of the Galaxy, there is about one atom in every cubic centimeter. Intergalactic space is filled so sparsely that to find one atom, on average, we must search through a cubic meter of space. Most of the universe is fantastically empty; places that are dense, such as the human body, are tremendously rare.

Even our most familiar solids are mostly space. If we could take apart such a solid, piece by piece, we would eventually reach the tiny molecules from which it is formed. Molecules are the smallest particles into which any matter can be divided while still retaining its chemical properties. A molecule of water (H_2O), for example, consists of two hydrogen atoms and one oxygen atom bonded together.

Molecules, in turn, are built of atoms, which are the smallest particles of an element that can still be identified as that element. For example, an atom of gold is the smallest possible piece of gold. Nearly 100 different kinds of atoms (elements) exist in nature. Most of them are rare, and only a handful account for more than 99% of everything with which we come in contact. The most abundant elements in the cosmos today are listed in Table 1.1; think of this table as the “greatest hits” of the universe when it comes to elements. Note that the list includes the four elements most common in life on Earth—hydrogen, carbon, nitrogen, and oxygen.

Table 1.1: The Cosmically Abundant Elements

Element[1]	Symbol	Number of Atoms per Million Hydrogen Atoms
Hydrogen	H	1,000,000
Helium	He	80,000
Carbon	C	450
Nitrogen	N	92
Oxygen	O	740
Neon	Ne	130
Magnesium	Mg	40
Silicon	Si	37
Sulfur	S	19
Iron	Fe	32

All atoms consist of a central, positively charged nucleus surrounded by negatively charged electrons. The bulk of the matter in each atom is found in the nucleus, which consists of positive protons and electrically neutral neutrons all bound tightly together in a very small space. Each element is defined by the number of protons in its atoms. Thus, any atom with 6 protons in its nucleus is called *carbon*, any with 50 protons is called *tin*, and any with 70 protons is called *ytterbium*. (For a list of the elements, see Appendix K.)

The distance from an atomic nucleus to its electrons is typically 100,000 times the size of the nucleus itself. This is why we say that even solid matter is mostly space. The typical atom is far emptier than the solar system out to Neptune. (The distance from Earth to the Sun, for example, is only 100 times the size of the Sun.) This is one reason atoms are not like miniature solar systems.

Remarkably, physicists have discovered that everything that happens in the universe, from the smallest atomic nucleus to the largest superclusters of galaxies, can be explained through the action of only four forces: gravity, electromagnetism (which combines the actions of electricity and magnetism), and two forces that act at the nuclear level. The fact that there are four forces (and not a million, or just one) has puzzled physicists and astronomers for many years and has led to a quest for a unified picture of nature.

Link to Learning

To construct an atom, particle by particle, check out this guided animation for building an atom.

Footnotes

- 1This list of elements is arranged in order of the atomic number, which is the number of protons in each nucleus.

This page titled 2.3: The Universe of the Very Small is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.9: The Universe of the Very Small by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

2.4: A Conclusion and a Beginning

If you are new to astronomy, you have probably reached the end of our brief tour in this chapter with mixed emotions. On the one hand, you may be fascinated by some of the new ideas you've read about and you may be eager to learn more. On the other hand, you may be feeling a bit overwhelmed by the number of topics we have covered, and the number of new words and ideas we have introduced. Learning astronomy is a little like learning a new language: at first it seems there are so many new expressions that you'll never master them all, but with practice, you soon develop facility with them.

At this point you may also feel a bit small and insignificant, dwarfed by the cosmic scales of distance and time. But, there is another way to look at what you have learned from our first glimpses of the cosmos. Let us consider the history of the universe from the Big Bang to today and compress it, for easy reference, into a single year. (We have borrowed this idea from Carl Sagan's 1977 Pulitzer Prize-winning book, *The Dragons of Eden*.)

On this scale, the Big Bang happened at the first moment of January 1, and this moment, when you are reading this chapter would be the end of the very last second of December 31. When did other events in the development of the universe happen in this "cosmic year?" Our solar system formed around September 10, and the oldest rocks we can date on Earth go back to the third week in September (Figure 1.15).

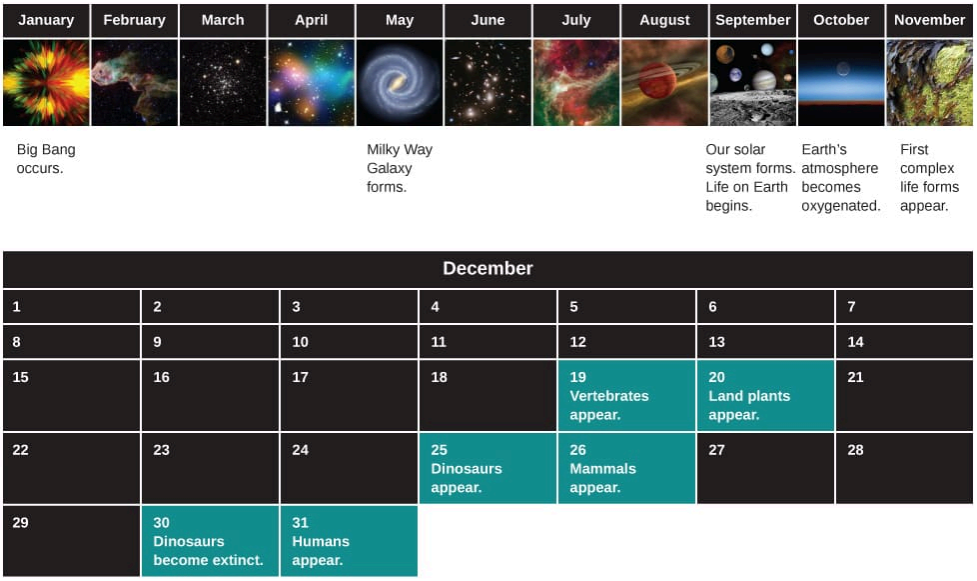


Figure 1.15 : Charting Cosmic Time. On a cosmic calendar, where the time since the Big Bang is compressed into 1 year, creatures we would call human do not emerge on the scene until the evening of December 31. (credit: February: modification of work by NASA, JPL-Caltech, W. Reach (SSC/Caltech); March: modification of work by ESA, Hubble and NASA, Acknowledgement: Giles Chapdelaine; April: modification of work by NASA, ESA, CFHT, CXO, M.J. Jee (University of California, Davis), A. Mahdavi (San Francisco State University); May: modification of work by NASA, JPL-Caltech; June: modification of work by NASA/ESA; July: modification of work by NASA, JPL-Caltech, Harvard-Smithsonian; August: modification of work by NASA, JPL-Caltech, R. Hurt (SSC-Caltech); September: modification of work by NASA; October: modification of work by NASA; November: modification of work by Dénes Emőke)

Where does the origin of human beings fall during the course of this cosmic year? The answer turns out to be the evening of December 31. The invention of the alphabet doesn't occur until the fiftieth second of 11:59 p.m. on December 31. And the beginnings of modern astronomy are a mere fraction of a second before the New Year. Seen in a cosmic context, the amount of time we have had to study the stars is minute, and our success in piecing together as much of the story as we have is remarkable.

Certainly our attempts to understand the universe are not complete. As new technologies and new ideas allow us to gather more and better data about the cosmos, our present picture of astronomy will very likely undergo many changes. Still, as you read our current progress report on the exploration of the universe, take a few minutes every once in a while just to savor how much you have already learned.

This page titled 2.4: A Conclusion and a Beginning is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 1.10: A Conclusion and a Beginning by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

3: ANCIENT ASTRONOMY

3.1: The Sky Above
3.2: Ancient Astronomy
3.3: Key Terms

This page titled 3: Ancient Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

3.1: The Sky Above

Learning Objectives

By the end of this section, you will be able to:

- Define the main features of the celestial sphere
- Explain the system astronomers use to describe the sky
- Describe how motions of the stars appear to us on Earth
- Describe how motions of the Sun, Moon, and planets appear to us on Earth
- Understand the modern meaning of the term *constellation*

Our senses suggest to us that Earth is the center of the universe—the hub around which the heavens turn. This geocentric (Earth-centered) view was what almost everyone believed until the European Renaissance. After all, it is simple, logical, and seemingly self-evident. Furthermore, the geocentric perspective reinforced those philosophical and religious systems that taught the unique role of human beings as the central focus of the cosmos. However, the geocentric view happens to be wrong. One of the great themes of our intellectual history is the overthrow of the geocentric perspective. Let us, therefore, take a look at the steps by which we reevaluated the place of our world in the cosmic order.

The Celestial Sphere

If you go on a camping trip or live far from city lights, your view of the sky on a clear night is pretty much identical to that seen by people all over the world before the invention of the telescope. Gazing up, you get the impression that the sky is a great hollow dome with you at the center (Figure 2.2), and all the stars are an equal distance from you on the surface of the dome. The top of that dome, the point directly above your head, is called the zenith, and where the dome meets Earth is called the horizon. From the sea or a flat prairie, it is easy to see the horizon as a circle around you, but from most places where people live today, the horizon is at least partially hidden by mountains, trees, buildings, or smog.

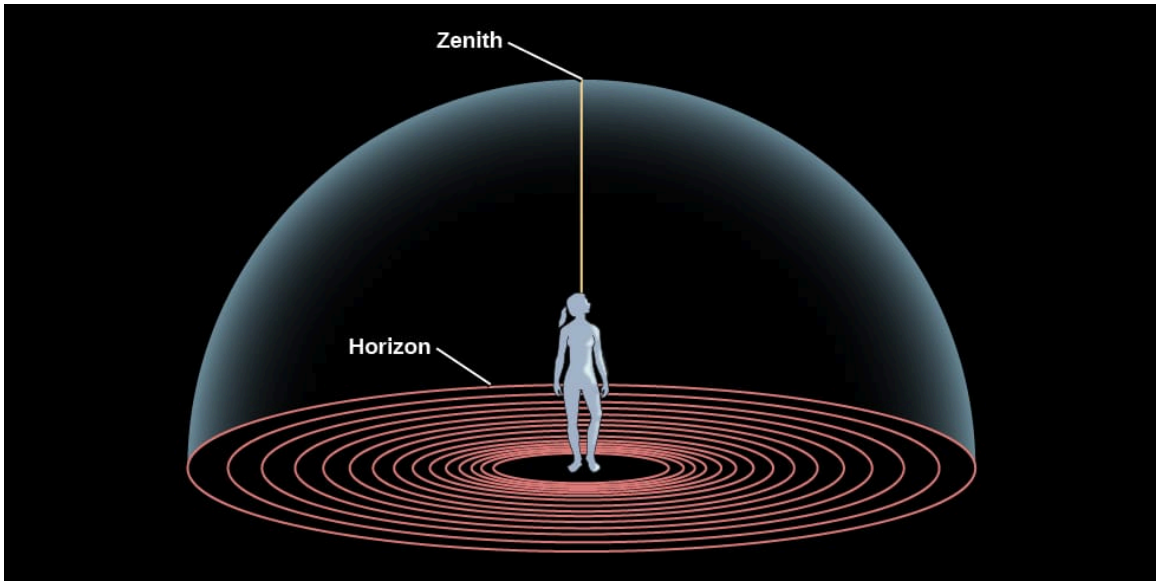


Figure 2.2 : The Sky around Us. The horizon is where the sky meets the ground; an observer's zenith is the point directly overhead.

If you lie back in an open field and observe the night sky for hours, as ancient shepherds and travelers regularly did, you will see stars rising on the eastern horizon (just as the Sun and Moon do), moving across the dome of the sky in the course of the night, and setting on the western horizon. Watching the sky turn like this night after night, you might eventually get the idea that the dome of the sky is really part of a great sphere that is turning around you, bringing different stars into view as it turns. The early Greeks regarded the sky as just such a celestial sphere (Figure 2.3). Some thought of it as an actual sphere of transparent

crystalline material, with the stars embedded in it like tiny jewels.

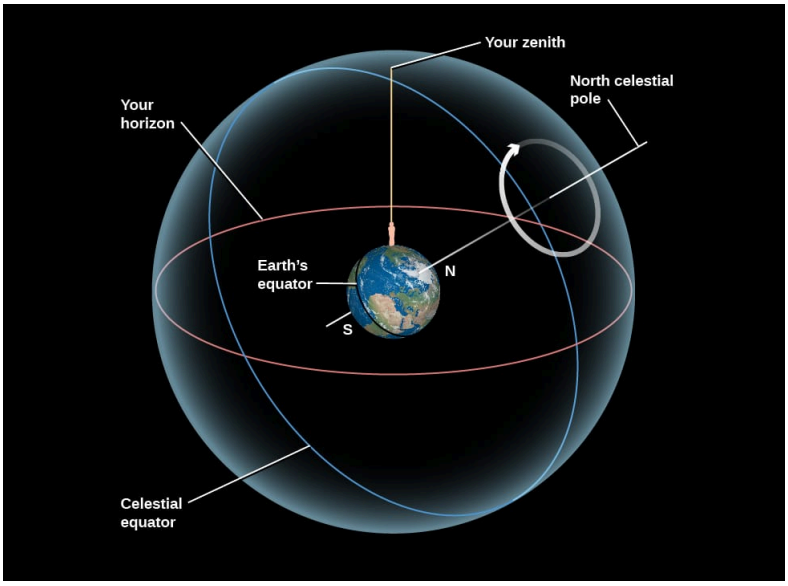


Figure 2.3 : Circles on the Celestial Sphere. Here we show the (imaginary) celestial sphere around Earth, on which objects are fixed, and which rotates around Earth on an axis. In reality, it is Earth that turns around this axis, creating the illusion that the sky revolves around us. Note that Earth in this picture has been tilted so that your location is at the top and the North Pole is where the N is. The apparent motion of celestial objects in the sky around the pole is shown by the circular arrow.

Today, we know that it is not the celestial sphere that turns as night and day proceed, but rather the planet on which we live. We can put an imaginary stick through Earth's North and South Poles, representing our planet's axis. It is because Earth turns on this axis every 24 hours that we see the Sun, Moon, and stars rise and set with clockwork regularity. Today, we know that these celestial objects are not really on a dome, but at greatly varying distances from us in space. Nevertheless, it is sometimes still convenient to talk about the celestial dome or sphere to help us keep track of objects in the sky. There is even a special theater, called a *planetarium*, in which we project a simulation of the stars and planets onto a white dome.

As the celestial sphere rotates, the objects on it maintain their positions with respect to one another. A grouping of stars such as the Big Dipper has the same shape during the course of the night, although it turns with the sky. During a single night, even objects we know to have significant motions of their own, such as the nearby planets, seem fixed relative to the stars. Only meteors—brief "shooting stars" that flash into view for just a few seconds—move appreciably with respect to other objects on the celestial sphere. (This is because they are not stars at all. Rather, they are small pieces of cosmic dust, burning up as they hit Earth's atmosphere.) We can use the fact that the entire celestial sphere seems to turn together to help us set up systems for keeping track of what things are visible in the sky and where they happen to be at a given time.

Celestial Poles and Celestial Equator

To help orient us in the turning sky, astronomers use a system that extends Earth's axis points into the sky. Imagine a line going through Earth, connecting the North and South Poles. This is Earth's axis, and Earth rotates about this line. If we extend this imaginary line outward from Earth, the points where this line intersects the celestial sphere are called the *north celestial pole* and the *south celestial pole*. As Earth rotates about its axis, the sky appears to turn in the opposite direction around those celestial poles (Figure 2.4). We also (in our imagination) throw Earth's equator onto the sky and call this the celestial equator. It lies halfway between the celestial poles, just as Earth's equator lies halfway between our planet's poles.



Figure 2.4 : Circling the South Celestial Pole. This long-exposure photo shows trails left by stars as a result of the apparent rotation of the celestial sphere around the south celestial pole. (In reality, it is Earth that rotates.) (Credit: ESO/Iztok Bončina)

Now let's imagine how riding on different parts of our spinning Earth affects our view of the sky. The apparent motion of the celestial sphere depends on your latitude (position north or south of the equator). First of all, notice that Earth's axis is pointing at the celestial poles, so these two points in the sky do not appear to turn.

If you stood at the North Pole of Earth, for example, you would see the north celestial pole overhead, at your zenith. The celestial equator, 90° from the celestial poles, would lie along your horizon. As you watched the stars during the course of the night, they would all circle around the celestial pole, with none rising or setting. Only that half of the sky north of the celestial equator is ever visible to an observer at the North Pole. Similarly, an observer at the South Pole would see only the southern half of the sky.

If you were at Earth's equator, on the other hand, you see the celestial equator (which, after all, is just an "extension" of Earth's equator) pass overhead through your zenith. The celestial poles, being 90° from the celestial equator, must then be at the north and south points on your horizon. As the sky turns, all stars rise and set; they move straight up from the east side of the horizon and set straight down on the west side. During a 24-hour period, all stars are above the horizon exactly half the time. (Of course, during some of those hours, the Sun is too bright for us to see them.)

What would an observer in the latitudes of the United States or Europe see? Remember, we are neither at Earth's pole nor at the equator, but in between them. For those in the continental United States and Europe, the north celestial pole is neither overhead nor on the horizon, but in between. It appears above the northern horizon at an angular height, or altitude, equal to the observer's latitude. In San Francisco, for example, where the latitude is 38° N, the north celestial pole is 38° above the northern horizon.

For an observer at 38° N latitude, the south celestial pole is 38° below the southern horizon and, thus, never visible. As Earth turns, the whole sky seems to pivot about the north celestial pole. For this observer, stars within 38° of the North Pole can never set. They are always above the horizon, day and night. This part of the sky is called the north circumpolar zone. For observers in the continental United States, the Big Dipper, Little Dipper, and Cassiopeia are examples of star groups in the north circumpolar zone. On the other hand, stars within 38° of the south celestial pole never rise. That part of the sky is the south circumpolar zone. To most U.S. observers, the Southern Cross is in that zone. (Don't worry if you are not familiar with the star groups just mentioned; we will introduce them more formally later on.)

At this particular time in Earth's history, there happens to be a star very close to the north celestial pole. It is called Polaris, the pole star, and has the distinction of being the star that moves the least amount as the northern sky turns each day. Because it moved so little while the other stars moved much more, it played a special role in the mythology of several Native American tribes, for example (some called it the "fastener of the sky").

Astronomy Basics: What's Your Angle?

Astronomers measure how far apart objects appear in the sky by using angles. By definition, there are 360° in a circle, so a circle stretching completely around the celestial sphere contains 360°. The half-sphere or dome of the sky then contains 180° from horizon to opposite horizon. Thus, if two stars are 18° apart, their separation spans about 1/10 of the dome of the sky. To give you a sense of how big a degree is, the full Moon is about half a degree across. This is about the width of your smallest finger (pinkie) seen at arm's length.

Rising and Setting of the Sun

We described the movement of stars in the night sky, but what about during the daytime? The stars continue to circle during the day, but the brilliance of the Sun makes them difficult to see. (The Moon can often be seen in the daylight, however.) On any given day, we can think of the Sun as being located at some position on the hypothetical celestial sphere. When the Sun rises—that is, when the rotation of Earth carries the Sun above the horizon—sunlight is scattered by the molecules of our atmosphere, filling our sky with light and hiding the stars above the horizon.

For thousands of years, astronomers have been aware that the Sun does more than just rise and set. It changes position gradually on the celestial sphere, moving each day about 1° to the east relative to the stars. Very reasonably, the ancients thought this meant the Sun was slowly moving around Earth, taking a period of time we call 1 year to make a full circle. Today, of course, we know it is Earth that is going around the Sun, but the effect is the same: the Sun's position in our sky changes day to day. We have a similar experience when we walk around a campfire at night; we see the flames appear in front of each person seated about the fire in turn.

The path the Sun appears to take around the celestial sphere each year is called the ecliptic (Figure 2.6). Because of its motion on the ecliptic, the Sun rises about 4 minutes later each day with respect to the stars. Earth must make just a bit more than one complete rotation (with respect to the stars) to bring the Sun up again.

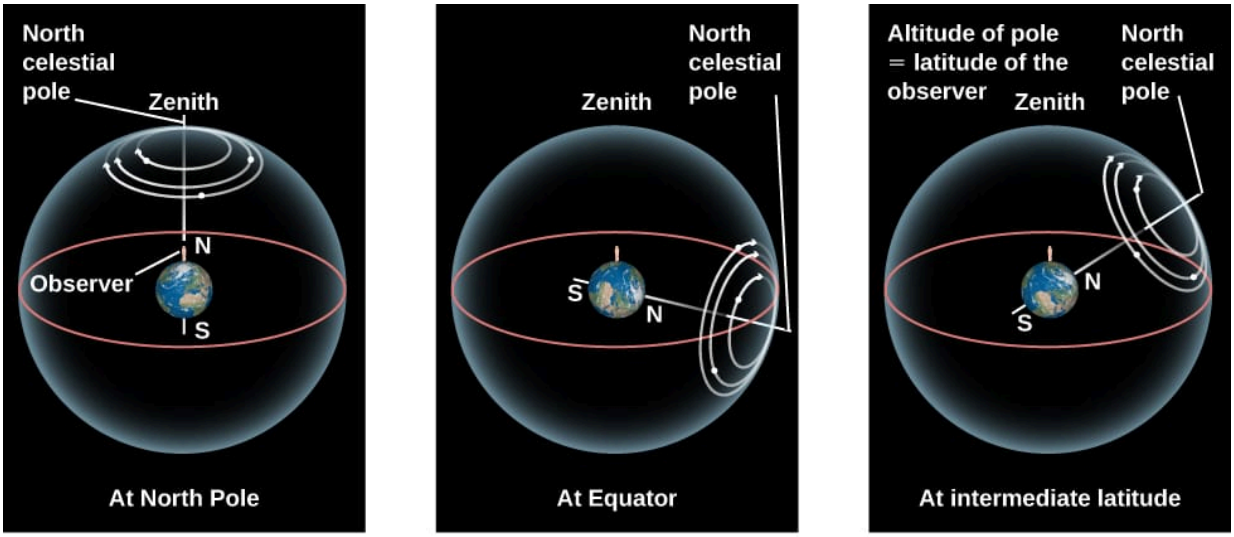


Figure 2.5 : Star Circles at Different Latitudes. The turning of the sky looks different depending on your latitude on Earth. The red circle in each case is your horizon. Your zenith is the point above your head. (a) At the North Pole, the stars circle the zenith and do not rise and set. (b) At the equator, the celestial poles are on the horizon, and the stars rise straight up and set straight down. (c) At intermediate latitudes, the north celestial pole is at some position between overhead and the horizon. Its angle above the horizon turns out to be equal to the observer's latitude. Stars rise and set at an angle to the horizon.

As the months go by and we look at the Sun from different places in our orbit, we see it projected against different places in our orbit, and thus against different stars in the background (Figure 2.6 and Table 2.1)—or we would, at least, if we could see the stars in the daytime. In practice, we must deduce which stars lie behind and beyond the Sun by observing the stars visible in the opposite direction at night. After a year, when Earth has completed one trip around the Sun, the Sun will appear to have completed one circuit of the sky along the ecliptic.

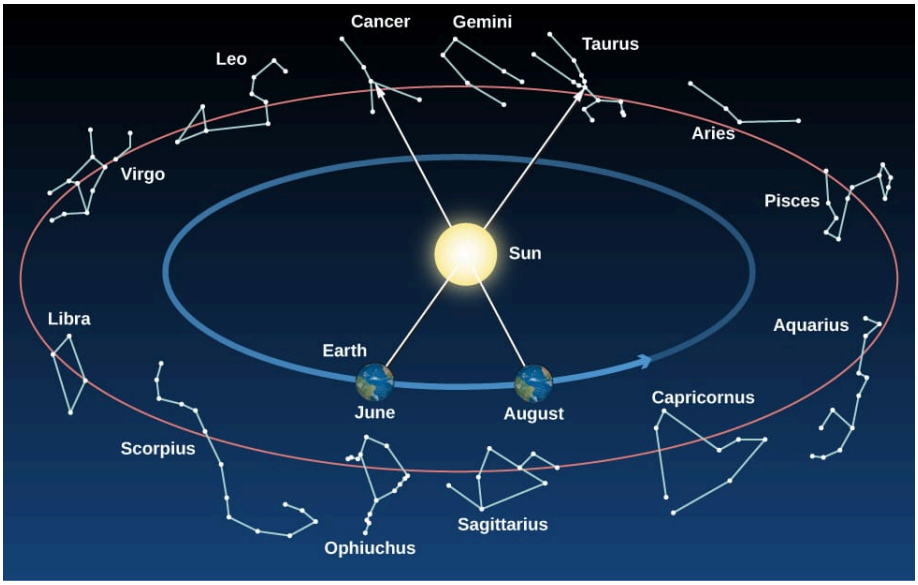


Figure 2.6 : Constellations on the Ecliptic. As Earth revolves around the Sun, we sit on "platform Earth" and see the Sun moving around the sky. The circle in the sky that the Sun appears to make around us in the course of a year is called the *ecliptic*. This circle (like all circles in the sky) goes through a set of constellations. The ancients thought these constellations, which the Sun (and the Moon and planets) visited, must be special and incorporated them into their system of astrology. Note that at any given time of the year, some of the constellations crossed by the ecliptic are visible in the night sky; others are in the day sky and are thus hidden by the brilliance of the Sun. (The term constellation will be defined more precisely in the next subsection.)

Table 2.1: Constellations on the Ecliptic

Constellation on the Ecliptic	Dates When the Sun Crosses It
Capricornus	January 21–February 16
Aquarius	February 16–March 11
Pisces	March 11–April 18
Aries	April 18–May 13
Taurus	May 13–June 22
Gemini	June 22–July 21
Cancer	July 21–August 10
Leo	August 10–September 16
Virgo	September 16–October 31
Libra	October 31–November 23
Scorpius	November 23–November 29
Ophiuchus	November 29–December 18
Sagittarius	December 18–January 21

The ecliptic does not lie along the celestial equator but is inclined to it at an angle of about 23.5°. In other words, the Sun's annual path in the sky is not linked with Earth's equator. This is because our planet's axis of rotation is tilted by about 23.5° from a vertical line sticking out of the plane of the ecliptic (Figure 2.7). Being tilted from "straight up" is not at all unusual among celestial bodies; Uranus and Pluto are actually tilted so much that they orbit the Sun "on their side."

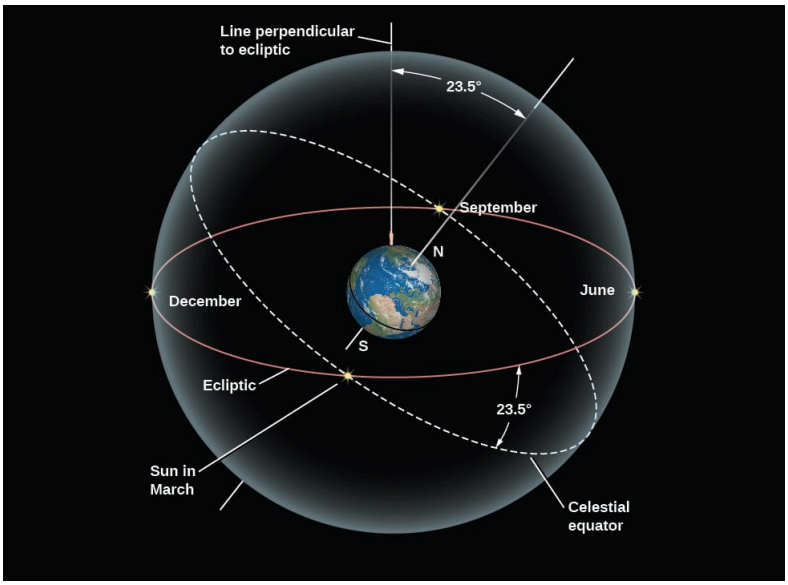


Figure 2.7 : The Celestial Tilt. The celestial equator is tilted by 23.5° to the ecliptic. As a result, North Americans and Europeans see the Sun north of the celestial equator and high in our sky in June, and south of the celestial equator and low in the sky in December.

The inclination of the ecliptic is the reason the Sun moves north and south in the sky as the seasons change. In Earth, Moon, and Sky, we discuss the progression of the seasons in more detail.

Fixed and Wandering Stars

The Sun is not the only object that moves among the fixed stars. The Moon and each of the planets that are visible to the unaided eye—Mercury, Venus, Mars, Jupiter, Saturn, and Uranus (although just barely)—also change their positions slowly from day to day. During a single day, the Moon and planets all rise and set as Earth turns, just as the Sun and stars do. But like the Sun, they have independent motions among the stars, superimposed on the daily rotation of the celestial sphere. Noticing these motions, the Greeks of 2000 years ago distinguished between what they called the *fixed stars*—those that maintain fixed patterns among themselves through many generations—and the *wandering stars*, or planets. The word “planet,” in fact, means “wanderer” in ancient Greek.

Today, we do not regard the Sun and Moon as planets, but the ancients applied the term to all seven of the moving objects in the sky. Much of ancient astronomy was devoted to observing and predicting the motions of these celestial wanderers. They even dedicated a unit of time, the week, to the seven objects that move on their own; that’s why there are 7 days in a week. The Moon, being Earth’s nearest celestial neighbor, has the fastest apparent motion; it completes a trip around the sky in about 1 month (or *moonth*). To do this, the Moon moves about 12°, or 24 times its own apparent width on the sky, each day.

Example 2.1: Angles in the Sky

A circle consists of 360 degrees (°). When we measure the angle in the sky that something moves, we can use this formula:

$$\text{speed} = \frac{\text{distance}}{\text{time}}$$

This is true whether the motion is measured in kilometers per hour or degrees per hour; we just need to use consistent units.

As an example, let’s say you notice the bright star Sirius due south from your observing location in the Northern Hemisphere. You note the time, and then later, you note the time that Sirius sets below the horizon. You find that Sirius has traveled an angular distance of about 75° in 5 h. About how many hours will it take for Sirius to return to its original location?

Solution

The speed of Sirius is $\frac{75^\circ}{5\,\mathrm{h}} = \frac{15^\circ}{1\,\mathrm{h}}$. If we want to know the time required for Sirius to return to its original location, we need to wait until it goes around a full circle, or 360°. Rearranging the formula for speed we were originally given, we find:

$$\text{time} = \frac{\text{distance}}{\text{speed}} = \frac{360^\circ}{15^\circ/\mathrm{h}} = 24\,\mathrm{h}$$

The actual time is a few minutes shorter than this, and we will explore why in a later chapter.

Exercise 3.1.1

The Moon moves in the sky relative to the background stars (in addition to moving with the stars as a result of Earth's rotation.) Go outside at night and note the position of the Moon relative to nearby stars. Repeat the observation a few hours later. How far has the Moon moved? (For reference, the diameter of the Moon is about 0.5°.) Based on your estimate of its motion, how long will it take for the Moon to return to the position relative to the stars in which you first observed it?

Answer available at the end of this section.

The individual paths of the Moon and planets in the sky all lie close to the ecliptic, although not exactly on it. This is because the paths of the planets about the Sun, and of the Moon about Earth, are all in nearly the same plane, as if they were circles on a huge sheet of paper. The planets, the Sun, and the Moon are thus always found in the sky within a narrow 18-degree-wide belt, centered on the ecliptic, called the zodiac (Figure 2.6). (The root of the term "zodiac" is the same as that of the word "zoo" and means a collection of animals; many of the patterns of stars within the zodiac belt reminded the ancients of animals, such as a fish or a goat.)

How the planets appear to move in the sky as the months pass is a combination of their actual motions plus the motion of Earth about the Sun; consequently, their paths are somewhat complex. As we will see, this complexity has fascinated and challenged astronomers for centuries.

Constellations

The backdrop for the motions of the "wanderers" in the sky is the canopy of stars. If there were no clouds in the sky and we were on a flat plain with nothing to obstruct our view, we could see about 3000 stars with the unaided eye. To find their way around such a multitude, the ancients found groupings of stars that made some familiar geometric pattern or (more rarely) resembled something they knew. Each civilization found its own patterns in the stars, much like a modern Rorschach test in which you are asked to discern patterns or pictures in a set of inkblots. The ancient Chinese, Egyptians, and Greeks, among others, found their own groupings—or constellations—of stars. These were helpful in navigating among the stars and in passing their star lore on to their children.

You may be familiar with some of the old star patterns we still use today, such as the Big Dipper, Little Dipper, and Orion the hunter, with his distinctive belt of three stars (Figure 2.8). However, many of the stars we see are not part of a distinctive star pattern at all, and a telescope reveals millions of stars too faint for the eye to see. Therefore, during the early decades of the 20th century, astronomers from many countries decided to establish a more formal system for organizing the sky.

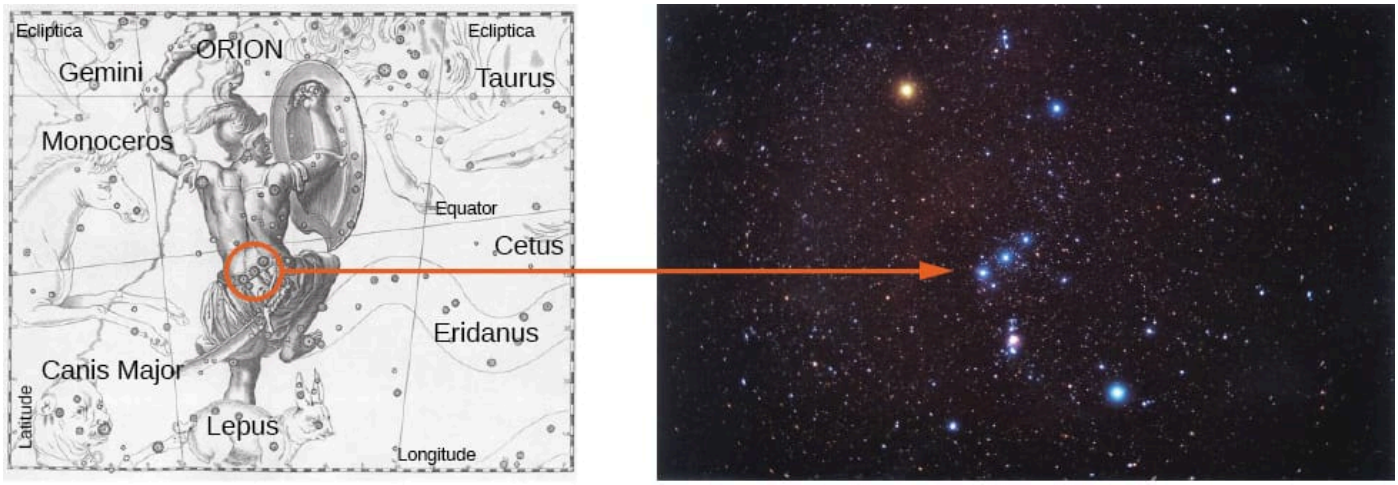


Figure 2.8 : Orion. (a) The winter constellation of Orion, the hunter, is surrounded by neighboring constellations, as illustrated in the seventeenth-century atlas by Hevelius. (b) A photograph shows the Orion region in the sky. Note the three blue stars that make up the belt of the hunter. The bright red star above the belt denotes his armpit and is called Betelgeuse (pronounced “Beetel-juice”). The bright blue star below the belt is his foot and is called Rigel. (credit a: modification of work by Johannes Hevelius; b: modification of work by Matthew Spinelli)

Today, we use the term *constellation* to mean one of 88 sectors into which we divide the sky, much as the United States is divided into 50 states. The modern boundaries between the constellations are imaginary lines in the sky running north–south and east–west, so that each point in the sky falls in a specific constellation, although, like the states, not all constellations are the same size. All the constellations are listed in Appendix L. Whenever possible, we have named each modern constellation after the Latin translations of one of the ancient Greek star patterns that lies within it. Thus, the modern constellation of Orion is a kind of box on the sky, which includes, among many other objects, the stars that made up the ancient picture of the hunter. Some people use the term *asterism* to denote an especially noticeable star pattern within a constellation (or sometimes spanning parts of several constellations). For example, the Big Dipper is an asterism within the constellation of Ursa Major, the Big Bear.

Students are sometimes puzzled because the constellations seldom resemble the people or animals for which they were named. In all likelihood, the Greeks themselves did not name groupings of stars because they looked like actual people or subjects (any more than the outline of Washington state resembles George Washington). Rather, they named sections of the sky in honor of the characters in their mythology and then fit the star configurations to the animals and people as best they could.

Link to Learning

This website about objects in the sky allows users to construct a detailed sky map showing the location and information about the Sun, Moon, planets, stars, constellations, and even satellites orbiting Earth. Begin by setting your observing location using the option in the menu in the upper right corner of the screen. An excellent website called Figures in the Sky shows the constellation figures imagined by cultures around the world.

Answers

Exercise 3.1.1

The speed of the moon is 0.5°/1 h. To move a full 360°, the moon needs 720 h: $\frac{0.5°}{1\,\text{h}} = \frac{360°}{720\,\text{h}}$. Dividing 720 h by the conversion factor of 24 h/day reveals the lunar cycle is about 30 days.

This page titled 3.1: The Sky Above is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 2.2: The Sky Above by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

3.2: Ancient Astronomy

Learning Objectives

By the end of this section, you will be able to:

- Describe early examples of astronomy around the world
- Explain how Greek astronomers were able to deduce that Earth is spherical
- Explain how Greek astronomers were able to calculate Earth's size
- Describe the motion of Earth called precession
- Describe Ptolemy's geocentric system of planetary motion

Let us now look briefly back into history. Much of modern Western civilization is derived in one way or another from the ideas of the ancient Greeks and Romans, and this is true in astronomy as well. However, many other ancient cultures also developed sophisticated systems for observing and interpreting the sky.

Astronomy around the World

Ancient Babylonian, Assyrian, and Egyptian astronomers knew the approximate length of the year. The Egyptians of 3000 years ago, for example, adopted a calendar based on a 365-day year. They kept careful track of the rising time of the bright star Sirius in the predawn sky, which has a yearly cycle that corresponded with the flooding of the Nile River. The Chinese also had a working calendar; they determined the length of the year at about the same time as the Egyptians. The Chinese also recorded comets, bright meteors, and dark spots on the Sun. (Many types of astronomical objects were introduced in Science and the Universe: A Brief Tour. If you are not familiar with terms like *comets* and *meteors*, you may want to review that chapter.) Later, Chinese astronomers kept careful records of "guest stars"—those that are normally too faint to see but suddenly flare up to become visible to the unaided eye for a few weeks or months. We still use some of these records in studying stars that exploded a long time ago.

The Mayan culture in Mexico and Central America developed a sophisticated calendar based on the planet Venus, and they made astronomical observations from sites dedicated to this purpose a thousand years ago. The Polynesians learned to navigate by the stars over hundreds of kilometers of open ocean—a skill that enabled them to colonize new islands far away from where they began.

In Britain, before the widespread use of writing, ancient people used stones to keep track of the motions of the Sun and Moon. We still find some of the great stone circles they built for this purpose, dating from as far back as 2800 BCE. The best known of these is Stonehenge, which is discussed in Earth, Moon, and Sky.[1]

Early Greek and Roman Cosmology

Our concept of the cosmos—its basic structure and origin—is called cosmology, a word with Greek roots. Before the invention of telescopes, humans had to depend on the simple evidence of their senses for a picture of the universe. The ancients developed cosmologies that combined their direct view of the heavens with a rich variety of philosophical and religious symbolism.

At least 2000 years before Columbus, educated people in the eastern Mediterranean region knew Earth was round. Belief in a spherical Earth may have stemmed from the time of Pythagoras, a philosopher and mathematician who lived 2500 years ago. He believed circles and spheres to be "perfect forms" and suggested that Earth should therefore be a sphere. As evidence that the gods liked spheres, the Greeks cited the fact that the Moon is a sphere, using evidence we describe later.

The writings of Aristotle (384–322 BCE), the tutor of Alexander the Great, summarize many of the ideas of his day. They describe how the progression of the Moon's phases—its apparent changing shape—results from our seeing different portions of the Moon's sunlit hemisphere as the month goes by (see Earth, Moon, and Sky). Aristotle also knew that the Sun has to be farther away from Earth than is the Moon because occasionally the Moon passed exactly between Earth and the Sun and hid the Sun temporarily from view. We call this a *solar eclipse*.

Aristotle cited convincing arguments that Earth must be round. First is the fact that as the Moon enters or emerges from Earth's shadow during an eclipse of the Moon, the shape of the shadow seen on the Moon is always round (Figure 2.9). Only a spherical object always produces a round shadow. If Earth were a disk, for example, there would be some occasions when the sunlight would strike it edge-on and its shadow on the Moon would be a line.



Figure 2.9 : Earth's Round Shadow. A lunar eclipse occurs when the Moon moves into and out of Earth's shadow. Note the curved shape of the shadow—evidence for a spherical Earth that has been recognized since antiquity. (credit: modification of work by Brian Paczkowski)

As a second argument, Aristotle explained that travelers who go south a significant distance are able to observe stars that are not visible farther north. And the height of the North Star—the star nearest the north celestial pole—decreases as a traveler moves south. On a flat Earth, everyone would see the same stars overhead. The only possible explanation is that the traveler must have moved over a curved surface on Earth, showing stars from a different angle. (See the How Do We Know Earth Is Round? feature for more ideas on proving Earth is round.)

One Greek thinker, Aristarchus of Samos (310–230 BCE), even suggested that Earth was moving around the Sun, but Aristotle and most of the ancient Greek scholars rejected this idea. One of the reasons for their conclusion was the thought that if Earth moved about the Sun, they would be observing the stars from different places along Earth's orbit. As Earth moved along, nearby stars should shift their positions in the sky relative to more distant stars. In a similar way, we see foreground objects appear to move against a more distant background whenever we are in motion. When we ride on a train, the trees in the foreground appear to shift their position relative to distant hills as the train rolls by. Unconsciously, we use this phenomenon all of the time to estimate distances around us.

The apparent shift in the direction of an object as a result of the motion of the observer is called parallax. We call the shift in the apparent direction of a star due to Earth's orbital motion *stellar parallax*. The Greeks made dedicated efforts to observe stellar parallax, even enlisting the aid of Greek soldiers with the clearest vision, but to no avail. The brighter (and presumably nearer) stars just did not seem to shift as the Greeks observed them in the spring and then again in the fall (when Earth is on the opposite side of the Sun).

This meant either that Earth was not moving or that the stars had to be so tremendously far away that the parallax shift was immeasurably small. A cosmos of such enormous extent required a leap of imagination that most ancient philosophers were not prepared to make, so they retreated to the safety of the Earth-centered view, which would dominate Western thinking for nearly two millennia.

Astronomy Basics: How Do We Know Earth Is Round?

In addition to the two ways (from Aristotle's writings) discussed in this chapter, you might also reason as follows:

1. Let's watch a ship leave its port and sail into the distance on a clear day. On a flat Earth, we would just see the ship get smaller and smaller as it sails away. But this isn't what we actually observe. Instead, ships sink below the horizon, with the hull disappearing first and the mast remaining visible for a while longer. Eventually, only the top of the mast can be seen as the ship sails around the curvature of Earth. Finally, the ship disappears under the horizon.
2. The International Space Station circles Earth once every 90 minutes or so. Photographs taken from the shuttle and other satellites show that Earth is round from every perspective.
3. Suppose you made a friend in each time zone of Earth. You call all of them at the same hour and ask, "Where is the Sun?" On a flat Earth, each caller would give you roughly the same answer. But on a round Earth you would find that, for some friends, the Sun would be high in the sky whereas for others it would be rising, setting, or completely out of sight (and this last group of friends would be upset with you for waking them up).

Measurement of Earth by Eratosthenes

The Greeks not only knew Earth was round, but also they were able to measure its size. The first fairly accurate determination of Earth's diameter was made in about 200 BCE by Eratosthenes (276–194 BCE), a Greek living in Alexandria, Egypt. His method was

a geometric one, based on observations of the Sun.

The Sun is so distant from us that all the light rays that strike our planet approach us along essentially parallel lines. To see why, look at Figure 2.10. Take a source of light near Earth—say, at position A. Its rays strike different parts of Earth along diverging paths. From a light source at B, or at C (which is still farther away), the angle between rays that strike opposite parts of Earth is smaller. The more distant the source, the smaller the angle between the rays. For a source infinitely distant, the rays travel along parallel lines.

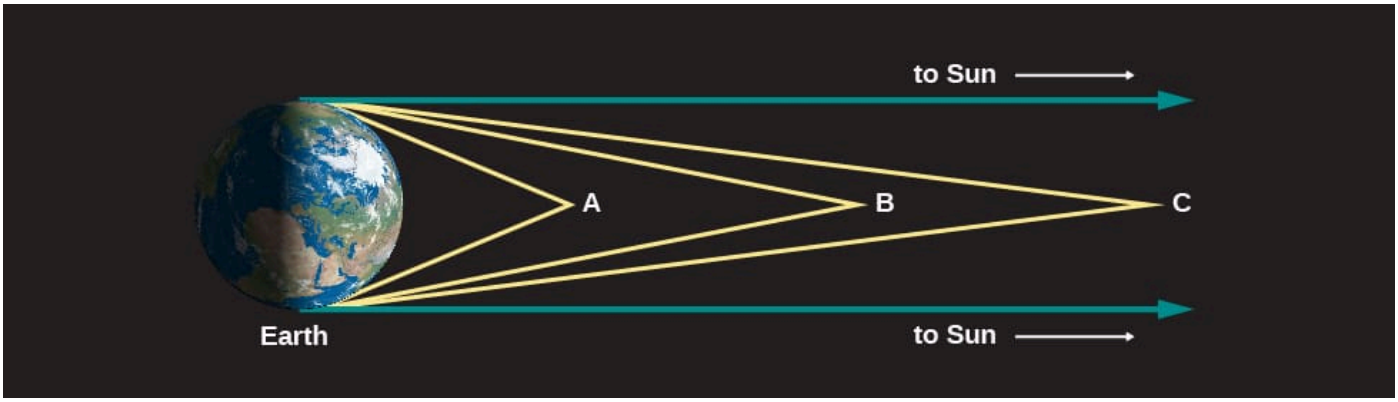


Figure 2.10 : Light Rays from Space. The more distant an object, the more nearly parallel the rays of light coming from it.

Of course, the Sun is not infinitely far away, but given its distance of 150 million kilometers, light rays striking Earth from a point on the Sun diverge from one another by an angle far too small to be observed with the unaided eye. As a consequence, if people all over Earth who could see the Sun were to point at it, their fingers would, essentially, all be parallel to one another. (The same is also true for the planets and stars—an idea we will use in our discussion of how telescopes work.)

Eratosthenes was told that on the first day of summer at Syene, Egypt (near modern Aswan), sunlight struck the bottom of a vertical well at noon. This indicated that the Sun was directly over the well—meaning that Syene was on a direct line from the center of Earth to the Sun. At the corresponding time and date in Alexandria, Eratosthenes observed the shadow a column made and saw that the Sun was not directly overhead, but was slightly south of the zenith, so that its rays made an angle with the vertical equal to about 1/50 of a circle (7°). Because the Sun's rays striking the two cities are parallel to one another, why would the two rays not make the same angle with Earth's surface? Eratosthenes reasoned that the curvature of the round Earth meant that "straight up" was not the same in the two cities. And the measurement of the angle in Alexandria, he realized, allowed him to figure out the size of Earth. Alexandria, he saw, must be 1/50 of Earth's circumference north of Syene (Figure 2.11). Alexandria had been measured to be 5000 stadia north of Syene. (The *stadium* was a Greek unit of length, derived from the length of the racetrack in a stadium.) Eratosthenes thus found that Earth's circumference must be 50×5000, or 250,000 stadia.

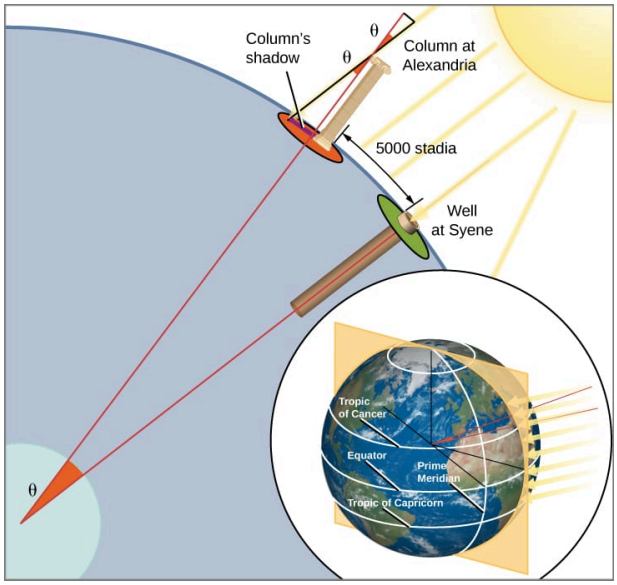


Figure 2.11 : How Eratosthenes Measured the Size of Earth. Eratosthenes measured the size of Earth by observing the angle at which the Sun's rays hit our planet's surface. The Sun's rays come in parallel, but because Earth's surface curves, a ray at Syene comes straight down whereas a ray at Alexandria makes an angle of 7° with the vertical. That means, in effect, that at Alexandria, Earth's surface has curved away from Syene by 7° of 360°, or 1/50 of a full circle. Thus, the distance between the two cities must be 1/50 the circumference of Earth. (credit: modification of work by NOAA Ocean Service Education)

It is not possible to evaluate precisely the accuracy of Eratosthenes solution because there is doubt about which of the various kinds of Greek stadia he used as his unit of distance. If it was the common Olympic stadium, his result is about 20% too large. According to another interpretation, he used a stadium equal to about 1/6 kilometer, in which case his figure was within 1% of the correct value of 40,000 kilometers. Even if his measurement was not exact, his success at measuring the size of our planet by using only shadows, sunlight, and the power of human thought was one of the greatest intellectual achievements in history.

Hipparchus and Precession

Perhaps the greatest astronomer of antiquity was Hipparchus, born in Nicaea in what is present-day Turkey. He erected an observatory on the island of Rhodes around 150 BCE, when the Roman Republic was expanding its influence throughout the Mediterranean region. There he measured, as accurately as possible, the positions of objects in the sky, compiling a pioneering star catalog with about 850 entries. He designated celestial coordinates for each star, specifying its position in the sky, just as we specify the position of a point on Earth by giving its latitude and longitude.

He also divided the stars into apparent magnitudes according to their apparent brightness. He called the brightest ones "stars of the first magnitude"; the next brightest group, "stars of the second magnitude"; and so forth. This rather arbitrary system, in modified form, still remains in use today (although it is less and less useful for professional astronomers).

By observing the stars and comparing his data with older observations, Hipparchus made one of his most remarkable discoveries: the position in the sky of the north celestial pole had altered over the previous century and a half. Hipparchus deduced correctly that this had happened not only during the period covered by his observations, but was in fact happening all the time: the direction around which the sky appears to rotate changes slowly but continuously. Recall from the section on celestial poles and the celestial equator that the north celestial pole is just the projection of Earth's North Pole into the sky. If the north celestial pole is wobbling around, then Earth itself must be doing the wobbling. Today, we understand that the direction in which Earth's axis points does indeed change slowly but regularly—a motion we call precession. If you have ever watched a spinning top wobble, you observed a similar kind of motion. The top's axis describes a path in the shape of a cone, as Earth's gravity tries to topple it (Figure 2.12).

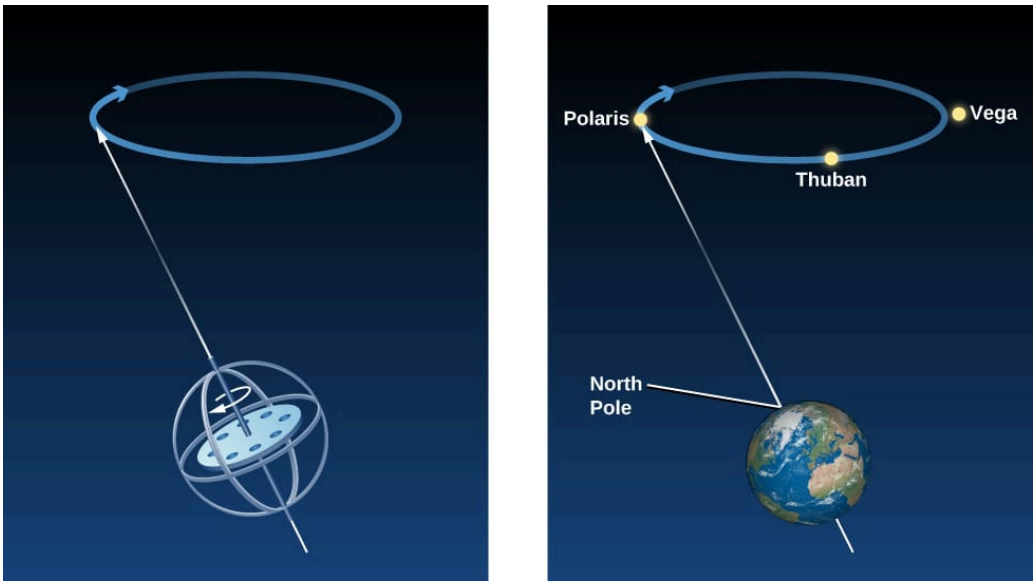


Figure 2.12 : Precession. Just as the axis of a rapidly spinning top wobbles slowly in a circle, so the axis of Earth wobbles in a 26,000-year cycle. Today the north celestial pole is near the star Polaris, but about 5000 years ago it was close to a star called Thuban, and in 14,000 years it will be closest to the star Vega.

Because our planet is not an exact sphere, but bulges a bit at the equator, the pulls of the Sun and Moon cause it to wobble like a top. It takes about 26,000 years for Earth's axis to complete one circle of precession. As a result of this motion, the point where our axis points in the sky changes as time goes on. While Polaris is the star closest to the north celestial pole today (it will reach its closest point around the year 2100), the star Vega in the constellation of Lyra will be the North Star in 14,000 years.

Ptolemy's Model of the Solar System

The last great astronomer of the Roman era was Claudius Ptolemy (or Ptolemaeus), who flourished in Alexandria in about the year 140. He wrote a mammoth compilation of astronomical knowledge, which today is called by its Arabic name, *Almagest* (meaning "The Greatest"). *Almagest* does not deal exclusively with Ptolemy's own work; it includes a discussion of the astronomical achievements of the past, principally those of Hipparchus. Today, it is our main source of information about the work of Hipparchus and other Greek astronomers.

Ptolemy's most important contribution was a geometric representation of the solar system that predicted the positions of the planets for any desired date and time. Hipparchus, not having enough data on hand to solve the problem himself, had instead amassed observational material for posterity to use. Ptolemy supplemented this material with new observations of his own and produced a cosmological model that endured more than a thousand years, until the time of Copernicus.

The complicating factor in explaining the motions of the planets is that their apparent wandering in the sky results from the combination of their own motions with Earth's orbital revolution. As we watch the planets from our vantage point on the moving Earth, it is a little like watching a car race while you are competing in it. Sometimes opponents' cars pass you, but at other times you pass them, making them appear to move backward for a while with respect to you.

Figure 2.13 shows the motion of Earth and a planet farther from the Sun—in this case, Mars. Earth travels around the Sun in the same direction as the other planet and in nearly the same plane, but its orbital speed is faster. As a result, it overtakes the planet periodically, like a faster race car on the inside track. The figure shows where we see the planet in the sky at different times. The path of the planet among the stars is illustrated in the star field on the right side of the figure.

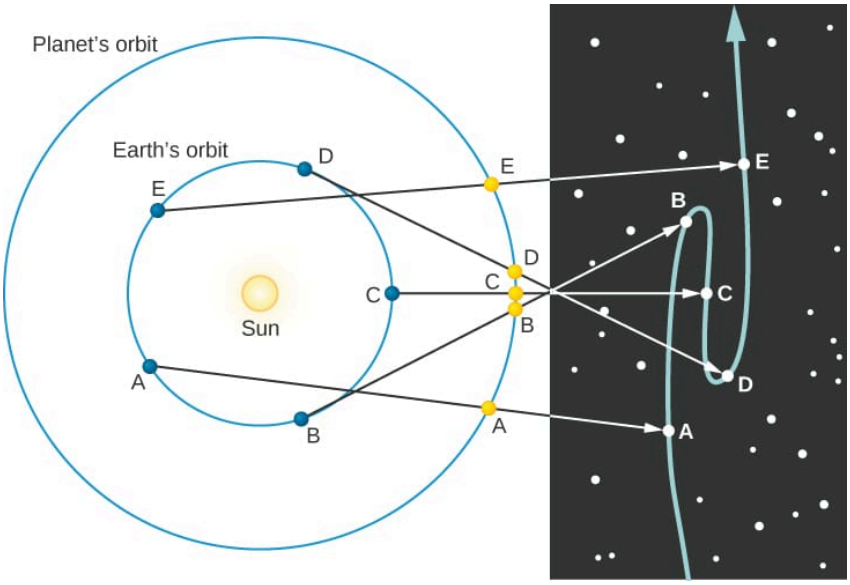


Figure 2.13 : Retrograde Motion of a Planet beyond Earth's Orbit. The letters on the diagram show where Earth and Mars are at different times. By following the lines from each Earth position through each corresponding Mars position, you can see how the retrograde path of Mars looks against the background stars.

Link to Learning

The planetary configurations simulator from Foothill AstroSims allows you to see the usual prograde and occasional retrograde motion of other planets. You can switch back and forth between viewing motion from Earth and Mars (as well as other planets).

Normally, planets move eastward in the sky over the weeks and months as they orbit the Sun, but from positions B to D in Figure 2.13, as Earth passes the planets in our example, it appears to drift backward, moving west in the sky. Even though it is actually moving to the east, the faster-moving Earth has overtaken it and seems, from our perspective, to be leaving it behind. As Earth rounds its orbit toward position E, the planet again takes up its apparent eastward motion in the sky. The temporary apparent westward motion of a planet as Earth swings between it and the Sun is called retrograde motion. Such backward motion is much easier for us to understand today, now that we know Earth is one of the moving planets and not the unmoving center of all creation. But Ptolemy was faced with the far more complex problem of explaining such motion while assuming a stationary Earth.

Furthermore, because the Greeks believed that celestial motions had to be circles, Ptolemy had to construct his model using circles alone. To do it, he needed dozens of circles, some moving around other circles, in a complex structure that makes a modern viewer dizzy. But we must not let our modern judgment cloud our admiration for Ptolemy's achievement. In his day, a complex universe centered on Earth was perfectly reasonable and, in its own way, quite beautiful. However, as Alfonso X, the King of Castile, was reported to have said after having the Ptolemaic system of planet motions explained to him, "If the Lord Almighty had consulted me before embarking upon Creation, I should have recommended something simpler."

Ptolemy solved the problem of explaining the observed motions of planets by having each planet revolve in a small orbit called an epicycle. The center of the epicycle then revolved about Earth on a circle called a *deferent* (Figure 2.14). When the planet is at position *x* in Figure 2.14 on the epicycle orbit, it is moving in the same direction as the center of the epicycle; from Earth, the planet appears to be moving eastward. When the planet is at *y*, however, its motion is in the direction opposite to the motion of the epicycle's center around Earth. By choosing the right combination of speeds and distances, Ptolemy succeeded in having the planet moving westward at the correct speed and for the correct interval of time, thus replicating retrograde motion with his model.

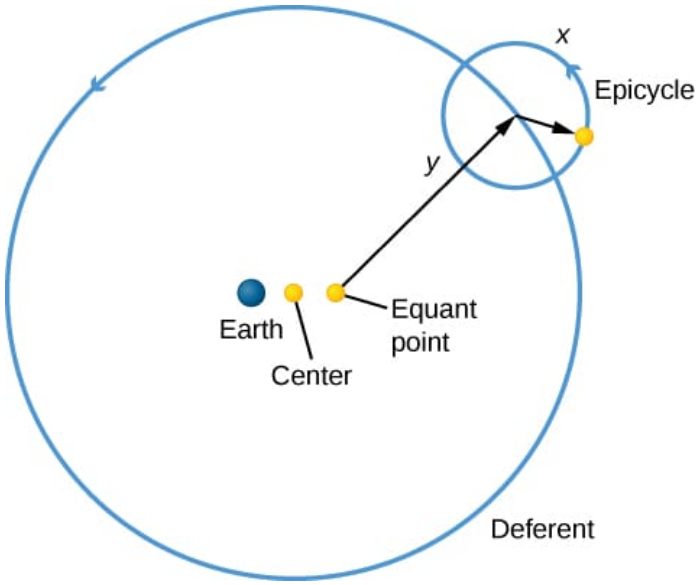


Figure 2.14 : Ptolemy's Complicated Cosmological System. Each planet orbits around a small circle called an *epicycle*. Each epicycle orbits on a larger circle called the *deferent*. This system is not centered exactly on Earth but on an offset point called the *equant*. The Greeks needed all this complexity to explain the actual motions in the sky because they believed that Earth was stationary and that all sky motions had to be circular.

Link to Learning

Use the Ptolemaic System simulator from Foothill AstroSims to explore how Ptolemy's system of deferents and epicycles explained the apparent motion of the planets.

However, we shall see in Orbits and Gravity that the planets, like Earth, travel about the Sun in orbits that are ellipses, not circles. Their actual behavior cannot be represented accurately by a scheme of uniform circular motions. In order to match the observed motions of the planets, Ptolemy had to center the deferent circles, not on Earth, but at points some distance from Earth. In addition, he introduced uniform circular motion around yet another axis, called the *equant point*. All of these considerably complicated his scheme.

It is a tribute to the genius of Ptolemy as a mathematician that he was able to develop such a complex system to account successfully for the observations of planets. It may be that Ptolemy did not intend for his cosmological model to describe reality, but merely to serve as a mathematical representation that allowed him to predict the positions of the planets at any time. Whatever his thinking, his model, with some modifications, was eventually accepted as authoritative in the Muslim world and (later) in Christian Europe.

Footnotes

- 1For resources providing more information about the astronomy of diverse cultures around the world, please see http://bit.ly/astrocultures.

This page titled 3.2: Ancient Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 2.3: Ancient Astronomy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

3.3: Key Terms

accelerate
to change velocity; to speed up, slow down, or change direction.

apparent magnitude
a measure of how bright a star looks in the sky; the larger the number, the dimmer the star appears to us

astrology
the pseudoscience that deals with the supposed influences on human destiny of the configurations and locations in the sky of the Sun, Moon, and planets

celestial equator
a great circle on the celestial sphere 90° from the celestial poles; where the celestial sphere intersects the plane of Earth's equator

celestial poles
points about which the celestial sphere appears to rotate; intersections of the celestial sphere with Earth's polar axis

celestial sphere
the apparent sphere of the sky; a sphere of large radius centered on the observer; directions of objects in the sky can be denoted by their position on the celestial sphere

circumpolar zone
those portions of the celestial sphere near the celestial poles that are either always above or always below the horizon

cosmology
the study of the organization and evolution of the universe

ecliptic
the apparent annual path of the Sun on the celestial sphere

epicycle
the circular orbit of a body in the Ptolemaic system, the center of which revolves about another circle (the deferent)

geocentric
centered on Earth

heliocentric
centered on the Sun

horizon (astronomical)
a great circle on the celestial sphere 90° from the zenith; more popularly, the circle around us where the dome of the sky meets Earth

horoscope
a chart used by astrologers that shows the positions along the zodiac and in the sky of the Sun, Moon, and planets at some given instant and as seen from a particular place on Earth—usually corresponding to the time and place of a person's birth

parallax
the apparent displacement of a nearby star that results from the motion of Earth around the Sun

planet
today, any of the larger objects revolving about the Sun or any similar objects that orbit other stars; in ancient times, any object that moved regularly among the fixed stars

precession (of Earth)
: the slow, conical motion of Earth's axis of rotation caused principally by the gravitational pull of the Moon and Sun on Earth's equatorial bulge

retrograde motion
: the apparent westward motion of a planet on the celestial sphere or with respect to the stars

year
: the period of revolution of Earth around the Sun

zenith
: the point on the celestial sphere opposite the direction of gravity; point directly above the observer

zodiac
: a belt around the sky about 18° wide centered on the ecliptic

This page titled 3.3: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

4: MODERN ASTRONOMY

4.1: The Birth of Modern Astronomy
4.2: The Laws of Planetary Motion

This page titled 4: Modern Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

4.1: The Birth of Modern Astronomy

Learning Objectives

By the end of this section, you will be able to:

- Explain how Copernicus developed the heliocentric model of the solar system
- Explain the Copernican model of planetary motion and describe evidence or arguments in favor of it
- Describe Galileo's discoveries concerning the study of motion and forces
- Explain how Galileo's discoveries tilted the balance of evidence in favor of the Copernican model

Astronomy made no major advances in strife-torn medieval Europe. The birth and expansion of Islam after the seventh century led to a flowering of Arabic and Jewish cultures that preserved, translated, and added to many of the astronomical ideas of the Greeks. Many of the names of the brightest stars, for example, are today taken from the Arabic, as are such astronomical terms as "zenith."

As European culture began to emerge from its long, dark age, trading with Arab countries led to a rediscovery of ancient texts such as *Almagest* and to a reawakening of interest in astronomical questions. This time of rebirth (in French, "*renaissance*") in astronomy was embodied in the work of Copernicus (Figure 2.16).

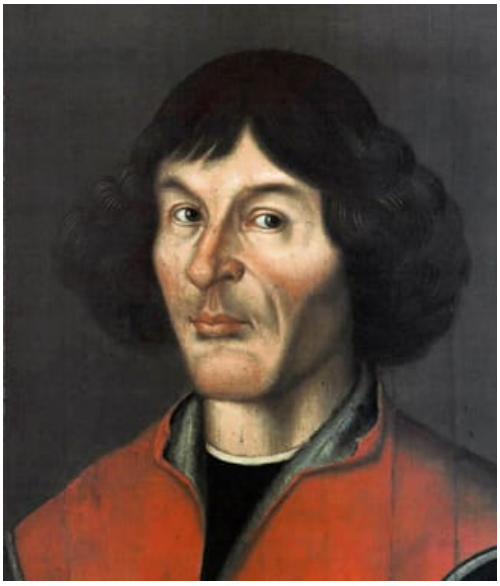


Figure 2.16 : Nicolaus Copernicus (1473–1543). Copernicus was a cleric and scientist who played a leading role in the emergence of modern science. Although he could not prove that Earth revolves about the Sun, he presented such compelling arguments for this idea that he turned the tide of cosmological thought and laid the foundations upon which Galileo and Kepler so effectively built in the following century.

Copernicus

One of the most important events of the Renaissance was the displacement of Earth from the center of the universe, an intellectual revolution initiated by a Polish cleric in the sixteenth century. Nicolaus Copernicus was born in Torun, a mercantile town along the Vistula River. His training was in law and medicine, but his main interests were astronomy and mathematics. His great contribution to science was a critical reappraisal of the existing theories of planetary motion and the development of a new Sun-centered, or heliocentric, model of the solar system. Copernicus concluded that Earth is a planet and that all the planets circle the Sun. Only the Moon orbits Earth (Figure 2.17).

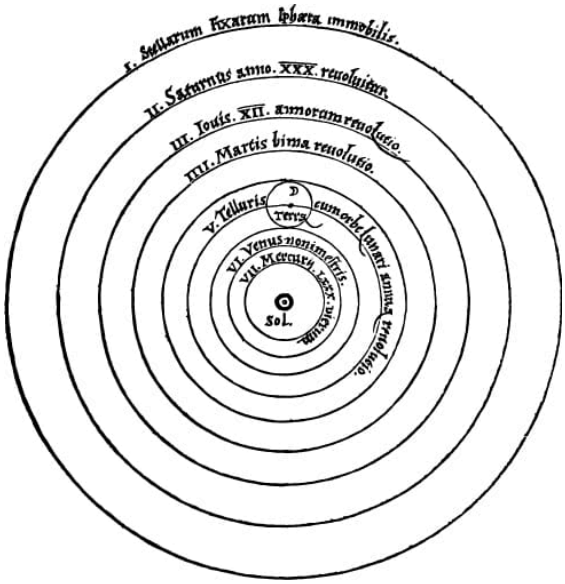


Figure 2.17 : Copernicus' System. Copernicus developed a heliocentric plan of the solar system. This system was published in the first edition of *De Revolutionibus Orbium Coelestium*. Notice the word *Sol* for "Sun" in the middle. (credit: Nicolai Copernici)

Copernicus described his ideas in detail in his book *De Revolutionibus Orbium Coelestium* (*On the Revolution of Celestial Orbs*), published in 1543, the year of his death. By this time, the old Ptolemaic system needed significant adjustments to predict the positions of the planets correctly. Copernicus wanted to develop an improved theory from which to calculate planetary positions, but in doing so, he was himself not free of all traditional prejudices.

He began with several assumptions that were common in his time, such as the idea that the motions of the heavenly bodies must be made up of combinations of uniform circular motions. But he did not assume (as most people did) that Earth had to be in the center of the universe, and he presented a defense of the heliocentric system that was elegant and persuasive. His ideas, although not widely accepted until more than a century after his death, were much discussed among scholars and, ultimately, had a profound influence on the course of world history.

One of the objections raised to the heliocentric theory was that if Earth were moving, we would all sense or feel this motion. Solid objects would be ripped from the surface, a ball dropped from a great height would not strike the ground directly below it, and so forth. But a moving person is not necessarily aware of that motion. We have all experienced seeing an adjacent train, bus, or ship appear to move, only to discover that it is we who are moving.

Copernicus argued that the apparent motion of the Sun about Earth during the course of a year could be represented equally well by a motion of Earth about the Sun. He also reasoned that the apparent rotation of the celestial sphere could be explained by assuming that Earth rotates while the celestial sphere is stationary. To the objection that if Earth rotated about an axis it would fly into pieces, Copernicus answered that if such motion would tear Earth apart, the still faster motion of the much larger celestial sphere required by the geocentric hypothesis would be even more devastating.

The Heliocentric Model

The most important idea in Copernicus' *De Revolutionibus* is that Earth is one of six (then-known) planets that revolve about the Sun. Using this concept, he was able to work out the correct general picture of the solar system. He placed the planets, starting nearest the Sun, in the correct order: Mercury, Venus, Earth, Mars, Jupiter, and Saturn. Further, he deduced that the nearer a planet is to the Sun, the greater its orbital speed. With his theory, he was able to explain the complex retrograde motions of the planets without epicycles and to work out a roughly correct scale for the solar system.

Copernicus could not prove that Earth revolves about the Sun. In fact, with some adjustments, the old Ptolemaic system could have accounted, as well, for the motions of the planets in the sky. But Copernicus pointed out that the Ptolemaic cosmology was clumsy and lacking the beauty and symmetry of its successor.

In Copernicus' time, in fact, few people thought there were ways to prove whether the heliocentric or the older geocentric system was correct. A long philosophical tradition, going back to the Greeks and defended by the Catholic Church, held that pure human

thought combined with divine revelation represented the path to truth. Nature, as revealed by our senses, was suspect. For example, Aristotle had reasoned that heavier objects (having more of the quality that made them heavy) must fall to Earth faster than lighter ones. This is absolutely incorrect, as any simple experiment dropping two balls of different weights shows. However, in Copernicus' day, experiments did not carry much weight (if you will pardon the expression); Aristotle's reasoning was more convincing.

In this environment, there was little motivation to carry out observations or experiments to distinguish between competing cosmological theories (or anything else). It should not surprise us, therefore, that the heliocentric idea was debated for more than half a century without any tests being applied to determine its validity. (In fact, in the North American colonies, the older geocentric system was still taught at Harvard University in the first years after it was founded in 1636.)

Contrast this with the situation today, when scientists rush to test each new hypothesis and do not accept any ideas until the results are in. For example, when two researchers at the University of Utah announced in 1989 that they had discovered a way to achieve nuclear fusion (the process that powers the stars) at room temperature, other scientists at more than 25 laboratories around the United States attempted to duplicate "cold fusion" within a few weeks—without success, as it turned out. The cold fusion theory soon went down in flames.

How would we look at Copernicus' model today? When a new hypothesis or theory is proposed in science, it must first be checked for consistency with what is already known. Copernicus' heliocentric idea passes this test, for it allows planetary positions to be calculated at least as well as does the geocentric theory. The next step is to determine which predictions the new hypothesis makes that differ from those of competing ideas. In the case of Copernicus, one example is the prediction that, if Venus circles the Sun, the planet should go through the full range of phases just as the Moon does, whereas if it circles Earth, it should not (Figure 2.18). Also, we should not be able to see the full phase of Venus from Earth because the Sun would then be between Venus and Earth. But in those days, before the telescope, no one imagined testing these predictions.

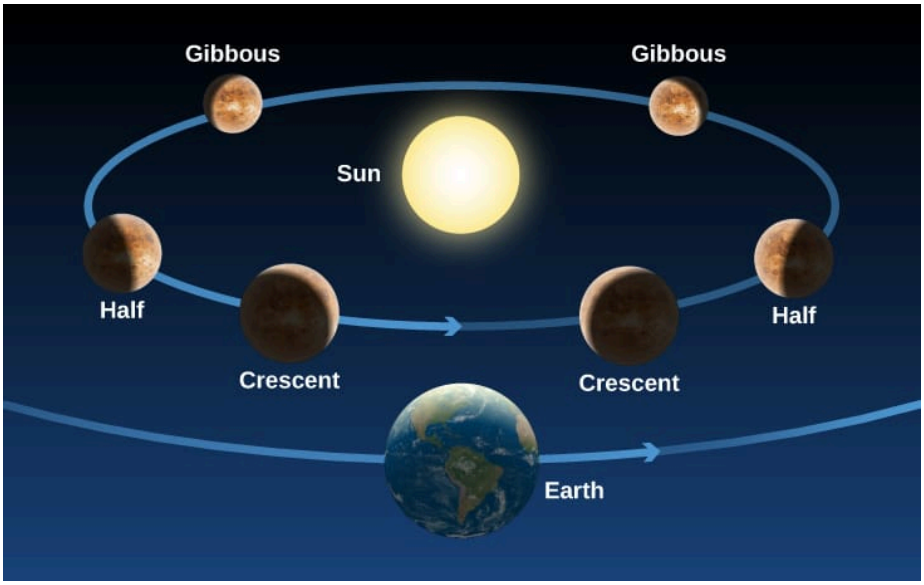


Figure 2.18 : Phases of Venus. As Venus moves around the Sun, we see changing illumination of its surface, just as we see the face of the Moon illuminated differently in the course of a month.

Link to Learning

This animation shows the phases of Venus. You can also see its distance from Earth as it orbits the Sun.

Galileo and the Beginning of Modern Science

Many of the modern scientific concepts of observation, experimentation, and the testing of hypotheses through careful quantitative measurements were pioneered by a man who lived nearly a century after Copernicus. Galileo Galilei (Figure 2.19), a contemporary of Shakespeare, was born in Pisa. Like Copernicus, he began training for a medical career, but he had little interest in the subject and later switched to mathematics. He held faculty positions at the University of Pisa and the University of Padua,

and eventually became mathematician to the Grand Duke of Tuscany in Florence.



Figure 2.19 : Galileo Galilei (1564–1642). Galileo advocated that we perform experiments or make observations to ask nature its ways. When Galileo turned the telescope to the sky, he found things were not the way philosophers had supposed.

Galileo's greatest contributions were in the field of mechanics, the study of motion and the actions of forces on bodies. It was familiar to all persons then, as it is to us now, that if something is at rest, it tends to remain at rest and requires some outside influence to start it in motion. Rest was thus generally regarded as the natural state of matter. Galileo showed, however, that rest is no more natural than motion.

If an object is slid along a rough horizontal floor, it soon comes to rest because friction between it and the floor acts as a retarding force. However, if the floor and the object are both highly polished, the object, given the same initial speed, will slide farther before stopping. On a smooth layer of ice, it will slide farther still. Galileo reasoned that if all resisting effects could be removed, the object would continue in a steady state of motion indefinitely. He argued that a force is required not only to start an object moving from rest but also to slow down, stop, speed up, or change the direction of a moving object. You will appreciate this if you have ever tried to stop a rolling car by leaning against it, or a moving boat by tugging on a line.

Galileo also studied the way objects accelerate—change their speed or direction of motion. Galileo watched objects as they fell freely or rolled down a ramp. He found that such objects accelerate uniformly; that is, in equal intervals of time they gain equal increments in speed. Galileo formulated these newly found laws in precise mathematical terms that enabled future experimenters to predict how far and how fast objects would move in various lengths of time.

Link to Learning

In theory, if Galileo is right, a feather and a hammer, dropped at the same time from a height, should land at the same moment. On Earth, this experiment is not possible because air resistance and air movements make the feather flutter, instead of falling straight down, accelerated only by the force of gravity. For generations, physics teachers had said that the place to try this experiment is somewhere where there is no air, such as the Moon. In 1971, *Apollo 15* astronaut David Scott took a hammer and feather to the Moon and tried it, to the delight of physics nerds everywhere. NASA provides the video of the hammer and feather as well as a brief explanation.

Sometime in the 1590s, Galileo adopted the Copernican hypothesis of a heliocentric solar system. In Roman Catholic Italy, this was not a popular philosophy, for Church authorities still upheld the ideas of Aristotle and Ptolemy, and they had powerful political and economic reasons for insisting that Earth was the center of creation. Galileo not only challenged this thinking but also had the audacity to write in Italian rather than scholarly Latin, and to lecture publicly on those topics. For him, there was no contradiction between the authority of the Church in matters of religion and morality, and the authority of nature (revealed by experiments) in matters of science. It was primarily because of Galileo and his "dangerous" opinions that, in 1616, the Church issued a prohibition decree stating that the Copernican doctrine was "false and absurd" and not to be held or defended.

Galileo's Astronomical Observations

It is not certain who first conceived of the idea of combining two or more pieces of glass to produce an instrument that enlarged images of distant objects, making them appear nearer. The first such "spyglasses" (now called *telescopes*) that attracted much notice were made in 1608 by the Dutch spectacle maker Hans Lippershey (1570–1619). Galileo heard of the discovery and, without ever having seen an assembled telescope, constructed one of his own with a three-power magnification (3×), which made distant objects appear three times nearer and larger (Figure 2.20).



Figure 2.20 : Telescope Used by Galileo. The telescope has a wooden tube covered with paper and a lens 26 millimeters across.

On August 25, 1609, Galileo demonstrated a telescope with a magnification of 9× to government officials of the city-state of Venice. By a magnification of 9×, we mean the linear dimensions of the objects being viewed appeared nine times larger or, alternatively, the objects appeared nine times closer than they really were. There were obvious military advantages associated with a device for seeing distant objects. For his invention, Galileo's salary was nearly doubled, and he was granted lifetime tenure as a professor. (His university colleagues were outraged, particularly because the invention was not even original.)

Others had used the telescope before Galileo to observe things on Earth. But in a flash of insight that changed the history of astronomy, Galileo realized that he could turn the power of the telescope toward the heavens. Before using his telescope for astronomical observations, Galileo had to devise a stable mount and improve the optics. He increased the magnification to 30×. Galileo also needed to acquire confidence in the telescope.

At that time, human eyes were believed to be the final arbiter of truth about size, shape, and color. Lenses, mirrors, and prisms were known to distort distant images by enlarging, reducing, or inverting them, or spreading the light into a spectrum (rainbow of colors). Galileo undertook repeated experiments to convince himself that what he saw through the telescope was identical to what he saw up close. Only then could he begin to believe that the miraculous phenomena the telescope revealed in the heavens were real.

Beginning his astronomical work late in 1609, Galileo found that many stars too faint to be seen with the unaided eye became visible with his telescope. In particular, he found that some nebulous blurs resolved into many stars, and that the Milky Way—the strip of whiteness across the night sky—was also made up of a multitude of individual stars.

Examining the planets, Galileo found four moons revolving about Jupiter in times ranging from just under 2 days to about 17 days. This discovery was particularly important because it showed that not everything has to revolve around Earth. Furthermore, it demonstrated that there could be centers of motion that are themselves in motion. Defenders of the geocentric view had argued that if Earth was in motion, then the Moon would be left behind because it could hardly keep up with a rapidly moving planet. Yet, here were Jupiter's moons doing exactly that. (To recognize this discovery and honor his work, NASA named a spacecraft that explored the Jupiter system Galileo.)

With his telescope, Galileo was able to carry out the test of the Copernican theory mentioned earlier, based on the phases of Venus. Within a few months, he had found that Venus goes through phases like the Moon, showing that it must revolve about the Sun, so that we see different parts of its daylight side at different times (see Figure 2.18.) These observations could not be reconciled with Ptolemy's model, in which Venus circled about Earth. In Ptolemy's model, Venus could also show phases, but they were the wrong phases in the wrong order from what Galileo observed.

Galileo also observed the Moon and saw craters, mountain ranges, valleys, and flat, dark areas that he thought might be water. These discoveries showed that the Moon might be not so dissimilar to Earth—suggesting that Earth, too, could belong to the realm of celestial bodies.

Link to Learning

For more information about the life and work of Galileo, see the Galileo Project at Rice University.

After Galileo's work, it became increasingly difficult to deny the Copernican view, and Earth was slowly dethroned from its central

position in the universe and given its rightful place as one of the planets attending the Sun. Initially, however, Galileo met with a great deal of opposition. The Roman Catholic Church, still reeling from the Protestant Reformation, was looking to assert its authority and chose to make an example of Galileo. He had to appear before the Inquisition to answer charges that his work was heretical, and he was ultimately condemned to house arrest. His books were on the Church's forbidden list until 1836, although in countries where the Roman Catholic Church held less sway, they were widely read and discussed. Not until 1992 did the Catholic Church admit publicly that it had erred in the matter of censoring Galileo's ideas.

The new ideas of Copernicus and Galileo began a revolution in our conception of the cosmos. It eventually became evident that the universe is a vast place and that Earth's role in it is relatively unimportant. The idea that Earth moves around the Sun like the other planets raised the possibility that they might be worlds themselves, perhaps even supporting life. As Earth was demoted from its position at the center of the universe, so, too, was humanity. The universe, despite what we may wish, does not revolve around us.

Most of us take these things for granted today, but four centuries ago such concepts were frightening and heretical for some, immensely stimulating for others. The pioneers of the Renaissance started the European world along the path toward science and technology that we still tread today. For them, nature was rational and ultimately knowable, and experiments and observations provided the means to reveal its secrets.

Seeing for Yourself: Observing the Planets

At most any time of the night, and at any season, you can spot one or more bright planets in the sky. All five of the planets known to the ancients—Mercury, Venus, Mars, Jupiter, and Saturn—are more prominent than any but the brightest stars, and they can be seen even from urban locations if you know where and when to look. One way to tell planets from bright stars is that planets twinkle less.

Venus, which stays close to the Sun from our perspective, appears either as an "evening star" in the west after sunset or as a "morning star" in the east before sunrise. It is the brightest object in the sky after the Sun and Moon. It far outshines any real star, and under the most favorable circumstances, it can even cast a visible shadow. Some young military recruits have tried to shoot Venus down as an approaching enemy craft or UFO.

Mars, with its distinctive red color, can be nearly as bright as Venus is when close to Earth, but normally it remains much less conspicuous. Jupiter is most often the second-brightest planet, approximately equaling in brilliance the brightest stars. Saturn is dimmer, and it varies considerably in brightness, depending on whether its large rings are seen nearly edge-on (faint) or more widely opened (bright).

Mercury is quite bright, but few people ever notice it because it never moves very far from the Sun (it's never more than 28° away in the sky) and is always seen against bright twilight skies.

True to their name, the planets "wander" against the background of the "fixed" stars. Although their apparent motions are complex, they reflect an underlying order upon which the heliocentric model of the solar system, as described in this chapter, was based. The positions of the planets are often listed in newspapers (sometimes on the weather page), and clear maps and guides to their locations can be found each month in such magazines as *Sky & Telescope* and *Astronomy* (available at most libraries and online). There are also a number of computer programs and phone and tablet apps that allow you to display where the planets are on any night.

This page titled 4.1: The Birth of Modern Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 2.5: The Birth of Modern Astronomy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

4.2: The Laws of Planetary Motion

Learning Objectives

By the end of this section, you will be able to:

- Describe how Tycho Brahe and Johannes Kepler contributed to our understanding of how planets move around the Sun
- Explain Kepler's three laws of planetary motion

At about the time that Galileo was beginning his experiments with falling bodies, the efforts of two other scientists dramatically advanced our understanding of the motions of the planets. These two astronomers were the observer Tycho Brahe and the mathematician Johannes Kepler. Together, they placed the speculations of Copernicus on a sound mathematical basis and paved the way for the work of Isaac Newton in the next century.

Tycho Brahe's Observatory

Three years after the publication of Copernicus' *De Revolutionibus*, Tycho Brahe was born to a family of Danish nobility. He developed an early interest in astronomy and, as a young man, made significant astronomical observations. Among these was a careful study of what we now know was an exploding star that flared up to great brilliance in the night sky. His growing reputation gained him the patronage of the Danish King Frederick II, and at the age of 30, Brahe was able to establish a fine astronomical observatory on the North Sea island of Hven (Figure 3.2). Brahe was the last and greatest of the pre-telescopic observers in Europe.



(a) (b)

Figure 3.2 : Tycho Brahe (1546–1601) and Johannes Kepler (1571–1630). (a) A stylized engraving shows Tycho Brahe using his instruments to measure the altitude of celestial objects above the horizon. The large curved instrument in the foreground allowed him to measure precise angles in the sky. Note that the scene includes hints of the grandeur of Brahe's observatory at Hven. (b) Kepler was a German mathematician and astronomer. His discovery of the basic laws that describe planetary motion placed the heliocentric cosmology of Copernicus on a firm mathematical basis.

At Hven, Brahe made a continuous record of the positions of the Sun, Moon, and planets for almost 20 years. His extensive and precise observations enabled him to note that the positions of the planets varied from those given in published tables, which were based on the work of Ptolemy. These data were extremely valuable, but Brahe didn't have the ability to analyze them and develop a better model than what Ptolemy had published. He was further inhibited because he was an extravagant and cantankerous fellow, and he accumulated enemies among government officials. When his patron, Frederick II, died in 1597, Brahe lost his political base and decided to leave Denmark. He took up residence in Prague, where he became court astronomer to Emperor Rudolf of

Bohemia. There, in the year before his death, Brahe found a most able young mathematician, Johannes Kepler, to assist him in analyzing his extensive planetary data.

Johannes Kepler

Johannes Kepler was born into a poor family in the German province of Württemberg and lived much of his life amid the turmoil of the Thirty Years' War (see Figure 3.2). He attended university at Tubingen and studied for a theological career. There, he learned the principles of the Copernican system and became converted to the heliocentric hypothesis. Eventually, Kepler went to Prague to serve as an assistant to Brahe, who set him to work trying to find a satisfactory theory of planetary motion—one that was compatible with the long series of observations made at Hven. Brahe was reluctant to provide Kepler with much material at any one time for fear that Kepler would discover the secrets of the universal motion by himself, thereby robbing Brahe of some of the glory. Only after Brahe's death in 1601 did Kepler get full possession of the priceless records. Their study occupied most of Kepler's time for more than 20 years.

Through his analysis of the motions of the planets, Kepler developed a series of principles, now known as *Kepler's three laws,* which described the behavior of planets based on their paths through space. The first two laws of planetary motion were published in 1609 in *The New Astronomy*. Their discovery was a profound step in the development of modern science.

The First Two Laws of Planetary Motion

The path of an object through space is called its orbit. Kepler initially assumed that the orbits of planets were circles, but doing so did not allow him to find orbits that were consistent with Brahe's observations. Working with the data for Mars, he eventually discovered that the orbit of that planet had the shape of a somewhat flattened circle, or ellipse. Next to the circle, the ellipse is the simplest kind of closed curve, belonging to a family of curves known as *conic sections* (Figure 3.3).

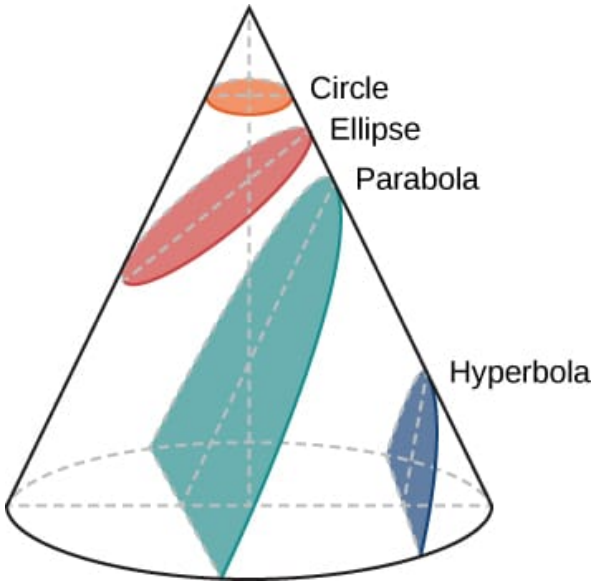


Figure 3.3 : Conic Sections. The circle, ellipse, parabola, and hyperbola are all formed by the intersection of a plane with a cone. This is why such curves are called conic sections.

You might recall from math classes that in a circle, the center is a special point. The distance from the center to anywhere on the circle is exactly the same. In an ellipse, the sum of the distance from two special points inside the ellipse to any point on the ellipse is always the same. These two points inside the ellipse are called its foci (singular: focus), a word invented for this purpose by Kepler.

This property suggests a simple way to draw an ellipse (Figure 3.4). We wrap the ends of a loop of string around two tacks pushed through a sheet of paper into a drawing board, so that the string is slack. If we push a pencil against the string, making the string taut, and then slide the pencil against the string all around the tacks, the curve that results is an ellipse. At any point where the pencil may be, the sum of the distances from the pencil to the two tacks is a constant length—the length of the string. The tacks are at the two foci of the ellipse.

The widest diameter of the ellipse is called its major axis. Half this distance—that is, the distance from the center of the ellipse to one end—is the semimajor axis, which is usually used to specify the size of the ellipse. For example, the semimajor axis of the orbit of Mars, which is also the planet's average distance from the Sun, is 228 million kilometers.

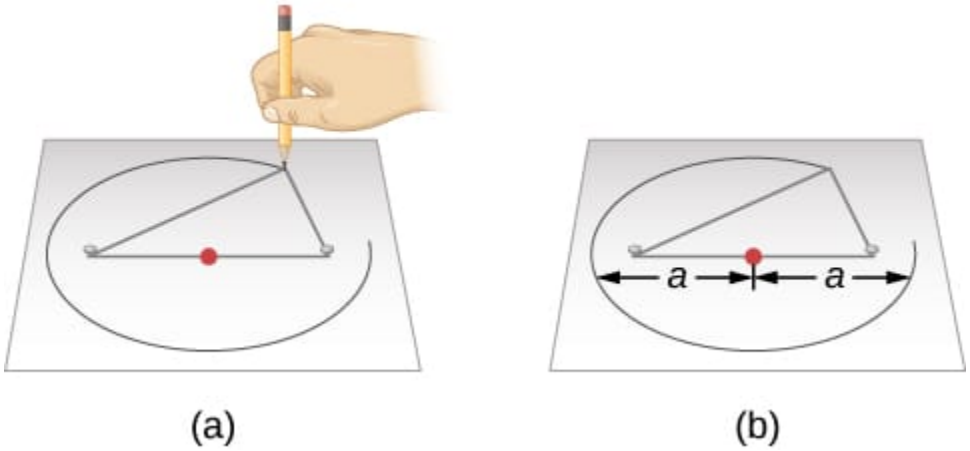


Figure 3.4 : Drawing an Ellipse. (a) We can construct an ellipse by pushing two tacks (the white objects) into a piece of paper on a drawing board, and then looping a string around the tacks. Each tack represents a focus of the ellipse, with one of the tacks being the Sun. Stretch the string tight using a pencil, and then move the pencil around the tacks. The length of the string remains the same, so that the sum of the distances from any point on the ellipse to the foci is always constant. (b) In this illustration, each semimajor axis is denoted by *a*. The distance *2a* is called the major axis of the ellipse.

The shape (roundness) of an ellipse depends on how close together the two foci are, compared with the major axis. The ratio of the distance between the foci to the length of the major axis is called the eccentricity of the ellipse.

If the foci (or tacks) are moved to the same location, then the distance between the foci would be zero. This means that the eccentricity is zero and the ellipse is just a circle; thus, a circle can be called an ellipse of zero eccentricity. In a circle, the semimajor axis would be the radius.

Next, we can make ellipses of various elongations (or extended lengths) by varying the spacing of the tacks (as long as they are not farther apart than the length of the string). The greater the eccentricity, the more elongated is the ellipse, up to a maximum eccentricity of 1.0, when the ellipse becomes "flat," the other extreme from a circle.

The size and shape of an ellipse are completely specified by its semimajor axis and its eccentricity. Using Brahe's data, Kepler found that Mars has an elliptical orbit, with the Sun at one focus (the other focus is empty). The eccentricity of the orbit of Mars is only about 0.1; its orbit, drawn to scale, would be practically indistinguishable from a circle, but the difference turned out to be critical for understanding planetary motions.

Kepler generalized this result in his first law and said that *the orbits of all the planets are ellipses*. Here was a decisive moment in the history of human thought: it was not necessary to have only circles in order to have an acceptable cosmos. The universe could be a bit more complex than the Greek philosophers had wanted it to be.

Kepler's second law deals with the speed with which each planet moves along its ellipse, also known as its orbital speed. Working with Brahe's observations of Mars, Kepler discovered that the planet speeds up as it comes closer to the Sun and slows down as it pulls away from the Sun. He expressed the precise form of this relationship by imagining that the Sun and Mars are connected by a straight, elastic line. When Mars is closer to the Sun (positions 1 and 2 in Figure 3.5), the elastic line is not stretched as much, and the planet moves rapidly. Farther from the Sun, as in positions 3 and 4, the line is stretched a lot, and the planet does not move so fast. As Mars travels in its elliptical orbit around the Sun, the elastic line sweeps out areas of the ellipse as it moves (the colored regions in our figure). Kepler found that in equal intervals of time (t), the areas swept out in space by this imaginary line are always equal; that is, the area of the region B from 1 to 2 is the same as that of region A from 3 to 4.

If a planet moves in a circular orbit, the elastic line is always stretched the same amount and the planet moves at a constant speed around its orbit. But, as Kepler discovered, in most orbits that speed of a planet orbiting its star (or moon orbiting its planet) tends to vary because the orbit is elliptical.

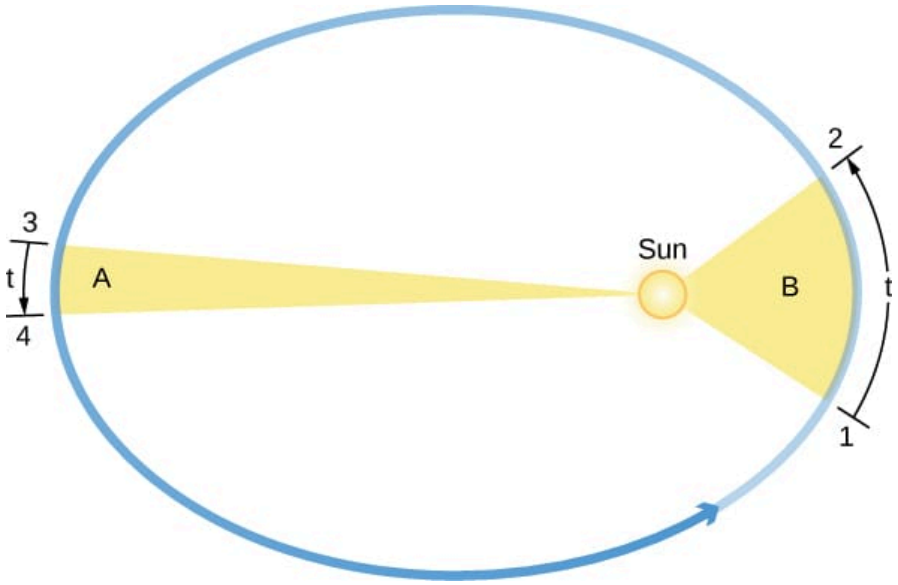


Figure 3.5 : Kepler's Second Law: The Law of Equal Areas. The orbital speed of a planet traveling around the Sun (the circular object inside the ellipse) varies in such a way that in equal intervals of time (t), a line between the Sun and a planet sweeps out equal areas (A and B). Note that the eccentricities of the planets' orbits in our solar system are substantially less than shown here.

Link to Learning

The Kepler's Second Law demonstrator from CCNY's ScienceSims project shows how an orbiting planet sweeps out the same area in the same time.

Kepler's Third Law

Kepler's first two laws of planetary motion describe the shape of a planet's orbit and allow us to calculate the speed of its motion at any point in the orbit. Kepler was pleased to have discovered such fundamental rules, but they did not satisfy his quest to fully understand planetary motions. He wanted to know why the orbits of the planets were spaced as they are and to find a mathematical pattern in their movements—a "harmony of the spheres" as he called it. For many years he worked to discover mathematical relationships governing planetary spacing and the time each planet took to go around the Sun.

In 1619, Kepler discovered a basic relationship to relate the planets' orbits to their relative distances from the Sun. We define a planet's orbital period, (***P***), as the time it takes a planet to travel once around the Sun. Also, recall that a planet's semimajor axis, *a,* is equal to its average distance from the Sun. The relationship, now known as *Kepler's third law*, says that a planet's orbital period squared is proportional to the semimajor axis of its orbit cubed, or

$$P^2 \propto a^3$$

When *P* (the orbital period) is measured in years, and *a* is expressed in a quantity known as an astronomical unit (AU), the two sides of the formula are not only proportional but equal. One AU is the average distance between Earth and the Sun and is approximately equal to 1.5×10^8 kilometers. In these units,

$$P^2 = a^3$$

Kepler's third law applies to all objects orbiting the Sun, including Earth, and provides a means for calculating their relative distances from the Sun from the time they take to orbit. Let's look at a specific example to illustrate how useful Kepler's third law is.

For instance, suppose you time how long Mars takes to go around the Sun (in Earth years). Kepler's third law can then be used to calculate Mars' average distance from the Sun. Mars' orbital period (1.88 Earth years) squared, or P^2, is $1.88^2 = 3.53$, and according to the equation for Kepler's third law, this equals the cube of its semimajor axis, or a^3. So what number must be cubed to give 3.53? The answer is $1.52 (\text{since} 1.52 \times 1.52 \times 1.52 = 3.53)$. Thus, Mars' semimajor axis in astronomical units must be 1.52 AU.

In other words, to go around the Sun in a little less than two years, Mars must be about 50% (half again) as far from the Sun as Earth is.

Example 3.1: Calculating Periods

Imagine an object is traveling around the Sun. What would be the orbital period of the object if its orbit has a semimajor axis of 50 AU?

Solution

From Kepler's third law, we know that (when we use units of years and AU)

$$P^2 = a^3$$

If the object's orbit has a semimajor axis of 50 AU (a = 50), we can cube 50 and then take the square root of the result to get P:

$$\begin{aligned} P &= \sqrt{a^3} \\ P &= \sqrt{50 \times 50 \times 50} = \sqrt{125{,}000} = 353.6 \text{ years} \end{aligned}$$

Therefore, the orbital period of the object is about 350 years. This would place our hypothetical object beyond the orbit of Pluto.

Exercise 4.2.1

What would be the orbital period of an asteroid (a rocky chunk between Mars and Jupiter) with a semimajor axis of 3 AU?

Answer available at the end of this section.

Kepler's three laws of planetary motion can be summarized as follows:

- Kepler's first law: Each planet moves around the Sun in an orbit that is an ellipse, with the Sun at one focus of the ellipse.
- Kepler's second law: The straight line joining a planet and the Sun sweeps out equal areas in space in equal intervals of time.
- Kepler's third law: The square of a planet's orbital period is directly proportional to the cube of the semimajor axis of its orbit.

Kepler's three laws provide a precise geometric description of planetary motion within the framework of the Copernican system. With these tools, it was possible to calculate planetary positions with greatly improved precision. Still, Kepler's laws are purely descriptive: they do not help us understand what forces of nature constrain the planets to follow this particular set of rules. That step was left to Isaac Newton.

Example 3.2: Applying Kepler's Third Law

Using the orbital periods and semimajor axes for Venus and Earth that are provided here, calculate p^2 and a^3, and verify that they obey Kepler's third law. Venus' orbital period is 0.62 year, and its semimajor axis is 0.72 AU. Earth's orbital period is 1.00 year, and its semimajor axis is 1.00 AU.

Solution

We can use the equation for Kepler's third law, $P^2 \propto a^3$. For Venus, $P^2 = 0.62 \times 0.62 = 0.38$ and $a^3 = 0.72 \times 0.72 \times 0.72 = .037$ (rounding numbers sometimes causes minor discrepancies like this). The square of the orbital period (0.38) approximates the cube of the semimajor axis (0.37). Therefore, Venus obeys Kepler's third law. For Earth, $P^2 = 1.00 \times 1.00 = 1.00$ and $a^3 = 1.00 \times 1.00 \times 1.00 = 1.00$. The square of the orbital period (1.00) approximates (in this case, equals) the cube of the semimajor axis (1.00). Therefore, Earth obeys Kepler's third law.

Exercise 4.2.1

Using the orbital periods and semimajor axes for Saturn and Jupiter that are provided here, calculate P^2 and a^3, and verify that they obey Kepler's third law. Saturn's orbital period is 29.46 years, and its semimajor axis is 9.54 AU. Jupiter's orbital period is 11.86 years, and its semimajor axis is 5.20 AU.

Answer available at the end of this section.

Link to Learning

In honor of the scientist who first devised the laws that govern the motions of planets, the team that built the first spacecraft to search for planets orbiting other stars decided to name the probe "Kepler." Visit NASA's Kepler website to learn more about Johannes Kepler's life and his laws of planetary motion. NASA's Kepler website and follow the links that interest you.

Answers

Exercise 4.2.1

$P = \sqrt{3 \times 3 \times 3} = \sqrt{27} = 5.2 \text{ years}$

Exercise 4.2.1

For Saturn, $P^2 = 29.46 \times 29.46 = 867.9$ and $a^3 = 9.54 \times 9.54 \times 9.54 = 868.3$. The square of the orbital period (867.9) approximates the cube of the semimajor axis (868.3). Therefore, Saturn obeys Kepler's third law.

This page titled 4.2: The Laws of Planetary Motion is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.2: The Laws of Planetary Motion by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5: PHYSICS AND GRAVITY

5.1: Newton’s Great Synthesis
5.2: Newton’s Universal Law of Gravitation
5.3: Orbits in the Solar System
5.4: Motions of Satellites and Spacecraft
5.5: Gravity with More Than Two Bodies
5.6: Key Terms

This page titled 5: Physics and Gravity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

5.1: Newton's Great Synthesis

Learning Objectives

By the end of this section, you will be able to:

- Describe Newton's three laws of motion
- Explain how Newton's three laws of motion relate to momentum
- Define mass, volume, and density and how they differ
- Define angular momentum

It was the genius of Isaac Newton that found a conceptual framework that completely explained the observations and rules assembled by Galileo, Brahe, Kepler, and others. Newton was born in Lincolnshire, England, in the year after Galileo's death (Figure 3.6). Against the advice of his mother, who wanted him to stay home and help with the family farm, he entered Trinity College at Cambridge in 1661 and eight years later was appointed professor of mathematics. Among Newton's contemporaries in England were architect Christopher Wren, authors Aphra Behn and Daniel Defoe, and composer G. F. Handel.

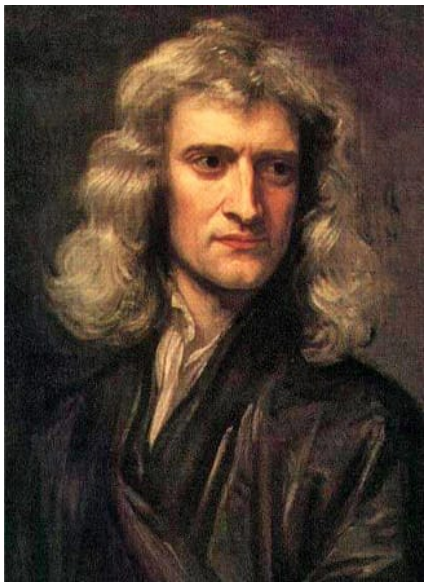


Figure 3.6 : Isaac Newton (1643–1727), 1689 Portrait by Sir Godfrey Kneller. Isaac Newton's work on the laws of motion, gravity, optics, and mathematics laid the foundations for much of physical science.

Newton's Laws of Motion

As a young man in college, Newton became interested in natural philosophy, as science was then called. He worked out some of his first ideas on machines and optics during the plague years of 1665 and 1666, when students were sent home from college. Newton, a moody and often difficult man, continued to work on his ideas in private, even inventing new mathematical tools to help him deal with the complexities involved. Eventually, his friend Edmund Halley (profiled in Comets and Asteroids: Debris of the Solar System) prevailed on him to collect and publish the results of his remarkable investigations on motion and gravity. The result was a volume that set out the underlying system of the physical world, *Philosophiae Naturalis Principia Mathematica*. The *Principia*, as the book is generally known, was published at Halley's expense in 1687.

At the very beginning of the *Principia*, Newton proposes three laws that would govern the motions of all objects:

- Newton's first law: Every object will continue to be in a state of rest or move at a constant speed in a straight line unless it is compelled to change by an outside force.
- Newton's second law: The change of motion of a body is proportional to and in the direction of the force acting on it.

- Newton's third law: For every action there is an equal and opposite reaction (*or:* the mutual actions of two bodies upon each other are always equal and act in opposite directions).

In the original Latin, the three laws contain only 59 words, but those few words set the stage for modern science. Let us examine them more carefully.

Interpretation of Newton's Laws

Newton's first law is a restatement of one of Galileo's discoveries, called the *conservation of momentum*. The law states that in the absence of any outside influence, there is a measure of a body's motion, called its momentum, that remains unchanged. You may have heard the term momentum used in everyday expressions, such as "This bill in Congress has a lot of momentum; it's going to be hard to stop."

Newton's first law is sometimes called the *law of inertia*, where inertia is the tendency of objects (and legislatures) to keep doing what they are already doing. In other words, a stationary object stays put, and a moving object keeps moving unless some force intervenes.

Let's define the precise meaning of momentum—it depends on three factors: (1) speed—how fast a body moves (zero if it is stationary), (2) the direction of its motion, and (3) its mass—a measure of the amount of matter in a body, which we will discuss later. Scientists use the term velocity to describe the speed and direction of motion. For example, 20 kilometers per hour due south is velocity, whereas 20 kilometers per hour just by itself is speed. Momentum then can be defined as an object's mass times its velocity.

It's not so easy to see this rule in action in the everyday world because of the many forces acting on a body at any one time. One important force is friction, which generally slows things down. If you roll a ball along the sidewalk, it eventually comes to a stop because the sidewalk exerts a rubbing force on the ball. But in the space between the stars, where there is so little matter that friction is insignificant, objects can in fact continue to move (to coast) indefinitely.

The momentum of a body can change only under the action of an outside influence. Newton's second law expresses *force* in terms of its ability to change momentum with time. A force (a push or a pull) has both size and direction. When a force is applied to a body, the momentum changes in the direction of the applied force. This means that a force is required to change either the speed or the direction of a body, or both—that is, to start it moving, to speed it up, to slow it down, to stop it, or to change its direction.

As you learned in Observing the Sky: The Birth of Astronomy, the rate of change in an object's velocity is called *acceleration*. Newton showed that the acceleration of a body was proportional to the force being applied to it. Suppose that after a long period of reading, you push an astronomy book away from you on a long, smooth table. (We use a smooth table so we can ignore friction.) If you push the book steadily, it will continue to speed up as long as you are pushing it. The harder you push the book, the larger its acceleration will be. How much a force will accelerate an object is also determined by the object's mass. If you kept pushing a pen with the same force with which you pushed the textbook, the pen—having less mass—would be accelerated to a greater speed.

Newton's third law is perhaps the most profound of the rules he discovered. Basically, it is a generalization of the first law, but it also gives us a way to define mass. If we consider a system of two or more objects isolated from outside influences, Newton's first law says that the total momentum of the objects should remain constant. Therefore, any change of momentum within the system must be balanced by another change that is equal and opposite so that the momentum of the entire system is not changed.

This means that forces in nature do not occur alone: we find that in each situation there is always a *pair* of forces that are equal to and opposite each other. If a force is exerted on an object, it must be exerted by something else, and the object will exert an equal and opposite force back on that something. We can look at a simple example to demonstrate this.

Suppose that a daredevil astronomy student—and avid skateboarder—wants to jump from his second-story dorm window onto his board below (we don't recommend trying this!). The force pulling him down after jumping (as we will see in the next section) is the force of gravity between him and Earth. Both he and Earth must experience the same total change of momentum because of the influence of these mutual forces. So, both the student and Earth are accelerated by each other's pull. However, the student does much more of the moving. Because Earth has enormously greater mass, it can experience the same change of momentum by accelerating only a very small amount. Things fall toward Earth all the time, but the acceleration of our planet as a result is far too small to be measured.

A more obvious example of the mutual nature of forces between objects is familiar to all who have batted a baseball. The recoil you feel as you swing your bat shows that the ball exerts a force on it during the impact, just as the bat does on the ball. Similarly, when a rifle you are bracing on your shoulder is discharged, the force pushing the bullet out of the muzzle is equal to the force pushing backward upon the gun and your shoulder.

This is the principle behind jet engines and rockets: the force that discharges the exhaust gases from the rear of the rocket is accompanied by the force that pushes the rocket forward. The exhaust gases need not push against air or Earth; a rocket actually operates best in a vacuum (Figure 3.7).



Figure 3.7 : Demonstrating Newton's Third Law. The U.S. Space Shuttle (here launching *Discovery*), powered by three fuel engines burning liquid oxygen and liquid hydrogen, with two solid fuel boosters, demonstrates Newton's third law. (credit: modification of work by NASA)

Link to Learning

For more about Isaac Newton's life and work, check out this timeline page with snapshots from his career, produced by the British Broadcasting Corporation (BBC).

Mass, Volume, and Density

Before we go on to discuss Newton's other work, we want to take a brief look at some terms that will be important to sort out clearly. We begin with mass, which is a measure of the amount of material within an object.

The *volume* of an object is the measure of the physical space it occupies. Volume is measured in cubic units, such as cubic centimeters or liters. The volume is the "size" of an object. A penny and an inflated balloon may both have the same mass, but they have very different volumes. The reason is that they also have very different *densities*, which is a measure of how much mass there is per unit volume. Specifically, density is the mass divided by the volume. Note that in everyday language we often use "heavy" and "light" as indications of density (rather than weight) as, for instance, when we say that iron is heavy or that whipped cream is light.

The units of density that will be used in this book are grams per cubic centimeter (g/cm^3).[1] If a block of some material has a mass of 300 grams and a volume of 100 cm^3, its density is 3 g/cm^3. Familiar materials span a considerable range in density, from artificial materials such as plastic insulating foam (less than 0.1 g/cm^3) to gold (19.3 g/cm^3). Table 3.1 gives the densities of some familiar materials. In the astronomical universe, much more remarkable densities can be found, all the way from a comet's tail (10^{-16} g/cm^3) to a collapsed "star corpse" called a neutron star (10^{15} g/cm^3).

Table 3.1: Densities of Common Materials

Material	Density (g/cm^3)
Gold	19.3
Lead	11.3
Iron	7.9
Earth (bulk)	5.5
Rock (typical)	2.5
Water	1
Wood (typical)	0.8
Insulating foam	0.1
Silica gel	0.02

To sum up, mass is *how much*, volume is *how big*, and density is *how tightly packed*.

Link to Learning

You can play with a simple animation demonstrating the relationship between the concepts of density, mass, and volume, and find out why objects like wood float in water.

Angular Momentum

A concept that is a bit more complex, but important for understanding many astronomical objects, is angular momentum, which is a measure of the rotation of a body as it revolves around some fixed point (an example is a planet orbiting the Sun). The angular momentum of an object is defined as the product of its mass, its velocity, and its distance from the fixed point around which it revolves.

If these three quantities remain constant—that is, if the motion of a particular object takes place at a constant velocity at a fixed distance from the spin center—then the angular momentum is also a constant. Kepler's second law is a consequence of the *conservation of angular momentum*. As a planet approaches the Sun on its elliptical orbit and the distance to the spin center decreases, the planet speeds up to conserve the angular momentum. Similarly, when the planet is farther from the Sun, it moves more slowly.

The conservation of angular momentum is illustrated by figure skaters, who bring their arms and legs in to spin more rapidly, and extend their arms and legs to slow down (Figure 3.8). You can duplicate this yourself on a well-oiled swivel stool by starting yourself spinning slowly with your arms extended and then pulling your arms in. Another example of the conservation of angular momentum is a shrinking cloud of dust or a star collapsing on itself (both are situations that you will learn about as you read on). As material moves to a lesser distance from the spin center, the speed of the material increases to conserve angular momentum.

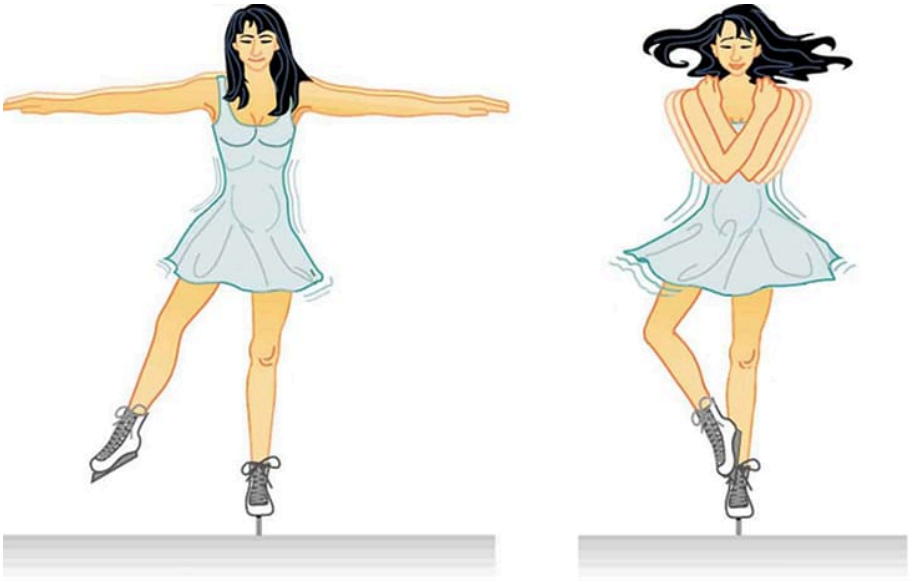


Figure 3.8 : Conservation of Angular Momentum. When a spinning figure skater brings in her arms, their distance from her spin center is smaller, so her speed increases. When her arms are out, their distance from the spin center is greater, so she slows down.

Footnotes

- 1Generally we use standard metric (or SI) units in this book. The proper metric unit of density in that system is kg/m^3. But to most people, g/cm^3 provides a more meaningful unit because the density of water is exactly 1 g/cm^3, and this is useful information for comparison. Density expressed in g/cm^3 is sometimes called specific density or specific weight.

This page titled 5.1: Newton's Great Synthesis is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.3: Newton's Great Synthesis by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5.2: Newton's Universal Law of Gravitation

Learning Objectives

By the end of this section, you will be able to:

- Explain what determines the strength of gravity
- Describe how Newton's universal law of gravitation extends our understanding of Kepler's laws

Newton's laws of motion show that objects at rest will stay at rest and those in motion will continue moving uniformly in a straight line unless acted upon by a force. Thus, it is the *straight line* that defines the most natural state of motion. But the planets move in ellipses, not straight lines; therefore, some force must be bending their paths. That force, Newton proposed, was gravity.

In Newton's time, gravity was something associated with Earth alone. Everyday experience shows us that Earth exerts a gravitational force upon objects at its surface. If you drop something, it accelerates toward Earth as it falls. Newton's insight was that Earth's gravity might extend as far as the Moon and produce the force required to curve the Moon's path from a straight line and keep it in its orbit. He further hypothesized that gravity is not limited to Earth, but that there is a general force of attraction between all material bodies. If so, the attractive force between the Sun and each of the planets could keep them in their orbits. (This may seem part of our everyday thinking today, but it was a remarkable insight in Newton's time.)

Once Newton boldly hypothesized that there was a universal attraction among all bodies everywhere in space, he had to determine the exact nature of the attraction. The precise mathematical description of that gravitational force had to dictate that the planets move exactly as Kepler had described them to (as expressed in Kepler's three laws). Also, that gravitational force had to predict the correct behavior of falling bodies on Earth, as observed by Galileo. How must the force of gravity depend on distance in order for these conditions to be met?

The answer to this question required mathematical tools that had not yet been developed, but this did not deter Isaac Newton, who invented what we today call calculus to deal with this problem. Eventually he was able to conclude that the magnitude of the force of gravity must decrease with increasing distance between the Sun and a planet (or between any two objects) in proportion to the inverse square of their separation. In other words, if a planet were twice as far from the Sun, the force would be $\left(\frac{1}{2}\right)^2$, or $\frac{1}{4}$ as large. Put the planet three times farther away, and the force is $\left(\frac{1}{3}\right)^2$, or $\frac{1}{9}$ as large.

Newton also concluded that the gravitational attraction between two bodies must be proportional to their masses. The more mass an object has, the stronger the pull of its gravitational force. The gravitational attraction between any two objects is therefore given by one of the most famous equations in all of science:

$$F_{\text{gravity}} =$$

where F_{gravity} is the gravitational force between two objects, M_1 and M_2 are the masses of the two objects, and R is their separation. G is a constant number known as the *universal gravitational constant*, and the equation itself symbolically summarizes Newton's *universal law of gravitation*. With such a force and the laws of motion, Newton was able to show mathematically that the only orbits permitted were exactly those described by Kepler's laws.

Newton's universal law of gravitation works for the planets, but is it really universal? The gravitational theory should also predict the observed acceleration of the Moon toward Earth as it orbits Earth, as well as of any object (say, an apple) dropped near Earth's surface. The falling of an apple is something we can measure quite easily, but can we use it to predict the motions of the Moon?

Recall that according to Newton's second law, forces cause acceleration. Newton's universal law of gravitation says that the force acting upon (and therefore the acceleration of) an object toward Earth should be inversely proportional to the square of its distance from the center of Earth. Objects like apples at the surface of Earth, at a distance of one Earth-radius from the center of Earth, are observed to accelerate downward at 9.8 meters per second per second (9.8m/s^2).

It is this force of gravity on the surface of Earth that gives us our sense of *weight*. Unlike your mass, which would remain the same on any planet or moon, your weight depends on the local force of gravity. So you would weigh less on Mars and the Moon than on Earth, even though there is no change in your mass. (Which means you would still have to go easy on the desserts in the college

cafeteria when you got back!)

The Moon is 60 Earth radii away from the center of Earth. If gravity (and the acceleration it causes) gets weaker with distance squared, the acceleration the Moon experiences should be a lot less than for the apple. The acceleration should be $(1/60)^2 = 1/3600$ (or 3600 times less—about $0.00272\mathrm{m/s^2}$). This is precisely the observed acceleration of the Moon in its orbit. (As we shall see, the Moon does not fall *to* Earth with this acceleration, but falls *around* Earth.) Imagine the thrill Newton must have felt to realize he had discovered, and verified, a law that holds for Earth, apples, the Moon, and, as far as he knew, everything in the universe.

Example 3.3: Calculating Weight

By what factor would a person's weight at the surface of Earth change if Earth had its present mass but eight times its present volume?

Solution

With eight times the volume, Earth's radius would double. This means the gravitational force at the surface would reduce by a factor of $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$, so a person would weigh only one-fourth as much.

Exercise 5.2.1

By what factor would a person's weight at the surface of Earth change if Earth had its present size but only one-third its present mass?

Answer available at the end of this section.

Gravity is a "built-in" property of mass. Whenever there are masses in the universe, they will interact via the force of gravitational attraction. The more mass there is, the greater the force of attraction. Here on Earth, the largest concentration of mass is, of course, the planet we stand on, and its pull dominates the gravitational interactions we experience. But everything with mass attracts everything else with mass anywhere in the universe.

Newton's law also implies that gravity never becomes zero. It quickly gets weaker with distance, but it continues to act to some degree no matter how far away you get. The pull of the Sun is stronger at Mercury than at Pluto, but it can be felt far beyond Pluto, where astronomers have good evidence that it continuously makes enormous numbers of smaller icy bodies move around huge orbits. And the Sun's gravitational pull joins with the pull of billions of others stars to create the gravitational pull of our Milky Way Galaxy. That force, in turn, can make other smaller galaxies orbit around the Milky Way, and so on.

Why is it then, you may ask, that the astronauts aboard the Space Shuttle appear to have no gravitational forces acting on them when we see images on television of the astronauts and objects floating in the spacecraft? After all, the astronauts in the shuttle are only a few hundred kilometers above the surface of Earth, which is not a significant distance compared to the size of Earth, so gravity is certainly not a great deal weaker that much farther away. The astronauts feel "weightless" (meaning that they don't feel the gravitational force acting on them) for the same reason that passengers in an elevator whose cable has broken or in an airplane whose engines no longer work feel weightless: they are falling (Figure 3.9).[2]



Figure 3.9 : Astronauts in Free Fall. While in space, astronauts are falling freely, so they experience “weightlessness.” Clockwise from top left: Tracy Caldwell Dyson (NASA), Naoko Yamazaki (JAXA), Dorothy Metcalf-Lindenburger (NASA), and Stephanie Wilson (NASA). (credit: NASA)

When *falling*, they are in free fall and accelerate at the same rate as everything around them, including their spacecraft or a camera with which they are taking photographs of Earth. When doing so, astronauts experience no additional forces and therefore feel “weightless.” Unlike the falling elevator passengers, however, the astronauts are falling *around* Earth, not *to* Earth; as a result they will continue to fall and are said to be “in orbit” around Earth (see the next section for more about orbits).

Orbital Motion and Mass

Kepler’s laws describe the orbits of the objects whose motions are described by Newton’s laws of motion and the law of gravity. Knowing that gravity is the force that attracts planets toward the Sun, however, allowed Newton to rethink Kepler’s third law. Recall that Kepler had found a relationship between the orbital period of a planet’s revolution and its distance from the Sun. But Newton’s formulation introduces the additional factor of the masses of the Sun (M_1) and the planet (M_2), both expressed in units of the Sun’s mass. Newton’s universal law of gravitation can be used to show mathematically that this relationship is actually

$$a^3 = (M_1 + M_2) \times P^2$$

where *a* is the semimajor axis and *P* is the orbital period.

How did Kepler miss this factor? In units of the Sun’s mass, the mass of the Sun is 1, and in units of the Sun’s mass, the mass of a typical planet is a negligibly small factor. This means that the sum of the Sun’s mass and a planet’s mass, $(M_1 + M_2)$, is very, very close to 1. This makes Newton’s formula appear almost the same as Kepler’s; the tiny mass of the planets compared to the Sun is the reason that Kepler did not realize that both masses had to be included in the calculation. There are many situations in astronomy, however, in which we *do* need to include the two mass terms—for example, when two stars or two galaxies orbit each other.

Including the mass term allows us to use this formula in a new way. If we can measure the motions (distances and orbital periods) of objects acting under their mutual gravity, then the formula will permit us to deduce their masses. For example, we can calculate the mass of the Sun by using the distances and orbital periods of the planets, or the mass of Jupiter by noting the motions of its moons.

Indeed, Newton’s reformulation of Kepler’s third law is one of the most powerful concepts in astronomy. Our ability to deduce the masses of objects from their motions is key to understanding the nature and evolution of many astronomical bodies. We will use this law repeatedly throughout this text in calculations that range from the orbits of comets to the interactions of galaxies.

Example 3.4: Calculating the Effects of Gravity

A planet like Earth is found orbiting its star at a distance of 1 AU in 0.71 Earth-year. Can you use Newton's version of Kepler's third law to find the mass of the star? (Remember that compared to the mass of a star, the mass of an earthlike planet can be considered negligible.)

Solution

In the formula $a^3 = (M_1 + M_2) \times P^2$, the factor $M_1 + M_2$ would now be approximately equal to M_1 (the mass of the star), since the planet's mass is so small by comparison. Then the formula becomes $a^3 = M_1 \times P^2$, and we can solve for M[1]:

$$M_1 = \frac{a^3}{P^2}$$

Since $a = 1, a^3 = 1$, so

$$M_1 = \frac{1}{P^2} = \frac{1}{0.71^2} = \frac{1}{0.5} = 2$$

So the mass of the star is twice the mass of our Sun. (Remember that this way of expressing the law has units in terms of Earth and the Sun, so masses are expressed in units of the mass of our Sun.)

Exercise 5.2.1

Suppose a star with twice the mass of our Sun had an earthlike planet that took 4 years to orbit the star. At what distance (semimajor axis) would this planet orbit its star?

Answer available at the end of this section.

Link to Learning

You might like to try a simulation that lets you move the Sun, Earth, Moon, and space station to see the effects of changing their distances on their gravitational forces and orbital paths. You can even turn off gravity and see what happens.

Footnotes

- 2In the film *Apollo 13*, the scenes in which the astronauts were "weightless" were actually filmed in a falling airplane. As you might imagine, the plane fell for only short periods before the engines engaged again.

Answers

Exercise 5.2.1

With one-third its present mass, the gravitational force at the surface would reduce by a factor of $\frac{1}{3}$, so a person would weight only one-third as much.

Exercise 5.2.1

Again, we can neglect the mass of the planet. So $M_1 = 2$ and $P = 4$ years. The formula is $a^3 = M_1 \times P^2$, so $a^3 = 2 \times 4^2 = 2 \times 16 = 32$. So *a* is the cube root of 32. To find this, you can just ask Google, "What is the cube root of 32?" and get the answer 3.2 AU.

This page titled 5.2: Newton's Universal Law of Gravitation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.4: Newton's Universal Law of Gravitation by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5.3: Orbits in the Solar System

Learning Objectives

By the end of this section, you will be able to:

- Compare the orbital characteristics of the planets in the solar system
- Compare the orbital characteristics of asteroids and comets in the solar system

Recall that the path of an object under the influence of gravity through space is called its orbit, whether that object is a spacecraft, planet, star, or galaxy. An orbit, once determined, allows the future positions of the object to be calculated.

Two points in any orbit in our solar system have been given special names. The place where the planet is closest to the Sun (*helios* in Greek) and moves the fastest is called the perihelion of its orbit, and the place where it is farthest away and moves the most slowly is the aphelion. For the Moon or a satellite orbiting Earth (*gee* in Greek), the corresponding terms are perigee and apogee. (In this book, we use the word *moon* for a natural object that goes around a planet and the word satellite to mean a human-made object that revolves around a planet.)

Orbits of the Planets

Today, Newton's work enables us to calculate and predict the orbits of the planets with marvelous precision. We know eight planets, beginning with Mercury closest to the Sun and extending outward to Neptune. The average orbital data for the planets are summarized in Table 3.2. (Ceres is the largest of the *asteroids,* now considered a dwarf planet.)

According to Kepler's laws, Mercury must have the shortest orbital period (88 Earth-days); thus, it has the highest orbital speed, averaging 48 kilometers per second. At the opposite extreme, Neptune has a period of 165 years and an average orbital speed of just 5 kilometers per second.

All the planets have orbits of rather low eccentricity. The most eccentric orbit is that of Mercury (0.21); the rest have eccentricities smaller than 0.1. It is fortunate that among the rest, Mars has an eccentricity greater than that of many of the other planets. Otherwise the pre-telescopic observations of Brahe would not have been sufficient for Kepler to deduce that its orbit had the shape of an ellipse rather than a circle.

The planetary orbits are also confined close to a common plane, which is near the plane of Earth's orbit (called the ecliptic). The strange orbit of the dwarf planet Pluto is inclined about 17° to the ecliptic, and that of the dwarf planet Eris (orbiting even farther away from the Sun than Pluto) by 44°, but all the major planets lie within 10° of the common plane of the solar system.

Link to Learning

JavaLab's solar system simulator allows you to explore the size and speed of the planets' orbits, and view the orbits from different perspectives.

Orbits of Asteroids and Comets

In addition to the eight planets, there are many smaller objects in the solar system. Some of these are moons (natural satellites) that orbit all the planets except Mercury and Venus. In addition, there are two classes of smaller objects in heliocentric orbits: *asteroids* and *comets*. Both asteroids and comets are believed to be small chunks of material left over from the formation process of the solar system.

In general, asteroids have orbits with smaller semimajor axes than do comets (Figure 3.10). The majority of them lie between 2.2 and 3.3 AU, in the region known as the asteroid belt (see Comets and Asteroids: Debris of the Solar System). As you can see in Table 3.2, the asteroid belt (represented by its largest member, Ceres) is in the middle of a gap between the orbits of Mars and

Jupiter. It is because these two planets are so far apart that stable orbits of small bodies can exist in the region between them.

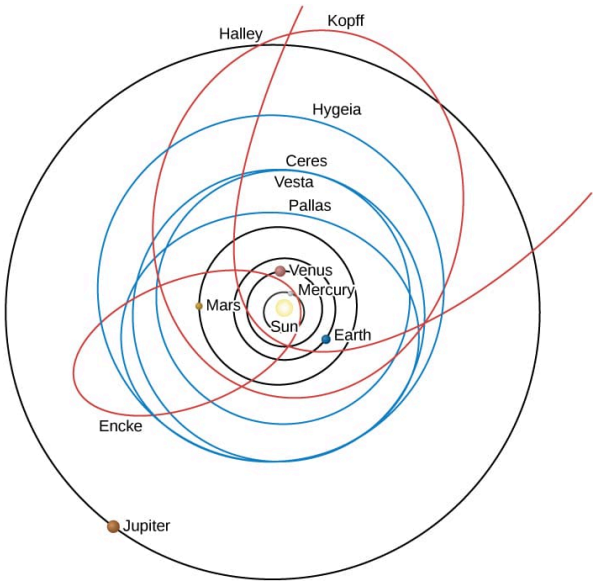


Figure 3.10 : Solar System Orbits. We see the orbits of typical comets and asteroids compared with those of the planets Mercury, Venus, Earth, Mars, and Jupiter (black circles). Shown in red are three comets: Halley, Kopff, and Encke. In blue are the four largest asteroids: Ceres, Pallas, Vesta, and Hygeia.

Table 3.2: Orbital Data for the Planets

Planet	Semimajor Axis (AU)	Period (y)	Eccentricity
Mercury	0.39	0.24	0.21
Venus	0.72	0.6	0.01
Earth	1	1.00	0.02
Mars	1.52	1.88	0.09
(Ceres)	2.77	4.6	0.08
Jupiter	5.20	11.86	0.05
Saturn	9.54	29.46	0.06
Uranus	19.19	84.01	0.05
Neptune	30.06	164.82	0.01

Comets generally have orbits of larger size and greater eccentricity than those of the asteroids. Typically, the eccentricity of their orbits is 0.8 or higher. According to Kepler's second law, therefore, they spend most of their time far from the Sun, moving very slowly. As they approach perihelion, the comets speed up and whip through the inner parts of their orbits more rapidly.

This page titled 5.3: Orbits in the Solar System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.5: Orbits in the Solar System by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5.4: Motions of Satellites and Spacecraft

Learning Objectives

By the end of this section, you will be able to:

- Explain how an object (such as a satellite) can be put into orbit around Earth
- Explain how an object (such as a planetary probe) can escape from orbit

Newton's universal law of gravitation and Kepler's laws describe the motions of Earth satellites and interplanetary spacecraft as well as the planets. Sputnik, the first artificial Earth satellite, was launched by what was then called the Soviet Union on October 4, 1957. Since that time, thousands of satellites have been placed into orbit around Earth, and spacecraft have also orbited the Moon, Venus, Mars, Jupiter, Saturn, and a number of asteroids and comets.

Once an artificial satellite is in orbit, its behavior is no different from that of a natural satellite, such as our Moon. If the satellite is high enough to be free of atmospheric friction, it will remain in orbit forever. However, although there is no difficulty in maintaining a satellite once it is in orbit, a great deal of energy is required to lift the spacecraft off Earth and accelerate it to orbital speed.

To illustrate how a satellite is launched, imagine a gun firing a bullet horizontally from the top of a high mountain, as in Figure 3.11, which has been adapted from a similar diagram by Newton. Imagine, further, that the friction of the air could be removed and that nothing gets in the bullet's way. Then the only force that acts on the bullet after it leaves the muzzle is the gravitational force between the bullet and Earth.

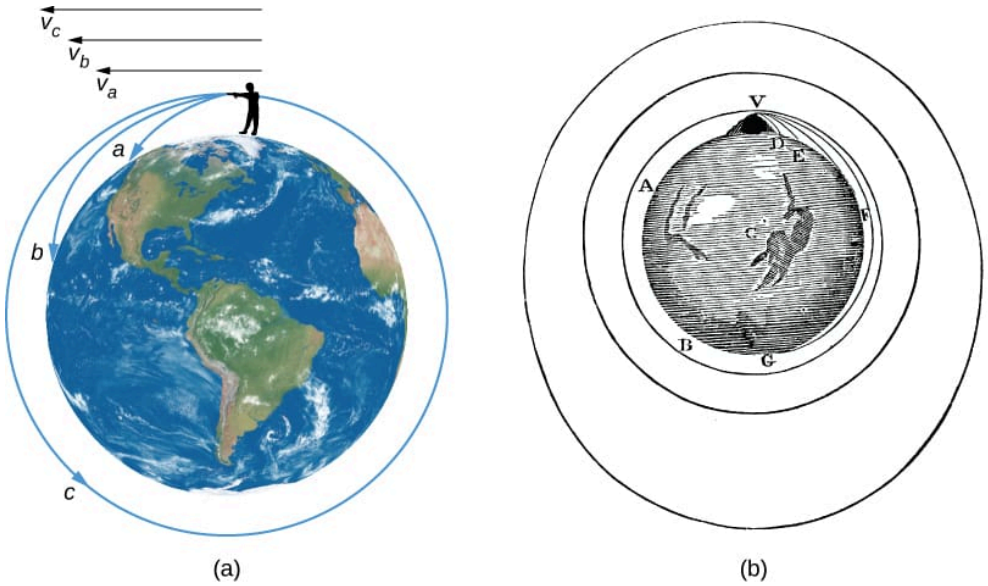


Figure 3.11 : Firing a Bullet into Orbit. (a) For paths *a* and *b*, the velocity is not enough to prevent gravity from pulling the bullet back to Earth; in case *c*, the velocity allows the bullet to fall completely around Earth. (b) This diagram by Newton in his *De Mundi Systemate*, 1731 edition, illustrates the same concept shown in (a).

If the bullet is fired with a velocity we can call v_a, the gravitational force acting upon it pulls it downward toward Earth, where it strikes the ground at point *a*. However, if it is given a higher muzzle velocity, v_b, its higher speed carries it farther before it hits the ground at point *b*.

If our bullet is given a high enough muzzle velocity, v_c, the curved surface of Earth causes the ground to remain the same distance from the bullet so that the bullet falls *around* Earth in a complete circle. The speed needed to do this—called the circular satellite velocity—is about 8 kilometers per second, or about 17,500 miles per hour in more familiar units.

Link to Learning

Use the Newton's Mountain simulator to see for yourself the effects of increasing an object's speed. You can raise the speed until you find the speed that is just fast enough for an object to orbit the Earth, the circular satellite velocity, and also the speed at which an object leaves the Earth forever, or the escape speed.

Each year, more than 50 new satellites are launched into orbit by such nations as Russia, the United States, China, Japan, India, and Israel, as well as by the European Space Agency (ESA), a consortium of European nations (Figure 3.12). Today, these satellites are used for weather tracking, ecology, global positioning systems, communications, and military purposes, to name a few uses. Most satellites are launched into low Earth orbit, since this requires the minimum launch energy. At the orbital speed of 8 kilometers per second, they circle the planet in about 90 minutes. Some of the very low Earth orbits are not indefinitely stable because, as Earth's atmosphere swells from time to time, a frictional drag is generated by the atmosphere on these satellites, eventually leading to a loss of energy and "decay" of the orbit.

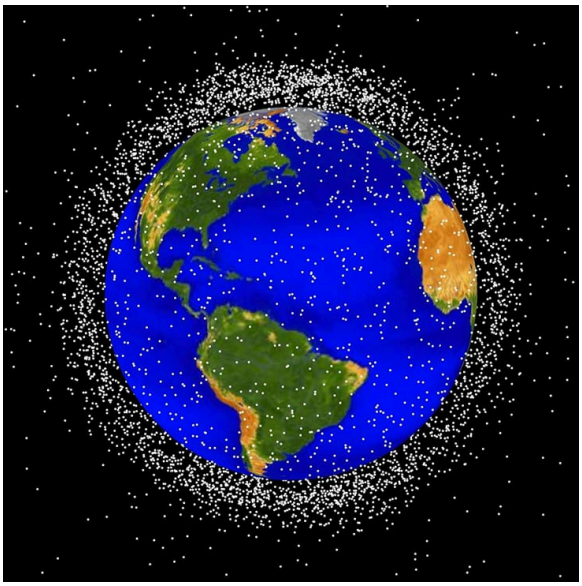


Figure 3.12 : Satellites in Earth Orbit. This figure shows the larger pieces of orbital debris that are being tracked by NASA in Earth's orbit. (credit: NASA/JSC)

Interplanetary Spacecraft

The exploration of the solar system has been carried out largely by robot spacecraft sent to the other planets. To escape Earth, these craft must achieve escape speed, the speed needed to move away from Earth forever, which is about 11 kilometers per second (about 25,000 miles per hour). After escaping Earth, these craft coast to their targets, subject only to minor trajectory adjustments provided by small thruster rockets on board. In interplanetary flight, these spacecraft follow orbits around the Sun that are modified only when they pass near one of the planets.

As it comes close to its target, a spacecraft is deflected by the planet's gravitational force into a modified orbit, either gaining or losing energy in the process. Spacecraft controllers have actually been able to use a planet's gravity to redirect a flyby spacecraft to a second target. For example, Voyager 2 used a series of gravity-assisted encounters to yield successive flybys of Jupiter (1979), Saturn (1980), Uranus (1986), and Neptune (1989). The Galileo spacecraft, launched in 1989, flew past Venus once and Earth twice to gain the energy required to reach its ultimate goal of orbiting Jupiter.

If we wish to orbit a planet, we must slow the spacecraft with a rocket when the spacecraft is near its destination, allowing it to be captured into an elliptical orbit. Additional rocket thrust is required to bring a vehicle down from orbit for a landing on the surface. Finally, if a return trip to Earth is planned, the landed payload must include enough propulsive power to repeat the entire process in reverse.

This page titled 5.4: Motions of Satellites and Spacecraft is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.6: Motions of Satellites and Spacecraft by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5.5: Gravity with More Than Two Bodies

Learning Objectives

By the end of this section, you will be able to:

- Explain how the gravitational interactions of many bodies can causes perturbations in their motions
- Explain how the planet Neptune was discovered

Until now, we have considered the Sun and a planet (or a planet and one of its moons) as nothing more than a pair of bodies revolving around each other. In fact, all the planets exert gravitational forces upon one another as well. These interplanetary attractions cause slight variations from the orbits than would be expected if the gravitational forces between planets were neglected. The motion of a body that is under the gravitational influence of two or more other bodies is very complicated and can be calculated properly only with large computers. Fortunately, astronomers have such computers at their disposal in universities and government research institutes.

The Interactions of Many Bodies

As an example, suppose you have a cluster of a thousand stars all orbiting a common center (such clusters are quite common, as we shall see in Star Clusters). If we know the exact position of each star at any given instant, we can calculate the combined gravitational force of the entire group on any one member of the cluster. Knowing the force on the star in question, we can therefore find how it will accelerate. If we know how it was moving to begin with, we can then calculate how it will move in the next instant of time, thus tracking its motion.

However, the problem is complicated by the fact that the other stars are also moving and thus changing the effect they will have on our star. Therefore, we must simultaneously calculate the acceleration of each star produced by the combination of the gravitational attractions of all the others in order to track the motions of all of them, and hence of any one. Such complex calculations have been carried out with modern computers to track the evolution of hypothetical clusters of stars with up to a million members (Figure 3.13).



Figure 3.13 : Modern Computing Power. These supercomputers at NASA's Ames Research Center are capable of tracking the motions of more than a million objects under their mutual gravitation. (credit: NASA Ames Research Center/Pleiades)

Within the solar system, the problem of computing the orbits of planets and spacecraft is somewhat simpler. We have seen that Kepler's laws, which do not take into account the gravitational effects of the other planets on an orbit, really work quite well.

This is because these additional influences are very small in comparison with the dominant gravitational attraction of the Sun. Under such circumstances, it is possible to treat the effects of other bodies as small perturbations (or disturbances). During the eighteenth and nineteenth centuries, mathematicians developed many elegant techniques for calculating perturbations, permitting them to predict very precisely the positions of the planets. Such calculations eventually led to the prediction and discovery of a new planet in 1846.

The Discovery of Neptune

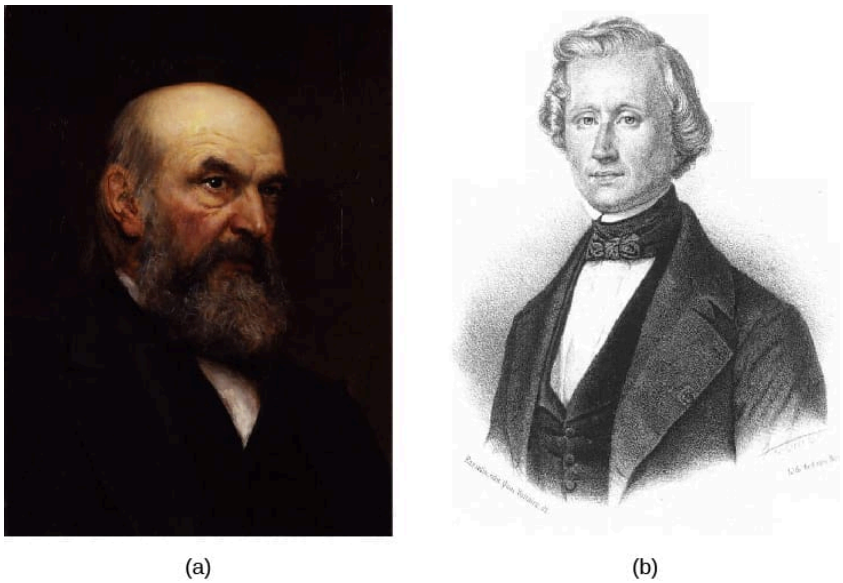
The discovery of the eighth planet, Neptune, was one of the high points in the development of gravitational theory. In 1781, William Herschel, a musician and amateur astronomer, accidentally discovered the seventh planet, Uranus. It happens that Uranus had been observed a century before, but in none of those earlier sightings was it recognized as a planet; rather, it was simply recorded as a star. Herschel's discovery showed that there could be planets in the solar system too dim to be visible to the unaided eye, but ready to be discovered with a telescope if we just knew where to look.

By 1790, an orbit had been calculated for Uranus using observations of its motion in the decade following its discovery. Even after allowance was made for the perturbing effects of Jupiter and Saturn, however, it was found that Uranus did not move on an orbit that exactly fit the earlier observations of it made since 1690. By 1840, the discrepancy between the positions observed for Uranus and those predicted from its computed orbit amounted to about 0.03°—an angle barely discernable to the unaided eye but still larger than the probable errors in the orbital calculations. In other words, Uranus just did not seem to move on the orbit predicted from Newtonian theory.

In 1843, John Couch Adams, a young Englishman who had just completed his studies at Cambridge, began a detailed mathematical analysis of the irregularities in the motion of Uranus to see whether they might be produced by the pull of an unknown planet. He hypothesized a planet more distant from the Sun than Uranus, and then determined the mass and orbit it had to have to account for the departures in Uranus' orbit. In October 1845, Adams delivered his results to George Airy, the British Astronomer Royal, informing him where in the sky to find the new planet. We now know that Adams' predicted position for the new body was correct to within 2°, but for a variety of reasons, Airy did not follow up right away.

Meanwhile, French mathematician Urbain Jean Joseph Le Verrier, unaware of Adams or his work, attacked the same problem and published its solution in June 1846. Airy, noting that Le Verrier's predicted position for the unknown planet agreed to within 1° with that of Adams, suggested to James Challis, Director of the Cambridge Observatory, that he begin a search for the new object. The Cambridge astronomer, having no up-to-date star charts of the Aquarius region of the sky where the planet was predicted to be, proceeded by recording the positions of all the faint stars he could observe with his telescope in that location. It was Challis' plan to repeat such plots at intervals of several days, in the hope that the planet would distinguish itself from a star by its motion. Unfortunately, he was negligent in examining his observations; although he had actually seen the planet, he did not recognize it.

About a month later, Le Verrier suggested to Johann Galle, an astronomer at the Berlin Observatory, that he look for the planet. Galle received Le Verrier's letter on September 23, 1846, and, possessing new charts of the Aquarius region, found and identified the planet that very night. It was less than a degree from the position Le Verrier predicted. The discovery of the eighth planet, now known as Neptune (the Latin name for the god of the sea), was a major triumph for gravitational theory for it dramatically confirmed the generality of Newton's laws. The honor for the discovery is properly shared by the two mathematicians, Adams and Le Verrier (Figure 3.14).



(a) (b)

Figure 3.14 : Mathematicians Who Discovered a Planet. (a) John Couch Adams (1819–1892) and (b) Urbain J. J. Le Verrier (1811–1877) share the credit for discovering the planet Neptune.

We should note that the discovery of Neptune was not a complete surprise to astronomers, who had long suspected the existence of the planet based on the "disobedient" motion of Uranus. On September 10, 1846, two weeks before Neptune was actually found, John Herschel, son of the discoverer of Uranus, remarked in a speech before the British Association, "We see [the new planet] as Columbus saw America from the shores of Spain. Its movements have been felt trembling along the far-reaching line of our analysis with a certainty hardly inferior to ocular demonstration."

This discovery was a major step forward in combining Newtonian theory with painstaking observations. Such work continues in our own times with the discovery of planets around other stars.

Link to Learning

For the fuller story of how Neptune was predicted and found (and the effect of the discovery on the search for Pluto), you can read this page on the mathematical discovery of planets.

Making Connections: Astronomy and the Poets

When Copernicus, Kepler, Galileo, and Newton formulated the fundamental rules that underlie everything in the physical world, they changed much more than the face of science. For some, they gave humanity the courage to let go of old superstitions and see the world as rational and manageable; for others, they upset comforting, ordered ways that had served humanity for centuries, leaving only a dry, "mechanical clockwork" universe in their wake.

Poets of the time reacted to such changes in their work and debated whether the new world picture was an appealing or frightening one. John Donne (1573–1631), in a poem called "Anatomy of the World," laments the passing of the old certainties:

> The new Philosophy [science] calls all in doubt,
> The element of fire is quite put out;
> The Sun is lost, and th' earth, and no man's wit
> Can well direct him where to look for it.

(Here the "element of fire" refers also to the sphere of fire, which medieval thought placed between Earth and the Moon.)

By the next century, however, poets like Alexander Pope were celebrating Newton and the Newtonian world view. Pope's famous couplet, written upon Newton's death, goes

> Nature, and nature's laws lay hid in night.
> God said, Let Newton be! And all was light.

In his 1733 poem, *An Essay on Man*, Pope delights in the complexity of the new views of the world, incomplete though they are:

> Of man, what see we, but his station here,
> From which to reason, to which refer? . . .
> He, who thro' vast immensity can pierce,
> See worlds on worlds compose one universe,
> Observe how system into system runs,
> What other planets circle other suns,
> What vary'd being peoples every star,
> May tell why Heav'n has made us as we are . . .
> All nature is but art, unknown to thee;
> All chance, direction, which thou canst not see;
> All discord, harmony not understood;
> All partial evil, universal good:
> And, in spite of pride, in erring reason's spite,
> One truth is clear, whatever is, is right.

Poets and philosophers continued to debate whether humanity was exalted or debased by the new views of science. The nineteenth-century poet Arthur Hugh Clough (1819–1861) cries out in his poem "The New Sinai":

> And as old from Sinai's top God said that God is one,
> By science strict so speaks He now to tell us, there is None!
> Earth goes by chemic forces; Heaven's a Mécanique Celeste!
> And heart and mind of humankind a watchwork as the rest!

(A "mécanique celeste" is a clockwork model to demonstrate celestial motions.)

The twentieth-century poet Robinson Jeffers (whose brother was an astronomer) saw it differently in a poem called "Star Swirls":

> There is nothing like astronomy to pull the stuff out of man.
> His stupid dreams and red-rooster importance:
> Let him count the star-swirls.

This page titled 5.5: Gravity with More Than Two Bodies is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 3.7: Gravity with More Than Two Bodies by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

5.6: Key Terms

angular momentum
: the measure of the motion of a rotating object in terms of its speed and how widely the object's mass is distributed around its axis

aphelion
: the point in its orbit where a planet (or other orbiting object) is farthest from the Sun

apogee
: the point in its orbit where an Earth satellite is farthest from Earth

asteroid belt
: the region of the solar system between the orbits of Mars and Jupiter in which most asteroids are located; the main belt, where the orbits are generally the most stable, extends from 2.2 to 3.3 AU from the Sun

astronomical unit (AU)
: the unit of length defined as the average distance between Earth and the Sun; this distance is about 1.5×10^8 kilometers

density
: the ratio of the mass of an object to its volume

eccentricity
: in an ellipse, the ratio of the distance between the foci to the major axis

ellipse
: a closed curve for which the sum of the distances from any point on the ellipse to two points inside (called the foci) is always the same

escape speed
: the speed a body must achieve to break away from the gravity of another body

focus
: (plural: foci) one of two fixed points inside an ellipse from which the sum of the distances to any point on the ellipse is constant

gravity
: the mutual attraction of material bodies or particles

Kepler's first law
: each planet moves around the Sun in an orbit that is an ellipse, with the Sun at one focus of the ellipse

Kepler's second law
: the straight line joining a planet and the Sun sweeps out equal areas in space in equal intervals of time

Kepler's third law
: the square of a planet's orbital period is directly proportional to the cube of the semimajor axis of its orbit

major axis
: the maximum diameter of an ellipse

mass
: a measure of the amount of material within an object

momentum
: the measure of the amount of motion of a body; the momentum of a body is the product of its mass and velocity; in the absence of an unbalanced force, momentum is conserved

Newton's first law
: every object will continue to be in a state of rest or move at a constant speed in a straight line unless it is compelled to change by an outside force

Newton's second law
: the change of motion of a body is proportional to and in the direction of the force acting on it

Newton's third law
: for every action there is an equal and opposite reaction (*or:* the mutual actions of two bodies upon each other are always equal and act in opposite directions)

orbit
: the path of an object that is in revolution about another object or point

orbital period (*P*)
: the time it takes an object to travel once around the Sun

orbital speed
: the speed at which an object (usually a planet) orbits around the mass of another object; in the case of a planet, the speed at which each planet moves along its ellipse

perigee
: the point in its orbit where an Earth satellite is closest to Earth

perihelion
: the point in its orbit where a planet (or other orbiting object) is nearest to the Sun

perturbation
: a small disturbing effect on the motion or orbit of a body produced by a third body

satellite
: an object that revolves around a planet

semimajor axis
: half of the major axis of a conic section, such as an ellipse

velocity
: the speed and direction a body is moving—for example, 44 kilometers per second toward the north galactic pole

This page titled 5.6: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

6: TIME

6.1: Earth and Sky
6.2: The Seasons
6.3: Keeping Time
6.4: The Calendar

This page titled 6: Time is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

6.1: Earth and Sky

Learning Objectives

By the end of this section, you will be able to:

- Describe how latitude and longitude are used to map Earth
- Explain how right ascension and declination are used to map the sky

In order to create an accurate map, a mapmaker needs a way to uniquely and simply identify the location of all the major features on the map, such as cities or natural landmarks. Similarly, astronomical mapmakers need a way to uniquely and simply identify the location of stars, galaxies, and other celestial objects. On Earth maps, we divide the surface of Earth into a grid, and each location on that grid can easily be found using its *latitude* and *longitude* coordinate. Astronomers have a similar system for objects on the sky. Learning about these can help us understand the apparent motion of objects in the sky from various places on Earth.

Locating Places on Earth

Let's begin by fixing our position on the surface of planet Earth. As we discussed in Observing the Sky: The Birth of Astronomy, Earth's axis of rotation defines the locations of its North and South Poles and of its equator, halfway between. Two other directions are also defined by Earth's motions: east is the direction toward which Earth rotates, and west is its opposite. At almost any point on Earth, the four directions—north, south, east, and west—are well defined, despite the fact that our planet is round rather than flat. The only exceptions are exactly at the North and South Poles, where the directions east and west are ambiguous (because points exactly at the poles do not turn).

We can use these ideas to define a system of coordinates attached to our planet. Such a system, like the layout of streets and avenues in Manhattan or Salt Lake City, helps us find where we are or want to go. Coordinates on a sphere, however, are a little more complicated than those on a flat surface. We must define circles on the sphere that play the same role as the rectangular grid that you see on city maps.

A great circle is any circle on the surface of a sphere whose center is at the center of the sphere. For example, Earth's equator is a great circle on Earth's surface, halfway between the North and South Poles. We can also imagine a series of great circles that pass through both the North and South Poles. Each of the circles is called a meridian; they are each perpendicular to the equator, crossing it at right angles.

Any point on the surface of Earth will have a meridian passing through it (Figure 4.2). The meridian specifies the east-west location, or longitude, of the place. By international agreement (and it took many meetings for the world's countries to agree), longitude is defined as the number of degrees of arc along the equator between your meridian and the one passing through Greenwich, England, which has been designated as the Prime Meridian. The longitude of the Prime Meridian is defined as 0°.

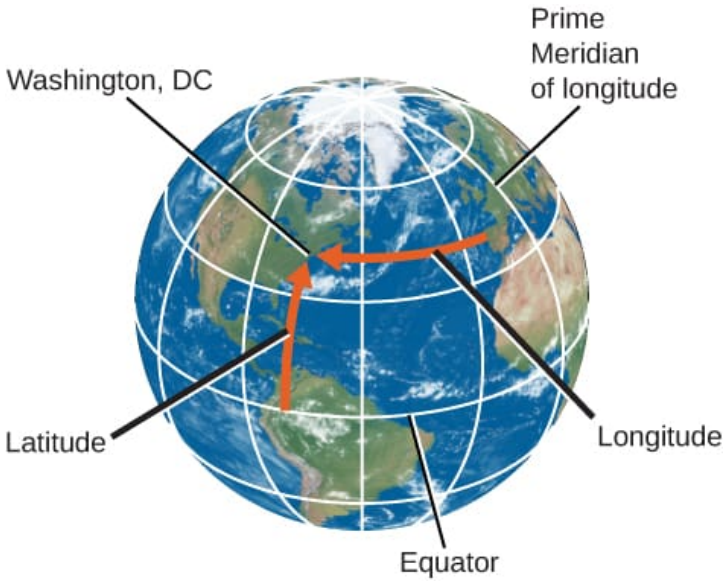


Figure 4.2 : Latitude and Longitude of Washington, DC. We use latitude and longitude to find cities like Washington, DC, on a globe. Latitude is the number of degrees north or south of the equator, and longitude is the number of degrees east or west of the Prime Meridian. Washington, DC's coordinates are 38° N and 77° W.

Why Greenwich, you might ask? Every country wanted 0° longitude to pass through its own capital. Greenwich, the site of the old Royal Observatory (Figure 4.3), was selected because it was between continental Europe and the United States, and because it was the site for much of the development of the method to measure longitude at sea. Longitudes are measured either to the east or to the west of the Greenwich meridian from 0° to 180°. As an example, the longitude of the clock-house benchmark of the U.S. Naval Observatory in Washington, DC, is 77.066° W.



Figure 4.3 : Royal Observatory in Greenwich, England. At the internationally agreed-upon zero point of longitude at the Royal Observatory Greenwich, tourists can stand and straddle the exact line where longitude "begins."(credit left: modification of work by "pdbreen"/Flickr; credit right: modification of work by Ben Sutherland)

Your latitude (or north-south location) is the number of degrees of arc you are away from the equator along your meridian. Latitudes are measured either north or south of the equator from 0° to 90°. (The latitude of the equator is 0°.) As an example, the latitude of the previously mentioned Naval Observatory benchmark is 38.921° N. The latitude of the South Pole is 90° S, and the latitude of the North Pole is 90° N.

Locating Places in the Sky

Positions in the sky are measured in a way that is very similar to the way we measure positions on the surface of Earth. Instead of latitude and longitude, however, astronomers use coordinates called declination and right ascension. To denote positions of objects in the sky, it is often convenient to make use of the fictitious celestial sphere. We saw in Observing the Sky: The Birth of Astronomy that the sky appears to rotate about points above the North and South Poles of Earth—points in the sky called the north celestial pole and the south celestial pole. Halfway between the celestial poles, and thus 90° from each pole, is the *celestial equator,* a great circle on the celestial sphere that is in the same plane as Earth's equator. We can use these markers in the sky to set up a system of celestial coordinates.

Declination on the celestial sphere is measured the same way that latitude is measured on the sphere of Earth: from the celestial equator toward the north (positive) or south (negative). So Polaris, the star near the north celestial pole, has a declination of almost +90°.

Right ascension (RA) is like longitude, except that instead of Greenwich, the arbitrarily chosen point where we start counting is the *vernal equinox*, a point in the sky where the *ecliptic* (the Sun's path) crosses the celestial equator. RA can be expressed either in units of angle (degrees) or in units of time. This is because the celestial sphere appears to turn around Earth once a day as our planet turns on its axis. Thus the 360° of RA that it takes to go once around the celestial sphere can just as well be set equal to 24 hours. Then each 15° of arc is equal to 1 hour of time. For example, the approximate celestial coordinates of the bright star Capella are RA $5h = 75°$ and declination $+50°$.

One way to visualize these circles in the sky is to imagine Earth as a transparent sphere with the terrestrial coordinates (latitude and longitude) painted on it with dark paint. Imagine the celestial sphere around us as a giant ball, painted white on the inside. Then imagine yourself at the center of Earth, with a bright light bulb in the middle, looking out through its transparent surface to the sky. The terrestrial poles, equator, and meridians will be projected as dark shadows on the celestial sphere, giving us the system of coordinates in the sky.

The Turning Earth

Why do many stars rise and set each night? Why, in other words, does the night sky seem to turn? We have seen that the apparent rotation of the celestial sphere could be accounted for either by a daily rotation of the sky around a stationary Earth or by the rotation of Earth itself. Since the seventeenth century, it has been generally accepted that it is Earth that turns, but not until the nineteenth century did the French physicist Jean Foucault provide an unambiguous demonstration of this rotation. In 1851, he suspended a 60-meter pendulum with a mass of about 25 kilograms from the dome of the Pantheon in Paris and started the pendulum swinging evenly. If Earth had not been turning, there would have been no alteration of the pendulum's plane of oscillation, and so it would have continued tracing the same path. Yet after a few minutes Foucault could see that the pendulum's plane of motion was turning. Foucault explained that it was not the pendulum that was shifting, but rather Earth that was turning beneath it (Figure 4.4). You can now find such pendulums in many science centers and planetariums around the world.

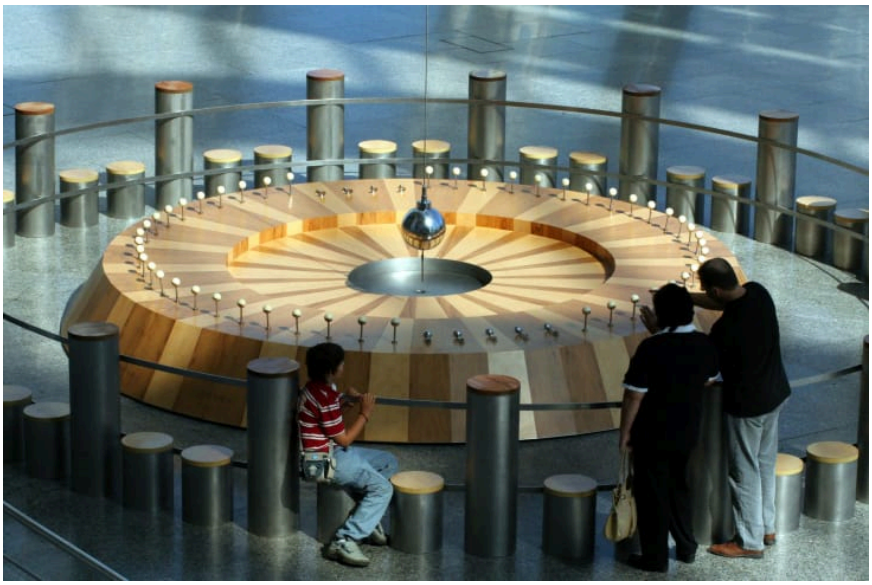


Figure 4.4 : Foucault's Pendulum. As Earth turns, the plane of oscillation of the Foucault pendulum shifts gradually so that over the course of 12 hours, all the targets in the circle at the edge of the wooden platform are knocked over in sequence. (credit: Manuel M. Vicente)

Can you think of other pieces of evidence that indicate that it is Earth and not the sky that is turning? (See Collaborative Group Activity A at the end of this chapter.)

This page titled 6.1: Earth and Sky is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.2: Earth and Sky by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

6.2: The Seasons

Learning Objectives

By the end of this section, you will be able to:

- Describe how the tilt of Earth's axis causes the seasons
- Explain how seasonal differences on Earth vary with latitude

One of the fundamental facts of life at Earth's midlatitudes, where most of this book's readers live, is that there are significant variations in the heat we receive from the Sun during the course of the year. We thus divide the year into *seasons*, each with its different amount of sunlight. The difference between seasons gets more pronounced the farther north or south from the equator we travel, and the seasons in the Southern Hemisphere are the opposite of what we find on the northern half of Earth. With these observed facts in mind, let us ask what causes the seasons.

Many people have believed that the seasons were the result of the changing distance between Earth and the Sun. This sounds reasonable at first: it should be colder when Earth is farther from the Sun. But the facts don't bear out this hypothesis. Although Earth's orbit around the Sun is an ellipse, its distance from the Sun varies by only about 3%. That's not enough to cause significant variations in the Sun's heating. To make matters worse for people in North America who hold this hypothesis, Earth is actually closest to the Sun in January, when the Northern Hemisphere is in the middle of winter. And if distance were the governing factor, why would the two hemispheres have opposite seasons? As we shall show, the seasons are actually caused by the 23.5° tilt of Earth's axis.

The Seasons and Sunshine

Figure 4.5 shows Earth's annual path around the Sun, with Earth's axis tilted by 23.5°. Note that our axis continues to point the same direction in the sky throughout the year. As Earth travels around the Sun, in June the Northern Hemisphere "leans into" the Sun and is more directly illuminated. In December, the situation is reversed: the Southern Hemisphere leans into the Sun, and the Northern Hemisphere leans away. In September and March, Earth leans "sideways"—neither into the Sun nor away from it—so the two hemispheres are equally favored with sunshine.

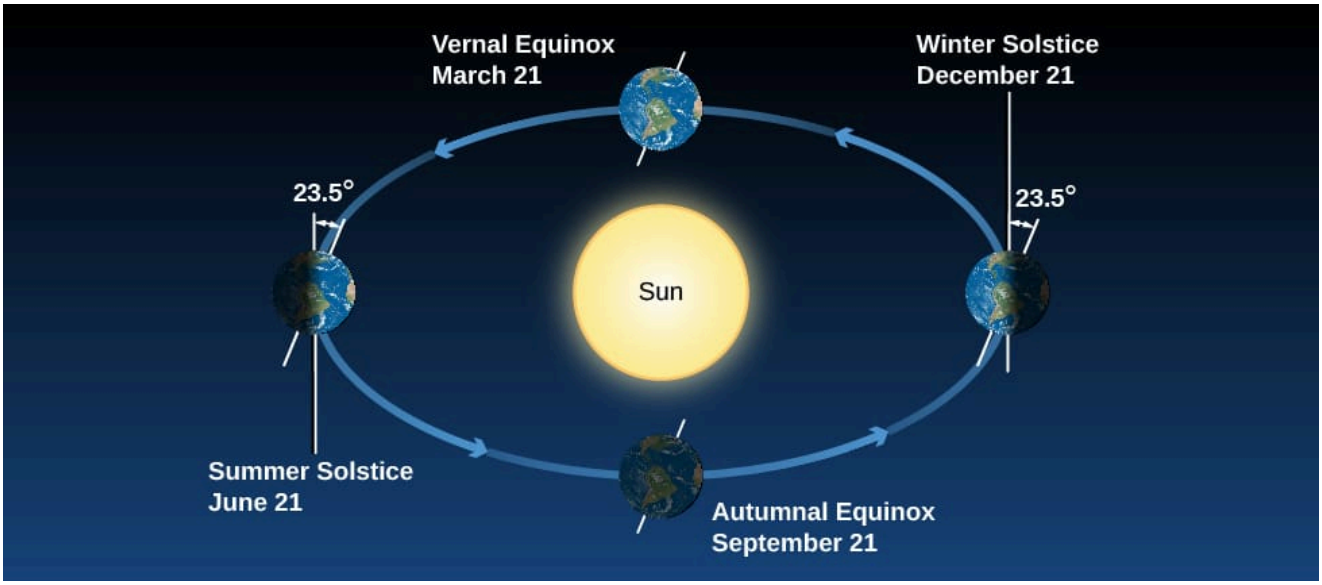


Figure 4.5 : Seasons. We see Earth at different seasons as it circles the Sun. In June, the Northern Hemisphere "leans into" the Sun, and those in the North experience summer and have longer days. In December, during winter in the Northern Hemisphere, the Southern Hemisphere "leans into" the Sun and is illuminated more directly. In spring and autumn, the two hemispheres receive more equal shares of sunlight.[1]

How does the Sun's favoring one hemisphere translate into making it warmer for us down on the surface of Earth? There are two effects we need to consider. When we lean into the Sun, sunlight hits us at a more direct angle and is more effective at heating Earth's surface (Figure 4.6). You can get a similar effect by shining a flashlight onto a wall. If you shine the flashlight straight on, you get an intense spot of light on the wall. But if you hold the flashlight at an angle (if the wall "leans out" of the beam), then the spot of light is more spread out. Like the straight-on light, the sunlight in June is more direct and intense in the Northern Hemisphere, and hence more effective at heating.

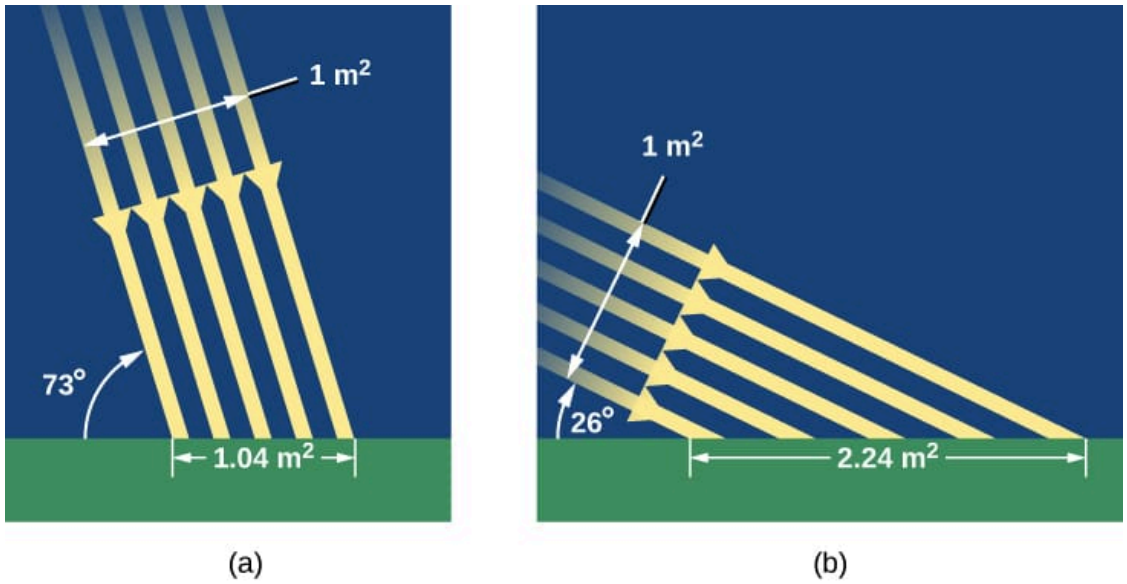


Figure 4.6 : The Sun's Rays in Summer and Winter. (a) In summer, the Sun appears high in the sky and its rays hit Earth more directly, spreading out less. (b) In winter, the Sun is low in the sky and its rays spread out over a much wider area, becoming less effective at heating the ground.

The second effect has to do with the length of time the Sun spends above the horizon (Figure 4.7). Even if you've never thought about astronomy before, we're sure you have observed that the hours of daylight increase in summer and decrease in winter. Let's see why this happens.

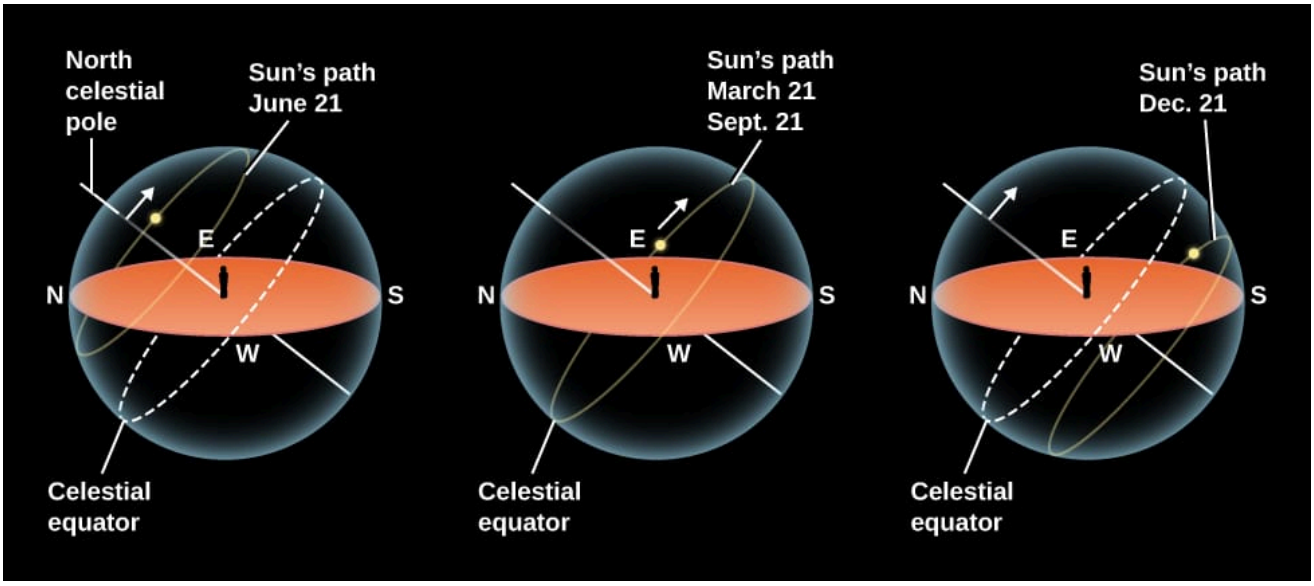


Figure 4.7 : The Sun's Path in the Sky for Different Seasons. On June 21, the Sun rises north of east and sets north of west. For observers in the Northern Hemisphere of Earth, the Sun spends about 15 hours above the horizon in the United States, meaning more hours of daylight. On December 21, the Sun rises south of east and sets south of west. It spends 9 hours above the horizon in the United States, which means fewer hours of daylight and more hours of night in northern lands (and a strong need for people to hold celebrations to cheer themselves up). On March 21 and September 21, the Sun spends equal amounts of time above and below the horizon in both hemispheres.

As we saw in Observing the Sky: The Birth of Astronomy, an equivalent way to look at our path around the Sun each year is to pretend that the Sun moves around Earth (on a circle called the ecliptic). Because Earth's axis is tilted, the ecliptic is tilted by about 23.5° relative to the celestial equator (review Figure 2.7). As a result, where we see the Sun in the sky changes as the year wears on.

In June, the Sun is north of the celestial equator and spends more time with those who live in the Northern Hemisphere. It rises high in the sky and is above the horizon in the United States for as long as 15 hours. Thus, the Sun not only heats us with more direct rays, but it also has more time to do it each day. (Notice in Figure 4.7 that the Northern Hemisphere's gain is the Southern Hemisphere's loss. There the June Sun is low in the sky, meaning fewer daylight hours. In Chile, for example, June is a colder, darker time of year.) In December, when the Sun is south of the celestial equator, the situation is reversed.

Link to Learning

The Motions of the Sun Simulator from Columbia University's Center for Teaching and Learning provides a demonstration of the Sun's apparent motion in the sky as the Earth rotates each day, and its changing altitude with latitude and time of year.

Let's look at what the Sun's illumination on Earth looks like at some specific dates of the year, when these effects are at their maximum. On or about June 21 (the date we who live in the Northern Hemisphere call the *summer solstice* or sometimes the first day of summer), the Sun shines down most directly upon the Northern Hemisphere of Earth. It appears about 23° north of the equator, and thus, on that date, it passes through the zenith of places on Earth that are at 23° N latitude. The situation is shown in detail in Figure 4.8. To a person at 23° N (near Hawaii, for example), the Sun is directly overhead at noon. This latitude, where the Sun can appear at the zenith at noon on the first day of summer, is called the *Tropic of Cancer*.

We also see in Figure 4.8 that the Sun's rays shine down all around the North Pole at the solstice. As Earth turns on its axis, the North Pole is continuously illuminated by the Sun; all places within 23° of the pole have sunshine for 24 hours. The Sun is as far north on this date as it can get; thus, 90° – 23° (or 67° N) is the southernmost latitude where the Sun can be seen for a full 24-hour period (sometimes called the "land of the midnight Sun"). That circle of latitude is called the *Arctic Circle*.

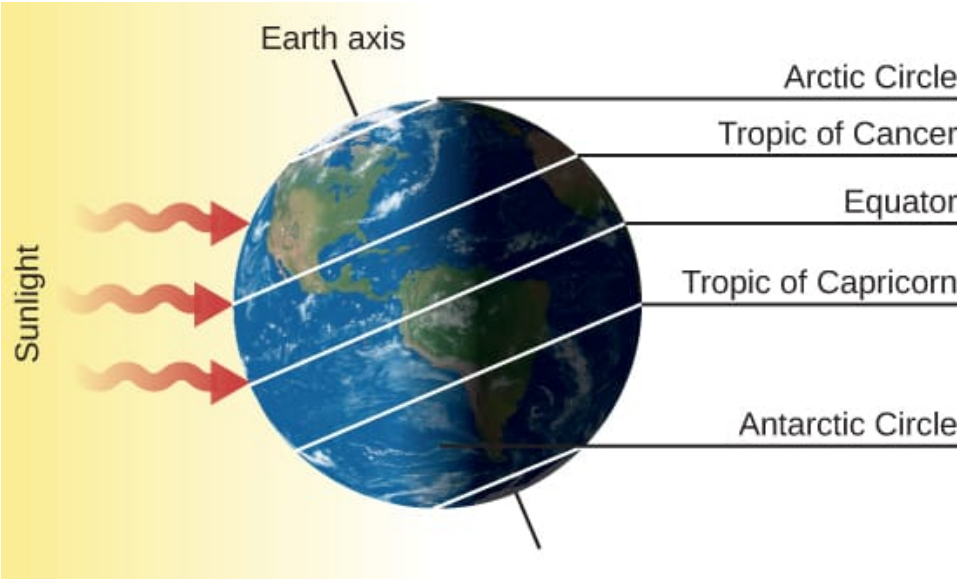


Figure 4.8 : Earth on June 21. This is the date of the summer solstice in the Northern Hemisphere. Note that as Earth turns on its axis (the line connecting the North and South Poles), the North Pole is in constant sunlight while the South Pole is veiled in 24 hours of darkness. The Sun is at the zenith for observers on the Tropic of Cancer.

Many early cultures scheduled special events around the summer solstice to celebrate the longest days and thank their gods for making the weather warm. This required people to keep track of the lengths of the days and the northward trek of the Sun in order to know the right day for the "party." (You can do the same thing by watching for several weeks, from the same observation point, where the Sun rises or sets relative to a fixed landmark. In spring, the Sun will rise farther and farther north of east, and set farther and farther north of west, reaching the maximum around the summer solstice.)

Now look at the South Pole in Figure 4.8. On June 21, all places within 23° of the South Pole—that is, south of what we call the *Antarctic Circle*—do not see the Sun at all for 24 hours.

The situation is reversed 6 months later, about December 21 (the date of the *winter solstice*, or the first day of winter in the Northern Hemisphere), as shown in Figure 4.9. Now it is the Arctic Circle that has the 24-hour night and the Antarctic Circle that has the midnight Sun. At latitude 23° S, called the *Tropic of Capricorn*, the Sun passes through the zenith at noon. Days are longer in the Southern Hemisphere and shorter in the north. In the United States and Southern Europe, there may be only 9 or 10 hours of sunshine during the day. It is winter in the Northern Hemisphere and summer in the Southern Hemisphere.

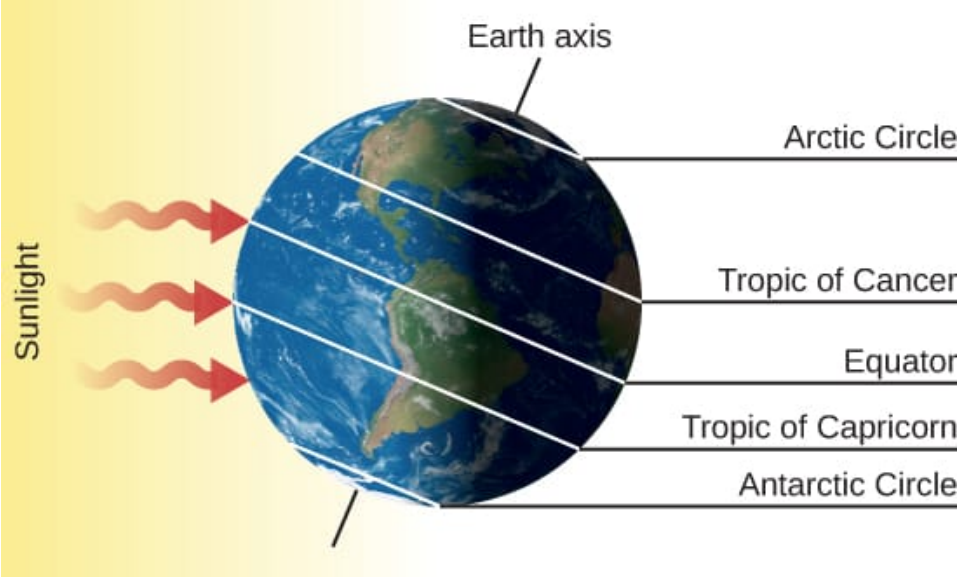


Figure 4.9 : Earth on December 21. This is the date of the winter solstice in the Northern Hemisphere. Now the North Pole is in darkness for 24 hours and the South Pole is illuminated. The Sun is at the zenith for observers on the Tropic of Capricorn and thus is low in the sky for the residents of the Northern Hemisphere.

Example 4.1: Seasonal Variations

As you can see in Figure 4.8, the Tropic of Cancer is the latitude for which the Sun is directly overhead on the summer solstice. At this time, the Sun is at a declination of 23° N of the celestial equator, and the corresponding latitude on Earth is 23° N of the equator. If Earth were tilted a bit less, then the Tropic of Cancer would be at a lower latitude, closer to the equator.

The Arctic Circle marks the southernmost latitude for which the day length is 24 hours on the day of the summer solstice. This is located at 90° – 23° = 67° N of Earth's equator. If Earth were tilted a bit less, then the Arctic Circle would move farther North. In the limit at which Earth is not tilted at all (its axis is perpendicular to the ecliptic), the Tropic of Cancer would be right on Earth's equator, and the Arctic Circle would simply be the North Pole. Suppose the tilt of Earth's axis were tilted only 5°. What would be the effect on the seasons and the locations of the Tropic of Cancer and Arctic Circle?

Solution

If Earth were tilted less, the seasons would be less extreme. The variation in day length and direct sunlight would be very small over the course of a year, and the Sun's daily path in the sky would not vary much. If Earth were tilted by 5°, the Sun's position on the day of the summer solstice would be 5° N of the celestial equator, so the Tropic of Cancer would be at the corresponding latitude on Earth of 5° N of the Equator. The Arctic Circle would be located at 90° – 5° = 85° N of the equator.

Exercise 6.2.1

Suppose the tilt of Earth's axis were 16°. What, then, would be the difference in latitude between the Arctic Circle and the Tropic of Cancer? What would be the effect on the seasons compared with that produced by the actual tilt of 23°?

Answer available at the end of this section.

Link to Learning

The Columbia University's Center for Teaching and Learning's Motions of the Sun Simulator allows you to see how the Sun's changing altitude over the course of the year changes the intensity of lighting experienced on Earth. Select the "Step by day" option before running the animation to illustrate this effect.

Many cultures that developed some distance north of the equator have a celebration around December 21 to help people deal with the depressing lack of sunlight and the often dangerously cold temperatures. Originally, this was often a time for huddling with family and friends, for sharing the reserves of food and drink, and for rituals asking the gods to return the light and heat and turn the cycle of the seasons around. Many cultures constructed elaborate devices for anticipating when the shortest day of the year was coming. Stonehenge in England, built long before the invention of writing, is probably one such device. In our own time, we continue the winter solstice tradition with various holiday celebrations around that December date.

Halfway between the solstices, on about March 21 and September 21, the Sun is on the celestial equator. From Earth, it appears above our planet's equator and favors neither hemisphere. Every place on Earth then receives roughly 12 hours of sunshine and 12 hours of night. The points where the Sun crosses the celestial equator are called the *vernal* (spring) and *autumnal* (fall) *equinoxes*.

The Seasons at Different Latitudes

The seasonal effects are different at different latitudes on Earth. Near the equator, for instance, all seasons are much the same. Every day of the year, the Sun is up half the time, so there are approximately 12 hours of sunshine and 12 hours of night. Local residents define the seasons by the amount of rain (wet season and dry season) rather than by the amount of sunlight. As we travel north or south, the seasons become more pronounced, until we reach extreme cases in the Arctic and Antarctic.

At the North Pole, all celestial objects that are north of the celestial equator are always above the horizon and, as Earth turns, circle around parallel to it. The Sun is north of the celestial equator from about March 21 to September 21, so at the North Pole, the Sun rises when it reaches the vernal equinox and sets when it reaches the autumnal equinox. Each year there are 6 months of sunshine at each pole, followed by 6 months of darkness.

Example 4.2: The Position of the Sun in the Sky

The Sun's coordinates on the celestial sphere range from a declination of 23° N of the celestial equator (or +23°) to a declination 23° S of the celestial equator (or –23°). So, the Sun's altitude at noon, when it crosses the meridian, varies by a total of 46°. What is the altitude of the Sun at noon on March 21, as seen from a place on Earth's equator? What is its altitude on June 21, as seen from a place on Earth's equator?

Solution

On Earth's equator, the celestial equator passes through the zenith. On March 21, the Sun is crossing the celestial equator, so it should be found at the zenith (90°) at noon. On June 21, the Sun is 23° N of the celestial equator, so it will be 23° away from the zenith at noon. The altitude above the horizon will be 23° less than the altitude of the zenith (90°), so it is 90° – 23° = 67° above the horizon.

Exercise 6.2.1

What is the altitude of the Sun at noon on December 21, as seen from a place on the Tropic of Cancer?

Answer available at the end of this section.

Clarifications about the Real World

In our discussions so far, we have been describing the rising and setting of the Sun and stars as they would appear if Earth had little or no atmosphere. In reality, however, the atmosphere has the curious effect of allowing us to see a little way "over the horizon." This effect is a result of *refraction*, the bending of light passing through air or water, something we will discuss in Astronomical Instruments. Because of this atmospheric refraction (and the fact that the Sun is not a point of light but a disk), the Sun appears to rise earlier and to set later than it would if no atmosphere were present.

In addition, the atmosphere scatters light and provides some twilight illumination even when the Sun is below the horizon. Astronomers define morning twilight as beginning when the Sun is 18° below the horizon, and evening twilight extends until the Sun sinks more than 18° below the horizon.

These atmospheric effects require small corrections in many of our statements about the seasons. At the equinoxes, for example, the Sun appears to be above the horizon for a few minutes longer than 12 hours, and below the horizon for fewer than 12 hours. These effects are most dramatic at Earth's poles, where the Sun actually can be seen more than a week before it reaches the celestial equator.

You probably know that the summer solstice (June 21) is not the warmest day of the year, even if it is the longest. The hottest months in the Northern Hemisphere are July and August. This is because our weather involves the air and water covering Earth's surface, and these large reservoirs do not heat up instantaneously. You have probably observed this effect for yourself; for example, a pond does not get warm the moment the Sun rises but is warmest late in the afternoon, after it has had time to absorb the Sun's heat. In the same way, Earth gets warmer after it has had a chance to absorb the extra sunlight that is the Sun's summer gift to us. And the coldest times of winter are a month or more after the winter solstice.

Footnotes

- 1Note that the dates indicated for the solstices and equinoxes are approximate; depending on the year, they may occur a day or two earlier or later.

Answers

Exercise 6.2.1

The Tropic of Cancer is at a latitude equal to Earth's tilt, so in this case, it would be at 16° N latitude. The Arctic Circle is at a latitude equal to 90° minus Earth's tilt, or 90° − 16° = 74°. The difference between these two latitudes is 74° − 16° = 58°. Since the tilt of Earth is less, there would be less variation in the tilt of Earth and less variation in the Sun's paths throughout the year, so there would be milder seasonal changes.

Exercise 6.2.1

On the day of the winter solstice, the Sun is located about 23° S of the celestial equator. From the Tropic of Cancer, a latitude of 23° N, the zenith would be a declination of 23° N. The difference in declination between zenith and the position of the Sun is 46°, so the Sun would be 46° away from the zenith. That means it would be at an altitude of 90° − 46° = 44°.

This page titled 6.2: The Seasons is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.3: The Seasons by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

6.3: Keeping Time

Learning Objectives

By the end of this section, you will be able to:

- Explain the difference between the solar day and the sidereal day
- Explain mean solar time and the reason for time zones

The measurement of time is based on the rotation of Earth. Throughout most of human history, time has been reckoned by positions of the Sun and stars in the sky. Only recently have mechanical and electronic clocks taken over this function in regulating our lives.

The Length of the Day

The most fundamental astronomical unit of time is the day, measured in terms of the rotation of Earth. There is, however, more than one way to define the day. Usually, we think of it as the rotation period of Earth with respect to the Sun, called the solar day. After all, for most people sunrise is more important than the rising time of Arcturus or some other star, so we set our clocks to some version of Sun-time. However, astronomers also use a sidereal day, which is defined in terms of the rotation period of Earth with respect to the stars.

A solar day is slightly longer than a sidereal day because (as you can see from Figure 4.10) Earth not only turns but also moves along its path around the Sun in a day. Suppose we start when Earth's orbital position is at day 1, with both the Sun and some distant star (located in the direction indicated by the long white arrow pointing left), directly in line with the zenith for the observer on Earth. When Earth has completed one rotation with respect to the distant star and is at day 2, the long arrow again points to the same distant star. However, notice that because of the movement of Earth along its orbit from day 1 to 2, the Sun has not yet reached a position above the observer. To complete a solar day, Earth must rotate an additional amount, equal to 1/365 of a full turn. The time required for this extra rotation is 1/365 of a day, or about 4 minutes. So the solar day is about 4 minutes longer than the sidereal day.

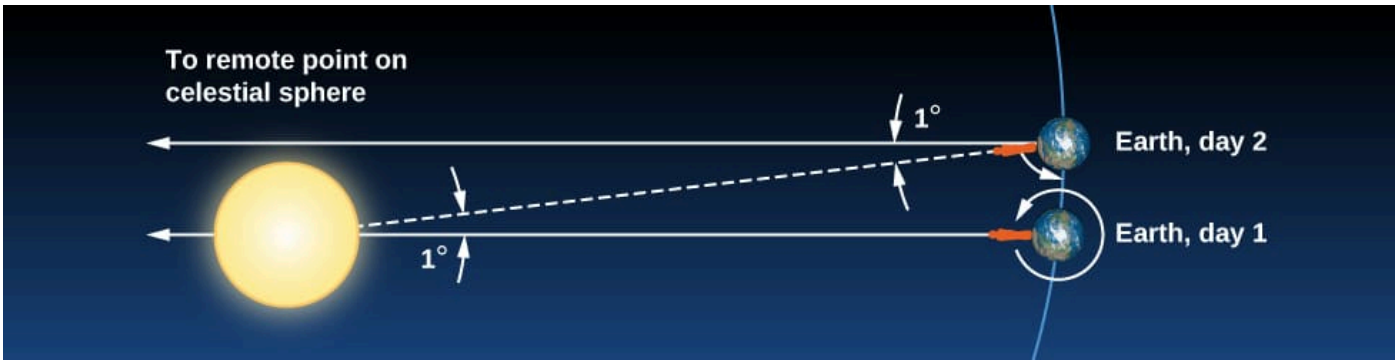


Figure 4.10 : Difference Between a Sidereal Day and a Solar Day. This is a top view, looking down as Earth orbits the Sun. Because Earth moves around the Sun (roughly 1° per day), after one complete rotation of Earth relative to the stars, we do not see the Sun in the same position.

Because our ordinary clocks are set to solar time, stars rise 4 minutes earlier each day. Astronomers prefer sidereal time for planning their observations because in that system, a star rises at the same time every day.

Example 4.3: Sidereal Time and Solar Time

The Sun makes a complete circle in the sky approximately every 24 hours, while the stars make a complete circle in the sky in 4 minutes less time, or 23 hours and 56 minutes. This causes the positions of the stars at a given time of day or night to change slightly each day. Since stars rise 4 minutes earlier each day, that works out to about 2 hours per month

$(4\text{minutes} \times 30 = 120\text{minutes or} 2\text{hours})$. So, if a particular constellation rises at sunset during the winter, you can be sure that by the summer, it will rise about 12 hours earlier, with the sunrise, and it will not be so easily visible in the night sky. Let's say that tonight the bright star Sirius rises at 7:00 p.m. from a given location so that by midnight, it is very high in the sky. At what time will Sirius rise in three months?

Solution

In three months' time, Sirius will be rising earlier by:

$$90 \text{ days} \times \frac{4 \text{ minutes}}{\text{day}} = 360 \text{ minutes or } 6$$

It will rise at about 1:00 p.m. and be high in the sky at around sunset instead of midnight. Sirius is the brightest star in the constellation of Canis Major (the big dog). So, some other constellation will be prominently visible high in the sky at this later date.

Exercise 6.3.1

If a star rises at 8:30 p.m. tonight, approximately what time will it rise two months from now?

Answer available at the end of this section.

Apparent Solar Time

We can define apparent solar time as time reckoned by the actual position of the Sun in the sky (or, during the night, its position below the horizon). This is the kind of time indicated by sundials, and it probably represents the earliest measure of time used by ancient civilizations. Today, we adopt the middle of the night as the starting point of the day and measure time in hours elapsed since midnight.

During the first half of the day, the Sun has not yet reached the meridian (the great circle in the sky that passes through our zenith). We designate those hours as before midday (*ante meridiem*, or a.m.), before the Sun reaches the local meridian. We customarily start numbering the hours after noon over again and designate them by p.m. (*post meridiem*), after the Sun reaches the local meridian.

Although apparent solar time seems simple, it is not really very convenient to use. The exact length of an apparent solar day varies slightly during the year. The eastward progress of the Sun in its annual journey around the sky is not uniform because the speed of Earth varies slightly in its elliptical orbit. Another complication is that Earth's axis of rotation is not perpendicular to the plane of its revolution. Thus, apparent solar time does not advance at a uniform rate. After the invention of mechanical clocks that run at a uniform rate, it became necessary to abandon the apparent solar day as the fundamental unit of time.

Mean Solar Time and Standard Time

Instead, we can consider the mean solar time, which is based on the average value of the solar day over the course of the year. A mean solar day contains exactly 24 hours and is what we use in our everyday timekeeping. Although mean solar time has the advantage of progressing at a uniform rate, it is still inconvenient for practical use because it is determined by the position of the Sun. For example, noon occurs when the Sun is highest in the sky on the meridian (but not necessarily at the zenith). But because we live on a round Earth, the exact time of noon is different as you change your longitude by moving east or west.

If mean solar time were strictly observed, people traveling east or west would have to reset their watches continually as the longitude changed, just to read the local mean time correctly. For instance, a commuter traveling from Oyster Bay on Long Island to New York City would have to adjust the time on the trip through the East River tunnel because Oyster Bay time is actually about 1.6 minutes more advanced than that of Manhattan. (Imagine an airplane trip in which an obnoxious flight attendant gets on the intercom every minute, saying, "Please reset your watch for local mean time.")

Until near the end of the nineteenth century, every city and town in the United States kept its own local mean time. With the

development of railroads and the telegraph, however, the need for some kind of standardization became evident. In 1883, the United States was divided into four standard time zones (now six, including Hawaii and Alaska), each with one system of time within that zone.

By 1900, most of the world was using the system of 24 standardized global time zones. Within each zone, all places keep the same *standard time*, with the local mean solar time of a standard line of longitude running more or less through the middle of each zone. Now travelers reset their watches only when the time change has amounted to a full hour. Pacific standard time is 3 hours earlier than eastern standard time, a fact that becomes painfully obvious in California when someone on the East Coast forgets and calls you at 5:00 a.m.

Globally, almost all countries have adopted one or more standard time zones, although one of the largest nations, India, has settled on a half-zone, being 5.5 hours from Greenwich standard. Also, China officially uses only one time zone, so all the clocks in that country keep the same time. In Tibet, for example, the Sun rises while the clocks (which keep Beijing time) say it is midmorning already.

Daylight saving time is simply the local standard time of the place plus 1 hour. It has been adopted for spring and summer use in most states in the United States, as well as in many countries, to prolong the sunlight into evening hours, on the apparent theory that it is easier to change the time by government action than it would be for individuals or businesses to adjust their own schedules to produce the same effect. It does not, of course, "save" any daylight at all—because the amount of sunlight is not determined by what we do with our clocks—and its observance is a point of legislative debate in some states.

The International Date Line

The fact that time is always advancing as you move toward the east presents a problem. Suppose you travel eastward around the world. You pass into a new time zone, on the average, about every 15° of longitude you travel, and each time you dutifully set your watch ahead an hour. By the time you have completed your trip, you have set your watch ahead a full 24 hours and thus gained a day over those who stayed at home.

The solution to this dilemma is the International Date Line, set by international agreement to run approximately along the 180° meridian of longitude. The date line runs down the middle of the Pacific Ocean, although it jogs a bit in a few places to avoid cutting through groups of islands and through Alaska (Figure 4.11). By convention, at the date line, the date of the calendar is changed by one day. Crossing the date line from west to east, thus advancing your time, you compensate by decreasing the date; crossing from east to west, you increase the date by one day. To maintain our planet on a rational system of timekeeping, we simply must accept that the date will differ in different cities at the same time. A good example is the date when the Imperial Japanese Navy bombed Pearl Harbor in Hawaii, known in the United States as Sunday, December 7, 1941, but taught to Japanese students as Monday, December 8.

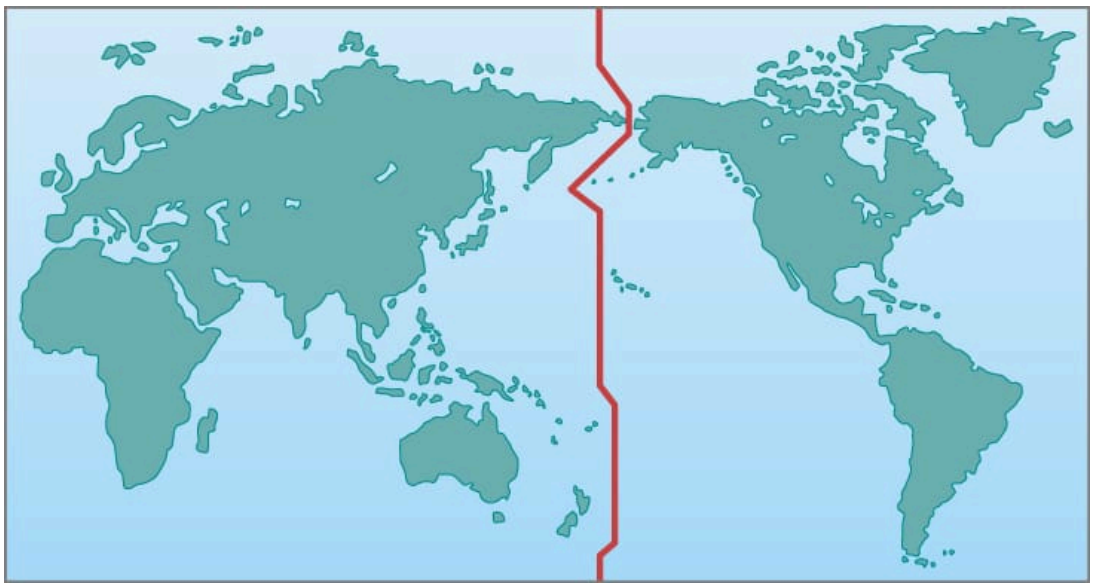


Figure 4.11 : Where the Date Changes. The International Date Line is an arbitrarily drawn line on Earth where the date changes. So that neighbors do not have different days, the line is located where Earth's surface is mostly water.

Answers

Exercise 6.3.1

In two months, the star will rise:
$60 \text{ days} \times \frac{4 \text{ minutes}}{\text{day}} = 240 \text{ minutes or } 4 \text{ hours earlier.}$
This means it will rise at 4:30 p.m.

This page titled 6.3: Keeping Time is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.4: Keeping Time by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

6.4: The Calendar

Learning Objectives

By the end of this section, you will be able to:

- Understand how calendars varied among different cultures
- Explain the origins of our modern calendar

"What's today's date?" is one of the most common questions you can ask (usually when signing a document or worrying about whether you should have started studying for your next astronomy exam). Long before the era of digital watches, smartphones, and fitness bands that tell the date, people used calendars to help measure the passage of time.

The Challenge of the Calendar

There are two traditional functions of any calendar. First, it must keep track of time over the course of long spans, allowing people to anticipate the cycle of the seasons and to honor special religious or personal anniversaries. Second, to be useful to a large number of people, a calendar must use natural time intervals that everyone can agree on—those defined by the motions of Earth, the Moon, and sometimes even the planets. The natural units of our calendar are the *day*, based on the period of rotation of Earth; the *month*, based on the cycle of the Moon's phases (see later in this chapter) about Earth; and the year, based on the period of revolution of Earth about the Sun. Difficulties have resulted from the fact that these three periods are not commensurable; that's a fancy way of saying that one does not divide evenly into any of the others.

The rotation period of Earth is, by definition, 1.0000 day (and here the solar day is used, since that is the basis of human experience). The period required by the Moon to complete its cycle of phases, called the *month*, is 29.5306 days. The basic period of revolution of Earth, called the *tropical year*, is 365.2422 days. The ratios of these numbers are not convenient for calculations. This is the historic challenge of the calendar, dealt with in various ways by different cultures.

Early Calendars

Even the earliest cultures were concerned with the keeping of time and the calendar. Some interesting examples include monuments left by Bronze Age people in northwestern Europe, especially the British Isles. The best preserved of the monuments is Stonehenge, about 13 kilometers from Salisbury in southwest England (Figure 4.12). It is a complex array of stones, ditches, and holes arranged in concentric circles. Carbon dating and other studies show that Stonehenge was built during three periods ranging from about 2800 to 1500 BCE. Some of the stones are aligned with the directions of the Sun and Moon during their risings and settings at critical times of the year (such as the summer and winter solstices), and it is generally believed that at least one function of the monument was connected with the keeping of a calendar.



Figure 4.12 : Stonehenge. The ancient monument known as Stonehenge was used to keep track of the motions of the Sun and Moon. (credit: modification of work by Adriano Aurelio Araujo)

The Maya in Central America, who thrived more than a thousand years ago, were also concerned with the keeping of time. Their calendar was as sophisticated as, and perhaps more complex than, contemporary calendars in Europe. The Maya did not attempt to correlate their calendar accurately with the length of the year or lunar month. Rather, their calendar was a system for keeping track of the passage of days and for counting time far into the past or future. Among other purposes, it was useful for predicting astronomical events, such as the position of Venus in the sky (Figure 4.13).



Figure 4.13 : El Caracol. This Mayan observatory at Chichen Itza in the Yucatan, Mexico, dates from around the year 1000. (credit: "wiredtourist.com"/Flickr)

The ancient Chinese developed an especially complex calendar, largely limited to a few privileged hereditary court astronomer-astrologers. In addition to the motions of Earth and the Moon, they were able to fit in the approximately 12-year cycle of Jupiter, which was central to their system of astrology. The Chinese still preserve some aspects of this system in their cycle of 12 "years"—the Year of the Dragon, the Year of the Pig, and so on—that are defined by the position of Jupiter in the zodiac.

Our Western calendar derives from a long history of timekeeping beginning with the Sumerians, dating back to at least the second millennium BCE, and continuing with the Egyptians and the Greeks around the eighth century BCE. These calendars led, eventually, to the *Julian calendar*, introduced by Julius Caesar, which approximated the year at 365.25 days, fairly close to the actual value of 365.2422. The Romans achieved this approximation by declaring years to have 365 days each, with the exception

of every fourth year. The *leap year* was to have one extra day, bringing its length to 366 days, and thus making the average length of the year in the Julian calendar 365.25 days.

In this calendar, the Romans had dropped the almost impossible task of trying to base their calendar on the Moon as well as the Sun, although a vestige of older lunar systems can be seen in the fact that our months have an average length of about 30 days. However, lunar calendars remained in use in other cultures, and Islamic calendars, for example, are still primarily lunar rather than solar.

The Gregorian Calendar

Although the Julian calendar (which was adopted by the early Christian Church) represented a great advance, its average year still differed from the true year by about 11 minutes, an amount that accumulates over the centuries to an appreciable error. By 1582, that 11 minutes per year had added up to the point where the first day of spring was occurring on March 11, instead of March 21. If the trend were allowed to continue, eventually the Christian celebration of Easter would be occurring in early winter. Pope Gregory XIII, a contemporary of Galileo, felt it necessary to institute further calendar reform.

The Gregorian calendar reform consisted of two steps. First, 10 days had to be dropped out of the calendar to bring the vernal equinox back to March 21; by proclamation, the day following October 4, 1582, became October 15. The second feature of the new Gregorian calendar was a change in the rule for leap year, making the average length of the year more closely approximate the tropical year. Gregory decreed that three of every four century years—all leap years under the Julian calendar—would be common years henceforth. The rule was that only century years divisible by 400 would be leap years. Thus, 1700, 1800, and 1900—all divisible by 4 but not by 400—were not leap years in the Gregorian calendar. On the other hand, the years 1600 and 2000, both divisible by 400, were leap years. The average length of this Gregorian year, 365.2425 mean solar days, is correct to about 1 day in 3300 years.

The Catholic countries immediately put the Gregorian reform into effect, but countries of the Eastern Church and most Protestant countries did not adopt it until much later. It was 1752 when England and the American colonies finally made the change. By parliamentary decree, September 2, 1752, was followed by September 14. Although special laws were passed to prevent such abuses as landlords collecting a full month's rent for September, there were still riots, and people demanded their 12 days back. Russia did not abandon the Julian calendar until the time of the Bolshevik revolution. The Russians then had to omit 13 days to come into step with the rest of the world. The anniversary of the October Revolution (old calendar) of 1917, bringing the communists to power, thus ended up being celebrated in November (new calendar), a difference that is perhaps not so important since the fall of communism.

This page titled 6.4: The Calendar is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.5: The Calendar by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

7: THE MOON

7.1: Phases and Motions of the Moon

This page titled 7: The Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

7.1: Phases and Motions of the Moon

Learning Objectives

By the end of this section, you will be able to:

- Explain the cause of the lunar phases
- Understand how the Moon rotates and revolves around Earth

After the Sun, the Moon is the brightest and most obvious object in the sky. Unlike the Sun, it does not shine under its own power, but merely glows with reflected sunlight. If you were to follow its progress in the sky for a month, you would observe a cycle of phases (different appearances), with the Moon starting dark and getting more and more illuminated by sunlight over the course of about two weeks. After the Moon's disk becomes fully bright, it begins to fade, returning to dark about two weeks later.

These changes fascinated and mystified many early cultures, which came up with marvelous stories and legends to explain the cycle of the Moon. Even in the modern world, many people don't understand what causes the phases, thinking that they are somehow related to the shadow of Earth. Let us see how the phases can be explained by the motion of the Moon relative to the bright light source in the solar system, the Sun.

Lunar Phases

Although we know that the Sun moves 1/12 of its path around the sky each month, for purposes of explaining the phases, we can assume that the Sun's light comes from roughly the same direction during the course of a four-week lunar cycle. The Moon, on the other hand, moves completely around Earth in that time. As we watch the Moon from our vantage point on Earth, how much of its face we see illuminated by sunlight depends on the angle the Sun makes with the Moon.

Here is a simple experiment to show you what we mean: stand about 6 feet in front of a bright electric light in a completely dark room (or outdoors at night) and hold in your hand a small round object such as a tennis ball or an orange. Your head can then represent Earth, the light represents the Sun, and the ball the Moon. Move the ball around your head (making sure you don't cause an eclipse by blocking the light with your head). You will see phases just like those of the Moon on the ball. (Another good way to get acquainted with the phases and motions of the Moon is to follow our satellite in the sky for a month or two, recording its shape, its direction from the Sun, and when it rises and sets.)

Let's examine the Moon's cycle of phases using Figure 4.14, which depicts the Moon's behavior for the entire month. The trick to this figure is that you must imagine yourself standing on Earth, facing the Moon in each of its phases. So, for the position labeled "New," you are on the right side of Earth and it's the middle of the day; for the position "Full," you are on the left side of Earth in the middle of the night. Note that in every position on Figure 4.14, the Moon is half illuminated and half dark (as a ball in sunlight should be). The difference at each position has to do with what part of the Moon faces Earth.

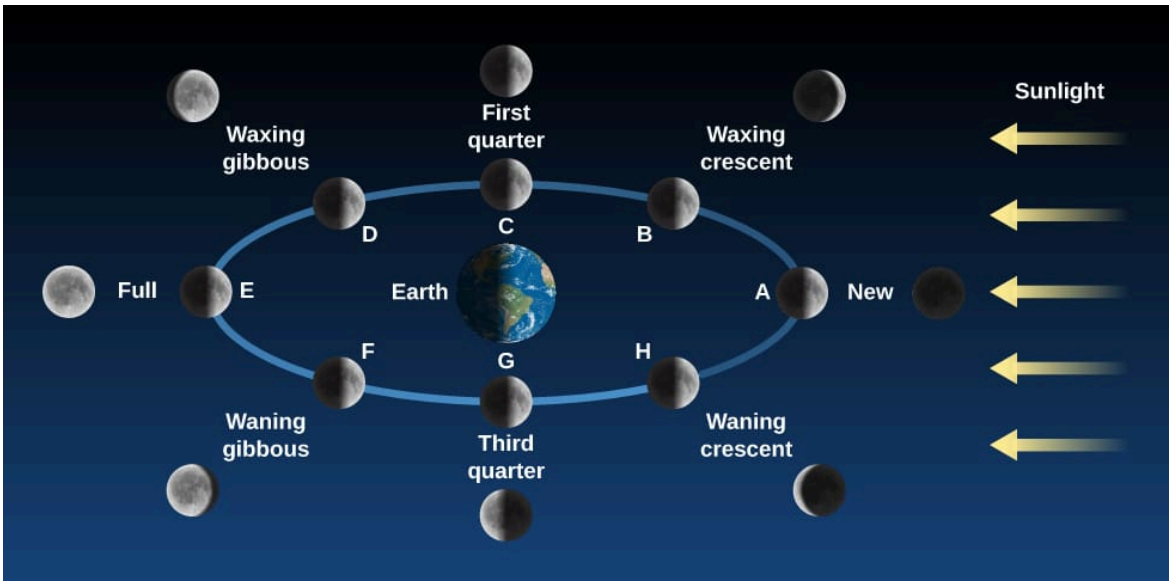


Figure 4.14 : Phases of the Moon. The appearance of the Moon changes over the course of a complete monthly cycle. The pictures of the Moon on the white circle show the perspective from space, with the Sun off to the right in a fixed position. The outer images show how the Moon appears to you in the sky from each point in the orbit. Imagine yourself standing on Earth, facing the Moon at each stage. In the position "New," for example, you are facing the Moon from the right side of Earth in the middle of the day. (Note that the distance of the Moon from Earth is not to scale in this diagram: the Moon is roughly 30 Earth-diameters away from us.) (credit: modification of work by NASA)

The Moon is said to be *new* when it is in the same general direction in the sky as the Sun (position A). Here, its illuminated (bright) side is turned away from us and its dark side is turned toward us. You might say that the Sun is shining on the "wrong " side of the Moon from our perspective. In this phase the Moon is invisible to us; its dark, rocky surface does not give off any light of its own. Because the new moon is in the same part of the sky as the Sun, it rises at sunrise and sets at sunset.

But the Moon does not remain in this phase long because it moves eastward each day in its monthly path around us. Since it takes about 30 days to orbit Earth and there are 360° in a circle, the Moon will move about 12° in the sky each day (or about 24 times its own diameter). A day or two after the new phase, the thin *crescent* first appears, as we begin to see a small part of the Moon's illuminated hemisphere. It has moved into a position where it now reflects a little sunlight toward us along one side. The bright crescent increases in size on successive days as the Moon moves farther and farther around the sky away from the direction of the Sun (position B). Because the Moon is moving eastward away from the Sun, it rises later and later each day (like a student during summer vacation).

After about one week, the Moon is one-quarter of the way around its orbit (position C) and so we say it is at the *first quarter* phase. Half of the Moon's illuminated side is visible to Earth observers. Because of its eastward motion, the Moon now lags about one-quarter of the day behind the Sun, rising around noon and setting around midnight.

During the week after the first quarter phase, we see more and more of the Moon's illuminated hemisphere (position D), a phase that is called *waxing* (or growing) gibbous (from the Latin *gibbus*, meaning hump). Eventually, the Moon arrives at position E in our figure, where it and the Sun are opposite each other in the sky. The side of the Moon turned toward the Sun is also turned toward Earth, and we have the *full* phase.

When the Moon is full, it is opposite the Sun in the sky. The Moon does the opposite of what the Sun does, rising at sunset and setting at sunrise. Note what that means in practice: the completely illuminated (and thus very noticeable) Moon rises just as it gets dark, remains in the sky all night long, and sets as the Sun's first rays are seen at dawn. Its illumination throughout the night helps lovers on a romantic stroll and students finding their way back to their dorms after a long night in the library or an off-campus party.

And when is the full moon highest in the sky and most noticeable? At midnight, a time made famous in generations of horror novels and films. (Note how the behavior of a vampire like Dracula parallels the behavior of the full Moon: Dracula rises at sunset, does his worst mischief at midnight, and must be back down in his coffin by sunrise. The old legends were a way of personifying the behavior of the Moon, which was a much more dramatic part of people's lives in the days before electric lights and television.)

Folklore has it that more crazy behavior is seen during the time of the full moon (the Moon even gives a name to crazy

behavior—"lunacy"). But, in fact, statistical tests of this "hypothesis" involving thousands of records from hospital emergency rooms and police files do not reveal any correlation of human behavior with the phases of the Moon. For example, homicides occur at the same rate during the new moon or the crescent moon as during the full moon. Most investigators believe that the real story is not that more crazy behavior happens on nights with a full moon, but rather that we are more likely to notice or remember such behavior with the aid of a bright celestial light that is up all night long.

During the two weeks following the full moon, the Moon goes through the same phases again in reverse order (points F, G, and H in Figure 4.14), returning to new phase after about 29.5 days. About a week after the full moon, for example, the Moon is at *third quarter*, meaning that it is three-quarters of the way around (not that it is three-quarters illuminated—in fact, half of the visible side of the Moon is again dark). At this phase, the Moon is now rising around midnight and setting around noon.

Note that there is one thing quite misleading about Figure 4.14. If you look at the Moon in position E, although it is full in theory, it appears as if its illumination would in fact be blocked by a big fat Earth, and hence we would not see anything on the Moon except Earth's shadow. In reality, the Moon is nowhere near as close to Earth (nor is its path so identical with the Sun's in the sky) as this diagram (and the diagrams in most textbooks) might lead you to believe.

The Moon is actually 30 *Earth-diameters* away from us; Science and the Universe: A Brief Tour contains a diagram that shows the two objects to scale. And, since the Moon's orbit is tilted relative to the path of the Sun in the sky, Earth's shadow misses the Moon most months. That's why we regularly get treated to a full moon. The times when Earth's shadow does fall on the Moon are called lunar eclipses and are discussed in Eclipses of the Sun and Moon.

Making Connections: Astronomy and the Days of the Week

The week seems independent of celestial motions, although its length may have been based on the time between quarter phases of the Moon. In Western culture, the seven days of the week are named after the seven "wanderers" that the ancients saw in the sky: the Sun, the Moon, and the five planets visible to the unaided eye (Mercury, Venus, Mars, Jupiter, and Saturn).

In English, we can easily recognize the names Sun-day (Sunday), Moon-day (Monday), and Saturn-day (Saturday), but the other days are named after the Norse equivalents of the Roman gods that gave their names to the planets. In languages more directly related to Latin, the correspondences are clearer. Wednesday, Mercury's day, for example, is *mercoledi* in Italian, *mercredi* in French, and *miércoles* in Spanish. Mars gives its name to Tuesday (*martes* in Spanish), Jupiter or Jove to Thursday (*giovedi* in Italian), and Venus to Friday (*vendredi* in French).

There is no reason that the week has to have seven days rather than five or eight. It is interesting to speculate that if we had lived in a planetary system where more planets were visible without a telescope, the Beatles could have been right and we might well have had "Eight Days a Week."

Link to Learning

This simulation of the Moon orbiting the Earth shows both a "top-down" view and the view from Earth's surface.

The Moon's Revolution and Rotation

The Moon's sidereal period—that is, the period of its revolution about Earth measured with respect to the stars—is a little over 27 days: the sidereal month is 27.3217 days to be exact. The time interval in which the phases repeat—say, from full to full—is the solar month, 29.5306 days. The difference results from Earth's motion around the Sun. The Moon must make more than a complete turn around the moving Earth to get back to the same phase with respect to the Sun. As we saw, the Moon changes its position on the celestial sphere rather rapidly: even during a single evening, the Moon creeps visibly eastward among the stars, traveling its own width in a little less than 1 hour. The delay in moonrise from one day to the next caused by this eastward motion averages about 50 minutes.

The Moon *rotates* on its axis in exactly the same time that it takes to *revolve* about Earth. As a consequence, the Moon always keeps the same face turned toward Earth (Figure 4.15). You can simulate this yourself by "orbiting" your roommate or another volunteer. Start by facing your roommate. If you make one rotation (spin) with your shoulders in the exact same time that you revolve around him or her, you will continue to face your roommate during the whole "orbit." As we will see in coming chapters,

our Moon is not the only world that exhibits this behavior, which scientists call synchronous rotation.

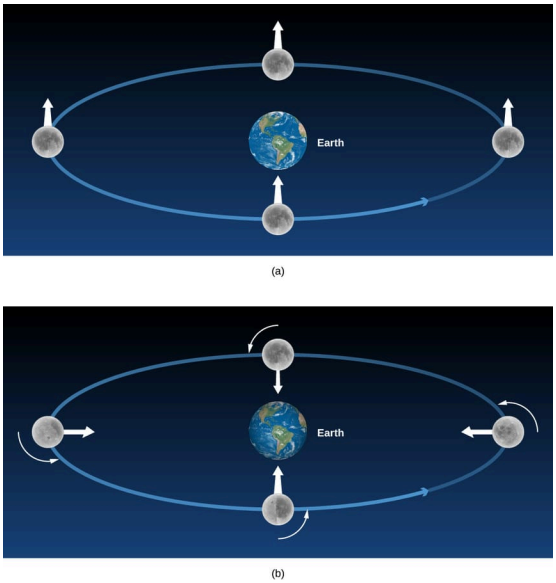


Figure 4.15 : The Moon without and with Rotation. In this figure, we stuck a white arrow into a fixed point on the Moon to keep track of its sides. (a) If the Moon did not rotate as it orbited Earth, it would present all of its sides to our view; hence the white arrow would point directly toward Earth only in the bottom position on the diagram. (b) Actually, the Moon rotates in the same period that it revolves, so we always see the same side (the white arrow keeps pointing to Earth).

The differences in the Moon's appearance from one night to the next are due to changing illumination by the Sun, not to its own rotation. You sometimes hear the back side of the Moon (the side we never see) called the "dark side." This is a misunderstanding of the real situation: which side is light and which is dark changes as the Moon moves around Earth. The back side is dark no more frequently than the front side. Since the Moon rotates, the Sun rises and sets on all sides of the Moon. With apologies to Pink Floyd, there is simply no regular "Dark Side of the Moon."

This page titled 7.1: Phases and Motions of the Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.6: Phases and Motions of the Moon by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

8: THE MOON 2

8.1: Ocean Tides and the Moon
8.2: Eclipses of the Sun and Moon
8.3: Key Terms

This page titled 8: The Moon 2 is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

8.1: Ocean Tides and the Moon

Learning Objectives

By the end of this section, you will be able to:

- Describe what causes tides on Earth
- Explain why the amplitude of tides changes during the course of a month

Anyone living near the sea is familiar with the twice-daily rising and falling of the tides. Early in history, it was clear that tides must be related to the Moon because the daily delay in high tide is the same as the daily delay in the Moon's rising. A satisfactory explanation of the tides, however, awaited the theory of gravity, supplied by Newton.

The Pull of the Moon on Earth

The gravitational forces exerted by the Moon at several points on Earth are illustrated in Figure 4.16. These forces differ slightly from one another because Earth is not a point, but has a certain size: all parts are not equally distant from the Moon, nor are they all in exactly the same direction from the Moon. Moreover, Earth is not perfectly rigid. As a result, the differences among the forces of the Moon's attraction on different parts of Earth (called *differential forces*) cause Earth to distort slightly. The side of Earth nearest the Moon is attracted toward the Moon more strongly than is the center of Earth, which in turn is attracted more strongly than is the side opposite the Moon. Thus, the differential forces tend to stretch Earth slightly into an *oblate spheroid* (which is something like an American football shape), with its long diameter pointed toward the Moon.

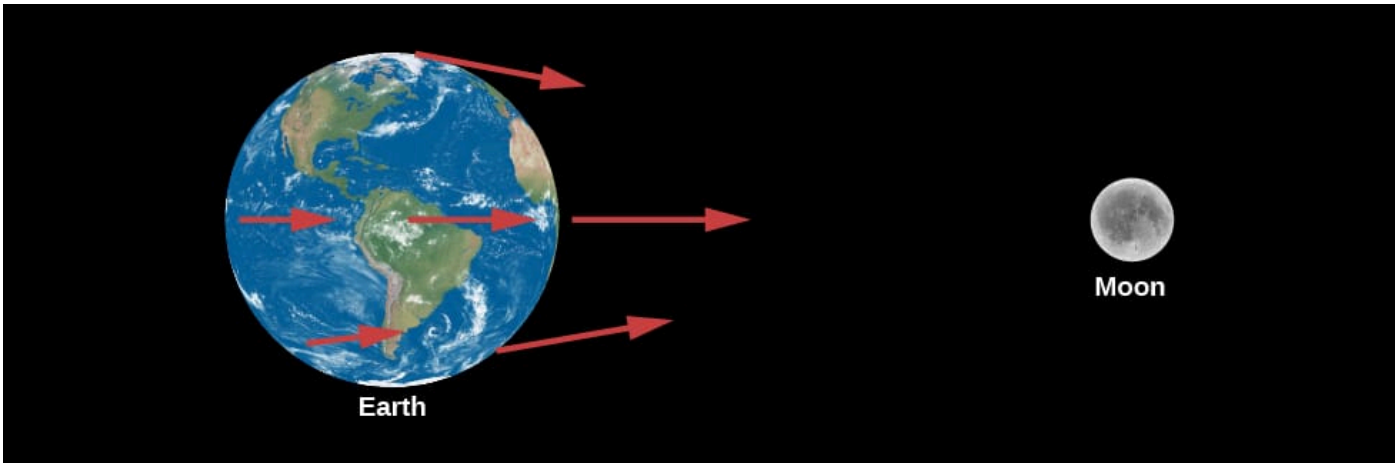


Figure 4.16 : Pull of the Moon. The Moon's differential attraction is shown on different parts of Earth. (Note that the differences have been exaggerated for educational purposes.)

If Earth were made of water, it would distort until the Moon's differential forces over different parts of its surface came into balance with Earth's own gravitational forces pulling it together. Calculations show that in this case, Earth would distort from a sphere by amounts ranging up to nearly 1 meter. Measurements of the actual deformation of Earth show that the solid Earth does distort, but only about one-third as much as water would, because of the greater rigidity of Earth's interior.

Because the tidal distortion of the solid Earth amounts—at its greatest—to only about 20 centimeters, Earth does not distort enough to balance the Moon's differential forces with its own gravity. Hence, objects at Earth's surface experience tiny horizontal tugs, tending to make them slide about. These *tide-raising forces* are too insignificant to affect solid objects like astronomy students or rocks in Earth's crust, but they do affect the waters in the oceans.

The Formation of Tides

The tide-raising forces, acting over a number of hours, produce motions of the water that result in measurable tidal bulges in the oceans. Water on the side of Earth facing the Moon flows toward it, with the greatest depths roughly at the point below the Moon. On the side of Earth opposite the Moon, water also flows to produce a tidal bulge (Figure 4.17).

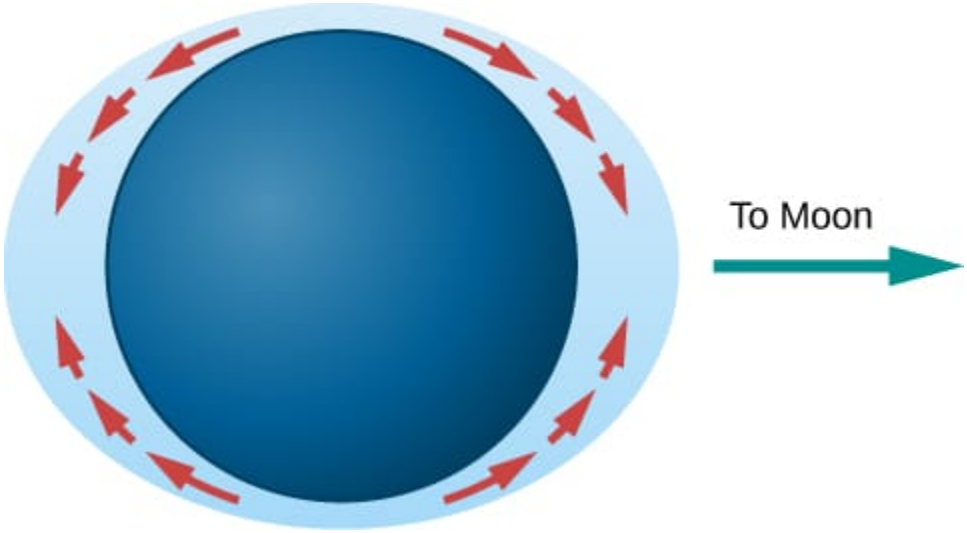


Figure 4.17 : Tidal Bulges in an “Ideal” Ocean. Differences in gravity cause tidal forces that push water in the direction of tidal bulges on Earth.

Link to Learning

You can run this animation for a visual demonstration of the tidal bulge.

Note that the tidal bulges in the oceans do not result from the Moon’s compressing or expanding the water, nor from the Moon’s lifting the water “away from Earth.” Rather, they result from an actual flow of water over Earth’s surface toward the two regions below and opposite the Moon, causing the water to pile up to greater depths at those places (Figure 4.18).



Figure 4.18 : High and Low Tides. This is a side-by-side comparison of the Bay of Fundy in Canada at high and low tides. (credit a, b: modification of work by Dylan Kereluk)

In the idealized (and, as we shall see, oversimplified) model just described, the height of the tides would be only a few feet. The rotation of Earth would carry an observer at any given place alternately into regions of deeper and shallower water. An observer being carried toward the regions under or opposite the Moon, where the water was deepest, would say, “The tide is coming in”; when carried away from those regions, the observer would say, “The tide is going out.” During a day, the observer would be carried

through two tidal bulges (one on each side of Earth) and so would experience two high tides and two low tides.

The Sun also produces tides on Earth, although it is less than half as effective as the Moon at tide raising. The actual tides we experience are a combination of the larger effect of the Moon and the smaller effect of the Sun. When the Sun and Moon are lined up (at new moon or full moon), the tides produced reinforce each other and so are greater than normal (Figure 4.19). These are called spring tides (the name is connected not to the season but to the idea that higher tides "spring up"). Spring tides are approximately the same, whether the Sun and Moon are on the same or opposite sides of Earth, because tidal bulges occur on both sides. When the Moon is at first quarter or last quarter (at right angles to the Sun's direction), the tides produced by the Sun partially cancel the tides of the Moon, making them lower than usual. These are called neap tides.

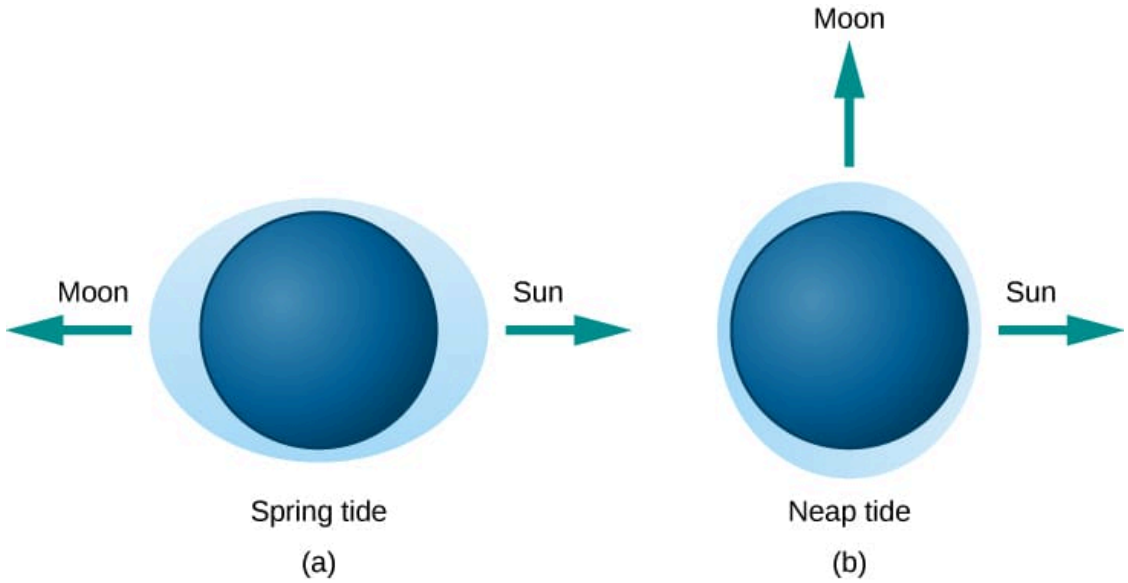


Figure 4.19 : Tides Caused by Different Alignments of the Sun and Moon. (a) In spring tides, the Sun's and Moon's pulls reinforce each other. (b) In neap tides, the Sun and the Moon pull at right angles to each other and the resulting tides are lower than usual.

The "simple" theory of tides, described in the preceding paragraphs, would be sufficient if Earth rotated very slowly and were completely surrounded by very deep oceans. However, the presence of land masses stopping the flow of water, the friction in the oceans and between oceans and the ocean floors, the rotation of Earth, the wind, the variable depth of the ocean, and other factors all complicate the picture. This is why, in the real world, some places have very small tides while in other places huge tides become tourist attractions. If you have been in such places, you may know that "tide tables" need to be computed and published for each location; one set of tide predictions doesn't work for the whole planet. In this introductory chapter, we won't delve further into these complexities.

Voyagers in Astronomy: George Darwin and the Slowing of Earth

The rubbing of water over the face of Earth involves an enormous amount of energy. Over long periods of time, the friction of the tides is slowing down the rotation of Earth. Our day gets longer by about 0.002 second each century. That seems very small, but such tiny changes can add up over millions and billions of years.

Although Earth's spin is slowing down, the angular momentum (see Orbits and Gravity) in a system such as the Earth-Moon system cannot change. Thus, some other spin motion must speed up to take the extra angular momentum. The details of what happens were worked out over a century ago by George Darwin, the son of naturalist Charles Darwin. George Darwin (see Figure 4.20) had a strong interest in science but studied law for six years and was admitted to the bar. However, he never practiced law, returning to science instead and eventually becoming a professor at Cambridge University. He was a protégé of Lord Kelvin, one of the great physicists of the nineteenth century, and he became interested in the long-term evolution of the solar system. He specialized in making detailed (and difficult) mathematical calculations of how orbits and motions change over geologic time.

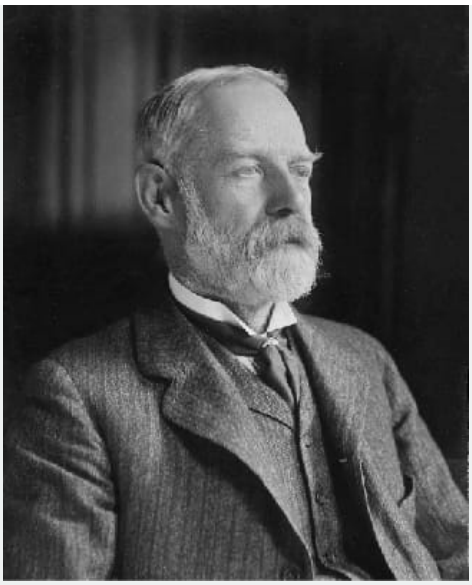


Figure 4.20 : George Darwin (1845–1912). George Darwin is best known for studying Earth's spin in relation to angular momentum.

What Darwin calculated for the Earth-Moon system was that the Moon will slowly spiral outward, away from Earth. As it moves farther away, it will orbit less quickly (just as planets farther from the Sun move more slowly in their orbits). Thus, the month will get longer. Also, because the Moon will be more distant, total eclipses of the Sun will no longer be visible from Earth.

Both the day and the month will continue to get longer, although bear in mind that the effects are very gradual. Darwin's calculations were confirmed by mirrors placed on the Moon by Apollo 11 astronauts. These show that the Moon is moving away by 3.8 centimeters per year, and that ultimately—billions of years in the future—the day and the month will be the same length (about 47 of our present days). At this point the Moon will be stationary in the sky over the same spot on Earth, meaning some parts of Earth will see the Moon and its phases and other parts will never see them. This kind of alignment is already true for Pluto's moon Charon (among others). Its rotation and orbital period are the same length as a day on Pluto.

This page titled 8.1: Ocean Tides and the Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.7: Ocean Tides and the Moon by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

8.2: Eclipses of the Sun and Moon

Learning Objectives

By the end of this section, you will be able to:

- Describe what causes lunar and solar eclipses
- Differentiate between a total and partial solar eclipse
- Explain why lunar eclipses are much more common than solar eclipses

One of the coincidences of living on Earth at the present time is that the two most prominent astronomical objects, the Sun and the Moon, have nearly the same apparent size in the sky. Although the Sun is about 400 times larger in diameter than the Moon, it is also about 400 times farther away, so both the Sun and the Moon have the same angular size—about 1/2°. As a result, the Moon, as seen from Earth, can appear to cover the Sun, producing one of the most impressive events in nature.

Any solid object in the solar system casts a shadow by blocking the light of the Sun from a region behind it. This shadow in space becomes apparent whenever another object moves into it. In general, an *eclipse* occurs whenever any part of either Earth or the Moon enters the shadow of the other. When the Moon's shadow strikes Earth, people within that shadow see the Sun at least partially covered by the Moon; that is, they witness a solar eclipse. When the Moon passes into the shadow of Earth, people on the night side of Earth see the Moon darken in what is called a lunar eclipse. Let's look at how these happen in more detail.

The shadows of Earth and the Moon consist of two parts: a cone where the shadow is darkest, called the *umbra*, and a lighter, more diffuse region of darkness called the *penumbra*. As you can imagine, the most spectacular eclipses occur when an object enters the umbra. Figure 4.21 illustrates the appearance of the Moon's shadow and what the Sun and Moon would look like from different points within the shadow.

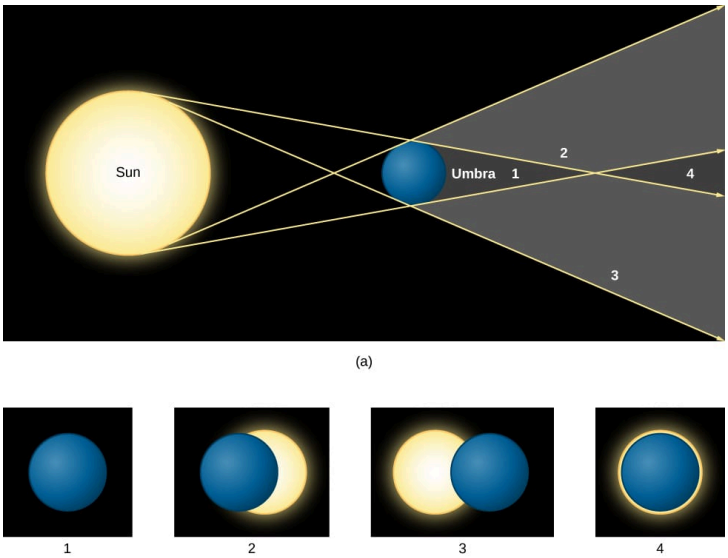


Figure 4.21 : Solar Eclipse. (a) The shadow cast by a spherical body (the Moon, for example) is shown. Notice the dark umbra and the lighter penumbra. Four points in the shadow are labeled with numbers. In (b) you see what the Sun and Moon would look like in the sky at the four labeled points. At position 1, you see a total eclipse. At positions 2 and 3, the eclipse is partial. At position 4, the Moon is farther away and thus cannot cover the Sun completely; a ring of light thus shows around the Sun, creating what is called an "annular" eclipse.

If the path of the Moon in the sky were identical to the path of the Sun (the ecliptic), we might expect to see an eclipse of the Sun and the Moon each month—whenever the Moon got in front of the Sun or into the shadow of Earth. However, as we mentioned, the Moon's orbit is tilted relative to the plane of Earth's orbit about the Sun by about 5° (imagine two hula hoops with a common center, but tilted a bit). As a result, during most months, the Moon is sufficiently above or below the ecliptic plane to avoid an eclipse. But when the two paths cross (twice a year), it is then "eclipse season" and eclipses are possible.

Eclipses of the Sun

The apparent or angular sizes of both the Sun and Moon vary slightly from time to time as their distances from Earth vary. (Figure 4.21 shows the distance of the observer varying at points 1–4, but the idea is the same.) Much of the time, the Moon looks slightly smaller than the Sun and cannot cover it completely, even if the two are perfectly aligned. In this type of “annular eclipse,” there is a ring of light around the dark sphere of the Moon.

However, if an eclipse of the Sun occurs when the Moon is somewhat nearer than its average distance, the Moon can completely hide the Sun, producing a *total* solar eclipse. Another way to say it is that a total eclipse of the Sun occurs at those times when the umbra of the Moon’s shadow reaches the surface of Earth.

The geometry of a total solar eclipse is illustrated in Figure 4.22. If the Sun and Moon are properly aligned, then the Moon’s darkest shadow intersects the ground at a small point on Earth’s surface. Anyone on Earth within the small area covered by the tip of the Moon’s shadow will, for a few minutes, be unable to see the Sun and will witness a total eclipse. At the same time, observers on a larger area of Earth’s surface who are in the penumbra will see only a part of the Sun eclipsed by the Moon: we call this a *partial* solar eclipse.

Between Earth’s rotation and the motion of the Moon in its orbit, the tip of the Moon’s shadow sweeps eastward at about 1500 kilometers per hour along a thin band across the surface of Earth. The thin zone across Earth within which a total solar eclipse is visible (weather permitting) is called the eclipse path. Within a region about 3000 kilometers on either side of the eclipse path, a partial solar eclipse is visible. It does not take long for the Moon’s shadow to sweep past a given point on Earth. The duration of totality may be only a brief instant; it can never exceed about 7 minutes.

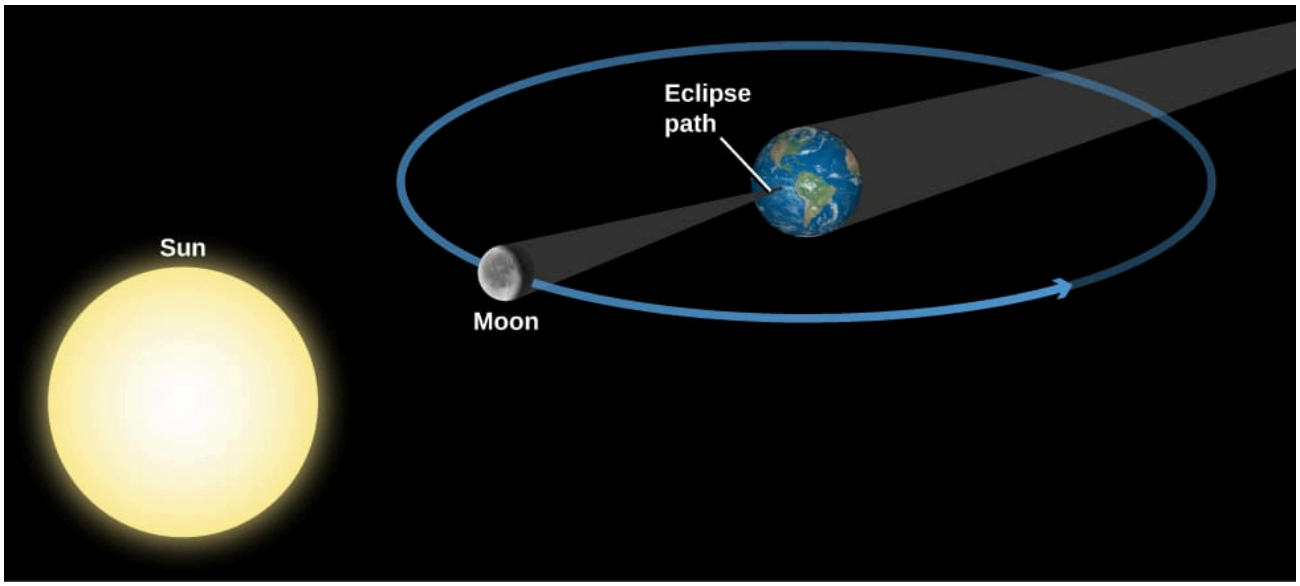


Figure 4.22 : Geometry of a Total Solar Eclipse. Note that our diagram is not to scale. The Moon blocks the Sun during new moon phase as seen from some parts of Earth and casts a shadow on our planet.

Because a total eclipse of the Sun is so spectacular, it is well worth trying to see one if you can. There are some people whose hobby is “eclipse chasing” and who brag about how many they have seen in their lifetimes. Because much of Earth’s surface is water, eclipse chasing can involve lengthy boat trips (and often requires air travel as well). As a result, eclipse chasing is rarely within the budget of a typical college student. Nevertheless, a list of future eclipses is given for your reference in Appendix H, just in case you strike it rich early. (And, as you can see in the Appendix, there was a total eclipse visible in the United States in 2017, and there will be another in 2024, to which college students may be able to afford travel.)

Appearance of a Total Eclipse

What can you see if you are lucky enough to catch a total eclipse? A solar eclipse starts when the Moon just begins to silhouette itself against the edge of the Sun’s disk. A partial phase follows, during which more and more of the Sun is covered by the Moon. About an hour after the eclipse begins, the Sun becomes completely hidden behind the Moon. In the few minutes immediately

before this period of totality begins, the sky noticeably darkens, some flowers close up, and chickens may go to roost. As an eerie twilight suddenly descends during the day, other animals (and people) may get disoriented. During totality, the sky is dark enough that planets become visible in the sky, and usually the brighter stars do as well.

As the bright disk of the Sun becomes entirely hidden behind the Moon, the Sun's remarkable corona flashes into view (Figure 4.23). The *corona* is the Sun's outer atmosphere, consisting of sparse gases that extend for millions of miles in all directions from the apparent surface of the Sun. It is ordinarily not visible because the light of the corona is feeble compared with the light from the underlying layers of the Sun. Only when the brilliant glare from the Sun's visible disk is blotted out by the Moon during a total eclipse is the pearly white corona visible. (We'll talk more about the corona in the chapter on The Sun: A Garden-Variety Star.)

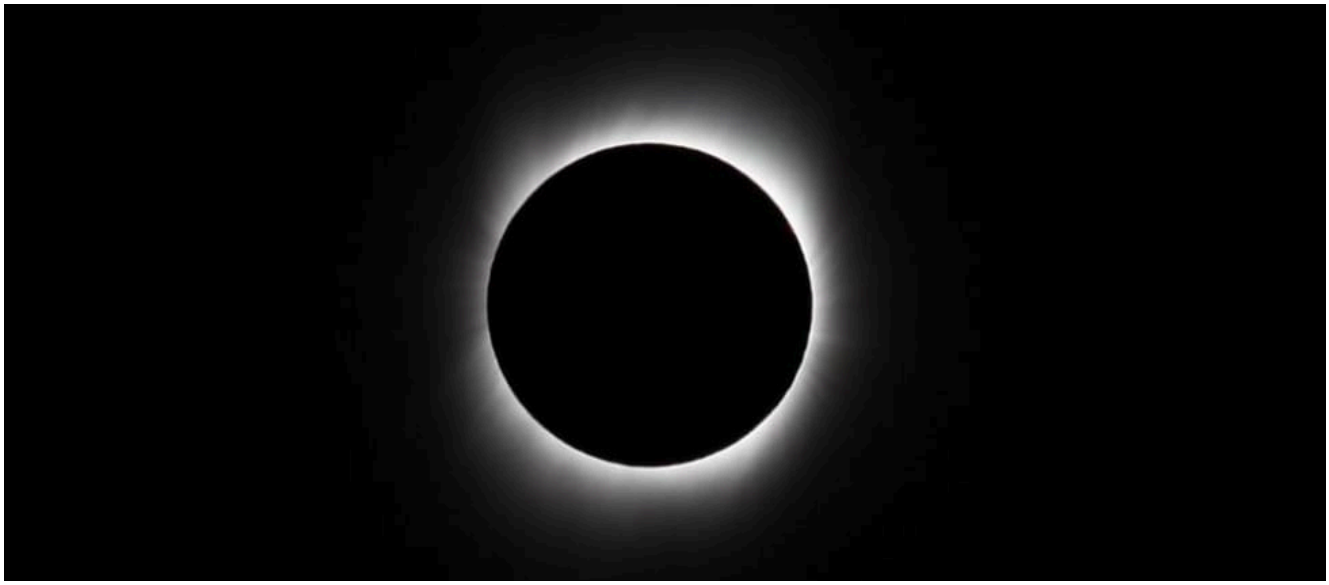


Figure 4.23 : The Sun's Corona. The corona (thin outer atmosphere) of the Sun is visible during a total solar eclipse. (It looks more extensive in photographs than it would to the unaided eye.) (credit: modification of work by Lutfar Rahman Nirjhar)

The total phase of the eclipse ends, as abruptly as it began, when the Moon begins to uncover the Sun. Gradually, the partial phases of the eclipse repeat themselves, in reverse order, until the Moon has completely uncovered the Sun. We should make one important safety point here: while the few minutes of the *total* eclipse are safe to look at, if any part of the Sun is uncovered, you must protect your eyes with safe eclipse glasses[2] or by projecting an image of the Sun (instead of looking at it directly). For more, read the How to Observe Solar Eclipses box in this chapter.

Eclipses of the Moon

A lunar eclipse occurs when the Moon enters the shadow of Earth. The geometry of a lunar eclipse is shown in Figure 4.24. Earth's dark shadow is about 1.4 million kilometers long, so at the Moon's distance (an average of 384,000 kilometers), it could cover about four full moons. Unlike a solar eclipse, which is visible only in certain local areas on Earth, a lunar eclipse is visible to everyone who can see the Moon. Because a lunar eclipse can be seen (weather permitting) from the entire night side of Earth, lunar eclipses are observed far more frequently from a given place on Earth than are solar eclipses.

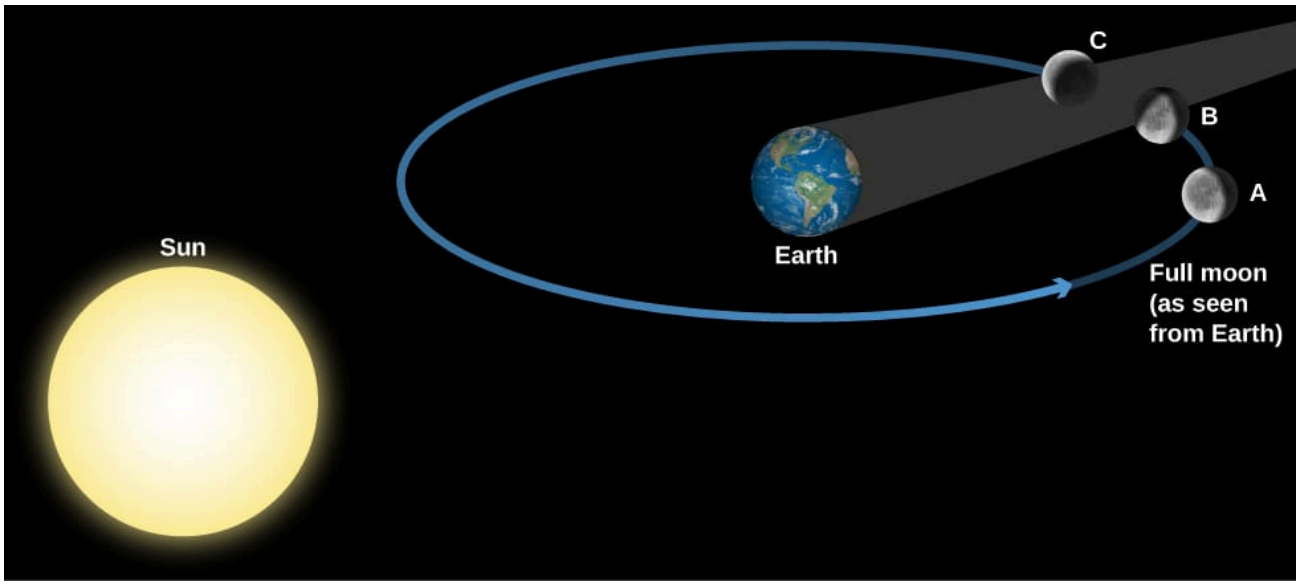


Figure 4.24 : Geometry of a Lunar Eclipse. The Moon is shown moving through the different parts of Earth's shadow during a total lunar eclipse. Note that the distance the Moon moves in its orbit during the eclipse has been exaggerated here for clarity.

An eclipse of the Moon is total only if the Moon's path carries it though Earth's umbra. If the Moon does not enter the umbra completely, we have a partial eclipse of the Moon. But because Earth is larger than the Moon, its umbra is larger, so that lunar eclipses last longer than solar eclipses, as we will discuss below.

A lunar eclipse can take place only when the Sun, Earth, and Moon are in a line. The Moon is opposite the Sun, which means the Moon will be in full phase before the eclipse, making the darkening even more dramatic. About 20 minutes before the Moon reaches the dark shadow, it dims somewhat as Earth partly blocks the sunlight. As the Moon begins to dip into the shadow, the curved shape of Earth's shadow upon it soon becomes apparent.

Even when totally eclipsed, the Moon is still faintly visible, usually appearing a dull coppery red. The illumination on the eclipsed Moon is sunlight that has been bent into Earth's shadow by passing through Earth's atmosphere.

After totality, the Moon moves out of the shadow and the sequence of events is reversed. The total duration of the eclipse depends on how closely the Moon's path approaches the axis of the shadow. For an eclipse where the Moon goes through the center of Earth's shadow, each partial phase consumes at least 1 hour, and totality can last as long as 1 hour and 40 minutes. Eclipses of the Moon are much more "democratic" than solar eclipses. Since the full moon is visible on the entire night side of Earth, the lunar eclipse is visible for all those who live in that hemisphere. (Recall that a total eclipse of the Sun is visible only in a narrow path where the shadow of the umbra falls.) Total eclipses of the Moon occur, on average, about once every two or three years. A list of future total eclipses of the Moon is in Appendix H. In addition, since the lunar eclipse happens to a full moon, and a full moon is not dangerous to look at, everyone can look at the Moon during all the parts of the eclipse without worrying about safety.

Thanks to our understanding of gravity and motion (see Orbits and Gravity), eclipses can now be predicted centuries in advance. We've come a long way since humanity stood frightened by the darkening of the Sun or the Moon, fearing the displeasure of the gods. Today, we enjoy the sky show with a healthy appreciation of the majestic forces that keep our solar system running.

Use the Solar and Lunar Eclipse simulation to explore the conditions of solar and lunar eclipses yourself.

Seeing for Yourself: How to Observe Solar Eclipses

A total eclipse of the Sun is a spectacular sight and should not be missed. However, it is extremely dangerous to look directly at the Sun: even a brief exposure can damage your eyes. Normally, few rational people are tempted to do this because it is painful (and something your mother told you never to do!). But during the partial phases of a solar eclipse, the temptation to

take a look is strong. Think before you give in. The fact that the Moon is covering part of the Sun doesn't make the uncovered part any less dangerous to look at. Still, there are perfectly safe ways to follow the course of a solar eclipse, if you are lucky enough to be in the path of the shadow.

The easiest technique is to make a pinhole projector. Take a piece of cardboard with a small (1 millimeter) hole punched in it, and hold it several feet above a light surface, such as a concrete sidewalk or a white sheet of paper, so that the hole is "aimed" at the Sun. The hole produces a fuzzy but adequate image of the eclipsed Sun. Alternatively, if it's the right time of year, you can let the tiny spaces between a tree's leaves form multiple pinhole images against a wall or sidewalk. Watching hundreds of little crescent Suns dancing in the breeze can be captivating. A kitchen colander also makes an excellent pinhole projector.

Although there are safe filters for looking at the Sun directly, people have suffered eye damage by looking through improper filters, or no filter at all. For example, neutral density photographic filters are not safe because they transmit infrared radiation that can cause severe damage to the retina. Also unsafe are smoked glass, completely exposed color film, sunglasses, and many other homemade filters. Safe filters include welders' goggles and specially designed eclipse glasses.

You should certainly look at the Sun directly when it is totally eclipsed, even through binoculars or telescopes. Unfortunately, the total phase, as we discussed, is all too brief. But if you know when it is coming and going, be sure you look, for it's an unforgettably beautiful sight. And, despite the ancient folklore that presents eclipses as dangerous times to be outdoors, the partial phases of eclipses—as long as you are not looking directly at the Sun—are not any more dangerous than being out in sunlight.

During past eclipses, unnecessary panic has been created by uninformed public officials acting with the best intentions. There were two marvelous total eclipses in Australia in the twentieth century during which townspeople held newspapers over their heads for protection and schoolchildren cowered indoors with their heads under their desks. What a pity that all those people missed what would have been one of the most memorable experiences of their lifetimes.

On August 21, 2017, a total eclipse of the Sun was visible across a large swath of the continental United States, and was seen by millions of people from around the country and the world (a).



Figure 4.25 : 2017 Total Solar and Lunar Eclipse (a) The eclipsed Sun August 21, 2017, showing remarkable detail in the Sun's outer atmosphere. This is a composite of short, medium, and long exposures, as no single exposure can capture the huge range of brightness the Sun exhibits. (b) A total eclipse of the Moon seen over California on January 31, 2018. The Moon moves slowly into the Earth's shadow, looks red when the eclipse is total and red sunlight refracts through the Earth's atmosphere, and then slowly moves out of the shadow. (credit a: modification of work by Rick Fienberg, American Astronomical Socity/TravelQuest International; credit b: modification of work by Brian Day.)

Link to Learning

For the August 2017 eclipse, the United States postal service issued a special commemorative stamp — the first ever with "thermochromic ink" which changes when you touch it.

Footnotes

- 2Eclipse glasses are available in many planetarium and observatory gift stores, and also from the two main U.S. manufacturers: American Paper Optics and Rainbow Symphony.

This page titled 8.2: Eclipses of the Sun and Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 4.8: Eclipses of the Sun and Moon by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

8.3: Key Terms

apparent solar time
time as measured by the position of the Sun in the sky (the time that would be indicated by a sundial)

declination
the angular distance north or south of the celestial equator

great circle
a circle on the surface of a sphere that is the curve of intersection of the sphere with a plane passing through its center

International Date Line
an arbitrary line on the surface of Earth near longitude 180° across which the date changes by one day

lunar eclipse
an eclipse of the Moon, in which the Moon moves into the shadow of Earth; lunar eclipses can occur only at the time of full moon

mean solar time
time based on the rotation of Earth; mean solar time passes at a constant rate, unlike apparent solar time

meridian
a great circle on the terrestrial or celestial sphere that passes through the poles

phases of the Moon
the different appearance of light and dark on the Moon as seen from Earth during its monthly cycle, from new moon to full moon and back to new moon

right ascension
the coordinate for measuring the east-west positions of celestial bodies; the angle measured eastward along the celestial equator from the vernal equinox to the hour circle passing through a body

sidereal day
Earth's rotation period as defined by the positions of the stars in the sky; the time between successive passages of the same star through the meridian

sidereal month
the period of the Moon's revolution about Earth measured with respect to the stars

solar day
Earth's rotation period as defined by the position of the Sun in the sky; the time between successive passages of the Sun through the meridian

solar eclipse
an eclipse of the Sun by the Moon, caused by the passage of the Moon in front of the Sun; solar eclipses can occur only at the time of the new moon

solar month
the time interval in which the phases repeat—say, from full to full phase

synchronous rotation
when a body (for example, the Moon) rotates at the same rate that it revolves around another body

tides
alternate rising and falling of sea level caused by the difference in the strength of the Moon's gravitational pull on different parts of Earth

This page titled 8.3: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

9: LIGHT

9.1: The Behavior of Light
9.2: The Electromagnetic Spectrum
9.3: Spectroscopy in Astronomy
9.4: The Structure of the Atom
9.5: Formation of Spectral Lines
9.6: The Doppler Effect
9.7: Key Terms

This page titled 9: Light is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

9.1: The Behavior of Light

Learning Objectives

By the end of this section, you will be able to:

- Explain the evidence for Maxwell's electromagnetic model of light
- Describe the relationship between wavelength, frequency, and speed of light
- Discuss the particle model of light and the definition of photon
- Explain how and why the amount of light we see from an object depends upon its distance

Coded into the light and other kinds of radiation that reach us from objects in the universe is a wide range of information about what those objects are like and how they work. If we can decipher this code and read the messages it contains, we can learn an enormous amount about the cosmos without ever having to leave Earth or its immediate environment.

The visible light and other radiation we receive from the stars and planets is generated by processes at the atomic level—by changes in the way the parts of an atom interact and move. Thus, to appreciate how light is generated, we must explore how atoms work. There is a bit of irony in the fact that in order to understand some of the largest structures in the universe, we must become acquainted with some of the smallest.

Notice that we have twice used the phrase "light and other radiation." One of the key ideas explored in this chapter is that visible light is not unique; it is merely the most familiar example of a much larger family of radiation that can carry information to us.

The word "radiation" will be used frequently in this book, so it is important to understand what it means. In everyday language, "radiation" is often used to describe certain kinds of energetic subatomic particles released by radioactive materials in our environment. (An example is the kind of radiation used to treat some cancers.) But this is not what we mean when we use the word "radiation" in an astronomy text. *Radiation*, as used in this book, is a general term for waves (including light waves) that *radiate* outward from a source.

As we saw in Orbits and Gravity, Newton's theory of gravity accounts for the motions of planets as well as objects on Earth. Application of this theory to a variety of problems dominated the work of scientists for nearly two centuries. In the nineteenth century, many physicists turned to the study of electricity and magnetism, which are intimately connected with the production of light.

The scientist who played a role in this field comparable to Newton's role in the study of gravity was physicist James Clerk Maxwell, born and educated in Scotland (Figure 5.2). Inspired by a number of ingenious experiments that showed an intimate relationship between electricity and magnetism, Maxwell developed a theory that describes both electricity and magnetism with only a small number of elegant equations. It is this theory that gives us important insights into the nature and behavior of light.



Figure 5.2 : James Clerk Maxwell (1831–1879). Maxwell unified the rules governing electricity and magnetism into a coherent theory.

Maxwell's Theory of Electromagnetism

We will look at the structure of the atom in more detail later, but we begin by noting that the typical atom consists of several types of particles, a number of which have not only mass but an additional property called electric charge. In the nucleus (central part) of every atom are *protons*, which are positively charged; outside the nucleus are electrons, which have a negative charge.

Maxwell's theory deals with these electric charges and their effects, especially when they are moving. In the vicinity of an electron charge, another charge feels a force of attraction or repulsion: opposite charges attract; like charges repel. When charges are not in motion, we observe only this electric attraction or repulsion. If charges are in motion, however (as they are inside every atom and in a wire carrying a current), then we measure another force called *magnetism*.

Magnetism was well known for much of recorded human history, but its cause was not understood until the nineteenth century. Experiments with electric charges demonstrated that magnetism was the result of moving charged particles. Sometimes, the motion is clear, as in the coils of heavy wire that make an industrial electromagnet. Other times, it is more subtle, as in the kind of magnet you buy in a hardware store, in which many of the electrons inside the atoms are spinning in roughly the same direction; it is the alignment of their motion that causes the material to become magnetic.

Physicists use the word *field* to describe the action of forces that one object exerts on other distant objects. For example, we say the Sun produces a *gravitational field* that controls Earth's orbit, even though the Sun and Earth do not come directly into contact. Using this terminology, we can say that stationary electric charges produce *electric fields*, and moving electric charges also produce *magnetic fields*.

Actually, the relationship between electric and magnetic phenomena is even more profound. Experiments showed that changing magnetic fields could produce electric currents (and thus changing electric fields), and changing electric currents could in turn produce changing magnetic fields. So once begun, electric and magnetic field changes could continue to trigger each other.

Maxwell analyzed what would happen if electric charges were oscillating (moving constantly back and forth) and found that the resulting pattern of electric and magnetic fields would spread out and travel rapidly through space. Something similar happens when a raindrop strikes the surface of water or a frog jumps into a pond. The disturbance moves outward and creates a pattern we call a *wave* in the water (Figure 5.3). You might, at first, think that there must be very few situations in nature where electric charges oscillate, but this is not at all the case. As we shall see, atoms and molecules (which consist of charged particles) oscillate back and forth all the time. The resulting electromagnetic disturbances are among the most common phenomena in the universe.



Figure 5.3 : Making Waves. An oscillation in a pool of water creates an expanding disturbance called a wave. (credit: modification of work by "vastateparksstaff"/Flickr)

Maxwell was able to calculate the speed at which an electromagnetic disturbance moves through space; he found that it is equal to the speed of light, which had been measured experimentally. On that basis, he speculated that light was one form of a family of possible electromagnetic disturbances called electromagnetic radiation, a conclusion that was again confirmed in laboratory experiments. When light (reflected from the pages of an astronomy textbook, for example) enters a human eye, its changing electric and magnetic fields stimulate nerve endings, which then transmit the information contained in these changing fields to the brain. The science of astronomy is primarily about analyzing radiation from distant objects to understand what they are and how they work.

The Wave-Like Characteristics of Light

The changing electric and magnetic fields in light are similar to the waves that can be set up in a quiet pool of water. In both cases, the disturbance travels rapidly outward from the point of origin and can use its energy to disturb other things farther away. (For example, in water, the expanding ripples moving away from our frog could disturb the peace of a dragonfly resting on a leaf in the same pool.) In the case of electromagnetic waves, the radiation generated by a transmitting antenna full of charged particles and moving electrons at your local radio station can, sometime later, disturb a group of electrons in your car radio antenna and bring you the news and weather while you are driving to class or work in the morning.

The waves generated by charged particles differ from water waves in some profound ways, however. Water waves require water to travel in. The sound waves we hear, to give another example, are pressure disturbances that require air to travel though. But electromagnetic waves do not require water or air: the fields generate each other and so can move through a vacuum (such as outer space). This was such a disturbing idea to nineteenth-century scientists that they actually made up a substance to fill all of space—one for which there was not a single shred of evidence—just so light waves could have something to travel through: they called it the *aether*. Today, we know that there is no aether and that electromagnetic waves have no trouble at all moving through empty space (as all the starlight visible on a clear night must surely be doing).

The other difference is that *all* electromagnetic waves move at the same speed in empty space (the speed of light—approximately 300,000 kilometers per second, or 300,000,000 meters per second, which can also be written as $3 \times 10^8 \mathrm{m}/$), which turns out to be the fastest possible speed in the universe. No matter where electromagnetic waves are generated from and no matter what other properties they have, when they are moving (and not interacting with matter), they move at the speed of light. Yet you know from everyday experience that there are different kinds of light. For example, we perceive that light waves differ from one another in a property we call color. Let's see how we can denote the differences among the whole broad family of electromagnetic waves.

The nice thing about a wave is that it is a repeating phenomenon. Whether it is the up-and-down motion of a water wave or the changing electric and magnetic fields in a wave of light, the pattern of disturbance repeats in a cyclical way. Thus, any wave motion can be characterized by a series of crests and troughs (Figure 5.4). Moving from one crest through a trough to the next crest completes one cycle. The horizontal length covered by one cycle is called the wavelength. Ocean waves provide an analogy:

the wavelength is the distance that separates successive wave crests.

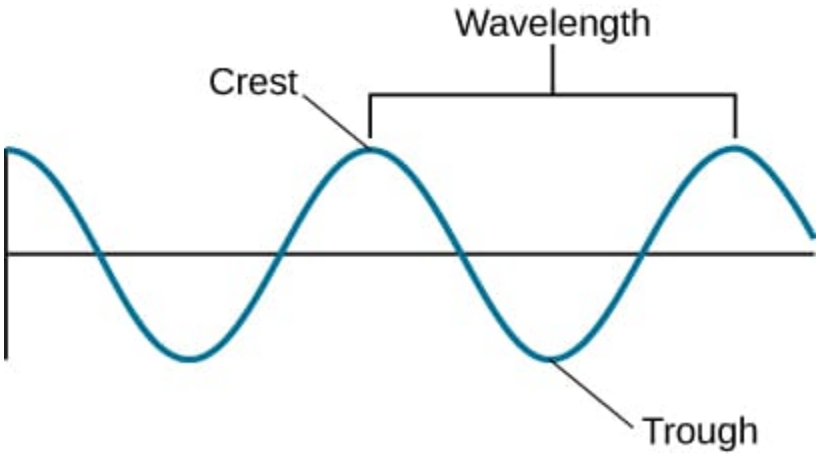


Figure 5.4 : Characterizing Waves. Electromagnetic radiation has wave-like characteristics. The wavelength (λ) is the distance between crests, the frequency (*f*) is the number of cycles per second, and the speed (*c*) is the distance the wave covers during a specified period of time (e.g., kilometers per second).

For visible light, our eyes perceive different wavelengths as different colors: red, for example, is the longest visible wavelength, and violet is the shortest. The main colors of visible light from longest to shortest wavelength can be remembered using the mnemonic ROY G BIV—for Red, Orange, Yellow, Green, Blue, Indigo, and Violet. Other invisible forms of electromagnetic radiation have different wavelengths, as we will see in the next section.

We can also characterize different waves by their frequency, the number of wave cycles that pass by per second. If you count 10 crests moving by each second, for example, then the frequency is 10 cycles per second (cps). In honor of Heinrich Hertz, the physicist who—inspired by Maxwell's work—discovered radio waves, a cps is also called a *hertz* (Hz). Take a look at your radio, for example, and you will see the channel assigned to each radio station is characterized by its frequency, usually in units of KHz (kilohertz, or thousands of hertz) or MHz (megahertz, or millions of hertz).

Wavelength (λ) and frequency (*f*) are related because all electromagnetic waves travel at the same speed. To see how this works, imagine a parade in which everyone is forced by prevailing traffic conditions to move at exactly the same speed. You stand on a corner and watch the waves of marchers come by. First you see row after row of miniature ponies. Because they are not very large and, therefore, have a shorter wavelength, a good number of the ponies can move past you each minute; we can say they have a high frequency. Next, however, come several rows of circus elephants. The elephants are large and marching at the same speed as the ponies, so far fewer of them can march past you per minute: Because they have a wider spacing (longer wavelength), they represent a lower frequency.

The formula for this relationship can be expressed as follows: for any wave motion, the speed at which a wave moves equals the frequency times the wavelength. Waves with longer wavelengths have lower frequencies. Mathematically, we can express this as

$$c = \lambda f$$

where the Greek letter for "l"—lambda, λ—is used to denote wavelength and *c* is the scientific symbol for the speed of light. Solving for the wavelength, this is expressed as:

$$\lambda = \frac{c}{f}.$$

Example 5.1: Deriving and Using the Wave Equation

The equation for the relationship between the speed and other characteristics of a wave can be derived from our basic understanding of motion. The average speed of anything that is moving is:

$$\text{average speed} = \frac{\text{distance}}{\text{time}}$$

(So, for example, a car on the highway traveling at a speed of 100 km/h covers 100 km during the time of 1 h.) For an electromagnetic wave to travel the distance of one of its wavelengths, λ, at the speed of light, *c*, we have $c = \frac{\lambda}{t}$. The frequency of a wave is the number of cycles per second. If a wave has a frequency of a million cycles per second, then the time for each cycle to go by is a millionth of a second. So, in general, $t = \frac{1}{f}$. Substituting into our wave equation, we get $c = \lambda \times f$. Now let's use this to calculate an example. What is the wavelength of visible light that has a frequency of $5.66 \times 10^{14}\text{Hz}$?

Solution

Solving the wave equation for wavelength, we find:

$$\lambda = \frac{c}{f}$$

Substituting our values gives:

$$\lambda = \frac{3.00 \times 10^{8}\,\text{m/s}}{5.66 \times 10^{14}\,\text{Hz}} = 5.30 \times 10^{-7}\,\text{m}$$

This answer can also be written as 530 nm, which is in the yellow-green part of the visible spectrum (nm stands for nanometers, where the term "nano" means "billionths").

Exercise 9.1.1

"Tidal waves," or tsunamis, are waves caused by earthquakes that travel rapidly through the ocean. If a tsunami travels at the speed of 600 km/h and approaches a shore at a rate of one wave crest every 15 min (4 waves/h), what would be the distance between those wave crests at sea?

Answer available at the end of this section.

Light as a Photon

The electromagnetic wave model of light (as formulated by Maxwell) was one of the great triumphs of nineteenth-century science. In 1887, when Heinrich Hertz actually made invisible electromagnetic waves (what today are called radio waves) on one side of a room and detected them on the other side, it ushered in a new era that led to the modern age of telecommunications. His experiment ultimately led to the technologies of television, cell phones, and today's wireless networks around the globe.

However, by the beginning of the twentieth century, more sophisticated experiments had revealed that light behaves in certain ways that cannot be explained by the wave model. Reluctantly, physicists had to accept that sometimes light behaves more like a "particle"—or at least a self-contained packet of energy—than a wave. We call such a packet of electromagnetic energy a photon.

The fact that light behaves like a wave in certain experiments and like a particle in others was a very surprising and unlikely idea. After all, our common sense says that waves and particles are opposite concepts. On one hand, a wave is a repeating disturbance that, by its very nature, is not in only one place, but spreads out. A particle, on the other hand, is something that can be in only one place at any given time. Strange as it sounds, though, countless experiments now confirm that electromagnetic radiation can sometimes behave like a wave and at other times like a particle.

Then, again, perhaps we shouldn't be surprised that something that always travels at the "speed limit" of the universe and doesn't need a medium to travel through might not obey our everyday common sense ideas. The confusion that this wave-particle duality of light caused in physics was eventually resolved by the introduction of a more complicated theory of waves and particles, now called quantum mechanics. (This is one of the most interesting fields of modern science, but it is mostly beyond the scope of our book. If you are interested in it, see some of the suggested resources at the end of this chapter.)

In any case, you should now be prepared when scientists (or the authors of this book) sometimes discuss electromagnetic radiation as if it consisted of waves and at other times refer to it as a stream of photons. A photon (being a packet of energy) carries a specific amount of energy. We can use the idea of energy to connect the photon and wave models. How much energy a photon has depends on its frequency when you think about it as a wave. A low-energy radio wave has a low frequency as a wave, while a high-energy X-ray at your dentist's office is a high-frequency wave. Among the colors of visible light, violet-light photons

have the highest energy and red-light photons have the lowest.

Test whether the connection between photons and waves is clear to you. In the above example, which photon would have the longer wavelength as a wave: the radio wave or the X-ray? If you answered the radio wave, you are correct. Radio waves have a lower frequency, so the wave cycles are longer (they are elephants, not miniature ponies).

Propagation of Light

Let's think for a moment about how light from a lightbulb moves through space. As waves expand, they travel away from the bulb, not just toward your eyes but in all directions. They must therefore cover an ever-widening space. Yet the total amount of light available can't change once the light has left the bulb. This means that, as the same expanding shell of light covers a larger and larger area, there must be less and less of it in any given place. Light (and all other electromagnetic radiation) gets weaker and weaker as it gets farther from its source.

The increase in the area that the light must cover is proportional to the square of the distance that the light has traveled (Figure 5.5). If we stand twice as far from the source, our eyes will intercept two-squared (2 × 2), or four times less light. If we stand 10 times farther from the source, we get 10-squared, or 100 times less light. You can see how this weakening means trouble for sources of light at astronomical distances. One of the nearest stars, Alpha Centauri A, emits about the same total energy as the Sun. But it is about 270,000 times farther away, and so it appears about 73 billion times fainter. No wonder the stars, which close-up would look more or less like the Sun, look like faint pinpoints of light from far away.

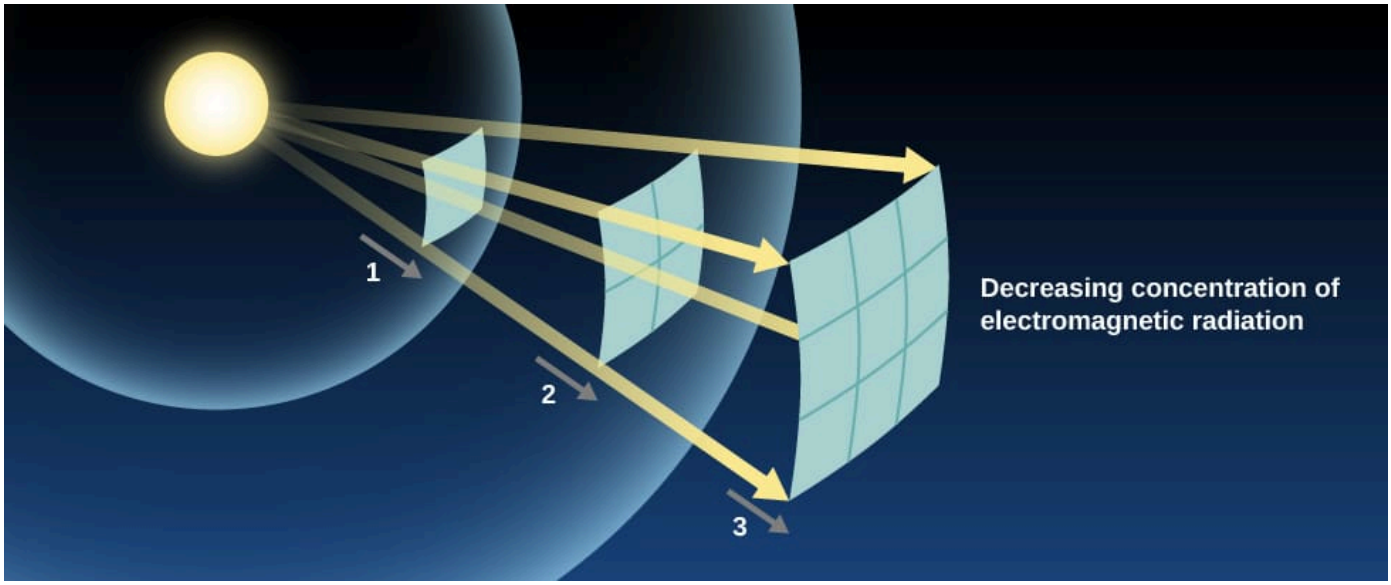


Figure 5.5 : Inverse Square Law for Light. As light radiates away from its source, it spreads out in such a way that the energy per unit area (the amount of energy passing through one of the small squares) decreases as the square of the distance from its source.

This idea—that the apparent brightness of a source (how bright it looks to us) gets weaker with distance in the way we have described—is known as the inverse square law for light propagation. In this respect, the propagation of light is similar to the effects of gravity. Remember that the force of gravity between two attracting masses is also inversely proportional to the square of their separation.

Example 5.2: The Inverse Square Law for Light

The intensity of a 120-W lightbulb observed from a distance 2 m away is 2.4 W/m^2. What would be the intensity if this distance was doubled?

Solution

If we move twice as far away, then the answer will change according to the inverse square of the distance, so the new intensity will be $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$ of the original intensity, or 0.6 W/m^2.

Exercise 9.1.1

How many times brighter or fainter would a star appear if it were moved to:

a. twice its present distance?
b. ten times its present distance?
c. half its present distance?

Answer available at the end of this section.

Answers

Exercise 9.1.1

$\lambda = \frac{600\ \text{km/h}}{4\ \text{waves/h}} = 150\ \text{km}$

Exercise 9.1.1

a. $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$; b. $\left(\frac{1}{10}\right)^2 = \frac{1}{100}$; c. $\left(\frac{1}{1/2}\right)^2 = 4$

This page titled 9.1: The Behavior of Light is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.2: The Behavior of Light by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.2: The Electromagnetic Spectrum

Learning Objectives

By the end of this section, you will be able to:

- Understand the bands of the electromagnetic spectrum and how they differ from one another
- Understand how each part of the spectrum interacts with Earth's atmosphere
- Explain how and why the light emitted by an object depends on its temperature

Objects in the universe send out an enormous range of electromagnetic radiation. Scientists call this range the electromagnetic spectrum, which they have divided into a number of categories. The spectrum is shown in Figure 5.6, with some information about the waves in each part or band.

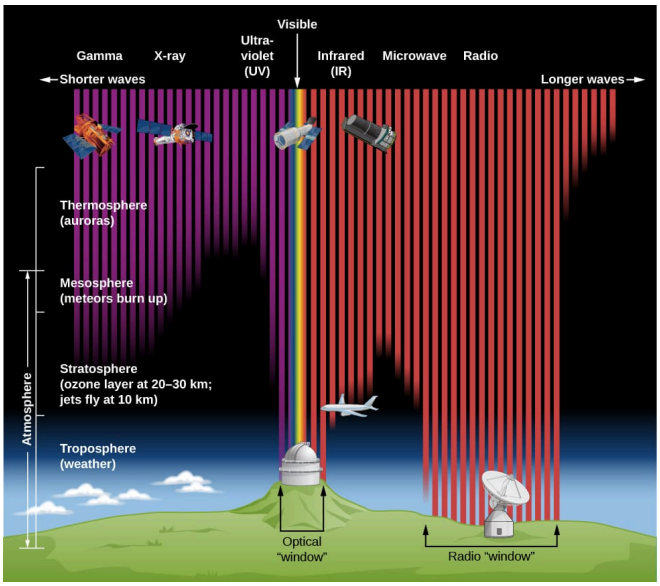


Figure 5.6 : Radiation and Earth's Atmosphere. This figure shows the bands of the electromagnetic spectrum and how well Earth's atmosphere transmits them. Note that high-frequency waves from space do not make it to the surface and must therefore be observed from space. Some infrared and microwaves are absorbed by water and thus are best observed from high altitudes. Low-frequency radio waves are blocked by Earth's ionosphere. (credit: modification of work by STScI/JHU/NASA)

Types of Electromagnetic Radiation

Electromagnetic radiation with the shortest wavelengths, no longer than 0.01 nanometer, is categorized as gamma rays (1 nanometer = 10^{-9} meters; see Appendix D). The name *gamma* comes from the third letter of the Greek alphabet: gamma rays were the third kind of radiation discovered coming from radioactive atoms when physicists first investigated their behavior. Because gamma rays carry a lot of energy, they can be dangerous for living tissues. Gamma radiation is generated deep in the interior of stars, as well as by some of the most violent phenomena in the universe, such as the deaths of stars and the merging of stellar corpses. Gamma rays coming to Earth are absorbed by our atmosphere before they reach the ground (which is a good thing for our health); thus, they can only be studied using instruments in space.

Electromagnetic radiation with wavelengths between 0.01 nanometer and 20 nanometers is referred to as X-rays. Being more energetic than visible light, X-rays are able to penetrate soft tissues but not bones, and so allow us to make images of the shadows of the bones inside us. While X-rays can penetrate a short length of human flesh, they are stopped by the large numbers of atoms in Earth's atmosphere with which they interact. Thus, X-ray astronomy (like gamma-ray astronomy) could not develop until we invented ways of sending instruments above our atmosphere (Figure 5.7).

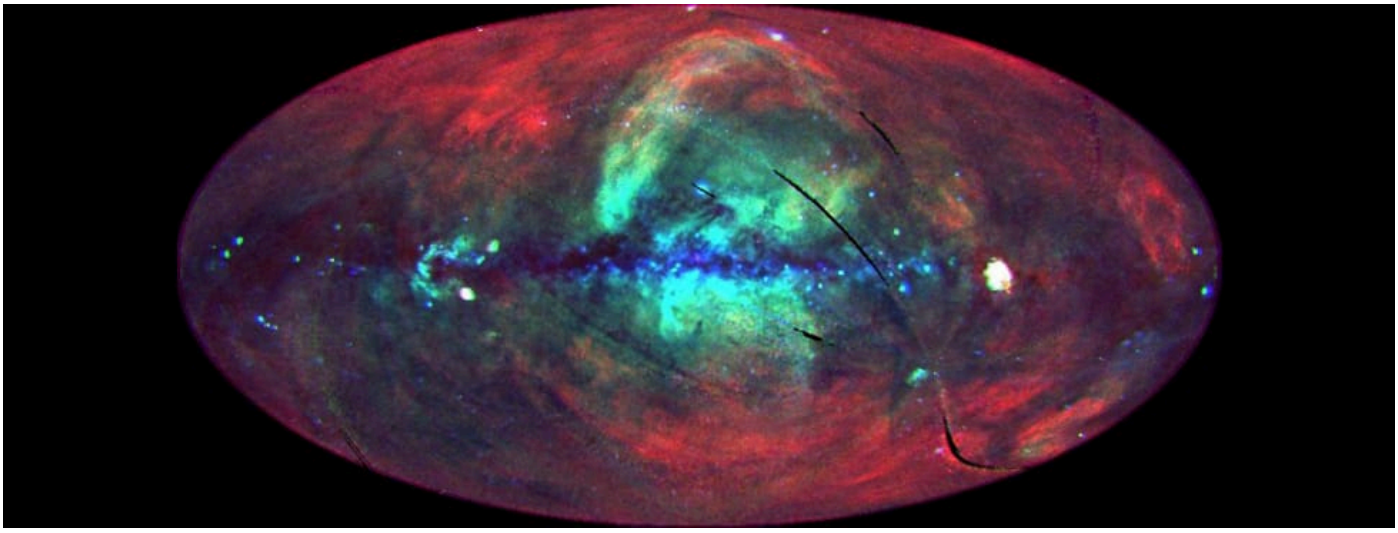


Figure 5.7 : X-Ray Sky. This is a map of the sky tuned to certain types of X-rays (seen from above Earth's atmosphere). The map tilts the sky so that the disk of our Milky Way Galaxy runs across its center. It was constructed and artificially colored from data gathered by the European ROSAT satellite. Each color (red, yellow, and blue) shows X-rays of different frequencies or energies. For example, red outlines the glow from a hot local bubble of gas all around us, blown by one or more exploding stars in our cosmic vicinity. Yellow and blue show more distant sources of X-rays, such as remnants of other exploded stars or the active center of our Galaxy (in the middle of the picture). (credit: modification of work by NASA)

Radiation intermediate between X-rays and visible light is ultraviolet (meaning higher energy than violet). Outside the world of science, ultraviolet light is sometimes called "black light" because our eyes cannot see it. Ultraviolet radiation is mostly blocked by the ozone layer of Earth's atmosphere, but a small fraction of ultraviolet rays from our Sun do penetrate to cause sunburn or, in extreme cases of overexposure, skin cancer in human beings. Ultraviolet astronomy is also best done from space.

Electromagnetic radiation with wavelengths between roughly 400 and 700 nm is called visible light because these are the waves that human vision can perceive. This is also the band of the electromagnetic spectrum that most readily reaches Earth's surface. These two observations are not coincidental: human eyes evolved to see the kinds of waves that arrive from the Sun most effectively. Visible light penetrates Earth's atmosphere effectively, except when it is temporarily blocked by clouds.

Between visible light and radio waves are the wavelengths of infrared or heat radiation. Astronomer William Herschel first discovered infrared in 1800 while trying to measure the temperatures of different colors of sunlight spread out into a spectrum. He noticed that when he accidently positioned his thermometer beyond the reddest color, it still registered heating due to some invisible energy coming from the Sun. This was the first hint about the existence of the other (invisible) bands of the electromagnetic spectrum, although it would take many decades for our full understanding to develop.

A heat lamp radiates mostly infrared radiation, and the nerve endings in our skin are sensitive to this band of the electromagnetic spectrum. Infrared waves are absorbed by water and carbon dioxide molecules, which are more concentrated low in Earth's atmosphere. For this reason, infrared astronomy is best done from high mountaintops, high-flying airplanes, and spacecraft.

After infrared comes the familiar microwave, used in short-wave communication and microwave ovens. (Wavelengths vary from 1 millimeter to 1 meter and are absorbed by water vapor, which makes them effective in heating foods.) The "micro-" prefix refers to the fact that microwaves are small in comparison to radio waves, the next on the spectrum. You may remember that tea—which is full of water—heats up quickly in your microwave oven, while a ceramic cup—from which water has been removed by baking—stays cool in comparison.

All electromagnetic waves longer than microwaves are called radio waves, but this is so broad a category that we generally divide it into several subsections. Among the most familiar of these are radar waves, which are used in radar guns by traffic officers to determine vehicle speeds, and AM radio waves, which were the first to be developed for broadcasting. The wavelengths of these different categories range from over a meter to hundreds of meters, and other radio radiation can have wavelengths as long as several kilometers.

With such a wide range of wavelengths, not all radio waves interact with Earth's atmosphere in the same way. FM and TV waves are not absorbed and can travel easily through our atmosphere. AM radio waves are absorbed or reflected by a layer in Earth's atmosphere called the ionosphere (the ionosphere is a layer of charged particles at the top of our atmosphere, produced by

interactions with sunlight and charged particles that are ejected from the Sun).

We hope this brief survey has left you with one strong impression: although visible light is what most people associate with astronomy, the light that our eyes can see is only a tiny fraction of the broad range of waves generated in the universe. Today, we understand that judging some astronomical phenomenon by using only the light we can see is like hiding under the table at a big dinner party and judging all the guests by nothing but their shoes. There's a lot more to each person than meets our eye under the table. It is very important for those who study astronomy today to avoid being "visible light chauvinists"—to respect only the information seen by their eyes while ignoring the information gathered by instruments sensitive to other bands of the electromagnetic spectrum.

Table 5.1 summarizes the bands of the electromagnetic spectrum and indicates the temperatures and typical astronomical objects that emit each kind of electromagnetic radiation. While at first, some of the types of radiation listed in the table may seem unfamiliar, you will get to know them better as your astronomy course continues. You can return to this table as you learn more about the types of objects astronomers study.

Table 5.1: Types of Electromagnetic Radiation

Type of Radiation	Wavelength Range (nm)	Radiated by Objects at This Temperature	Typical Sources
Gamma rays	Less than 0.01	More than 10^8 K	Produced in nuclear reactions; require very high-energy processes
X-rays	0.01–20	10^6–10^8 K	Gas in clusters of galaxies, supernova remnants, solar corona
Ultraviolet	20–400	10^4–10^6 K	Supernova remnants, very hot stars
Visible	400–700	10^3–10^4 K	Stars
Infrared	10^3–10^6	10–10^3 K	Cool clouds of dust and gas, planets, moons
Microwave	10^6–10^9	Less than 10 K	Active galaxies, pulsars, cosmic background radiation
Radio	More than 10^9	Less than 10 K	Supernova remnants, pulsars, cold gas

Radiation and Temperature

Some astronomical objects emit mostly infrared radiation, others mostly visible light, and still others mostly ultraviolet radiation. What determines the type of electromagnetic radiation emitted by the Sun, stars, and other dense astronomical objects? The answer often turns out to be their *temperature*.

At the microscopic level, everything in nature is in motion. A solid is composed of molecules and atoms in continuous vibration: they move back and forth in place, but their motion is much too small for our eyes to make out. A gas consists of atoms and/or molecules that are flying about freely at high speed, continually bumping into one another and bombarding the surrounding matter. The hotter the solid or gas, the more rapid the motion of its molecules or atoms. The temperature of something is thus a measure of the average motion energy of the particles that make it up.

This motion at the microscopic level is responsible for much of the electromagnetic radiation on Earth and in the universe. As atoms and molecules move about and collide, or vibrate in place, their electrons give off electromagnetic radiation. The characteristics of this radiation are determined by the temperature of those atoms and molecules. In a hot material, for example, the individual particles vibrate in place or move rapidly from collisions, so the emitted waves are, on average, more energetic. And recall that higher energy waves have a higher frequency. In very cool material, the particles have low-energy atomic and molecular motions and thus generate lower-energy waves.

Link to Learning

Check out the NASA briefing or NASA's 5-minute introductory video to learn more about the electromagnetic spectrum.

Radiation Laws

To understand, in more quantitative detail, the relationship between temperature and electromagnetic radiation, we imagine an idealized object called a blackbody. Such an object (unlike your sweater or your astronomy instructor's head) does not reflect or scatter any radiation, but absorbs all the electromagnetic energy that falls onto it. The energy that is absorbed causes the atoms and molecules in it to vibrate or move around at increasing speeds. As it gets hotter, this object will radiate electromagnetic waves until absorption and radiation are in balance. We want to discuss such an idealized object because, as you will see, stars behave in very nearly the same way.

The radiation from a blackbody has several characteristics, as illustrated in Figure 5.8. The graph shows the power emitted at each wavelength by objects of different temperatures. In science, the word *power* means the energy coming off per second (and it is typically measured in *watts*, which you are probably familiar with from buying lightbulbs).

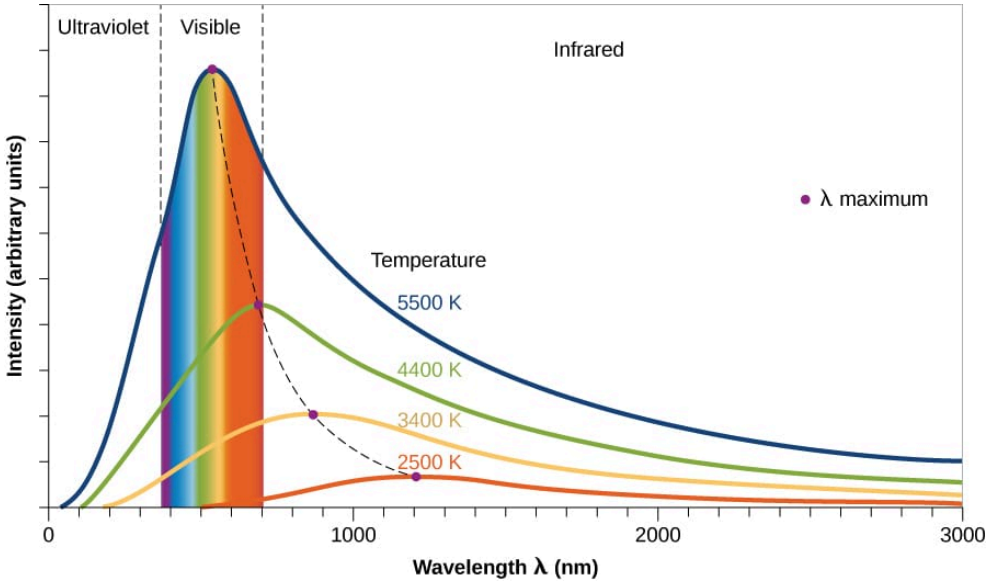


Figure 5.8 : Radiation Laws Illustrated. This graph shows in arbitrary units how many photons are given off at each wavelength for objects at four different temperatures. The wavelengths corresponding to visible light are shown by the colored bands. Note that at hotter temperatures, more energy (in the form of photons) is emitted at all wavelengths. The higher the temperature, the shorter the wavelength at which the peak amount of energy is radiated (this is known as Wien's law).

First of all, notice that the curves show that, at each temperature, our blackbody object emits radiation (photons) at all wavelengths (all colors). This is because in any solid or denser gas, some molecules or atoms vibrate or move between collisions slower than average and some move faster than average. So when we look at the electromagnetic waves emitted, we find a broad range, or spectrum, of energies and wavelengths. More energy is emitted at the average vibration or motion rate (the highest part of each curve), but if we have a large number of atoms or molecules, some energy will be detected at each wavelength.

Second, note that an object at a higher temperature emits more power at all wavelengths than does a cooler one. In a hot gas (the taller curves in Figure 5.8), for example, the atoms have more collisions and give off more energy. In the real world of stars, this means that hotter stars give off more energy at every wavelength than do cooler stars.

Third, the graph shows us that the higher the temperature, the shorter the wavelength at which the maximum power is emitted. Remember that a shorter wavelength means a higher frequency and energy. It makes sense, then, that hot objects give off a larger fraction of their energy at shorter wavelengths (higher energies) than do cool objects. You may have observed examples of this rule in everyday life. When a burner on an electric stove is turned on low, it emits only heat, which is infrared radiation, but does not glow with visible light. If the burner is set to a higher temperature, it starts to glow a dull red. At a still-higher setting, it glows a brighter orange-red (shorter wavelength). At even higher temperatures, which cannot be reached with ordinary stoves, metal can appear brilliant yellow or even blue-white.

We can use these ideas to come up with a rough sort of "thermometer" for measuring the temperatures of stars. Because many stars give off most of their energy in visible light, the color of light that dominates a star's appearance is a rough indicator of its temperature. If one star looks red and another looks blue, which one has the higher temperature? Because blue is the shorter-wavelength color, it is the sign of a hotter star. (Note that the temperatures we associate with different colors in science are not

the same as the ones artists use. In art, red is often called a “hot” color and blue a “cool” color. Likewise, we commonly see red on faucet or air conditioning controls to indicate hot temperatures and blue to indicate cold temperatures. Although these are common uses to us in daily life, in nature, it’s the other way around.)

We can develop a more precise star thermometer by measuring how much energy a star gives off at each wavelength and by constructing diagrams like Figure 5.8. The location of the peak (or maximum) in the power curve of each star can tell us its temperature. The average temperature at the surface of the Sun, which is where the radiation that we see is emitted, turns out to be 5800 K. (Throughout this text, we use the kelvin or absolute temperature scale. On this scale, water freezes at 273 K and boils at 373 K. All molecular motion ceases at 0 K. The various temperature scales are described in Appendix D.) There are stars cooler than the Sun and stars hotter than the Sun.

The wavelength at which maximum power is emitted can be calculated according to the equation

$$\lambda_{\text{max}} = \frac{3 \times 10^6}{T}$$

where the wavelength is in nanometers (one billionth of a meter) and the temperature is in K (the constant Math input error

Example 5.3: Calculating the Temperature of a Blackbody

We can use Wien’s law to calculate the temperature of a star provided we know the wavelength of peak intensity for its spectrum. If the emitted radiation from a red dwarf star has a wavelength of maximum power at 1200 nm, what is the temperature of this star, assuming it is a blackbody?

Solution

Solving Wien’s law for temperature gives:

$$T = \frac{3 \times 10^6 \text{ nm K}}{\lambda_{\text{max}}} = \frac{3 \times 10^6 \text{ nm K}}{1200 \text{ nm}} = 2500 \text{ K}$$

Exercise 9.2.1

What is the temperature of a star whose maximum light is emitted at a much shorter wavelength of 290 nm?

Answer available at the end of this section.

Since this star has a peak wavelength that is at a shorter wavelength (in the ultraviolet part of the spectrum) than that of our Sun (in the visible part of the spectrum), it should come as no surprise that its surface temperature is much hotter than our Sun’s.

We can also describe our observation that hotter objects radiate more power at all wavelengths in a mathematical form. If we sum up the contributions from all parts of the electromagnetic spectrum, we obtain the total energy emitted by a blackbody. What we usually measure from a large object like a star is the energy flux, the power emitted per square meter. The word *flux* means “flow” here: we are interested in the flow of power into an area (like the area of a telescope mirror). It turns out that the energy flux from a blackbody at temperature *T* is proportional to the fourth power of its absolute temperature. This relationship is known as the Stefan-Boltzmann law and can be written in the form of an equation as

$$F = \sigma T^4$$

where *F* stands for the energy flux (in units of watts per square meter), T is given in Kelvins, and σ (Greek letter sigma) is a constant number $\left(5.67 \times 10^{-8}\right)$.

Notice how impressive this result is. Increasing the temperature of a star would have a tremendous effect on the power it radiates. If the Sun, for example, were twice as hot—that is, if it had a temperature of 11,600 K—it would radiate 2^4, or 16 times more power than it does now. Tripling the temperature would raise the power output 81 times. Hot stars really shine away a tremendous amount of energy.

Example 5.4: Calculating the Power of a Star

While energy flux tells us how much power a star emits per square meter, we would often like to know how much total power is emitted by the star. We can determine that by multiplying the energy flux by the number of square meters on the surface of the star. Stars are mostly spherical, so we can use the formula $4\pi R^2$ for the surface area, where R is the radius of the star. The total power emitted by the star (which we call the star's "absolute luminosity") can be found by multiplying the formula for energy flux and the formula for the surface area:

$$L = 4\pi R^2 \sigma T^4$$

Two stars have the same size and are the same distance from us. Star A has a surface temperature of 6000 K, and star B has a surface temperature twice as high, 12,000 K. How much more luminous is star B compared to star A?

Solution

$$L_\mathrm{A} = 4\pi {R_\mathrm{A}}^2 \sigma {T_\mathrm{A}}^4 \text{ and } L_\mathrm{B} = 4\pi {R_\mathrm{B}}^2 \sigma {T_\mathrm{B}}^4$$

Take the ratio of the luminosity of Star A to Star B:

$$\frac{L_\mathrm{B}}{L_\mathrm{A}} = \frac{4\pi {R_\mathrm{B}}^2 \sigma {T_\mathrm{B}}^4}{4\pi {R_\mathrm{A}}^2 \sigma {T_\mathrm{A}}^4} = \frac{{R_\mathrm{B}}^2 {T_\mathrm{B}}^4}{{R_\mathrm{A}}^2 {T_\mathrm{A}}^4}$$

Because the two stars are the same size, $R_\mathrm{A} = R_\mathrm{B}$, leaving

$$\frac{{T_\mathrm{B}}^4}{{T_\mathrm{A}}^4} = \frac{(12{,}000\ \mathrm{K})^4}{(6000\ \mathrm{K})^4} = 2^4 = 16$$

Exercise 9.2.1

Two stars with identical diameters are the same distance away. One has a temperature of 8700 K and the other has a temperature of 2900 K. Which is brighter? How much brighter is it?

Answer available at the end of this section.

Link to Learning

You can use the Blackbody Spectrum simulator to explore the relationship between an object's temperature, peak wavelength, overall brightness, and apparent color.

Answers

Exercise 9.2.1

$T = \frac{3 \times 10^6\ \mathrm{nm\ K}}{\lambda_\mathrm{max}} = \frac{3 \times 10^6\ \mathrm{nm\ K}}{290\ \mathrm{nm}} = 10{,}300\ \mathrm{K}$

Exercise 9.2.1

The 8700 K star has triple the temperature, so it is 3^4 = 81 times brighter.

This page titled 9.2: The Electromagnetic Spectrum is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.3: The Electromagnetic Spectrum by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.3: Spectroscopy in Astronomy

Learning Objectives

By the end of this section, you will be able to:

- Describe the properties of light
- Explain how astronomers learn the composition of a gas by examining its spectral lines
- Discuss the various types of spectra

Electromagnetic radiation carries a lot of information about the nature of stars and other astronomical objects. To extract this information, however, astronomers must be able to study the amounts of energy we receive at different wavelengths of light in fine detail. Let's examine how we can do this and what we can learn.

Properties of Light

Light exhibits certain behaviors that are important to the design of telescopes and other instruments. For example, light can be *reflected* from a surface. If the surface is smooth and shiny, as with a mirror, the direction of the reflected light beam can be calculated accurately from knowledge of the shape of the reflecting surface. Light is also bent, or *refracted*, when it passes from one kind of transparent material into another—say, from the air into a glass lens.

Reflection and refraction of light are the basic properties that make possible all *optical* instruments (devices that help us to see things better)—from eyeglasses to giant astronomical telescopes. Such instruments are generally combinations of glass lenses, which bend light according to the principles of refraction, and curved mirrors, which depend on the properties of reflection. Small optical devices, such as eyeglasses or binoculars, generally use lenses, whereas large telescopes depend almost entirely on mirrors for their main optical elements. We will discuss astronomical instruments and their uses more fully in Astronomical Instruments. For now, we turn to another behavior of light, one that is essential for the decoding of light.

In 1672, in the first paper that he submitted to the Royal Society, Sir Isaac Newton described an experiment in which he permitted sunlight to pass through a small hole and then through a prism. Newton found that sunlight, which looks white to us, is actually made up of a mixture of all the colors of the rainbow (Figure 5.9).

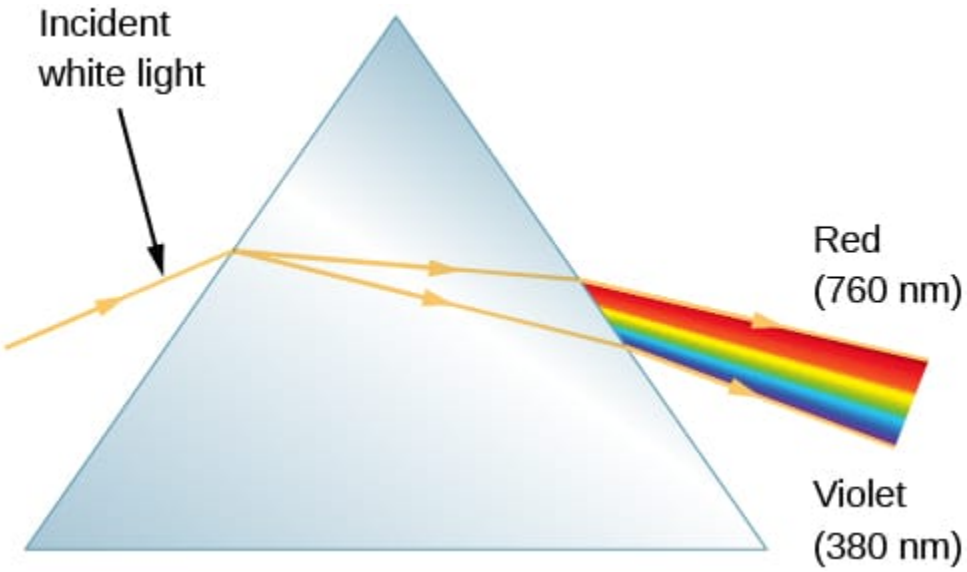


Figure 5.9 : Action of a Prism. When we pass a beam of white sunlight through a prism, we see a rainbow-colored band of light that we call a continuous spectrum.

Figure 5.9 shows how light is separated into different colors with a prism—a piece of glass in the shape of a triangle with refracting surfaces. Upon entering one face of the prism, the path of the light is refracted (bent), but not all of the colors are bent by the same amount. The bending of the beam depends on the wavelength of the light as well as the properties of the material, and as a result, different wavelengths (or colors of light) are bent by different amounts and therefore follow slightly different paths through the prism. The violet light is bent more than the red. This phenomenon is called dispersion and explains Newton's rainbow experiment.

Upon leaving the opposite face of the prism, the light is bent again and further dispersed. If the light leaving the prism is focused on a screen, the different wavelengths or colors that make up white light are lined up side by side just like a rainbow (Figure 5.10). (In fact, a rainbow is formed by the dispersion of light though raindrops; see The Rainbow feature box.) Because this array of colors is a spectrum of light, the instrument used to disperse the light and form the spectrum is called a spectrometer.

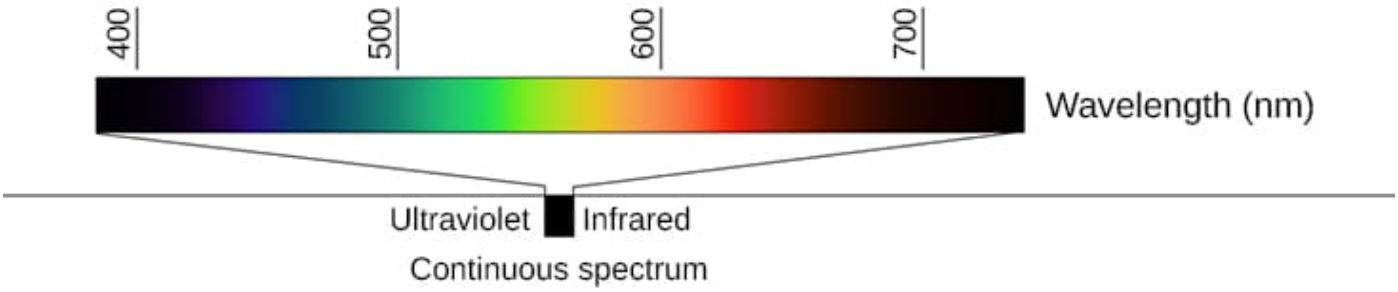


Figure 5.10 : Continuous Spectrum. When white light passes through a prism, it is dispersed and forms a continuous spectrum of all the colors. Although it is hard to see in this printed version, in a well-dispersed spectrum, many subtle gradations in color are visible as your eye scans from one end (violet) to the other (red).

The Value of Stellar Spectra

When Newton described the laws of refraction and dispersion in optics, and observed the solar spectrum, all he could see was a continuous band of colors. If the spectrum of the white light from the Sun and stars were simply a continuous rainbow of colors, astronomers would have little interest in the detailed study of a star's spectrum once they had learned its average surface temperature. In 1802, however, William Wollaston built an improved spectrometer that included a lens to focus the Sun's spectrum on a screen. With this device, Wollaston saw that the colors were not spread out uniformly, but instead, some ranges of color were missing, appearing as dark bands in the solar spectrum. He mistakenly attributed these lines to natural boundaries between the colors. In 1815, German physicist Joseph Fraunhofer, upon a more careful examination of the solar spectrum, found about 600 such dark lines (missing colors), which led scientists to rule out the boundary hypothesis (Figure 5.11).

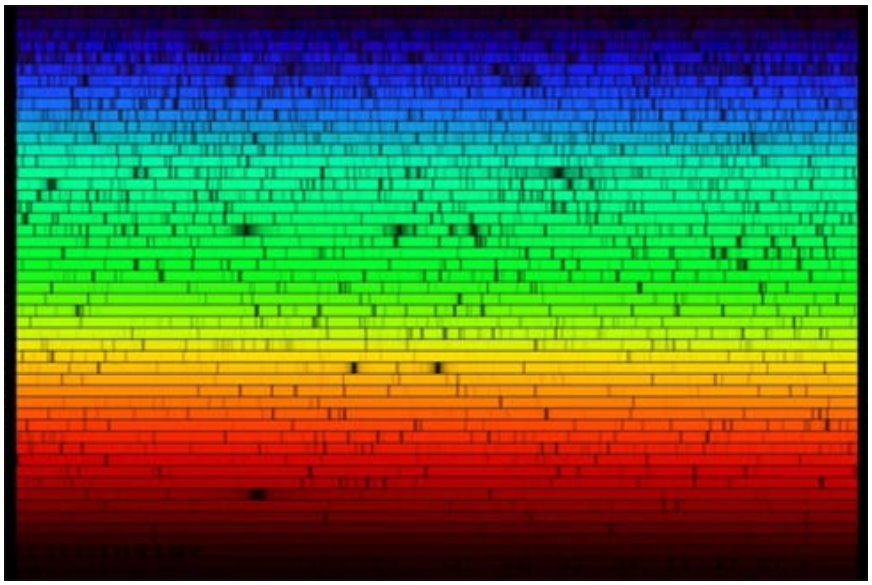


Figure 5.11 : Visible Spectrum of the Sun. Our star's spectrum is crossed by dark lines produced by atoms in the solar atmosphere that absorb light at certain wavelengths. (credit: modification of work by Nigel Sharp, NOAO/National Solar Observatory at Kitt Peak/AURA, and the National Science Foundation)

Later, researchers found that similar dark lines could be produced in the spectra ("spectra" is the plural of "spectrum") of artificial light sources. They did this by passing their light through various apparently transparent substances—usually containers with just a bit of thin gas in them.

These gases turned out not to be transparent at *all* colors: they were quite opaque at a few sharply defined wavelengths. Something in each gas had to be absorbing just a few colors of light and no others. All gases did this, but each different element absorbed a different set of colors and thus showed different dark lines. If the gas in a container consisted of two elements, then light passing through it was missing the colors (showing dark lines) for both of the elements. So it became clear that certain lines in the spectrum "go with" certain elements. This discovery was one of the most important steps forward in the history of astronomy.

What would happen if there were no continuous spectrum for our gases to remove light from? What if, instead, we heated the same thin gases until they were hot enough to glow with their own light? When the gases were heated, a spectrometer revealed no continuous spectrum, but several separate bright lines. That is, these hot gases emitted light only at certain specific wavelengths or colors.

When the gas was pure hydrogen, it would emit one pattern of colors; when it was pure sodium, it would emit a different pattern. A mixture of hydrogen and sodium emitted both sets of spectral lines. The colors the gases emitted when they were heated were the very same colors as those they had absorbed when a continuous source of light was behind them. From such experiments, scientists began to see that different substances showed distinctive *spectral signature*s by which their presence could be detected (Figure 5.12). Just as your signature allows the bank to identify you, the unique pattern of colors for each type of atom (its spectrum) can help us identify which element or elements are in a gas.

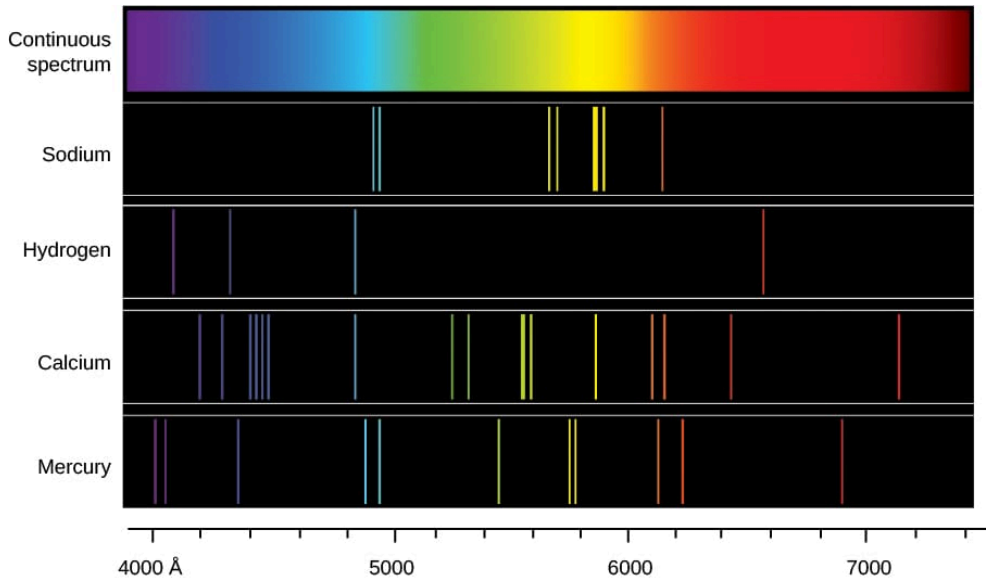


Figure 5.12 : Continuous Spectrum and Line Spectra from Different Elements. Each type of glowing gas (each element) produces its own unique pattern of lines, so the composition of a gas can be identified by its spectrum. The spectra of sodium, hydrogen, calcium, and mercury gases are shown here.

Types of Spectra

In these experiments, then, there were three different types of spectra. A continuous spectrum (formed when a solid or very dense gas gives off radiation) is an array of all wavelengths or colors of the rainbow. A continuous spectrum can serve as a backdrop from which the atoms of much less dense gas can absorb light. A dark line, or absorption spectrum, consists of a series or pattern of dark lines—missing colors—superimposed upon the continuous spectrum of a source. A bright line, or emission spectrum, appears as a pattern or series of bright lines; it consists of light in which only certain discrete wavelengths are present. (Figure 5.11 shows an absorption spectrum, whereas Figure 5.12 shows the emission spectrum of a number of common elements along with an example of a continuous spectrum.)

When we have a hot, thin gas, each particular chemical element or compound produces its own characteristic pattern of spectral lines—its spectral signature. No two types of atoms or molecules give the same patterns. In other words, each particular gas can absorb or emit only certain wavelengths of the light peculiar to that gas. In contrast, absorption spectra occur when passing white light through a cool, thin gas. The temperature and other conditions determine whether the lines are bright or dark (whether light is absorbed or emitted), but the wavelengths of the lines for any element are the same in either case. It is the precise pattern of wavelengths that makes the signature of each element unique. Liquids and solids can also generate spectral lines or bands, but they are broader and less well defined—and hence, more difficult to interpret. Spectral analysis, however, can be quite useful. It can, for example, be applied to light reflected off the surface of a nearby asteroid as well as to light from a distant galaxy.

The dark lines in the solar spectrum thus give evidence of certain chemical elements between us and the Sun absorbing those wavelengths of sunlight. Because the space between us and the Sun is pretty empty, astronomers realized that the atoms doing the absorbing must be in a thin atmosphere of cooler gas around the Sun. This outer atmosphere is not all that different from the rest of the Sun, just thinner and cooler. Thus, we can use what we learn about its composition as an indicator of what the whole Sun is made of. Similarly, we can use the presence of absorption and emission lines to analyze the composition of other stars and clouds of gas in space.

Such analysis of spectra is the key to modern astronomy. Only in this way can we “sample” the stars, which are too far away for us to visit. Encoded in the electromagnetic radiation from celestial objects is clear information about the chemical makeup of these objects. Only by understanding what the stars were made of could astronomers begin to form theories about what made them shine and how they evolved.

In 1860, German physicist Gustav Kirchhoff became the first person to use spectroscopy to identify an element in the Sun when he found the spectral signature of sodium gas. In the years that followed, astronomers found many other chemical elements in the Sun and stars. In fact, the element helium was found first in the Sun from its spectrum and only later identified on Earth. (The word “helium” comes from *helios*, the Greek name for the Sun.)

Why are there specific lines for each element? The answer to that question was not found until the twentieth century; it required the development of a model for the atom. We therefore turn next to a closer examination of the atoms that make up all matter.

Link to Learning

You can use the Spectrum Constructor to explore the three types of spectra and how they appear through a diffraction grating when the brightness of each wavelength is measured with a spectrometer.

Making Connections: The Rainbow

Rainbows are an excellent illustration of the dispersion of sunlight. You have a good chance of seeing a rainbow any time you are between the Sun and a rain shower, as illustrated in Figure 5.13. The raindrops act like little prisms and break white light into the spectrum of colors. Suppose a ray of sunlight encounters a raindrop and passes into it. The light changes direction—is refracted—when it passes from air to water; the blue and violet light are refracted more than the red. Some of the light is then reflected at the backside of the drop and reemerges from the front, where it is again refracted. As a result, the white light is spread out into a rainbow of colors.



Figure 5.13 : Rainbow Refraction. (a) This diagram shows how light from the Sun, which is located behind the observer, can be refracted by raindrops to produce (b) a rainbow. (c) Refraction separates white light into its component colors.

Note that violet light lies above the red light after it emerges from the raindrop. When you look at a rainbow, however, the red light is higher in the sky. Why? Look again at Figure 5.13. If the observer looks at a raindrop that is high in the sky, the violet light passes over her head and the red light enters her eye. Similarly, if the observer looks at a raindrop that is low in the sky, the violet light reaches her eye and the drop appears violet, whereas the red light from that same drop strikes the ground and is not seen. Colors of intermediate wavelengths are refracted to the eye by drops that are intermediate in altitude between the drops that appear violet and the ones that appear red. Thus, a single rainbow always has red on the outside and violet on the inside.

This page titled 9.3: Spectroscopy in Astronomy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.4: Spectroscopy in Astronomy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.4: The Structure of the Atom

Learning Objectives

By the end of this section, you will be able to:

- Describe the structure of atoms and the components of nuclei
- The idea that matter is composed of tiny particles called atoms is at least 25 centuries old. It took until the twentieth century, however, for scientists to invent instruments that permitted them to probe inside an atom and find that it is not, as had been thought, hard and indivisible. Instead, the atom is a complex structure composed of still smaller particles.

Probing the Atom

The first of these smaller particles was discovered by British physicist James (J. J.) Thomson in 1897. Named the *electron*, this particle is negatively charged. (It is the flow of these particles that produces currents of electricity, whether in lightning bolts or in the wires leading to your lamp.) Because an atom in its normal state is electrically neutral, each electron in an atom must be balanced by the same amount of positive charge.

The next step was to determine where in the atom the positive and negative charges are located. In 1911, British physicist Ernest Rutherford devised an experiment that provided part of the answer to this question. He bombarded an extremely thin piece of gold foil, only about 400 atoms thick, with a beam of alpha particles (Figure 5.14). *Alpha particles* (α particles) are helium atoms that have lost their electrons and thus are positively charged. Most of these particles passed though the gold foil just as if it and the atoms in it were nearly empty space. About 1 in 8000 of the alpha particles, however, completely reversed direction and bounced backward from the foil. Rutherford wrote, “It was quite the most incredible event that has ever happened to me in my life. It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you.”

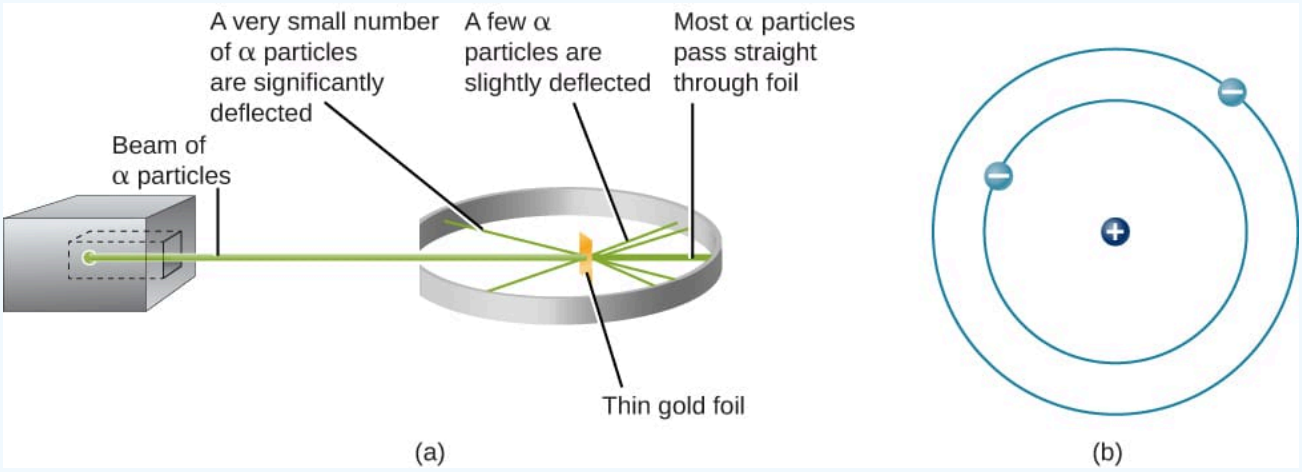


Figure 5.14 : Rutherford’s Experiment. (a) When Rutherford allowed α particles from a radioactive source to strike a target of gold foil, he found that, although most of them went straight through, some rebounded back in the direction from which they came. (b) From this experiment, he concluded that the atom must be constructed like a miniature solar system, with the positive charge concentrated in the nucleus and the negative charge orbiting in the large volume around the nucleus. Note that this drawing is not to scale; the electron orbits are much larger relative to the size of the nucleus.

The only way to account for the particles that reversed direction when they hit the gold foil was to assume that nearly all of the mass, as well as all of the positive charge in each individual gold atom, is concentrated in a tiny center or nucleus. When a positively charged alpha particle strikes a nucleus, it reverses direction, much as a cue ball reverses direction when it strikes another billiard ball. Rutherford’s model placed the other type of charge—the negative electrons—in orbit around this nucleus.

Rutherford's model required that the electrons be in motion. Positive and negative charges attract each other, so stationary electrons would fall into the positive nucleus. Also, because both the electrons and the nucleus are extremely small, most of the atom is empty, which is why nearly all of Rutherford's particles were able to pass right through the gold foil without colliding with anything. Rutherford's model was a very successful explanation of the experiments he conducted, although eventually scientists would discover that even the nucleus itself has structure.

The Atomic Nucleus

The simplest possible atom (and the most common one in the Sun and stars) is hydrogen. The nucleus of ordinary hydrogen contains a single proton. Moving around this proton is a single electron. The mass of an electron is nearly 2000 times smaller than the mass of a proton; the electron carries an amount of charge exactly equal to that of the proton but opposite in sign (Figure 5.15). Opposite charges attract each other, so it is an electromagnetic force that holds the proton and electron together, just as gravity is the force that keeps planets in orbit around the Sun.

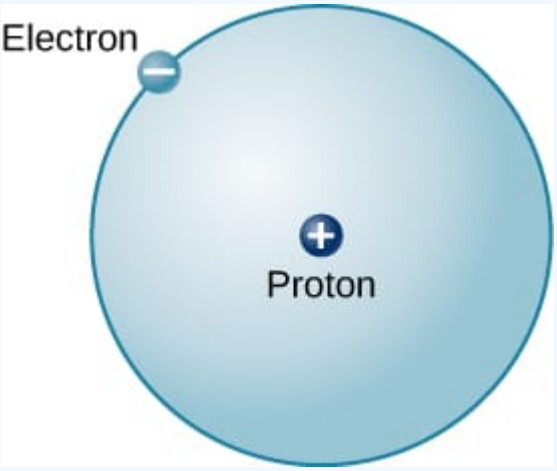


Figure 5.15 : Hydrogen Atom. This is a schematic diagram of a hydrogen atom in its lowest energy state, also called the ground state. The proton and electron have equal but opposite charges, which exert an electromagnetic force that binds the hydrogen atom together. In the illustration, the size of the particles is exaggerated so that you can see them; they are not to scale. They are also shown much closer than they would actually be as it would take more than an entire page to show their actual distance to scale.

There are many other types of atoms in nature. Helium, for example, is the second-most abundant element in the Sun. Helium has two protons in its nucleus instead of the single proton that characterizes hydrogen. In addition, the helium nucleus contains two neutrons, particles with a mass comparable to that of the proton but with no electric charge. Moving around this nucleus are two electrons, so the total net charge of the helium atom is also zero (Figure 5.16).

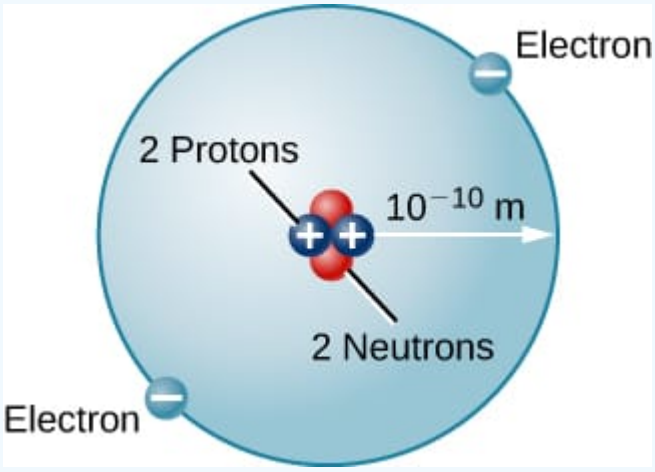


Figure 5.16 : Helium Atom. Here we see a schematic diagram of a helium atom in its lowest energy state. Two protons are present in the nucleus of all helium atoms. In the most common variety of helium, the nucleus also contains two neutrons, which have nearly the same mass as the proton but carry no charge. Two electrons orbit the nucleus.

From this description of hydrogen and helium, perhaps you have guessed the pattern for building up all the elements (different types of atoms) that we find in the universe. The type of element is determined by the number of protons in the nucleus of the atom. For example, any atom with six protons is the element carbon, with eight protons is oxygen, with 26 is iron, and with 92 is uranium. On Earth, a typical atom has the same number of electrons as protons, and these electrons follow complex orbital patterns around the nucleus. Deep inside stars, however, it is so hot that the electrons get loose from the nucleus and (as we shall see) lead separate yet productive lives.

The ratio of neutrons to protons increases as the number of protons increases, but each element is unique. The number of neutrons is not necessarily the same for all atoms of a given element. For example, most hydrogen atoms contain no neutrons at all. There are, however, hydrogen atoms that contain one proton and one neutron, and others that contain one proton and two neutrons. The various types of hydrogen nuclei with different numbers of neutrons are called isotopes of hydrogen (Figure 5.17), and all other elements have isotopes as well. You can think of isotopes as siblings in the same element "family"—closely related but with different characteristics and behaviors.

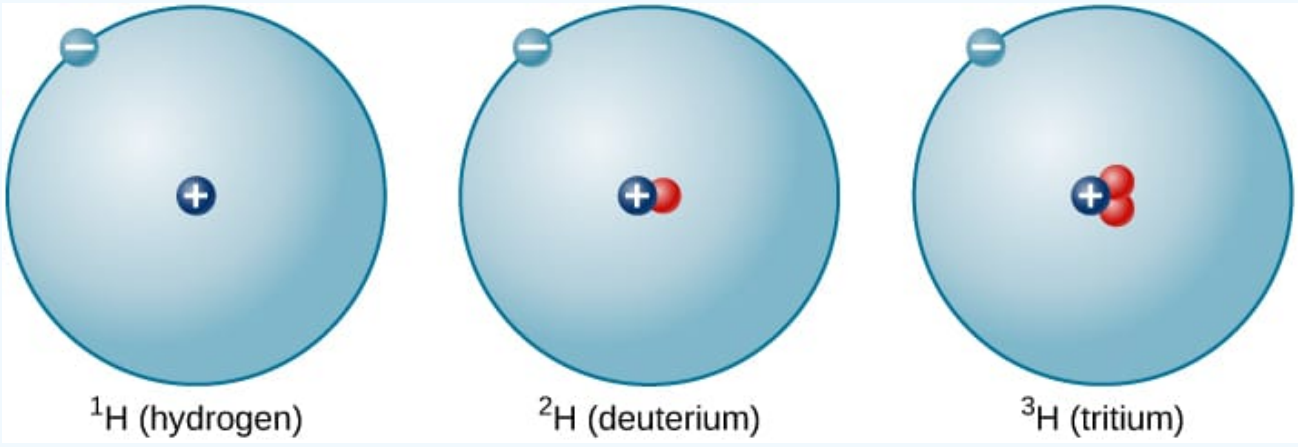


Figure 5.17 : Isotopes of Hydrogen. A single proton in the nucleus defines the atom to be hydrogen, but there may be zero, one, or two neutrons. The most common isotope of hydrogen is the one with only a single proton and no neutrons.

Link to Learning

To explore the structure of atoms, go to the PhET Build an Atom website where you can add protons, neutrons, or electrons to a model and the name of the element you have created will appear. You can also see the net charge, the mass number, whether it is stable or unstable, and whether it is an ion or a neutral atom.

The Bohr Atom

Rutherford's model for atoms has one serious problem. Maxwell's theory of electromagnetic radiation says that when electrons change either speed or the direction of motion, they must emit energy. Orbiting electrons constantly change their direction of motion, so they should emit a constant stream of energy. Applying Maxwell's theory to Rutherford's model, all electrons should spiral into the nucleus of the atom as they lose energy, and this collapse should happen very quickly—in about 10^{-16} seconds.

It was Danish physicist Niels Bohr (1885–1962) who solved the mystery of how electrons remain in orbit. He was trying to develop a model of the atom that would also explain certain regularities observed in the spectrum of hydrogen. He suggested that the spectrum of hydrogen can be understood if we assume that orbits of only certain sizes are possible for the electron. Bohr further assumed that as long as the electron moves in only one of these allowed orbits, it radiates no energy: its energy would change only if it moved from one orbit to another.

This suggestion, in the words of science historian Abraham Pais, was "one of the most audacious hypotheses ever introduced in physics." If something equivalent were at work in the everyday world, you might find that, as you went for a walk after astronomy class, nature permitted you to walk two steps per minute, five steps per minute, and 12 steps per minute, but no speeds in between. No matter how you tried to move your legs, only certain walking speeds would be permitted. To make things more bizarre, it would take no effort to walk at any one of the allowed speeds, but it would

be difficult to change from one speed to another. Luckily, no such rules apply at the level of human behavior. But at the microscopic level of the atom, experiment after experiment has confirmed the validity of Bohr's strange idea. Bohr's suggestions became one of the foundations of the new (and much more sophisticated) model of the subatomic world called quantum mechanics.

In Bohr's model, if the electron moves from one orbit to another closer to the atomic nucleus, it must give up some energy in the form of electromagnetic radiation. If the electron goes from an inner orbit to one farther from the nucleus, however, it requires some additional energy. One way to obtain the necessary energy is to absorb electromagnetic radiation that may be streaming past the atom from an outside source.

A key feature of Bohr's model is that each of the permitted electron orbits around a given atom has a certain energy value; we therefore can think of each orbit as an energy level. To move from one orbit to another (which will have its own specific energy value) requires a change in the electron's energy—a change determined by the difference between the two energy values. If the electron goes to a lower level, the energy difference will be given off; if the electron goes to a higher level, the energy difference must be obtained from somewhere else. Each jump (or transition) to a different level has a fixed and definite energy change associated with it.

A crude analogy for this situation might be life in a tower of luxury apartments where the rent is determined by the quality of the view. Such a building has certain, definite numbered levels or floors on which apartments are located. No one can live on floor 5.37 or 22.5. In addition, the rent gets higher as you go up to higher floors. If you want to exchange an apartment on the twentieth floor for one on the second floor, you will not owe as much rent. However, if you want to move from the third floor to the twenty-fifth floor, your rent will increase. In an atom, too, the "cheapest" place for an electron to live is the lowest possible level, and energy is required to move to a higher level.

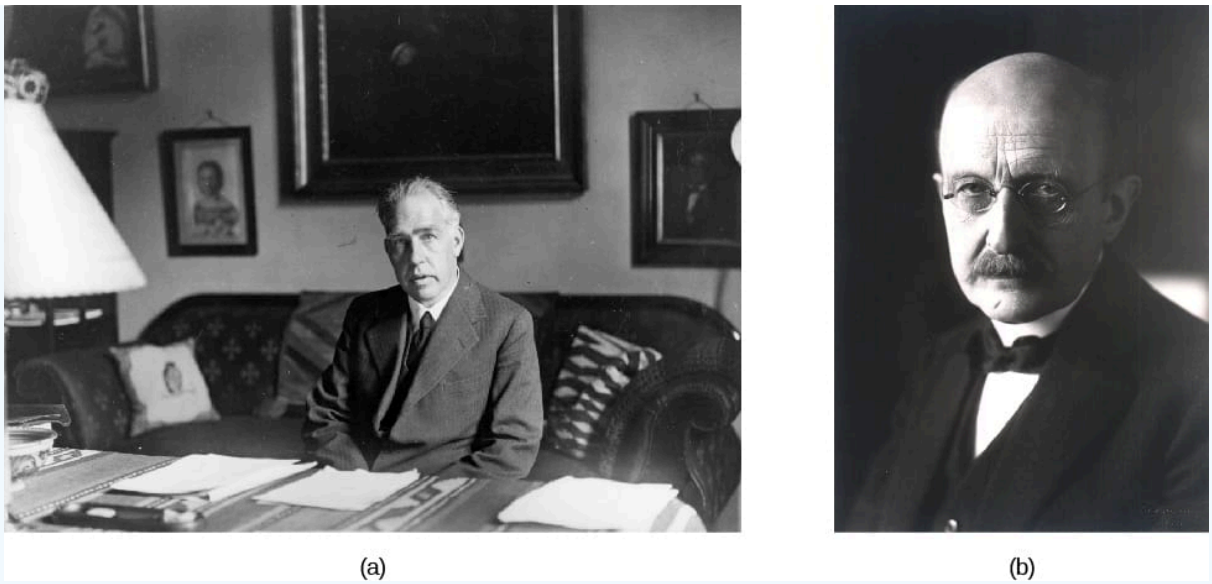
Here we have one of the situations where it is easier to think of electromagnetic radiation as particles (photons) rather than as waves. As electrons move from one level to another, they give off or absorb little packets of energy. When an electron moves to a higher level, it absorbs a photon of just the right energy (provided one is available). When it moves to a lower level, it emits a photon with the exact amount of energy it no longer needs in its "lower-cost living situation."

The photon and wave perspectives must be equivalent: light is light, no matter how we look at it. Thus, each photon carries a certain amount of energy that is proportional to the frequency (f) of the wave it represents. The value of its energy (E) is given by the formula

$$E = hf$$

where the constant of proportionality, h, is called Planck's constant.

The constant is named for Max Planck, the German physicist who was one of the originators of the quantum theory (Figure 5.18). If metric units are used (that is, if energy is measured in joules and frequency in hertz), then Planck's constant has the value $h = 6.626 \times 10^{-34}$ joule-seconds (J-s). Higher-energy photons correspond to higher-frequency waves (which have a shorter wavelength); lower-energy photons are waves of lower frequency.



(a) (b)

Figure 5.18 : Niels Bohr (1885–1962) and Max Planck (1858–1947). (a) Bohr, shown at his desk in this 1935 photograph, and (b) Planck helped us understand the energy behavior of photons.

To take a specific example, consider a calcium atom inside the Sun's atmosphere in which an electron jumps from a lower level to a higher level. To do this, it needs about 5×10^{-19} joules of energy, which it can conveniently obtain by absorbing a passing photon of that energy coming from deeper inside the Sun. This photon is equivalent to a wave of light whose frequency is about 7.5×10^{14} hertz and whose wavelength is about 3.9×10^{-7} meters (393 nanometers), in the deep violet part of the visible light spectrum. Although it may seem strange at first to switch from picturing light as a photon (or energy packet) to picturing it as a wave, such switching has become second nature to astronomers and can be a handy tool for doing calculations about spectra.

Example 5.5: The Energy of a Photon

Now that we know how to calculate the wavelength and frequency of a photon, we can use this information, along with Planck's constant, to determine how much energy each photon carries. How much energy does a red photon of wavelength 630 nm have?

Solution

First, as we learned earlier, we can find the frequency of the photon:

$$f = \frac{c}{\lambda} = \frac{3 \times 10^{8}\,\text{m/s}}{630 \times 10^{-9}\,\text{m}} = 4.8 \times 10^{14}\,\text{Hz}$$

Next, we can use Planck's constant to determine the energy (remember that a Hz is the same as 1/s):

$$E = hf = \left(6.626 \times 10^{-34}\,\text{J-s}\right)\left(4.8 \times 10^{14}\,\text{Hz}(1/\text{s})\right) = 3.2 \times 10^{-19}\,\text{J}$$

Exercise 9.4.1

What is the energy of a yellow photon with a frequency of 5.5×10^{14} Hz?

Answer available at the end of this section.

Answers

Exercise 9.4.1

$E = hf = \left(6.626 \times 10^{-34}\text{J-s}\right)\left(5.5 \times 10^{14}\text{Hz}\right) = 3.6 \times 10^{-19}\text{ J}$

This page titled 9.4: The Structure of the Atom is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.5: The Structure of the Atom by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.5: Formation of Spectral Lines

Learning Objectives

By the end of this section, you will be able to:

- Explain how emission line spectra and absorption line spectra are formed
- Describe what ions are and how they are formed
- Explain how spectral lines and ionization levels in a gas can help us determine its temperature

We can use Bohr's model of the atom to understand how spectral lines are formed. The concept of energy levels for the electron orbits in an atom leads naturally to an explanation of why atoms absorb or emit only specific energies or wavelengths of light.

The Hydrogen Spectrum

Let's look at the hydrogen atom from the perspective of the Bohr model. Suppose a beam of white light (which consists of photons of all visible wavelengths) shines through a gas of atomic hydrogen. A photon of wavelength 656 nanometers has just the right energy to raise an electron in a hydrogen atom from the second to the third orbit. Thus, as all the photons of different energies (or wavelengths or colors) stream by the hydrogen atoms, photons with *this* particular wavelength can be absorbed by those atoms whose electrons are orbiting on the second level. When they are absorbed, the electrons on the second level will move to the third level, and a number of the photons of this wavelength and energy will be missing from the general stream of white light.

Other photons will have the right energies to raise electrons from the second to the fourth orbit, or from the first to the fifth orbit, and so on. Only photons with these exact energies can be absorbed. All of the other photons will stream past the atoms untouched. Thus, hydrogen atoms absorb light at only certain wavelengths and produce dark lines at those wavelengths in the spectrum we see.

Suppose we have a container of hydrogen gas through which a whole series of photons is passing, allowing many electrons to move up to higher levels. When we turn off the light source, these electrons "fall" back down from larger to smaller orbits and emit photons of light—but, again, only light of those energies or wavelengths that correspond to the energy difference between permissible orbits. The orbital changes of hydrogen electrons that give rise to some spectral lines are shown in Figure 5.19.

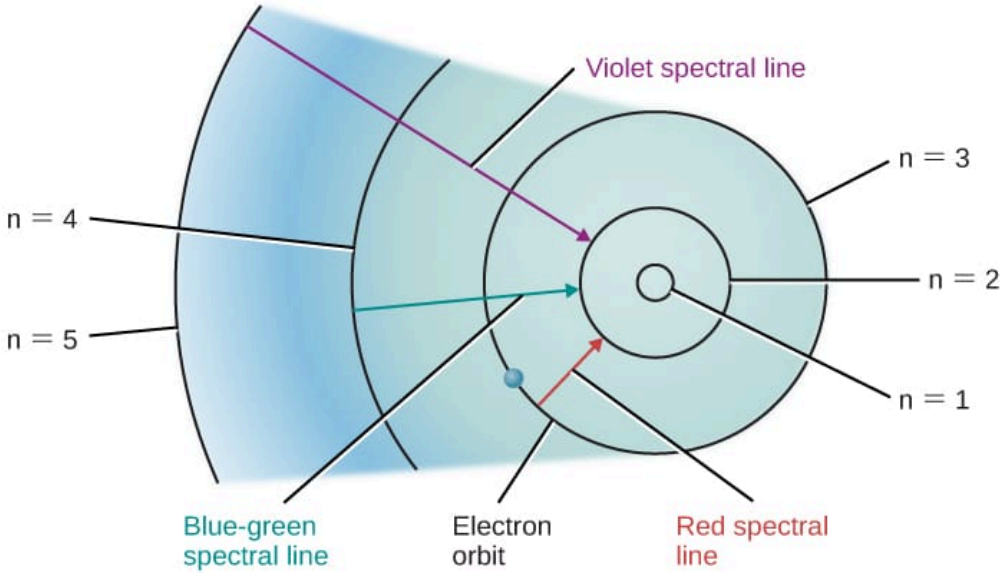


Figure 5.19 : Bohr Model for Hydrogen. In this simplified model of a hydrogen atom, the concentric circles shown represent permitted orbits or energy levels. An electron in a hydrogen atom can only exist in one of these energy levels (or states). The closer the electron is to the nucleus, the more tightly bound the electron is to the nucleus. By absorbing energy, the electron can move to energy levels farther from the nucleus (and even escape if enough energy is absorbed).

Similar pictures can be drawn for atoms other than hydrogen. However, because these other atoms ordinarily have more than one electron each, the orbits of their electrons are much more complicated, and the spectra are more complex as well. For our purposes, the key conclusion is this: *each type of atom has its own unique pattern of electron orbits, and no two sets of orbits are exactly alike*. This means that each type of atom shows its own unique set of spectral lines, produced by electrons moving between its unique set of orbits.

Astronomers and physicists have worked hard to learn the lines that go with each element by studying the way atoms absorb and emit light in laboratories here on Earth. Then they can use this knowledge to identify the elements in celestial bodies. In this way, we now know the chemical makeup of not just any star, but even galaxies of stars so distant that their light started on its way to us long before Earth had even formed.

Energy Levels and Excitation

Bohr's model of the hydrogen atom was a great step forward in our understanding of the atom. However, we know today that atoms cannot be represented by quite so simple a picture. For example, the concept of sharply defined electron orbits is not really correct; however, at the level of this introductory course, the notion that only certain discrete energies are allowable for an atom is very useful. The energy levels we have been discussing can be thought of as representing certain average distances of the electron's possible orbits from the atomic nucleus.

Ordinarily, an atom is in the state of lowest possible energy, its ground state. In the Bohr model of the hydrogen atom, the ground state corresponds to the electron being in the innermost orbit. An atom can absorb energy, which raises it to a higher energy level (corresponding, in the simple Bohr picture, to an electron's movement to a larger orbit)—this is referred to as excitation. The atom is then said to be in an *excited state*. Generally, an atom remains excited for only a very brief time. After a short interval, typically a hundred-millionth of a second or so, it drops back spontaneously to its ground state, with the simultaneous emission of light. The atom may return to its lowest state in one jump, or it may make the transition in steps of two or more jumps, stopping at intermediate levels on the way down. With each jump, it emits a photon of the wavelength that corresponds to the energy difference between the levels at the beginning and end of that jump.

An energy-level diagram for a hydrogen atom and several possible atomic transitions are shown in Figure 5.20. When we measure the energies involved as the atom jumps between levels, we find that the transitions to or from the ground state, called the *Lyman series* of lines, result in the emission or absorption of ultraviolet photons. But the transitions to or from the first excited state (labeled n = 2 in part (a) of Figure 5.20), called the Balmer series, produce emission or absorption in visible light. In fact, it was to explain this Balmer series that Bohr first suggested his model of the atom.

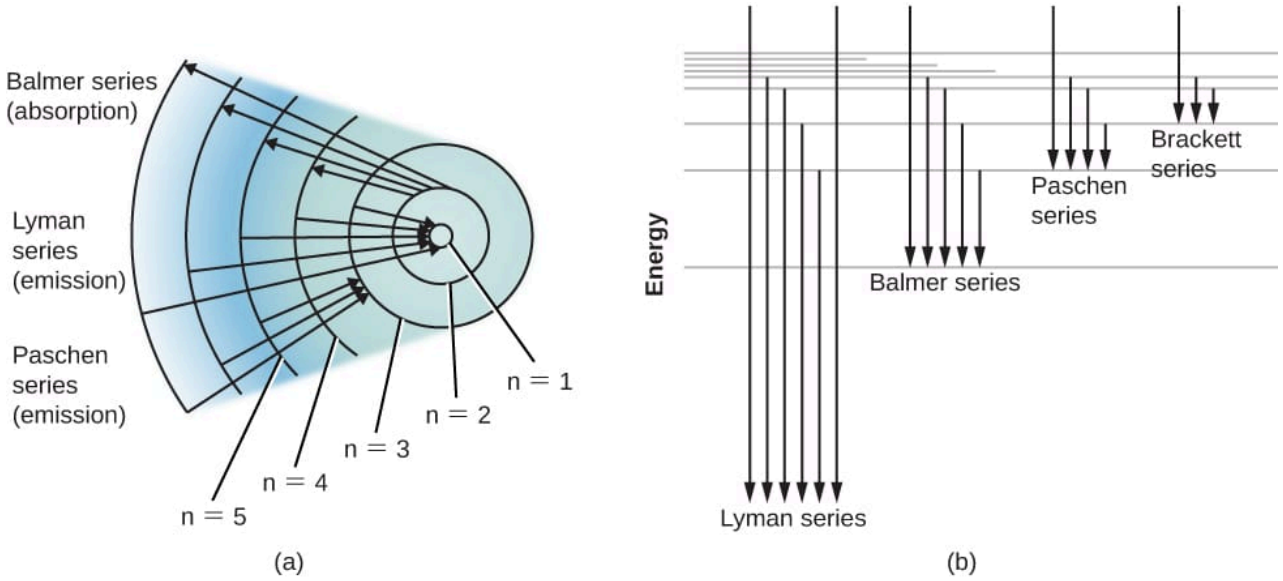


Figure 5.20 : Energy-Level Diagrams for Hydrogen. (a) Here we follow the emission or absorption of photons by a hydrogen atom according to the Bohr model. Several different series of spectral lines are shown, corresponding to transitions of electrons from or to certain allowed orbits. Each series of lines that terminates on a specific inner orbit is named for the physicist who studied it. At the top, for example, you see the Balmer series, and arrows show electrons jumping from the second orbit (n = 2) to the third, fourth, fifth, and sixth orbits. Each time a "poor" electron from a lower level wants to rise to a higher position in life, it must absorb energy to do so. It can absorb the energy it needs from passing waves (or photons) of light. The next set of arrows (Lyman series) show electrons falling down to the first orbit from different (higher) levels. Each time a "rich" electron goes downward toward the nucleus, it can afford to give off (emit) some energy it no longer needs. (In this idealized diagram, the energy levels are shown equally spaced; in real life, they are not.) (b) At higher and higher energy levels, the levels become more and more crowded together, approaching a limit. The region above the top line represents energies at which the atom is ionized (the electron is no longer attached to the atom). Each series of arrows represents electrons falling from higher levels to lower ones, releasing photons or waves of energy in the process.

Atoms that have absorbed specific photons from a passing beam of white light and have thus become excited generally de-excite themselves and emit that light again in a very short time. You might wonder, then, why *dark* spectral lines are ever produced. In other words, why doesn't this reemitted light quickly "fill in" the darker absorption lines?

Imagine a beam of white light coming toward you through some cooler gas. Some of the reemitted light *is* actually returned to the beam of white light you see, but this fills in the absorption lines only to a slight extent. The reason is that the atoms in the gas reemit light *in all directions*, and only a small fraction of the reemitted light is in the direction of the original beam (toward you). In a star, much of the reemitted light actually goes in directions leading back into the star, which does observers outside the star no good whatsoever.

Figure 5.21 summarizes the different kinds of spectra we have discussed. An incandescent lightbulb produces a continuous spectrum. When that continuous spectrum is viewed through a thinner cloud of gas, an absorption line spectrum can be seen superimposed on the continuous spectrum. If we look only at a cloud of excited gas atoms (with no continuous source seen behind it), we see that the excited atoms give off an emission line spectrum.

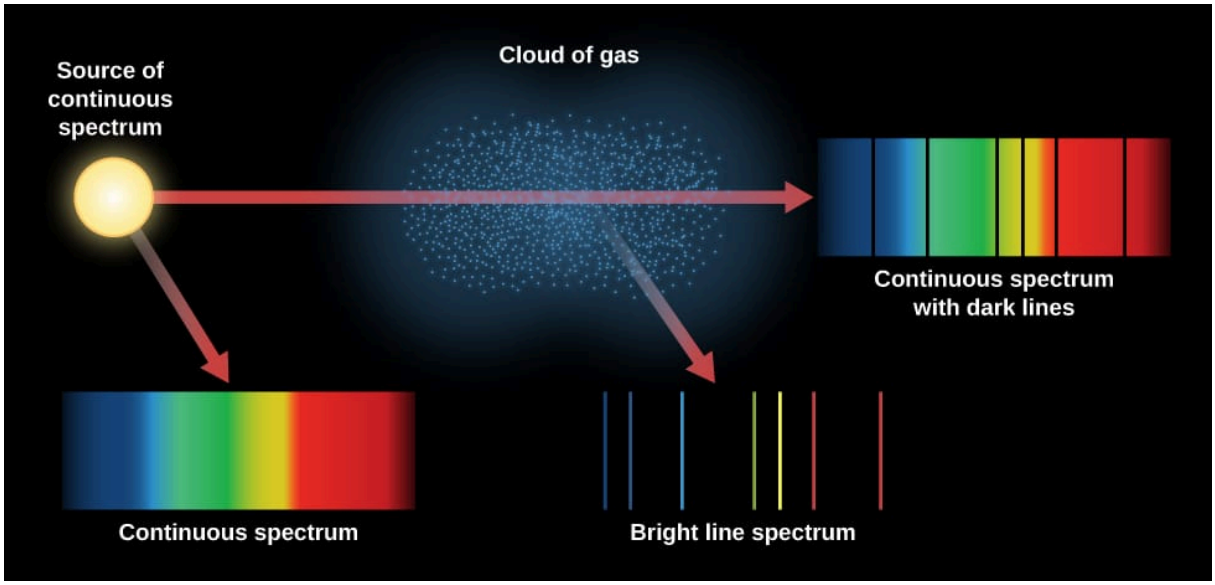


Figure 5.21 : Three Kinds of Spectra. When we see a lightbulb or other source of continuous radiation, all the colors are present. When the continuous spectrum is seen through a thinner gas cloud, the cloud's atoms produce absorption lines in the continuous spectrum. When the excited cloud is seen without the continuous source behind it, its atoms produce emission lines. We can learn which types of atoms are in the gas cloud from the pattern of absorption or emission lines.

Atoms in a hot gas are moving at high speeds and continually colliding with one another and with any loose electrons. They can be excited (electrons moving to a higher level) and de-excited (electrons moving to a lower level) by these collisions as well as by absorbing and emitting light. The speed of atoms in a gas depends on the temperature. When the temperature is higher, so are the speed and energy of the collisions. The hotter the gas, therefore, the more likely that electrons will occupy the outermost orbits, which correspond to the highest energy levels. This means that the level where electrons *start* their upward jumps in a gas can serve as an indicator of how hot that gas is. In this way, the absorption lines in a spectrum give astronomers information about the temperature of the regions where the lines originate.

Link to Learning

Use this simulation to play with a hydrogen atom and see what happens when electrons move to higher levels and then give off photons as they go to a lower level.

Ionization

We have described how certain discrete amounts of energy can be absorbed by an atom, raising it to an excited state and moving one of its electrons farther from its nucleus. If enough energy is absorbed, the electron can be completely removed from the atom—this is called ionization. The atom is then said to be ionized. The minimum amount of energy required to remove one electron from an atom in its ground state is called its ionization energy.

Still-greater amounts of energy must be absorbed by the now-ionized atom (called an ion) to remove an additional electron deeper in the structure of the atom. Successively greater energies are needed to remove the third, fourth, fifth—and so on—electrons from the atom. If enough energy is available, an atom can become completely ionized, losing all of its electrons. A hydrogen atom, having only one electron to lose, can be ionized only once; a helium atom can be ionized twice; and an oxygen atom up to eight times. When we examine regions of the cosmos where there is a great deal of energetic radiation, such as the neighborhoods where hot young stars have recently formed, we see a lot of ionization going on.

An atom that has become positively ionized has lost a negative charge—the missing electron—and thus is left with a net positive charge. It therefore exerts a strong attraction on any free electron. Eventually, one or more electrons will be captured and the atom will become neutral (or ionized to one less degree) again. During the electron-capture process, the atom emits one or more photons. Which photons are emitted depends on whether the electron is captured at once to the lowest energy level of the atom or stops at one or more intermediate levels on its way to the lowest available level.

Just as the excitation of an atom can result from a collision with another atom, ion, or electron (collisions with electrons are usually most important), so can ionization. The rate at which such collisional ionizations occur depends on the speeds of the atoms and hence on the temperature of the gas—the hotter the gas, the more of its atoms will be ionized.

The rate at which ions and electrons recombine also depends on their relative speeds—that is, on the temperature. In addition, it depends on the density of the gas: the higher the density, the greater the chance for recapture, because the different kinds of particles are crowded more closely together. From a knowledge of the temperature and density of a gas, it is possible to calculate the fraction of atoms that have been ionized once, twice, and so on. In the Sun, for example, we find that most of the hydrogen and helium atoms in its atmosphere are neutral, whereas most of the calcium atoms, as well as many other heavier atoms, are ionized once.

The energy levels of an ionized atom are entirely different from those of the same atom when it is neutral. Each time an electron is removed from the atom, the energy levels of the ion, and thus the wavelengths of the spectral lines it can produce, change. This helps astronomers differentiate the ions of a given element. Ionized hydrogen, having no electron, can produce no absorption lines.

This page titled 9.5: Formation of Spectral Lines is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.6: Formation of Spectral Lines by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.6: The Doppler Effect

Learning Objectives

By the end of this section, you will be able to:

- Explain why the spectral lines of photons we observe from an object will change as a result of the object's motion toward or away from us
- Describe how we can use the Doppler effect to deduce how fast astronomical objects are moving through space

The last two sections introduced you to many new concepts, and we hope that through those, you have seen one major idea emerge. Astronomers can learn about the elements in stars and galaxies by decoding the information in their spectral lines. There is a complicating factor in learning how to decode the message of starlight, however. If a star is moving toward or away from us, its lines will be in a slightly different place in the spectrum from where they would be in a star at rest. And most objects in the universe do have some motion relative to the Sun.

Motion Affects Waves

In 1842, Christian Doppler first measured the effect of motion on waves by hiring a group of musicians to play on an open railroad car as it was moving along the track. He then applied what he learned to all waves, including light, and pointed out that if a light source is approaching or receding from the observer, the light waves will be, respectively, crowded more closely together or spread out. The general principle, now known as the Doppler effect, is illustrated in Figure 5.22.

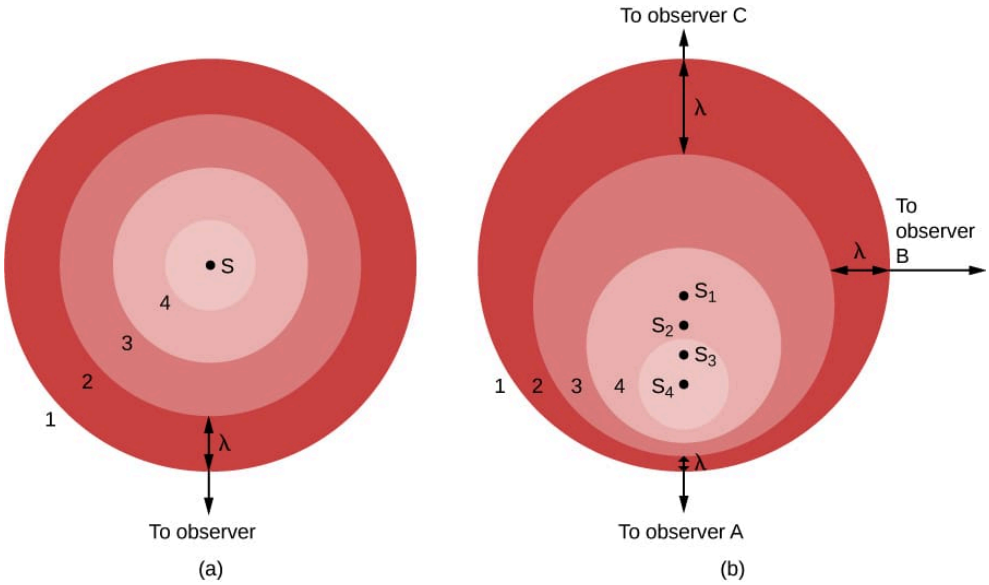


Figure 5.22 : Doppler Effect. (a) A source, S, makes waves whose numbered crests (1, 2, 3, and 4) wash over a stationary observer. (b) The source S now moves toward observer *A* and away from observer *C*. Wave crest 1 was emitted when the source was at position S1, crest 2 at position S2, and so forth. Observer *A* sees waves compressed by this motion and sees a blueshift (if the waves are light). Observer *C* sees the waves stretched out by the motion and sees a redshift. Observer *B*, whose line of sight is perpendicular to the source's motion, sees no change in the waves (and feels left out).

In part (a) of the figure, the light source (S) is at rest with respect to the observer. The source gives off a series of waves, whose crests we have labeled 1, 2, 3, and 4. The light waves spread out evenly in all directions, like the ripples from a splash in a pond. The crests are separated by a distance, λ, where λ is the wavelength. The observer, who happens to be located in the direction of the bottom of the image, sees the light waves coming nice and evenly, one wavelength apart. Observers located anywhere else would see the same thing.

On the other hand, if the source of light is moving with respect to the observer, as seen in part (b), the situation is more complicated. Between the time one crest is emitted and the next one is ready to come out, the source has moved a bit, toward the bottom of the page. From the point of view of observer *A*, this motion of the source has decreased the distance between crests—it's squeezing the crests together, this observer might say.

In part (b), we show the situation from the perspective of three observers. The source is seen in four positions, S_1, S_2, S_3, and S_4, each corresponding to the emission of one wave crest. To observer *A*, the waves seem to follow one another more closely, at a decreased wavelength and thus increased frequency. (Remember, all light waves travel at the speed of light through empty space, no matter what. This means that motion cannot affect the speed, but only the wavelength and the frequency. As the wavelength decreases, the frequency must increase. If the waves are shorter, more will be able to move by during each second.)

The situation is not the same for other observers. Let's look at the situation from the point of view of observer *C*, located opposite observer *A* in the figure. For her, the source is moving away from her location. As a result, the waves are not squeezed together but instead are spread out by the motion of the source. The crests arrive with an increased wavelength and decreased frequency. To observer *B*, in a direction at right angles to the motion of the source, no effect is observed. The wavelength and frequency remain the same as they were in part (a) of the figure.

We can see from this illustration that the Doppler effect is produced only by a motion toward or away from the observer, a motion called radial velocity. Sideways motion does not produce such an effect. Observers between *A* and *B* would observe some shortening of the light waves for that part of the motion of the source that is along their line of sight. Observers between *B* and *C* would observe lengthening of the light waves that are along their line of sight.

You may have heard the Doppler effect with sound waves. When a train whistle or police siren approaches you and then moves away, you will notice a decrease in the pitch (which is how human senses interpret sound wave frequency) of the sound waves. Compared to the waves at rest, they have changed from slightly more frequent when coming toward you, to slightly less frequent when moving away from you.

Link to Learning

A nice example of this change in the sound of a train whistle can be heard at the end of the classic Beach Boys song "Caroline, No" on their album *Pet Sounds*. To hear this sound, go to this YouTube version of the song. The sound of the train begins at approximately 2:20.

Color Shifts

When the source of waves moves toward you, the wavelength decreases a bit. If the waves involved are visible light, then the colors of the light change slightly. As wavelength decreases, they shift toward the blue end of the spectrum: astronomers call this a *blueshift* (since the end of the spectrum is really violet, the term should probably be *violetshift*, but blue is a more common color). When the source moves away from you and the wavelength gets longer, we call the change in colors a *redshift*. Because the Doppler effect was first used with visible light in astronomy, the terms "blueshift" and "redshift" became well established. Today, astronomers use these words to describe changes in the wavelengths of radio waves or X-rays as comfortably as they use them to describe changes in visible light.

The greater the motion toward or away from us, the greater the Doppler shift. If the relative motion is entirely along the line of sight, the formula for the Doppler shift of light is

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}$$

where λ is the wavelength emitted by the source, $\Delta\lambda$ is the difference between λ and the wavelength measured by the observer, *c* is the speed of light, and *v* is the relative speed of the observer and the source in the line of sight. The variable *v* is counted as positive if the velocity is one of recession, and negative if it is one of approach. Solving this equation for the velocity, we find $v = c \times \Delta\lambda/\lambda$.

Link to Learning

Click-and-drag the object emitting waves or the object receiving waves in this simulator to experiment with the Doppler effect yourself. The plots along the top show how the detected wavelengths change if the objects are approaching each other or moving further apart.

If a star approaches or recedes from us, the wavelengths of light in its continuous spectrum appear shortened or lengthened, respectively, as do those of the dark lines. However, unless its speed is tens of thousands of kilometers per second, the star does not appear noticeably bluer or redder than normal. The Doppler shift is thus not easily detected in a continuous spectrum and cannot be measured accurately in such a spectrum. The wavelengths of the absorption lines can be measured accurately, however, and their Doppler shift is relatively simple to detect.

Example 5.6: The Doppler Effect

We can use the Doppler effect equation to calculate the radial velocity of an object if we know three things: the speed of light, the original (unshifted) wavelength of the light emitted, and the difference between the wavelength of the emitted light and the wavelength we observe. For particular absorption or emission lines, we usually know exactly what wavelength the line has in our laboratories on Earth, where the source of light is not moving. We can measure the new wavelength with our instruments at the telescope, and so we know the difference in wavelength due to Doppler shifting. Since the speed of light is a universal constant, we can then calculate the radial velocity of the star.

A particular emission line of hydrogen is originally emitted with a wavelength of 656.3 nm from a gas cloud. At our telescope, we observe the wavelength of the emission line to be 656.6 nm. How fast is this gas cloud moving toward or away from Earth?

Solution

Because the light is shifted to a longer wavelength (redshifted), we know this gas cloud is moving away from us. The speed can be calculated using the Doppler shift formula:

$$\begin{aligned} \nu &= c \times \frac{\Delta\lambda}{\lambda} = \left(3.0 \times 10^8\ \mathrm{m/s}\right)\left(\frac{0.3\ \mathrm{nm}}{656.3\ \mathrm{nm}}\right) = \left(3.0 \times 10^8\ \mathrm{m/s}\right)\left(\frac{0.3 \times 10^{-9}\ \mathrm{m}}{656.3 \times 10^{-9}\ \mathrm{m}}\right) \\ &= 140{,}000\ \mathrm{m/s} = 140\ \mathrm{km/s} \end{aligned}$$

Exercise 9.6.1

Suppose a spectral line of hydrogen, normally at 500 nm, is observed in the spectrum of a star to be at 500.1 nm. How fast is the star moving toward or away from Earth?

Answer available at the end of this section.

You may now be asking: if all the stars are moving and motion changes the wavelength of each spectral line, won't this be a disaster for astronomers trying to figure out what elements are present in the stars? After all, it is the precise wavelength (or color) that tells astronomers which lines belong to which element. And we first measure these wavelengths in containers of gas in our laboratories, which are not moving. If every line in a star's spectrum is now shifted by its motion to a different wavelength (color), how can we be sure which lines and which elements we are looking at in a star whose speed we do not know?

Take heart. This situation sounds worse than it really is. Astronomers rarely judge the presence of an element in an astronomical object by a single line. It is the *pattern* of lines unique to hydrogen or calcium that enables us to determine that those elements are part of the star or galaxy we are observing. The Doppler effect does not change the pattern of lines from a given element—it only shifts the whole pattern slightly toward redder or bluer wavelengths. The shifted pattern is still quite easy to recognize. Best of all, when we do recognize a familiar element's pattern, we get a bonus: the amount the pattern is shifted can enable us to determine the speed of the objects in our line of sight.

The training of astronomers includes much work on learning to decode light (and other electromagnetic radiation). A skillful "decoder" can learn the temperature of a star, what elements are in it, and even its speed in a direction toward us or away from

us. That's really an impressive amount of information for stars that are light-years away.

Answers

Exercise 9.6.1

Because the light is shifted to a longer wavelength, the star is moving away from us:
$\nu = c \times \frac{\Delta\lambda}{\lambda} = (3.0 \times 10^8\,\mathrm{m/s})\left(\frac{0.1\,\mathrm{nm}}{500\,\mathrm{nm}}\right) = (3.0 \times 10^8\,\mathrm{m/s})\left(\frac{0.1 \times 10^{-9}\,\mathrm{m}}{500 \times 10^{-9}\,\mathrm{m}}\right) = 60{,}000\,\mathrm{m/s}$. Its speed is 60,000 m/s.

This page titled 9.6: The Doppler Effect is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 5.7: The Doppler Effect by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

9.7: Key Terms

absorption spectrum
: a series or pattern of dark lines superimposed on a continuous spectrum

blackbody
: an idealized object that absorbs all electromagnetic energy that falls onto it

continuous spectrum
: a spectrum of light composed of radiation of a continuous range of wavelengths or colors, rather than only certain discrete wavelengths

dispersion
: separation of different wavelengths of white light through refraction of different amounts

Doppler effect
: the apparent change in wavelength or frequency of the radiation from a source due to its relative motion away from or toward the observer

electromagnetic radiation
: radiation consisting of waves propagated through regularly varying electric and magnetic fields and traveling at the speed of light

electromagnetic spectrum
: the whole array or family of electromagnetic waves, from radio to gamma rays

emission spectrum
: a series or pattern of bright lines superimposed on a continuous spectrum

energy flux
: the amount of energy passing through a unit area (for example, 1 square meter) per second; the units of flux are watts per square meter

energy level
: a particular level, or amount, of energy possessed by an atom or ion above the energy it possesses in its least energetic state; also used to refer to the states of energy an electron can have in an atom

excitation
: the process of giving an atom or an ion an amount of energy greater than it has in its lowest energy (ground) state

frequency
: the number of waves that cross a given point per unit time (in radiation)

gamma rays
: photons (of electromagnetic radiation) of energy with wavelengths no longer than 0.01 nanometer; the most energetic form of electromagnetic radiation

ground state
: the lowest energy state of an atom

infrared
: electromagnetic radiation of wavelength 10^3–10^6 nanometers; longer than the longest (red) wavelengths that can be perceived by the eye, but shorter than radio wavelengths

inverse square law
: (for light) the amount of energy (light) flowing through a given area in a given time decreases in proportion to the square of the distance from the source of energy or light

ion
an atom that has become electrically charged by the addition or loss of one or more electrons

ionization
the process by which an atom gains or loses electrons

isotope
any of two or more forms of the same element whose atoms have the same number of protons but different numbers of neutrons

microwave
electromagnetic radiation of wavelengths from 1 millimeter to 1 meter; longer than infrared but shorter than radio waves

nucleus (of an atom)
the massive part of an atom, composed mostly of protons and neutrons, and about which the electrons revolve

photon
a discrete unit (or "packet") of electromagnetic energy

radial velocity
motion toward or away from the observer; the component of relative velocity that lies in the line of sight

radio waves
all electromagnetic waves longer than microwaves, including radar waves and AM radio waves

spectrometer
an instrument for obtaining a spectrum; in astronomy, usually attached to a telescope to record the spectrum of a star, galaxy, or other astronomical object

Stefan-Boltzmann law
a formula from which the rate at which a blackbody radiates energy can be computed; the total rate of energy emission from a unit area of a blackbody is proportional to the fourth power of its absolute temperature: $F = \sigma T^4$

ultraviolet
electromagnetic radiation of wavelengths 10 to 400 nanometers; shorter than the shortest visible wavelengths

visible light
electromagnetic radiation with wavelengths of roughly 400–700 nanometers; visible to the human eye

wavelength
the distance from crest to crest or trough to trough in a wave

Wien's law
formula that relates the temperature of a blackbody to the wavelength at which it emits the greatest intensity of radiation

X-rays
electromagnetic radiation with wavelengths between 0.01 nanometer and 20 nanometers; intermediate between those of ultraviolet radiation and gamma rays

This page titled 9.7: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

10: OBSERVING

10.1: Telescopes
10.2: Telescopes Today
10.3: Visible-Light Detectors and Instruments
10.4: Radio Telescopes
10.5: Observations outside Earth's Atmosphere
10.6: The Future of Large Telescopes
10.7: Key Terms

This page titled 10: Observing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

10.1: Telescopes

Learning Objectives

By the end of this section, you will be able to:

- Describe the three basic components of a modern system for measuring astronomical sources
- Describe the main functions of a telescope
- Describe the two basic types of visible-light telescopes and how they form images

Systems for Measuring Radiation

There are three basic components of a modern system for measuring radiation from astronomical sources. First, there is a telescope, which serves as a "bucket" for collecting visible light (or radiation at other wavelengths, as shown in (Figure 6.2). Just as you can catch more rain with a garbage can than with a coffee cup, large telescopes gather much more light than your eye can. Second, there is an instrument attached to the telescope that sorts the incoming radiation by wavelength. Sometimes the sorting is fairly crude. For example, we might simply want to separate blue light from red light so that we can determine the temperature of a star. But at other times, we want to see individual spectral lines to determine what an object is made of, or to measure its speed (as explained in the Radiation and Spectra chapter). Third, we need some type of detector, a device that senses the radiation in the wavelength regions we have chosen and permanently records the observations.

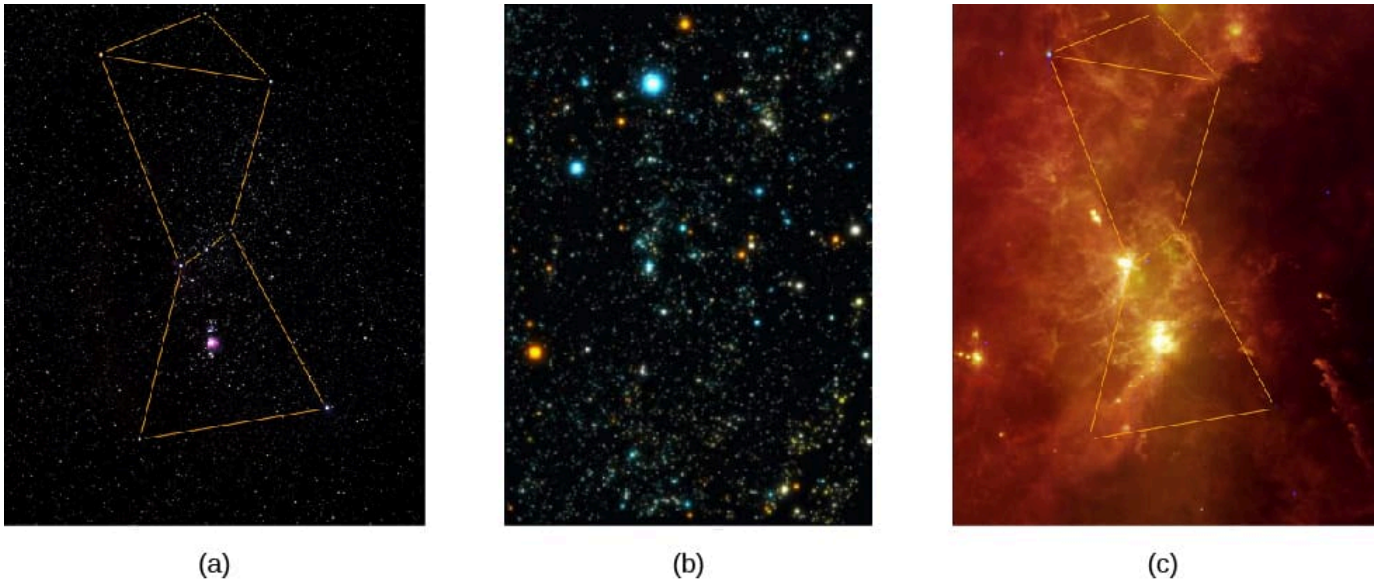


Figure 6.2 : Orion Region at Different Wavelengths. The same part of the sky looks different when observed with instruments that are sensitive to different bands of the spectrum. (a) Visible light: this shows part of the Orion region as the human eye sees it, with dotted lines added to show the figure of the mythical hunter, Orion. (b) X-rays: here, the view emphasizes the point-like X-ray sources nearby. The colors are artificial, changing from yellow to white to blue with increasing energy of the X-rays. The bright, hot stars in Orion are still seen in this image, but so are many other objects located at very different distances, including other stars, star corpses, and galaxies at the edge of the observable universe. (c) Infrared radiation: here, we mainly see the glowing dust in this region. (credit a: modification of work by Howard McCallon/NASA/IRAS; credit b: modification of work by Howard McCallon/NASA/IRAS; credit c: modification of work by Michael F. Corcoran)

The history of the development of astronomical telescopes is about how new technologies have been applied to improve the efficiency of these three basic components: the telescopes, the wavelength-sorting device, and the detectors. Let's first look at the development of the telescope.

Many ancient cultures built special sites for observing the sky (Figure 6.3). At these ancient *observatories*, they could measure the positions of celestial objects, mostly to keep track of time and date. Many of these ancient observatories had religious and

ritual functions as well. The eye was the only device available to gather light, all of the colors in the light were observed at once, and the only permanent record of the observations was made by human beings writing down or sketching what they saw.



(a) (b)

Figure 6.3 : Two Pre-Telescopic Observatories. (a) Machu Picchu is a fifteenth century Incan site located in Peru. (b) Stonehenge, a prehistoric site (3000–2000 BCE), is located in England. (credit a: modification of work by Allard Schmidt)

While Hans Lippershey, Zaccharias Janssen, and Jacob Metius are all credited with the invention of the telescope around 1608—applying for patents within weeks of each other—it was Galileo who, in 1610, used this simple tube with lenses (which he called a spyglass) to observe the sky and gather more light than his eyes alone could. Even his small telescope—used over many nights—revolutionized ideas about the nature of the planets and the position of Earth.

How Telescopes Work

Telescopes have come a long way since Galileo's time. Now they tend to be huge devices; the most expensive cost hundreds of millions to billions of dollars. (To provide some reference point, however, keep in mind that just renovating college football stadiums typically costs hundreds of millions of dollars—with the most expensive recent renovation, at Texas A&M University's Kyle Field, costing $450 million.) The reason astronomers keep building bigger and bigger telescopes is that celestial objects—such as planets, stars, and galaxies—send much more light to Earth than any human eye (with its tiny opening) can catch, and bigger telescopes can detect fainter objects. If you have ever watched the stars with a group of friends, you know that there's plenty of starlight to go around; each of you can see each of the stars. If a thousand more people were watching, each of them would also catch a bit of each star's light. Yet, as far as you are concerned, the light not shining into your eye is wasted. It would be great if some of this "wasted" light could also be captured and brought to your eye. This is precisely what a telescope does.

The most important functions of a telescope are (1) to *collect* the faint light from an astronomical source and (2) to *focus* all the light into a point or an image. Most objects of interest to astronomers are extremely faint: the more light we can collect, the better we can study such objects. (And remember, even though we are focusing on visible light first, there are many telescopes that collect other kinds of electromagnetic radiation.)

Telescopes that collect visible radiation use a lens or mirror to gather the light. Other types of telescopes may use collecting devices that look very different from the lenses and mirrors with which we are familiar, but they serve the same function. In all types of telescopes, the light-gathering ability is determined by the area of the device acting as the light-gathering "bucket." Since most telescopes have mirrors or lenses, we can compare their light-gathering power by comparing the apertures, or diameters, of the opening through which light travels or reflects.

The amount of light a telescope can collect increases with the size of the aperture. A telescope with a mirror that is 4 meters in diameter can collect 16 times as much light as a telescope that is 1 meter in diameter. (The diameter is squared because the area of a circle equals $\frac{\pi d^2}{4}$, where d is the diameter of the circle.)

Example 6.1: Calculating the Light-Collecting Area

What is the area of a 1-m diameter telescope? A 4-m diameter one?

Solution

Using the equation for the area of a circle,

$$A =$$

the area of a 1-m telescope is

$$\frac{\pi d^2}{4} = \frac{\pi(1\,\mathrm{m})^2}{4} = 0.79\,\mathrm{m}^2$$

and the area of a 4-m telescope is

$$\frac{\pi d^2}{4} = \frac{\pi(4\,\mathrm{m})^2}{4} = 12.6\,\mathrm{m}^2$$

Exercise 10.1.1

Show that the ratio of the two areas is 16:1.

Answer available at the end of this section.

After the telescope forms an image, we need some way to detect and record it so that we can measure, reproduce, and analyze the image in various ways. Before the nineteenth century, astronomers simply viewed images with their eyes and wrote descriptions of what they saw. This was very inefficient and did not lead to a very reliable long-term record; you know from crime shows on television that eyewitness accounts are often inaccurate.

In the nineteenth century, the use of photography became widespread. In those days, photographs were a chemical record of an image on a specially treated glass plate. Today, the image is generally detected with sensors similar to those in digital cameras, recorded electronically, and stored in computers. This permanent record can then be used for detailed and quantitative studies. Professional astronomers rarely look through the large telescopes that they use for their research.

Formation of an Image by a Lens or a Mirror

Whether or not you wear glasses, you see the world through lenses; they are key elements of your eyes. A lens is a transparent piece of material that bends the rays of light passing through it. If the light rays are parallel as they enter, the lens brings them together in one place to form an image (Figure 6.4). If the curvatures of the lens surfaces are just right, all parallel rays of light (say, from a star) are bent, or *refracted*, in such a way that they converge toward a point, called the focus of the lens. At the focus, an image of the light source appears. In the case of parallel light rays, the distance from the lens to the location where the light rays focus, or image, behind the lens is called the *focal length* of the lens.

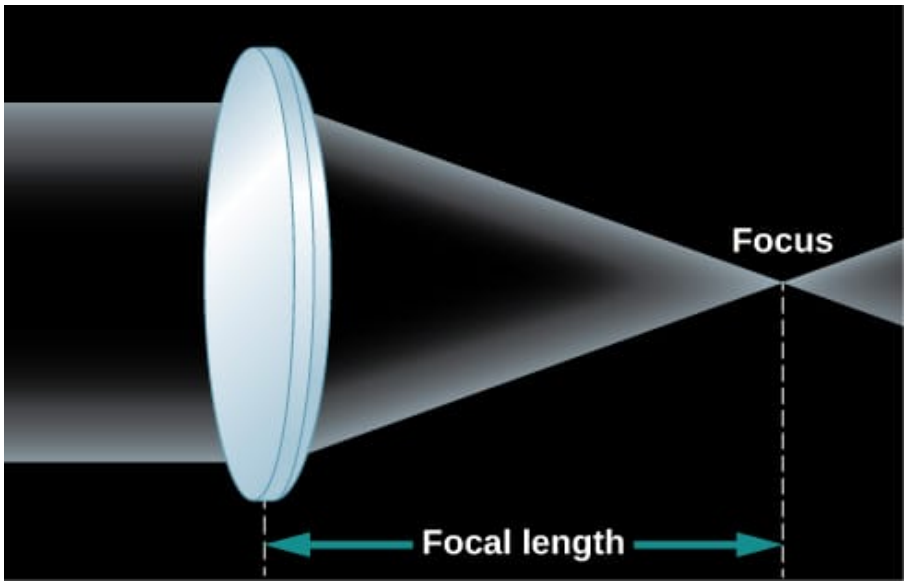


Figure 6.4 : Formation of an Image by a Simple Lens. Parallel rays from a distant source are bent by the convex lens so that they all come together in a single place (the focus) to form an image.

As you look at Figure 6.4, you may ask why two rays of light from the same star would be parallel to each other. After all, if you draw a picture of star shining in all directions, the rays of light coming from the star don't look parallel at all. But remember that the stars (and other astronomical objects) are all extremely far away. By the time the few rays of light pointed toward us actually arrive at Earth, they are, for all practical purposes, parallel to each other. Put another way, any rays that were *not* parallel to the ones pointed at Earth are now heading in some very different direction in the universe.

To view the image formed by the lens in a telescope, we use an additional lens called an eyepiece. The eyepiece focuses the image at a distance that is either directly viewable by a human or at a convenient place for a detector. Using different eyepieces, we can change the *magnification* (or size) of the image and also redirect the light to a more accessible location. Stars look like points of light, and magnifying them makes little difference, but the image of a planet or a galaxy, which has structure, can often benefit from being magnified.

Many people, when thinking of a telescope, picture a long tube with a large glass lens at one end. This design, which uses a lens as its main optical element to form an image, as we have been discussing, is known as a *refractor* (Figure 6.5), and a telescope based on this design is called a refracting telescope. Galileo's telescopes were refractors, as are today's binoculars and field glasses. However, there is a limit to the size of a refracting telescope. The largest one ever built was a 49-inch refractor built for the Paris 1900 Exposition, and it was dismantled after the Exposition. Currently, the largest refracting telescope is the 40-inch refractor at Yerkes Observatory in Wisconsin.

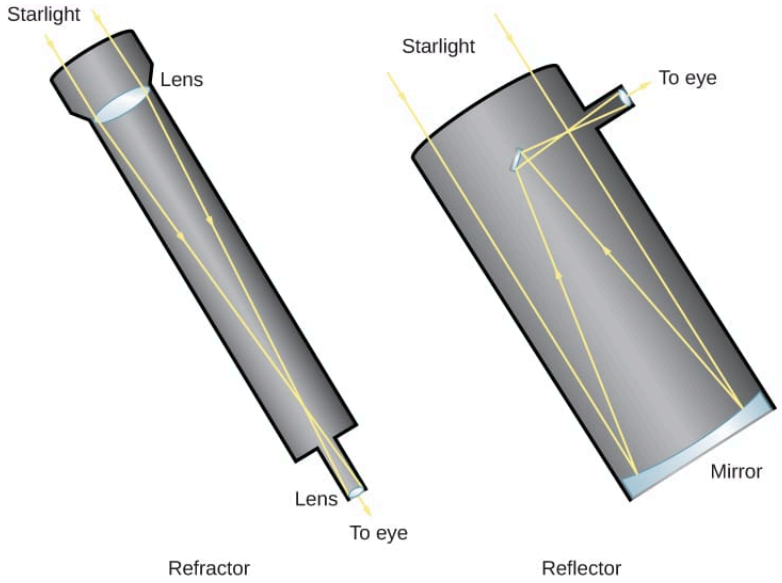


Figure 6.5 : Refracting and Reflecting Telescopes. Light enters a refracting telescope through a lens at the upper end, which focuses the light near the bottom of the telescope. An eyepiece then magnifies the image so that it can be viewed by the eye, or a detector like a photographic plate can be placed at the focus. The upper end of a reflecting telescope is open, and the light passes through to the mirror located at the bottom of the telescope. The mirror then focuses the light at the top end, where it can be detected. Alternatively, as in this sketch, a second mirror may reflect the light to a position outside the telescope structure, where an observer can have easier access to it. Professional astronomers' telescopes are more complicated than this, but they follow the same principles of reflection and refraction.

One problem with a refracting telescope is that the light must pass *through* the lens of a refractor. That means the glass must be perfect all the way through, and it has proven very difficult to make large pieces of glass without flaws and bubbles in them. Also, optical properties of transparent materials change a little bit with the wavelengths (or colors) of light, so there is some additional distortion, known as chromatic aberration. Each wavelength focuses at a slightly different spot, causing the image to appear blurry.

In addition, since the light must pass through the lens, the lens can only be supported around its edges (just like the frames of our eyeglasses). The force of gravity will cause a large lens to sag and distort the path of the light rays as they pass through it. Finally, because the light passes through it, both sides of the lens must be manufactured to precisely the right shape in order to produce a sharp image.

A different type of telescope uses a concave *primary mirror* as its main optical element. The mirror is curved like the inner surface of a sphere, and it reflects light in order to form an image (Figure 6.5). Telescope mirrors are coated with a shiny metal, usually silver, aluminum, or, occasionally, gold, to make them highly reflective. If the mirror has the correct shape, all parallel rays are reflected back to the same point, the focus of the mirror. Thus, images are produced by a mirror exactly as they are by a lens.

Telescopes designed with mirrors avoid the problems of refracting telescopes. Because the light is reflected from the front surface only, flaws and bubbles within the glass do not affect the path of the light. In a telescope designed with mirrors, only the front surface has to be manufactured to a precise shape, and the mirror can be supported from the back. For these reasons, most astronomical telescopes today (both amateur and professional) use a mirror rather than a lens to form an image; this type of telescope is called a reflecting telescope. The first successful reflecting telescope was built by Isaac Newton in 1668.

In a reflecting telescope, the concave mirror is placed at the bottom of a tube or open framework. The mirror reflects the light back up the tube to form an image near the front end at a location called the prime focus. The image can be observed at the prime focus, or additional mirrors can intercept the light and redirect it to a position where the observer can view it more easily (Figure 6.6). Since an astronomer at the prime focus can block much of the light coming to the main mirror, the use of a small *secondary mirror* allows more light to get through the system.

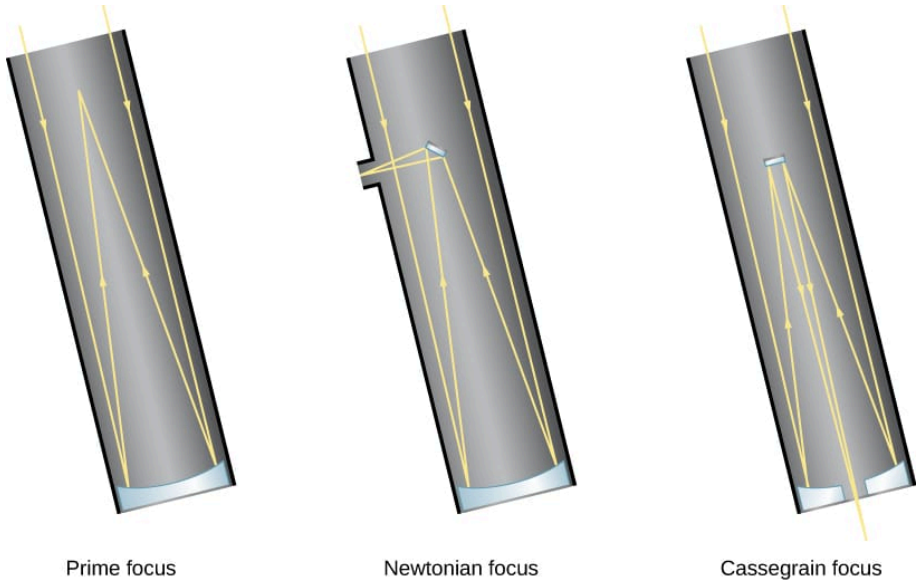


Figure 6.6 : Focus Arrangements for Reflecting Telescopes. Reflecting telescopes have different options for where the light is brought to a focus. With prime focus, light is detected where it comes to a focus after reflecting from the primary mirror. With Newtonian focus, light is reflected by a small secondary mirror off to one side, where it can be detected (see also Figure 6.5). Most large professional telescopes have a Cassegrain focus in which light is reflected by the secondary mirror down through a hole in the primary mirror to an observing station below the telescope.

Making Connections: Choosing Your Own Telescope

If the astronomy course you are taking whets your appetite for exploring the sky further, you may be thinking about buying your own telescope. Many excellent amateur telescopes are available, and some research is required to find the best model for your needs. Some good sources of information about personal telescopes are the two popular US magazines aimed at amateur astronomers: *Sky & Telescope* and *Astronomy*. Both carry regular articles with advice, reviews, and advertisements from reputable telescope dealers.

Some of the factors that determine which telescope is right for you depend upon your preferences:

- Will you be setting up the telescope in one place and leaving it there, or do you want an instrument that is portable and can come with you on outdoor excursions? How portable should it be, in terms of size and weight?
- Do you want to observe the sky with your eyes only, or do you want to take photographs? (Long-exposure photography, for example, requires a good clock drive to turn your telescope to compensate for Earth's rotation.)
- What types of objects will you be observing? Are you interested primarily in comets, planets, star clusters, or galaxies, or do you want to observe all kinds of celestial sights?

You may not know the answers to some of these questions yet. For this reason, you may want to "test-drive" some telescopes first. Most communities have amateur astronomy clubs that sponsor star parties open to the public. The members of those clubs often know a lot about telescopes and can share their ideas with you. Your instructor may know where the nearest amateur astronomy club meets; or, to find a club near you, use the websites suggested in Appendix B.

Furthermore, you may already have an instrument like a telescope at home (or have access to one through a relative or friend). Many amateur astronomers recommend starting your survey of the sky with a good pair of binoculars. These are easily carried around and can show you many objects not visible (or clear) to the unaided eye.

When you are ready to purchase a telescope, you might find the following ideas useful:

- The key characteristic of a telescope is the aperture of the main mirror or lens; when someone says they have a 6-inch or 8-inch telescope, they mean the diameter of the collecting surface. The larger the aperture, the more light you can gather, and the fainter the objects you can see or photograph.

- Telescopes of a given aperture that use lenses (refractors) are typically more expensive than those using mirrors (reflectors) because both sides of a lens must be polished to great accuracy. And, because the light passes through it, the lens must be made of high-quality glass throughout. In contrast, only the front surface of a mirror must be accurately polished.
- Magnification is not one of the criteria on which to base your choice of a telescope. As we discussed, the magnification of the image is done by a smaller eyepiece, so the magnification can be adjusted by changing eyepieces. However, a telescope will magnify not only the astronomical object you are viewing but also the turbulence of Earth's atmosphere. If the magnification is too high, your image will shimmer and shake and be difficult to view. A good telescope will come with a variety of eyepieces that stay within the range of useful magnification.
- The mount of a telescope (the structure on which it rests) is one of its most critical elements. Because a telescope shows a tiny field of view, which is magnified significantly, even the smallest vibration or jarring of the telescope can move the object you are viewing around or out of your field of view. A sturdy and stable mount is essential for serious viewing or photography (although it clearly affects how portable your telescope can be).
- A telescope requires some practice to set up and use effectively. Don't expect everything to go perfectly on your first try. Take some time to read the instructions. If a local amateur astronomy club is nearby, use it as a resource.

Answers

Exercise 10.1.1

$\frac{12.6\,\mathrm{m}^2}{0.79\,\mathrm{m}^2} = 16$. Therefore, with 16 times the area, a 4-m telescope collects 16 times the light of a 1-m telescope.

This page titled 10.1: Telescopes is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.2: Telescopes by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.2: Telescopes Today

Learning Objectives

By the end of this section, you will be able to:

- Recognize the largest visible-light and infrared telescopes in operation today
- Discuss the factors relevant to choosing an appropriate telescope site
- Define the technique of adaptive optics and describe the effects of the atmosphere on astronomical observations

Since Newton's time, when the sizes of the mirrors in telescopes were measured in inches, reflecting telescopes have grown ever larger. In 1948, US astronomers built a telescope with a 5-meter (200-inch) diameter mirror on Palomar Mountain in Southern California. It remained the largest visible-light telescope in the world for several decades. The giants of today, however, have primary mirrors (the largest mirrors in the telescope) that are 8- to 10-meters in diameter, and larger ones are being built (Figure 6.7).

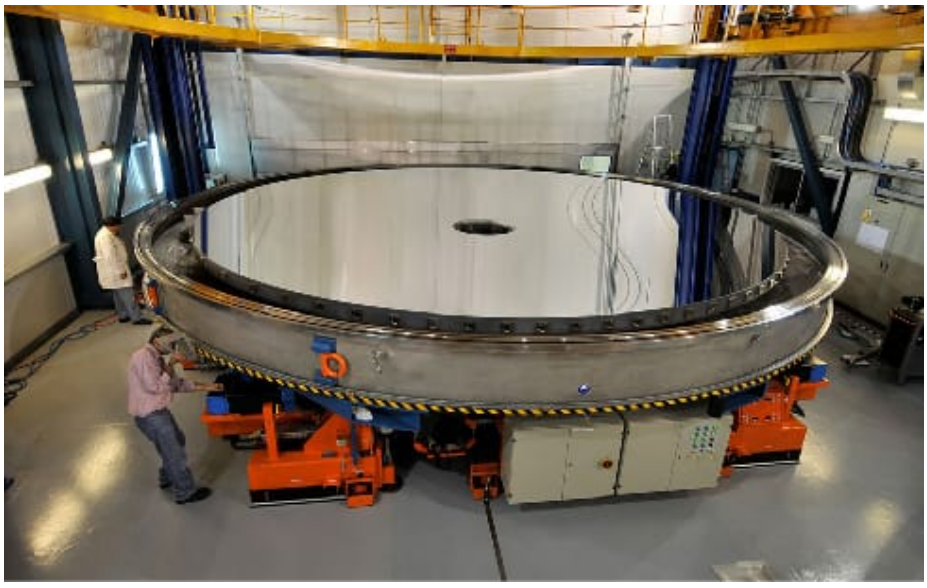


Figure 6.7 : Large Telescope Mirror. This image shows one of the primary mirrors of the European Southern Observatory's Very Large Telescope, named Yepun, just after it was recoated with aluminum. The mirror is a little over 8 meters in diameter. (credit: ESO/G. Huedepohl)

Modern Visible-Light and Infrared Telescopes

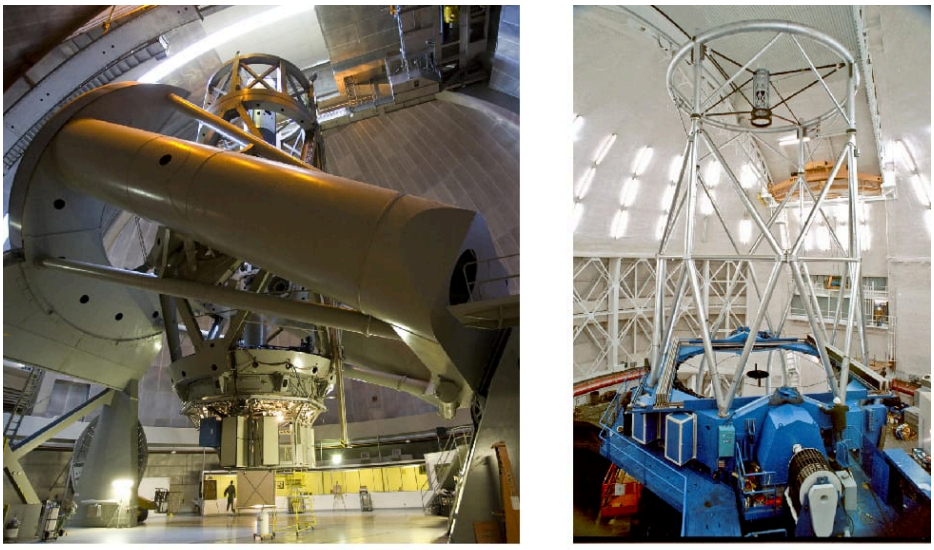
The decades starting in 1990 saw telescope building around the globe grow at an unprecedented rate. (See Table 6.1, which also includes websites for each telescope in case you want to visit or learn more about them.) Technological advancements had finally made it possible to build telescopes significantly larger than the 5-meter telescope at Palomar at a reasonable cost. New technologies have also been designed to work well in the infrared, and not just visible, wavelengths.

Table 6.1: Large Single-Dish Visible-Light and Infrared Telescopes

Aperture (m)	Telescope Name	Location	Status	Website
39	European Extremely Large Telescope (E-ELT)	Cerro Armazonas, Chile	First light 2025 (estimated)	www.eso.org/sci/facilities/eelt
30	Thirty-Meter Telescope (TMT)	Maunakea, HI	First light 2025 (estimated)	www.tmt.org

Aperture (m)	Telescope Name	Location	Status	Website
24.5	Giant Magellan Telescope (GMT)	Las Campanas Observatory, Chile	First light 2025 (estimated)	www.gmto.org
11.1×9.9	Southern African Large Telescope (SALT)	Sutherland, South Africa	2005	www.salt.ac.za
10.4	Gran Telescopio Canarias (GTC)	La Palma, Canary Islands	First light 2007	http://www.gtc.iac.es
10.0	Keck I and II (two telescopes)	Maunakea, HI	Completed 1993–96	www.keckobservatory.org
9.1	Hobby–Eberly Telescope (HET)	Mount Locke, TX	Completed 1997	www.as.utexas.edu/mcdonald/het
8.4	Large Binocular Telescope (LBT) (two telescopes)	Mount Graham, AZ	First light 2004	www.lbto.org
8.4	Large Synoptic Survey Telescope (Rubin Obs.)	The Cerro Pachón, Chile	First light 2021	www.lsst.org
8.3	Subaru Telescope	Maunakea, HI	First light 1998	www.naoj.org
8.2	Very Large Telescope (VLT)	Cerro Paranal, Chile	All four telescopes completed 2000	www.eso.org/public/teles-instr/paranal
8.1	Gemini North and Gemini South	Maunakea, HI (North) and Cerro Pachón, Chile (South)	First light 1999 (North), First light 2000 (South)	www.gemini.edu
6.5	Magellan Telescopes (two telescopes: Baade and Landon Clay)	Las Campanas, Chile	First light 2000 and 2002	obs.carnegiescience.edu/Magellan
6.5	Multi-Mirror Telescope (MMT)	Mount Hopkins, AZ	Completed 1979	www.mmto.org
6.0	Big Telescope Altazimuth (BTA-6)	Mount Pastukhov, Russia	Completed 1976	w0.sao.ru/Doc-en/Telescopes/bta/descrip.html
5.1	Hale Telescope	Mount Palomar, CA	Completed 1948	www.astro.caltech.edu/palomar/about/telescopes/hale.html

The differences between the Palomar telescope and the modern Gemini North telescope (to take an example) are easily seen in Figure 6.8. The Palomar telescope is a massive steel structure designed to hold the 14.5-ton primary mirror with a 5-meter diameter. Glass tends to sag under its own weight; hence, a huge steel structure is needed to hold the mirror. A mirror 8 meters in diameter, the size of the Gemini North telescope, if it were built using the same technology as the Palomar telescope, would have to weigh at least eight times as much and would require an enormous steel structure to support it.



(a) (b)

Figure 6.8 : Modern Reflecting Telescopes. (a) The Palomar 5-meter reflector: The Hale telescope on Palomar Mountain has a complex mounting structure that enables the telescope (in the open “tube” pointing upward in this photo) to swing easily into any position. (b) The Gemini North 8-meter telescope: The Gemini North mirror has a larger area than the Palomar mirror, but note how much less massive the whole instrument seems. (credit a: modification of work by Caltech/Palomar Observatory; credit b: modification of work by Gemini Observatory/AURA)

The 8-meter Gemini North telescope looks like a featherweight by contrast, and indeed it is. The mirror is only about 8 inches thick and weighs 24.5 tons, less than twice as much as the Palomar mirror. The Gemini North telescope was completed about 50 years after the Palomar telescope. Engineers took advantage of new technologies to build a telescope that is much lighter in weight relative to the size of the primary mirror. The Gemini mirror does sag, but with modern computers, it is possible to measure that sag many times each second and apply forces at 120 different locations to the back of the mirror to correct the sag, a process called *active control*. Seventeen telescopes with mirrors 6.5 meters in diameter and larger have been constructed since 1990.

The twin 10-meter Keck telescopes on Maunakea, which were the first of these new-technology instruments, use precision control in an entirely novel way. Instead of a single primary mirror 10 meters in diameter, each Keck telescope achieves its larger aperture by combining the light from 36 separate hexagonal mirrors, each 1.8 meters wide (Figure 6.9). Computer-controlled actuators (motors) constantly adjust these 36 mirrors so that the overall reflecting surface acts like a single mirror with just the right shape to collect and focus the light into a sharp image.

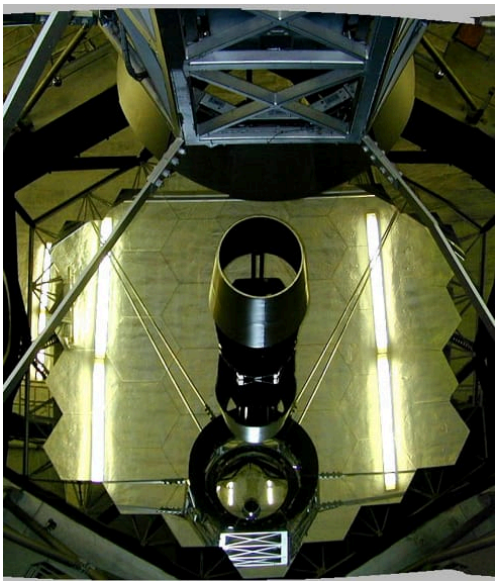


Figure 6.9 : Thirty-Six Eyes Are Better Than One. The mirror of the 10-meter Keck telescope is composed of 36 hexagonal sections. (credit: NASA)

Link to Learning

Learn more about the Keck Observatory on Maunakea through this History Channel clip on the telescopes and the work that they do.

In addition to holding the mirror, the steel structure of a telescope is designed so that the entire telescope can be pointed quickly toward any object in the sky. Since Earth is rotating, the telescope must have a motorized drive system that moves it very smoothly from east to west at exactly the same rate that Earth is rotating from west to east, so it can continue to point at the object being observed. All this machinery must be housed in a dome to protect the telescope from the elements. The dome has an opening in it that can be positioned in front of the telescope and moved along with it, so that the light from the objects being observed is not blocked.

Voyagers in Astronomy: George Ellery Hale: Master Telescope Builder

George Ellery Hale (Figure 6.10) was a giant among early telescope builders. Not once, but four times, he initiated projects that led to the construction of what was the world's largest telescope at the time. And he was a master at winning over wealthy benefactors to underwrite the construction of these new instruments.



Figure 6.10 : George Ellery Hale (1868–1938). Hale's work led to the construction of several major telescopes, including the 40-inch refracting telescope at Yerkes Observatory, and three reflecting telescopes: the 60-inch Hale and 100-inch Hooker telescopes at Mount Wilson Observatory, and the 200-inch Hale Telescope at Palomar Observatory.

Hale's training and early research were in solar physics. In 1892, at age 24, he was named associate professor of astral physics and director of the astronomical observatory at the University of Chicago. At the time, the largest telescope in the world was the 36-inch refractor at the Lick Observatory near San Jose, California. Taking advantage of an existing glass blank for a 40-inch telescope, Hale set out to raise money for a larger telescope than the one at Lick. One prospective donor was Charles T. Yerkes, who, among other things, ran the trolley system in Chicago.

Hale wrote to Yerkes, encouraging him to support the construction of the giant telescope by saying that "the donor could have no more enduring monument. It is certain that Mr. Lick's name would not have been nearly so widely known today were it not for the famous observatory established as a result of his munificence." Yerkes agreed, and the new telescope was completed in May 1897; it remains the largest refractor in the world (Figure 6.11).



Figure 6.11 : World's Largest Refractor. The Yerkes 40-inch (1-meter) telescope.

Even before the completion of the Yerkes refractor, Hale was not only dreaming of building a still larger telescope but was also taking concrete steps to achieve that goal. In the 1890s, there was a major controversy about the relative quality of refracting and reflecting telescopes. Hale realized that 40 inches was close to the maximum feasible aperture for refracting telescopes. If telescopes with significantly larger apertures were to be built, they would have to be reflecting telescopes.

Using funds borrowed from his own family, Hale set out to construct a 60-inch reflector. For a site, he left the Midwest for the much better conditions on Mount Wilson—at the time, a wilderness peak above the small city of Los Angeles. In 1904, at the age of 36, Hale received funds from the Carnegie Foundation to establish the Mount Wilson Observatory. The 60-inch mirror was placed in its mount in December 1908.

Two years earlier, in 1906, Hale had already approached John D. Hooker, who had made his fortune in hardware and steel pipe, with a proposal to build a 100-inch telescope. The technological risks were substantial. The 60-inch telescope was not yet complete, and the usefulness of large reflectors for astronomy had yet to be demonstrated. George Ellery Hale's brother called him "the greatest gambler in the world." Once again, Hale successfully obtained funds, and the 100-inch telescope was completed in November 1917. (It was with this telescope that Edwin Hubble was able to establish that the spiral nebulae were separate islands of stars—or galaxies—quite removed from our own Milky Way.)

Hale was not through dreaming. In 1926, he wrote an article in *Harper's Magazine* about the scientific value of a still larger telescope. This article came to the attention of the Rockefeller Foundation, which granted $6 million for the construction of a 200-inch telescope. Hale died in 1938, but the 200-inch (5-meter) telescope on Palomar Mountain was dedicated 10 years later and is now named in Hale's honor.

Picking the Best Observing Sites

A telescope like the Gemini or Keck telescope costs about $100 million to build. That kind of investment demands that the telescope be placed in the best possible site. Since the end of the nineteenth century, astronomers have realized that the best observatory sites are on mountains, far from the lights and pollution of cities. Although a number of urban observatories remain, especially in the large cities of Europe, they have become administrative centers or museums. The real action takes place far away, often on desert mountains or isolated peaks in the Atlantic and Pacific Oceans, where we find the staff's living quarters, computers, electronic and machine shops, and of course the telescopes themselves. A large observatory today requires a supporting staff of 20 to 100 people in addition to the astronomers.

The performance of a telescope is determined not only by the size of its mirror but also by its location. Earth's atmosphere, so vital to life, presents challenges for the observational astronomer. In at least four ways, our air imposes limitations on the usefulness of telescopes:

1. The most obvious limitation is weather conditions such as clouds, wind, and rain. At the best sites, the weather is clear as much as 75% of the time.
2. Even on a clear night, the atmosphere filters out a certain amount of starlight, especially in the infrared, where the absorption is due primarily to water vapor. Astronomers therefore prefer dry sites, generally found at high altitudes.
3. The sky above the telescope should be dark. Near cities, the air scatters the glare from lights, producing an illumination that hides the faintest stars and limits the distances that can be probed by telescopes. (Astronomers call this effect light pollution.) Observatories are best located at least 100 miles from the nearest large city.
4. Finally, the air is often unsteady; light passing through this turbulent air is disturbed, resulting in blurred star images. Astronomers call these effects "bad seeing." When seeing is bad, images of celestial objects are distorted by the constant twisting and bending of light rays by turbulent air.

The best observatory sites are therefore high, dark, and dry. The world's largest telescopes are found in such remote mountain locations as the Andes Mountains of Chile (Figure 6.12), the desert peaks of Arizona, the Canary Islands in the Atlantic Ocean, and Maunakea in Hawaii, a dormant volcano with an altitude of 13,700 feet (4200 meters).

Link to Learning

Light pollution is a problem not just for professional astronomers but for everyone who wants to enjoy the beauty of the night sky. In addition research is now showing that it can disrupt the life cycle of animals with whom we share the urban and suburban landscape. And the light wasted shining into the sky leads to unnecessary municipal expenses and use of fossil fuels. Concerned people have formed an organization, the International Dark-Sky Association, whose website is full of good information. A citizen science project called *Globe at Night* allows you to measure the light levels in your community by counting stars and to compare it to others around the world. Recently, a new threat to dark skies has emerged. A number of commercial companies are planning and launching large swarms of satellites into low-Earth orbit to provide internet access around the world. Unfortunately, these satellites reflect sunlight and have already begun to ruin astronomical photographs

the world over. See the Further Exploration section at the end of this chapter for more resources on this topic. If you get interested in this topic and want to do a paper for your astronomy course or another course while you are in college, the *Dark Night Skies* guide can point you to a variety of resources on the topic.



Figure 6.12 : High and Dry Site. Cerro Paranal, a mountain summit 2.7 kilometers above sea level in Chile's Atacama Desert, is the site of the European Southern Observatory's Very Large Telescope. This photograph shows the four 8-meter telescope buildings on the site and vividly illustrates that astronomers prefer high, dry sites for their instruments. The 4.1-meter Visible and Infrared Survey Telescope for Astronomy (VISTA) can be seen in the distance on the next mountain peak. (credit: ESO)

The Resolution of a Telescope

In addition to gathering as much light as they can, astronomers also want to have the sharpest images possible. Resolution refers to the precision of detail present in an image: that is, the smallest features that can be distinguished. Astronomers are always eager to make out more detail in the images they study, whether they are following the weather on Jupiter or trying to peer into the violent heart of a "cannibal galaxy" that recently ate its neighbor for lunch.

One factor that determines how good the resolution will be is the size of the telescope. Larger apertures produce sharper images. Until very recently, however, visible-light and infrared telescopes on Earth's surface could not produce images as sharp as the theory of light said they should.

The problem—as we saw earlier in this chapter—is our planet's atmosphere, which is turbulent. It contains many small-scale blobs or cells of gas that range in size from inches to several feet. Each cell has a slightly different temperature from its neighbor, and each cell acts like a lens, bending (refracting) the path of the light by a small amount. This bending slightly changes the position where each light ray finally reaches the detector in a telescope. The cells of air are in motion, constantly being blown through the light path of the telescope by winds, often in different directions at different altitudes. As a result, the path followed by the light is constantly changing.

For an analogy, think about watching a parade from a window high up in a skyscraper. You decide to throw some confetti down toward the marchers. Even if you drop a handful all at the same time and in the same direction, air currents will toss the pieces around, and they will reach the ground at different places. As we described earlier, we can think of the light from the stars as a series of parallel beams, each making its way through the atmosphere. Each path will be slightly different, and each will reach the detector of the telescope at a slightly different place. The result is a blurred image, and because the cells are being blown by the wind, the nature of the blur will change many times each second. You have probably noticed this effect as the "twinkling" of stars seen from Earth. The light beams are bent enough that part of the time they reach your eye, and part of the time some of them miss, thereby making the star seem to vary in brightness. In space, however, the light of the stars is steady.

Astronomers search the world for locations where the amount of atmospheric blurring, or turbulence, is as small as possible. It turns out that the best sites are in coastal mountain ranges and on isolated volcanic peaks in the middle of an ocean. Air that has

flowed long distances over water before it encounters land is especially stable.

The resolution of an image is measured in units of angle on the sky, typically in units of arcseconds. One arcsecond is 1/3600 degree, and there are 360 degrees in a full circle. So we are talking about tiny angles on the sky. To give you a sense of just how tiny, we might note that 1 arcsecond is how big a quarter would look when seen from a distance of 5 kilometers. The best images obtained from the ground with traditional techniques reveal details as small as several tenths of an arcsecond across. This image size is remarkably good. One of the main reasons for launching the Hubble Space Telescope was to escape Earth's atmosphere and obtain even sharper images.

But since we can't put every telescope into space, astronomers have devised a technique called adaptive optics that can beat Earth's atmosphere at its own game of blurring. This technique (which is most effective in the infrared region of the spectrum with our current technology) makes use of a small flexible mirror placed in the beam of a telescope. A sensor measures how much the atmosphere has distorted the image, and as often as 500 times per second, it sends instructions to the flexible mirror on how to change shape in order to compensate for distortions produced by the atmosphere. The light is thus brought back to an almost perfectly sharp focus at the detector. Figure 6.13 shows just how effective this technique is. With adaptive optics, ground-based telescopes can achieve resolutions of 0.1 arcsecond or a little better in the infrared region of the spectrum. This impressive figure is the equivalent of the resolution that the Hubble Space Telescope achieves in the visible-light region of the spectrum.

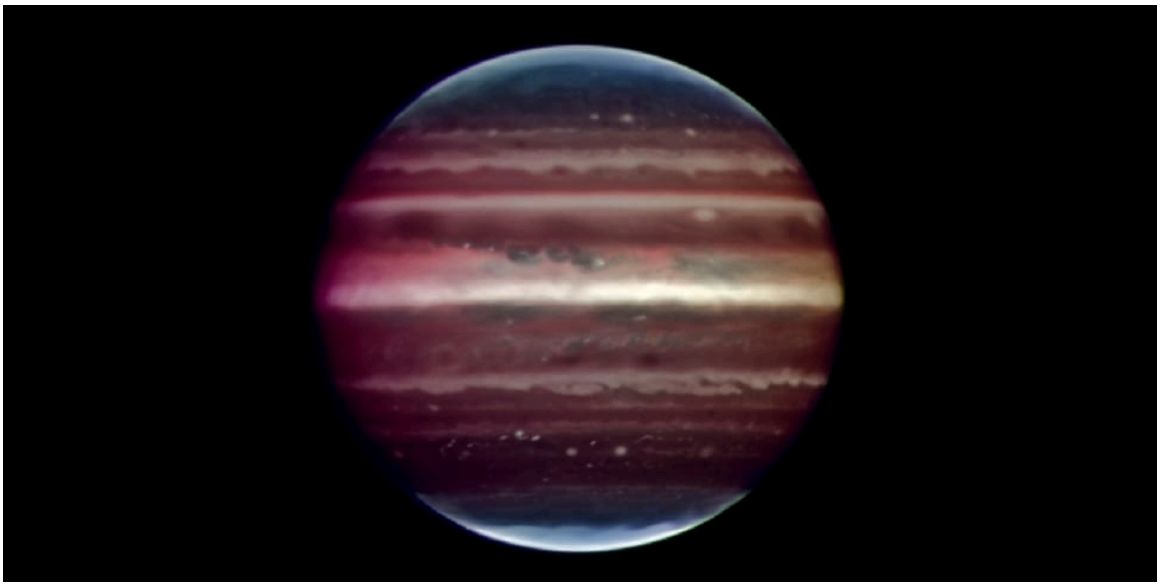


Figure 6.13 : Power of Adaptive Optics. One of the clearest pictures of Jupiter ever taken from the ground, this image was produced with adaptive optics using an 8-meter-diameter telescope at the Very Large Telescope in Chile. Adaptive optics uses infrared wavelengths to remove atmospheric blurring, resulting in a much clearer image. (credit: modification of work by ESO, F.Marchis, M.Wong (UC Berkeley); E.Marchetti, P.Amico, S.Tordo (ESO))

Astronomy Basics: How Astronomers Really Use Telescopes

In the popular view (and some bad movies), an astronomer spends most nights in a cold observatory peering through a telescope, but this is not very accurate today. Most astronomers do not live at observatories, but near the universities or laboratories where they work. An astronomer might spend only a week or so each year observing at the telescope and the rest of the time measuring or analyzing the data acquired from large project collaborations and dedicated surveys. Many astronomers use radio telescopes for space experiments, which work just as well during the daylight hours. Still others work at purely theoretical problems using supercomputers and never observe at a telescope of any kind.

Even when astronomers are observing with large telescopes, they seldom peer through them. Electronic detectors permanently record the data for detailed analysis later. At some observatories, observations may be made remotely, with the astronomer sitting at a computer thousands of miles away from the telescope.

Time on major telescopes is at a premium, and an observatory director will typically receive many more requests for telescope time than can be accommodated during the year. Astronomers must therefore write a convincing proposal explaining how they would like to use the telescope and why their observations will be important to the progress of astronomy. A committee

of astronomers is then asked to judge and rank the proposals, and time is assigned only to those with the greatest merit. Even if your proposal is among the high-rated ones, you may have to wait many months for your turn. If the skies are cloudy on the nights you have been assigned, it may be more than a year before you get another chance.

Some older astronomers still remember long, cold nights spent alone in an observatory dome, with only music from a tape recorder or an all-night radio station for company. The sight of the stars shining brilliantly hour after hour through the open slit in the observatory dome was unforgettable. So, too, was the relief as the first pale light of dawn announced the end of a 12-hour observation session. Astronomy is much easier today, with teams of observers working together, often at their computers, in a warm room. Those who are more nostalgic, however, might argue that some of the romance has gone from the field, too.

This page titled 10.2: Telescopes Today is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.3: Telescopes Today by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.3: Visible-Light Detectors and Instruments

Learning Objectives

By the end of this section, you will be able to:

- Describe the difference between photographic plates and charge-coupled devices
- Describe the unique difficulties associated with infrared observations and their solutions
- Describe how a spectrometer works

After a telescope collects radiation from an astronomical source, the radiation must be *detected* and measured. The first detector used for astronomical observations was the human eye, but it suffers from being connected to an imperfect recording and retrieving device—the human brain. Photography and modern electronic detectors have eliminated the quirks of human memory by making a permanent record of the information from the cosmos.

The eye also suffers from having a very short *integration time*; it takes only a fraction of a second to add light energy together before sending the image to the brain. One important advantage of modern detectors is that the light from astronomical objects can be collected by the detector over longer periods of time; this technique is called "taking a long exposure." Exposures of several hours are required to detect very faint objects in the cosmos.

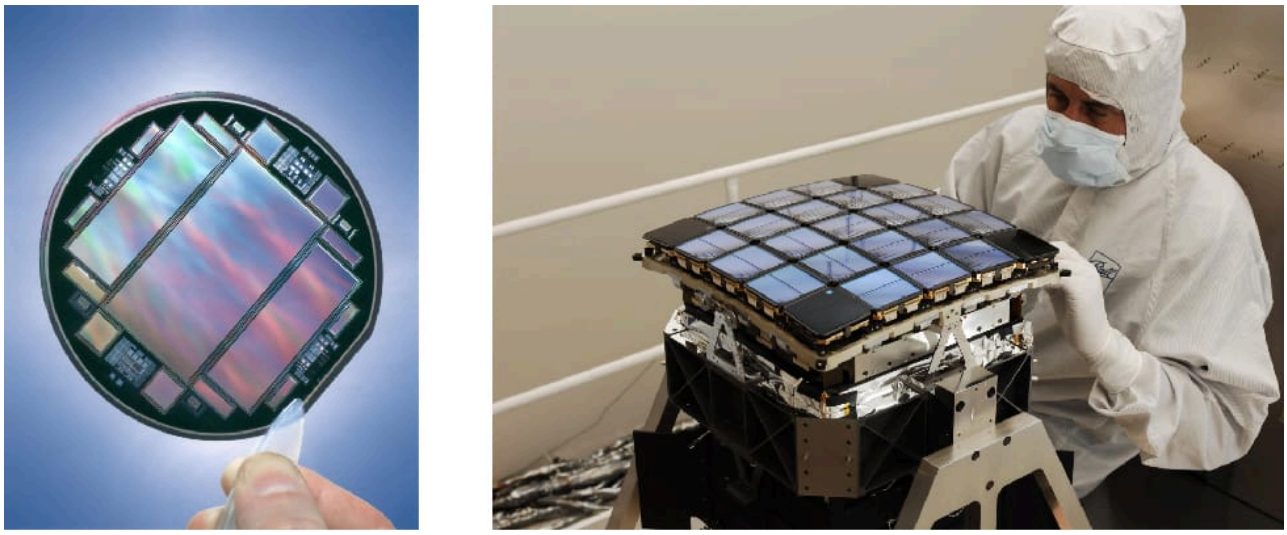
Before the light reaches the detector, astronomers today normally use some type of instrument to sort the light according to wavelength. The instrument may be as simple as colored filters, which transmit light within a specified range of wavelengths. A red transparent plastic is an everyday example of a filter that transmits only the red light and blocks the other colors. After the light passes through a filter, it forms an image that astronomers can then use to measure the apparent brightness and color of objects. We will show you many examples of such images in the later chapters of this book, and we will describe what we can learn from them.

Alternatively, the instrument between telescope and detector may be one of several devices that spread the light out into its full rainbow of colors so that astronomers can measure individual lines in the spectrum. Such an instrument (which you learned about in the chapter on Radiation and Spectra) is called a *spectrometer* because it allows astronomers to measure (to meter) the spectrum of a source of radiation. Whether a filter or a spectrometer, both types of wavelength-sorting instruments still have to use detectors to record and measure the properties of light.

Photographic and Electronic Detectors

Throughout most of the twentieth century, photographic film or *glass plates* served as the prime astronomical detectors, whether for photographing spectra or direct images of celestial objects. In a photographic plate, a light-sensitive chemical coating is applied to a piece of glass that, when developed, provides a lasting record of the image. At observatories around the world, vast collections of photographs preserve what the sky has looked like during the past 100 years. Photography represents a huge improvement over the human eye, but it still has limitations. Photographic films are inefficient: only about 1% of the light that actually falls on the film contributes to the chemical change that makes the image; the rest is wasted.

Astronomers today have much more efficient electronic detectors to record astronomical images. Most often, these are charge-coupled devices (CCDs), which are similar to the detectors used in video camcorders or in digital cameras (like the one more and more students have on their cell phones) (see Figure 6.14). In a CCD, photons of radiation hitting any part of the detector generate a stream of charged particles (electrons) that are stored and counted at the end of the exposure. Each place where the radiation is counted is called a pixel (picture element), and modern detectors can count the photons in millions of pixels (megapixels, or MPs).



(a) (b)

Figure 6.14 : Charge-Coupled Devices (CCDs). (a) This CCD is a mere 300-micrometers thick (thinner than a human hair) yet holds more than 21 million pixels. (b) This matrix of 42 CCDs served the Kepler telescope. (credit a: modification of work by US Department of Energy; credit b: modification of work by NASA and Ball Aerospace)

Because CCDs typically record as much as 60–70% of all the photons that strike them, and the best silicon and infrared CCDs exceed 90% sensitivity, we can detect much fainter objects. Among these are many small moons around the outer planets, icy dwarf planets beyond Pluto, and dwarf galaxies of stars. CCDs also provide more accurate measurements of the brightness of astronomical objects than photography, and their output is digital—in the form of numbers that can go directly into a computer for analysis.

Infrared Observations

Observing the universe in the infrared band of the spectrum presents some additional challenges. The infrared region extends from wavelengths near 1 micrometer (μm), which is about the long wavelength sensitivity limit of both CCDs and photography, to 100 micrometers or longer. Recall from the discussion on radiation and spectra that infrared is “heat radiation” (given off at temperatures that we humans are comfortable with). The main challenge to astronomers using infrared is to distinguish between the tiny amount of heat radiation that reaches Earth from stars and galaxies, and the much greater heat radiated by the telescope itself and our planet’s atmosphere.

Typical temperatures on Earth’s surface are near 300 K, and the atmosphere through which observations are made is only a little cooler. According to Wien’s law (from the chapter on Radiation and Spectra), the telescope, the observatory, and even the sky are radiating infrared energy with a peak wavelength of about 10 micrometers. To infrared eyes, everything on Earth is brightly aglow—including the telescope and camera (Figure 6.15). The challenge is to detect faint cosmic sources against this sea of infrared light. Another way to look at this is that an astronomer using infrared must always contend with the situation that a visible-light observer would face if working in broad daylight with a telescope and optics lined with bright fluorescent lights.



Figure 6.15 : Infrared Eyes. Infrared waves can penetrate places in the universe from which light is blocked, as shown in this infrared image where the plastic bag blocks visible light but not infrared. (credit: NASA/JPL-Caltech/R. Hurt (SSC))

To solve this problem, astronomers must protect the infrared detector from nearby radiation, just as you would shield photographic film from bright daylight. Since anything warm radiates infrared energy, the detector must be isolated in very cold surroundings; often, it is held near absolute zero (1 to 3 K) by immersing it in liquid helium. The second step is to reduce the radiation emitted by the telescope structure and optics, and to block this heat from reaching the infrared detector.

Link to Learning

Check out The Infrared Zoo to get a sense of what familiar objects look like with infrared radiation. Slide the slider to change the wavelength of radiation for the picture, and click the arrow to see other animals.

Spectroscopy

Spectroscopy is one of the astronomer's most powerful tools, providing information about the composition, temperature, motion, and other characteristics of celestial objects. More than half of the time spent on most large telescopes is used for spectroscopy.

The many different wavelengths present in light can be separated by passing them through a spectrometer to form a spectrum. The design of a simple spectrometer is illustrated in Figure 6.16. Light from the source (actually, the image of a source produced by the telescope) enters the instrument through a small hole or narrow slit, and is collimated (made into a beam of parallel rays) by a lens. The light then passes through a prism, producing a spectrum: different wavelengths leave the prism in different directions because each wavelength is bent by a different amount when it enters and leaves the prism. A second lens placed behind the prism focuses the many different images of the slit or entrance hole onto a CCD or other detecting device. This collection of images (spread out by color) is the spectrum that astronomers can then analyze at a later point. As spectroscopy spreads the light out into more and more collecting bins, fewer photons go into each bin, so either a larger telescope is needed or the integration time must be greatly increased—usually both.

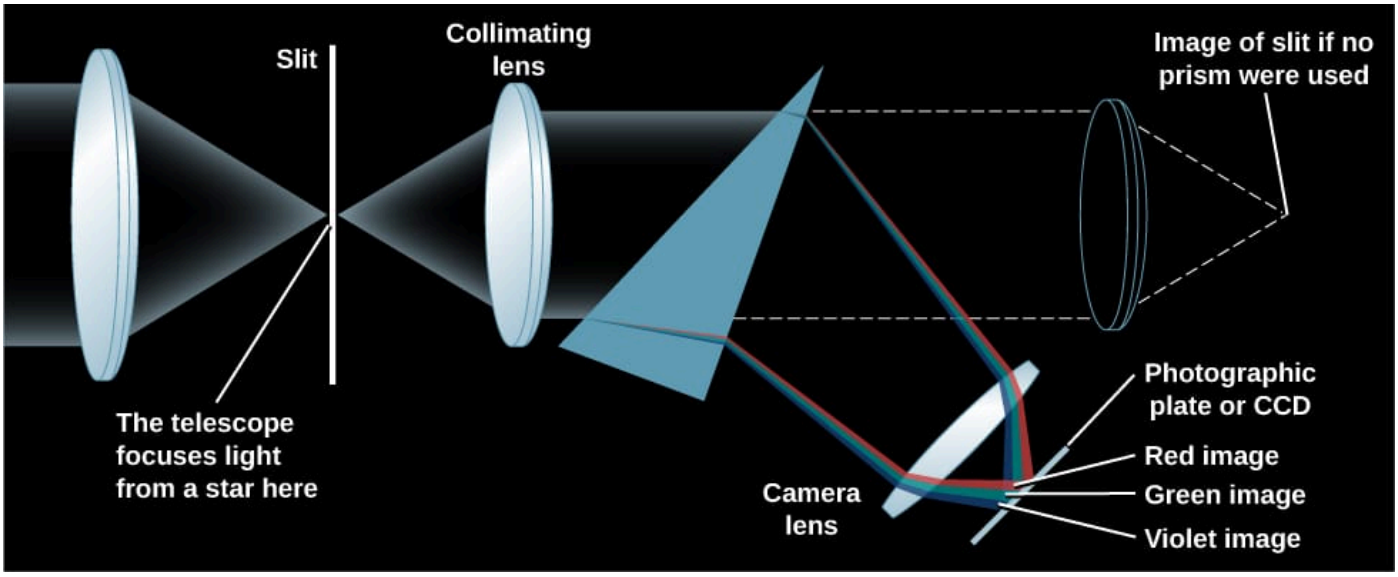


Figure 6.16 : Prism Spectrometer. The light from the telescope is focused on a slit. A prism (or grating) disperses the light into a spectrum, which is then photographed or recorded electronically.

In practice, astronomers today are more likely to use a different device, called a *grating*, to disperse the spectrum. A grating is a piece of material with thousands of grooves on its surface. While it functions completely differently, a grating, like a prism, also spreads light out into a spectrum.

This page titled 10.3: Visible-Light Detectors and Instruments is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.4: Visible-Light Detectors and Instruments by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.4: Radio Telescopes

Learning Objectives

By the end of this section, you will be able to:

- Describe how radio waves from space are detected
- Identify the world's largest radio telescopes
- Define the technique of interferometry and discuss the benefits of interferometers over single-dish telescopes

In addition to visible and infrared radiation, radio waves from astronomical objects can also be detected from the surface of Earth. In the early 1930s, Karl G. Jansky, an engineer at Bell Telephone Laboratories, was experimenting with antennas for long-range radio communication when he encountered some mysterious static—radio radiation coming from an unknown source (Figure 6.17). He discovered that this radiation came in strongest about four minutes earlier on each successive day and correctly concluded that since Earth's sidereal rotation period (how long it takes us to rotate relative to the stars) is four minutes shorter than a solar day, the radiation must be originating from some region fixed on the celestial sphere. Subsequent investigation showed that the source of this radiation was part of the Milky Way Galaxy; Jansky had discovered the first source of cosmic radio waves.

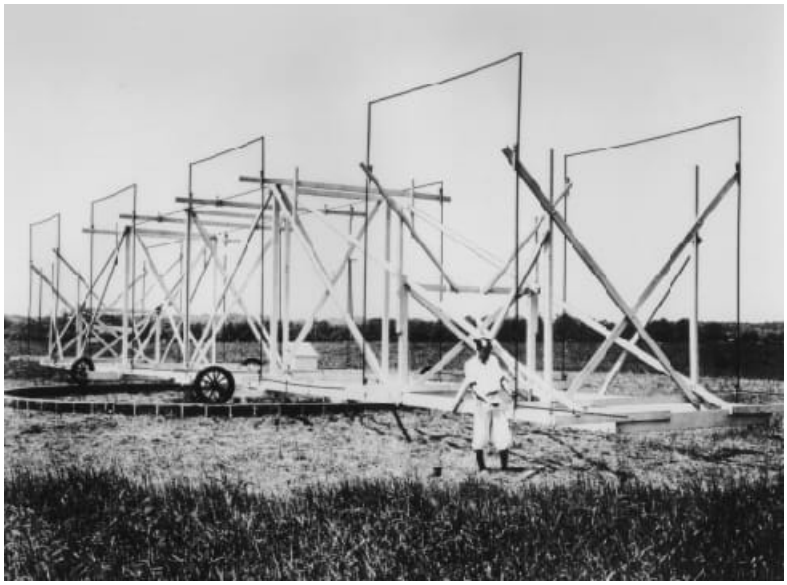


Figure 6.17 : First Radio Telescope. This rotating radio antenna was used by Jansky in his serendipitous discovery of radio radiation from the Milky Way.

In 1936, Grote Reber, who was an amateur astronomer interested in radio communications, used galvanized iron and wood to build the first antenna specifically designed to receive cosmic radio waves. Over the years, Reber built several such antennas and used them to carry out pioneering surveys of the sky for celestial radio sources; he remained active in radio astronomy for more than 30 years. During the first decade, he worked practically alone because professional astronomers had not yet recognized the vast potential of radio astronomy.

Detection of Radio Energy from Space

It is important to understand that radio waves cannot be "heard": they are not the sound waves you hear coming out of the radio receiver in your home or car. Like light, radio waves are a form of electromagnetic radiation, but unlike light, we cannot detect them with our senses—we must rely on electronic equipment to pick them up. In commercial radio broadcasting, we encode sound information (music or a newscaster's voice) into radio waves. These must be decoded at the other end and then turned back into sound by speakers or headphones.

The radio waves we receive from space do not, of course, have music or other program information encoded in them. If cosmic radio signals were translated into sound, they would sound like the static you hear when scanning between stations. Nevertheless, there is information in the radio waves we receive—information that can tell us about the chemistry and physical conditions of the sources of the waves.

Just as vibrating charged particles can produce electromagnetic waves (see the Radiation and Spectra chapter), electromagnetic waves can make charged particles move back and forth. Radio waves can produce a current in conductors of electricity such as metals. An antenna is such a conductor: it intercepts radio waves, which create a feeble current in it. The current is then amplified in a radio receiver until it is strong enough to measure or record. Like your television or radio, receivers can be tuned to select a single frequency (channel). In astronomy, however, it is more common to use sophisticated data-processing techniques that allow thousands of separate frequency bands to be detected simultaneously. Thus, the astronomical radio receiver operates much like a spectrometer on a visible-light or infrared telescope, providing information about how much radiation we receive at each wavelength or frequency. After computer processing, the radio signals are recorded on magnetic disks for further analysis.

Radio waves are reflected by conducting surfaces, just as light is reflected from a shiny metallic surface, and according to the same laws of optics. A radio-reflecting telescope consists of a concave metal reflector (called a *dish*), analogous to a telescope mirror. The radio waves collected by the dish are reflected to a focus, where they can then be directed to a receiver and analyzed. Because humans are such visual creatures, radio astronomers often construct a pictorial representation of the radio sources they observe. Figure 6.18 shows such a radio image of a distant galaxy, where radio telescopes reveal vast jets and complicated regions of radio emissions that are completely invisible in photographs taken with light.

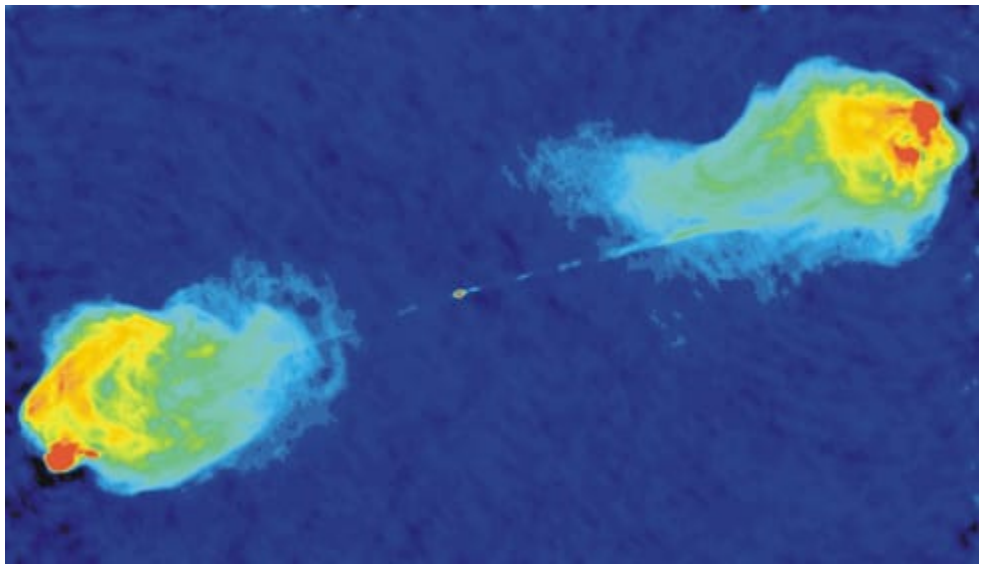


Figure 6.18 : Radio Image. This image has been constructed of radio observations at the Very Large Array of a galaxy called Cygnus A. Colors have been added to help the eye sort out regions of different radio intensities. Red regions are the most intense, blue the least. The visible galaxy would be a small dot in the center of the image. The radio image reveals jets of expelled material (more than 160,000 light-years long) on either side of the galaxy. (credit: NRAO/AUI)

Radio astronomy is a young field compared with visible-light astronomy, but it has experienced tremendous growth in recent decades. The world's largest radio reflectors that can be pointed to any direction in the sky have apertures of 100 meters. One of these has been built at the US National Radio Astronomy Observatory in West Virginia (Figure 6.19). Table 6.2 lists some of the major radio telescopes of the world.



Figure 6.19 : Robert C. Byrd Green Bank Telescope. This fully steerable radio telescope in West Virginia went into operation in August 2000. Its dish is about 100 meters across. (credit: modification of work by “b3nscott”/Flickr)

Table 6.2: Major Radio Observatories of the World

Observatory	Location	Description	Website
Individual Radio Dishes			
Five-hundred-meter Aperture Spherical radio Telescope (FAST)	Guizhou, China	500-m fixed dish	fast.bao.ac.cn/en/
Arecibo Observatory	Arecibo, Puerto Rico	305-m fixed dish	www.naic.edu
Green Bank Telescope (GBT)	Green Bank, WV	110 × 100-m steerable dish	www.science.nrao.edu/facilities/gbt
Effelsberg 100-m Telescope	Bonn, Germany	100-m steerable dish	www.mpifr-bonn.mpg.de/en/effelsberg
Lovell Telescope	Manchester, England	76-m steerable dish	www.jb.man.ac.uk/aboutus/lovell
Canberra Deep Space Communication Complex (CDSCC)	Tidbinbilla, Australia	70-m steerable dish	www.cdscc.nasa.gov
Goldstone Deep Space Communications Complex (GDSCC)	Barstow, CA	70-m steerable dish	www.gdscc.nasa.gov
Parkes Observatory	Parkes, Australia	64-m steerable dish	www.parkes.atnf.csiro.au
Arrays of Radio Dishes			
Square Kilometre Array (SKA)	South Africa and Western Australia	Thousands of dishes, km^2 collecting area, partial array in 2020	www.skatelescope.org
Atacama Large Millimeter/ submillimeter Array (ALMA)	Atacama desert, Northern Chile	66 7-m and 12-m dishes	www.almaobservatory.org
Jansky Very Large Array (VLA)	Socorro, New Mexico	27-element array of 25-m dishes (36-km baseline)	www.science.nrao.edu/facilities/vla
Westerbork Synthesis Radio Telescope (WSRT)	Westerbork, the Netherlands	12-element array of 25-m dishes (1.6-km baseline)	www.astron.nl/radio-observatory/public/public-0

Observatory	Location	Description	Website
Very Long Baseline Array (VLBA)	Ten US sites, HI to the Virgin Islands	10-element array of 25-m dishes (9000 km baseline)	www.science.nrao.edu/facilities/vlba
Australia Telescope Compact Array (ATCA)	Several sites in Australia	8-element array (seven 22-m dishes plus Parkes 64 m)	www.narrabri.atnf.csiro.au
Multi-Element Radio Linked Interferometer Network (MERLIN)	Cambridge, England, and other British sites	Network of seven dishes (the largest is 32 m)	www.e-merlin.ac.uk
Millimeter-wave Telescopes			
IRAM	Granada, Spain	30-m steerable mm-wave dish	www.iram-institute.org
James Clerk Maxwell Telescope (JCMT)	Maunakea, HI	15-m steerable mm-wave dish	www.eaobservatory.org/jcmt
Nobeyama Radio Observatory (NRO)	Minamimaki, Japan	6-element array of 10-m wave dishes	www.nro.nao.ac.jp/en
Hat Creek Radio Observatory (HCRO)	Cassel, CA	6-element array of 5-m wave dishes	www.sri.com/research-development/specialized-facilities/hat-creek-radio-observatory

Radio Interferometry

As we discussed earlier, a telescope's ability to show us fine detail (its resolution) depends upon its aperture, but it also depends upon the wavelength of the radiation that the telescope is gathering. The longer the waves, the harder it is to resolve fine detail in the images or maps we make. Because radio waves have such long wavelengths, they present tremendous challenges for astronomers who need good resolution. In fact, even the largest radio dishes on Earth, operating alone, cannot make out as much detail as the typical small visible-light telescope used in a college astronomy lab. To overcome this difficulty, radio astronomers have learned to sharpen their images by linking two or more radio telescopes together electronically. Two or more telescopes linked together in this way are called an interferometer.

"Interferometer" may seem like a strange term because the telescopes in an interferometer work cooperatively; they don't "interfere" with each other. Interference, however, is a technical term for the way that multiple waves interact with each other when they arrive in our instruments, and this interaction allows us to coax more detail out of our observations. The resolution of an interferometer depends upon the separation of the telescopes, not upon their individual apertures. Two telescopes separated by 1 kilometer provide the same resolution as would a single dish 1 kilometer across (although they are not, of course, able to collect as much radiation as a radio-wave bucket that is 1 kilometer across).

To get even better resolution, astronomers combine a large number of radio dishes into an interferometer array. In effect, such an array works like a large number of two-dish interferometers, all observing the same part of the sky together. Computer processing of the results permits the reconstruction of a high-resolution radio image. The most extensive such instrument in the United States is the National Radio Astronomy Observatory's Jansky Very Large Array (VLA) near Socorro, New Mexico. It consists of 27 movable radio telescopes (on railroad tracks), each having an aperture of 25 meters, spread over a total span of about 36 kilometers. By electronically combining the signals from all of its individual telescopes, this array permits the radio astronomer to make pictures of the sky at radio wavelengths comparable to those obtained with a visible-light telescope, with a resolution of about 1 arcsecond.

The Atacama Large Millimeter/submillimeter array (ALMA) in the Atacama Desert of Northern Chile (Figure 6.20), at an altitude of 16,400 feet, consists of 12 7-meter and 54 12-meter telescopes, and can achieve baselines up to 16 kilometers. Since it became operational in 2013, it has made observations at resolutions down to 6 milliarcseconds (0.006 arcseconds), a remarkable achievement for radio astronomy.



Figure 6.20 : Atacama Large Millimeter/Submillimeter Array (ALMA). Located in the Atacama Desert of Northern Chile, ALMA currently provides the highest resolution for radio observations. (credit: ESO/S. Guisard)

Link to Learning

Watch this documentary that explains the work that went into designing and building ALMA, discusses some of its first images, and explores its future.

Initially, the size of interferometer arrays was limited by the requirement that all of the dishes be physically wired together. The maximum dimensions of the array were thus only a few tens of kilometers. However, larger interferometer separations can be achieved if the telescopes do not require a physical connection. Astronomers, with the use of current technology and computing power, have learned to time the arrival of electromagnetic waves coming from space very precisely at each telescope and combine the data later. If the telescopes are as far apart as California and Australia, or as West Virginia and Crimea in Ukraine, the resulting resolution far surpasses that of visible-light telescopes.

The United States operates the Very Long Baseline Array (VLBA), made up of 10 individual telescopes stretching from the Virgin Islands to Hawaii (Figure 6.21). The VLBA, completed in 1993, can form astronomical images with a resolution of 0.0001 arcseconds, permitting features as small as 10 astronomical units (AU) to be distinguished at the center of our Galaxy.



Figure 6.21 : Very Long Baseline Array. This map shows the distribution of 10 antennas that constitute an array of radio telescopes stretching across the United States and its territories.

Recent advances in technology have also made it possible to do interferometry at visible-light and infrared wavelengths. At the beginning of the twenty-first century, three observatories with multiple telescopes each began using their dishes as interferometers, combining their light to obtain a much greater resolution. In addition, a dedicated interferometric array was built on Mt. Wilson in California. Just as in radio arrays, these observations allow astronomers to make out more detail than a single telescope could provide.

Table 6.3: Visible-Light Interferometers

Longest Baseline (m)	Telescope Name	Location	Mirrors	Status
400	CHARA Array (Center for High Angular Resolution Astronomy)	Mount Wilson, CA	Six 1-m telescopes	Operational since 2004

Longest Baseline (m)	Telescope Name	Location	Mirrors	Status
200	Very Large Telescope	Cerro Paranal, Chile	Four 8.2-m telescopes	Completed 2000
85	Keck I and II telescopes	Maunakea, HI	Two 10-m telescopes	Operated from 2001 to 2012
22.8	Large Binocular Telescope	Mount Graham, AZ	Two 8.4-m telescopes	First light 2004

Radar Astronomy

Radar is the technique of transmitting radio waves to an object in our solar system and then detecting the radio radiation that the object reflects back. The time required for the round trip can be measured electronically with great precision. Because we know the speed at which radio waves travel (the speed of light), we can determine the distance to the object or a particular feature on its surface (such as a mountain).

Radar observations have been used to determine the distances to planets and how fast things are moving in the solar system (using the Doppler effect, discussed in the Radiation and Spectra chapter). Radar waves have played important roles in navigating spacecraft throughout the solar system. In addition, as will be discussed in later chapters, radar observations have determined the rotation periods of Venus and Mercury, probed tiny Earth-approaching asteroids, and allowed us to investigate the mountains and valleys on the surfaces of Mercury, Venus, Mars, and the large moons of Jupiter.

Any radio dish can be used as a radar telescope if it is equipped with a powerful transmitter as well as a receiver. For many years, the most spectacular facility in the world for radar astronomy was the 1000-foot (305-meter) telescope at Arecibo in Puerto Rico (Figure 6.22). The Arecibo telescope was too large to be pointed directly at different parts of the sky. Instead, it was constructed in a huge natural “bowl” (more than a mere dish) formed by several hills, and it was lined with reflecting metal panels. A limited ability to track astronomical sources was achieved by moving the receiver system, which was suspended on cables 100 meters above the surface of the bowl. Unfortunately, the telescope was seriously damaged in the powerful storms of 2020 and had to be decommissioned. An even larger (500-meter) radar telescope has recently gone into operation in China and is called the Five-hundred-meter Aperture Spherical Telescope (FAST).



Figure 6.22 : Largest Radio and Radar Dish. The Arecibo Observatory in Puerto Rico was the largest and most powerful astronomical radar faculty in the world and was often featured in films. In November 2020 it collapsed, damaging the 300-meter diameter “dish” and destroying the radar transmitter and receiver. (credit: National Astronomy and Ionosphere Center, Cornell U., NSF)

This page titled 10.4: Radio Telescopes is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.5: Radio Telescopes by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.5: Observations outside Earth's Atmosphere

Learning Objectives

By the end of this section, you will be able to:

- List the advantages of making astronomical observations from space
- Explain the importance of the Hubble Space Telescope
- Describe some of the major space-based observatories astronomers use

Earth's atmosphere blocks most radiation at wavelengths shorter than visible light, so we can only make direct ultraviolet, X-ray, and gamma ray observations from space (though indirect gamma ray observations can be made from Earth). Getting above the distorting effects of the atmosphere is also an advantage at visible and infrared wavelengths. The stars don't "twinkle" in space, so the amount of detail you can observe is limited only by the size of your instrument. On the other hand, it is expensive to place telescopes into space, and repairs can present a major challenge. This is why astronomers continue to build telescopes for use on the ground as well as for launching into space.

Airborne and Space Infrared Telescopes

Water vapor, the main source of atmospheric interference for making infrared observations, is concentrated in the lower part of Earth's atmosphere. For this reason, a gain of even a few hundred meters in elevation can make an important difference in the quality of an infrared observatory site. Given the limitations of high mountains, most of which attract clouds and violent storms, and the fact that the ability of humans to perform complex tasks degrades at high altitudes, it was natural for astronomers to investigate the possibility of observing infrared waves from airplanes and ultimately from space.

Infrared observations from airplanes have been made since the 1960s, starting with a 15-centimeter telescope on board a Learjet. From 1974 through 1995, NASA operated a 0.9-meter airborne telescope flying regularly out of the Ames Research Center south of San Francisco. Observing from an altitude of 12 kilometers, the telescope was above 99% of the atmospheric water vapor. More recently, NASA (in partnership with the German Aerospace Center) has constructed a much larger 2.5-meter telescope, called the Stratospheric Observatory for Infrared Astronomy (SOFIA), which flies in a modified Boeing 747SP (Figure 6.23).



Figure 6.23 : Stratospheric Observatory for Infrared Astronomy (SOFIA). SOFIA allows observations to be made above most of Earth's atmospheric water vapor. (credit: NASA)

Link to Learning

To find out more about SOFIA, watch this video provided by NASA's Armstrong Flight Research Center.

Getting even higher and making observations from space itself have important advantages for infrared astronomy. First is the elimination of all interference from the atmosphere. Equally important is the opportunity to cool the entire optical system of the instrument in order to nearly eliminate infrared radiation from the telescope itself. If we tried to cool a telescope within the atmosphere, it would quickly become coated with condensing water vapor and other gases, making it useless. Only in the vacuum of space can optical elements be cooled to hundreds of degrees below freezing and still remain operational.

The first orbiting infrared observatory, launched in 1983, was the Infrared Astronomical Satellite (IRAS), built as a joint project by the United States, the Netherlands, and Britain. IRAS was equipped with a 0.6-meter telescope cooled to a temperature of less than 10 K. For the first time, the infrared sky could be seen as if it were night, rather than through a bright foreground of atmospheric and telescope emissions. IRAS carried out a rapid but comprehensive survey of the entire infrared sky over a 10-month period, cataloging about 350,000 sources of infrared radiation. Since then, several other infrared telescopes have operated in space with much better sensitivity and resolution due to improvements in infrared detectors. The most powerful of these infrared telescopes is the 0.85-meter Spitzer Space Telescope, which launched in 2003. A few of its observations are shown in Figure 6.24. With infrared observations, astronomers can detect cooler parts of cosmic objects, such as the dust clouds around star nurseries and the remnants of dying stars, that visible-light images don't reveal.

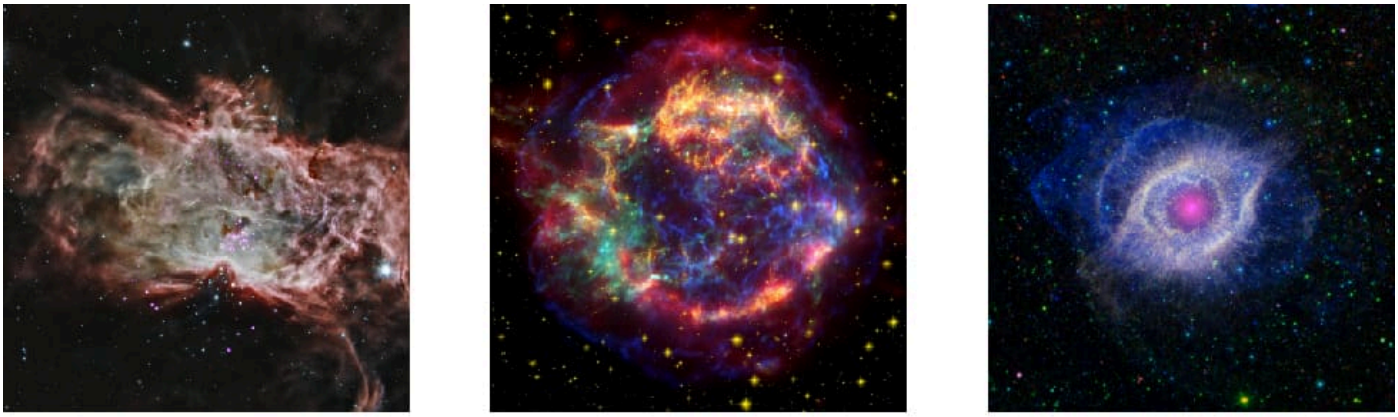


Figure 6.24 : Observations from the Spitzer Space Telescope (SST). These infrared images—a region of star formation, the remnant of an exploded star, and a region where an old star is losing its outer shell—show just a few of the observations made and transmitted back to Earth from the SST. Since our eyes are not sensitive to infrared rays, we don't perceive colors from them. The colors in these images have been selected by astronomers to highlight details like the composition or temperature in these regions. (credit "Flame nebula": modification of work by NASA (X-ray: NASA/CXC/PSU/K.Getman, E.Feigelson, M.Kuhn & the MYStIX team; Infrared:NASA/JPL-Caltech); credit "Cassiopeia A": modification of work by NASA/JPL-Caltech; credit "Helix nebula": modification of work by NASA/JPL-Caltech)

Hubble Space Telescope

In April 1990, a great leap forward in astronomy was made with the launch of the Hubble Space Telescope (HST). With an aperture of 2.4 meters, this is the largest telescope put into space so far. (Its aperture was limited by the size of the payload bay in the Space Shuttle that served as its launch vehicle.) It was named for Edwin Hubble, the astronomer who discovered the expansion of the universe in the 1920s (whose work we will discuss in the chapters on Galaxies).

HST is operated jointly by NASA's Goddard Space Flight Center and the Space Telescope Science Institute in Baltimore. It was the first orbiting observatory designed to be serviced by Shuttle astronauts and, over the years since it was launched, they made several visits to improve or replace its initial instruments and to repair some of the systems that operate the spacecraft (Figure 6.1)—though this repair program has now been discontinued, and no more visits or improvements will be made.

With the Hubble, astronomers have obtained some of the most detailed images of astronomical objects from the solar system outward to the most distant galaxies. Among its many great achievements is the Hubble Ultra-Deep Field, an image of a small

region of the sky observed for almost 100 hours. It contains views of about 10,000 galaxies, some of which formed when the universe was just a few percent of its current age (Figure 6.25).

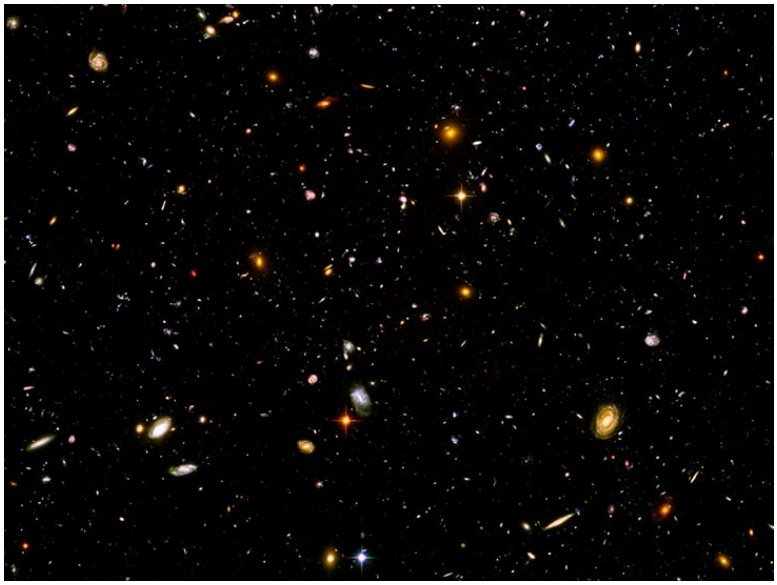


Figure 6.25 : Hubble Ultra-Deep Field (HUDF). The Hubble Space Telescope has provided an image of a specific region of space built from data collected between September 24, 2003, and January 16, 2004. These data allow us to search for galaxies that existed approximately 13 billion years ago. (credit: modification of work by NASA)

The HST's mirror was ground and polished to a remarkable degree of accuracy. If we were to scale up its 2.4-meter mirror to the size of the entire continental United States, there would be no hill or valley larger than about 6 centimeters in its smooth surface. Unfortunately, after it was launched, scientists discovered that the primary mirror had a slight error in its *shape,* equal to roughly 1/50 the width of a human hair. Small as that sounds, it was enough to ensure that much of the light entering the telescope did not come to a clear focus and that all the images were blurry. (In a misplaced effort to save money, a complete test of the optical system had not been carried out before launch, so the error was not discovered until HST was in orbit.)

The solution was to do something very similar to what we do for astronomy students with blurry vision: put corrective optics in front of their eyes. In December 1993, in one of the most exciting and difficult space missions ever flown, astronauts captured the orbiting telescope and brought it back into the shuttle payload bay. There they installed a package containing compensating optics as well as a new, improved camera before releasing HST back into orbit. The telescope now works as it was intended to, and further missions to it were able to install even more advanced instruments to take advantage of its capabilities.

High-Energy Observatories

Ultraviolet, X-ray, and direct gamma-ray (high-energy electromagnetic wave) observations can be made only from space. Such observations first became possible in 1946, with V2 rockets captured from Germany after World War II. The US Naval Research Laboratory put instruments on these rockets for a series of pioneering flights, used initially to detect ultraviolet radiation from the Sun. Since then, many other rockets have been launched to make X-ray and ultraviolet observations of the Sun, and later of other celestial objects.

Beginning in the 1960s, a steady stream of high-energy observatories has been launched into orbit to reveal and explore the universe at short wavelengths. Among recent X-ray telescopes is the Chandra X-ray Observatory, which was launched in 1999 (Figure 6.26). It is producing X-ray images with unprecedented resolution and sensitivity. Designing instruments that can collect and focus energetic radiation like X-rays and gamma rays is an enormous technological challenge. The 2002 Nobel Prize in physics was awarded to Riccardo Giacconi, a pioneer in the field of building and launching sophisticated X-ray instruments. In 2008, NASA launched the Fermi Gamma-ray Space Telescope, designed to measure cosmic gamma rays at energies greater than any previous telescope, and thus able to collect radiation from some of the most energetic events in the universe.



Figure 6.26 : Chandra X-Ray Satellite. Chandra, the world's most powerful X-ray telescope, was developed by NASA and launched in July 1999. (credit: modification of work by NASA)

One major challenge is to design "mirrors" to reflect such penetrating radiation as X-rays and gamma rays, which normally pass straight through matter. However, although the technical details of design are more complicated, the three basic components of an observing system, as we explained earlier in this chapter, are the same at all wavelengths: a telescope to gather up the radiation, filters or instruments to sort the radiation according to wavelength, and some method of detecting and making a permanent record of the observations. Table 6.4 lists some of the most important active space observatories that humanity has launched.

Gamma-ray detections can also be made from Earth's surface by using the atmosphere as the primary detector. When a gamma ray hits our atmosphere, it accelerates charged particles (mostly electrons) in the atmosphere. Those energetic particles hit other particles in the atmosphere and give off their own radiation. The effect is a cascade of light and energy that can be detected on the ground. The VERITAS array in Arizona and the H.E.S.S. array in Namibia are two such ground-based gamma-ray observatories.

Table 6.4: Recent Observatories in Space

Observatory	Date Operation Began	Bands of the Spectrum	Notes	Website
Hubble Space Telescope (HST)	1990	visible, UV, IR	2.4-m mirror; images and spectra	www.hubblesite.org
Chandra X-Ray Observatory	1999	X-rays	X-ray images and spectra	www.chandra.si.edu
XMM-Newton	1999	X-rays	X-ray spectroscopy	http://www.cosmos.esa.int/web/xmm-newton
International Gamma-Ray Astrophysics Laboratory (INTEGRAL)	2002	X- and gamma-rays	higher resolution gamma-ray images	http://sci.esa.int/integral/
Spitzer Space Telescope	2003	IR	0.85-m telescope	www.spitzer.caltech.edu
Fermi Gamma-ray Space Telescope	2008	gamma-rays	first high-energy gamma-ray observations	fermi.gsfc.nasa.gov
Kepler	2009	visible-light	planet finder	kepler.nasa.gov
Wide-field Infrared Survey Explorer (WISE)	2009	IR	whole-sky map, asteroid searches	www.nasa.gov/mission_pages/WISE/main
Gaia	2013	visible-light	Precise map of the Milky Way	http://sci.esa.int/gaia/

Observatory	Date Operation Began	Bands of the Spectrum	Notes	Website
Transiting Exoplanet Survey Satellite (TESS)	2018	visible-light	Planet finder	http://tess.mit.edu

This page titled 10.5: Observations outside Earth's Atmosphere is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.6: Observations outside Earth's Atmosphere by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.6: The Future of Large Telescopes

Learning Objectives

By the end of this section, you will be able to:

- Describe the next generation of ground- and space-based observatories
- Explain some of the challenges involved in building these observatories

If you've ever gone on a hike, you have probably been eager to see what lies just around the next bend in the path. Researchers are no different, and astronomers and engineers are working on the technologies that will allow us to explore even more distant parts of the universe and to see them more clearly.

The premier space facility planned for the next decade is the James Webb Space Telescope (Figure 6.27), which was launched on December 25, 2021. James Webb was one of the early administrators in NASA. As we write this (in December 2021), the telescope is on its way to a stable orbit point, some 1.5 million kilometers from Earth (where no astronauts can currently travel if the facility needs repair). During its 29-day journey, a number of activities must be successfully executed in order to prepare the JWST to make observations. A solar array will be deployed to provide energy, and a sunshield will be stretched out to protect the telescope from the intense heat of the Sun. The secondary and primary mirrors will then be unfolded. The primary mirror is 6 meters in diameter, made up, like the Keck telescopes, of 36 small hexagons. JWST has been designed to operate at infrared wavelengths and will have the sensitivity needed to detect the very first generation of stars, formed when the universe was only a few hundred million years old.



Figure 6.27 : James Webb Space Telescope (JWST). This image shows some of the mirrors of the JWST as they underwent cryogenic testing. The mirrors were exposed to extreme temperatures in order to gather accurate measurements on changes in their shape as they heated and cooled. (credit: NASA/MSFC/David Higginbotham/Emmett Given)

Link to Learning

Watch this video to learn more about the James Webb Space Telescope and how it will build upon the work that Hubble has allowed us to begin in exploring the universe. Visit NASA's James Webb Space Telescope site to keep track of how the JWST is doing.

On the ground, astronomers have started building the Vera Rubin Observatory[1], which has an 8.4-meter telescope with a

significantly larger field of view than any existing telescopes. It will rapidly scan the sky to find *transients*, phenomena that change quickly, such as exploding stars and chunks of rock that orbit near Earth. It is expected to see first light in 2022.

The international gamma-ray community is planning the Cherenkov Telescope Array (CTA), two arrays of telescopes, one in each hemisphere, which will indirectly measure gamma rays from the ground. The CTA will measure gamma-ray energies a thousand times as great as the Fermi telescope can detect.

Several groups of astronomers around the globe interested in studying visible light and the infrared are exploring the feasibility of building ground-based telescopes with mirrors larger than 30 meters across. Stop and think what this means: 30 meters is one-third the length of a football field. It is technically impossible to build and transport a single astronomical mirror that is 30 meters or larger in diameter. The primary mirror of these giant telescopes will consist of smaller mirrors, all aligned so that they act as a very large mirror in combination.

The most ambitious of these projects is the European Extremely Large Telescope (ELT) (Figure 6.28). (Astronomers try to outdo each other not only with the size of these telescopes, but also their names!) The design of the European ELT calls for a 39.3-meter primary mirror, which will follow the Keck design and be made up of 798 hexagonal mirrors, each 1.4 meters in diameter and all held precisely in position so that they form a continuous surface. Construction on the site in the Atacama Desert in Northern Chile started in 2014, and operations are expected to begin in about 2025.

International consortia with major contributions from U.S. astronomers have developed plans for the construction of two large new telescopes. One is a Thirty-Meter Telescope (TMT) for which the preferred site is Maunakea in Hawaii. The design of this telescope is similar to that of the European ELT and will make use of 492 hexagonal elements. Each segment is about 1.44 meters (56.6 inches) across corners. The segments are closely spaced, with gaps between the segments only 2.5 mm (0.1 inch) wide.

The Giant Magellan Telescope (GMT) is the second ELT project with major participation by U.S. astronomers. The GMT is also a segmented mirror telescope that employs seven stiff monolith 8.4-meter mirrors as segments. Construction has started at the selected site, which is near the Las Campanas Observatory on the southern edge of the Atacama Desert.

These giant telescopes will combine light-gathering power with high-resolution imaging. These powerful new instruments will enable astronomers to tackle many important astronomical problems. As just one example, they provide us images and spectra of planets around other stars and thus, perhaps, give us the first real evidence (from the chemistry of these planets' atmospheres) that life exists elsewhere.

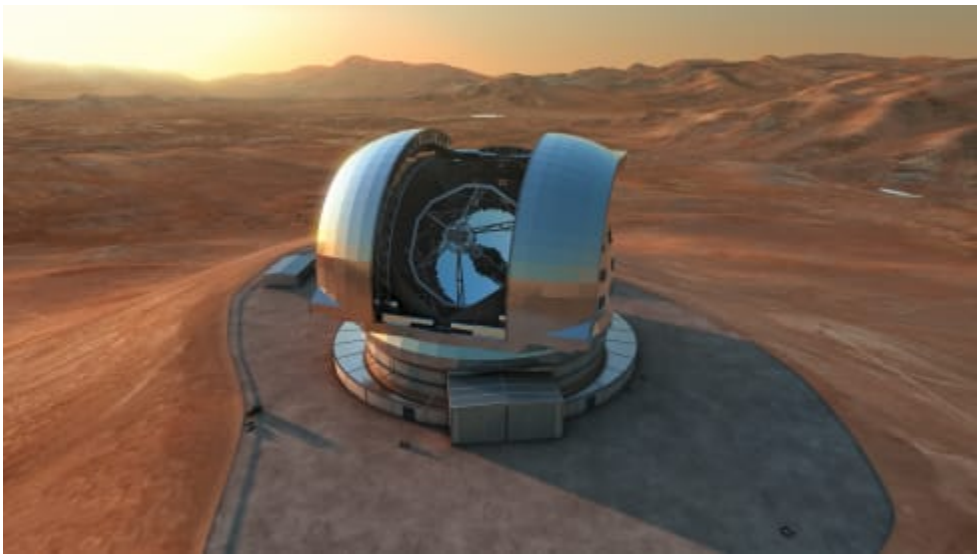


Figure 6.28 : Artist's Conception of the European Extremely Large Telescope. The primary mirror in this telescope is 39.3 meters across. The telescope is under construction in the Atacama Desert in Northern Chile. (credit: ESO/L. Calçada)

Link to Learning

Check out this fun diagram comparing the sizes of the largest planned and existing telescopes to a regulation basketball and tennis court.

Footnotes

- 1The observatory is named after the American astronomer whose work led us to the understanding that much of the universe is made of a mysterious substance that scientists call dark matter (which we explain in The Mass of the Galaxy).

This page titled 10.6: The Future of Large Telescopes is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 6.7: The Future of Large Telescopes by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

10.7: Key Terms

adaptive optics
systems used with telescopes that can compensate for distortions in an image introduced by the atmosphere, thus resulting in sharper images

aperture
diameter of the primary lens or mirror of a telescope

charge-coupled device (CCD)
array of high-sensitivity electronic detectors of electromagnetic radiation, used at the focus of a telescope (or camera lens) to record an image or spectrum

chromatic aberration
distortion that causes an image to appear fuzzy when each wavelength coming into a transparent material focuses at a different spot

detector
device sensitive to electromagnetic radiation that makes a record of astronomical observations

eyepiece
magnifying lens used to view the image produced by the objective lens or primary mirror of a telescope

focus
(of telescope) point where the rays of light converged by a mirror or lens meet

interference
process in which waves mix together such that their crests and troughs can alternately reinforce and cancel one another

interferometer
instrument that combines electromagnetic radiation from one or more telescopes to obtain a resolution equivalent to what would be obtained with a single telescope with a diameter equal to the baseline separating the individual separate telescopes

interferometer array
combination of multiple radio dishes to, in effect, work like a large number of two-dish interferometers

prime focus
point in a telescope where the objective lens or primary mirror focuses the light

radar
technique of transmitting radio waves to an object and then detecting the radiation that the object reflects back to the transmitter; used to measure the distance to, and motion of, a target object or to form images of it

reflecting telescope
telescope in which the principal light collector is a concave mirror

refracting telescope
telescope in which the principal light collector is a lens or system of lenses

resolution
detail in an image; specifically, the smallest angular (or linear) features that can be distinguished

seeing
unsteadiness of Earth's atmosphere, which blurs telescopic images; good seeing means the atmosphere is steady

telescope
instrument for collecting visible-light or other electromagnetic radiation

This page titled 10.7: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

11: SOLAR SYSTEM

11.1: Overview of Our Planetary System
11.2: Compostion and Structure of Planets
11.3: Dating Planetary Surfaces
11.4: Origin of the Solar System
11.5: Summary

This page titled 11: Solar System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

11.1: Overview of Our Planetary System

Learning Objectives

By the end of this section, you will be able to:

- Describe how the objects in our solar system are identified, explored, and characterized
- Describe the types of small bodies in our solar system, their locations, and how they formed
- Model the solar system with distances from everyday life to better comprehend distances in space

The solar system[1] consists of the Sun and many smaller objects: the planets, their moons and rings, and such "debris" as asteroids, comets, and dust. Decades of observation and spacecraft exploration have revealed that most of these objects formed together with the Sun about 4.5 billion years ago. They represent clumps of material that condensed from an enormous cloud of gas and dust. The central part of this cloud became the Sun, and a small fraction of the material in the outer parts eventually formed the other objects.

During the past 50 years, we have learned more about the solar system than anyone imagined before the space age. In addition to gathering information with powerful new telescopes, we have sent spacecraft directly to many members of the planetary system. (Planetary astronomy is the only branch of our science in which we can, at least vicariously, travel to the objects we want to study.) With evocative names such as *Voyager*, *Pioneer*, *Curiosity*, and *Pathfinder*, our robot explorers have flown past, orbited, or landed on every planet, returning images and data that have dazzled both astronomers and the public. In the process, we have also investigated two dwarf planets, hundreds of fascinating moons, four ring systems, a dozen asteroids, and several comets (smaller members of our solar system that we will discuss later).

Our probes have penetrated the atmosphere of Jupiter and landed on the surfaces of Venus, Mars, our Moon, Saturn's moon Titan, the asteroids Eros, Itokawa, Ryugu, and Bennu, and the Comet Churyumov-Gerasimenko (usually referred to as 67P). Humans have set foot on the Moon and returned samples of its surface soil for laboratory analysis (Figure 7.2). We have flown a helicopter drone on Mars. We have even discovered other places in our solar system that might be able to support some kind of life.



Figure 7.2 : Astronauts on the Moon. The lunar lander and surface rover from the Apollo 15 mission are seen in this view of the one place beyond Earth that has been explored directly by humans. (credit: modification of work by David R. Scott, NASA)

Link to Learning

View this gallery of NASA images that trace the history of the Apollo mission.

An Inventory

The Sun, a star that is brighter than about 80% of the stars in the Galaxy, is by far the most massive member of the solar system, as shown in Table 7.1. It is an enormous ball about 1.4 million kilometers in diameter, with surface layers of incandescent gas and an interior temperature of millions of degrees. The Sun will be discussed in later chapters as our first, and best-studied, example of a star.

Table 7.1: Mass of Members of the Solar System

Object	Percentage of Total Mass of Solar System
Sun	99.80
Jupiter	0.10
Comets	0.0005–0.03 (estimate)
All other planets and dwarf planets	0.04
Moons and rings	0.00005
Asteroids	0.000002 (estimate)
Cosmic dust	0.0000001 (estimate)

Table 7.1 also shows that most of the material of the planets is actually concentrated in the largest one, Jupiter, which is more massive than all the rest of the planets combined. Astronomers were able to determine the masses of the planets centuries ago using Kepler's laws of planetary motion and Newton's law of gravity to measure the planets' gravitational effects on one another or on moons that orbit them (see Orbits and Gravity). Today, we make even more precise measurements of their masses by tracking their gravitational effects on the motion of spacecraft that pass near them.

Beside Earth, five other planets were known to the ancients—Mercury, Venus, Mars, Jupiter, and Saturn—and two were discovered after the invention of the telescope: Uranus and Neptune. The eight planets all revolve in the same direction around the Sun. They orbit in approximately the same plane, like cars traveling on concentric tracks on a giant, flat racecourse. Each planet stays in its own "traffic lane," following a nearly circular orbit about the Sun and obeying the "traffic" laws discovered by Galileo, Kepler, and Newton. Besides these planets, we have also been discovering smaller worlds beyond Neptune that are called trans-Neptunian objects or TNOs (see Figure 7.3). The first to be found, in 1930, was Pluto, but others have been discovered during the twenty-first century. One of them, Eris, is about the same size as Pluto and has at least one moon (Pluto has five known moons.) The largest TNOs are also classed as *dwarf planets,* as is the largest asteroid, Ceres. (Dwarf planets will be discussed further in the chapter on Rings, Moons, and Pluto). To date, more than 2600 of these TNOs have been discovered, and one, called Arrokoth, was explored by the New Horizons spacecraft.

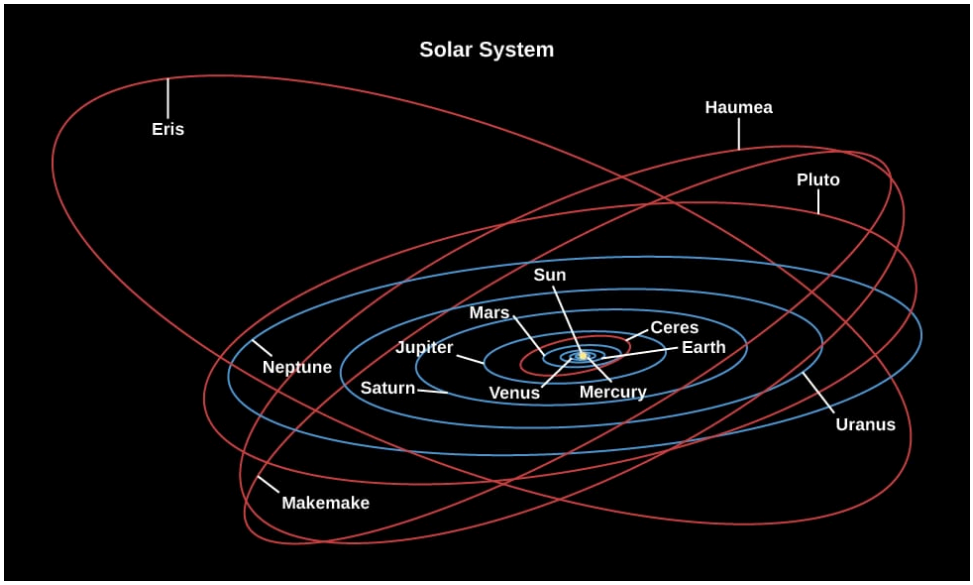


Figure 7.3 : Orbits of the Planets. All eight major planets orbit the Sun in roughly the same plane. The five currently known dwarf planets are also shown: Eris, Haumea, Pluto, Ceres, and Makemake. Note that Pluto's orbit is not in the plane of the planets.

Each of the planets and dwarf planets also rotates (spins) about an axis running through it, and in most cases the direction of rotation is the same as the direction of revolution about the Sun. The exceptions are Venus, which rotates backward very slowly (that is, in a retrograde direction), and Uranus and Pluto, which also have strange rotations, each spinning about an axis tipped nearly on its side. We do not yet know the spin orientations of Eris, Haumea, and Makemake.

The four planets closest to the Sun (Mercury through Mars) are called the inner or terrestrial planets. Often, the Moon is also discussed as a part of this group, bringing the total of terrestrial objects to five. (We generally call Earth's satellite "the Moon," with a capital M, and the other satellites "moons," with lowercase m's.) The terrestrial planets are relatively small worlds, composed primarily of rock and metal. All of them have solid surfaces that bear the records of their geological history in the forms of craters, mountains, and volcanoes (Figure 7.4).

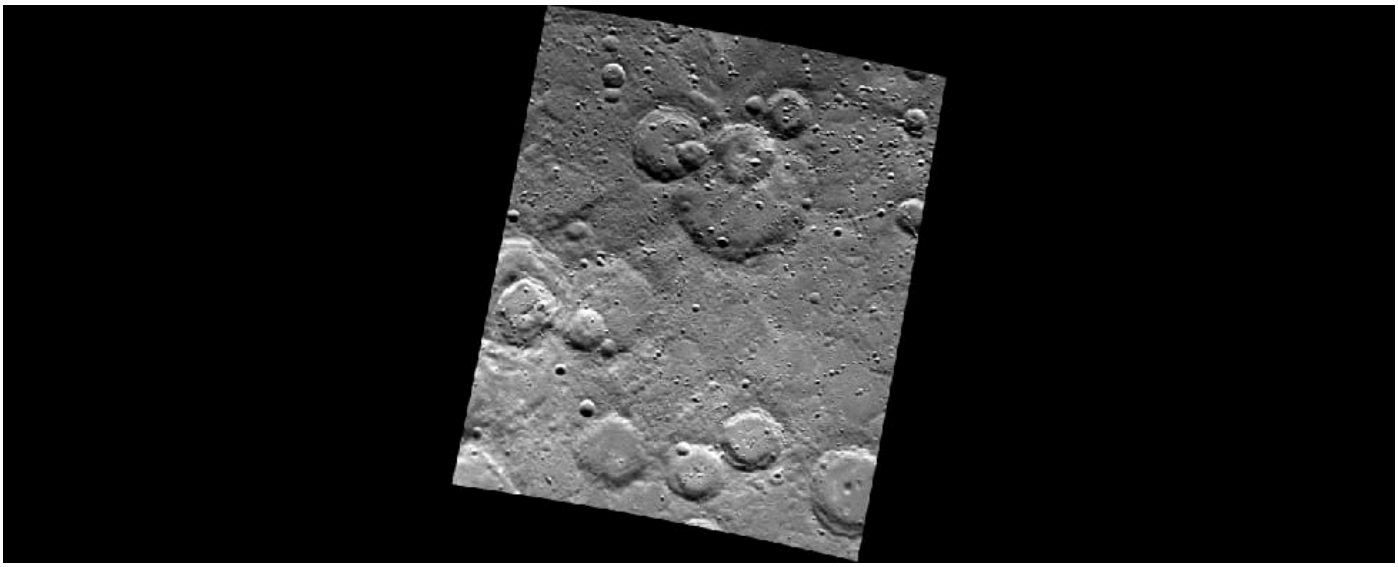


Figure 7.4 : Surface of Mercury. The pockmarked face of the terrestrial world of Mercury is more typical of the inner planets than the watery surface of Earth. This black-and-white image, taken with the Mariner 10 spacecraft, shows a region more than 400 kilometers wide. (credit: modification of work by NASA/John Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington)

The next four planets (Jupiter through Neptune) are much larger and are composed primarily of lighter ices, liquids, and gases. We call these four the jovian planets (after "Jove," another name for Jupiter in mythology) or giant planets—a name they richly deserve (Figure 7.5). About 1,300 Earths could fit inside Jupiter, for example. These planets do not have solid surfaces on which future explorers might land. They are more like vast, spherical oceans with much smaller, dense cores.

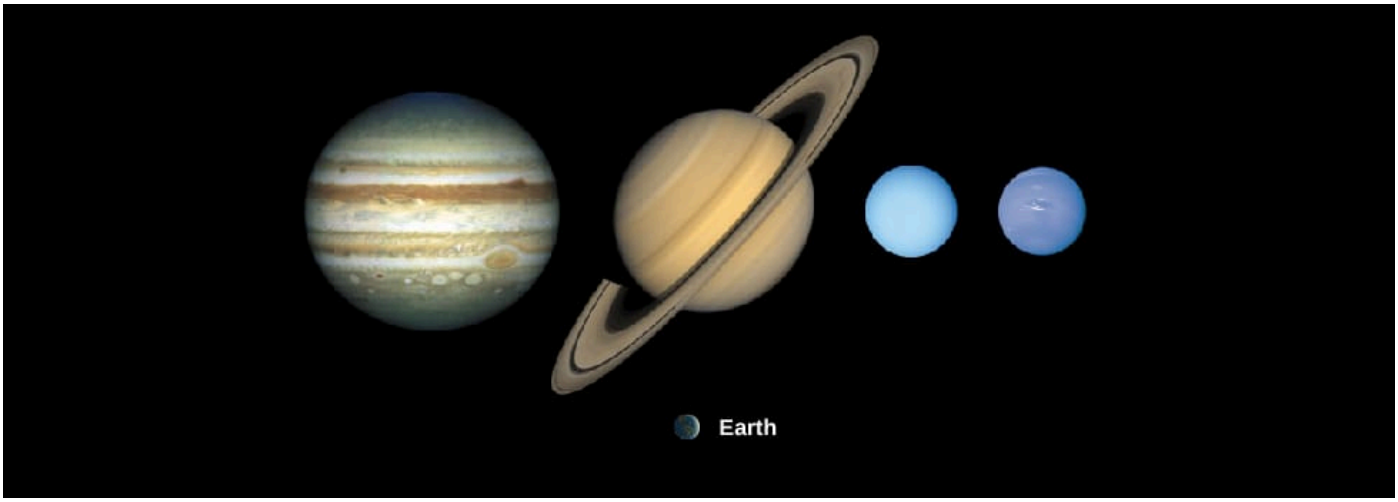


Figure 7.5 : The Four Giant Planets. This montage shows the four giant planets: Jupiter, Saturn, Uranus, and Neptune. Below them, Earth is shown to scale. (credit: modification of work by NASA, Solar System Exploration)

Near the outer edge of the system lies Pluto, which was the first of the distant icy worlds to be discovered beyond Neptune (Pluto was visited by a spacecraft, the NASA New Horizons mission, in 2015 [see Figure 7.6]). Table 7.2 summarizes some of the main facts about the planets.

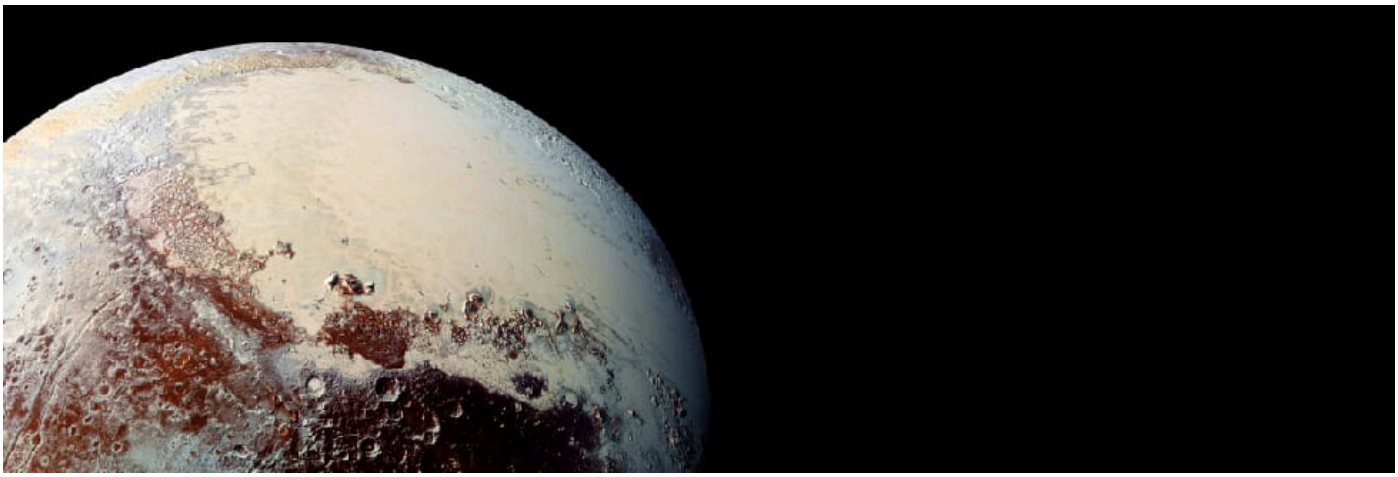


Figure 7.6 : Pluto Close-up. This intriguing image from the New Horizons spacecraft, taken when it flew by the dwarf planet in July 2015, shows some of its complex surface features. The rounded white area is called the Sputnik Plain, after humanity's first spacecraft. (credit: modification of work by NASA/Johns Hopkins University Applied Physics Laboratory/Southwest Research Institute)

Table 7.2: The Planets

Name	Distance from Sun (AU)[2]	Revolution Period (y)	Diameter (km)	Mass (10^{23} kg)	Density (g/cm^3)[3]
Mercury	0.39	0.24	4,878	3.3	5.4
Venus	0.72	0.62	12,120	48.7	5.2
Earth	1.00	1.00	12,756	59.8	5.5
Mars	1.52	1.88	6,787	6.4	3.9
Jupiter	5.20	11.86	142,984	18,991	1.3
Saturn	9.54	29.46	120,536	5686	0.7
Uranus	19.18	84.07	51,118	866	1.3
Neptune	30.06	164.82	49,660	1030	1.6

Example 7.1: Comparing Densities

Let's compare the densities of several members of the solar system. The density of an object equals its mass divided by its volume. The volume (*V*) of a sphere (like a planet) is calculated using the equation

$$V = \frac{4}{3}\pi R^3$$

where π (the Greek letter pi) has a value of approximately 3.14. Although planets are not perfect spheres, this equation works well enough. The masses and diameters of the planets are given in Table 7.2. For data on selected moons, see Appendix G. Let's use Saturn's moon Mimas as our example, with a mass of 4×10^{19} kg and a diameter of approximately 400 km (radius, $200\text{km} = 2 \times 10^5 m$).

Solution

The volume of Mimas is

$$\frac{4}{3} \times 3.14 \times \left(2 \times 10^5\ \text{m}\right)^3 = 3.3 \times 10^{16}\ \text{m}^3.$$

Density is mass divided by volume:

$$\frac{4 \times 10^{19}\ \text{kg}}{3.3 \times 10^{16}\ \text{m}^3} = 1.2 \times 10^3\ \text{kg/m}^3.$$

Note that the density of water in these units is 1000 kg/m^3, so Mimas must be made mainly of ice, not rock. (Note that the density of Mimas given in Appendix G is 1.2, but the units used there are different. In that table, we give density in units of g/cm^3, for which the density of water equals 1. Can you show, by converting units, that 1 g/cm^3 is the same as 1000 kg/m^3?)

Exercise 11.1.1

Calculate the average density of our own planet, Earth. Show your work. How does it compare to the density of an ice moon like Mimas? See Table 7.2 for data.

Answer available at the end of this section.

Link to Learning

Learn more about NASA's mission to Pluto and see high-resolution images of Pluto and its moon Charon.

Smaller Members of the Solar System

Most of the planets are accompanied by one or more moons; only Mercury and Venus move through space alone. There are more than 210 known moons orbiting planets and dwarf planets (see Appendix G for a listing of the larger ones), and undoubtedly many other small ones remain undiscovered. The largest of the moons are as big as small planets and just as interesting. In addition to our Moon, they include the four largest moons of Jupiter (called the Galilean moons, after their discoverer) and the largest moons of Saturn and Neptune (confusingly named Titan and Triton).

Each of the giant planets also has rings made up of countless small bodies ranging in size from mountains to mere grains of dust, all in orbit about the equator of the planet. The bright rings of Saturn are, by far, the easiest to see. They are among the most beautiful sights in the solar system (Figure 7.7). But, all four ring systems are interesting to scientists because of their complicated forms, influenced by the pull of the moons that also orbit these giant planets.

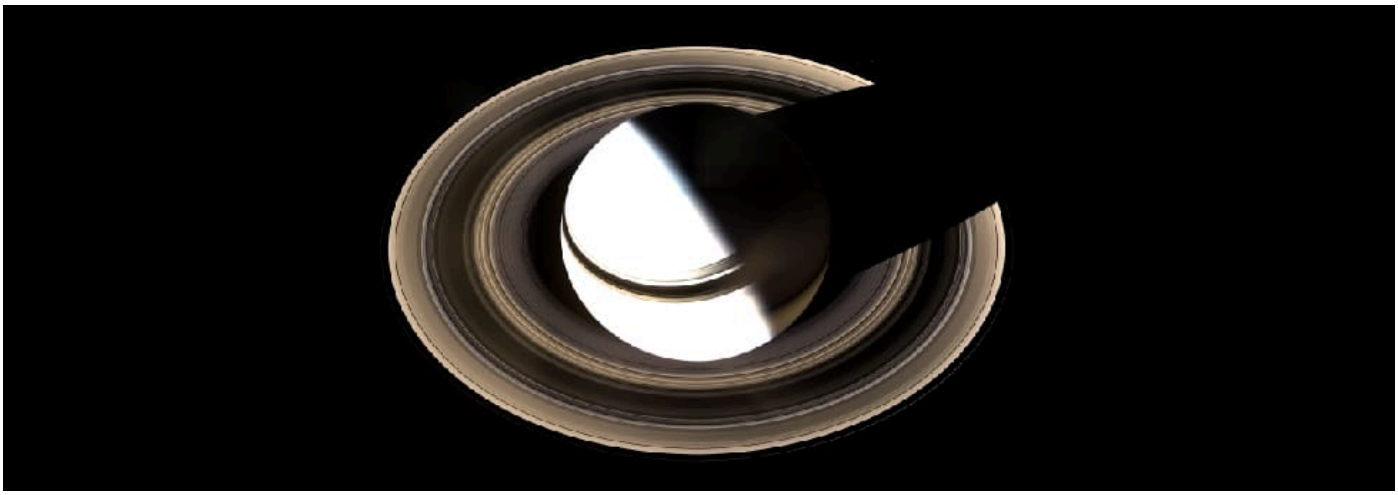


Figure 7.7 : Saturn and Its Rings. This 2007 Cassini image shows Saturn and its complex system of rings, taken from a distance of about 1.2 million kilometers. This natural-color image is a composite of 36 images taken over the course of 2.5 hours. (credit: modification of work by NASA/JPL/Space Science Institute)

The solar system has many other less-conspicuous members. Another group is the asteroids, rocky bodies that orbit the Sun like miniature planets, mostly in the space between Mars and Jupiter (although some do cross the orbits of planets like Earth—see Figure 7.8). Most asteroids are remnants of the initial population of the solar system that existed before the planets themselves formed. Some of the smallest moons of the planets, such as the moons of Mars, are very likely captured asteroids.

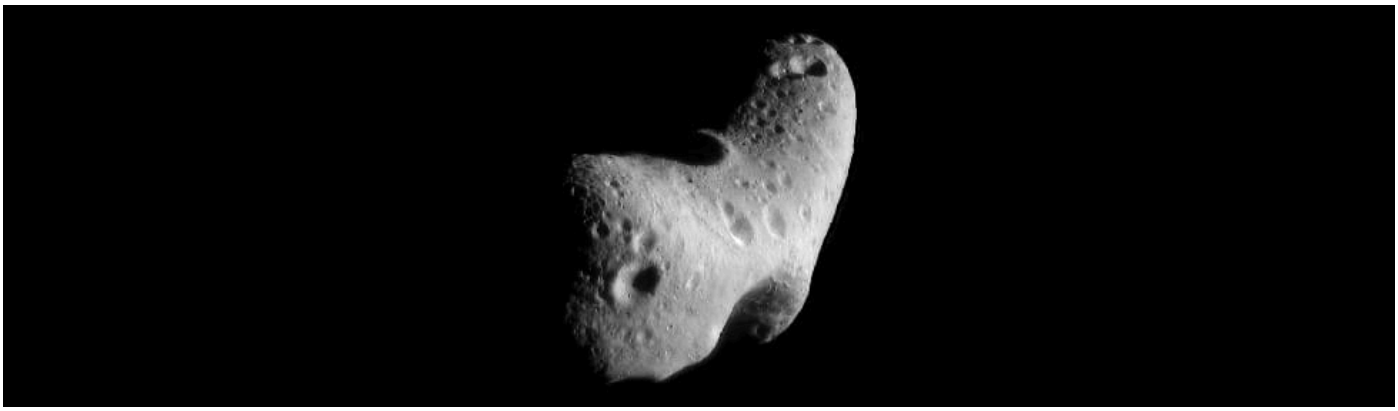


Figure 7.8 : Asteroid Eros. This small Earth-crossing asteroid image was taken by the NEAR-Shoemaker spacecraft from an altitude of about 100 kilometers. This view of the heavily cratered surface is about 10 kilometers wide. The spacecraft orbited Eros for a year before landing gently on its surface. (credit: modification of work by NASA/JHUAPL)

Another class of small bodies is composed mostly of ice, made of frozen gases such as water, carbon dioxide, and carbon monoxide; these objects are called comets (see Figure 7.9). Comets also are remnants from the formation of the solar system, but they were formed and continue (with rare exceptions) to orbit the Sun in distant, cooler regions—stored in a sort of cosmic deep freeze. This is also the realm of the larger icy worlds, called dwarf planets.

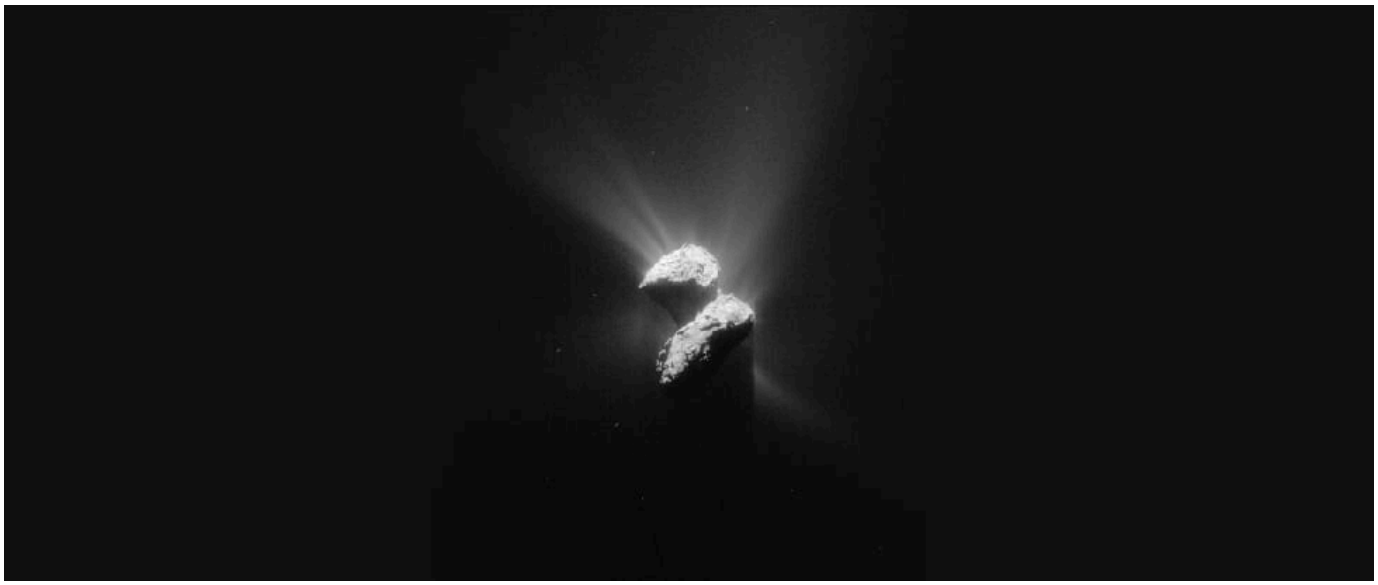


Figure 7.9 : Comet Churyumov-Gerasimenko (67P). This image shows Comet Churyumov-Gerasimenko, also known as 67P, near its closest approach to the Sun in 2015, as seen from the *Rosetta* spacecraft. Note the jets of gas escaping from the solid surface. (credit: modification of work by ESA/Rosetta/NAVACAM, CC BY-SA IGO 3.0)

Finally, there are countless grains of broken rock, which we call cosmic dust, scattered throughout the solar system. When these particles enter Earth's atmosphere (as millions do each day) they burn up, producing a brief flash of light in the night sky known as a meteor (meteors are often referred to as shooting stars). Occasionally, some larger chunk of rocky or metallic material survives its passage through the atmosphere and lands on Earth. Any piece that strikes the ground is known as a meteorite. (You can see meteorites on display in many natural history museums and can sometimes even purchase pieces of them from gem and mineral dealers.)

Voyagers in Astronomy: Carl Sagan: Solar System Advocate

The best-known astronomer in the world during the 1970s and 1980s, Carl Sagan devoted most of his professional career to studying the planets and considerable energy to raising public awareness of what we can learn from exploring the solar system (see Figure 7.10). Born in Brooklyn, New York, in 1934, Sagan became interested in astronomy as a youngster; he also credits science fiction stories for sustaining his fascination with what's out in the universe.



Figure 7.10 : Carl Sagan (1934–1996) and Neil deGrasse Tyson. Sagan was Tyson's inspiration to become a scientist. (credit "Sagan": modification of work by NASA, JPL; credit "Tyson": modification of work by Bruce F. Press)

In the early 1960s, when many scientists still thought Venus might turn out to be a hospitable place, Sagan calculated that the thick atmosphere of Venus could act like a giant greenhouse, keeping the heat in and raising the temperature enormously. He showed that the seasonal changes astronomers had seen on Mars were caused, not by vegetation, but by wind-blown dust. He was a member of the scientific teams for many of the robotic missions that explored the solar system and was instrumental in getting NASA to put a message-bearing plaque aboard the Pioneer spacecraft, as well as audio-video records on the Voyager spacecraft—all of them destined to leave our solar system entirely and send these little bits of Earth technology out among the stars.

To encourage public interest and public support of planetary exploration, Sagan helped found The Planetary Society, now the largest space-interest organization in the world. He was a tireless and eloquent advocate of the need to study the solar system close-up and the value of learning about other worlds in order to take better care of our own.

Sagan simulated conditions on early Earth to demonstrate how some of life's fundamental building blocks might have formed from the "primordial soup" of natural compounds on our planet. In addition, he and his colleagues developed computer models showing the consequences of nuclear war for Earth would be even more devastating than anyone had thought (this is now called the nuclear winter hypothesis) and demonstrating some of the serious consequences of continued pollution of our atmosphere.

Sagan was perhaps best known, however, as a brilliant popularizer of astronomy and the author of many books on science, including the best-selling *Cosmos*, and several evocative tributes to solar system exploration such as *The Cosmic Connection* and *Pale Blue Dot*. His book *The Demon Haunted World*, completed just before his death in 1996, is perhaps the best antidote to fuzzy thinking about pseudo-science and irrationality in print today. An intriguing science fiction novel he wrote, titled *Contact*, which became a successful film as well, is still recommended by many science instructors as a scenario for making contact with life elsewhere that is much more reasonable than most science fiction.

Sagan was a master, too, of the television medium. His 13-part public television series, *Cosmos*, was seen by an estimated 500 million people in 60 countries and has become one of the most-watched series in the history of public broadcasting. A few astronomers scoffed at a scientist who spent so much time in the public eye, but it is probably fair to say that Sagan's enthusiasm and skill as an explainer won more friends for the science of astronomy than anyone or anything else in the second half of the twentieth century.

In the two decades since Sagan's death, no other scientist has achieved the same level of public recognition. Perhaps closest is the director of the Hayden Planetarium, Neil deGrasse Tyson, who followed in Sagan's footsteps by making an updated version of the *Cosmos* program in 2014. Tyson is quick to point out that Sagan was his inspiration to become a scientist, telling how Sagan invited him to visit for a day at Cornell when he was a high school student looking for a career. However, the media environment has fragmented a great deal since Sagan's time. It is interesting to speculate whether Sagan could have adapted his communication style to the world of cable television, Twitter, Facebook, and podcasts.

Link to Learning

Two imaginative videos provide a tour of the solar system objects we have been discussing. Shane Gellert's I Need Some Space uses NASA photography and models to show the various worlds with which we share our system. In the more science fiction-oriented Wanderers video, we see some of the planets and moons as tourist destinations for future explorers, with commentary taken from recordings by Carl Sagan.

A Scale Model of the Solar System

Astronomy often deals with dimensions and distances that far exceed our ordinary experience. What does 1.4 billion kilometers—the distance from the Sun to Saturn—really mean to anyone? It can be helpful to visualize such large systems in terms of a scale model.

In our imaginations, let us build a scale model of the solar system, adopting a scale factor of 1 billion (10^9)—that is, reducing the actual solar system by dividing every dimension by a factor of 10^9. Earth, then, has a diameter of 1.3 centimeters, about the size of a grape. The Moon is a pea orbiting this at a distance of 40 centimeters, or a little more than a foot away. The Earth-Moon system fits into a standard backpack.

In this model, the Sun is nearly 1.5 meters in diameter, about the average height of an adult, and our Earth is at a distance of 150 meters—about one city block—from the Sun. Jupiter is five blocks away from the Sun, and its diameter is 15 centimeters, about the size of a very large grapefruit. Saturn is 10 blocks from the Sun; Uranus, 20 blocks; and Neptune, 30 blocks. Pluto, with a distance that varies quite a bit during its 249-year orbit, is currently just beyond 30 blocks and getting farther with time. Most of the moons of the outer solar system are the sizes of various kinds of seeds orbiting the grapefruit, oranges, and lemons that represent the outer planets.

In our scale model, a human is reduced to the dimensions of a single atom, and cars and spacecraft to the size of molecules. Sending the Voyager spacecraft to Neptune involves navigating a single molecule from the Earth–grape toward a lemon 5 kilometers away with an accuracy equivalent to the width of a thread in a spider's web.

If that model represents the solar system, where would the nearest stars be? If we keep the same scale, the closest stars would be tens of thousands of kilometers away. If you built this scale model in the city where you live, you would have to place the representations of these stars on the other side of Earth or beyond.

By the way, model solar systems like the one we just presented have been built in cities throughout the world. In Sweden, for example, Stockholm's huge Globe Arena has become a model for the Sun, and Pluto is represented by a 12-centimeter sculpture in the small town of Delsbo, 300 kilometers away. Another model solar system is in Washington on the Mall between the White House and Congress (perhaps proving they are worlds apart?).

Link to Learning

This model of the solar system shows all orbits and sizes to scale, and it lets you fly between the planets at an enhanced speed.

Making Connections: Names in the Solar System

We humans just don't feel comfortable until something has a name. Types of butterflies, new elements, and the mountains of Venus all need names for us to feel we are acquainted with them. How do we give names to objects and features in the solar system?

Planets and moons are named after gods and heroes in Greek and Roman mythology (with a few exceptions among the moons of Uranus, which have names drawn from English literature). When William Herschel, a German immigrant to England, first discovered the planet we now call Uranus, he wanted to name it Georgium Sidus (George's star) after King George III of his adopted country. This caused such an outcry among astronomers in other nations, however, that the classic tradition was upheld—and has been maintained ever since. Luckily, there were a lot of minor gods in the ancient pantheon, so plenty of names are left for the many small moons we are discovering around the giant planets. (Appendix G lists the larger moons). More recently, the names of dwarf planets and their moons have been drawn from the mythology of other cultures besides Greek and Roman.

Comets are often named after their discoverers (offering an extra incentive to comet hunters). Asteroids are named by their discoverers after just about anyone or anything they want. Recently, asteroid names have been used to recognize people who have made significant contributions to astronomy, including the three senior authors of this book.

That was pretty much all the naming that was needed while our study of the solar system was confined to Earth. But now, our spacecraft have surveyed and photographed many worlds in great detail, and each world has a host of features that also need names. To make sure that naming things in space remains multinational, rational, and somewhat dignified, astronomers have given the responsibility of approving names to a special committee of the International Astronomical Union (IAU), the body that includes scientists from every country that does astronomy.

This IAU committee has developed a set of rules for naming features on other worlds. For example, craters on Venus are named for women who have made significant contributions to human knowledge and welfare. Volcanic features on Jupiter's moon Io, which is in a constant state of volcanic activity, are named after gods of fire and thunder from the mythologies of many cultures. Craters on Mercury commemorate famous novelists, playwrights, artists, and composers. On Saturn's moon Tethys, all the features are named after characters and places in Homer's great epic poem, *The Odyssey*. As we explore

further, it may well turn out that more places in the solar system need names than Earth history can provide. Perhaps by then, explorers and settlers on these worlds will be ready to develop their own names for the places they may (if but for a while) call home.

You may be surprised to know that the meaning of the word *planet* has recently become controversial because we have discovered many other planetary systems that don't look very much like our own. Even within our solar system, the planets differ greatly in size and chemical properties. The biggest dispute concerns Pluto, which is much smaller than the other eight major planets. The category of dwarf planet was invented to include Pluto and similar icy objects beyond Neptune. But is a dwarf planet also a planet? Logically, it should be, but even this simple issue of grammar has been the subject of heated debate among both astronomers and the general public.

Footnotes

- 1The generic term for a group of planets and other bodies circling a star is *planetary system*. Ours is called the *solar system* because our Sun is sometimes called *Sol*. Strictly speaking, then, there is only one solar system; planets orbiting other stars are in planetary systems.
- 2An AU (or astronomical unit) is the distance from Earth to the Sun.
- 3We give densities in units where the density of water is 1 g/cm^3. To get densities in units of kg/m^3, multiply the given value by 1000.

Answers

Exercise 11.1.1

For a sphere,
$\text{density} = \frac{\text{mass}}{(\frac{4}{3}\pi R^3)}\ \text{kg/m}^3$.
For Earth, then,
$\text{density} = \frac{6\times10^{24}\,\text{kg}}{4.2\times2.6\times10^{20}\,\text{m}^3} = 5.5\ \times\ 10^3\,\text{kg/m}^3$.
This density is four to five times greater than Mimas'. In fact, Earth is the densest of the planets.

This page titled 11.1: Overview of Our Planetary System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 7.2: Overview of Our Planetary System by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

11.2: Compostion and Structure of Planets

Learning Objectives

By the end of this section, you will be able to:

- Describe the characteristics of the giant planets, terrestrial planets, and small bodies in the solar system
- Explain what influences the temperature of a planet's surface
- Explain why there is geological activity on some planets and not on others

The fact that there are two distinct kinds of planets—the rocky terrestrial planets and the gas-rich jovian planets—leads us to believe that they formed under different conditions. Certainly their compositions are dominated by different elements. Let us look at each type in more detail.

The Giant Planets

The two largest planets, Jupiter and Saturn, have nearly the same chemical makeup as the Sun; they are composed primarily of the two elements hydrogen and helium, with 75% of their mass being hydrogen and 25% helium. On Earth, both hydrogen and helium are gases, so Jupiter and Saturn are sometimes called gas planets. But, this name is misleading. Jupiter and Saturn are so large that the gas is compressed in their interior until the hydrogen becomes a liquid. Because the bulk of both planets consists of compressed, liquefied hydrogen, we should really call them liquid planets.

Under the force of gravity, the heavier elements sink toward the inner parts of a liquid or gaseous planet. Both Jupiter and Saturn, therefore, have cores composed of heavier rock, metal, and ice, but we cannot see these regions directly. In fact, when we look down from above, all we see is the atmosphere with its swirling clouds (Figure 7.11). We must infer the existence of the denser core inside these planets from studies of each planet's gravity.



Figure 7.11 : Jupiter. This true-color image of Jupiter was taken from the Cassini spacecraft in 2000. The dark spot is the shadow of one of the giant planet's moons. (credit: modification of work by NASA/JPL/University of Arizona)

Uranus and Neptune are much smaller than Jupiter and Saturn, but each also has a core of rock, metal, and ice. Uranus and Neptune were less efficient at attracting hydrogen and helium gas, so they have much smaller atmospheres in proportion to their cores.

Chemically, each giant planet is dominated by hydrogen and its many compounds. Nearly all the oxygen present is combined

chemically with hydrogen to form water (H_2O). Chemists call such a hydrogen-dominated composition *reduced*. Throughout the outer solar system, we find abundant water (mostly in the form of ice) and reducing chemistry.

The Terrestrial Planets

The terrestrial planets are quite different from the giants. In addition to being much smaller, they are composed primarily of rocks and metals. These, in turn, are made of elements that are less common in the universe as a whole. The most abundant rocks, called silicates, are made of silicon and oxygen, and the most common metal is iron. We can tell from their densities (see Table 7.2) that Mercury has the greatest proportion of metals (which are denser) and the Moon has the lowest. Earth, Venus, and Mars all have roughly similar bulk compositions: about one third of their mass consists of iron-nickel or iron-sulfur combinations; two thirds is made of silicates. Because these planets are largely composed of oxygen compounds (such as the silicate minerals of their crusts), their chemistry is said to be *oxidized*.

When we look at the internal structure of each of the terrestrial planets, we find that the densest metals are in a central core, with the lighter silicates near the surface. If these planets were liquid, like the giant planets, we could understand this effect as the result the sinking of heavier elements due to the pull of gravity. This leads us to conclude that, although the terrestrial planets are solid today, at one time they must have been hot enough to melt.

Differentiation is the process by which gravity helps separate a planet's interior into layers of different compositions and densities. The heavier metals sink to form a core, while the lightest minerals float to the surface to form a crust. Later, when the planet cools, this layered structure is preserved. In order for a rocky planet to differentiate, it must be heated to the melting point of rocks, which is typically more than 1300 K.

Moons, Asteroids, and Comets

Chemically and structurally, Earth's Moon is like the terrestrial planets, but most moons are in the outer solar system, and they have compositions similar to the cores of the giant planets around which they orbit. The three largest moons—Ganymede and Callisto in the jovian system, and Titan in the saturnian system—are composed half of frozen water, and half of rocks and metals. Most of these moons differentiated during formation, and today they have cores of rock and metal, with upper layers and crusts of very cold and—thus very hard—ice (Figure 7.12).



Figure 7.12 : Ganymede. This view of Jupiter's moon Ganymede was taken in June 1996 by the Galileo spacecraft. The brownish gray color of the surface indicates a dusty mixture of rocky material and ice. The bright spots are places where recent impacts have uncovered fresh ice from underneath. (credit: modification of work by NASA/JPL)

Most of the asteroids and comets, as well as the smallest moons, were probably never heated to the melting point. However, some of the largest asteroids, such as Vesta, appear to be differentiated; others are fragments from differentiated bodies. Many of the smaller objects seem to be fragments or rubble piles that are the result of collisions. Because most asteroids and comets retain their original composition, they represent relatively unmodified material dating back to the time of the formation of the solar

system. In a sense, they act as chemical fossils, helping us to learn about a time long ago whose traces have been erased on larger worlds.

Temperatures: Going to Extremes

Generally speaking, the farther a planet or moon is from the Sun, the cooler its surface. The planets are heated by the radiant energy of the Sun, which gets weaker with the square of the distance. You know how rapidly the heating effect of a fireplace or an outdoor radiant heater diminishes as you walk away from it; the same effect applies to the Sun. Mercury, the closest planet to the Sun, has a blistering surface temperature that ranges from 280–430 °C on its sunlit side, whereas the surface temperature on Pluto is only about –220 °C, colder than liquid air.

Mathematically, the temperatures decrease approximately in proportion to the square root of the distance from the Sun. Pluto is about 30 AU at its closest to the Sun (or 100 times the distance of Mercury) and about 49 AU at its farthest from the Sun. Thus, Pluto's temperature is less than that of Mercury by the square root of 100, or a factor of 10: from 500 K to 50 K.

In addition to its distance from the Sun, the surface temperature of a planet can be influenced strongly by its atmosphere. Without our atmospheric insulation (the greenhouse effect, which keeps the heat in), the oceans of Earth would be permanently frozen. Conversely, if Mars once had a larger atmosphere in the past, it could have supported a more temperate climate than it has today. Venus is an even more extreme example, where its thick atmosphere of carbon dioxide acts as insulation, reducing the escape of heat built up at the surface, resulting in temperatures greater than those on Mercury. Today, Earth is the only planet where surface temperatures generally lie between the freezing and boiling points of water. As far as we know, Earth is the only planet to support life.

Astronomy Basics: There's No Place Like Home

In the classic film *The Wizard of Oz*, Dorothy, the heroine, concludes after her many adventures in "alien" environments that "there's no place like home." The same can be said of the other worlds in our solar system. There are many fascinating places, large and small, that we might like to visit, but humans could not survive on any without a great deal of artificial assistance.

A thick carbon dioxide atmosphere keeps the surface temperature on our neighbor Venus at a sizzling 700 K (near 900 °F). Mars, on the other hand, has temperatures generally below freezing, with air (also mostly carbon dioxide) so thin that it resembles that found at an altitude of 30 kilometers (100,000 feet) in Earth's atmosphere. And the red planet is so dry that it has not had any rain for billions of years.

The outer layers of the jovian planets are neither warm enough nor solid enough for human habitation. Any bases we build in the systems of the giant planets may well have to be in space or one of their moons—none of which is particularly hospitable to a luxury hotel with a swimming pool and palm trees. Perhaps we will find warmer havens deep inside the clouds of Jupiter or in the ocean under the frozen ice of its moon Europa.

All of this suggests that we had better take good care of Earth because it is the only site where life as we know it could survive. Recent human activity may be reducing the habitability of our planet by adding pollutants to the atmosphere, especially the potent greenhouse gas carbon dioxide. Human civilization is changing our planet dramatically, and these changes are not necessarily for the better. In a solar system that seems unready to receive us, making Earth less hospitable to life may be a grave mistake.

Geological Activity

The crusts of all of the terrestrial planets, as well as of the larger moons, and Pluto, have been modified over their histories by both internal and external forces. Externally, each has been battered by a slow rain of projectiles from space, leaving their surfaces pockmarked by impact craters of all sizes (see Figure 7.4). We have good evidence that this bombardment was far greater in the early history of the solar system, but it certainly continues to this day, even if at a lower rate. The collision of more than 20 large pieces of Comet Shoemaker–Levy 9 with Jupiter in the summer of 1994 (see Figure 7.13) is one dramatic example of this process.

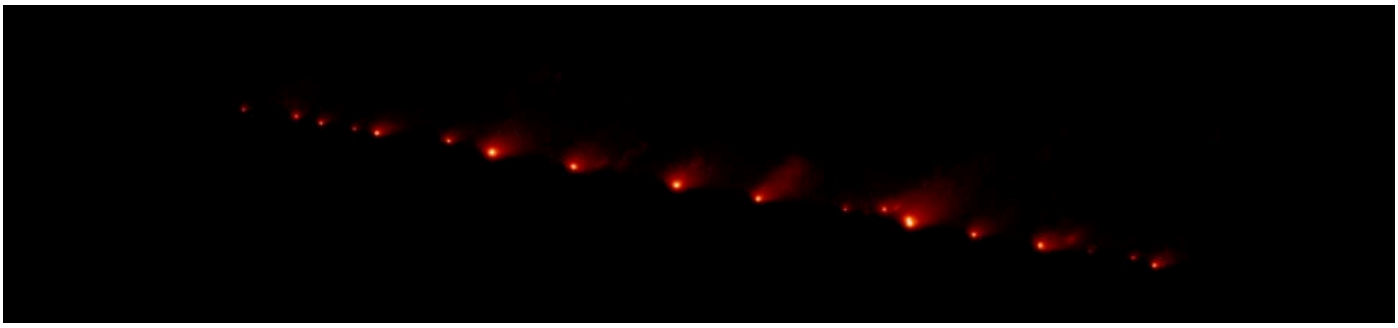


Figure 7.13 : Comet Shoemaker–Levy 9. In this image of Comet Shoemaker–Levy 9 taken on May 17, 1994, by NASA's Hubble Space Telescope, you can see about 20 icy fragments into which the comet broke. The comet was approximately 660 million kilometers from Earth, heading on a collision course with Jupiter. (credit: modification of work by NASA, ESA, H. Weaver (STScI), E. Smith (STScI))

Figure 7.14 shows the aftermath of these collisions, when debris clouds larger than Earth could be seen in Jupiter's atmosphere.

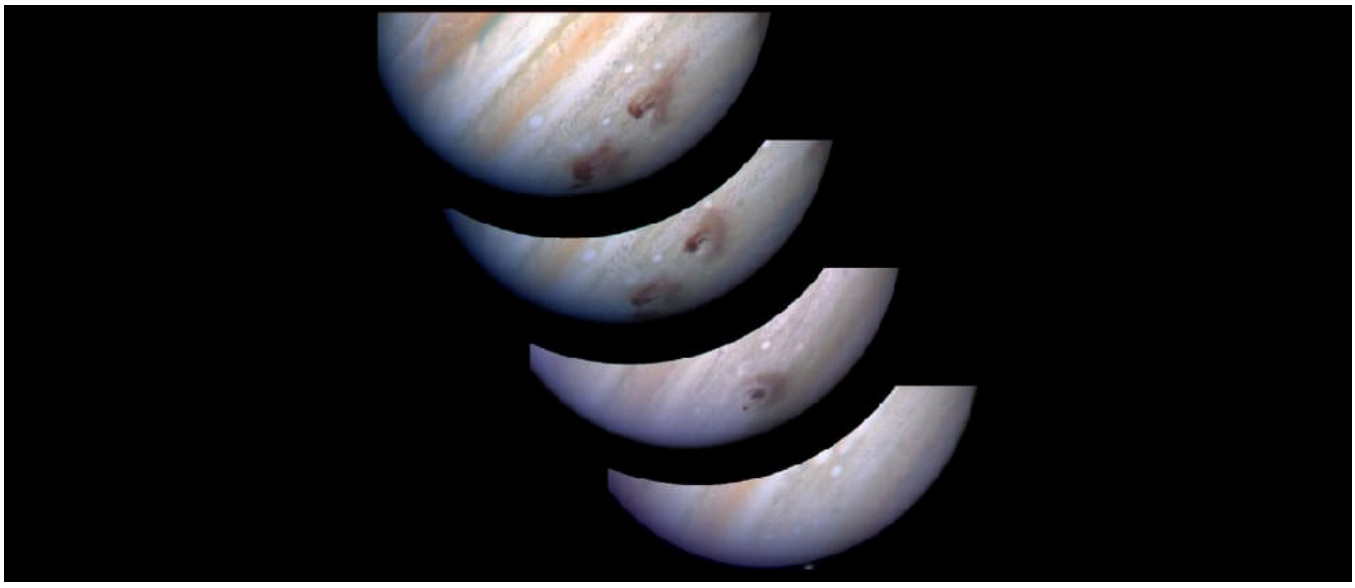


Figure 7.14 : Jupiter with Huge Dust Clouds. The Hubble Space Telescope took this sequence of images of Jupiter in summer 1994, when fragments of Comet Shoemaker–Levy 9 collided with the giant planet. Here we see the site hit by fragment G, from five minutes to five days after impact. Several of the dust clouds generated by the collisions became larger than Earth. (credit: modification of work by H. Hammel, NASA)

During the time all the planets have been subject to such impacts, internal forces on the terrestrial planets have buckled and twisted their crusts, built up mountain ranges, erupted as volcanoes, and generally reshaped the surfaces in what we call geological activity. (The prefix *geo* means "Earth," so this is a bit of an "Earth-chauvinist" term, but it is so widely used that we bow to tradition.) Among the terrestrial planets, Earth and Venus have experienced the most geological activity over their histories, although some of the moons in the outer solar system are also surprisingly active. In contrast, our own Moon is a dead world where geological activity ceased billions of years ago.

Geological activity on a planet is the result of a hot interior. The forces of volcanism and mountain building are driven by heat escaping from the interiors of planets. As we will see, each of the planets was heated at the time of its birth, and this primordial heat initially powered extensive volcanic activity, even on our Moon. But, small objects such as the Moon soon cooled off. The larger the planet or moon, the longer it retains its internal heat, and therefore the more we expect to see surface evidence of continuing geological activity. The effect is similar to our own experience with a hot baked potato: the larger the potato, the more slowly it cools. If we want a potato to cool quickly, we cut it into small pieces.

For the most part, the history of volcanic activity on the terrestrial planets conforms to the predictions of this simple theory. The Moon, the smallest of these objects, is a geologically dead world. Although we know less about Mercury, it seems likely that this planet, too, ceased most volcanic activity about the same time the Moon did. Mars represents an intermediate case. It has been much more active than the Moon, but less so than Earth. Earth and Venus, the largest terrestrial planets, still have molten interiors even today, some 4.5 billion years after their birth.

This page titled 11.2: Compostion and Structure of Planets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 7.3: Compostion and Structure of Planets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

11.3: Dating Planetary Surfaces

Learning Objectives

By the end of this section, you will be able to:

- Explain how astronomers can tell whether a planetary surface is geologically young or old
- Describe different methods for dating planets

How do we know the age of the surfaces we see on planets and moons? If a world has a surface (as opposed to being mostly gas and liquid), astronomers have developed some techniques for estimating how long ago that surface solidified. Note that the age of these surfaces is not necessarily the age of the planet as a whole. On geologically active objects (including Earth), vast outpourings of molten rock or the erosive effects of water and ice, which we call planet weathering, have erased evidence of earlier epochs and present us with only a relatively young surface for investigation.

Counting the Craters

One way to estimate the age of a surface is by counting the number of impact craters. This technique works because the rate at which impacts have occurred in the solar system has been roughly constant for several billion years. Thus, in the absence of forces to eliminate craters, the number of craters is simply proportional to the length of time the surface has been exposed. This technique has been applied successfully to many solid planets and moons (Figure 7.15).

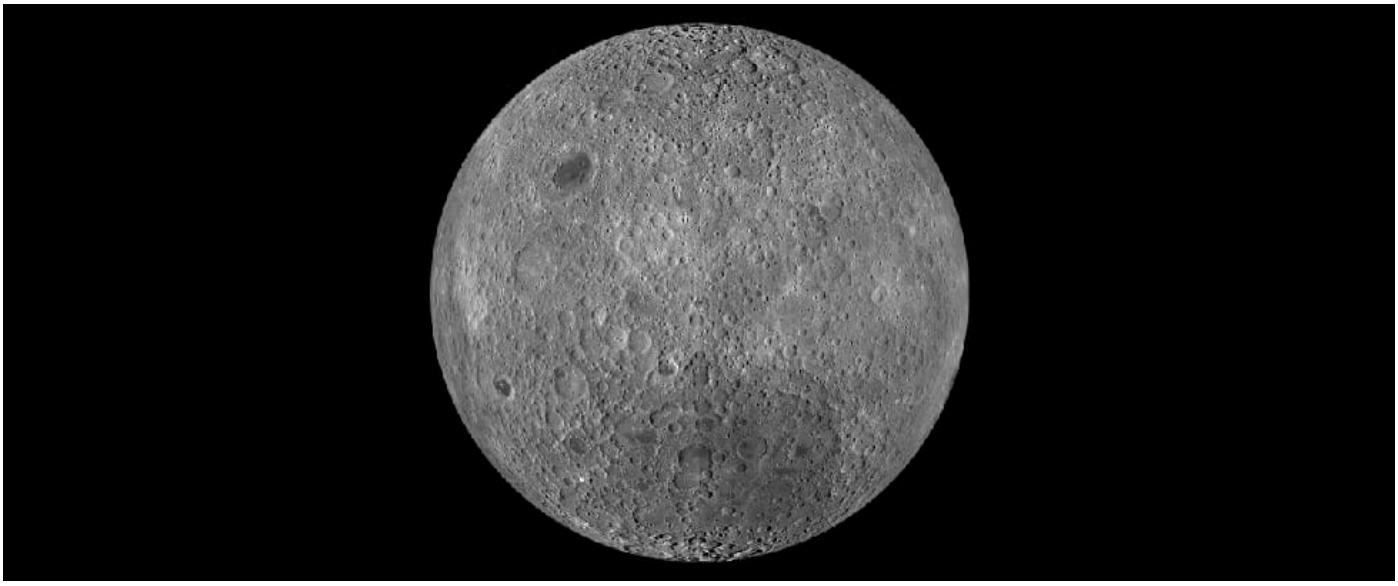


Figure 7.15 : Our Cratered Moon. This composite image of the Moon's surface was made from many smaller images taken between November 2009 and February 2011 by the Lunar Reconnaissance Orbiter (LRO) and shows craters of many different sizes. (credit: modification of work by NASA/GSFC/Arizona State University)

Bear in mind that crater counts can tell us only the time since the surface experienced a major change that could modify or erase preexisting craters. Estimating ages from crater counts is a little like walking along a sidewalk in a snowstorm after the snow has been falling steadily for a day or more. You may notice that in front of one house the snow is deep, while next door the sidewalk may be almost clear. Do you conclude that less snow has fallen in front of Ms. Jones' house than Mr. Smith's? More likely, you conclude that Jones has recently swept the walk clean and Smith has not. Similarly, the numbers of craters indicate how long it has been since a planetary surface was last "swept clean" by ongoing lava flows or by molten materials ejected when a large impact happened nearby.

Still, astronomers can use the numbers of craters on different parts of the same world to provide important clues about how

regions on that world evolved. On a given planet or moon, the more heavily cratered terrain will generally be older (that is, more time will have elapsed there since something swept the region clean).

Radioactive Rocks

Another way to trace the history of a solid world is to measure the age of individual rocks. After samples were brought back from the Moon by Apollo astronauts, the techniques that had been developed to date rocks on Earth were applied to rock samples from the Moon to establish a geological chronology for the Moon. Furthermore, a few samples of material from the Moon, Mars, and the large asteroid Vesta have fallen to Earth as meteorites and can be examined directly (see the chapter on Cosmic Samples and the Origin of the Solar System).

Scientists measure the age of rocks using the properties of natural radioactivity. Around the beginning of the twentieth century, physicists began to understand that some atomic nuclei are not stable but can split apart (decay) spontaneously into smaller nuclei. The process of radioactive decay involves the emission of particles such as electrons, or of radiation in the form of gamma rays (see the chapter on Radiation and Spectra).

For any one radioactive nucleus, it is not possible to predict when the decay process will happen. Such decay is random in nature, like the throw of dice: as gamblers have found all too often, it is impossible to say just when the dice will come up 7 or 11. But, for a very large number of dice tosses, we can calculate the odds that 7 or 11 will come up. Similarly, if we have a very large number of radioactive atoms of one type (say, uranium), there is a specific time period, called its half-life, during which the chances are fifty-fifty that decay will occur for any of the nuclei.

A particular nucleus may last a shorter or longer time than its half-life, but in a large sample, almost exactly half of the nuclei will have decayed after a time equal to one half-life. Half of the remaining nuclei will have decayed after two half-lives pass, leaving only one half of a half—or one quarter—of the original sample (Figure 7.16).

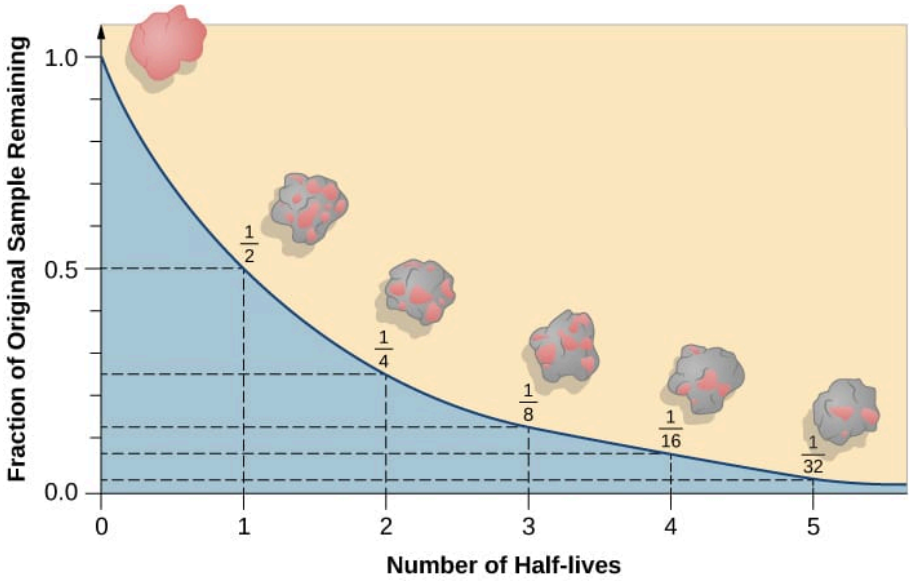


Figure 7.16 : Radioactive Decay. This graph shows (in pink) the amount of a radioactive sample that remains after several half-lives have passed. After one half-life, half the sample is left; after two half-lives, one half of the remainder (or one quarter) is left; and after three half-lives, one half of that (or one eighth) is left. Note that, in reality, the decay of radioactive elements in a rock sample would not cause any visible change in the appearance of the rock; the splashes of color are shown here for conceptual purposes only.

If you had 1 gram of pure radioactive nuclei with a half-life of 100 years, then after 100 years you would have 1/2 gram; after 200 years, 1/4 gram; after 300 years, only 1/8 gram; and so forth. However, the material does not disappear. Instead, the radioactive atoms are replaced with their decay products. Sometimes the radioactive atoms are called *parents* and the decay products are called *daughter* elements.

In this way, radioactive elements with half-lives we have determined can provide accurate nuclear clocks. By comparing how much of a radioactive parent element is left in a rock to how much of its daughter products have accumulated, we can learn how long the decay process has been going on and hence how long ago the rock formed. Table 7.3 summarizes the decay reactions used

most often to date lunar and terrestrial rocks.

Table 7.3: Radioactive Decay Reaction Used to Date Rocks[4]

Parent	Daughter	Half-Life (billions of years)
Samarium-147	Neodymium-143	106
Rubidium-87	Strontium-87	48.8
Thorium-232	Lead-208	14.0
Uranium-238	Lead-206	4.47
Potassium-40	Argon-40	1.31

Link to Learning

PBS provides an evolution series excerpt that explains how we use radioactive elements to date Earth.

This Science Channel video features Bill Nye the Science Guy showing how scientists have used radioactive dating to determine the age of Earth.

This radioactive decay simulator lets you examine how the number of atoms of a radioactive element decreases over time and see how the ratio of parent to child atoms could be used to figure out the age of the whole set of atoms.

When astronauts first flew to the Moon, one of their most important tasks was to bring back lunar rocks for radioactive age-dating. Until then, astronomers and geologists had no reliable way to measure the age of the lunar surface. Counting craters had let us calculate relative ages (for example, the heavily cratered lunar highlands were older than the dark lava plains), but scientists could not measure the actual age in years. Some thought that the ages were as young as those of Earth's surface, which has been resurfaced by many geological events. For the Moon's surface to be so young would imply active geology on our satellite. Only in 1969, when the first Apollo samples were dated, did we learn that the Moon is an ancient, geologically dead world. Using such dating techniques, we have been able to determine the ages of both Earth and the Moon: each was formed about 4.5 billion years ago (although, as we shall see, Earth probably formed earlier than the Moon).

We should also note that the decay of radioactive nuclei generally releases energy in the form of heat. Although the energy from a single nucleus is not very large (in human terms), the enormous numbers of radioactive nuclei in a planet or moon (especially early in its existence) can be a significant source of internal energy for that world. Geologists estimate that about half of Earth's current internal heat budget comes from the decay of radioactive isotopes in its interior.

Footnotes

- 4The number after each element is its atomic weight, equal to the number of protons plus neutrons in its nucleus. This specifies the *isotope* of the element; different isotopes of the same element differ in the number of neutrons.

This page titled 11.3: Dating Planetary Surfaces is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 7.4: Dating Planetary Surfaces by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

11.4: Origin of the Solar System

Learning Objectives

By the end of this section, you will be able to:

- Describe the characteristics of planets that are used to create formation models of the solar system
- Describe how the characteristics of extrasolar systems help us to model our own solar system
- Explain the importance of collisions in the formation of the solar system

Much of astronomy is motivated by a desire to understand the origin of things: to find at least partial answers to age-old questions of where the universe, the Sun, Earth, and we ourselves came from. Each planet and moon is a fascinating place that may stimulate our imagination as we try to picture what it would be like to visit. Taken together, the members of the solar system preserve patterns that can tell us about the formation of the entire system. As we begin our exploration of the planets, we want to introduce our modern picture of how the solar system formed.

The recent discovery of thousands of planets in orbit around other stars has shown astronomers that many exoplanetary systems can be quite different from our own solar system. For example, it is common for these systems to include planets intermediate in size between our terrestrial and giant planets. These are often called *superearths*. Some exoplanet systems even have giant planets close to the star, reversing the order we see in our system. In The Birth of Stars and the Discovery of Planets outside the Solar System, we will look at these exoplanet systems. But for now, let us focus on theories of how our own particular system has formed and evolved.

Looking for Patterns

One way to approach our question of origin is to look for regularities among the planets. We found, for example, that all the planets lie in nearly the same plane and revolve in the same direction around the Sun. The Sun also spins in the same direction about its own axis. Astronomers interpret this pattern as evidence that the Sun and planets formed together from a spinning cloud of gas and dust that we call the solar nebula (Figure 7.17).



Figure 7.17 : Solar Nebula. This artist's conception of the solar nebula shows the flattened cloud of gas and dust from which our planetary system formed. Icy and rocky planetesimals (precursors of the planets) can be seen in the foreground. The bright center is where the Sun is forming. (credit: William K. Hartmann, Planetary Science Institute)

The composition of the planets gives another clue about origins. Spectroscopic analysis allows us to determine which elements

are present in the Sun and the planets. The Sun has the same hydrogen-dominated composition as Jupiter and Saturn, and therefore appears to have been formed from the same reservoir of material. In comparison, the terrestrial planets and our Moon are relatively deficient in the light gases and the various ices that form from the common elements oxygen, carbon, and nitrogen. Instead, on Earth and its neighbors, we see mostly the rarer heavy elements such as iron and silicon. This pattern suggests that the processes that led to planet formation in the inner solar system must somehow have excluded much of the lighter materials that are common elsewhere. These lighter materials must have escaped, leaving a residue of heavy stuff.

The reason for this is not hard to guess, bearing in mind the heat of the Sun. The inner planets and most of the asteroids are made of rock and metal, which can survive heat, but they contain very little ice or gas, which evaporate when temperatures are high. (To see what we mean, just compare how long a rock and an ice cube survive when they are placed in the sunlight.) In the outer solar system, where it has always been cooler, the planets and their moons, as well as icy dwarf planets and comets, are composed mostly of ice and gas.

The Evidence from Far Away

A second approach to understanding the origins of the solar system is to look outward for evidence that other systems of planets are forming elsewhere. We cannot look back in time to the formation of our own system, but many stars in space are much younger than the Sun. In these systems, the processes of planet formation might still be accessible to direct observation. We observe that there are many other “solar nebulas” or *circumstellar disks*—flattened, spinning clouds of gas and dust surrounding young stars. These disks resemble our own solar system’s initial stages of formation billions of years ago (Figure 7.18).



Figure 7.18 : Atlas of Planetary Nurseries. These Hubble Space Telescope photos show sections of the Orion Nebula, a relatively close-by region where stars are currently forming. Each image shows an embedded circumstellar disk orbiting a very young star. Seen from different angles, some are energized to glow by the light of a nearby star while others are dark and seen in silhouette against the bright glowing gas of the Orion Nebula. Each is a contemporary analog of our own solar nebula—a location where planets are probably being formed today. (credit: modification of work by NASA/ESA, L. Ricci (ESO))

Building Planets

Circumstellar disks are a common occurrence around very young stars, suggesting that disks and stars form together. Astronomers can use theoretical calculations to see how solid bodies might form from the gas and dust in these disks as they cool. These models show that material begins to coalesce first by forming smaller objects, precursors of the planets, which we call planetesimals.

Today’s fast computers can simulate the way millions of planetesimals, probably no larger than 100 kilometers in diameter, might gather together under their mutual gravity to form the planets we see today. We are beginning to understand that this process was a violent one, with planetesimals crashing into each other and sometimes even disrupting the growing planets themselves. As a consequence of those violent impacts (and the heat from radioactive elements in them), all the planets were heated until they were liquid and gas, and therefore differentiated, which helps explain their present internal structures.

The process of impacts and collisions in the early solar system was complex and, apparently, often random. The solar nebula model can explain many of the regularities we find in the solar system, but the random collisions of massive collections of planetesimals could be the reason for some exceptions to the “rules” of solar system behavior. For example, why do the planets Uranus and Pluto spin on their sides? Why does Venus spin slowly and in the opposite direction from the other planets? Why

does the composition of the Moon resemble Earth in many ways and yet exhibit substantial differences? The answers to such questions probably lie in enormous collisions that took place in the solar system long before life on Earth began.

Today, some 4.5 billion years after its origin, the solar system is—thank goodness—a much less violent place. As we will see, however, some planetesimals have continued to interact and collide, and their fragments move about the solar system as roving "transients" that can make trouble for the established members of the Sun's family, such as our own Earth. (We discuss this "troublemaking" in Comets and Asteroids: Debris of the Solar System.)

Link to Learning

A great variety of infographics at space.com let you explore what it would be like to live on various worlds in the solar system.

This page titled 11.4: Origin of the Solar System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 7.5: Origin of the Solar System by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

11.5: Summary

Our solar system currently consists of the Sun, eight planets, five dwarf planets, nearly 200 known moons, and a host of smaller objects. The planets can be divided into two groups: the inner terrestrial planets and the outer giant planets. Pluto, Eris, Haumea, and Makemake do not fit into either category; as icy dwarf planets, they exist in an ice realm on the fringes of the main planetary system. The giant planets are composed mostly of liquids and gases. Smaller members of the solar system include asteroids (including the dwarf planet Ceres), which are rocky and metallic objects found mostly between Mars and Jupiter; comets, which are made mostly of frozen gases and generally orbit far from the Sun; and countless smaller grains of cosmic dust. When a meteor survives its passage through our atmosphere and falls to Earth, we call it a meteorite.

The giant planets have dense cores roughly 10 times the mass of Earth, surrounded by layers of hydrogen and helium. The terrestrial planets consist mostly of rocks and metals. They were once molten, which allowed their structures to differentiate (that is, their denser materials sank to the center). The Moon resembles the terrestrial planets in composition, but most of the other moons—which orbit the giant planets—have larger quantities of frozen ice within them. In general, worlds closer to the Sun have higher surface temperatures. The surfaces of terrestrial planets have been modified by impacts from space and by varying degrees of geological activity.

The ages of the surfaces of objects in the solar system can be estimated by counting craters: on a given world, a more heavily cratered region will generally be older than one that is less cratered. We can also use samples of rocks with radioactive elements in them to obtain the time since the layer in which the rock formed last solidified. The half-life of a radioactive element is the time it takes for half the sample to decay; we determine how many half-lives have passed by how much of a sample remains the radioactive element and how much has become the decay product. In this way, we have estimated the age of the Moon and Earth to be roughly 4.5 billion years.

Regularities among the planets have led astronomers to hypothesize that the Sun and the planets formed together in a giant, spinning cloud of gas and dust called the solar nebula. Astronomical observations show tantalizingly similar circumstellar disks around other stars. Within the solar nebula, material first coalesced into planetesimals; many of these gathered together to make the planets and moons. The remainder can still be seen as comets and asteroids. Probably all planetary systems have formed in similar ways, but many exoplanet systems have evolved along quite different paths, as we will see in Cosmic Samples and the Origin of the Solar System.

This page titled 11.5: Summary is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

- 7.7: Summary by OpenStax is licensed CC BY 4.0.

12: EARTH

12.1: The Global Perspective
12.2: Earth's Crust
12.3: Earth's Atmosphere
12.4: Life, Chemical Evolution, and Climate Change
12.5: Cosmic Influences on the Evolution of Earth
12.6: Key Terms

This page titled 12: Earth is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

12.1: The Global Perspective

Learning Objectives

By the end of this section, you will be able to:

- Describe the components of Earth's interior and explain how scientists determined its structure
- Specify the origin, size, and extent of Earth's magnetic field

Earth is a medium-size planet with a diameter of approximately 12,760 kilometers (Figure 8.2). As one of the inner or terrestrial planets, it is composed primarily of heavy elements such as iron, silicon, and oxygen—very different from the composition of the Sun and stars, which are dominated by the light elements hydrogen and helium. Earth's orbit is nearly circular, and Earth is warm enough to support liquid water on its surface. It is the only planet in our solar system that is neither too hot nor too cold, but "just right" for the development of life as we know it. Some of the basic properties of Earth are summarized in Table 8.1.



Figure 8.2 : Blue Marble. This image of Earth from space, taken by the Apollo 17 astronauts, is known as the "Blue Marble." This is one of the rare images of a full Earth taken during the Apollo program; most images show only part of Earth's disk in sunlight. (credit: modification of work by NASA)

Table 8.1: Some Properties of Earth

Property	Measurement
Semimajor axis	1.00 AU
Period	1.00 year
Mass	$5.98 \times 10^{24}\text{kg}$
Diameter	12,756 km
Radius	6378 km
Escape velocity	11.2 km/s
Rotational period	23 h 56 m 4 s
Surface area	$5.1 \times 10^{8}\,\text{km}^2$
Density	5.514 g/cm^3

Property	Measurement
Atmospheric pressure	1.00 bar

Earth's Interior

The interior of a planet—even our own Earth—is difficult to study, and its composition and structure must be determined indirectly. Our only direct experience is with the outermost skin of Earth's crust, a layer no more than a few kilometers deep. It is important to remember that, in many ways, we know less about our own planet 5 kilometers beneath our feet than we do about the surfaces of Venus and Mars.

Earth is composed largely of metal and silicate rock (see the Composition and Structure of Planets section). Most of this material is in a solid state, but some of it is hot enough to be molten. The structure of material in Earth's interior has been probed in considerable detail by measuring the transmission of seismic waves through Earth. These are waves that spread through the interior of Earth from earthquakes or explosion sites.

Seismic waves travel through a planet rather like sound waves through a struck bell. Just as the sound frequencies vary depending on the material the bell is made of and how it is constructed, so a planet's response depends on its composition and structure. By monitoring the seismic waves in different locations, scientists can learn about the layers through which the waves have traveled. Some of these vibrations travel along the surface; others pass directly through the interior. Seismic studies have shown that Earth's interior consists of several distinct layers with different compositions, illustrated in Figure 8.3. As waves travel through different materials in Earth's interior, the waves—just like light waves in telescope lenses—bend (or refract) so that some seismic stations on Earth receive the waves and others are in "shadows." Detecting the waves in a network of seismographs helps scientists construct a model of Earth's interior, showing liquid and solid layers. This type of seismic imaging is not unlike that used in ultrasound, a type of imaging used to see inside the body.

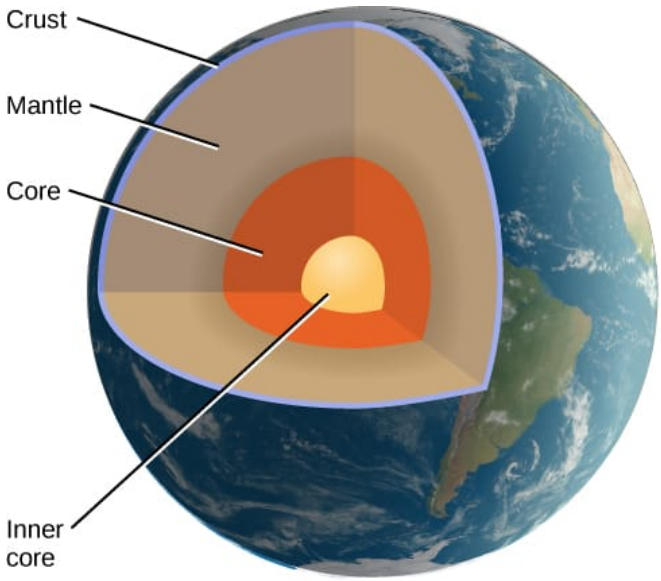


Figure 8.3 : Interior Structure of Earth. The crust, mantle, and inner and outer cores (solid and liquid, respectively) as shown as revealed by seismic studies.

The top layer is the crust, the part of Earth we know best (Figure 8.4). Oceanic crust covers 55% of Earth's surface and lies mostly submerged under the oceans. It is typically about 6 kilometers thick and is composed of volcanic rocks called basalt. Produced by the cooling of volcanic lava, basalts are made primarily of the elements silicon, oxygen, iron, aluminum, and magnesium. The continental crust covers 45% of the surface, some of which is also beneath the oceans. The continental crust is 20 to 70 kilometers thick and is composed predominantly of a different volcanic class of silicates (rocks made of silicon and oxygen) called granite. These crustal rocks, both oceanic and continental, typically have densities of about 3 g/cm^3. (For comparison, the density of water is 1 g/cm^3.) The crust is the easiest layer for geologists to study, but it makes up only about 0.3% of the total mass of Earth.

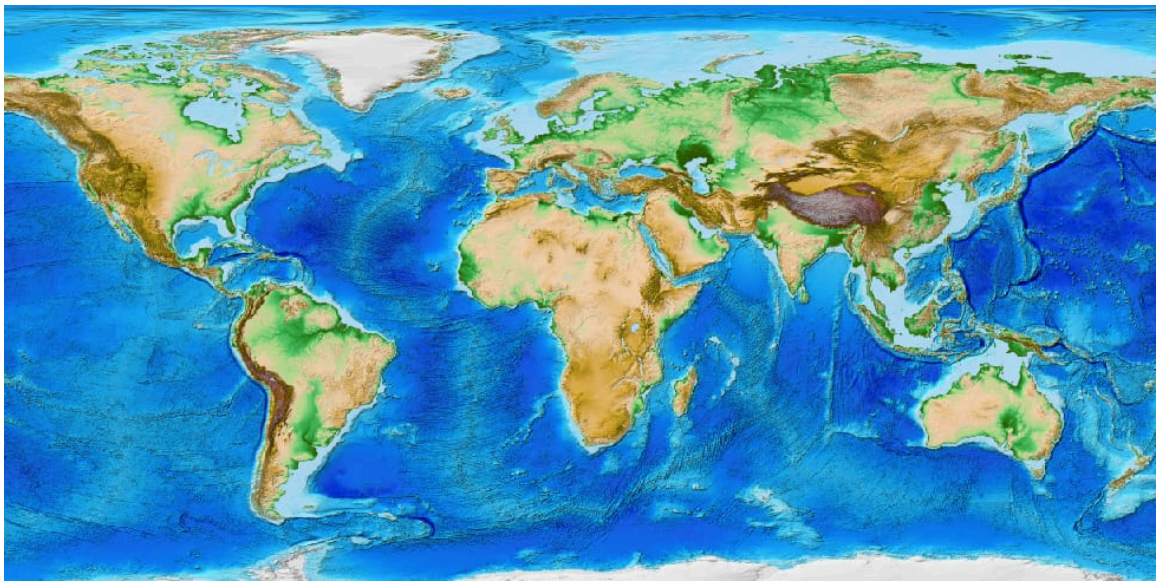


Figure 8.4 : Earth's Crust. This computer-generated image shows the surface of Earth's crust as determined from satellite images and ocean floor radar mapping. Oceans and lakes are shown in blue, with darker areas representing depth. Dry land is shown in shades of green and brown, and the Greenland and Antarctic ice sheets are depicted in shades of white. (credit: modification of work by C. Amante, B. W. Eakins, National Geophysical Data Center, NOAA)

The largest part of the solid Earth, called the mantle, stretches from the base of the crust downward to a depth of 2900 kilometers. The mantle is more or less solid, but at the temperatures and pressures found there, mantle rock can deform and flow slowly. The density in the mantle increases downward from about 3.5 g/cm^3 to more than 5 g/cm^3 as a result of the compression produced by the weight of overlying material. Samples of upper mantle material are occasionally ejected from volcanoes, permitting a detailed analysis of its chemistry.

Beginning at a depth of 2900 kilometers, we encounter the dense metallic core of Earth. With a diameter of 7000 kilometers, our core is substantially larger than the entire planet Mercury. The outer core is liquid, but the innermost part of the core (about 2400 kilometers in diameter) is probably solid. In addition to iron, the core probably also contains substantial quantities of nickel and sulfur, all compressed to a very high density.

The separation of Earth into layers of different densities is an example of *differentiation,* the process of sorting the major components of a planet by density. The fact that Earth is differentiated suggests that it was once warm enough for its interior to melt, permitting the heavier metals to sink to the center and form the dense core. Evidence for differentiation comes from comparing the planet's bulk density (5.5 g/cm^3) with the surface materials (3 g/cm^3) to suggest that denser material must be buried in the core.

Magnetic Field and Magnetosphere

We can find additional clues about Earth's interior from its magnetic field. Our planet behaves in some ways as if a giant bar magnet were inside it, aligned approximately with the rotational poles of Earth. This magnetic field is generated by moving material in Earth's liquid metallic core. As the liquid metal inside Earth circulates, it sets up a circulating electric current. When many charged particles are moving together like that—in the laboratory or on the scale of an entire planet—they produce a magnetic field.

Earth's magnetic field extends into surrounding space. When a charged particle encounters a magnetic field in space, it becomes trapped in the magnetic zone. Above Earth's atmosphere, our field is able to trap small quantities of electrons and other atomic particles. This region, called the magnetosphere, is defined as the zone within which Earth's magnetic field dominates over the weak interplanetary magnetic field that extends outward from the Sun (Figure 8.5).

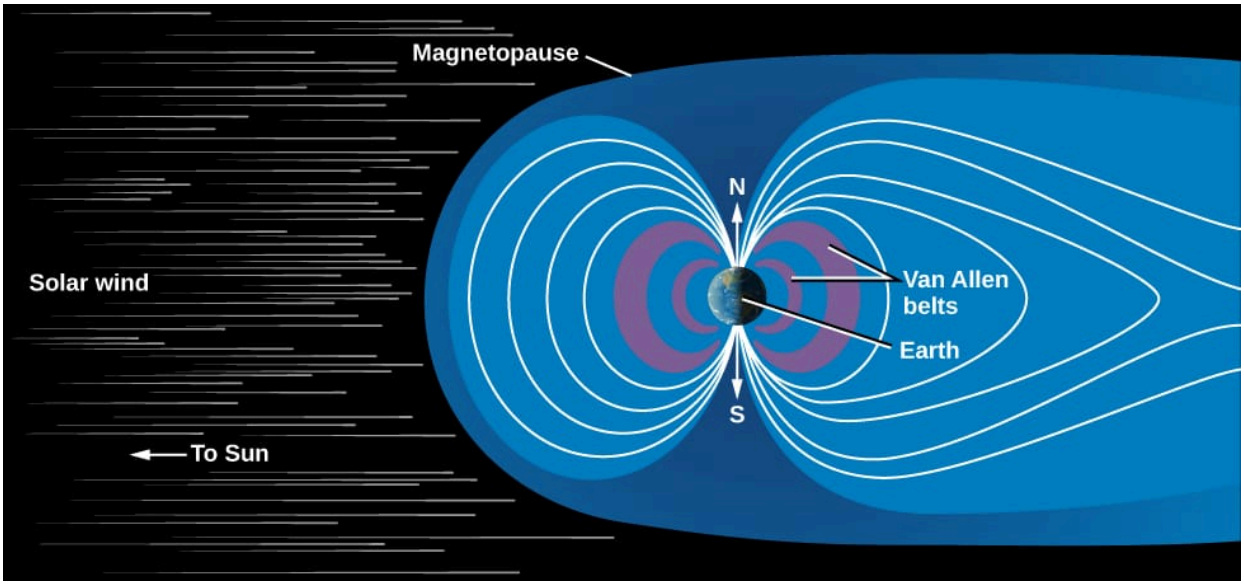


Figure 8.5 : Earth's Magnetosphere. A cross-sectional view of our magnetosphere (or zone of magnetic influence), as revealed by numerous spacecraft missions. Note how the wind of charged particles from the Sun "blows" the magnetic field outward like a wind sock.

Where do the charged particles trapped in our magnetosphere come from? They flow outward from the hot surface of the Sun; this is called the *solar wind*. It not only provides particles for Earth's magnetic field to trap, it also stretches our field in the direction pointing away from the Sun. Typically, Earth's magnetosphere extends about 60,000 kilometers, or 10 Earth radii, in the direction of the Sun. But, in the direction away from the Sun, the magnetic field can reach as far as the orbit of the Moon, and sometimes farther.

The magnetosphere was discovered in 1958 by instruments on the first US Earth satellite, *Explorer 1*, which recorded the ions (charged particles) trapped in its inner part. The regions of high-energy ions in the magnetosphere are often called the *Van Allen belts* in recognition of the University of Iowa professor who built the scientific instrumentation for *Explorer 1*. Since 1958, hundreds of spacecraft have explored various regions of the magnetosphere. You can read more about its interaction with the Sun in a later chapter.

This page titled 12.1: The Global Perspective is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 8.2: The Global Perspective by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

12.2: Earth's Crust

Learning Objectives

By the end of this section, you will be able to:

- Denote the primary types of rock that constitute Earth's crust
- Explain the theory of plate tectonics
- Describe the difference between rift and subduction zones
- Describe the relationship between fault zones and mountain building
- Explain the various types of volcanic activity occurring on Earth

Let us now examine our planet's outer layers in more detail. Earth's crust is a dynamic place. Volcanic eruptions, erosion, and large-scale movements of the continents rework the surface of our planet constantly. Geologically, ours is the most active planet. Many of the geological processes described in this section have taken place on other planets as well, but usually in their distant pasts. Some of the moons of the giant planets also have impressive activity levels. For example, Jupiter's moon Io has a remarkable number of active volcanoes.

Composition of the Crust

Earth's crust is largely made up of oceanic basalt and continental granite. These are both igneous rock, the term used for any rock that has cooled from a molten state. All volcanically produced rock is igneous (Figure 8.6).



Figure 8.6 : Formation of Igneous Rock as Liquid Lava Cools and Freezes. This is a lava flow from a basaltic eruption. Basaltic lava flows quickly and can move easily over distances of more than 20 kilometers. (credit: USGS)

Two other kinds of rock are familiar to us on Earth, although it turns out that neither is common on other planets. Sedimentary rocks are made of fragments of igneous rock or the shells of living organisms deposited by wind or water and cemented together without melting. On Earth, these rocks include the common sandstones, shales, and limestones. Metamorphic rocks are produced when high temperature or pressure alters igneous or sedimentary rock physically or chemically (the word *metamorphic* means "changed in form"). Metamorphic rocks are produced on Earth because geological activity carries surface rocks down to considerable depths and then brings them back up to the surface. Without such activity, these changed rocks would not exist at the surface.

There is a fourth very important category of rock that can tell us much about the early history of the planetary system: primitive

rock, which has largely escaped chemical modification by heating. Primitive rock represents the original material out of which the planetary system was made. No primitive material is left on Earth because the entire planet was heated early in its history. To find primitive rock, we must look to smaller objects such as comets, asteroids, and small planetary moons. We can sometimes see primitive rock in samples that fall to Earth from these smaller objects.

A block of quartzite on Earth is composed of materials that have gone through all four of these states. Beginning as primitive material before Earth was born, it was heated in the early Earth to form igneous rock, transformed chemically and redeposited (perhaps many times) to form sedimentary rock, and finally changed several kilometers below Earth's surface into the hard, white metamorphic stone we see today.

Plate Tectonics

Geology is the study of Earth's crust and the processes that have shaped its surface throughout history. (Although *geo-* means "related to Earth," astronomers and planetary scientists also talk about the geology of other planets.) Heat escaping from the interior provides energy for the formation of our planet's mountains, valleys, volcanoes, and even the continents and ocean basins themselves. But not until the middle of the twentieth century did geologists succeed in understanding just how these landforms are created.

Plate tectonics is an idea that explains how slow motions within the mantle of Earth move large segments of the crust, resulting in a gradual "drifting" of the continents as well as the formation of mountains and other large-scale geological features. Plate tectonics is a concept as basic to geology as evolution by natural selection is to biology or gravity is to understanding the orbits of planets. Looking at it from a different perspective, plate tectonics is a mechanism for Earth to transport heat efficiently from the interior, where it has accumulated, out to space. It is a cooling system for the planet. All planets develop a heat transfer process as they evolve; mechanisms may differ from that on Earth as a result of chemical makeup and other constraints.

Earth's crust and upper mantle (to a depth of about 60 kilometers) are divided into about a dozen tectonic plates that fit together like the pieces of a jigsaw puzzle (Figure 8.7). In some places, such as the Atlantic Ocean, the plates are moving apart; in others, such as off the western coast of South America, they are being forced together. The power to move the plates is provided by slow convection of the mantle, a process by which heat escapes from the interior through the upward flow of warmer material and the slow sinking of cooler material. (Convection, in which energy is transported from a warm region, such as the interior of Earth, to a cooler region, such as the upper mantle, is a process we encounter often in astronomy—in stars as well as planets. It is also important in boiling water for coffee while studying for astronomy exams.)

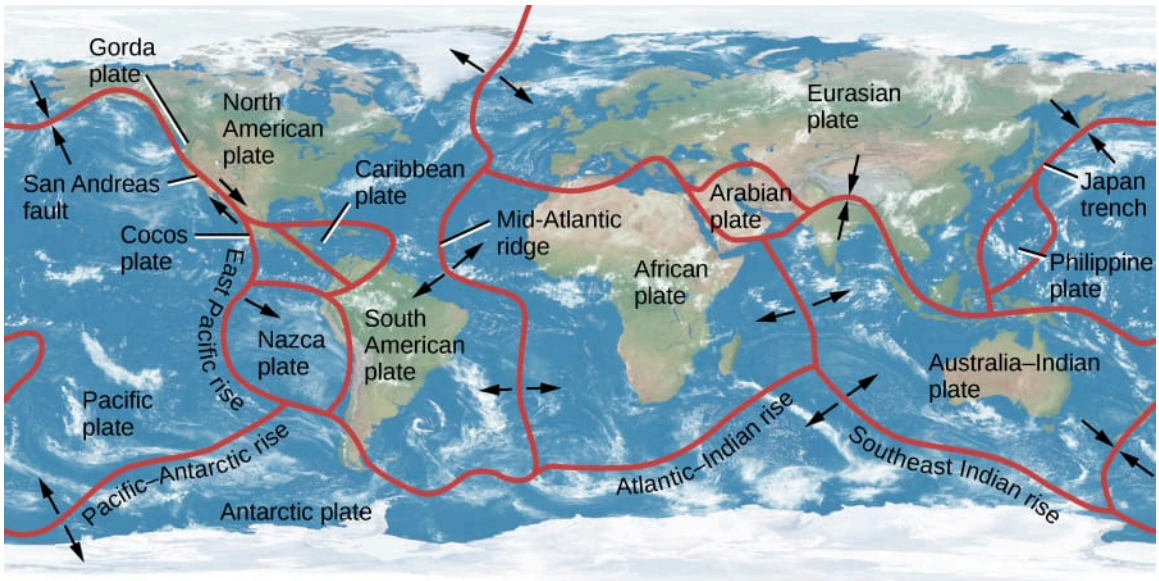


Figure 8.7 : Earth's Continental Plates. This map shows the major plates into which the crust of Earth is divided. Arrows indicate the motion of the plates at average speeds of 4 to 5 centimeters per year, similar to the rate at which your hair grows.

Link to Learning

The US Geological Survey provides a map of recent earthquakes and shows the boundaries of the tectonic plates and where earthquakes occur in relation to these boundaries. You can look close-up at the United States or zoom out for a global view.

As the plates slowly move, they bump into each other and cause dramatic changes in Earth's crust over time. Four basic kinds of interactions between crustal plates are possible at their boundaries: (1) they can pull apart, (2) one plate can burrow under another, (3) they can slide alongside each other, or (4) they can jam together. Each of these activities is important in determining the geology of Earth.

Voyagers in Astronomy: Alfred Wegener: Catching the Drift of Plate Tectonics

When studying maps or globes of Earth, many students notice that the coast of North and South America, with only minor adjustments, could fit pretty well against the coast of Europe and Africa. It seems as if these great landmasses could once have been together and then were somehow torn apart. The same idea had occurred to others (including Francis Bacon as early as 1620), but not until the twentieth century could such a proposal be more than speculation. The scientist who made the case for continental drift in 1920 was a German meteorologist and astronomer named Alfred Wegener (Figure 8.8).



Figure 8.8 : Alfred Wegener (1880–1930). Wegener proposed a scientific theory for the slow shifting of the continents.

Born in Berlin in 1880, Wegener was, from an early age, fascinated by Greenland, the world's largest island, which he dreamed of exploring. He studied at the universities in Heidelberg, Innsbruck, and Berlin, receiving a doctorate in astronomy by reexamining thirteenth-century astronomical tables. But, his interests turned more and more toward Earth, particularly its weather. He carried out experiments using kites and balloons, becoming so accomplished that he and his brother set a world record in 1906 by flying for 52 hours in a balloon.

Wegener first conceived of continental drift in 1910 while examining a world map in an atlas, but it took 2 years for him to assemble sufficient data to propose the idea in public. He published the results in book form in 1915. Wegener's evidence went far beyond the congruence in the shapes of the continents. He proposed that the similarities between fossils found only in South America and Africa indicated that these two continents were joined at one time. He also showed that resemblances among living animal species on different continents could best be explained by assuming that the continents were once connected in a supercontinent he called *Pangaea* (from Greek elements meaning "all land").

Wegener's suggestion was met with a hostile reaction from most scientists. Although he had marshaled an impressive list of arguments for his hypothesis, he was missing a *mechanism*. No one could explain *how* solid continents could drift over thousands of miles. A few scientists were sufficiently impressed by Wegener's work to continue searching for additional evidence, but many found the notion of moving continents too revolutionary to take seriously. Developing an understanding of the mechanism (plate tectonics) would take decades of further progress in geology, oceanography, and geophysics.

Wegener was disappointed in the reception of his suggestion, but he continued his research and, in 1924, he was appointed to a special meteorology and geophysics professorship created especially for him at the University of Graz (where he was, however, ostracized by most of the geology faculty). Four years later, on his fourth expedition to his beloved Greenland, he celebrated his fiftieth birthday with colleagues and then set off on foot toward a different camp on the island. He never made it; he was found a few days later, dead of an apparent heart attack.

Critics of science often point to the resistance to the continental drift hypothesis as an example of the flawed way that scientists regard new ideas. (Many people who have advanced crackpot theories have claimed that they are being ridiculed unjustly, just as Wegener was.) But we think there is a more positive light in which to view the story of Wegener's suggestion. Scientists in his day maintained a skeptical attitude because they needed more evidence and a clear mechanism that would fit what they understood about nature. Once the evidence and the mechanism were clear, Wegener's hypothesis quickly became the centerpiece of our view of a dynamic Earth.

Link to Learning

See how the drift of the continents has changed the appearance of our planet's crust.

Rift and Subduction Zones

Plates pull apart from each other along rift zones, such as the Mid-Atlantic ridge, driven by upwelling currents in the mantle (Figure 8.9). A few rift zones are found on land. The best known is the central African rift—an area where the African continent is slowly breaking apart. Most rift zones, however, are in the oceans. Molten rock rises from below to fill the space between the receding plates; this rock is basaltic lava, the kind of igneous rock that forms most of the ocean basins.

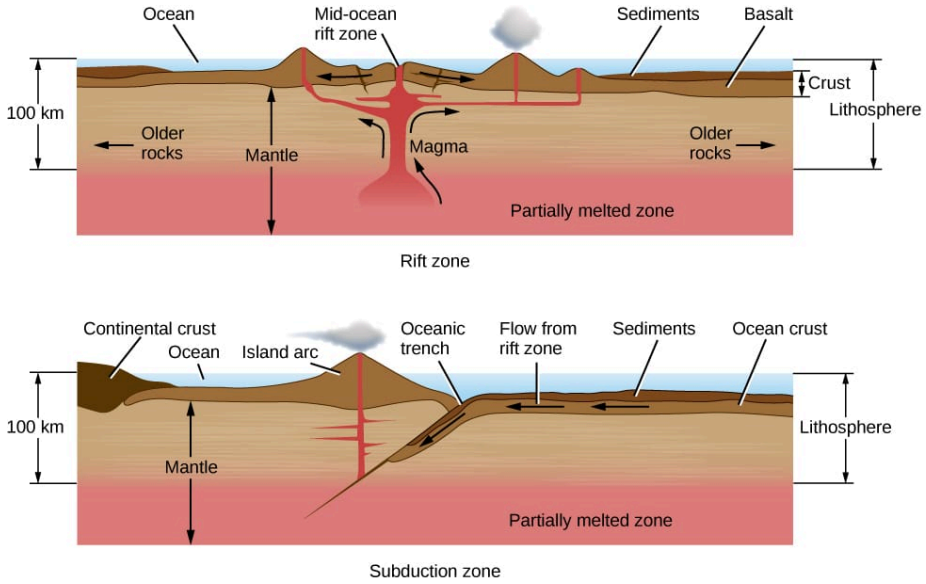


Figure 8.9 : Rift Zone and Subduction Zone. Rift and subduction zones are the regions (mostly beneath the oceans) where new crust is formed and old crust is destroyed as part of the cycle of plate tectonics.

From a knowledge of how the seafloor is spreading, we can calculate the average age of the oceanic crust. About 60,000 kilometers of active rifts have been identified, with average separation rates of about 5 centimeters per year. The new area added to Earth each year is about 2 square kilometers, enough to renew the entire oceanic crust in a little more than 100 million years. This is a very short interval in geological time—less than 3% of the age of Earth. The present ocean basins thus turn out to be

among the youngest features on our planet.

As new crust is added to Earth, the old crust must go somewhere. When two plates come together, one plate is often forced beneath another in what is called a subduction zone (Figure 8.9). In general, the thick continental masses cannot be subducted, but the thinner oceanic plates can be rather readily thrust down into the upper mantle. Often a subduction zone is marked by an ocean trench; a fine example of this type of feature is the deep Japan trench along the coast of Asia. The subducted plate is forced down into regions of high pressure and temperature, eventually melting several hundred kilometers below the surface. Its material is recycled into a downward-flowing convection current, ultimately balancing the flow of material that rises along rift zones. The amount of crust destroyed at subduction zones is approximately equal to the amount formed at rift zones.

All along the subduction zone, earthquakes and volcanoes mark the death throes of the plate. Some of the most destructive earthquakes in history have taken place along subduction zones, including the 1923 Yokohama earthquake and fire that killed 100,000 people, the 2004 Sumatra earthquake and tsunami that killed more than 200,000 people, and the 2011 Tohoku earthquake that resulted in the meltdown of three nuclear power reactors in Japan.

Fault Zones and Mountain Building

Along much of their length, the crustal plates slide parallel to each other. These plate boundaries are marked by cracks or faults. Along active fault zones, the motion of one plate with respect to the other is several centimeters per year, about the same as the spreading rates along rifts.

One of the most famous faults is the San Andreas Fault in California, which lies at the boundary between the Pacific plate and the North American plate (Figure 8.10). This fault runs from the Gulf of California to the Pacific Ocean northwest of San Francisco. The Pacific plate, to the west, is moving northward, carrying Los Angeles, San Diego, and parts of the southern California coast with it. In several million years, Los Angeles may be an island off the coast of San Francisco.



Figure 8.10 : San Andreas Fault. We see part of a very active region in California where one crustal plate is sliding sideways with respect to the other. The fault is marked by the valley running up the right side of the photo. Major slippages along this fault can produce extremely destructive earthquakes. (credit: John Wiley)

Unfortunately for us, the motion along fault zones does not take place smoothly. The creeping motion of the plates against each other builds up stresses in the crust that are released in sudden, violent slippages that generate earthquakes. Because the average motion of the plates is constant, the longer the interval between earthquakes, the greater the stress and the more energy released when the surface finally moves.

For example, the part of the San Andreas Fault near the central California town of Parkfield has slipped every 25 years or so during the past century, moving an average of about 1 meter each time. In contrast, the average interval between major earthquakes in the Los Angeles region is about 150 years, and the average motion is about 7 meters. The last time the San Andreas fault slipped in this area was in 1857; tension has been building ever since, and sometime soon it is bound to be released. Sensitive

instruments placed within the Los Angeles basin show that the basin is distorting and contracting in size as these tremendous pressures build up beneath the surface.

Example 8.1: Fault Zones and Plate Motion

After scientists mapped the boundaries between tectonic plates in Earth's crust and measured the annual rate at which the plates move (which is about 5 cm/year), we could estimate quite a lot about the rate at which the geology of Earth is changing. As an example, let's suppose that the next slippage along the San Andreas Fault in southern California takes place in the year 2027 and that it completely relieves the accumulated strain in this region. How much slippage is required for this to occur?

Solution

The speed of motion of the Pacific plate relative to the North American plate is 5 cm/y. That's 500 cm (or 5 m) per century. The last southern California earthquake was in 1857. The time from 1857 to 2027 is
170 y, or 1.7 centuries, so the slippage to relieve the strain completely would be
$5\text{m/century} \times 1.7\text{centuries} = 8.5\text{m}.$

Exercise 12.2.1

If the next major southern California earthquake occurs in 2047 and only relieves one-half of the accumulated strain, how much slippage will occur?

Answer available at the end of this section.

When two continental masses are moving on a collision course, they push against each other under great pressure. Earth buckles and folds, dragging some rock deep below the surface and raising other folds to heights of many kilometers. This is the way many, but not all, of the mountain ranges on Earth were formed. The Alps, for example, are a result of the African plate bumping into the Eurasian plate. As we will see, however, quite different processes produced the mountains on other planets.

Once a mountain range is formed by upthrusting of the crust, its rocks are subject to erosion by water and ice. The sharp peaks and serrated edges have little to do with the forces that make the mountains initially. Instead, they result from the processes that tear down mountains. Ice is an especially effective sculptor of rock (Figure 8.11). In a world without moving ice or running water (such as the Moon or Mercury), mountains remain smooth and dull.



Figure 8.11 : Mountains on Earth. The Torres del Paine are a young region of Earth's crust where sharp mountain peaks are being sculpted by glaciers. We owe the beauty of our young, steep mountains to the erosion by ice and water. (credit: David Morrison)

Volcanoes

Volcanoes mark locations where lava rises to the surface. One example is mid ocean ridges, which are long undersea mountain ranges formed by lava rising from Earth's mantle at plate boundaries. A second major kind of volcanic activity is associated with subduction zones, and volcanoes sometimes also appear in regions where continental plates are colliding. In each case, the volcanic activity gives us a way to sample some of the material from deeper within our planet.

Other volcanic activity occurs above mantle "hot spots"—areas far from plate boundaries where heat is nevertheless rising from the interior of Earth. One of the best-known hot spots is under the island of Hawaii, where it currently supplies the heat to maintain three active volcanoes, two on land and one under the ocean. The Hawaii hot spot has been active for at least 100 million years. As Earth's plates have moved during that time, the hot spot has generated a 3500-kilometer-long chain of volcanic islands. The tallest Hawaiian volcanoes are among the largest individual mountains on Earth, more than 100 kilometers in diameter and rising 9 kilometers above the ocean floor. One of the Hawaiian volcanic mountains, the now-dormant Maunakea, has become one of the world's great sites for doing astronomy.

Link to Learning

The US Geological Service provides an interactive map of the famous "ring of fire," which is the chain of volcanoes surrounding the Pacific Ocean, and shows the Hawaiian "hot spot" enclosed within.

Not all volcanic eruptions produce mountains. If lava flows rapidly from long cracks, it can spread out to form lava plains. The largest known terrestrial eruptions, such as those that produced the Snake River basalts in the northwestern United States or the Deccan plateau in India, are of this type. Similar lava plains are found on the Moon and the other terrestrial planets.

Answers

Exercise 12.2.1

The difference in time from 1857 to 2047 is 190 y, or 1.9 centuries. Because only half the strain is released, this is equivalent to half the annual rate of motion. The total slippage comes to
$0.5 \times 5\text{m/century} \times 1.9\text{centuries} = 4.75\text{m}.$

This page titled 12.2: Earth's Crust is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 8.3: Earth's Crust by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

12.3: Earth's Atmosphere

Learning Objectives

By the end of this section, you will be able to:

- Differentiate between Earth's various atmospheric layers
- Describe the chemical composition and possible origins of our atmosphere
- Explain the difference between weather and climate

We live at the bottom of the ocean of air that envelops our planet. The atmosphere, weighing down upon Earth's surface under the force of gravity, exerts a pressure at sea level that scientists define as 1 bar (a term that comes from the same root as *barometer*, an instrument used to measure atmospheric pressure). A bar of pressure means that each square centimeter of Earth's surface has a weight equivalent to 1.03 kilograms pressing down on it. Humans have evolved to live at this pressure; make the pressure a lot lower or higher and we do not function well.

The total mass of Earth's atmosphere is about 5×10^{18} kilograms. This sounds like a large number, but it is only about a millionth of the total mass of Earth. The atmosphere represents a smaller fraction of Earth than the fraction of your mass represented by the hair on your head.

Structure of the Atmosphere

The structure of the atmosphere is illustrated in Figure 8.12. Most of the atmosphere is concentrated near the surface of Earth, within about the bottom 10 kilometers where clouds form and airplanes fly. Within this region—called the troposphere—warm air, heated by the surface, rises and is replaced by descending currents of cooler air; this is an example of convection. This circulation generates clouds and wind. Within the troposphere, temperature decreases rapidly with increasing elevation to values near 50 °C below freezing at its upper boundary, where the stratosphere begins. Most of the stratosphere, which extends to about 50 kilometers above the surface, is cold and free of clouds.

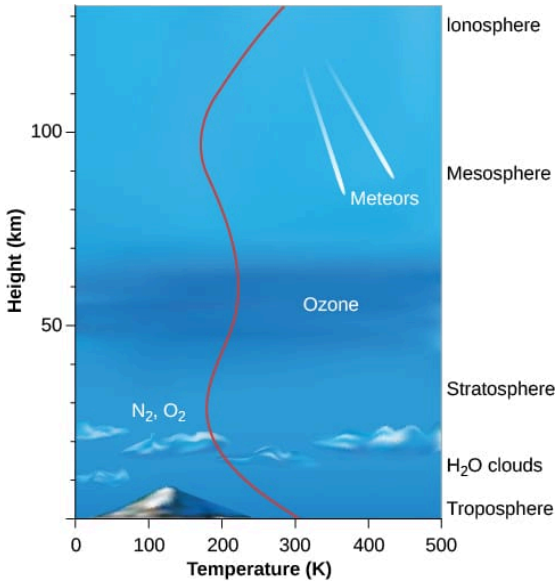


Figure 8.12 : Structure of Earth's Atmosphere. Height increases up the left side of the diagram, and the names of the different atmospheric layers are shown at the right. In the upper ionosphere, ultraviolet radiation from the Sun can strip electrons from their atoms, leaving the atmosphere ionized. The curving red line shows the temperature (see the scale on the *x*-axis).

Near the top of the stratosphere is a layer of ozone (O_3), a form of oxygen with three atoms per molecule instead of the usual two. Because ozone is a good absorber of ultraviolet light, it protects the surface from some of the Sun's dangerous

ultraviolet radiation, making it possible for life to exist on Earth. The breakup of ozone adds heat to the stratosphere, reversing the decreasing temperature trend in the troposphere. Because ozone is essential to our survival, we reacted with justifiable concern to evidence that became clear in the 1980s that atmospheric ozone was being destroyed by human activities. By international agreement, the production of industrial chemicals that cause ozone depletion, called chlorofluorocarbons, or CFCs, has been phased out. As a result, ozone loss has stopped and the "ozone hole" over the Antarctic is shrinking gradually. This is an example of how concerted international action can help maintain the habitability of Earth.

Link to Learning

Visit NASA's scientific visualization studio for a short video of what would have happened to Earth's ozone layer by 2065 if CFCs had not been regulated.

At heights above 100 kilometers, the atmosphere is so thin that orbiting satellites can pass through it with very little friction. Many of the atoms are ionized by the loss of an electron, and this region is often called the ionosphere. At these elevations, individual atoms can occasionally escape completely from the gravitational field of Earth. There is a continuous, slow leaking of atmosphere—especially of lightweight atoms, which move faster than heavy ones. Earth's atmosphere cannot, for example, hold on for long to hydrogen or helium, which escape into space. Earth is not the only planet to experience atmosphere leakage. Atmospheric leakage also created Mars' thin atmosphere. Venus' dry atmosphere evolved because its proximity to the Sun vaporized and dissociated any water, with the component gases lost to space.

Atmospheric Composition and Origin

At Earth's surface, the atmosphere consists of 78% nitrogen (N_2), 21% oxygen (O_2), and 1% argon (Ar), with traces of water vapor (H_2O), carbon dioxide (CO_2), and other gases. Variable amounts of dust particles and water droplets are also found suspended in the air.

A complete census of Earth's volatile materials, however, should look at more than the gas that is now present. *Volatile* materials are those that evaporate at a relatively low temperature. If Earth were just a little bit warmer, some materials that are now liquid or solid might become part of the atmosphere. Suppose, for example, that our planet were heated to above the boiling point of water (100 °C, or 373 K); that's a large change for humans, but a small change compared to the range of possible temperatures in the universe. At 100 °C, the oceans would boil and the resulting water vapor would become a part of the atmosphere.

To estimate how much water vapor would be released, note that there is enough water to cover the entire Earth to a depth of about 300 meters. Because the pressure exerted by 10 meters of water is equal to about 1 bar, the average pressure at the ocean floor is about 300 bars. Water weighs the same whether in liquid or vapor form, so if the oceans boiled away, the atmospheric pressure of the water would still be 300 bars. Water would therefore greatly dominate Earth's atmosphere, with nitrogen and oxygen reduced to the status of trace constituents.

On a warmer Earth, another source of additional atmosphere would be found in the sedimentary carbonate rocks of the crust. These minerals contain abundant carbon dioxide. If all these rocks were heated, they would release about 70 bars of CO_2, far more than the current CO_2 pressure of only 0.0005 bar. Thus, the atmosphere of a warm Earth would be dominated by water vapor and carbon dioxide, with a surface pressure nearing 400 bars.

Several lines of evidence show that the composition of Earth's atmosphere has changed over our planet's history. Scientists can infer the amount of atmospheric oxygen, for example, by studying the chemistry of minerals that formed at various times. We examine this issue in more detail later in this chapter.

Today we see that CO_2, H_2O, sulfur dioxide (SO_2), and other gases are released from deeper within Earth through the action of volcanoes. (For CO_2, the primary source today is the burning of fossil fuels, which releases far more CO_2 than that from volcanic eruptions.) Much of this apparently new gas, however, is recycled material that has been subducted through plate tectonics. But where did our planet's original atmosphere come from?

Three possibilities exist for the original source of Earth's atmosphere and oceans: (1) the atmosphere could have been formed with the rest of Earth as it accumulated from debris left over from the formation of the Sun; (2) it could have been released from the interior through volcanic activity, subsequent to the formation of Earth; or (3) it may have been derived from impacts by comets and asteroids from the outer parts of the solar system. Current evidence favors a combination of the interior and impact

sources.

Weather and Climate

All planets with atmospheres have *weather*, which is the name we give to the circulation of the atmosphere. The energy that powers the weather is derived primarily from the sunlight that heats the surface. Both the rotation of the planet and slower seasonal changes cause variations in the amount of sunlight striking different parts of Earth. The atmosphere and oceans redistribute the heat from warmer to cooler areas. Weather on any planet represents the response of its atmosphere to changing inputs of energy from the Sun (see Figure 8.13 for a dramatic example).



Figure 8.13 : Storm from Space. This satellite image shows Hurricane Irene in 2011, shortly before the storm hit land in New York City. The combination of Earth's tilted axis of rotation, moderately rapid rotation, and oceans of liquid water can lead to violent weather on our planet. (credit: NASA/NOAA GOES Project)

Climate is a term used to refer to the effects of the atmosphere that last through decades and centuries. Changes in climate (as opposed to the random variations in weather from one year to the next) are often difficult to detect over short time periods, but as they accumulate, their effect can be devastating. One saying is that "Climate is what you expect, and weather is what you get." Modern farming is especially sensitive to temperature and rainfall; for example, calculations indicate that a drop of only 2 °C throughout the growing season would cut the wheat production by half in Canada and the United States. At the other extreme, an increase of 2 °C in the average temperature of Earth would be enough to melt many glaciers, including much of the ice cover of Greenland, raising sea level by as much as 10 meters, flooding many coastal cities and ports, and putting small islands completely under water.

The best documented changes in Earth's climate are the great ice ages, which have lowered the temperature of the Northern Hemisphere periodically over the past half million years or so (Figure 8.14). The last ice age, which ended about 14,000 years ago, lasted some 20,000 years. At its height, the ice was almost 2 kilometers thick over Boston and stretched as far south as New York City.

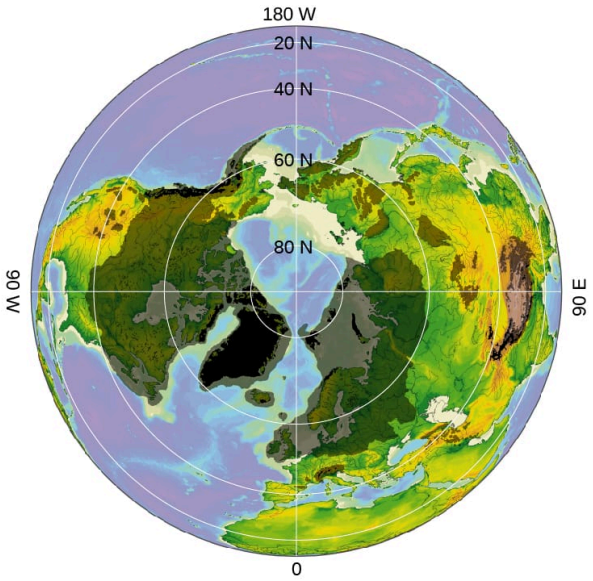


Figure 8.14 : Ice Age. This computer-generated image shows the frozen areas of the Northern Hemisphere during past ice ages from the vantage point of looking down on the North Pole. The area in black indicates the most recent glaciation (coverage by glaciers), and the area in gray shows the maximum level of glaciation ever reached. (credit: modification of work by Hannes Grobe/AWI)

These ice ages were primarily the result of changes in the tilt of Earth's rotational axis, produced by the gravitational effects of the other planets. We are less certain about evidence that at least once (and perhaps twice) about a billion years ago, the entire ocean froze over, a situation called *snowball Earth*.

The development and evolution of life on Earth has also produced changes in the composition and temperature of our planet's atmosphere, as we shall see in the next section.

This page titled 12.3: Earth's Atmosphere is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 8.4: Earth's Atmosphere by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

12.4: Life, Chemical Evolution, and Climate Change

Learning Objectives

By the end of this section, you will be able to:

- Outline the origins and subsequent diversity of life on Earth
- Explain the ways that life and geological activity have influenced the evolution of the atmosphere
- Describe the causes and effects of the atmospheric greenhouse effect and global warming
- Describe the impact of human activity on our planet's atmosphere and ecology

As far as we know, Earth seems to be the only planet in the solar system with life. The origin and development of life are an important part of our planet's story. Life arose early in Earth's history, and living organisms have been interacting with their environment for billions of years. We recognize that life-forms have evolved to adapt to the environment on Earth, and we are now beginning to realize that Earth itself has been changed in important ways by the presence of living matter. The study of the coevolution of life and our planet is one of the subjects of the modern science of *astrobiology*.

The Origin of Life

The record of the birth of life on Earth has been lost in the restless motions of the crust. According to chemical evidence, by the time the oldest surviving rocks were formed about 3.9 billion years ago, life already existed. At 3.5 billion years ago, life had achieved the sophistication to build large colonies called *stromatolites*, a form so successful that stromatolites still grow on Earth today (Figure 8.15). But, few rocks survive from these ancient times, and abundant fossils have been preserved only during the past 600 million years—less than 15% of our planet's history.



Figure 8.15 : Cross-Sections of Fossil Stromatolites. This polished cross-section of a fossilized colony of stromatolites dates to the Precambrian Era. The layered, domelike structures are mats of sediment trapped in shallow waters by large numbers of blue-green bacteria that can photosynthesize. Such colonies of microorganisms date back more than 3 billion years. (credit: James St. John)

There is little direct evidence about the actual origin of life. We know that the atmosphere of early Earth, unlike today's, contained abundant carbon dioxide and some methane, but no oxygen gas. In the absence of oxygen, many complex chemical reactions are possible that lead to the production of amino acids, proteins, and other chemical building blocks of life. Therefore, it seems likely that these chemical building blocks were available very early in Earth's history and they would have combined to make living organisms.

For tens of millions of years after Earth's formation, life (perhaps little more than large molecules, like the viruses of today) probably existed in warm, nutrient-rich seas, living off accumulated organic chemicals. When this easily accessible food became depleted, life began the long evolutionary road that led to the vast numbers of different organisms on Earth today. As it did so, life began to influence the chemical composition of the atmosphere.

In addition to the study of life's history as revealed by chemical and fossil evidence in ancient rocks, scientists use tools from the rapidly advancing fields of genetics and *genomics*—the study of the genetic code that is shared by all life on Earth. While each individual has a unique set of genes (which is why genetic "fingerprinting" is so useful for the study of crime), we also have many genetic traits in common. Your *genome*, the complete map of the DNA in your body, is identical at the 99.9% level to that of Julius Caesar or Marie Curie. At the 99% level, human and chimpanzee genomes are the same. By looking at the gene sequences of many organisms, we can determine that all life on Earth is descended from a common ancestor, and we can use the genetic variations among species as a measure of how closely different species are related.

These genetic analysis tools have allowed scientists to construct what is called the "tree of life" (Figure 8.16). This diagram illustrates the way organisms are related by examining one sequence of the nucleic acid RNA that all species have in common. This figure shows that life on Earth is dominated by microscopic creatures that you have probably never heard of. Note that the plant and animal kingdoms are just two little branches at the far right. Most of the diversity of life, and most of our evolution, has taken place at the microbial level. Indeed, it may surprise you to know that there are more microbes in a bucket of soil than there are stars in the Galaxy. You may want to keep this in mind when, later in this book, we turn to the search for life on other worlds. The "aliens" that are most likely to be out there are microbes.

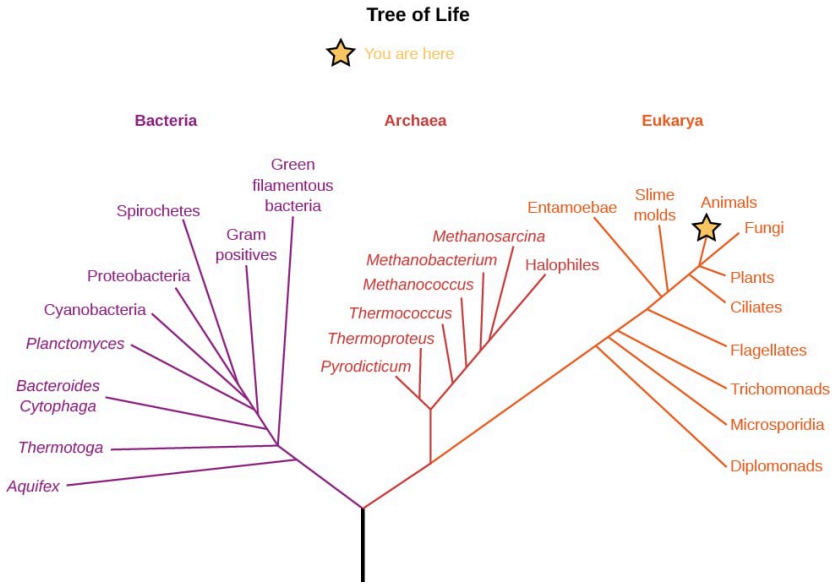


Figure 8.16 : Tree of Life. This chart shows the main subdivisions of life on Earth and how they are related. Note that the animal and plant kingdoms are just short branches on the far right, along with the fungi. The most fundamental division of Earth's living things is onto three large domains called bacteria, archaea, and eukarya. Most of the species listed are microscopic. (credit: modification of work by Eric Gaba)

Such genetic studies lead to other interesting conclusions as well. For example, it appears that the earliest surviving terrestrial life-forms were all adapted to live at high temperatures. Some biologists think that life might actually have begun in locations on our planet that were extremely hot. Yet another intriguing possibility is that life began on Mars (which cooled sooner) rather than Earth and was "seeded" onto our planet by meteorites traveling from Mars to Earth. Mars rocks are still making their way to Earth, but so far none has shown evidence of serving as a "spaceship" to carry microorganisms from Mars to Earth.

The Evolution of the Atmosphere

One of the key steps in the evolution of life on Earth was the development of blue-green algae, a very successful life-form that takes in carbon dioxide from the environment and releases oxygen as a waste product. These successful microorganisms proliferated, giving rise to all the lifeforms we call plants. Since the energy for making new plant material from chemical building blocks comes from sunlight, we call the process photosynthesis.

Studies of the chemistry of ancient rocks show that Earth's atmosphere lacked abundant free oxygen until about 2 billion years

ago, despite the presence of plants releasing oxygen by photosynthesis. Apparently, chemical reactions with Earth's crust removed the oxygen gas as quickly as it formed. Slowly, however, the increasing evolutionary sophistication of life led to a growth in the plant population and thus increased oxygen production. At the same time, it appears that increased geological activity led to heavy erosion on our planet's surface. This buried much of the plant carbon before it could recombine with oxygen to form CO_2.

Free oxygen began accumulating in the atmosphere about 2 billion years ago, and the increased amount of this gas led to the formation of Earth's ozone layer (recall that ozone is a triple molecule of oxygen, O_3), which protects the surface from deadly solar ultraviolet light. Before that, it was unthinkable for life to venture outside the protective oceans, so the landmasses of Earth were barren.

The presence of oxygen, and hence ozone, thus allowed colonization of the land. It also made possible a tremendous proliferation of animals, which lived by taking in and using the organic materials produced by plants as their own energy source.

As animals evolved in an environment increasingly rich in oxygen, they were able to develop techniques for breathing oxygen directly from the atmosphere. We humans take it for granted that plenty of free oxygen is available in Earth's atmosphere, and we use it to release energy from the food we take in. Although it may seem funny to think of it this way, we are lifeforms that have evolved to breathe in the waste product of plants. It is plants and related microbes that are the primary producers, using sunlight to create energy-rich "food" for the rest of us.

On a planetary scale, one of the consequences of life has been a decrease in atmospheric carbon dioxide. In the absence of life, Earth would probably have an atmosphere dominated by CO_2, like Mars or Venus. But living things, in combination with high levels of geological activity, have effectively stripped our atmosphere of most of this gas.

The Greenhouse Effect and Global Warming

We have a special interest in the carbon dioxide content of the atmosphere because of the key role this gas plays in retaining heat from the Sun through a process called the greenhouse effect. To understand how the greenhouse effect works, consider the fate of sunlight that strikes the surface of Earth. The light penetrates our atmosphere, is absorbed by the ground, and heats the surface layers. At the temperature of Earth's surface, that energy is then reemitted as infrared or heat radiation (Figure 8.17). However, the molecules of our atmosphere, which allow visible light through, are good at absorbing infrared energy. As a result, CO_2 (along with methane and water vapor) acts like a blanket, trapping heat in the atmosphere and impeding its flow back to space. To maintain an energy balance, the temperature of the surface and lower atmosphere must increase until the total energy radiated by Earth to space equals the energy received from the Sun. The more CO_2 there is in our atmosphere, the higher the temperature at which Earth's surface reaches a new balance.

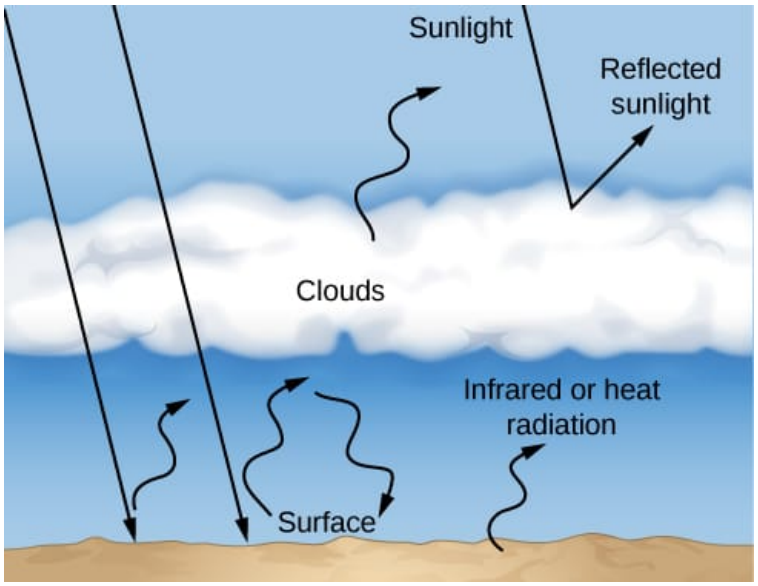


Figure 8.17 : How the Greenhouse Effect Works. Sunlight that penetrates to Earth's lower atmosphere and surface is reradiated as infrared or heat radiation, which is trapped by greenhouse gases such as water vapor, methane, and CO_2 in the atmosphere. The result is a higher surface temperature for our planet.

The greenhouse effect in a planetary atmosphere is similar to the heating of a gardener's greenhouse or the inside of a car left out in the Sun with the windows rolled up. In these examples, the window glass plays the role of greenhouse gases, letting sunlight in but reducing the outward flow of heat radiation. As a result, a greenhouse or car interior winds up much hotter than would be expected from the heating of sunlight alone. On Earth, the current greenhouse effect elevates the surface temperature by about 23 °C. Without this greenhouse effect, the average surface temperature would be well below freezing and Earth would be locked in a global ice age.

That's the good news; the bad news is that the heating due to the greenhouse effect is increasing. Modern industrial society depends on energy extracted from burning fossil fuels. In effect, we are exploiting the energy-rich material created by photosynthesis tens of millions of years ago. As these ancient coal and oil deposits are oxidized (burned using oxygen), large quantities of carbon dioxide are released into the atmosphere. The problem is exacerbated by the widespread destruction of tropical forests, which we depend on to extract CO_2 from the atmosphere and replenish our supply of oxygen. In the past century of increased industrial and agricultural development, the amount of CO_2 in the atmosphere increased by about 30% and continues to rise at more than 0.5% per year.

Before the end of the present century, Earth's CO_2 level is predicted to reach twice the value it had before the industrial revolution (Figure 8.18). The consequences of such an increase for Earth's surface and atmosphere (and the creatures who live there) are likely to be complex changes in climate, and may be catastrophic for many species. Many groups of scientists are now studying the effects of such global warming with elaborate computer models, and climate change has emerged as the greatest known threat (barring nuclear war) to both industrial civilization and the ecology of our planet.

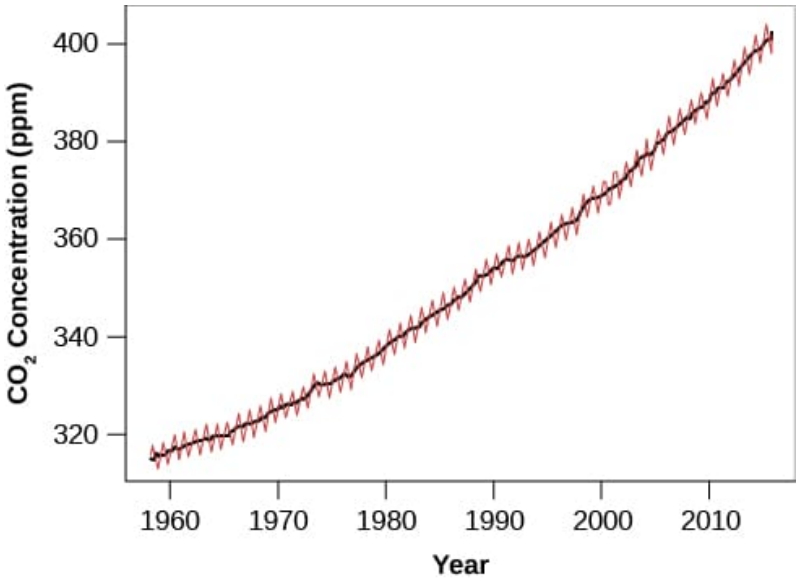


Figure 8.18 : Increase of Atmospheric Carbon Dioxide over Time. Scientists expect that the amount of CO_2 will double its preindustrial level before the end of the twenty-first century. Measurements of the isotopic signatures of this added CO_2 demonstrate that it is mostly coming from burning fossil fuels. (credit: modification of work by NOAA)

Link to Learning

This short PBS video explains the physics of the greenhouse effect.

Already climate change is widely apparent. Around the world, temperature records are constantly set and broken; all but one of the hottest recorded years have taken place since 2000. Glaciers are retreating, and the Arctic Sea ice is now much thinner than when it was first explored with nuclear submarines in the 1950s. Rising sea levels (from both melting glaciers and expansion of the water as its temperature rises) pose one of the most immediate threats, and many coastal cities have plans to build dikes or seawalls to hold back the expected flooding. The rate of temperature increase is without historical precedent, and we are rapidly entering "unknown territory" where human activities are leading to the highest temperatures on Earth in more than 50 million years.

Human Impacts on Our Planet

Earth is so large and has been here for so long that some people have trouble accepting that humans are really changing the planet, its atmosphere, and its climate. They are surprised to learn, for example, that the carbon dioxide released from burning fossil fuels is 100 times greater than that emitted by volcanoes. But, the data clearly tell the story that our climate is changing rapidly, and that almost all of the change is a result of human activity.

This is not the first time that humans have altered our environment dramatically. Some of the greatest changes were caused by our ancestors, before the development of modern industrial society. If aliens had visited Earth 50,000 years ago, they would have seen much of the planet supporting large animals of the sort that now survive only in Africa. The plains of Australia were occupied by giant marsupials such as diprotodon and zygomaturus (the size of our elephants today), and a species of kangaroo that stood 10 feet high. North America and North Asia hosted mammoths, saber tooth cats, mastodons, giant sloths, and even camels. The Islands of the Pacific teemed with large birds, and vast forests covered what are now the farms of Europe and China. Early human hunters killed many large mammals and marsupials, early farmers cut down most of the forests, and the Polynesian expansion across the Pacific doomed the population of large birds.

An even greater mass extinction is underway as a result of rapid climate change. In recognition of our impact on the environment, scientists have proposed giving a new name to the current epoch, the *anthropocene*, when human activity started to have a significant global impact. Although not an officially approved name, the concept of “anthropocene” is useful for recognizing that we humans now represent the dominant influence on our planet’s atmosphere and ecology, for better or for worse.

This page titled 12.4: Life, Chemical Evolution, and Climate Change is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 8.5: Life, Chemical Evolution, and Climate Change by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

12.5: Cosmic Influences on the Evolution of Earth

Learning Objectives

By the end of this section, you will be able to:

- Explain the scarcity of impact craters on Earth compared with other planets and moons
- Describe the evidence for recent impacts on Earth
- Detail how a massive impact changed the conditions for life on Earth, leading to the extinction of the dinosaurs
- Describe how impacts have influenced the evolution of life on Earth
- Discuss the search for objects that could potentially collide with our planet

In discussing Earth's geology earlier in this chapter, we dealt only with the effects of internal forces, expressed through the processes of plate tectonics and volcanism. On the Moon, in contrast, we see primarily craters, produced by the impacts of interplanetary debris such as asteroids and comets. Why don't we see more evidence on Earth of the kinds of impact craters that are so prominent on the Moon and other worlds?

Where Are the Craters on Earth?

It is not possible that Earth escaped being struck by the interplanetary debris that has pockmarked the Moon. From a cosmic perspective, the Moon is almost next door. Our atmosphere does make small pieces of cosmic debris burn up (which we see as *meteors*—commonly called shooting stars). But, the layers of our air provide no shield against the large impacts that form craters several kilometers in diameter and are common on the Moon.

In the course of its history, Earth must therefore have been impacted as heavily as the Moon. The difference is that, on Earth, these craters are destroyed by our active geology before they can accumulate. As plate tectonics constantly renews our crust, evidence of past cratering events is slowly erased. Only in the past few decades have geologists succeeded in identifying the eroded remnants of many impact craters (Figure 8.19). Even more recent is our realization that, over the history of Earth, these impacts have had an important influence on the evolution of life.

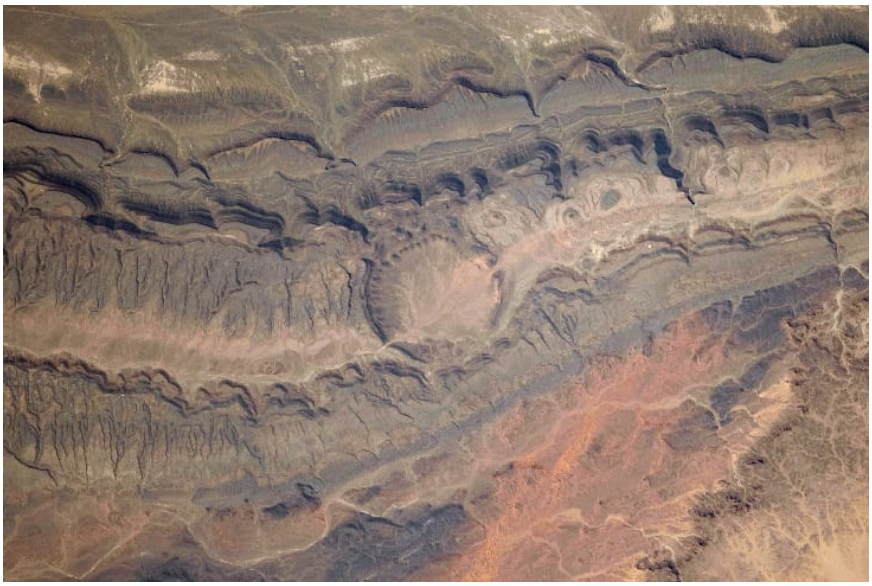


Figure 8.19 : Ouarkziz Impact Crater. Located in Algeria, this 4-km-wide crater (the round feature in the center) is the result of a meteor impact during the Cretaceous period. Although the crater has experienced heavy erosion, this image from the International Space Station shows the circular pattern resulting from impact. (credit: modification of work by NASA)

Recent Impacts

The collision of interplanetary debris with Earth is not a hypothetical idea. Evidence of relatively recent impacts can be found on our planet's surface. One well-studied historic collision took place on June 30, 1908, near the Tunguska River in Siberia. In this desolate region, there was a remarkable explosion in the atmosphere about 8 kilometers above the surface. The shock wave flattened more than a thousand square kilometers of forest (Figure 8.20). Herds of reindeer and other animals were killed, and a man at a trading post 80 kilometers from the blast was thrown from his chair and knocked unconscious. The blast wave spread around the world, as recorded by instruments designed to measure changes in atmospheric pressure.



Figure 8.20 : Aftermath of the Tunguska Explosion. This photograph, taken 19 years after the blast, shows a part of the forest that was destroyed by the 5-megaton explosion, resulting when a stony projectile about the size of a small office building (40 meters in diameter) collided with our planet. (credit: modification of work by Leonid Kulik)

Despite this violence, no craters were formed by the Tunguska explosion. Shattered by atmospheric pressure, the stony projectile with a diameter of approximately 60 meters disintegrated above our planet's surface to create a blast equivalent to a 10-megaton nuclear bomb. Had it been smaller or more fragile, the impacting body would have dissipated its energy at high altitude and probably attracted no attention. Today, such high-altitude atmospheric explosions are monitored regularly by military surveillance systems.

If it had been larger or made of stronger material (such as metal), the Tunguska projectile would have penetrated all the way to the surface of Earth and exploded to form a crater. Instead, only the heat and shock of the atmospheric explosion reached the surface, but the devastation it left behind in Siberia bore witness to the power of such impacts. Imagine if the same rocky impactor had exploded over New York City in 1908; history books might today record it as one of the most deadly events in human history.

Tens of thousands of people witnessed directly the explosion of a smaller (20-meter) projectile over the Russian city of Chelyabinsk on an early winter morning in 2013. It exploded at a height of 21 kilometers in a burst of light brighter than the Sun, and the shockwave of the 0.5-megaton explosion broke tens of thousands of windows and sent hundreds of people to the hospital. Rock fragments (meteorites) were easily collected by people in the area after the blast because they landed on fresh snow.

Link to Learning

Dr. David Morrison, one of the original authors of this textbook, provides a nontechnical talk about the Chelyabinsk explosion, and impacts in general.

The best-known recent crater on Earth was formed about 50,000 years ago in Arizona. The projectile in this case was a lump of iron about 40 meters in diameter. Now called *Meteor Crater* and a major tourist attraction on the way to the Grand Canyon, the

crater is about a mile across and has all the features associated with similar-size lunar impact craters (Figure 8.21). Meteor Crater is one of the few impact features on Earth that remains relatively intact; some older craters are so eroded that only a trained eye can distinguish them. Nevertheless, more than 150 have been identified. (See the list of suggested online sites at the end of this chapter if you want to find out more about these other impact scars.)



Figure 8.21 : Meteor Crater in Arizona. Here we see a 50,000-year-old impact crater made by the collision of a 40-meter lump of iron with our planet. Although impact craters are common on less active bodies such as the Moon, this is one of the very few well-preserved craters on Earth. (modification of work by D. Roddy/USGS)

Mass Extinction

The impact that produced Meteor Crater would have been dramatic indeed to any humans who witnessed it (from a safe distance) since the energy release was equivalent to a 10-megaton nuclear bomb. But such explosions are devastating only in their local areas; they have no *global* consequences. Much larger (and rarer) impacts, however, can disturb the ecological balance of the entire planet and thus influence the course of evolution.

The best-documented large impact took place 65 million years ago, at the end of what is now called the Cretaceous period of geological history. This time in the history of life on Earth was marked by a mass extinction, in which more than half of the species on our planet died out. There are a dozen or more mass extinctions in the geological record, but this particular event (nicknamed the “great dying”) has always intrigued paleontologists because it marks the end of the dinosaur age. For tens of millions of years these great creatures had flourished and dominated. Then, they suddenly disappeared (along with many other species), and thereafter mammals began the development and diversification that ultimately led to all of us.

The object that collided with Earth at the end of the Cretaceous period struck a shallow sea in what is now the Yucatán peninsula of Mexico. Its mass must have been more than a trillion tons, determined from study of a worldwide layer of sediment deposited from the dust cloud that enveloped the planet after its impact. First identified in 1979, this sediment layer is rich in the rare metal iridium and other elements that are relatively abundant in asteroids and comets, but exceedingly rare in Earth’s crust. Even though it was diluted by the material that the explosion excavated from the surface of Earth, this cosmic component can still be identified. In addition, this layer of sediment contains many minerals characteristic of the temperatures and pressures of a gigantic explosion.

The impact that led to the extinction of dinosaurs released energy equivalent to 5 billion Hiroshima-size nuclear bombs and excavated a crater 200 kilometers across and deep enough to penetrate through Earth’s crust. This large crater, named Chicxulub for a small town near its center, has subsequently been buried in sediment, but its outlines can still be identified (Figure 8.22). The explosion that created the Chicxulub crater lifted about 100 trillion tons of dust into the atmosphere. We can determine this amount by measuring the thickness of the sediment layer that formed when this dust settled to the surface.



Figure 8.22 : Site of the Chicxulub Crater. This map shows the location of the impact crater created 65 million years ago on Mexico's Yucatán peninsula. The crater is now buried under more than 500 meters of sediment. (credit: modification of work by "Carport"/Wikimedia)

Such a quantity of airborne material would have blocked sunlight completely, plunging Earth into a period of cold and darkness that lasted several months. Many plants dependent on sunlight would have died, leaving plant-eating animals without a food supply. Other worldwide effects included large-scale fires (started by the hot, flying debris from the explosion) that destroyed much of the planet's forests and grasslands, and a long period in which rainwater around the globe was acidic. It was these environmental effects, rather than the explosion itself, that were responsible for the mass extinction, including the demise of the dinosaurs.

Impacts and the Evolution of Life

It is becoming clear that many—perhaps most—mass extinctions in Earth's long history resulted from a variety of other causes, but in the case of the dinosaur killer, the cosmic impact certainly played a critical role and may have been the "final straw" in a series of climactic disturbances that resulted in the "great dying."

A catastrophe for one group of living things, however, may create opportunities for another group. Following each mass extinction, there is a sudden evolutionary burst as new species develop to fill the ecological niches opened by the event. Sixty-five million years ago, our ancestors, the mammals, began to thrive when so many other species died out. We are the lucky beneficiaries of this process.

Impacts by comets and asteroids represent the only mechanisms we know of that could cause truly global catastrophes and seriously influence the evolution of life all over the planet. As paleontologist Stephen Jay Gould of Harvard noted, such a perspective changes fundamentally our view of biological evolution. The central issues for the survival of a species must now include more than just its success in competing with other species and adapting to slowly changing environments, as envisioned by Darwin's idea of natural selection. Also required is an ability to survive random global catastrophes due to impacts.

Still earlier in its history, Earth was subject to even larger impacts from the leftover debris of planet formation. We know that the Moon was struck repeatedly by objects larger than 100 kilometers in diameter—1000 times more massive than the object that wiped out most terrestrial life 65 million years ago. Earth must have experienced similar large impacts during its first 700 million years of existence. Some of them were probably violent enough to strip the planet of most its atmosphere and to boil away its oceans. Such events would sterilize the planet, destroying any life that had begun. Life may have formed and been wiped out several times before our own microbial ancestors took hold sometime about 4 billion years ago.

The fact that the oldest surviving microbes on Earth are thermophiles (adapted to very high temperatures) can also be explained by such large impacts. An impact that was just a bit too small to sterilize the planet would still have destroyed anything that lived in what we consider "normal" environments, and only the creatures adapted to high temperatures would survive. Thus, the oldest surviving terrestrial lifeforms are probably the remnants of a sort of evolutionary bottleneck caused by repeated large impacts early in the planet's history.

Impacts in Our Future?

The impacts by asteroids and comets that have had such a major influence on life are not necessarily a thing of the past. In the full scope of planetary history, 65 million years ago was just yesterday. Earth actually orbits the Sun within a sort of cosmic shooting gallery, and although major impacts are rare, they are by no means over. Humanity could suffer the same fate as the dinosaurs, or lose a city to the much more frequent impacts like the one over Tunguska, unless we figure out a way to predict the next big impact and to protect our planet. The fact that our solar system is home to some very large planets in outer orbits may be beneficial to us; the gravitational fields of those planets can be very effective at pulling in cosmic debris and shielding us from larger, more frequent impacts.

Beginning in the 1990s, a few astronomers began to analyze the cosmic impact hazard and to persuade the government to support a search for potentially hazardous asteroids. Several small but sophisticated wide-field telescopes are now used for this search, which is called the NASA Spaceguard Survey. Already we know that there are currently no asteroids on a collision course with Earth that are as big (10–15 kilometers) as the one that killed the dinosaurs. The Spaceguard Survey now concentrates on finding smaller potential impactors. By 2021, the search had netted more than 98 percent of the hazardous near-Earth asteroids larger than 1 kilometer. The Spaceguard Survey will soon be augmented by the Vera Rubin Telescope that will scan the full sky on a regular basis, and an orbiting search telescope is under development by NASA. Of course, we cannot make a similar statement about the asteroids that have not yet been discovered, but these will be found and evaluated one by one for their potential hazard. These asteroid surveys are one of the few really life-and-death projects carried out by astronomers, with a potential to help to save our planet from future major impacts. Scientists are now studying how we could protect ourselves from an impact if a large asteroid were discovered on a collision course. The NASA spacecraft called DART (Double Asteroid Redirection Test), launched in late 2021, will demonstrate one example of such technology by crashing into the small asteroid Diomorphos to slightly alter its orbit.

Link to Learning

The Torino Impact Hazard Scale is a method for categorizing the impact hazard associated with near-Earth objects such as asteroids and comets. It is a communication tool for astronomers and the public to assess the seriousness of collision predictions by combining probability statistics and known kinetic damage potentials into a single threat value.

Purdue University's "Impact: Earth" calculator lets you input the characteristics of an approaching asteroid to determine the effect of its impact on our planet.

This page titled 12.5: Cosmic Influences on the Evolution of Earth is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 8.6: Cosmic Influences on the Evolution of Earth by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

12.6: Key Terms

bar
: a force of 100,000 Newtons acting on a surface area of 1 square meter; the average pressure of Earth's atmosphere at sea level is 1.013 bars

basalt
: igneous rock produced by the cooling of lava; makes up most of Earth's oceanic crust and is found on other planets that have experienced extensive volcanic activity

convection
: movement caused within a gas or liquid by the tendency of hotter, and therefore less dense material, to rise and colder, denser material to sink under the influence of gravity, which consequently results in transfer of heat

core
: the central part of the planet; consists of higher density material

crust
: the outer layer of a terrestrial planet

fault
: in geology, a crack or break in the crust of a planet along which slippage or movement can take place, accompanied by seismic activity

granite
: a type of igneous silicate rock that makes up most of Earth's continental crust

greenhouse effect
: the blanketing (absorption) of infrared radiation near the surface of a planet—for example, by CO_2 in its atmosphere

greenhouse gas
: a gas in an atmosphere that absorbs and emits radiation within the thermal infrared range; on Earth, these atmospheric gases primarily include carbon dioxide, methane, and water vapor

igneous rock
: rock produced by cooling from a molten state

magnetosphere
: the region around a planet in which its intrinsic magnetic field dominates the interplanetary field carried by the solar wind; hence, the region within which charged particles can be trapped by the planetary magnetic field

mantle
: the largest part of Earth's interior; lies between the crust and the core

mass extinction
: the sudden disappearance in the fossil record of a large number of species of life, to be replaced by fossils of new species in subsequent layers; mass extinctions are indicators of catastrophic changes in the environment, such as might be produced by a large impact on Earth

metamorphic rock
: rock produced by physical and chemical alteration (without melting) under high temperature and pressure

ozone
: (O_3) a heavy molecule of oxygen that contains three atoms rather than the more normal two

photosynthesis
: a complex sequence of chemical reactions through which some living things can use sunlight to manufacture products that store energy (such as carbohydrates), releasing oxygen as one by-product

plate tectonics
: the motion of segments or plates of the outer layer of a planet over the underlying mantle

primitive rock
: rock that has not experienced great heat or pressure and therefore remains representative of the original condensed materials from the solar nebula

rift zone
: in geology, a place where the crust is being torn apart by internal forces generally associated with the injection of new material from the mantle and with the slow separation of tectonic plates

sedimentary rock
: rock formed by the deposition and cementing of fine grains of material, such as pieces of igneous rock or the shells of living things

seismic wave
: a vibration that travels through the interior of Earth or any other object; on Earth, these are generally caused by earthquakes

stratosphere
: the layer of Earth's atmosphere above the troposphere and below the ionosphere

subduction
: the sideways and downward movement of the edge of a plate of Earth's crust into the mantle beneath another plate

troposphere
: the lowest level of Earth's atmosphere, where most weather takes place

volcano
: a place where material from a planet's mantle erupts on its surface

This page titled 12.6: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

13: THE MOON AND MERCURY

The Moon is the only other world human beings have ever visited. What is it like to stand on the surface of our natural satellite? And what can we learn from going there and bringing home pieces of a different world?

We begin our discussion of the planets as cratered worlds with two relatively simple objects: the Moon and Mercury. Unlike Earth, the Moon is geologically dead, a place that has exhausted its internal energy sources. Because its airless surface preserves events that happened long ago, the Moon provides a window on earlier epochs of solar system history. The planet Mercury is in many ways similar to the Moon, which is why the two are discussed together: both are relatively small, lacking in atmospheres, deficient in geological activity, and dominated by the effects of impact cratering. Still, the processes that have molded their surfaces are not unique to these two worlds. We shall see that they have acted on many other members of the planetary system as well.

13.1: General Properties of the Moon

13.2: The Lunar Surface

13.3: Impact Craters

13.4: The Origin of the Moon

13.5: Mercury

13.6: Key Terms

This page titled 13: The Moon and Mercury is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9: Cratered Worlds by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.1: General Properties of the Moon

Learning Objectives

By the end of this section, you will be able to:

- Discuss what has been learned from both manned and robotic lunar exploration
- Describe the composition and structure of the Moon

The Moon has only one-eightieth the mass of Earth and about one-sixth Earth's surface gravity—too low to retain an atmosphere (Figure 9.2). Moving molecules of a gas can escape from a planet just the way a rocket does, and the lower the gravity, the easier it is for the gas to leak away into space. While the Moon can acquire a temporary atmosphere from impacting comets, this atmosphere is quickly lost by freezing onto the surface or by escape to surrounding space. The Moon today is dramatically deficient in a wide range of *volatiles*, those elements and compounds that evaporate at relatively low temperatures. Some of the Moon's properties are summarized in Table 9.1, along with comparative values for Mercury.

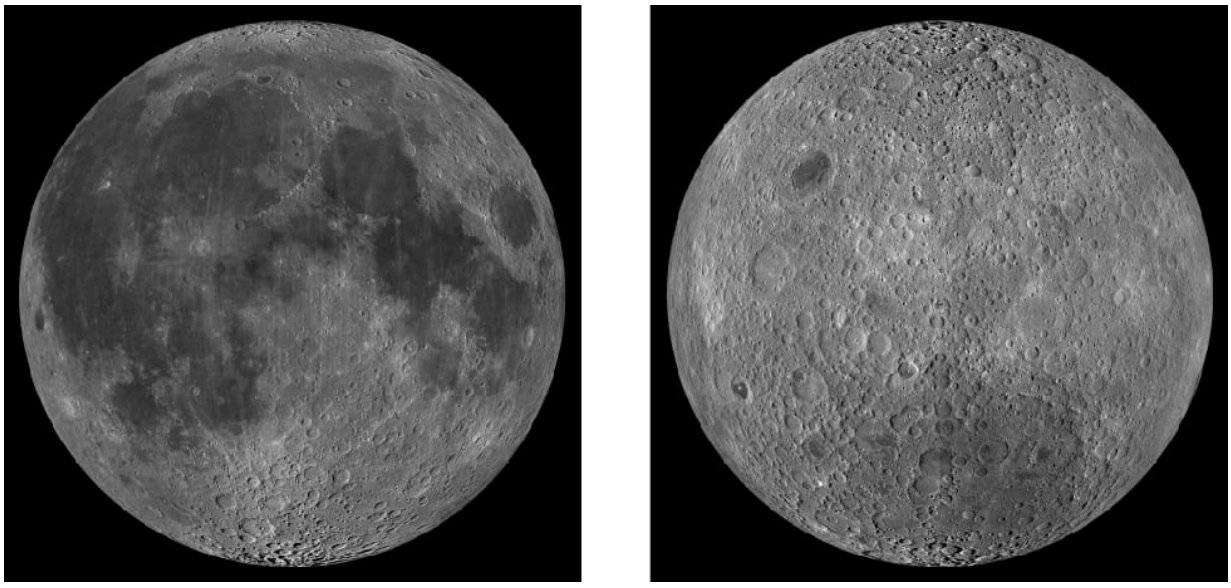


Figure 9.2 : Two Sides of the Moon. The left image shows part of the hemisphere that faces Earth; several dark maria are visible. The right image shows part of the hemisphere that faces away from Earth; it is dominated by highlands. The resolution of this image is several kilometers, similar to that of high-powered binoculars or a small telescope. (credit: modification of work by NASA/GSFC/Arizona State University)

Table 9.1: Properties of the Moon and Mercury

Property	Moon	Mercury
Mass (Earth = 1)	0.0123	0.055
Diameter (km)	3476	4878
Density (g/cm^3)	3.3	5.4
Surface gravity (Earth = 1)	0.17	0.38
Escape velocity (km/s)	2.4	4.3
Rotation period (days)	27.3	58.65
Surface area (Earth = 1)	0.27	0.38

Exploration of the Moon

Most of what we know about the Moon today derives from the US Apollo program, which sent nine piloted spacecraft to our satellite between 1968 and 1972, landing 12 astronauts on its surface (Figure 9.1). Before the era of spacecraft studies, astronomers had mapped the side of the Moon that faces Earth with telescopic resolution of about 1 kilometer, but lunar geology hardly existed as a scientific subject. All that changed beginning in the early 1960s. Initially, Russia took the lead in lunar exploration with Luna 3, which returned the first photos of the lunar far side in 1959, and then with Luna 9, which landed on the surface in 1966 and transmitted pictures and other data to Earth. However, these efforts were overshadowed on July 20, 1969, when the first American astronaut set foot on the Moon.

Table 9.2 summarizes the nine Apollo flights: six that landed and three others that circled the Moon but did not land. The initial landings were on flat plains selected for safety reasons. But with increasing experience and confidence, NASA targeted the last three missions to more geologically interesting locales. The level of scientific exploration also increased with each mission, as the astronauts spent longer times on the Moon and carried more elaborate equipment. Finally, on the last Apollo landing, NASA included one scientist, geologist Jack Schmitt, among the astronauts (Figure 9.3).

Table 9.2: Apollo Flights to the Moon

Flight	Date	Landing Site	Main Accomplishment
Apollo 8	Dec. 1968	—	First humans to fly around the Moon
Apollo 10	May 1969	—	First spacecraft rendezvous in lunar orbit
Apollo 11	July 1969	Mare Tranquillitatis	First human landing on the Moon; 22 kilograms of samples returned
Apollo 12	Nov. 1969	Oceanus Procellarum	First Apollo Lunar Surface Experiment Package (ALSEP); visit to Surveyor 3 lander
Apollo 13	Apr. 1970	—	Landing aborted due to explosion in service module
Apollo 14	Jan. 1971	Mare Nubium	First “rickshaw” on the Moon
Apollo 15	July 1971	Mare Imbrium/Hadley	First “rover;” visit to Hadley Rille; astronauts traveled 24 kilometers
Apollo 16	Apr. 1972	Descartes	First landing in highlands; 95 kilograms of samples returned
Apollo 17	Dec. 1972	Taurus-Littrow highlands	Geologist among the crew; 111 kilograms of samples returned



Figure 9.3 : Scientist on the Moon. Geologist (and later US senator) Harrison “Jack” Schmitt in front of a large boulder in the Littrow Valley at the edge of the lunar highlands. Note how black the sky is on the airless Moon. No stars are visible because the surface is brightly lit by the Sun, and the exposure therefore is not long enough to reveal stars.

In addition to landing on the lunar surface and studying it at close range, the Apollo missions accomplished three objectives of major importance for lunar science. First, the astronauts collected nearly 400 kilograms of samples for detailed laboratory analysis on Earth (Figure 9.4). These samples have revealed as much about the Moon and its history as all other lunar studies combined. Second, each Apollo landing after the first one deployed an Apollo Lunar Surface Experiment Package (ALSEP), which continued to operate for years after the astronauts departed. Third, the orbiting Apollo command modules carried a wide range of instruments to photograph and analyze the lunar surface from above.



Figure 9.4 : Handling Moon Rocks. Lunar samples collected in the Apollo Project are analyzed and stored in NASA facilities at the Johnson Space Center in Houston, Texas. Here, a technician examines a rock sample using gloves in a sealed environment to avoid contaminating the sample. (credit: NASA JSC)

The last human left the Moon in December 1972, just a little more than three years after Neil Armstrong took his “giant leap for mankind.” The program of lunar exploration was cut off midstride due to political and economic pressures. It had cost just about $100 per American, spread over 10 years—the equivalent of one large pizza per person per year. Yet for many people, the Moon landings were one of the central events in twentieth-century history.

The giant Apollo rockets built to travel to the Moon were left to rust on the lawns of NASA centers in Florida, Texas, and Alabama,

although recently, some have at least been moved indoors to museums (Figure 9.5). It is more than 50 years since astronauts landed on the Moon, and there are no firm plans for future flights. However, both the United States and China are developing the capability to land on the Moon, and there may be a new "space race" in the late 2020s. (In a bizarre piece of irony, a few people even question whether we went to the Moon at all, proposing instead that the Apollo program was a fake, filmed on a Hollywood sound stage. See the Link to Learning box below for some scientists' replies to such claims.) However, scientific interest in the Moon is stronger than ever, and more than half a dozen scientific spacecraft—sent from NASA, ESA, Japan, India, and China—have orbited or landed on our nearest neighbor during the past two decades.x

Link to Learning

Read The Great Moon Hoax about the claim that NASA never succeeded in putting people on the Moon.

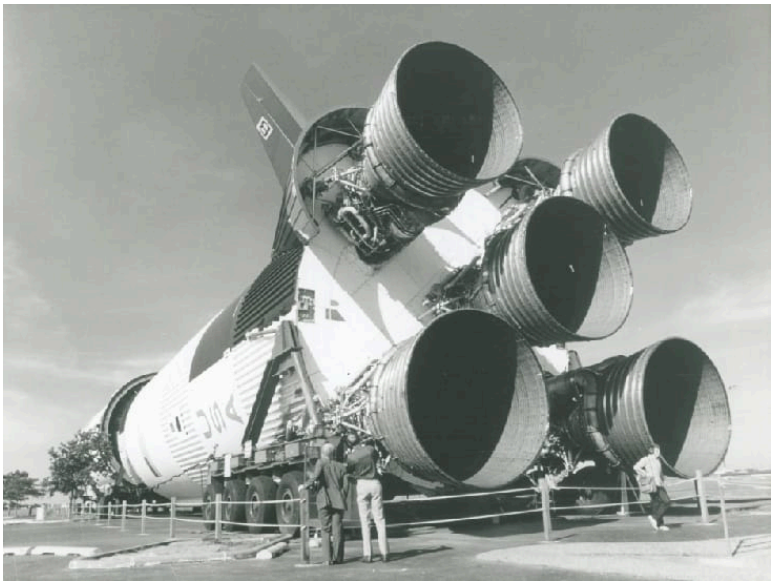


Figure 9.5 : Moon Rocket on Display. One of the unused Saturn 5 rockets built to go to the Moon is now a tourist attraction at NASA's Johnson Space Center in Houston, although it has been moved indoors since this photo was taken. (credit: modification of work by David Morrison)

Lunar exploration has become an international enterprise with many robotic spacecraft focusing on lunar science. The USSR sent a number in the 1960s, including robot sample returns, and recently China has been active, with three landers and a sample return mission. Of special interest is the search for accessible water ice near the lunar South Pole. Table 9.3 lists some of the most recent lunar missions.

Table 9.3: Some International Missions to the Moon

Launch Year	Spacecraft	Type of Mission	Agency
1994	Clementine	Orbiter	US (USAF/NASA)
1998	Lunar Prospector	Orbiter	US (NASA)
2003	SMART-1	Orbiter	Europe (ESA)
2007	SELENE 1	Orbiter	Japan (JAXA)
2007	Chang'e 1	Orbiter	China (CNSA)
2008	Chandrayaan-1	Orbiter	India (ISRO)
2009	LRO	Orbiter	US (NASA)
2009	LCROSS	Impactor	US (NASA)
2010	Chang'e 2	Orbiter	China (CNSA)
2011	GRAIL	Twin orbiters	US (NASA)

Launch Year	Spacecraft	Type of Mission	Agency
2013	LADEE	Orbiter	US (NASA)
2013	Chang'e 3	Lander/Rover	China (CNSA)
2019	Chang-e 4	Lander/Rover (far side)	China(CNSA)
2020	Chang-e 5	Lander/Sample Return	China(CNSA)

Composition and Structure of the Moon

The composition of the Moon is not the same as that of Earth. With an average density of only 3.3 g/cm^3, the Moon must be made almost entirely of silicate rock. Compared to Earth, it is depleted in iron and other metals. It is as if the Moon were composed of the same silicates as Earth's mantle and crust, with the metals and the volatiles selectively removed. These differences in composition between Earth and Moon provide important clues about the origin of the Moon, a topic we will cover in detail later in this chapter.

Studies of the Moon's interior carried out with seismometers taken to the Moon as part of the Apollo program confirm the absence of a large metal core. The twin GRAIL spacecraft launched into lunar orbit in 2011 provided even more precise tracking of the interior structure. We also know from the study of lunar samples that water and other volatiles have been depleted from the lunar crust. The tiny amounts of water detected in these samples were originally attributed to small leaks in the container seal that admitted water vapor from Earth's atmosphere. However, scientists have now concluded that some chemically bound water is present in the lunar rocks.

Most dramatically, water ice has been detected in permanently shadowed craters near the lunar poles. In 2009, NASA crashed a small spacecraft called the Lunar Crater Observation and Sensing Satellite (LCROSS) into the crater Cabeus near the Moon's south pole. The impact at 9,000 kilometers per hour released energy equivalent to 2 tons of dynamite, blasting a plume of water vapor and other chemicals high above the surface. This plume was visible to telescopes in orbit around the Moon, and the LCROSS spacecraft itself made measurements as it flew through the plume. A NASA spacecraft called the Lunar Reconnaissance Orbiter (LRO) also measured the very low temperatures inside several lunar craters, and its sensitive cameras were even able to image crater interiors by starlight.

The total quantity of water ice in the Moon's polar craters is estimated to be hundreds of billions of tons. As liquid, this would only be enough water to fill a lake 100 miles across, but compared with the rest of the dry lunar crust, so much water is remarkable. Presumably, this polar water was carried to the Moon by comets and asteroids that hit its surface. Some small fraction of the water froze in a few extremely cold regions (cold traps) where the Sun never shines, such as the bottom of deep craters at the Moon's poles. One reason this discovery could be important is that it raises the possibility of future human habitation near the lunar poles, or even of a lunar base as a way-station on routes to Mars and the rest of the solar system. If the ice could be mined, it would yield both water and oxygen for human support, and it could be broken down into hydrogen and oxygen, a potent rocket fuel.

This page titled 13.1: General Properties of the Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9.2: General Properties of the Moon by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.2: The Lunar Surface

Learning Objectives

By the end of this section, you will be able to:

- Differentiate between the major surface features of the Moon
- Describe the history of the lunar surface
- Describe the properties of the lunar “soil”

General Appearance

If you look at the Moon through a telescope, you can see that it is covered by impact craters of all sizes. The most conspicuous of the Moon’s surface features—those that can be seen with the unaided eye and that make up the feature often called “the man in the Moon”—are vast splotches of darker lava flows.

Centuries ago, early lunar observers thought that the Moon had continents and oceans and that it was a possible abode of life. They called the dark areas “seas” (*maria* in Latin, or *mare* in the singular, pronounced “mah ray”). Their names, Mare Nubium (Sea of Clouds), Mare Tranquillitatis (Sea of Tranquility), and so on, are still in use today. In contrast, the “land” areas between the seas are not named. Thousands of individual craters have been named, however, mostly for great scientists and philosophers (Figure 9.6). Among the most prominent craters are those named for Plato, Copernicus, Tycho, and Kepler. Galileo only has a small crater, however, reflecting his low standing among the Vatican scientists who made some of the first lunar maps.

We know today that the resemblance of lunar features to terrestrial ones is superficial. Even when they look somewhat similar, the origins of lunar features such as craters and mountains are very different from their terrestrial counterparts. The Moon’s relative lack of internal activity, together with the absence of air and water, make most of its geological history unlike anything we know on Earth.

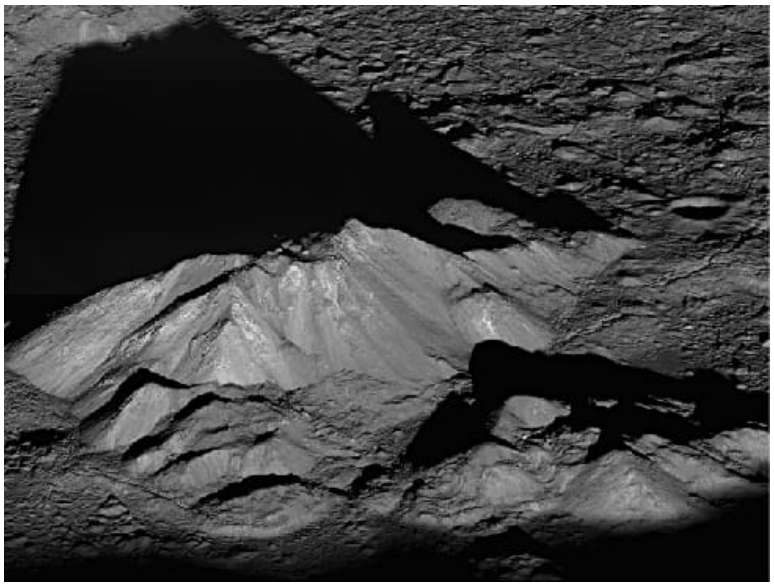


Figure 9.6 : Sunrise on the Central Mountain Peaks of Tycho Crater, as Imaged by the NASA Lunar Reconnaissance Orbiter. Tycho, about 82 kilometers in diameter, is one of the youngest of the very large lunar craters. The central mountain rises 2 kilometers above the crater floor. (credit: modification of work by NASA/Goddard/Arizona State University)

Lunar History

To trace the detailed history of the Moon or of any planet, we must be able to estimate the ages of individual rocks. Once lunar samples were brought back by the Apollo astronauts, the radioactive dating techniques that had been developed for Earth were applied to them. The solidification ages of the samples ranged from about 3.3 to 4.4 billion years old, substantially older than most of the rocks on Earth. For comparison, as we saw in the chapter on Earth, Moon, and Sky, both Earth and the Moon were formed between 4.5 and 4.6 billion years ago.

Most of the crust of the Moon (83%) consists of silicate rocks called *anorthosites*; these regions are known as the lunar highlands. They are made of relatively low-density rock that solidified on the cooling Moon like slag floating on the top of a smelter. Because they formed so early in lunar history (between 4.1 and 4.4 billion years ago), the highlands are also extremely heavily cratered, bearing the scars of all those billions of years of impacts by interplanetary debris (Figure 9.7).



Figure 9.7 : Lunar Highlands. The old, heavily cratered lunar highlands make up 83% of the Moon's surface. (credit: Apollo 11 Crew, NASA)

Unlike the mountains on Earth, the Moon's highlands do not have any sharp folds in their ranges. The highlands have low, rounded profiles that resemble the oldest, most eroded mountains on Earth (Figure 9.8). Because there is no atmosphere or water on the Moon, there has been no wind, water, or ice to carve them into cliffs and sharp peaks, the way we have seen them shaped on Earth. Their smooth features are attributed to gradual erosion, mostly due to impact cratering.

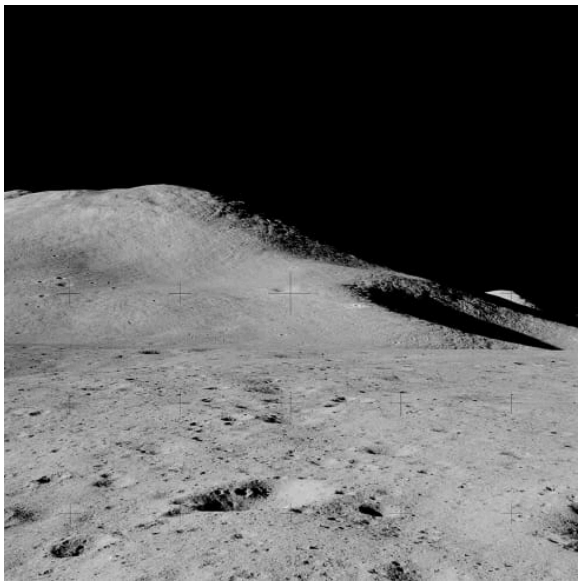


Figure 9.8 : Lunar Mountain. This photo of Mt. Hadley on the edge of Mare Imbrium was taken by Dave Scott, one of the Apollo 15 astronauts. Note the smooth contours of the lunar mountains, which have not been sculpted by water or ice. (credit: NASA/Apollo Lunar Surface Journal)

The maria are much less cratered than the highlands, and cover just 17% of the lunar surface, mostly on the side of the Moon that faces Earth (Figure 9.9).

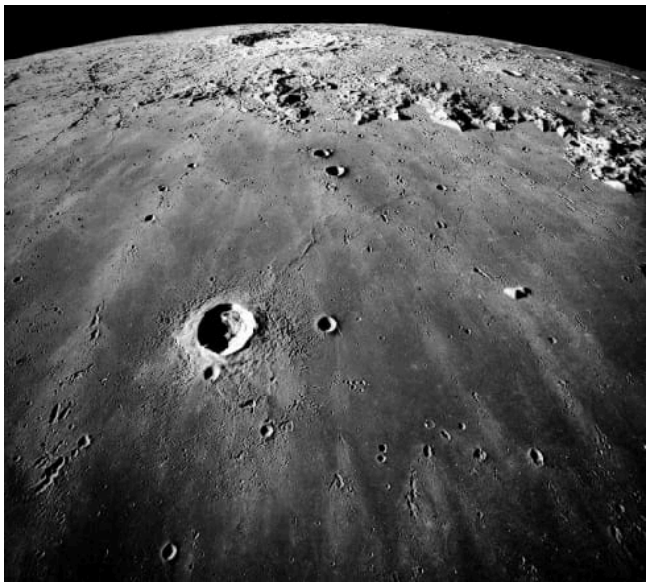


Figure 9.9 : Lunar Maria. About 17% of the Moon's surface consists of the maria—flat plains of basaltic lava. This view of Mare Imbrium also shows numerous secondary craters and evidence of material ejected from the large crater Copernicus on the upper horizon. Copernicus is an impact crater almost 100 kilometers in diameter that was formed long after the lava in Imbrium had already been deposited. (credit: NASA, Apollo 17)

Today, we know that the maria consist mostly of dark-colored basalt (volcanic lava) laid down in volcanic eruptions billions of years ago. Eventually, these lava flows partly filled the huge depressions called *impact basins*, which had been produced by collisions of large chunks of material with the Moon relatively early in its history. The basalt on the Moon (Figure 9.10) is very similar in composition to the crust under the oceans of Earth or to the lavas erupted by many terrestrial volcanoes. The youngest of the lunar impact basins is Mare Orientale, shown in Figure 9.11.



Figure 9.10 : Rock from a Lunar Mare. In this sample of basalt from the mare surface, you can see the holes left by gas bubbles, which are characteristic of rock formed from lava. All lunar rocks are chemically distinct from terrestrial rocks, a fact that has allowed scientists to identify a few lunar samples among the thousands of meteorites that reach Earth. (credit: modification of work by NASA)

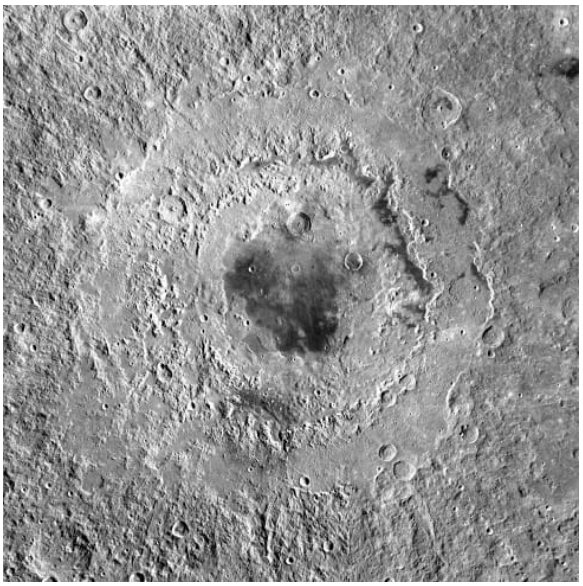


Figure 9.11 : Mare Orientale. The youngest of the large lunar impact basins is Orientale, formed 3.8 billion years ago. Its outer ring is about 1000 kilometers in diameter, roughly the distance between New York City and Detroit, Michigan. Unlike most of the other basins, Orientale has not been completely filled in with lava flows, so it retains its striking "bull's-eye" appearance. It is located on the edge of the Moon as seen from Earth. (credit: NASA)

Volcanic activity may have begun very early in the Moon's history, although most evidence of the first half billion years is lost. What we do know is that the major mare volcanism, which involved the release of lava from hundreds of kilometers below the surface, ended about 3.3 billion years ago. After that, the Moon's interior cooled, and volcanic activity was limited to a very few small areas. The primary forces altering the surface come from the outside, not the interior.

On the Lunar Surface

"The surface is fine and powdery. I can pick it up loosely with my toe. But I can see the footprints of my boots and the treads in the fine sandy particles." —Neil Armstrong, Apollo 11 astronaut, immediately after stepping onto the Moon for the first time.

The surface of the Moon is buried under a fine-grained soil of tiny, shattered rock fragments. The dark basaltic dust of the lunar maria was kicked up by every astronaut footstep, and thus eventually worked its way into all of the astronauts' equipment. The

upper layers of the surface are porous, consisting of loosely packed dust into which their boots sank several centimeters (Figure 9.12). This lunar dust, like so much else on the Moon, is the product of impacts. Each cratering event, large or small, breaks up the rock of the lunar surface and scatters the fragments. Ultimately, billions of years of impacts have reduced much of the surface layer to particles about the size of dust or sand.



Figure 9.12 : Footprint on Moon Dust. Apollo photo of an astronaut's boot print in the lunar soil. (credit: NASA)

In the absence of any air, the lunar surface experiences much greater temperature extremes than the surface of Earth, even though Earth is virtually the same distance from the Sun. Near local noon, when the Sun is highest in the sky, the temperature of the dark lunar soil rises above the boiling point of water. During the long lunar night (which, like the lunar day, lasts two Earth weeks[1]), the temperature drops to about 100 K (–173 °C). The extreme cooling is a result not only of the absence of air but also of the porous nature of the Moon's dusty soil, which cools more rapidly than solid rock would.

Link to Learning

Learn how the moon's craters and maria were formed by watching a video produced by NASA's Lunar Reconnaissance Orbiter (LRO) team about the evolution of the Moon, tracing it from its origin about 4.5 billion years ago to the Moon we see today. See a simulation of how the Moon's craters and maria were formed through periods of impact, volcanic activity, and heavy bombardment.

Footnotes

- 1You can see the cycle of day and night on the side of the Moon facing us in the form of the Moon's phases. It takes about 14 days for the side of the Moon facing us to go from full moon (all lit up) to new moon (all dark). There is more on this in Chapter 4: Earth, Moon, and Sky.

This page titled 13.2: The Lunar Surface is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9.3: The Lunar Surface by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.3: Impact Craters

Learning Objectives

By the end of this section, you will be able to:

- Compare and contrast ideas about how lunar craters form
- Explain the process of impact crater formation
- Discuss the use of crater counts to determine relative ages of lunar landforms

Volcanic Versus Impact Origin of Craters

The Moon provides an important benchmark for understanding the history of our planetary system. Most solid worlds show the effects of impacts, often extending back to the era when a great deal of debris from our system's formation process was still present. On Earth, this long history has been erased by our active geology. On the Moon, in contrast, most of the impact history is preserved. If we can understand what has happened on the Moon, we may be able to apply this knowledge to other worlds. The Moon is especially interesting because it is not just any moon, but *our* Moon—a nearby world that has shared the history of Earth for more than 4 billion years and preserved a record that, for Earth, has been destroyed by our active geology.

Until the middle of the twentieth century, scientists did not generally recognize that lunar craters were the result of impacts. Since impact craters are extremely rare on Earth, geologists did not expect them to be the major feature of lunar geology. They reasoned (perhaps unconsciously) that since the craters we have on Earth are volcanic, the lunar craters must have a similar origin.

One of the first geologists to propose that lunar craters were the result of impacts was Grove K. Gilbert, a scientist with the US Geological Survey in the 1890s. He pointed out that the large lunar craters—mountain-rimmed, circular features with floors generally below the level of the surrounding plains—are larger and have different shapes from known volcanic craters on Earth. Terrestrial volcanic craters are smaller and deeper and almost always occur at the tops of volcanic mountains (Figure 9.13). The only alternative to explain the Moon's craters was an impact origin. His careful reasoning, although not accepted at the time, laid the foundations for the modern science of lunar geology.

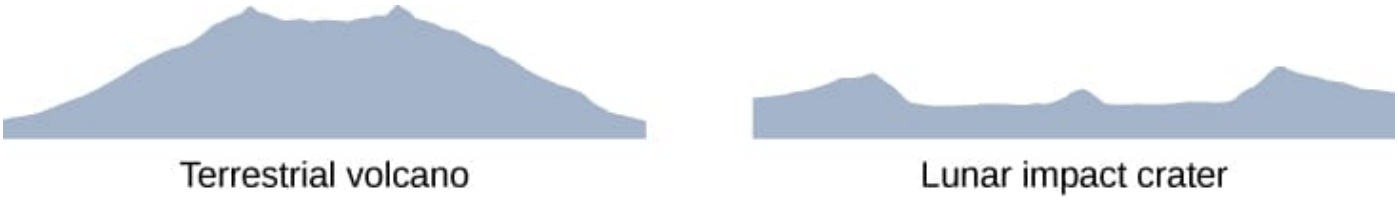


Figure 9.13 : Volcanic and Impact Craters. Profiles of a typical terrestrial volcanic crater and a typical lunar impact crater are quite different.

Gilbert concluded that the lunar craters were produced by impacts, but he didn't understand why all of them were circular and not oval. The reason lies in the escape velocity, the minimum speed that a body must reach to permanently break away from the gravity of another body; it is also the minimum speed that a projectile approaching Earth or the Moon will hit with. Attracted by the gravity of the larger body, the incoming chunk strikes with at least escape velocity, which is 11 kilometers per second for Earth and 2.4 kilometers per second (5400 miles per hour) for the Moon. To this escape velocity is added whatever speed the projectile already had with respect to Earth or Moon, typically 10 kilometers per second or more.

At these speeds, the energy of impact produces a violent *explosion* that excavates a large volume of material in a symmetrical way. Photographs of bomb and shell craters on Earth confirm that explosion craters are always essentially circular. Only following World War I did scientists recognize the similarity between impact craters and explosion craters, but, sadly, Gilbert did not live to see his impact hypothesis widely accepted.

The Cratering Process

Let's consider how an impact at these high speeds produces a crater. When such a fast projectile strikes a planet, it penetrates

two or three times its own diameter before stopping. During these few seconds, its energy of motion is transferred into a shock wave (which spreads through the target body) and into heat (which vaporizes most of the projectile and some of the surrounding target). The shock wave fractures the rock of the target, while the expanding silicate vapor generates an explosion similar to that of a nuclear bomb detonated at ground level (Figure 9.14). The size of the excavated crater depends primarily on the speed of impact, but generally it is 10 to 15 times the diameter of the projectile.

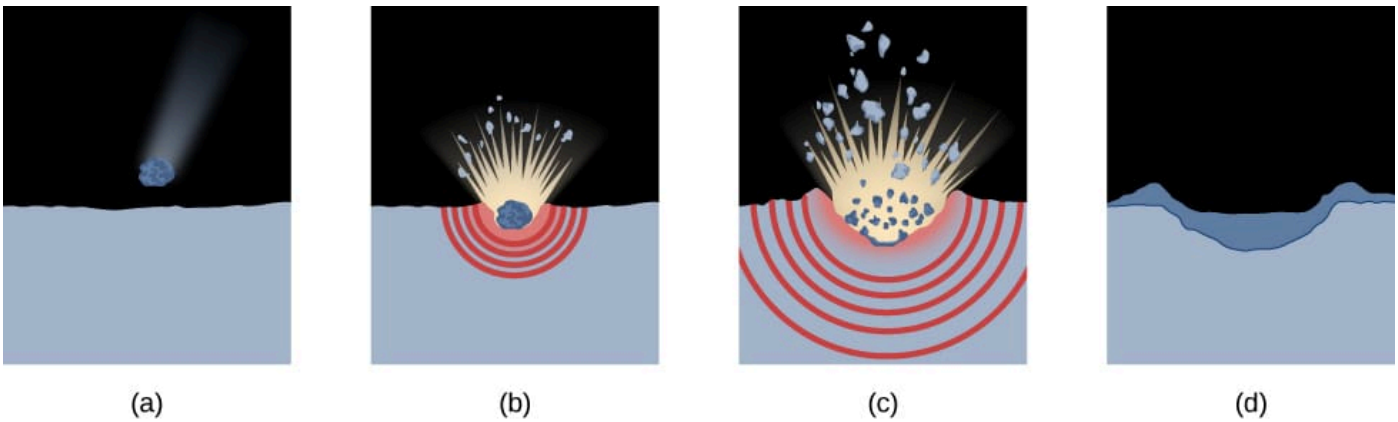


Figure 9.14 : Stages in the Formation of an Impact Crater. (a) The impact occurs. (b) The projectile vaporizes and a shock wave spreads through the lunar rock. (c) Ejecta are thrown out of the crater. (d) Most of the ejected material falls back to fill the crater, but some of it is found outside the crater in the form of an ejecta blanket.

An impact explosion of the sort described above leads to a characteristic kind of crater, as shown in Figure 9.15. The central cavity is initially bowl-shaped (the word “crater” comes from the Greek word for “bowl”), but the rebound of the crust partially fills it in, producing a flat floor and sometimes creating a central peak. Around the rim, landslides create a series of terraces.

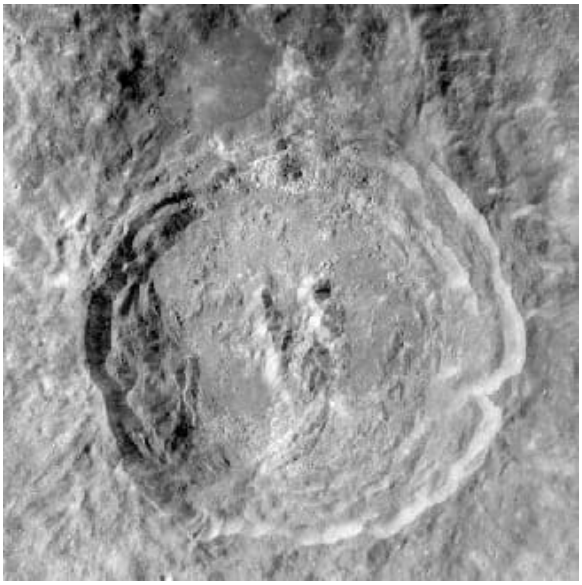


Figure 9.15 : Typical Impact Crater. King Crater on the far side of the Moon, a fairly recent lunar crater 75 kilometers in diameter, shows most of the features associated with large impact structures. (credit: NASA/JSC/Arizona State University)

The rim of the crater is turned up by the force of the explosion, so it rises above both the floor and the adjacent terrain. Surrounding the rim is an *ejecta blanket* consisting of material thrown out by the explosion. This debris falls back to create a rough, hilly region, typically about as wide as the crater diameter. Additional, higher-speed ejecta fall at greater distances from the crater, often digging small *secondary craters* where they strike the surface (Figure 9.9).

Some of these streams of ejecta can extend for hundreds or even thousands of kilometers from the crater, creating the bright *crater rays* that are prominent in lunar photos taken near full phase. The brightest lunar crater rays are associated with large young craters such as Kepler and Tycho.

Seeing for Yourself: Observing the Moon

The Moon is one of the most beautiful sights in the sky, and it is the only object close enough to reveal its *topography* (surface features such as mountains and valleys) without a visit from a spacecraft. A fairly small amateur telescope easily shows craters and mountains on the Moon as small as a few kilometers across.

Even as seen through a good pair of binoculars, we can observe that the appearance of the Moon's surface changes dramatically with its phase. At full phase, it shows almost no topographic detail, and you must look closely to see more than a few craters. This is because sunlight illuminates the surface straight on, and in this flat lighting, no shadows are cast. Much more revealing is the view near first or third quarter, when sunlight streams in from the side, causing topographic features to cast sharp shadows. It is almost always more rewarding to study a planetary surface under such oblique lighting, when the maximum information about surface relief can be obtained.

The flat lighting at full phase does, however, accentuate brightness contrasts on the Moon, such as those between the maria and highlands. Notice in Figure 9.16 that several of the large mare craters seem to be surrounded by white material and that the light streaks or rays that can stretch for hundreds of kilometers across the surface are clearly visible. These lighter features are ejecta, splashed out from the crater-forming impact.



Figure 9.16 : Appearance of the Moon at Different Phases. (a) Illumination from the side brings craters and other topographic features into sharp relief, as seen on the far left side. (b) At full phase, there are no shadows, and it is more difficult to see such features. However, the flat lighting at full phase brings out some surface features, such as the bright rays of ejecta that stretch out from a few large young craters. (credit: modification of work by Luc Viatour)

By the way, there is no danger in looking at the Moon with binoculars or telescopes. The reflected sunlight is never bright enough to harm your eyes. In fact, the sunlit surface of the Moon has about the same brightness as a sunlit landscape of dark rock on Earth. Although the Moon looks bright in the night sky, its surface is, on average, much less reflective than Earth's, with its atmosphere and white clouds. This difference is nicely illustrated by the photo of the Moon passing in front of Earth taken from the Deep Space Climate Observatory spacecraft (Figure 9.17). Since the spacecraft took the image from a position inside the orbit of Earth, we see both objects fully illuminated (full Moon and full Earth). By the way, you cannot see much detail on the Moon because the exposure has been set to give a bright image of Earth, not the Moon.

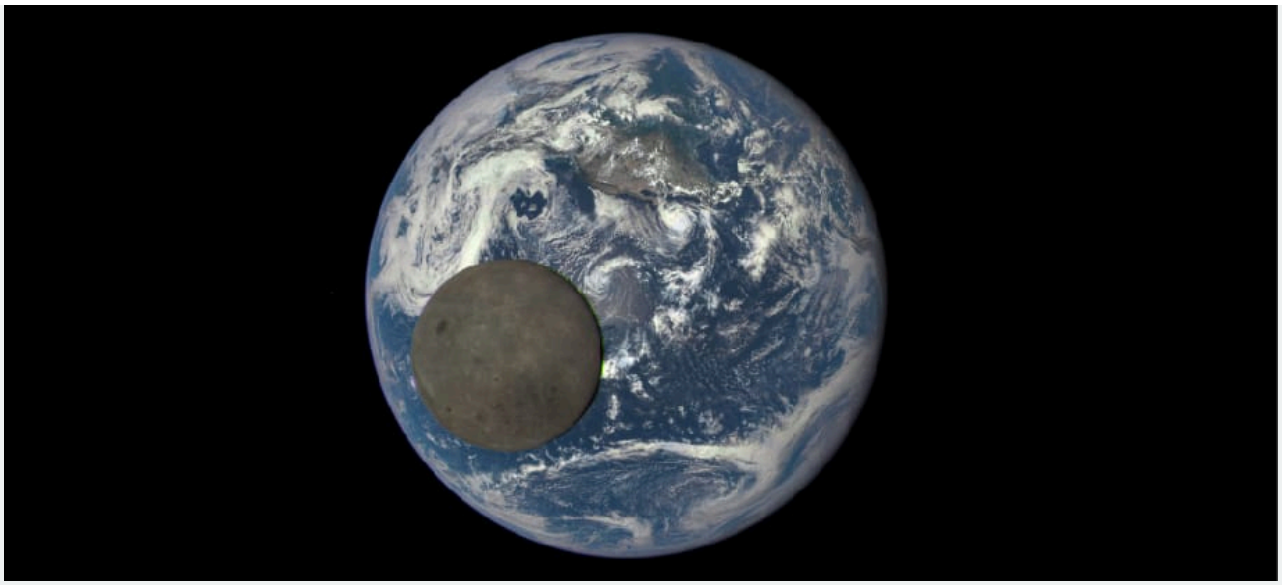


Figure 9.17 : The Moon Crossing the Face of Earth. In this 2015 image from the Deep Space Climate Observatory spacecraft, both objects are fully illuminated, but the Moon looks darker because it has a much lower average reflectivity than Earth. (credit: modification of work by NASA, DSCOVR EPIC team)

One interesting thing about the Moon that you can see without binoculars or telescopes is popularly called “the new Moon in the old Moon’s arms.” Look at the Moon when it is a thin crescent, and you can often make out the faint circle of the entire lunar disk, even though the sunlight shines on only the crescent. The rest of the disk is illuminated not by sunlight but by earthlight—sunlight reflected from Earth. The light of the full Earth on the Moon is about 50 times brighter than that of the full Moon shining on Earth.

Using Crater Counts

If a world has had little erosion or internal activity, like the Moon during the past 3 billion years, it is possible to use the number of impact craters on its surface to estimate the age of that surface. By “age” here we mean the time since a major disturbance occurred on that surface (such as the volcanic eruptions that produced the lunar maria).

We cannot directly measure the rate at which craters are being formed on Earth and the Moon, since the average interval between large crater-forming impacts is longer than the entire span of human history. Our best-known example of such a large crater, Meteor Crater in Arizona (Figure 9.18), is about 50,000 years old. However, the cratering rate can be estimated from the number of craters on the lunar maria or calculated from the number of potential “projectiles” (asteroids and comets) present in the solar system today. Both lines of reasoning lead to about the same estimations.



Figure 9.18 : Meteor Crater. This aerial photo of Meteor Crater in Arizona shows the simple form of a meteorite impact crater. The crater's rim diameter is about 1.2 kilometers. (credit: Shane Torgerson)

For the Moon, these calculations indicate that a crater 1 kilometer in diameter should be produced about every 200,000 years, a 10-kilometer crater every few million years, and one or two 100-kilometer craters every billion years. If the cratering rate has stayed the same, we can figure out how long it must have taken to make all the craters we see in the lunar maria. Our calculations show that it would have taken several billion years. This result is similar to the age determined for the maria from radioactive dating of returned samples—3.3 to 3.8 billion years old.

The fact that these two calculations agree suggests that astronomers' original assumption was right: comets and asteroids in approximately their current numbers have been impacting planetary surfaces for billions of years. Calculations carried out for other planets (and their moons) indicate that they also have been subject to about the same number of interplanetary impacts during this time.

We have good reason to believe, however, that earlier than 3.8 billion years ago, the impact rates must have been a great deal higher. This becomes immediately evident when comparing the numbers of craters on the lunar highlands with those on the maria. Typically, there are 10 times more craters on the highlands than on a similar area of maria. Yet the radioactive dating of highland samples showed that they are only a little older than the maria, typically 4.2 billion years rather than 3.8 billion years. If the rate of impacts had been constant throughout the Moon's history, the highlands would have had to be at least 10 times older. They would thus have had to form 38 billion years ago—long before the universe itself began.

In science, when an assumption leads to an implausible conclusion, we must go back and re-examine that assumption—in this case, the constant impact rate. The contradiction is resolved if the impact rate varied over time, with a much heavier bombardment earlier than 3.8 billion years ago (Figure 9.19). This "heavy bombardment" produced most of the craters we see today in the highlands.

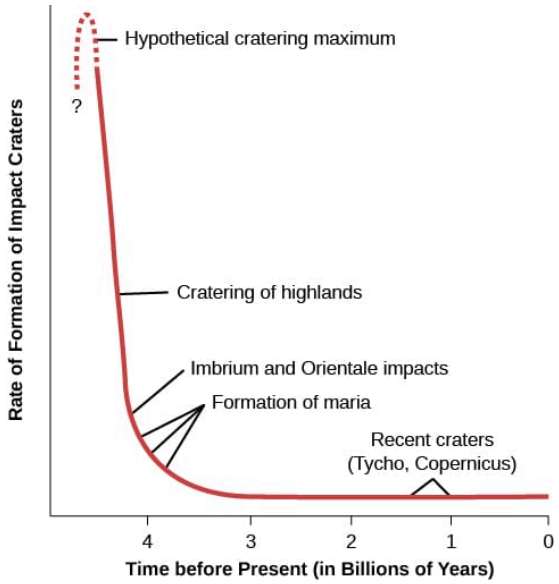


Figure 9.19 : Cratering Rates over Time. The number of craters being made on the Moon's surface has varied with time over the past 4.3 billion years.

This idea we have been exploring—that large impacts (especially during the early history of the solar system) played a major role in shaping the worlds we see—is not unique to our study of the Moon. As you read through the other chapters about the planets, you will see further indications that a number of the present-day characteristics of our system may be due to its violent past.

This page titled 13.3: Impact Craters is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9.4: Impact Craters by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.4: The Origin of the Moon

Learning Objectives

By the end of this section, you will be able to:

- Describe the top three early hypotheses of the formation of the Moon
- Summarize the current “giant impact” concept of how the Moon formed

Ideas for the Origin of the Moon

It is characteristic of modern science to ask how things originated. Understanding the origin of the Moon has proven to be challenging for planetary scientists, however. Part of the difficulty is simply that we know so much about the Moon (quite the opposite of our usual problem in astronomy). As we will see, one key problem is that the Moon is both tantalizingly similar to Earth and frustratingly different.

Most of the earlier hypotheses for the Moon’s origin followed one of three general ideas:

1. The fission theory—the Moon was once part of Earth, but somehow separated from it early in their history.
2. The sister theory—the Moon formed together with (but independent of) Earth, as we believe many moons of the outer planets formed.
3. The capture theory—the Moon formed elsewhere in the solar system and was captured by Earth.

Unfortunately, there seem to be fundamental problems with each of these ideas. Perhaps the easiest hypothesis to reject is the capture theory. Its primary drawback is that no one knows of any way that early Earth could have captured such a large moon from elsewhere. One body approaching another cannot go into orbit around it without a substantial loss of energy; this is the reason that spacecraft destined to orbit other planets are equipped with retro-rockets. Furthermore, if such a capture did take place, the captured object would go into a very eccentric orbit rather than the nearly circular orbit our Moon occupies today. Finally, there are too many compositional similarities between Earth and the Moon, particularly an identical fraction of the major isotopes[2] of oxygen, to justify seeking a completely independent origin.

The fission hypothesis, which states that the Moon separated from Earth, was suggested in the late nineteenth century. Modern calculations have shown that this sort of spontaneous fission or splitting is impossible. Furthermore, it is difficult to understand how a Moon made out of terrestrial material in this way could have developed the many distinctive chemical differences now known to characterize our neighbor.

Scientists were therefore left with the sister hypothesis—that the Moon formed alongside Earth—or with some modification of the fission hypothesis that can find a more acceptable way for the lunar material to have separated from Earth. But the more we learned about our Moon, the less these old ideas seem to fit the bill.

The Giant Impact Hypothesis

In an effort to resolve these apparent contradictions, scientists developed a fourth hypothesis for the origin of the Moon, one that involves a giant impact early in Earth’s history. There is increasing evidence that large chunks of material—objects of essentially planetary mass—were orbiting in the inner solar system at the time that the terrestrial planets formed. The giant impact hypothesis envisions Earth being struck obliquely by an object approximately one-tenth Earth’s mass—a “bullet” about the size of Mars. This is very nearly the largest impact Earth could experience without being shattered.

Such an impact would disrupt much of Earth and eject a vast amount of material into space, releasing almost enough energy to break the planet apart. Computer simulations indicate that material totaling several percent of Earth’s mass could be ejected in such an impact. Most of this material would be from the stony mantles of Earth and the impacting body, not from their metal cores. This ejected rock vapor then cooled and formed a ring of material orbiting Earth. It was this ring that ultimately condensed into the Moon.

While we do not have any current way of showing that the giant impact hypothesis is the correct model of the Moon's origin, it does offer potential solutions to most of the major problems raised by the chemistry of the Moon. First, since the Moon's raw material is derived from the mantles of Earth and the projectile, the absence of metals is easily understood. Second, most of the volatile elements would have been lost during the high-temperature phase following the impact, explaining the lack of these materials on the Moon. Yet, by making the Moon primarily of terrestrial mantle material, it is also possible to understand similarities such as identical abundances of various oxygen isotopes.

Footnotes

- 2Remember from the Radiation and Spectra chapter that the term isotope means a different "version" of an element. Specifically, different isotopes of the same element have equal numbers of protons but different numbers of neutrons (as in carbon-12 versus carbon-14.)

This page titled 13.4: The Origin of the Moon is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9.5: The Origin of the Moon by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.5: Mercury

Learning Objectives

By the end of this section, you will be able to:

- Characterize the orbit of Mercury around the Sun
- Describe Mercury's structure and composition
- Explain the relationship between Mercury's orbit and rotation
- Describe the topography and features of Mercury's surface
- Summarize our ideas about the origin and evolution of Mercury

Mercury's Orbit

The planet Mercury is similar to the Moon in many ways. Like the Moon, it has no atmosphere, and its surface is heavily cratered. As described later in this chapter, it also shares with the Moon the likelihood of a violent birth.

Mercury is the nearest planet to the Sun, and, in accordance with Kepler's third law, it has the shortest period of revolution about the Sun (88 of our days) and the highest average orbital speed (48 kilometers per second). It is appropriately named for the fleet-footed messenger god of the Romans. Because Mercury remains close to the Sun, it can be difficult to pick out in the sky. As you might expect, it's best seen when its eccentric orbit takes it as far from the Sun as possible.

The semimajor axis of Mercury's orbit—that is, the planet's average distance from the Sun—is 58 million kilometers, or 0.39 AU. However, because its orbit has the high eccentricity of 0.206, Mercury's actual distance from the Sun varies from 46 million kilometers at perihelion to 70 million kilometers at aphelion (the ideas and terms that describe orbits were introduced in Orbits and Gravity).

Composition and Structure

Mercury's mass is one-eighteenth that of Earth, making it the smallest terrestrial planet. Mercury is the smallest planet (except for the dwarf planets), having a diameter of 4878 kilometers, less than half that of Earth. Mercury's density is 5.4 g/cm^3, much greater than the density of the Moon, indicating that the composition of those two objects differs substantially.

Mercury's composition is one of the most interesting things about it and makes it unique among the planets. Mercury's high density tells us that it must be composed largely of heavier materials such as metals. The most likely models for Mercury's interior suggest a metallic iron-nickel core amounting to 60% of the total mass, with the rest of the planet made up primarily of silicates. The core has a diameter of 3500 kilometers and extends out to within 700 kilometers of the surface. We could think of Mercury as a metal ball the size of the Moon surrounded by a rocky crust 700 kilometers thick (Figure 9.20). Unlike the Moon, Mercury does have a weak magnetic field. The existence of this field is consistent with the presence of a large metal core, and it suggests that at least part of the core must be liquid in order to generate the observed magnetic field.[3]

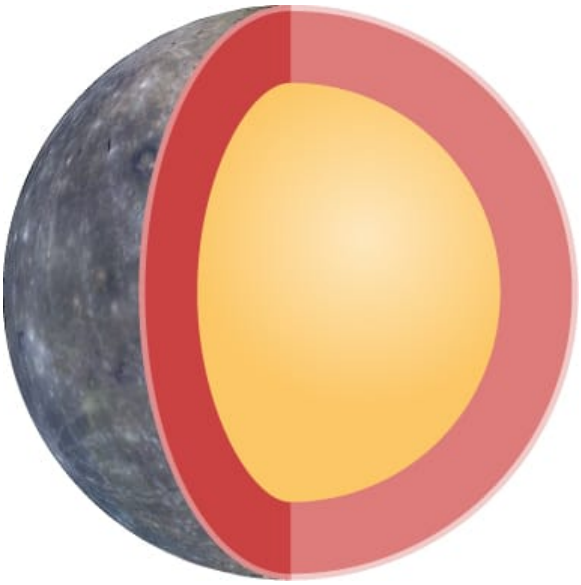


Figure 9.20 : Mercury's Internal Structure. The interior of Mercury is dominated by a metallic core about the same size as our Moon.

Example 9.1: Densities of Worlds

The average density of a body equals its mass divided by its volume. For a sphere, density is:

$$\text{density} =$$

Astronomers can measure both mass and radius accurately when a spacecraft flies by a body.

Using the information in this chapter, we can calculate the approximate average density of the Moon.

Solution

For a sphere,

$$\text{density} = \frac{\text{mass}}{\frac{4}{3}\pi R^3} = \frac{7.35 \times 10^{22}\,\text{kg}}{4.2 \times 5.2 \times 10^{18}\,\text{m}^3} = 3.4 \times 10^3$$ ■

Table 9.1 gives a value of 3.3g/cm^3, which is $3.3 \times 10^3\text{kg/m}^3$.

Exercise 13.5.1

Using the information in this chapter, calculate the average density of Mercury. Show your work. Does your calculation agree with the figure we give in this chapter?

Answer available at the end of this section.

That matches the value given in Table 9.1 when g/cm^3 is converted into kg/m^3.

Mercury's Strange Rotation

Visual studies of Mercury's indistinct surface markings were once thought to indicate that the planet kept one face to the Sun (as the Moon does to Earth). Thus, for many years, it was widely believed that Mercury's rotation period was equal to its revolution period of 88 days, making one side perpetually hot while the other was always cold.

Radar observations of Mercury in the mid-1960s, however, showed conclusively that Mercury does not keep one side fixed toward

the Sun. If a planet is turning, one side seems to be approaching Earth while the other is moving away from it. The resulting Doppler shift spreads or broadens the precise transmitted radar-wave frequency into a range of frequencies in the reflected signal (Figure 9.21). The degree of broadening provides an exact measurement of the rotation rate of the planet.

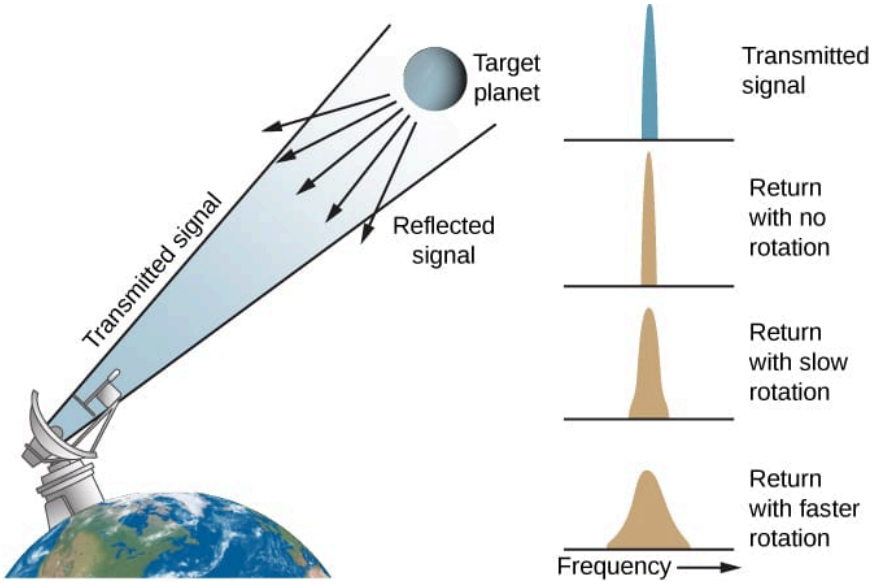


Figure 9.21 : Doppler Radar Measures Rotation. When a radar beam is reflected from a rotating planet, the motion of one side of the planet's disk toward us and the other side away from us causes Doppler shifts in the reflected signal. The effect is to cause both a redshift and a blueshift, widening the spread of frequencies in the radio beam.

Mercury's period of rotation (how long it takes to turn with respect to the distant stars) is 59 days, which is just two-thirds of the planet's period of revolution. Subsequently, astronomers found that a situation where the spin and the orbit of a planet (its year) are in a 2:3 ratio turns out to be stable. (See What a Difference a Day Makes for more on the effects of having such a long day on Mercury.)

Mercury, being close to the Sun, is very hot on its daylight side; but because it has no appreciable atmosphere, it gets surprisingly cold during the long nights. The temperature on the surface climbs to 700 K (430 °C) at noontime. After sunset, however, the temperature drops, reaching 100 K (–170 °C) just before dawn. (It is even colder in craters near the poles that receive no sunlight at all.) The range in temperature on Mercury is thus 600 K (or 600 °C), a greater difference than on any other planet.

Making Connections: What a Difference a Day Makes

Mercury rotates three times for each two orbits around the Sun. It is the only planet that exhibits this relationship between its spin and its orbit, and there are some interesting consequences for any observers who might someday be stationed on the surface of Mercury.

Here on Earth, we take for granted that days are much shorter than years. Therefore, the two astronomical ways of defining the local "day"—how long the planet takes to rotate and how long the Sun takes to return to the same position in the sky—are the same on Earth for most practical purposes. But this is not the case on Mercury. While Mercury rotates (spins once) in 59 Earth days, the time for the Sun to return to the same place in Mercury's sky turns out to be two Mercury years, or 176 Earth days. (Note that this result is not intuitively obvious, so don't be upset if you didn't come up with it.) Thus, if one day at noon a Mercury explorer suggests to her companion that they should meet at noon the next day, this could mean a very long time apart!

To make things even more interesting, recall that Mercury has an eccentric orbit, meaning that its distance from the Sun varies significantly during each mercurian year. By Kepler's law, the planet moves fastest in its orbit when closest to the Sun. Let's examine how this affects the way we would see the Sun in the sky during one 176-Earth-day cycle. We'll look at the situation as if we were standing on the surface of Mercury in the center of a giant basin that astronomers call Caloris (Figure 9.23).

At the location of Caloris, Mercury is most distant from the Sun at sunrise; this means the rising Sun looks smaller in the sky (although still more than twice the size it appears from Earth). As the Sun rises higher and higher, it looks bigger and bigger; Mercury is now getting closer to the Sun in its eccentric orbit. At the same time, the apparent motion of the Sun slows down as Mercury's faster motion in orbit begins to catch up with its rotation.

At noon, the Sun is now three times larger than it looks from Earth and hangs almost motionless in the sky. As the afternoon wears on, the Sun appears smaller and smaller, and moves faster and faster in the sky. At sunset, a full Mercury year (or 88 Earth days after sunrise), the Sun is back to its smallest apparent size as it dips out of sight. Then it takes another Mercury year before the Sun rises again. (By the way, sunrises and sunsets are much more sudden on Mercury, since there is no atmosphere to bend or scatter the rays of sunlight.)

Astronomers call locations like the Caloris Basin the "hot longitudes" on Mercury because the Sun is closest to the planet at noon, just when it is lingering overhead for many Earth days. This makes these areas the hottest places on Mercury.

We bring all this up not because the exact details of this scenario are so important but to illustrate how many of the things we take for granted on Earth are not the same on other worlds. As we've mentioned before, one of the best things about taking an astronomy class should be ridding you forever of any "Earth chauvinism" you might have. The way things are on our planet is just one of the many ways nature can arrange reality.

Link to Learning

The Day on Mercury visualization shows how Mercury's slow rotation leads to very long days and how this combines with its highly elliptical orbit to cause some unusual motions of the sun in the sky. Note the stick figure on the surface of Mercury; the Day/Night box shows what that stick figure would see in the sky.

The Surface of Mercury

The first close-up look at Mercury came in 1974, when the US spacecraft Mariner 10 passed 9500 kilometers from the surface of the planet and transmitted more than 2000 photographs to Earth, revealing details with a resolution down to 150 meters. Subsequently, the planet was mapped in great detail by the MESSENGER spacecraft, which was launched in 2004 and made multiple flybys of Earth, Venus, and Mercury before settling into orbit around Mercury in 2011. It ended its life in 2015, when it was commanded to crash into the surface of the planet.

Mercury's surface strongly resembles the Moon in appearance (Figure 9.22 and Figure 9.23). It is covered with thousands of craters and larger basins up to 1300 kilometers in diameter. Some of the brighter craters are rayed, like Tycho and Copernicus on the Moon, and many have central peaks. There are also *scarps* (cliffs) more than a kilometer high and hundreds of kilometers long, as well as ridges and plains.

MESSENGER instruments measured the surface composition and mapped past volcanic activity. One of its most important discoveries was the verification of water ice (first detected by radar) in craters near the poles, similar to the situation on the Moon, and the unexpected discovery of organic (carbon-rich) compounds mixed with the water ice.

Link to Learning

Scientists working with data from the MESSENGER mission put together a rotating globe of Mercury, in false color, showing some of the variations in the composition of the planet's surface. You can watch it spin.

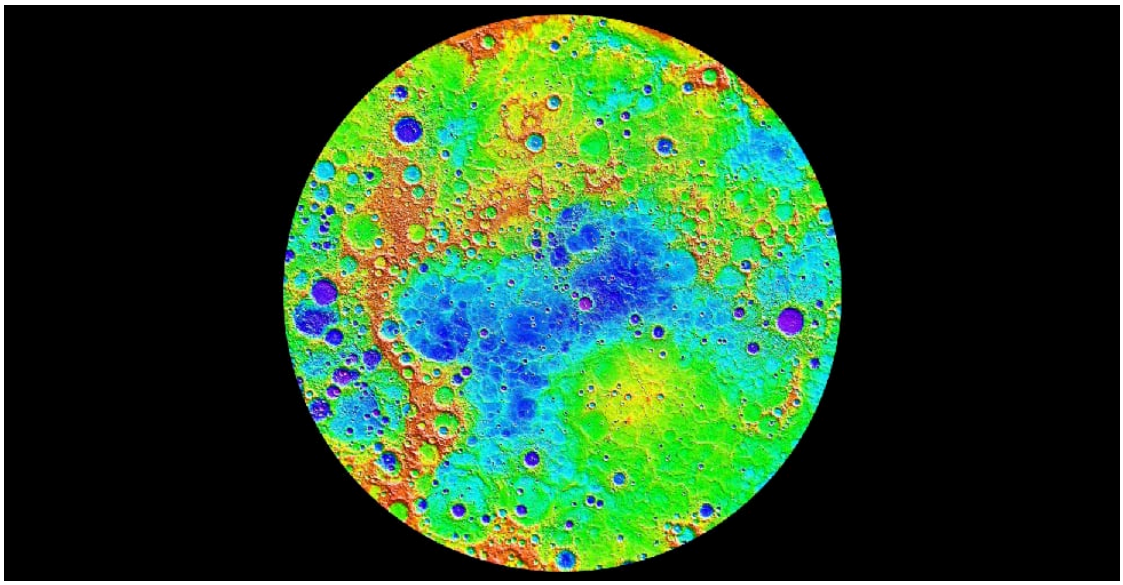


Figure 9.22 : Mercury's Topography. The topography of Mercury's northern hemisphere is mapped in great detail from MESSENGER data. The lowest regions are shown in purple and blue, and the highest regions are shown in red. The difference in elevation between the lowest and highest regions shown here is roughly 10 kilometers. The permanently shadowed low-lying craters near the north pole contain radar-bright water ice. (credit: modification of work by NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington)

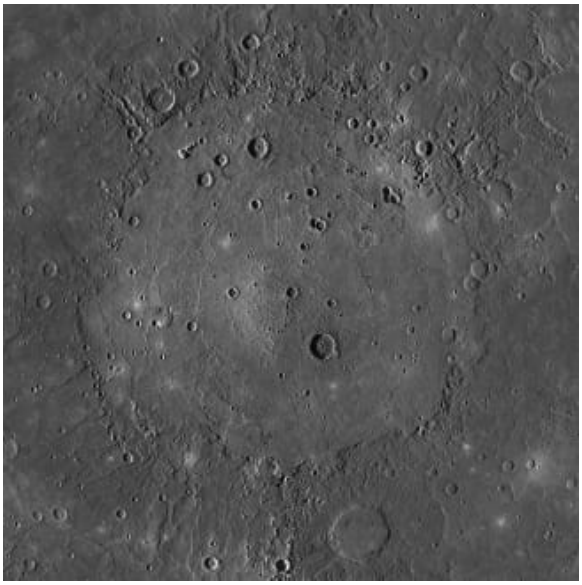


Figure 9.23 : Caloris Basin. This partially flooded impact basin is the largest known structural feature on Mercury. The smooth plains in the interior of the basin have an area of almost two million square kilometers. Compare this photo with Figure 9.11, the Orientale Basin on the Moon. (credit: NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington)

Most of the mercurian features have been named in honor of artists, writers, composers, and other contributors to the arts and humanities, in contrast with the scientists commemorated on the Moon. Among the named craters are Bach, Shakespeare, Tolstoy, Van Gogh, and Scott Joplin.

There is no evidence of plate tectonics on Mercury. However, the planet's distinctive long scarps can sometimes be seen cutting across craters; this means the scarps must have formed later than the craters (Figure 9.24). These long, curved cliffs appear to have their origin in the slight compression of Mercury's crust. Apparently, at some point in its history, the planet shrank, wrinkling the crust, and it must have done so after most of the craters on its surface had already formed.

If the standard cratering chronology applies to Mercury, this shrinkage must have taken place during the last 4 billion years and not during the solar system's early period of heavy bombardment.

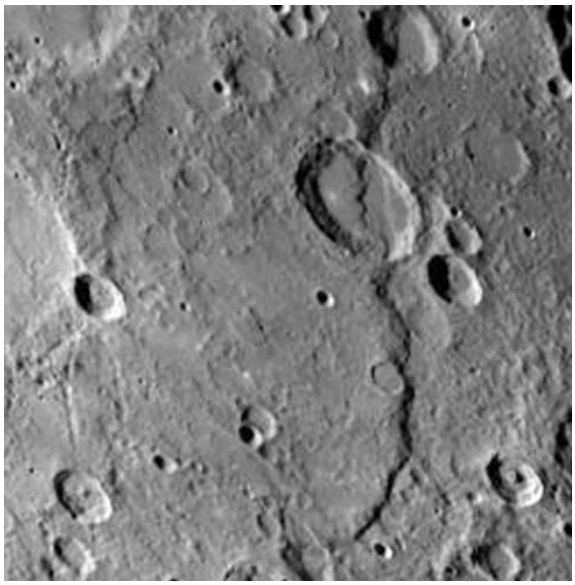


Figure 9.24 : Discovery Scarp on Mercury. This long cliff, nearly 1 kilometer high and more than 100 kilometers long, cuts across several craters. Astronomers conclude that the compression that made “wrinkles” like this in the planet's surface must have taken place after the craters were formed. (credit: modification of work by NASA/JPL/Northwestern University)

The Origin of Mercury

The problem with understanding how Mercury formed is the reverse of the problem posed by the composition of the Moon. We have seen that, unlike the Moon, Mercury is composed mostly of metal. However, astronomers think that Mercury should have formed with roughly the same ratio of metal to silicate as that found on Earth or Venus. How did it lose so much of its rocky material?

The most probable explanation for Mercury’s silicate loss may be similar to the explanation for the Moon’s lack of a metal core. Mercury is likely to have experienced several giant impacts very early in its youth, and one or more of these may have torn away a fraction of its mantle and crust, leaving a body dominated by its iron core. A recent alternative explanation suggests that, early on, heat from the young Sun evaporated much of the lighter, more volatile, silicates in the material that formed Mercury, leaving the planet with a majority of metals.

Link to Learning

You can follow some of NASA’s latest research on Mercury and see some helpful animations on the MESSENGER web page.

Today, astronomers recognize that the early solar system was a chaotic place, with the final stages of planet formation characterized by impacts of great violence. Some objects of planetary mass have been destroyed, whereas others could have fragmented and then re-formed, perhaps more than once. Both the Moon and Mercury, with their strange compositions, bear testimony to the catastrophes that must have characterized the solar system during its youth.

Footnotes

- 3Recall from the Radiation and Spectra chapter that magnetism is an effect of moving electric charges. In atoms of metals, the outer electrons are easier to dislodge and they can form a current when the metal is in liquid form and can flow.

Answers

Exercise 13.5.1

$$\text{density} = \frac{\text{mass}}{\frac{4}{3}\pi R^3} = \frac{3.3 \times 10^{23}\ \text{kg}}{4.2 \times 1.45 \times 10^{19}\ \text{m}^3} = 5.4\ \times\ 10^3$$

This page titled 13.5: Mercury is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 9.6: Mercury by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

13.6: Key Terms

highlands
: the lighter, heavily cratered regions of the Moon, which are generally several kilometers higher than the maria

mare
: (plural: maria) Latin for “sea;” the name applied to the dark, relatively smooth features that cover 17% of the Moon’s surface

This page titled 13.6: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

14: VENUS

14.1: The Nearest Planets - An Overview
14.2: The Geology of Venus
14.3: The Massive Atmosphere of Venus

This page titled 14: Venus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

14.1: The Nearest Planets - An Overview

Learning Objectives

By the end of this section, you will be able to:

- Explain why it's difficult to learn about Venus from Earth-based observation alone
- Describe the history of our interest in Mars before the Space Age
- Compare the basic physical properties of Earth, Mars, and Venus, including their orbits

As you might expect from close neighbors, Mars and Venus are among the brightest objects in the night sky. The average distance of Mars from the Sun is 227 million kilometers (1.52 AU), or about half again as far from the Sun as Earth. Venus' orbit is very nearly circular, at a distance of 108 million kilometers (0.72 AU) from the Sun. Like Mercury, Venus sometimes appears as an "evening star" and sometimes as a "morning star." Venus approaches Earth more closely than does any other planet: at its nearest, it is only 40 million kilometers from us. The closest Mars ever gets to Earth is about 56 million kilometers.

Appearance

Venus appears very bright, and even a small telescope reveals that it goes through phases like the Moon. Galileo discovered that Venus displays a full range of phases, and he used this as an argument to show that Venus must circle the Sun and not Earth. The planet's actual surface is not visible because it is shrouded by dense clouds that reflect about 70% of the sunlight that falls on them, frustrating efforts to study the underlying surface, even with cameras in orbit around the planet (Figure 10.2).

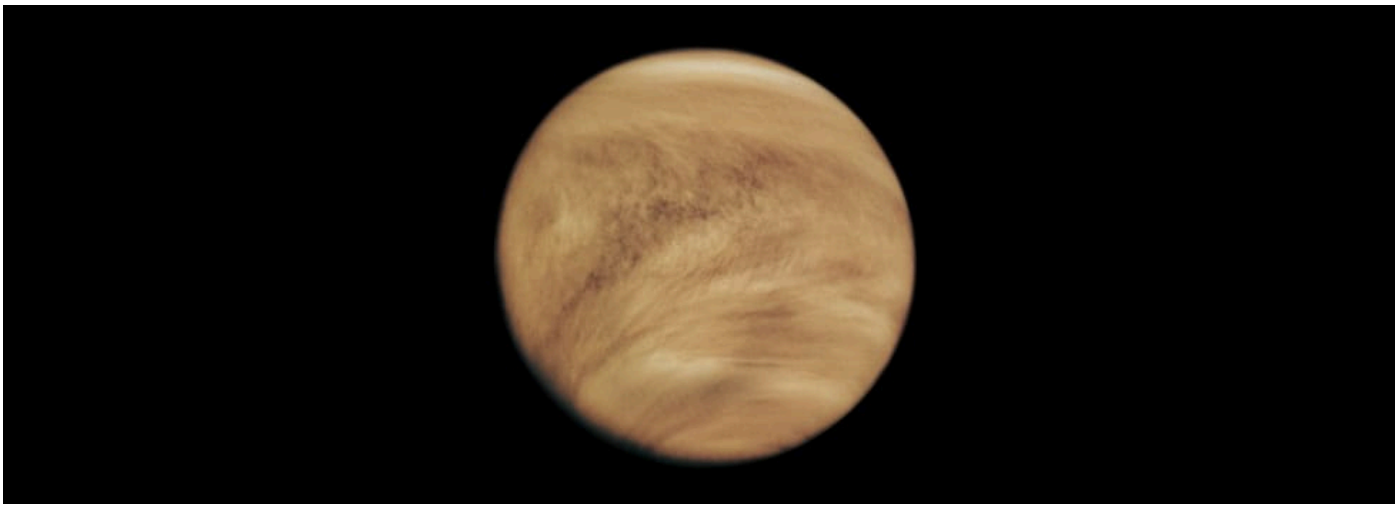


Figure 10.2 : Venus as Photographed by the Pioneer Venus Orbiter. This ultraviolet image shows an upper-atmosphere cloud structure that would be invisible at visible wavelengths. Note that there is not even a glimpse of the planet's surface. (credit: modification of work by NASA)

In contrast, Mars is more tantalizing as seen through a telescope (Figure 10.3). The planet is distinctly red, due (as we now know) to the presence of iron oxides in its soil. This color may account for its association with war (and blood) in the legends of early cultures. The best resolution obtainable from telescopes on the ground is about 100 kilometers, or about the same as what we can see on the Moon with the unaided eye. At this resolution, no hint of topographic structure can be detected: no mountains, no valleys, not even impact craters. On the other hand, bright polar ice caps can be seen easily, together with dusky surface markings that sometimes change in outline and intensity from season to season.

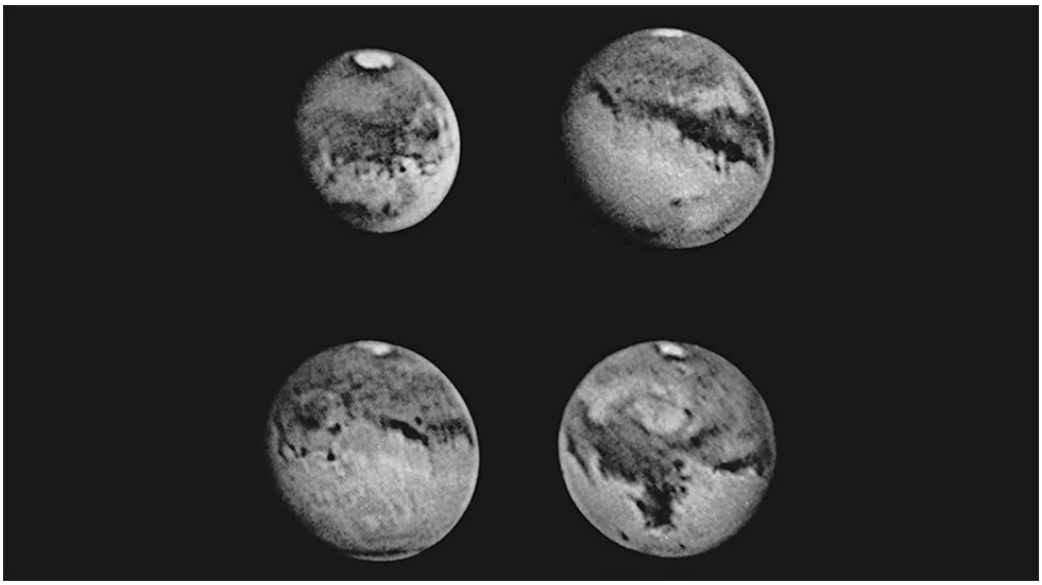


Figure 10.3 : Mars as Seen from Earth's Surface. These are among the best Earth-based photos of Mars, taken in 1988 when the planet was exceptionally close to Earth. The polar caps and dark surface markings are evident, but not topographic features. (credit: modification of work by Steve Larson, Lunar and Planetary Laboratory, University of Arizona)

For a few decades around the turn of the twentieth century, some astronomers believed that they saw evidence of an intelligent civilization on Mars. The controversy began in 1877, when Italian astronomer Giovanni Schiaparelli (1835–1910) announced that he could see long, faint, straight lines on Mars that he called *canale*, or channels. In English-speaking countries, the term was mistakenly translated as "canals," implying an artificial origin.

Even before Schiaparelli's observations, astronomers had watched the bright polar caps change size with the seasons and had seen variations in the dark surface features. With a little imagination, it was not difficult to picture the canals as long fields of crops bordering irrigation ditches that brought water from the melting polar ice to the parched deserts of the red planet. (They assumed the polar caps were composed of water ice, which isn't exactly true, as we will see shortly.)

Until his death in 1916, the most effective proponent of intelligent life on Mars was Percival Lowell, a self-made American astronomer and member of the wealthy Lowell family of Boston (see the feature box on Percival Lowell: Dreaming of an Inhabited Mars). A skilled author and speaker, Lowell made what seemed to the public to be a convincing case for intelligent Martians, who had constructed the huge canals to preserve their civilization in the face of a deteriorating climate (Figure 10.4).

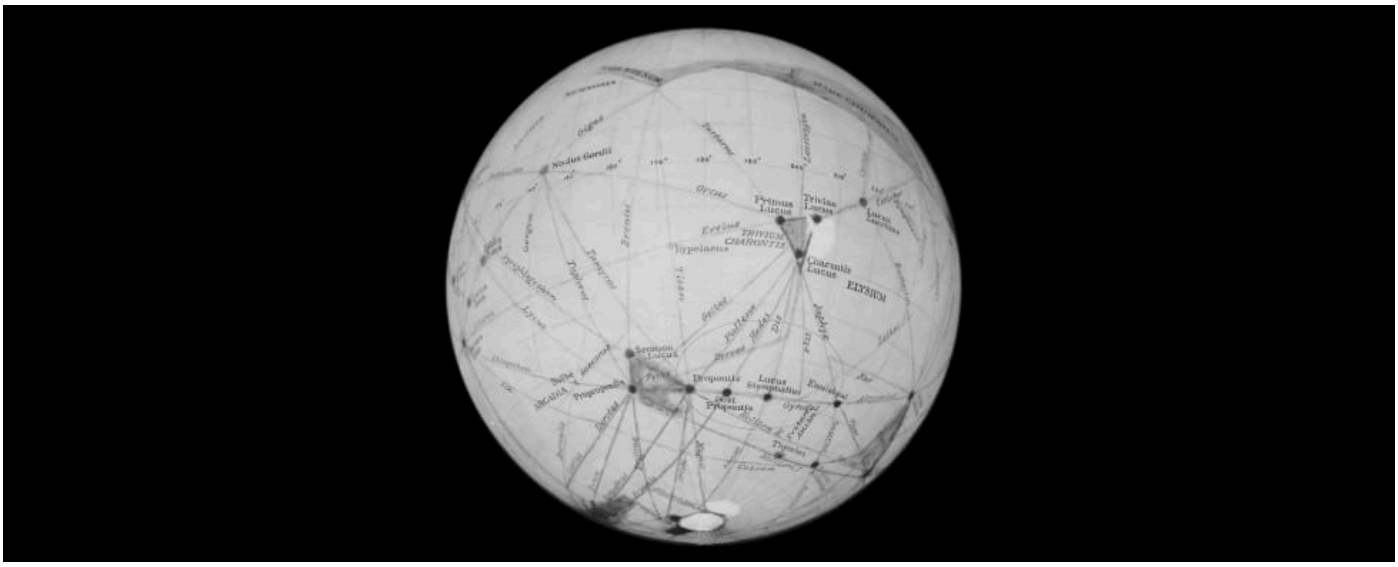


Figure 10.4 : Lowell's Mars Globe. One of the remarkable globes of Mars prepared by Percival Lowell, showing a network of dozens of canals, oases, and triangular water reservoirs that he claimed were visible on the red planet.

The argument for a race of intelligent Martians, however, hinged on the reality of the canals, a matter that remained in serious dispute among astronomers. The canal markings were always difficult to study, glimpsed only occasionally because atmospheric conditions caused the tiny image of Mars to shimmer in the telescope. Lowell saw canals everywhere (even a few on Venus), but many other observers could not see them at all and remained unconvinced of their existence. When telescopes larger than Lowell's failed to confirm the presence of canals, the skeptics felt vindicated. Now it is generally accepted that the straight lines were an optical illusion, the result of the human mind's tendency to see order in random features that are glimpsed dimly at the limits of the eye's resolution. When we see small, dim dots of surface markings, our minds tend to connect those dots into straight lines.

Voyagers in Astronomy: Percival Lowell: Dreaming of an Inhabited Mars

Percival Lowell was born into the well-to-do Massachusetts family about whom John Bossidy made the famous toast:

> And this is good old Boston,
> The home of the bean and the cod,
> Where the Lowells talk to the Cabots
> And the Cabots talk only to God.

Percival's brother Lawrence became president of Harvard University, and his sister, Amy, became a distinguished poet. Percival was already interested in astronomy as a boy: he made observations of Mars at age 13. His undergraduate thesis at Harvard dealt with the origin of the solar system, but he did not pursue this interest immediately. Instead, he entered the family business and traveled extensively in Asia. In 1892, however, he decided to dedicate himself to carrying on Schiaparelli's work and solving the mysteries of the martian canals.

In 1894, with the help of astronomers at Harvard but using his own funds, Lowell built an observatory on a high plateau in Flagstaff, Arizona, where he hoped the seeing would be clear enough to show him Mars in unprecedented detail. He and his assistants quickly accumulated a tremendous number of drawings and maps, purporting to show a vast network of martian canals (see Figure 10.4). He elaborated his ideas about the inhabitants of the red planet in several books, including *Mars* (1895) and *Mars and Its Canals* (1906), and in hundreds of articles and speeches.

As Lowell put it,

> A mind of no mean order would seem to have presided over the system we see—a mind certainly of considerably more comprehensiveness than that which presides over the various departments of our own public works. Party politics, at all events, have had no part in them; for the system is planet-wide. . . . Certainly what we see hints at the existence of beings who are in advance of, not behind us, in the journey of life.

Lowell's views captured the public imagination and inspired many novels and stories, the most famous of which was H. G. Wells' *War of the Worlds* (1897). In this famous "invasion" novel, the thirsty inhabitants of a dying planet Mars (based entirely on Lowell's ideas) come to conquer Earth with advanced technology.

Although the Lowell Observatory first became famous for its work on the martian canals, both Lowell and the observatory eventually turned to other projects as well. He became interested in the search for a ninth (and then undiscovered) planet in the solar system. In 1930, Pluto was found at the Lowell Observatory, and it is not a coincidence that the name selected for the new planet starts with Lowell's initials. It was also at the Lowell Observatory that the first measurements were made of the great speed at which galaxies are moving away from us, observations that would ultimately lead to our modern view of an expanding universe.

Lowell (Figure 10.5) continued to live at his observatory, marrying at age 53 and publishing extensively. He relished the debate his claims about Mars caused far more than the astronomers on the other side, who often complained that Lowell's work was making planetary astronomy a less respectable field. At the same time, the public fascination with the planets fueled by Lowell's work (and its interpreters) may, several generations later, have helped fan support for the space program and the many missions whose results grace the pages of our text.



Figure 10.5 : Percival Lowell (1855–1916). This 1914 photograph shows Percival Lowell observing Venus with his 24-inch telescope at Flagstaff, Arizona.

Link to Learning

In October 1938, the Mercury Theater of the Air on radio dramatized *The War of the Worlds* as a series of radio news reports. This broadcast scared many people into thinking that Lowell's Martians were really invading New Jersey, and caused something of a panic. You can listen to the original radio broadcast if you scroll down to "War of the Worlds."

Rotation of the Planets

Astronomers have determined the rotation period of Mars with great accuracy by watching the motion of permanent surface markings; its sidereal day is 24 hours 37 minutes 23 seconds, just a little longer than the rotation period of Earth. This high precision is not obtained by watching Mars for a single rotation, but by noting how many turns it makes over a long period of time. Good observations of Mars date back more than 200 years, a period during which tens of thousands of martian days have passed. As a result, the rotation period can be calculated to within a few hundredths of a second.

The rotational axis of Mars has a tilt of about 25°, similar to the tilt of Earth's axis. Thus, Mars experiences seasons very much like those on Earth. Because of the longer martian year (almost two Earth years), however, each season there lasts about six of our months.

The situation with Venus is different. Since no surface detail can be seen through Venus' clouds, its rotation period can be found only by bouncing radar signals off the planet (as explained for Mercury in the Cratered Worlds chapter). The first radar observations of Venus' rotation were made in the early 1960s. Later, topographical surface features were identified on the planet that showed up in the reflected radar signals. The rotation period of Venus, precisely determined from the motion of such "radar features" across its disk, is 243 days. Even more surprising than how *long* Venus takes to rotate is the fact that it spins in a backward or retrograde direction (east to west).

Stop for a moment and think about how odd this slow rotation makes the calendar on Venus. The planet takes 225 Earth days to orbit the Sun and 243 Earth days to spin on its axis. So the day on Venus (as defined by its spinning once) is longer than the year! As a result, the time the Sun takes to return to the same place in Venus' sky—another way we might define the meaning of a day—turns out to be 117 Earth days. (If you say "See you tomorrow" on Venus, you'll have a long time to wait.) Although we do not know the reason for Venus' slow backward rotation, we can guess that it may have suffered one or more extremely powerful collisions during the formation process of the solar system.

Basic Properties of Venus and Mars

Before discussing each planet individually, let us compare some of their basic properties with each other and with Earth (Table 10.1). Venus is in many ways Earth's twin, with a mass 0.82 times the mass of Earth and an almost identical density. The average amount of geological activity has been also relatively high, almost as high as on Earth. On the other hand, with a surface pressure nearly 100 times greater than ours, Venus' atmosphere is not at all like that of Earth. The surface of Venus is also remarkably hot, with a temperature of 730 K (over 850 °F), hotter than the self-cleaning cycle of your oven. One of the major challenges presented by Venus is to understand why the atmosphere and surface environment of this twin have diverged so sharply from those of our own planet.

Table 10.1: Properties of Earth, Venus, and Mars

Property	Earth	Venus	Mars
Semimajor axis (AU)	1.00	0.72	1.52
Period (year)	1.00	0.61	1.88
Mass (Earth = 1)	1.00	0.82	0.11
Diameter (km)	12,756	12,102	6,790
Density (g/cm^3)	5.5	5.3	3.9
Surface gravity (Earth = 1)	1.00	0.91	0.38
Escape velocity (km/s)	11.2	10.4	5.0
Rotation period (hours or days)	23.9 h	243 d	24.6 h
Surface area (Earth = 1)	1.00	0.90	0.28
Atmospheric pressure (bar)	1.00	90	0.007

Mars, by contrast, is rather small, with a mass only 0.11 times the mass of Earth. It is larger than either the Moon or Mercury, however, and, unlike them, it retains a thin atmosphere. Mars is also large enough to have supported considerable geological activity in the distant past. But the most fascinating thing about Mars is that long ago it probably had a thick atmosphere and seas of liquid water—the conditions we associate with development of life. There is even a chance that some form of life persists today in protected environments below the martian surface.

This page titled 14.1: The Nearest Planets - An Overview is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.2: The Nearest Planets - An Overview by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

14.2: The Geology of Venus

Learning Objectives

By the end of this section, you will be able to:

- Describe the general features of the surface of Venus
- Explain what the study of craters on Venus tells us about the age of its surface
- Compare tectonic activity and volcanoes on Venus with those of Earth
- Explain why the surface of Venus is inhospitable to human life

Since Venus has about the same size and composition as Earth, we might expect its geology to be similar. This is partly true, but Venus does not exhibit the same kind of *plate tectonics* as Earth, and we will see that its lack of erosion results in a very different surface appearance.

Spacecraft Exploration of Venus

Nearly 50 spacecraft have been launched to Venus, but only about half were successful. Although the 1962 US Mariner 2 flyby was the first, the Soviet Union launched most of the subsequent missions to Venus. In 1970, Venera 7 became the first probe to land and broadcast data from the surface of Venus. It operated for 23 minutes before succumbing to the high surface temperature. Additional Venera probes and landers followed, photographing the surface and analyzing the atmosphere and soil.

To understand the geology of Venus, however, we needed to make a global study of its surface, a task made very difficult by the perpetual cloud layers surrounding the planet. The problem resembles the challenge facing air traffic controllers at an airport, when the weather is so cloudy or smoggy that they can't locate the incoming planes visually. The solution is similar in both cases: use a radar instrument to probe through the obscuring layer.

The first global radar map was made by the US Pioneer Venus orbiter in the late 1970s, followed by better maps from the twin Soviet Venera 15 and 16 radar orbiters in the early 1980s. However, most of our information on the geology of Venus is derived from the US *Magellan* spacecraft, which mapped Venus with a powerful *imaging radar*. *Magellan* produced images with a resolution of 100 meters, much better than that of previous missions, yielding our first detailed look at the surface of our sister planet (Figure 10.6). (The *Magellan* spacecraft returned more data to Earth than all previous planetary missions combined; each 100 minutes of data transmission from the spacecraft provided enough information, if translated into characters, to fill two 30-volume encyclopedias.)

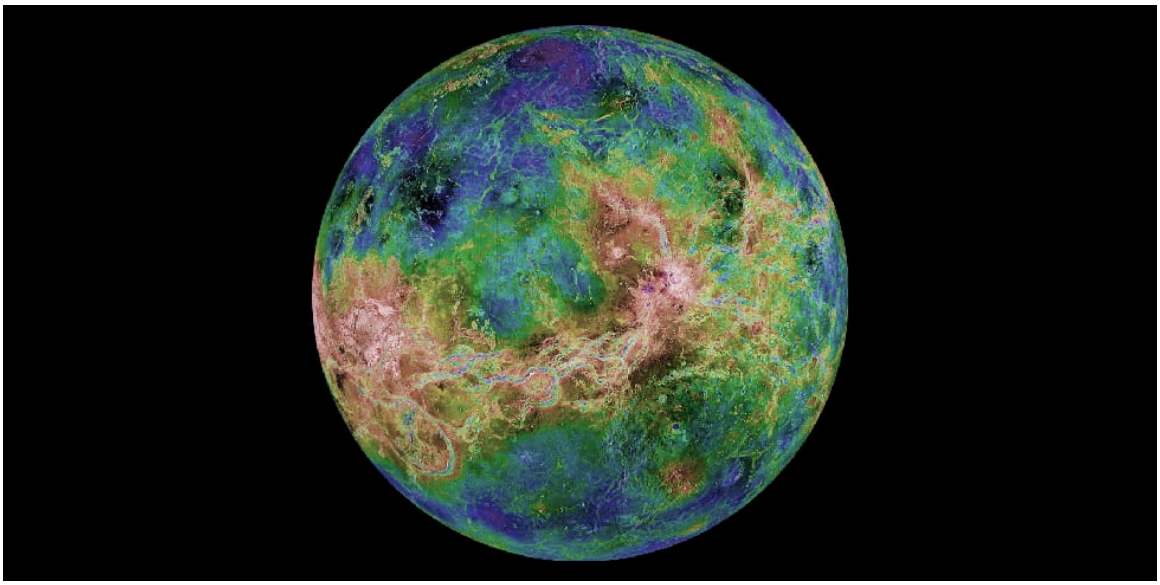


Figure 10.6 : Radar Map of Venus. This composite image has a resolution of about 3 kilometers. Colors have been added to indicate elevation, with blue meaning low and brown and white high. The large continent Aphrodite stretches around the equator, where the bright (therefore rough) surface has been deformed by tectonic forces in the crust of Venus. (credit: modification of work by NASA/JPL/USGS)

Consider for a moment how good *Magellan*'s resolution of 100 meters really is. It means the radar images from Venus can show anything on the surface larger than a football field. Suddenly, a whole host of topographic features on Venus became accessible to our view. As you look at the radar images throughout this chapter, bear in mind that these are constructed from radar reflections, not from visible-light photographs. For example, bright features on these radar images are an indication of rough terrain, whereas darker regions are smoother.

Probing Through the Clouds of Venus

The radar maps of Venus reveal a planet that looks much the way Earth might look if our planet's surface were not constantly being changed by erosion and deposition of sediment. Because there is no water or ice on Venus and the surface wind speeds are low, almost nothing obscures or erases the complex geological features produced by the movements of Venus' crust, by volcanic eruptions, and by impact craters. Having finally penetrated below the clouds of Venus, we find its surface to be naked, revealing the history of hundreds of millions of years of geological activity.

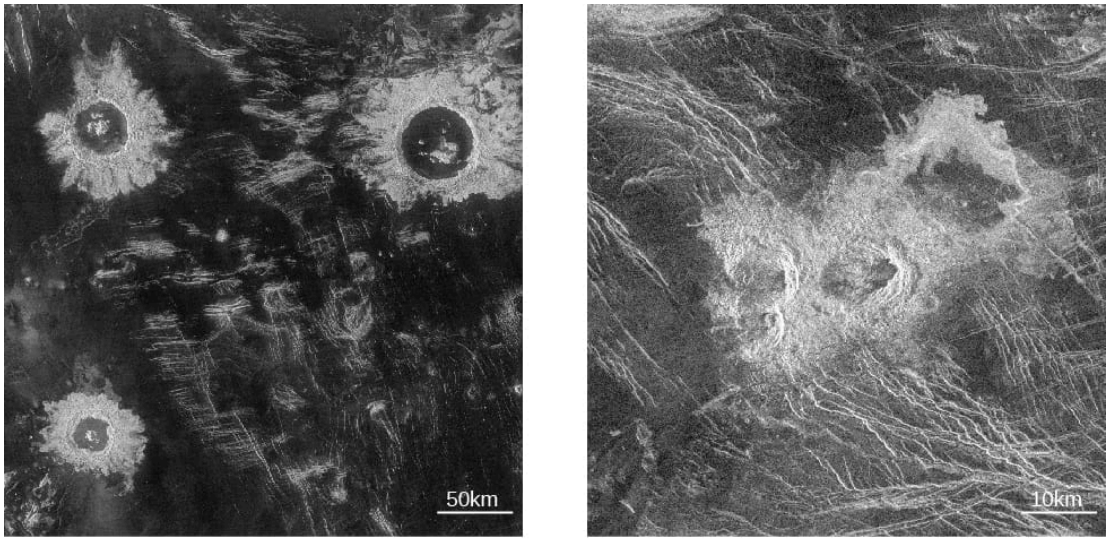
About 75% of the surface of Venus consists of lowland lava plains. Superficially, these plains resemble the basaltic ocean basins of Earth, but they were not produced in quite the same way. There is no evidence of subduction zones on Venus, indicating that, unlike Earth, this planet never experienced plate tectonics. Although *convection* (the rising of hot materials) in its mantle generated great stresses in the crust of Venus, they did not start large continental plates moving. The formation of the lava plains of Venus more nearly resembles that of the lunar maria. Both were the result of widespread lava eruptions without the crustal spreading associated with plate tectonics.

Rising above the lowland lava plains are two full-scale continents of mountainous terrain. The largest continent on Venus, called Aphrodite, is about the size of Africa (you can see it stand out in Figure 10.6). Aphrodite stretches along the equator for about one-third of the way around the planet. Next in size is the northern highland region Ishtar, which is about the size of Australia. Ishtar contains the highest region on the planet, the Maxwell Mountains, which rise 11 kilometers above the surrounding lowlands. (The Maxwell Mountains are the only feature on Venus named after a man. They commemorate James Clerk Maxwell, whose theory of electromagnetism led to the invention of radar. All other features are named for women, either from history or mythology.)

Craters and the Age of the Venus Surface

One of the first questions astronomers addressed with the high-resolution *Magellan* images was the age of the surface of Venus. Remember that the age of a planetary surface is rarely the age of the world it is on. A young age merely implies an active geology in that location. Such ages can be derived from counting impact craters. Figure 10.7 is an example of what these craters look like on the Venus radar images. The more densely cratered the surface, the greater its age. The largest crater on Venus (called Mead)

is 275 kilometers in diameter, slightly larger than the largest known terrestrial crater (Chicxulub), but much smaller than the lunar impact basins.



(a) (b)

Figure 10.7 : Impact Craters on Venus. (a) These large impact craters are in the Lavinia region of Venus. Because they are rough, the crater rims and ejecta appear brighter in these radar images than do the smoother surrounding lava plains. The largest of these craters has a diameter of 50 kilometers. (b) This small, complex crater is named after writer Gertrude Stein. The triple impact was caused by the breaking apart of the incoming asteroid during its passage through the thick atmosphere of Venus. The projectile had an initial diameter of between 1 and 2 kilometers. (credit a: modification of work by NASA/JPL; credit b: modification of work by NASA/JPL)

You might think that the thick atmosphere of Venus would protect the surface from impacts, burning up the projectiles long before they could reach the surface. But this is the case for only smaller projectiles. Crater statistics show very few craters less than 10 kilometers in diameter, indicating that projectiles smaller than about 1 kilometer (the size that typically produces a 10-kilometer crater) were stopped by the atmosphere. Those craters with diameters from 10 to 30 kilometers are frequently distorted or multiple, apparently because the incoming projectile broke apart in the atmosphere before it could strike the ground as shown in the Stein crater in Figure 10.7. If we limit ourselves to impacts that produce craters with diameters of 30 kilometers or larger, however, then crater counts are as useful on Venus for measuring surface age as they are on airless bodies such as the Moon.

The large craters in the venusian plains indicate an average surface age that is only between 300 and 600 million years. These results indicate that Venus is indeed a planet with persistent geological activity, intermediate between that of Earth's ocean basins (which are younger and more active) and that of its continents (which are older and less active).

Almost all of the large craters on Venus look fresh, with little degradation or filling in by either lava or windblown dust. This is one way we know that the rates of erosion or sediment deposition are very low. We have the impression that relatively little has happened since the venusian plains were last resurfaced by large-scale volcanic activity. Apparently Venus experienced some sort of planet-wide volcanic convulsion between 300 and 600 million years ago, a mysterious event that is unlike anything in terrestrial history.

Volcanoes on Venus

Like Earth, Venus is a planet that has experienced widespread volcanism. In the lowland plains, volcanic eruptions are the principal way the surface is renewed, with large flows of highly fluid lava destroying old craters and generating a fresh surface. In addition, numerous younger volcanic mountains and other structures are associated with surface hot spots—places where convection in the planet's mantle transports the interior heat to the surface.

The largest individual volcano on Venus, called Sif Mons, is about 500 kilometers across and 3 kilometers high—broader but lower than the Hawaiian volcano Mauna Loa. At its top is a volcanic crater, or *caldera*, about 40 kilometers across, and its slopes show individual lava flows up to 500 kilometers long. Thousands of smaller volcanoes dot the surface, down to the limit of visibility of the *Magellan* images, which correspond to cones or domes about the size of a shopping mall parking lot. Most of these seem similar to terrestrial volcanoes. Other volcanoes have unusual shapes, such as the "pancake domes" illustrated in Figure 10.8.

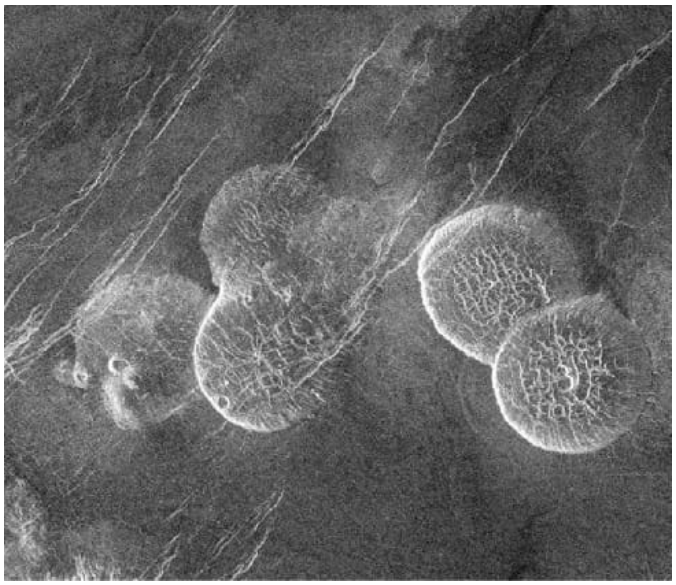


Figure 10.8 : Pancake-Shaped Volcanoes on Venus. These remarkable circular domes, each about 25 kilometers across and about 2 kilometers tall, are the result of eruptions of highly viscous (sludgy) lava that spreads out evenly in all directions. (credit: modification of work by NASA/JPL)

All of the volcanism is the result of eruption of lava onto the surface of the planet. But the hot lava rising from the interior of a planet does not always make it to the surface. On both Earth and Venus, this upwelling lava can collect to produce bulges in the crust. Many of the granite mountain ranges on Earth, such as the Sierra Nevada in California, involve such subsurface volcanism. These bulges are common on Venus, where they produce large circular or oval features called *coronae* (singular: corona) (Figure 10.9).

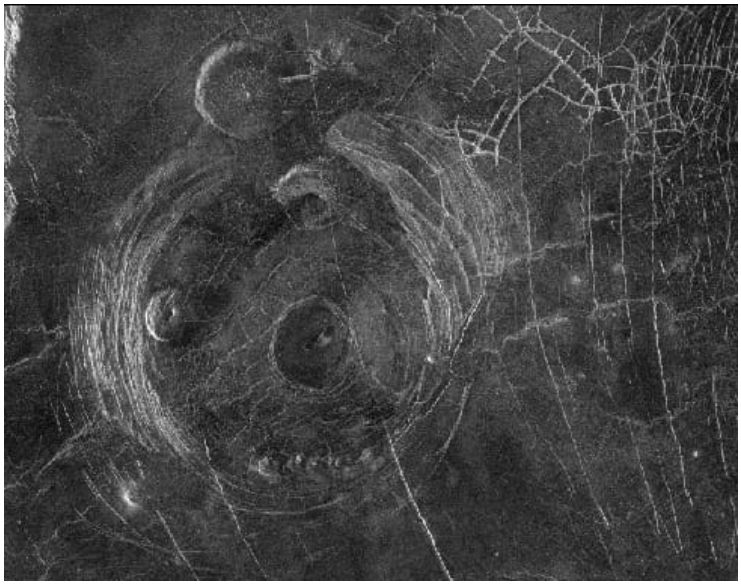


Figure 10.9 : The “Miss Piggy” Corona. Fotla Corona is located in the plains to the south of Aphrodite Terra. Curved fracture patterns show where the material beneath has put stress on the surface. A number of pancake and dome volcanoes are also visible. Fotla was a Celtic fertility goddess. Some students see a resemblance between this corona and Miss Piggy of the Muppets (her left ear, at the top of the picture, is the pancake volcano in the upper center of the image). (credit: NASA/JPL)

Tectonic Activity

Convection currents of molten material in the mantle of Venus push and stretch the crust. Such forces are called tectonic, and the geological features that result from these forces are called *tectonic features*. On Venus’ lowland plains, tectonic forces have broken the lava surface to create remarkable patterns of ridges and cracks (Figure 10.10). In a few places, the crust has even

torn apart to generate rift valleys. The circular features associated with coronae are tectonic ridges and cracks, and most of the mountains of Venus also owe their existence to tectonic forces.

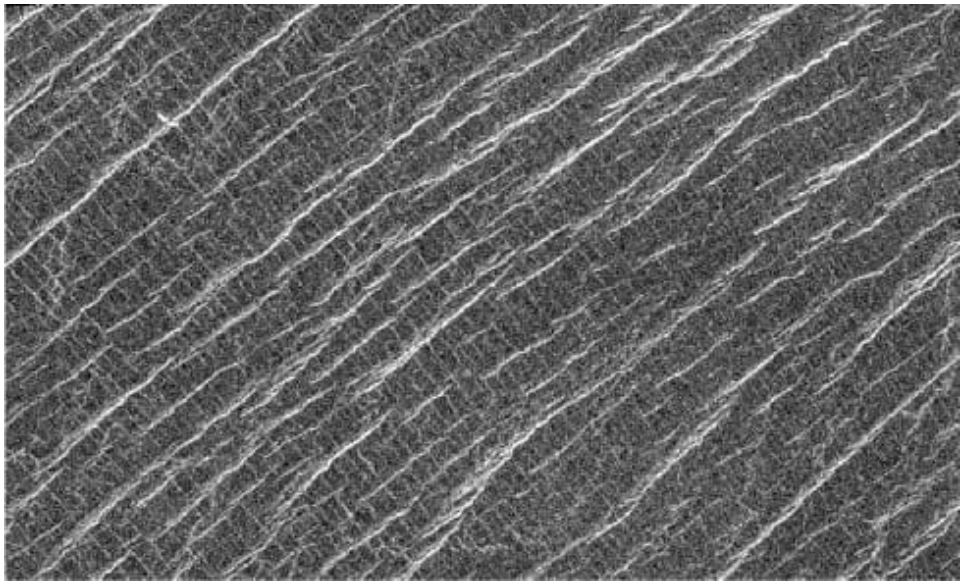


Figure 10.10 : Ridges and Cracks. This region of the Lakshmi Plains on Venus has been fractured by tectonic forces to produce a cross-hatched grid of cracks and ridges. Be sure to notice the fainter linear features that run perpendicular to the brighter ones. As this is a radar image, the brightness of the ridges indicates their relative height. This image shows a region about 80 kilometers wide and 37 kilometers high. Lakshmi is a Hindu goddess of prosperity. (credit: modification of work by Magellan Team, JPL, NASA)

The Ishtar continent, which has the highest elevations on Venus, is the most dramatic product of these tectonic forces. Ishtar and its tall Maxwell Mountains resemble the Tibetan Plateau and Himalayan Mountains on Earth. Both are the product of compression of the crust, and both are maintained by the continuing forces of mantle convection.

On Venus' Surface

The successful Venera landers of the 1970s found themselves on an extraordinarily inhospitable planet, with a surface pressure of 90 bars and a temperature hot enough to melt lead and zinc. Despite these unpleasant conditions, the spacecraft were able to photograph their surroundings and collect surface samples for chemical analysis before their instruments gave out. The diffuse sunlight striking the surface was tinted red by the clouds, and the illumination level was equivalent to a heavy overcast on Earth.

The probes found that the rock in the landing areas is igneous, primarily basalts. Examples of the Venera photographs are shown in Figure 10.11. Each picture shows a flat, desolate landscape with a variety of rocks, some of which may be ejecta from impacts. Other areas show flat, layered lava flows. There have been no further landings on Venus since the 1970s. However, NASA has plans for a radar orbiter with much higher resolution than Magellan, that will focus on the complex structures in the Aphrodite region. Some geologists think this area shows evidence of geologic plates, with past activity analogous to the plate tectonics on Earth.

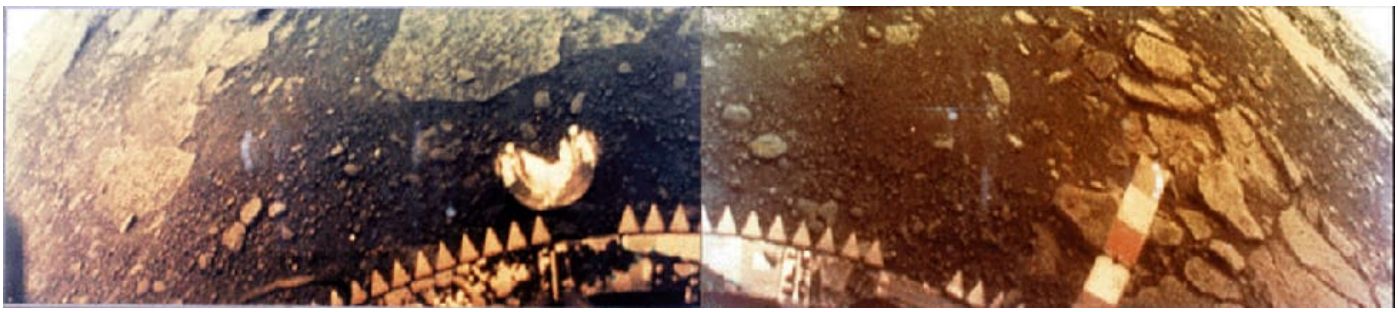


Figure 10.11 : Surface of Venus. These views of the surface of Venus are from the Venera 13 spacecraft. Everything is orange because the thick atmosphere of Venus absorbs the bluer colors of light. The horizon is visible in the upper corner of each image. (credit: NASA)

This page titled 14.2: The Geology of Venus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.3: The Geology of Venus by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

14.3: The Massive Atmosphere of Venus

Learning Objectives

By the end of this section, you will be able to:

- Describe the general composition and structure of the atmosphere on Venus
- Explain how the greenhouse effect has led to high temperatures on Venus

The thick atmosphere of Venus produces the high surface temperature and shrouds the surface in a perpetual red twilight. Sunlight does not penetrate directly through the heavy clouds, but the surface is fairly well lit by diffuse light (about the same as the light on Earth under a heavy overcast). The weather at the bottom of this deep atmosphere remains perpetually hot and dry, with calm winds. Because of the heavy blanket of clouds and atmosphere, one spot on the surface of Venus is similar to any other as far as weather is concerned.

Composition and Structure of the Atmosphere

The most abundant gas on Venus is carbon dioxide (CO_2), which accounts for 96% of the atmosphere. The second most common gas is nitrogen. The predominance of carbon dioxide over nitrogen is not surprising when you recall that Earth's atmosphere would also be mostly carbon dioxide if this gas were not locked up in marine sediments (see the discussion of Earth's atmosphere in Earth as a Planet).

Table 10.2 compares the compositions of the atmospheres of Venus, Mars, and Earth. Expressed in this way, as percentages, the proportions of the major gases are very similar for Venus and Mars, but in total quantity, their atmospheres are dramatically different. With its surface pressure of 90 bars, the venusian atmosphere is more than 10,000 times more massive than its martian counterpart. Overall, the atmosphere of Venus is very dry; the absence of water is one of the important ways that Venus differs from Earth.

Table 10.2: Atmospheric Composition of Earth, Venus, and Mars

Gas	Earth	Venus	Mars
Carbon dioxide (CO_2)	0.03%	96%	95.3%
Nitrogen (N_2)	78.1%	3.5%	2.7%
Argon (Ar)	0.93%	0.006%	1.6%
Oxygen (O_2)	21.0%	0.003%	0.15%
Neon (Ne)	0.002%	0.001%	0.0003%

The atmosphere of Venus has a huge troposphere (region of convection) that extends up to at least 50 kilometers above the surface (Figure 10.12). Within the troposphere, the gas is heated from below and circulates slowly, rising near the equator and descending over the poles. Being at the base of the atmosphere of Venus is something like being a kilometer or more below the ocean surface on Earth. There, the mass of water evens out temperature variations and results in a uniform environment—the same effect the thick atmosphere has on Venus.

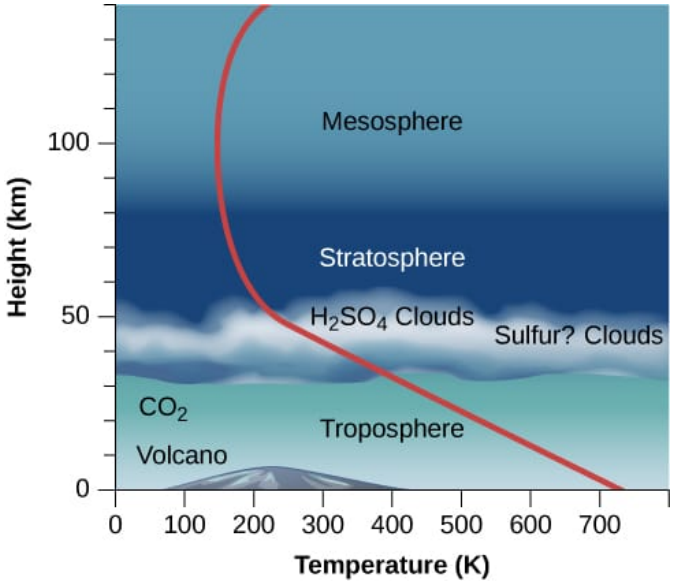


Figure 10.12 : Venus' Atmosphere. The layers of the massive atmosphere of Venus shown here are based on data from the Pioneer and Venera entry probes. Height is measured along the left axis, the bottom scale shows temperature, and the red line allows you to read off the temperature at each height. Notice how steeply the temperature rises below the clouds, thanks to the planet's huge greenhouse effect.

In the upper troposphere, between 30 and 60 kilometers above the surface, a thick cloud layer is composed primarily of sulfuric acid droplets. Sulfuric acid (H_2SO_4) is formed from the chemical combination of sulfur dioxide (SO_2) and water (H_2O). In the atmosphere of Earth, sulfur dioxide is one of the primary gases emitted by volcanoes, but it is quickly diluted and washed out by rainfall. In the dry atmosphere of Venus, this unpleasant substance is apparently stable. Below 30 kilometers, the Venus atmosphere is clear of clouds.

Surface Temperature on Venus

The high surface temperature of Venus was discovered by radio astronomers in the late 1950s and confirmed by the Mariner and Venera probes. How can our neighbor planet be so hot? Although Venus is somewhat closer to the Sun than is Earth, its surface is hundreds of degrees hotter than you would expect from the extra sunlight it receives. Scientists wondered what could be heating the surface of Venus to a temperature above 700 K. The answer turned out to be the *greenhouse effect*.

The greenhouse effect works on Venus just as it does on Earth, but since Venus has so much more CO_2—almost a million times more—the effect is much stronger. The thick CO_2 acts as a blanket, making it very difficult for the infrared (heat) radiation from the ground to get back into space. As a result, the surface heats up. The energy balance is only restored when the planet is radiating as much energy as it receives from the Sun, but this can happen only when the temperature of the lower atmosphere is very high. One way of thinking of greenhouse heating is that it must raise the surface temperature of Venus until this energy balance is achieved.

Has Venus always had such a massive atmosphere and high surface temperature, or might it have evolved to such conditions from a climate that was once more nearly earthlike? The answer to this question is of particular interest to us as we look at the increasing levels of CO_2 in Earth's atmosphere. As the greenhouse effect becomes stronger on Earth, are we in any danger of transforming our own planet into a hellish place like Venus?

Let us try to reconstruct the possible evolution of Venus from an earthlike beginning to its present state. Venus may once have had a climate similar to that of Earth, with moderate temperatures, water oceans, and much of its CO_2 dissolved in the ocean or chemically combined with the surface rocks. Then we allow for modest additional heating—by gradual increase in the energy output of the Sun, for example. When we calculate how Venus' atmosphere would respond to such effects, it turns out that even a small amount of extra heat can lead to increased evaporation of water from the oceans and the release of gas from surface rocks.

This in turn means a further increase in the atmospheric CO_2 and H_2O, gases that would amplify the greenhouse effect in Venus' atmosphere. That would lead to still more heat near Venus' surface and the release of further CO_2 and H_2O. Unless some other processes intervene, the temperature thus continues to rise. Such a situation is called the runaway greenhouse effect.

We want to emphasize that the runaway greenhouse effect is not just a large greenhouse effect; it is an evolutionary *process*. The atmosphere evolves from having a small greenhouse effect, such as on Earth, to a situation where greenhouse warming is a major factor, as we see today on Venus. Once the large greenhouse conditions develop, the planet establishes a new, much hotter equilibrium near its surface.

Reversing the situation is difficult because of the role water plays. On Earth, most of the CO_2 is either chemically bound in the rocks of our crust or dissolved by the water in our oceans. As Venus got hotter and hotter, its oceans evaporated, eliminating that safety valve. But the water vapor in the planet's atmosphere will not last forever in the presence of ultraviolet light from the Sun. The light element hydrogen can escape from the atmosphere, leaving the oxygen behind to combine chemically with surface rock. The loss of water is therefore an irreversible process: once the water is gone, it cannot be restored. There is evidence that this is just what happened to the water once present on Venus.

We don't know if the same runaway greenhouse effect could one day happen on Earth. Although we are uncertain about the point at which a stable greenhouse effect breaks down and turns into a runaway greenhouse effect, Venus stands as clear testament to the fact that a planet cannot continue heating indefinitely without a major change in its oceans and atmosphere. It is a conclusion that we and our descendants will surely want to pay close attention to.

This page titled 14.3: The Massive Atmosphere of Venus is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.4: The Massive Atmosphere of Venus by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

15: MARS

15.1: The Geology of Mars
15.2: Water and Life on Mars
15.3: Divergent Planetary Evolution
15.4: Key Terms

This page titled 15: Mars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

15.1: The Geology of Mars

Learning Objectives

By the end of this section, you will be able to:

- Discuss the main missions that have explored Mars
- Explain what we have learned from examination of meteorites from Mars
- Describe the various features found on the surface of Mars
- Compare the volcanoes and canyons on Mars with those of Earth
- Describe the general conditions on the surface of Mars

Mars is more interesting to most people than Venus because it is more hospitable. Even from the distance of Earth, we can see surface features on Mars and follow the seasonal changes in its polar caps (Figure 10.13). Although the surface today is dry and cold, evidence collected by spacecraft suggests that Mars once had blue skies and lakes of liquid water. Even today, it is the sort of place we can imagine astronauts visiting and perhaps even setting up permanent bases.



Figure 10.13 : Mars Photographed by the Hubble Space Telescope. This is one of the best photos of Mars taken from our planet, obtained in June 2001 when Mars was only 68 million kilometers away. The resolution is about 20 kilometers—much better than can be obtained with ground-based telescopes but still insufficient to reveal the underlying geology of Mars. (credit: modification of work by NASA and the Hubble Heritage Team (STScI/AURA))

Spacecraft Exploration of Mars

Mars has been intensively investigated by spacecraft. More than 50 spacecraft have been launched toward Mars, but only about half were fully successful. The first visitor was the US Mariner 4, which flew past Mars in 1965 and transmitted 22 photos to Earth. These pictures showed an apparently bleak planet with abundant impact craters. In those days, craters were unexpected; some people who were romantically inclined still hoped to see canals or something like them. In any case, newspaper headlines sadly announced that Mars was a “dead planet.”

In 1971, NASA’s Mariner 9 became the first spacecraft to orbit another planet, mapping the entire surface of Mars at a resolution of about 1 kilometer and discovering a great variety of geological features, including volcanoes, huge canyons, intricate layers on the polar caps, and channels that appeared to have been cut by running water. Geologically, Mars didn’t look so dead after all.

The twin Viking spacecraft of the 1970s were among the most ambitious and successful of all planetary missions. Two *orbiters*

surveyed the planet and served to relay communications for two *landers* on the surface. After an exciting and sometimes frustrating search for a safe landing spot, the Viking 1 lander touched down on the surface of Chryse Planitia (the Plains of Gold) on July 20, 1976, exactly 7 years after Neil Armstrong's historic first step on the Moon. Two months later, Viking 2 landed with equal success in another plain farther north, called Utopia. The landers photographed the surface with high resolution and carried out complex experiments searching for evidence of life, while the orbiters provided a global perspective on Mars geology.

Mars languished unvisited for two decades after Viking. Two more spacecraft were launched toward Mars, by NASA and the Russian Space Agency, but both failed before reaching the planet.

The situation changed, beginning in the 1990s, as NASA began a new exploration program using spacecraft that were smaller and less expensive than Viking. The first of the new missions, appropriately called Pathfinder, landed the first wheeled, solar-powered rover on the martian surface on July 4, 1997 (Figure 10.14). An orbiter called *Mars Global Surveyor* (MGS) arrived a few months later and began high-resolution photography of the entire surface over more than one martian year. The most dramatic discovery by this spacecraft, which continued to operate until 2006, was evidence of gullies apparently cut by surface water, as we will discuss later. These missions were followed in 2003 by the NASA *Mars Odyssey* orbiter, and the ESA *Mars Express* orbiter, both carrying high-resolution cameras. A gamma-ray spectrometer on *Odyssey* discovered a large amount of subsurface hydrogen (probably in the form of frozen water). Subsequent orbiters included the NASA *Mars Reconnaissance Orbiter* to evaluate future landing sites, MAVEN to study the upper atmosphere, and India's *Mangalayaan*, also focused on study of Mars' thin layers of air. Several of these orbiters are also equipped to communicate with landers and rovers on the surface and serve as data relays to Earth.

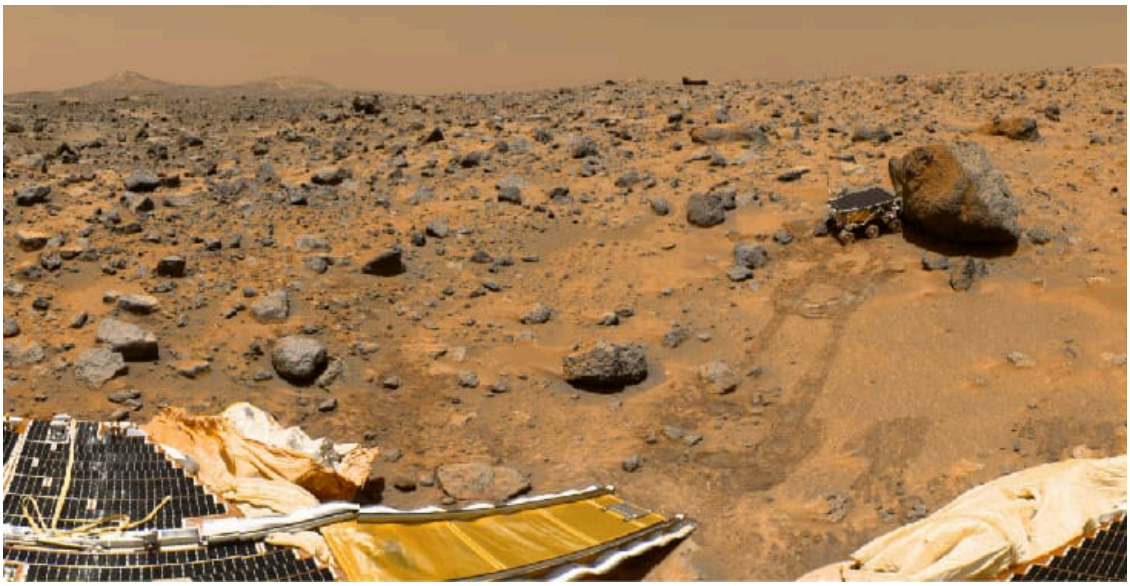
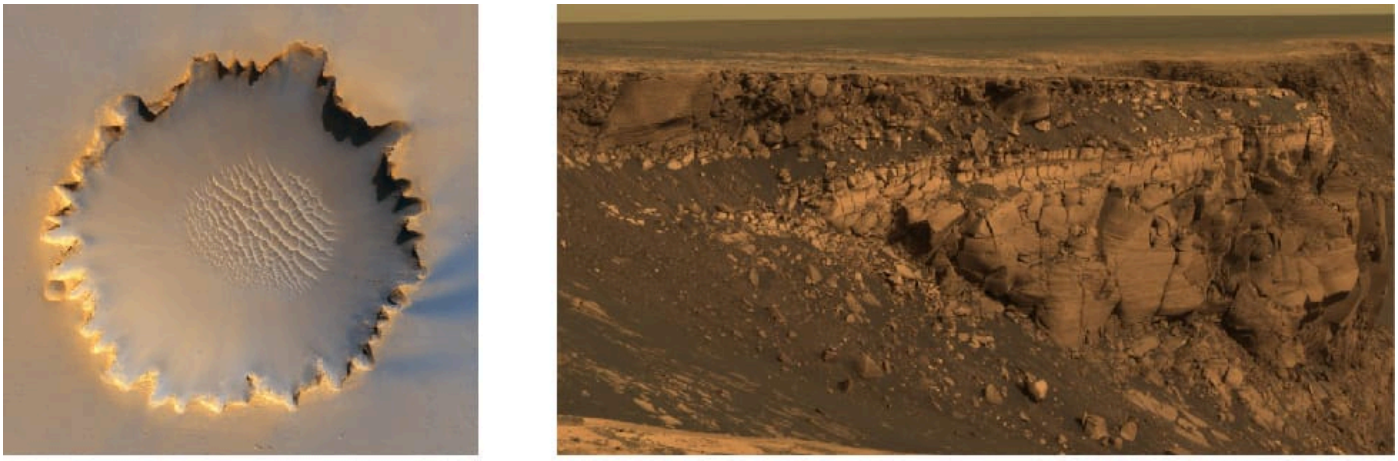


Figure 10.14 : Surface View from Mars Pathfinder. The scene from the Pathfinder lander shows a windswept plain, sculpted long ago when water flowed out of the martian highlands and into the depression where the spacecraft landed. The *Sojourner* rover, the first wheeled vehicle on Mars, is about the size of a microwave oven. Its flat top contains solar cells that provided electricity to run the vehicle. You can see the ramp from the lander and the path the rover took to the larger rock that the mission team nicknamed "Yogi." (credit: NASA/JPL)

In 2003, NASA began a series of highly successful Mars landers. Twin Mars Exploration Rovers (MER), named *Spirit* and *Opportunity*, have been successful far beyond their planned lifetimes. The design goal for the rovers was 600 meters of travel; in fact, they have traveled jointly more than 50 kilometers. After scouting around its rim, *Opportunity* drove down the steep walls into an impact crater called Victoria, then succeeded with some difficulty in climbing back out to resume its route (Figure 10.15). Dust covering the rovers' solar cells caused a drop in power, but when a seasonal dust storm blew away the dust, the rovers resumed full operation. In order to survive winter, the rovers were positioned on slopes to maximize solar heating and power generation. In 2006, *Spirit* lost power on one of its wheels, and subsequently became stuck in the sand, where it continued operation as a fixed ground station. Meanwhile, in 2008, *Phoenix* (a spacecraft "reborn" of spare parts from a previous Mars mission that had failed) landed near the edge of the north polar cap, at latitude 68°, and directly measured water ice in the soil.



(a) (b)

Figure 10.15 : Victoria Crater. (a) This crater in Meridiani Planum is 800 meters wide, making it slightly smaller than Meteor crater on Earth. Note the dune field in the interior. (b) This image shows the view from the *Opportunity* rover as it scouted the rim of Victoria crater looking for a safe route down into the interior. (credit a: modification of work by NASA/JPL-Caltech/University of Arizona/Cornell/Phio State University; credit b: modification of work by NASA/JPL/Cornell)

In 2011, NASA launched its largest (and most expensive) Mars mission since Viking (see Figure 10.1). The 1-ton rover *Curiosity*, the size of a subcompact car, has plutonium-powered electrical generators, so that it is not dependent on sunlight for power. *Curiosity* made a pinpoint landing on the floor of Gale crater, a site selected for its complex geology and evidence that it had been submerged by water in the past. Previously, Mars landers had been sent to flat terrains with few hazards, as required by their lower targeting accuracy. The scientific goals of *Curiosity* include investigations of climate and geology, and assessment of the habitability of past and present Mars environments. In 2018, NASA's InSight Lander touched down on Mars, carrying a suite of scientific instruments. These include a package (nicknamed "the mole") that will dig into the surface of Mars 1 mm at a time, hoping to reach a depth of 5 meters with heat sensors. Neither of these missions carries a specific life detection instrument, however. So far, scientists have not been able to devise a simple instrument that could distinguish living from nonliving materials on Mars.

The focus on the possibility of life returned with the 2020 launch of the large rover *Perseverance*, the near-twin of *Curiosity*. It landed in a former lakebed, and part of its mission is to drill into ancient sedimentary rocks and collect small samples of these rocks to be returned (by a later mission) to Earth for laboratory study. NASA's remarkable helicopter drone, named *Ingenuity*, has acted like an airborne scout for this mission.

Link to Learning

The *Curiosity* rover required a remarkably complex landing sequence and NASA made a video about it called "7 Minutes of Terror" that went viral on the Internet.

A dramatic video summary of the first two years of *Curiosity*'s exploration of the martian surface can be viewed as well.

Martian Samples

Much of what we know of the Moon, including the circumstances of its origin, comes from studies of lunar samples, but spacecraft have not yet returned martian samples to Earth for laboratory analysis. It is with great interest, therefore, that scientists have discovered that samples of martian material are nevertheless already here on Earth, available for study. These are all members of a rare class of *meteorites* (Figure 10.16)—rocks that have fallen from space.



Figure 10.16 : Martian Meteorite. This fragment of basalt, ejected from Mars in a crater-forming impact, eventually arrived on Earth's surface. (credit: NASA)

How would rocks have escaped from Mars? Many impacts have occurred on the red planet, as shown by its heavily cratered surface. Fragments blasted from large impacts can escape from Mars, whose surface gravity is only 38% of Earth's. A long time later (typically a few million years), a very small fraction of these fragments collide with Earth and survive their passage through our atmosphere, just like other meteorites. (We'll discuss meteorites in more detail in the chapter on Cosmic Samples and the Origin of the Solar System.) By the way, rocks from the Moon have also reached our planet as meteorites, although we were able to demonstrate their lunar origin only by comparison with samples returned by the Apollo missions

Most of the martian meteorites are volcanic basalts; most of them are also relatively young—about 1.3 billion years old. We know from details of their composition that they are not from Earth or the Moon. Besides, there was no volcanic activity on the Moon to form them as recently as 1.3 billion years ago. It would be very difficult for ejecta from impacts on Venus to escape through its thick atmosphere. By the process of elimination, the only reasonable origin seems to be Mars, where the Tharsis volcanoes were active at that time.

The martian origin of these meteorites was confirmed by the analysis of tiny gas bubbles trapped inside several of them. These bubbles match the atmospheric properties of Mars as first measured directly by Viking. It appears that some atmospheric gas was trapped in the rock by the shock of the impact that ejected it from Mars and started it on its way toward Earth.

One of the most exciting results from analysis of these martian samples has been the discovery of both water and organic (carbon-based) compounds in them, which suggests that Mars may once have had oceans and perhaps even life on its surface. As we have already hinted, there is other evidence for the presence of flowing water on Mars in the remote past, and even extending to the present.

In this and the following sections, we will summarize the picture of Mars as revealed by all these exploratory missions and by about 40 samples from Mars.

Global Properties of Mars

Mars has a diameter of 6790 kilometers, just over half the diameter of Earth, giving it a total surface area very nearly equal to the continental (land) area of our planet. Its overall density of 3.9 g/cm^3 suggests a composition consisting primarily of silicates but with a small metal core. The planet has no global magnetic field, although there are areas of strong surface magnetization that indicate that there was a global field billions of years ago. Apparently, the red planet has no liquid material in its core today that would conduct electricity.

Thanks to the *Mars Global Surveyor*, we have mapped the entire planet, as shown in Figure 10.17. A laser altimeter on board made millions of separate measurements of the surface topography to a precision of a few meters—good enough to show even the annual deposition and evaporation of the polar caps. Like Earth, the Moon, and Venus, the surface of Mars has continental or highland areas as well as widespread volcanic plains. The total range in elevation from the top of the highest mountain (Olympus Mons) to the bottom of the deepest basin (Hellas) is 31 kilometers.

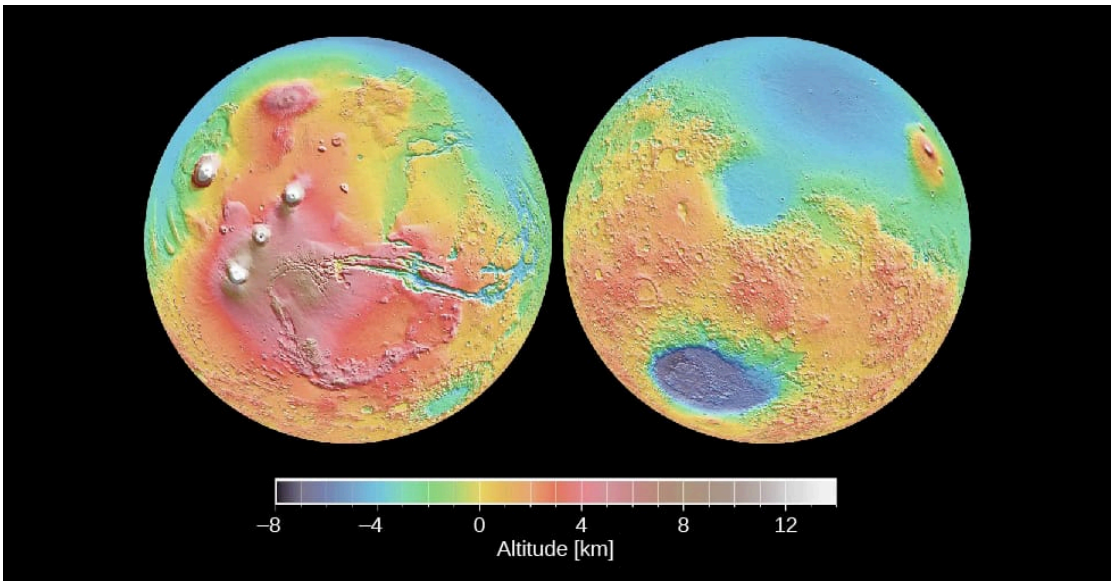


Figure 10.17 : Mars Map from Laser Ranging. These globes are highly precise topographic maps, reconstructed from millions of individual elevation measurements made with the *Mars Global Surveyor*. Color is used to indicate elevation. The hemisphere on the left includes the Tharsis bulge and Olympus Mons, the highest mountain on Mars; the hemisphere on the right includes the Hellas basin, which has the lowest elevation on Mars. (credit: modification of work by NASA/JPL)

Approximately half the planet consists of heavily cratered highland terrain, found primarily in the southern hemisphere. The other half, which is mostly in the north, contains younger, lightly cratered volcanic plains at an average elevation about 5 kilometers lower than the highlands. Remember that we saw a similar pattern on Earth, the Moon, and Venus. A geological division into older highlands and younger lowland plains seems to be characteristic of all the terrestrial planets except Mercury.

Lying across the north-south division of Mars is an uplifted continent the size of North America. This is the 10-kilometer-high Tharsis bulge, a volcanic region crowned by four great volcanoes that rise still higher into the martian sky.

Volcanoes on Mars

The lowland plains of Mars look very much like the lunar maria, and they have about the same density of impact craters. Like the lunar maria, they probably formed between 3 and 4 billion years ago. Apparently, Mars experienced extensive volcanic activity at about the same time the Moon did, producing similar basaltic lavas.

The largest volcanic mountains of Mars are found in the Tharsis area (you can see them in Figure 10.17), although smaller volcanoes dot much of the surface. The most dramatic volcano on Mars is Olympus Mons (Mount Olympus), with a diameter larger than 500 kilometers and a summit that towers more than 20 kilometers above the surrounding plains—three times higher than the tallest mountain on Earth (Figure 10.18). The volume of this immense volcano is nearly 100 times greater than that of Mauna Loa in Hawaii. Placed on Earth's surface, Olympus would more than cover the entire state of Missouri.



Figure 10.18 : Olympus Mons. The largest volcano on Mars, and probably the largest in the solar system, is Olympus Mons, illustrated in this computer-generated rendering based on data from the *Mars Global Surveyor's* laser altimeter. Placed on Earth, the base of Olympus Mons would completely cover the state of Missouri; the caldera, the circular opening at the top, is 65 kilometers across, about the size of Los Angeles. (credit: NASA/Corbis)

Images taken from orbit allow scientists to search for impact craters on the slopes of these volcanoes in order to estimate their age. Many of the volcanoes show a fair number of such craters, suggesting that they ceased activity a billion years or more ago. However, Olympus Mons has very, very few impact craters. Its present surface cannot be more than about 100 million years old; it may even be much younger. Some of the fresh-looking lava flows might have been formed a hundred years ago, or a thousand, or a million, but geologically speaking, they are quite young. This leads geologists to the conclusion that Olympus Mons possibly remains intermittently active today—something future Mars land developers may want to keep in mind.

Martian Cracks and Canyons

The Tharsis bulge has many interesting geological features in addition to its huge volcanoes. In this part of the planet, the surface itself has bulged upward, forced by great pressures from below, resulting in extensive tectonic cracking of the crust. Among the most spectacular tectonic features on Mars are the canyons called the Valles Marineris (or Mariner Valleys, named after Mariner 9, which first revealed them to us), which are shown in Figure 10.19. They extend for about 5000 kilometers (nearly a quarter of the way around Mars) along the slopes of the Tharsis bulge. If it were on Earth, this canyon system would stretch all the way from Los Angeles to Washington, DC. The main canyon is about 7 kilometers deep and up to 100 kilometers wide, large enough for the Grand Canyon of the Colorado River to fit comfortably into one of its side canyons. Viewers of the movie "The Martian" can see a recreation of these canyonlands, as the film's hero takes a long trip through a spectacular (and somewhat exaggerated) presentation of this part of Mars.

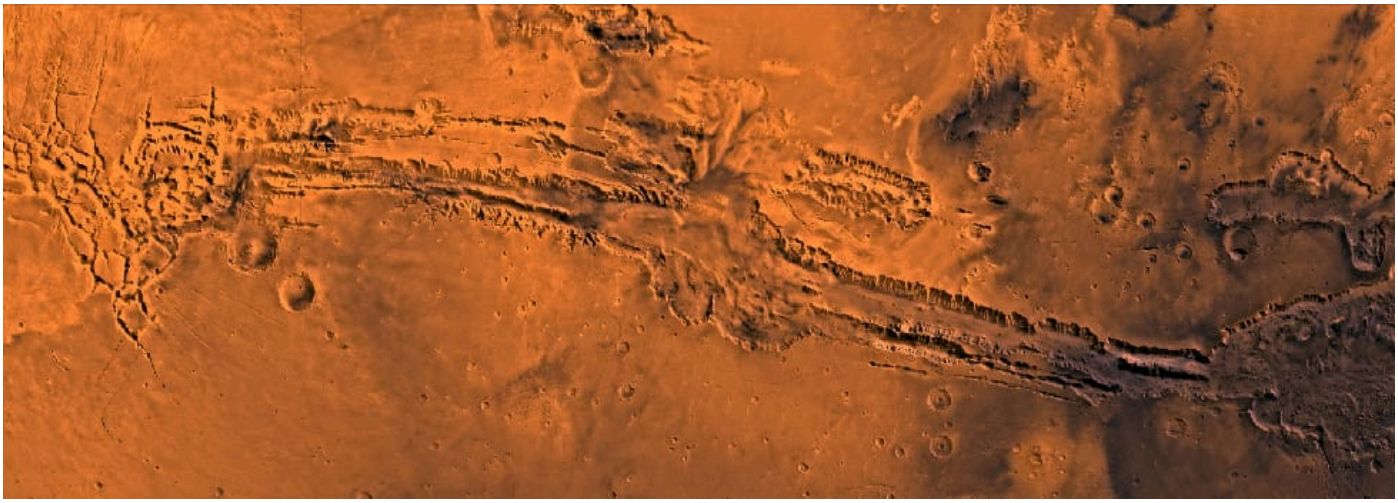


Figure 10.19 : Heavily Eroded Canyonlands on Mars. This image shows the Valles Marineris canyon complex, which is 3000 kilometers long and 8 kilometers deep. (credit: NASA/JPL/USGS)

Link to Learning

An excellent 4-minute video tour of Valles Marineris, narrated by planetary scientist Phil Christensen, is available for viewing.

The term “canyon” is somewhat misleading here because the Valles Marineris canyons have no outlets and were not cut by running water. They are basically tectonic cracks, produced by the same crustal tensions that caused the Tharsis uplift. However, water has played a later role in shaping the canyons, primarily by seeping from deep springs and undercutting the cliffs. This undercutting led to landslides that gradually widened the original cracks into the great valleys we see today (Figure 10.20). Today, the primary form of erosion in the canyons is probably wind.



Figure 10.20 : Martian Landslides. This Viking orbiter image shows Ophir Chasma, one of the connected valleys of the Valles Marineris canyon system. Look carefully and you can see enormous landslides whose debris is piled up underneath the cliff wall, which tower up to 10 kilometers above the canyon floor. (credit: modification of work by NASA/JPL/USGS)

While the Tharsis bulge and Valles Marineris are impressive, in general, we see fewer tectonic structures on Mars than on Venus. In part, this may reflect a lower general level of geological activity, as would be expected for a smaller planet. But it is also possible that evidence of widespread faulting has been buried by wind-deposited sediment over much of Mars. Like Earth, Mars may have hidden part of its geological history under a cloak of soil.

The View on the Martian Surface

The first spacecraft to land successfully on Mars were Vikings 1 and 2 and Mars Pathfinder. All sent back photos that showed a desolate but strangely beautiful landscape, including numerous angular rocks interspersed with dune like deposits of fine-grained, reddish soil (Figure 10.21).



Figure 10.21 : Three Martian Landing Sites. The Mars landers Viking 1 in Chryse, Pathfinder in Ares Valley, and Viking 2 in Utopia, all photographed their immediate surroundings. It is apparent from the similarity of these three photos that each spacecraft touched down on a flat, windswept plain littered with rocks ranging from tiny pebbles up to meter-size boulders. It is probable that most of Mars looks like this on the surface. (credit “Viking 1”: modification of work by Van der Hoorn/NASA; credit “Pathfinder”: modification of work by NASA; credit “Viking 2”: modification of work by NASA; credit Mars: modification of work by NASA/Goddard Space Flight Center)

All three of these landers were targeted to relatively flat, lowland terrain. Instruments on the landers found that the soil consisted of clays and iron oxides, as had long been expected from the red color of the planet. All the rocks measured appeared to be of volcanic origin and roughly the same composition. Later landers were targeted to touch down in areas that apparently were flooded sometime in the past, where sedimentary rock layers, formed in the presence of water, are common. (Although we should note that nearly all the planet is blanketed in at least a thin layer of wind-blown dust).

The Viking landers included weather stations that operated for several years, providing a perspective on martian weather. The temperatures they measured varied greatly with the seasons, due to the absence of moderating oceans and clouds. Typically, the summer maximum at Viking 1 was 240 K (–33 °C), dropping to 190 K (–83 °C) at the same location just before dawn. The lowest air temperatures, measured farther north by Viking 2, were about 173 K (–100 °C). During the winter, Viking 2 also photographed water frost deposits on the ground (Figure 10.22). We make a point of saying “water frost” here because at some locations on Mars, it gets cold enough for carbon dioxide (dry ice) to freeze out of the atmosphere as well.

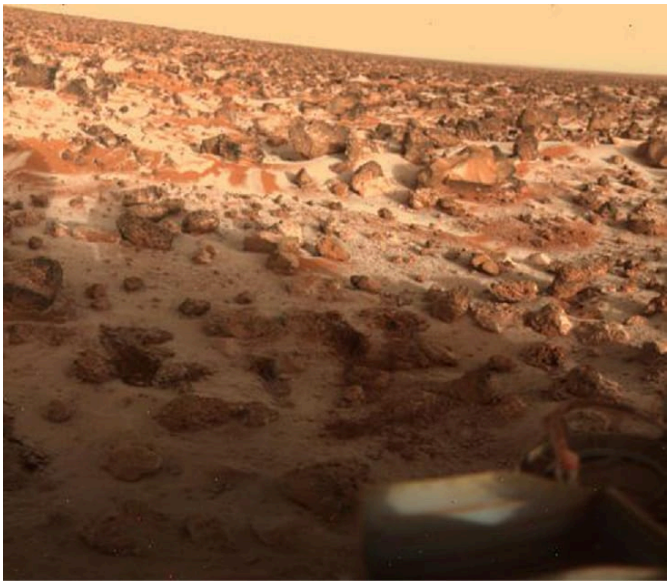


Figure 10.22 : Water Frost in Utopia. This image of surface frost was photographed at the Viking 2 landing site during late winter. (credit: NASA/JPL)

Most of the winds measured on Mars are only a few kilometers per hour. However, Mars is capable of great windstorms that can shroud the entire planet with windblown dust. Such high winds can strip the surface of some of its loose, fine dust, leaving the rock exposed. The later rovers found that each sunny afternoon the atmosphere became turbulent as heat rose off the surface. This turbulence generated dust devils, which play an important role in lifting the fine dust into the atmosphere. As the dust devils strip off the top layer of light dust and expose darker material underneath, they can produce fantastic patterns on the ground (Figure 10.23).

Wind on Mars plays an important role in redistributing surface material. Figure 10.23 shows a beautiful area of dark sand dunes on top of lighter material. Much of the material stripped out of the martian canyons has been dumped in extensive dune fields like this, mostly at high latitudes.



(a) (b)

Figure 10.23 : Dust Devil Tracks and Sand Dunes. (a) This high-resolution photo from the *Mars Global Surveyor* shows the dark tracks of several dust devils that have stripped away a thin coating of light-colored dust. This view is of an area about 3 kilometers across. Dust devils are one of the most important ways that dust gets redistributed by the martian winds. They may also help keep the solar panels of our rovers free of dust. (b) These windblown sand dunes on Mars overlay a lighter sandy surface. Each dune in this high-resolution view is about 1 kilometer across. (credit a: modification of work by NASA/JPL/University of Arizona; credit b: modification of work by NASA/JPL-Caltech/University of Arizona)

This page titled 15.1: The Geology of Mars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.5: The Geology of Mars by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

15.2: Water and Life on Mars

Learning Objectives

By the end of this section, you will be able to:

- Describe the general composition of the atmosphere on Mars
- Explain what we know about the polar ice caps on Mars and how we know it
- Describe the evidence for the presence of water in the past history of Mars
- Summarize the evidence for and against the possibility of life on Mars

Of all the planets and moons in the solar system, Mars seems to be the most promising place to look for life, both fossil microbes and (we hope) some forms of life deeper underground that still survive today. But where (and how) should we look for life? We know that the one requirement shared by all life on Earth is liquid water. Therefore, the guiding principle in assessing habitability on Mars and elsewhere has been to "follow the water." That is the perspective we take in this section, to follow the water on the red planet and hope it will lead us to life.

Atmosphere and Clouds on Mars

The atmosphere of Mars today has an average surface pressure of only 0.007 bar, less than 1% that of Earth. (This is how thin the air is about 30 kilometers above Earth's surface.) Martian air is composed primarily of carbon dioxide (95%), with about 3% nitrogen and 2% argon. The proportions of different gases are similar to those in the atmosphere of Venus (see Table 10.2), but a lot less of each gas is found in the thin air on Mars.

While winds on Mars can reach high speeds, they exert much less force than wind of the same velocity would on Earth because the atmosphere is so thin. The wind is able, however, to loft very fine dust particles, which can sometimes develop planet-wide dust storms. It is this fine dust that coats almost all the surface, giving Mars its distinctive red color. In the absence of surface water, wind erosion plays a major role in sculpting the martian surface (Figure 10.24).

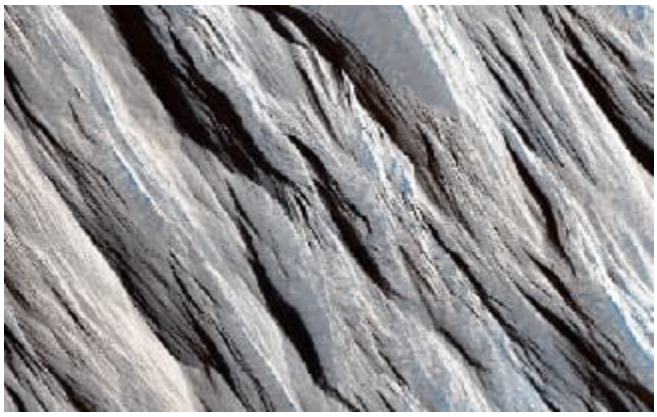


Figure 10.24 : Wind Erosion on Mars. These long straight ridges, called yardangs, are aligned with the dominant wind direction. This is a high-resolution image from the *Mars Reconnaissance Orbiter* and is about 1 kilometer wide. (credit: NASA/JPL-Caltech/University of Arizona)

Link to Learning

The issue of how strong the winds on Mars can be plays a big role in the 2015 hit movie *The Martian* in which the main character is stranded on Mars after being buried in the sand in a windstorm so great that his fellow astronauts have to leave the planet so their ship is not damaged. Astronomers have noted that the martian winds could not possibly be as forceful as depicted in the film because the air pressure is so low. In most ways, however, the depiction of Mars in this movie is remarkably accurate.

Although the atmosphere contains small amounts of water vapor and occasional clouds of water ice, liquid water is not stable under present conditions on Mars. Part of the problem is the low temperatures on the planet. But even if the temperature on a sunny summer day rises above the freezing point, the low pressure means that liquid water still cannot exist on the surface, except at the lowest elevations. At a pressure of less than 0.006 bar, the boiling point is as low or lower than the freezing point, and water changes directly from solid to vapor without an intermediate liquid state (as does "dry ice," carbon dioxide, on Earth). However, salts dissolved in water lower its freezing point, as we know from the way salt is used to thaw roads after snow and ice forms during winter on Earth. Salty water is therefore sometimes able to exist in liquid form on the martian surface, under the right conditions.

Several types of clouds can form in the martian atmosphere. First there are dust clouds, discussed above. Second are water-ice clouds similar to those on Earth. These often form around mountains, just as happens on our planet. Finally, the CO_2 of the atmosphere can itself condense at high altitudes to form hazes of dry ice crystals. The CO_2 clouds have no counterpart on Earth, since on our planet temperatures never drop low enough (down to about 150 K or about –125 °C) for this gas to condense.

The Polar Caps

Through a telescope, the most prominent surface features on Mars are the bright polar caps, which change with the seasons, similar to the seasonal snow cover on Earth. We do not usually think of the winter snow in northern latitudes as a part of our polar caps, but seen from space, the thin winter snow merges with Earth's thick, permanent ice caps to create an impression much like that seen on Mars (Figure 10.25).

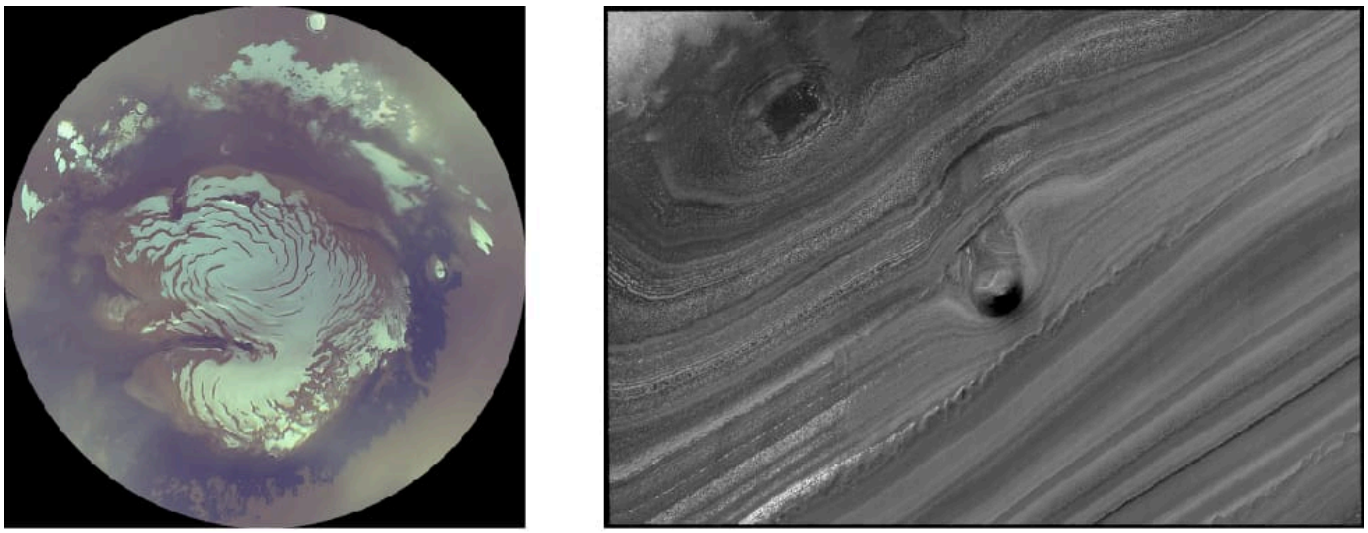


Figure 10.25 : Martian North Polar Cap. (a) This is a composite image of the north pole in summer, obtained in October 2006 by the *Mars Reconnaissance Orbiter*. It shows the mostly water-ice residual cap sitting atop light, tan-colored, layered sediments. Note that although the border of this photo is circular, it shows only a small part of the planet. (b) Here we see a small section of the layered terrain near the martian north pole. There is a mound about 40 meters high that is sticking out of a trough in the center of the picture. (credit a: modification of work by NASA/JPL/MSSS; credit b: modification of work by NASA/JPL-Caltech/University of Arizona)

The *seasonal caps* on Mars are composed not of ordinary snow but of frozen CO_2 (dry ice). These deposits condense directly from the atmosphere when the surface temperature drops below about 150 K. The caps develop during the cold martian winters and extend down to about 50° latitude by the start of spring.

Quite distinct from these thin seasonal caps of CO_2 are the *permanent* or *residual caps* that are always present near the poles. The southern permanent cap has a diameter of 350 kilometers and is composed of frozen CO_2 deposits together with a great deal of water ice. Throughout the southern summer, it remains at the freezing point of CO_2, 150 K, and this cold reservoir is thick enough to survive the summer heat intact.

The northern permanent cap is different. It is much larger, never shrinking to a diameter less than 1000 kilometers, and is composed of water ice. Summer temperatures in the north are too high for the frozen CO_2 to be retained. Measurements from the *Mars Global Surveyor* have established the exact elevations in the north polar region of Mars, showing that it is a large basin

about the size of our own Arctic Ocean basin. The ice cap itself is about 3 kilometers thick, with a total volume of about 10 million km^3 (similar to that of Earth's Mediterranean Sea). If Mars ever had extensive liquid water, this north polar basin would have contained a shallow sea. There is some indication of ancient shorelines visible, but better images will be required to verify this suggestion.

Images taken from orbit also show a distinctive type of terrain surrounding the permanent polar caps, as shown in Figure 10.25. At latitudes above 80° in both hemispheres, the surface consists of recent layered deposits that cover the older cratered ground below. Individual layers are typically ten to a few tens of meters thick, marked by alternating light and dark bands of sediment. Probably the material in the polar deposits includes dust carried by wind from the equatorial regions of Mars.

What do these terraced layers tell us about Mars? Some cyclic process is depositing dust and ice over periods of time. The time scales represented by the polar layers are tens of thousands of years. Apparently the martian climate experiences periodic changes at intervals similar to those between ice ages on Earth. Calculations indicate that the causes are probably also similar: the gravitational pull of the other planets produces variations in Mars' orbit and tilt as the great clockwork of the solar system goes through its paces.

The *Phoenix* spacecraft landed near the north polar cap in summer (Figure 10.26). Controllers knew that it would not be able to survive a polar winter, but directly measuring the characteristics of the polar region was deemed important enough to send a dedicated mission. The most exciting discovery came when the spacecraft tried to dig a shallow trench under the spacecraft. When the overlying dust was stripped off, they saw bright white material, apparently some kind of ice. From the way this ice sublimated over the next few days, it was clear that it was frozen water.

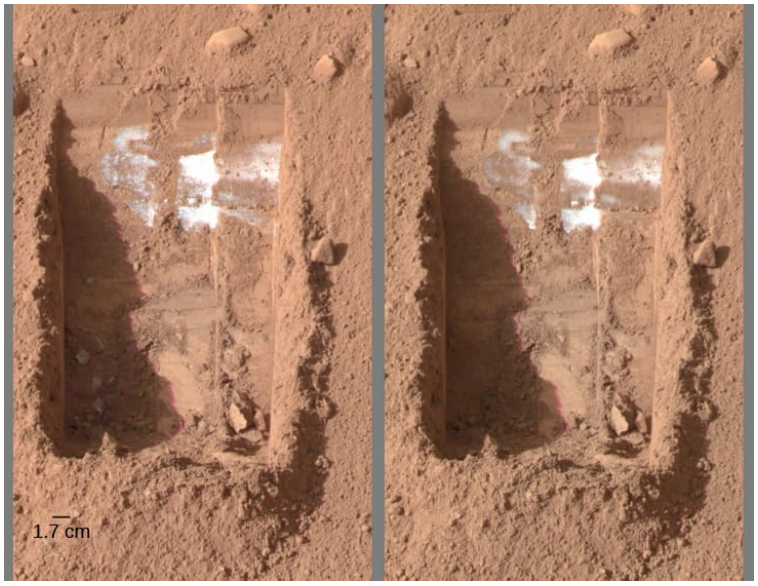


Figure 10.26 : Evaporating Ice on Mars. We see a trench dug by the *Phoenix* lander in the north polar region four martian days apart in June 2008. If you look at the shadowed region in the bottom left of the trench, you can see three spots of ice in the left image which have sublimated away in the right image. (credit: modification of work by NASA/JPL-Caltech/University of Arizona/Texas A&M University)

Example 10.1: Comparing the Amount of Water on Mars and Earth

It is interesting to estimate the amount of water (in the form of ice) on Mars and to compare this with the amount of water on Earth. In each case, we can find the total volume of a layer on a sphere by multiplying the area of the sphere Math input error

Solution

The volume of Earth's water is therefore the area $4\pi R^2$

$$4\pi\left(6.378 \times 10^6 \mathrm{m}\right)^2 = 5.1 \times 10^{14}\ \mathrm{m}^2$$

multiplied by the thickness of 3000 m:

$$5.1 \times 10^{14}\,\mathrm{m}^2 \times 3000\,\mathrm{m} = 1.5 \times 10^{18}\,\mathrm{m}^3$$

This gives $1.5 \times 10^{18}\mathrm{m}^3$ of water. Since water has a density of 1 ton per cubic meter (1000 kg/m^3), we can calculate the mass:

$$1.5 \times 10^{18}\,\mathrm{m}^3 \times 1\,\mathrm{ton/m}^3 = 1.5 \times 10^{18}\,\mathrm{tons}$$

For Mars, the ice doesn't cover the whole planet, only the caps; the polar cap area is

$$\pi R^2 = \pi(4 \times 10^5\,\mathrm{m})^2 = 5 \times 10^{11}\,\mathrm{m}^2$$

(Note that we converted kilometers to meters.)

The volume = area × height, so we have:

$$(2 \times 10^3\,\mathrm{m})(5 \times 10^{11}\,\mathrm{m}^2) = 1 \times 10^{15}\,\mathrm{m}^3 = 10^{15}\,\mathrm{m}^3$$

Therefore, the mass is:

$$10^{15}\,\mathrm{m}^3 \times 1\,\mathrm{ton/m}^3 = 10^{15}\,\mathrm{tons}$$

This is about 0.1% that of Earth's oceans.

Exercise 15.2.1

A better comparison might be to compare the amount of ice in the Mars polar ice caps to the amount of ice in the Greenland ice sheet on Earth, which has been estimated as 2.85 × 10^{15} m^3. How does this compare with the ice on Mars?

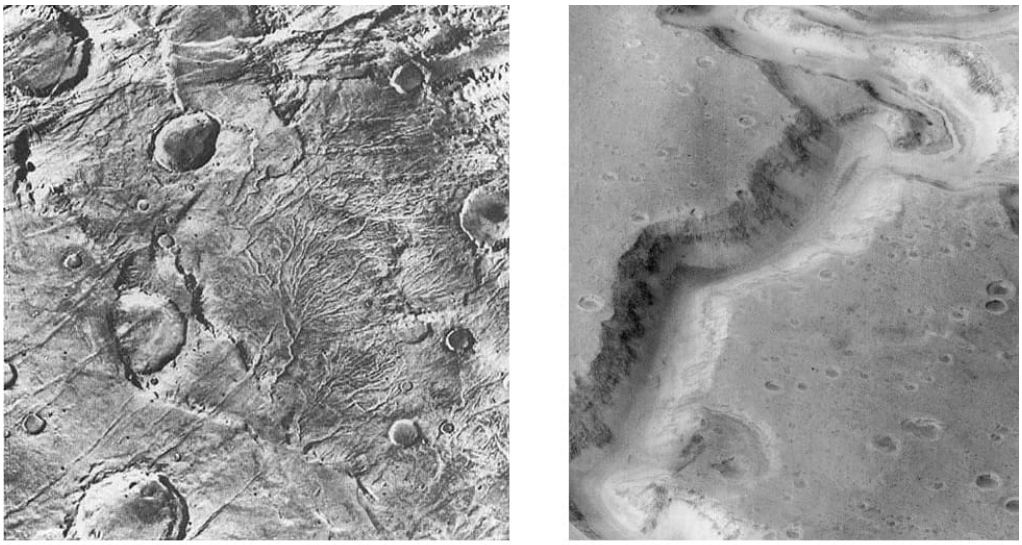
Answer available at the end of this section.

Channels and Gullies on Mars

Although no bodies of liquid water exist on Mars today, evidence has accumulated that rivers flowed on the red planet long ago. Two kinds of geological features appear to be remnants of ancient watercourses, while a third class—smaller gullies—suggests intermittent outbreaks of liquid water even today. We will examine each of these features in turn.

In the highland equatorial plains, there are multitudes of small, sinuous (twisting) channels—typically a few meters deep, some tens of meters wide, and perhaps 10 or 20 kilometers long (Figure 10.27). They are called runoff channels because they look like what geologists would expect from the surface runoff of ancient rain storms. These runoff channels seem to be telling us that the planet had a very different climate long ago. To estimate the age of these channels, we look at the cratering record. Crater counts show that this part of the planet is more cratered than the lunar maria but less cratered than the lunar highlands. Thus, the runoff channels are probably older than the lunar maria, presumably about 4 billion years old.

The second set of water-related features we see are *outflow channels* (Figure 10.27) are much larger than the runoff channels. The largest of these, which drain into the Chryse basin where Pathfinder landed, are 10 kilometers or more wide and hundreds of kilometers long. Many features of these outflow channels have convinced geologists that they were carved by huge volumes of running water, far too great to be produced by ordinary rainfall. Where could such floodwater have come from on Mars?



(a) (b)

Figure 10.27 : Runoff and Outflow Channels. (a) These runoff channels in the old martian highlands are interpreted as the valleys of ancient rivers fed by either rain or underground springs. The width of this image is about 200 kilometers. (b) This intriguing channel, called Nanedi Valles, resembles Earth riverbeds in some (but not all) ways. The tight curves and terraces seen in the channel certainly suggest the sustained flow of a fluid like water. The channel is about 2.5 kilometers across. (credit a: modification of work by Jim Secosky/NASA; credit b: modification of work by Jim Secosky/NASA)

As far we can tell, the regions where the outflow channels originate contained abundant water frozen in the soil as permafrost. Some local source of heating must have released this water, leading to a period of rapid and catastrophic flooding. Perhaps this heating was associated with the formation of the volcanic plains on Mars, which date back to roughly the same time as the outflow channels.

Note that neither the runoff channels nor the outflow channels are wide enough to be visible from Earth, nor do they follow straight lines. They could not have been the “canals” Percival Lowell imagined seeing on the red planet.

The third type of water feature, the smaller *gullies*, was discovered by the *Mars Global Surveyor* (Figure 10.28). The *Mars Global Surveyor*’s camera images achieved a resolution of a few meters, good enough to see something as small as a truck or bus on the surface. On the steep walls of valleys and craters at high latitudes, there are many erosional features that look like gullies carved by flowing water. These gullies are very young: not only are there no superimposed impact craters, but in some instances, the gullies seem to cut across recent wind-deposited dunes. Perhaps there is liquid water underground that can occasionally break out to produce short-lived surface flows before the water can freeze or evaporate.

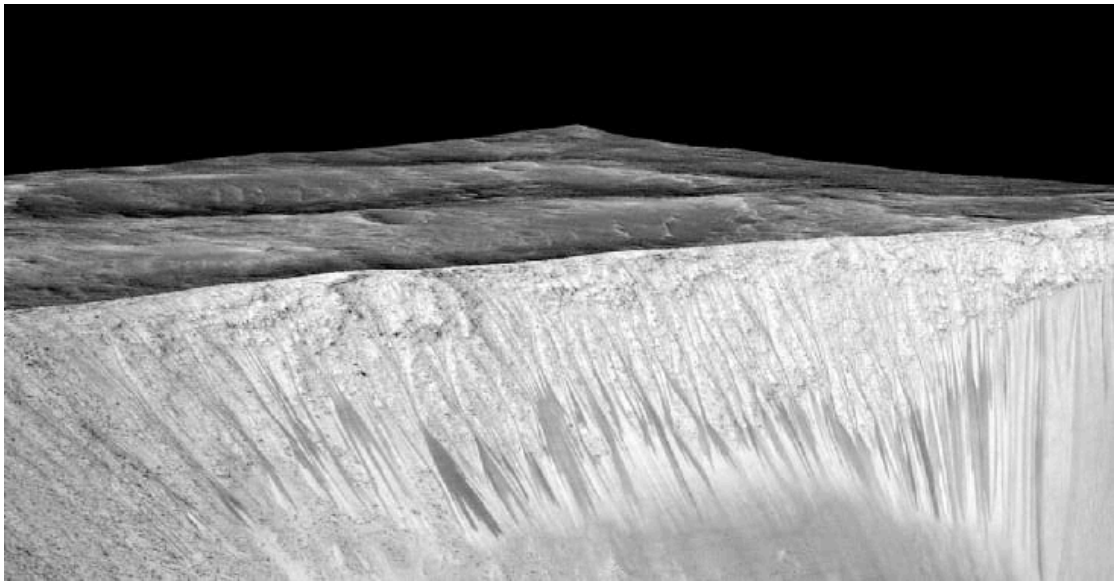


Figure 10.28 : Gullies on the Wall of Garni Crater. This high-resolution image is from the *Mars Reconnaissance Orbiter*. The dark streaks, which are each several hundred meters long, change in a seasonal pattern that suggests they are caused by the temporary flow of surface water. (credit: NASA/JPL-Caltech/University of Arizona)

The gullies also have the remarkable property of changing regularly with the martian seasons. Many of the dark streaks (visible in Figure 10.28) elongate within a period of a few days, indicating that something is flowing downhill—either water or dark sediment. If it is water, it requires a continuing source, either from the atmosphere or from springs that tap underground water layers (aquifers.) Underground water would be the most exciting possibility, but this explanation seems inconsistent with the fact that many of the dark streaks start at high elevations on the walls of craters.

Additional evidence that the dark streaks (called by the scientists *recurring slope lineae*) are caused by water was found in 2015 when spectra were obtained of the dark streaks (Figure 10.29). These showed the presence of hydrated salts produced by the evaporation of salty water. If the water is salty, it could remain liquid long enough to flow downstream for distances of a hundred meters or more, before it either evaporates or soaks into the ground. However, this discovery still does not identify the ultimate source of the water.

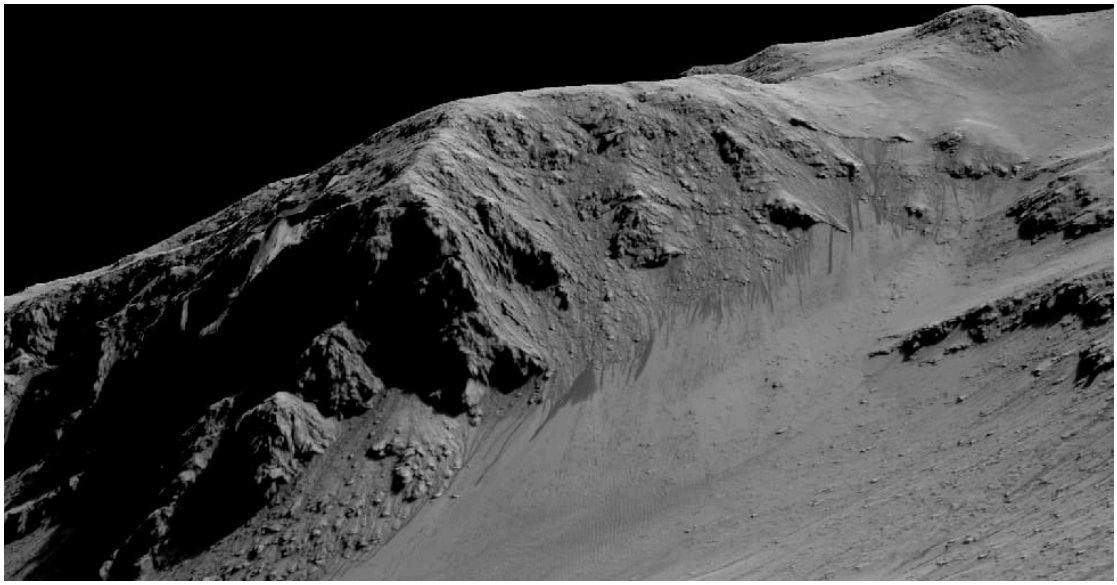


Figure 10.29 : Evidence for Liquid Water on Mars. The dark streaks in Horowitz crater, which move downslope, have been called recurring slope lineae. The streaks in the center of the image go down the wall of the crater for about a distance of 100 meters. Spectra taken of this region indicate that these are locations where salty liquid water flows on or just below the surface of Mars. (The vertical dimension is exaggerated by a factor of 1.5 compared to horizontal dimensions.) (credit: NASA/JPL-Caltech/University of Arizona)

Ancient Lakes and Glaciers

The rovers (*Spirit*, *Opportunity*, *Curiosity*, and *Perseverance,*, that have operated on the surface of Mars have been used to hunt for additional evidence of water. They could not reach the most interesting sites, such as the gullies, which are located on steep slopes. Instead, they explored sites that might be dried-out lake beds, dating back to a time when the climate on Mars was warmer and the atmosphere thicker—allowing water to be liquid on the surface.

Spirit was specifically targeted to explore what looked like an ancient lake-bed in Gusev crater, with an outflow channel emptying into it. However, when the spacecraft landed, it found that the former lakebed had been covered by thin lava flows, blocking the rover from access to the sedimentary rocks it had hoped to find. However, *Opportunity* had better luck. Peering at the walls of a small crater, it detected layered sedimentary rock. These rocks contained chemical evidence of evaporation, suggesting there had been a shallow salty lake in that location. In these sedimentary rocks were also small spheres that were rich in the mineral hematite, which forms only in watery environments. Apparently this very large basin had once been underwater.

Link to Learning

The small spherical rocks were nicknamed “blueberries” by the science team and the discovery of a whole “berry-bowl” of them was announced in this interesting news release from NASA.

The *Curiosity* rover landed inside Gale crater, where photos taken from orbit also suggested past water erosion. It discovered numerous sedimentary rocks, some in the form of mudstones from an ancient lakebed; it also found indications of rocks formed by the action of shallow water at the time the sediment formed (Figure 10.30).

Even today there is evidence of large quantities of ice just below the surface of Mars. In the mid-latitudes, high-resolution photos from orbit have revealed glaciers covered with dirt and dust. In some cliffs, the ice is observed directly (see Figure 10.30). *Perseverance* is exploring the 45-km-wide crater Jezero, which is a dusty former lakebed and river delta with evidence of large rocks deposited by ancient floods. These glaciers are thought to have formed during warm periods, when the atmospheric pressure was greater and snow and ice could precipitate. They also suggest readily available frozen water that could support future human exploration of the planet.

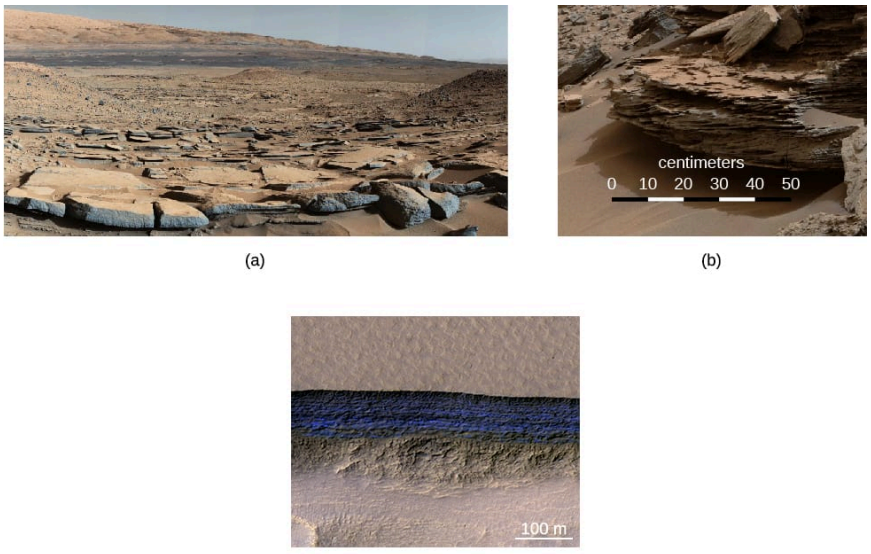


Figure 10.30 : Gale Crater and Underground Ice Deposits. (a) This scene, photographed by the *Curiosity* rover, shows an ancient lakebed of cracked mudstones. (b) Geologists working with the *Curiosity* rover interpret this image of cross-bedded sandstone in Gale crater as evidence of liquid water passing over a loose bed of sediment at the time this rock formed. (c) Ice bands a hundred meters tall are visible in blue in a cliff-face on Mars, suggesting large deposits of frozen water buried just a few meters below the surface. Note that the blue color has been exaggerated in this photo, taken by the Mars Reconnaissance Orbiter spacecraft. (credit a: modification of work by NASA/JPL-Caltech/MSSS; credit b: modification of work by NASA/JPL-Caltech/MSSS; credit c: modification of work by NASA/JPL-Caltech/UA/USGS)

Making Connections: Astronomy and Pseudoscience: The “Face on Mars”

People like human faces. We humans have developed great skill in recognizing people and interpreting facial expressions. We also have a tendency to see faces in many natural formations, from clouds to the man in the Moon. One of the curiosities that emerged from the Viking orbiters’ global mapping of Mars was the discovery of a strangely shaped mesa in the Cydonia region that resembled a human face. Despite later rumors of a cover-up, the “Face on Mars” was, in fact, recognized by Viking scientists and included in one of the early mission press releases. At the low resolution and oblique lighting under which the Viking image was obtained, the mile-wide mesa had something of a Sphinx-like appearance.

Unfortunately, a small band of individuals decided that this formation was an artificial, carved sculpture of a human face placed on Mars by an ancient civilization that thrived there hundreds of thousands of years ago. A band of “true believers” grew around the face and tried to deduce the nature of the “sculptors” who made it. This group also linked the face to a variety of other pseudoscientific phenomena such as crop circles (patterns in fields of grain, mostly in Britain, now known to be the work of pranksters).

Members of this group accused NASA of covering up evidence of intelligent life on Mars, and they received a great deal of help in publicizing their perspective from tabloid media. Some of the believers picketed the Jet Propulsion Laboratory at the time of the failure of the *Mars Observer* spacecraft, circulating stories that the “failure” of the *Mars Observer* was itself a fake, and that its true (secret) mission was to photograph the face.

The high-resolution *Mars Observer* camera (MOC) was reflown on the *Mars Global Surveyor* mission, which arrived at Mars in 1997. On April 5, 1998, in Orbit 220, the MOC obtained an oblique image of the face at a resolution of 4 meters per pixel, a factor-of-10 improvement in resolution over the Viking image. Another image in 2001 had even higher resolution. Immediately released by NASA, the new images showed a low mesa-like hill cut crossways by several roughly linear ridges and depressions, which were misidentified in the 1976 photo as the eyes and mouth of a face. Only with an enormous dose of imagination can any resemblance to a face be seen in the new images, demonstrating how dramatically our interpretation of geology can change with large improvements in resolution. The original and the higher resolution images can be seen in Figure 10.31.

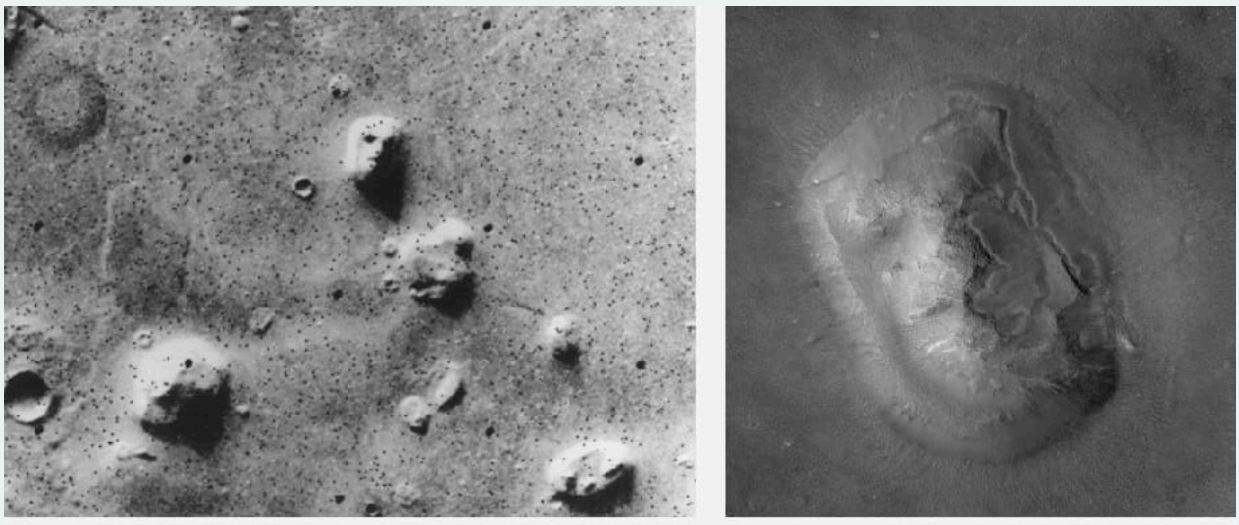


Figure 10.31 : Face on Mars. The so-called “Face on Mars” is seen (a) in low resolution from Viking (the “face” is in the upper part of the picture) and (b) with 20 times better resolution from the *Mars Global Surveyor*. (credit a: modification of work NASA/JPL; credit b: modification of work by NASA/JPL/MSSS)

After 20 years of promoting pseudoscientific interpretations and various conspiracy theories, can the “Face on Mars” believers now accept reality? Unfortunately, it does not seem so. They have accused NASA of faking the new picture. They also suggest that the secret mission of the *Mars Observer* included a nuclear bomb used to destroy the face before it could be photographed in greater detail by the *Mars Global Surveyor*.

Space scientists find these suggestions incredible. NASA is spending increasing sums for research on life in the universe, and a major objective of current and upcoming Mars missions is to search for evidence of past microbial life on Mars. Conclusive evidence of extraterrestrial life would be one of the great discoveries of science and incidentally might well lead to increased funding for NASA. The idea that NASA or other government agencies would (or could) mount a conspiracy to suppress such welcome evidence is truly bizarre.

Alas, the "Face on Mars" story is only one example of a whole series of conspiracy theories that are kept before the public by dedicated believers, by people out to make a fast buck, and by irresponsible media attention. Others include the "urban legend" that the Air Force has the bodies of extraterrestrials at a secret base, the widely circulated report that UFOs crashed near Roswell, New Mexico (actually it was a balloon carrying scientific instruments to find evidence of Soviet nuclear tests), or the notion that alien astronauts helped build the Egyptian pyramids and many other ancient monuments because our ancestors were too stupid to do it alone.

In response to the increase in publicity given to these "fiction science" ideas, a group of scientists, educators, scholars, and magicians (who know a good hoax when they see one) have formed the Committee for Skeptical Inquiry. Two of the original authors of your book are active on the committee. For more information about its work delving into the rational explanations for paranormal claims, see their excellent magazine, *The Skeptical Inquirer*, or check out their website at https://skepticalinquirer.org/.

Climate Change on Mars

The evidence about ancient rivers and lakes of water on Mars discussed so far suggests that, billions of years ago, martian temperatures must have been warmer and the atmosphere must have been more substantial than it is today. But what could have changed the climate on Mars so dramatically?

We presume that, like Earth and Venus, Mars probably formed with a higher surface temperature thanks to the greenhouse effect. But Mars is a smaller planet, and its lower gravity means that atmospheric gases could escape more easily than from Earth and Venus. As more and more of the atmosphere escaped into space, the temperature on the surface gradually fell.

Eventually Mars became so cold that most of the water froze out of the atmosphere, further reducing its ability to retain heat. The planet experienced a sort of *runaway refrigerator effect*, just the opposite of the runaway greenhouse effect that occurred on Venus. Probably, this loss of atmosphere took place within less than a billion years after Mars formed. The result is the cold, dry Mars we see today.

Conditions a few meters below the martian surface, however, may be much different. There, liquid water (especially salty water) might persist, kept warm by the internal heat of Mars or the insulating layers solid and rock. Even on the surface, there may be ways to change the martian atmosphere temporarily.

Mars is likely to experience long-term climate cycles, which may be caused by the changing orbit and tilt of the planet. At times, one or both of the polar caps might melt, releasing a great deal of water vapor into the atmosphere. Perhaps an occasional impact by a comet might produce a temporary atmosphere that is thick enough to permit liquid water on the surface for a few weeks or months. Some have even suggested that future technology might allow us to *terraform* Mars—that is, to engineer its atmosphere and climate in ways that might make the planet more hospitable for long-term human habitation.

The Search for Life on Mars

If there was running water on Mars in the past, perhaps there was life as well. Could life, in some form, remain in the martian soil today? Testing this possibility, however unlikely, was one of the primary objectives of the Viking landers in 1976. These landers carried miniature biological laboratories to test for microorganisms in the martian soil. Martian soil was scooped up by the spacecraft's long arm and placed into the experimental chambers, where it was isolated and incubated in contact with a variety of gases, radioactive isotopes, and nutrients to see what would happen. The experiments looked for evidence of *respiration* by living animals, *absorption* of *nutrients* offered to organisms that might be present, and an *exchange of gases* between the soil and its surroundings for any reason whatsoever. A fourth instrument pulverized the soil and analyzed it carefully to determine what organic (carbon-bearing) material it contained.

The Viking experiments were so sensitive that, had one of the spacecraft landed anywhere on Earth (with the possible exception

of Antarctica), it would easily have detected life. But, to the disappointment of many scientists and members of the public, no life was detected on Mars. The soil tests for absorption of nutrients and gas exchange did show some activity, but this was most likely caused by chemical reactions that began as water was added to the soil and had nothing to do with life. In fact, these experiments showed that martian soil seems much more chemically active than terrestrial soils because of its exposure to solar ultraviolet radiation (since Mars has no ozone layer).

The organic chemistry experiment showed no trace of organic material, which is apparently destroyed on the martian surface by the sterilizing effect of this ultraviolet light. While the possibility of life on the surface has not been eliminated, most experts consider it negligible. Although Mars has the most earthlike environment of any planet in the solar system, the sad fact is that nobody seems to be home today, at least on the surface.

However, there is no reason to think that life could not have begun on Mars about 4 billion years ago, at the same time it started on Earth. The two planets had very similar surface conditions then. Thus, the attention of scientists has shifted to the search for *fossil* life on Mars. One of the primary questions to be addressed by future spacecraft is whether Mars once supported its own life forms and, if so, how this martian life compared with that on our own planet. Future missions will include the return of martian samples selected from sedimentary rocks at sites that once held water and thus perhaps ancient life. The most powerful searches for martian life (past or present) will thus be carried out in our laboratories here on Earth.

Making Connections: Planetary Protection

When scientists begin to search for life on another planet, they must make sure that we do not contaminate the other world with life carried from Earth. At the very beginning of spacecraft exploration on Mars, an international agreement specified that all landers were to be carefully sterilized to avoid accidentally transplanting terrestrial microbes to Mars. In the case of Viking, we know the sterilization was successful. Viking's failure to detect martian organisms also implies that these experiments did not detect hitchhiking terrestrial microbes.

As we have learned more about the harsh conditions on the martian surface, the sterilization requirements have been somewhat relaxed. It is evident that no terrestrial microbes could grow on the martian surface, with its low temperature, absence of water, and intense ultraviolet radiation. Microbes from Earth might survive in a dormant, dried state, but they cannot grow and proliferate on Mars.

The problem of contaminating Mars will become more serious, however, as we begin to search for life below the surface, where temperatures are higher and no ultraviolet light penetrates. The situation will be even more daunting if we consider human flights to Mars. Any humans will carry with them a multitude of terrestrial microbes of all kinds, and it is hard to imagine how we can effectively keep the two biospheres isolated from each other if Mars has indigenous life. Perhaps the best situation could be one in which the two life-forms are so different that each is effectively invisible to the other—not recognized on a chemical level as living or as potential food.

The most immediate issue of public concern is not with the contamination of Mars but with any dangers associated with returning Mars samples to Earth. NASA is committed to the complete biological isolation of returned samples until they are demonstrated to be safe. Even though the chances of contamination are extremely low, it is better to be safe than sorry.

Most likely there is no danger, even if there is life on Mars and alien microbes hitch a ride to Earth inside some of the returned samples. In fact, Mars is sending samples to Earth all the time in the form of the Mars meteorites. Since some of these microbes (if they exist) could probably survive the trip to Earth inside their rocky home, we may have been exposed many times over to martian microbes. Either they do not interact with our terrestrial life, or in effect our planet has already been inoculated against such alien bugs.

Link to Learning

More than any other planet, Mars has inspired science fiction writers over the years. You can find scientifically reasonable stories about Mars in a subject index of such stories online. If you click on Mars as a topic, you will find stories by a number of space scientists, including William Hartmann, Geoffrey Landis, and Ludek Pesek.

Answers

Exercise 15.2.1

The Greenland ice sheet has about 2.85 times as much ice as in the polar ice caps on Mars. They are about the same to the nearest power of 10.

This page titled 15.2: Water and Life on Mars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.6: Water and Life on Mars by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

15.3: Divergent Planetary Evolution

Learning Objectives

By the end of this section, you will be able to:

- Compare the planetary evolution of Venus, Earth, and Mars

Venus, Mars, and our own planet Earth form a remarkably diverse triad of worlds. Although all three orbit in roughly the same inner zone around the Sun and all apparently started with about the same chemical mix of silicates and metals, their evolutionary paths have diverged. As a result, Venus became hot and dry, Mars became cold and dry, and only Earth ended up with what we consider a hospitable climate.

We have discussed the runaway greenhouse effect on Venus and the runaway refrigerator effect on Mars, but we do not understand exactly what started these two planets down these separate evolutionary paths. Was Earth ever in danger of a similar fate? Or might it still be diverted onto one of these paths, perhaps due to stress on the atmosphere generated by human pollutants? One of the reasons for studying Venus and Mars is to seek insight into these questions.

Some people have even suggested that if we understood the evolution of Mars and Venus better, we could possibly reverse their evolution and restore more earthlike environments. While it seems unlikely that humans could ever make either Mars or Venus into a replica of Earth, considering such possibilities is a useful part of our more general quest to understand the delicate environmental balance that distinguishes our planet from its two neighbors. In Cosmic Samples and the Origin of the Solar System, we return to the comparative study of the terrestrial planets and their divergent evolutionary histories.

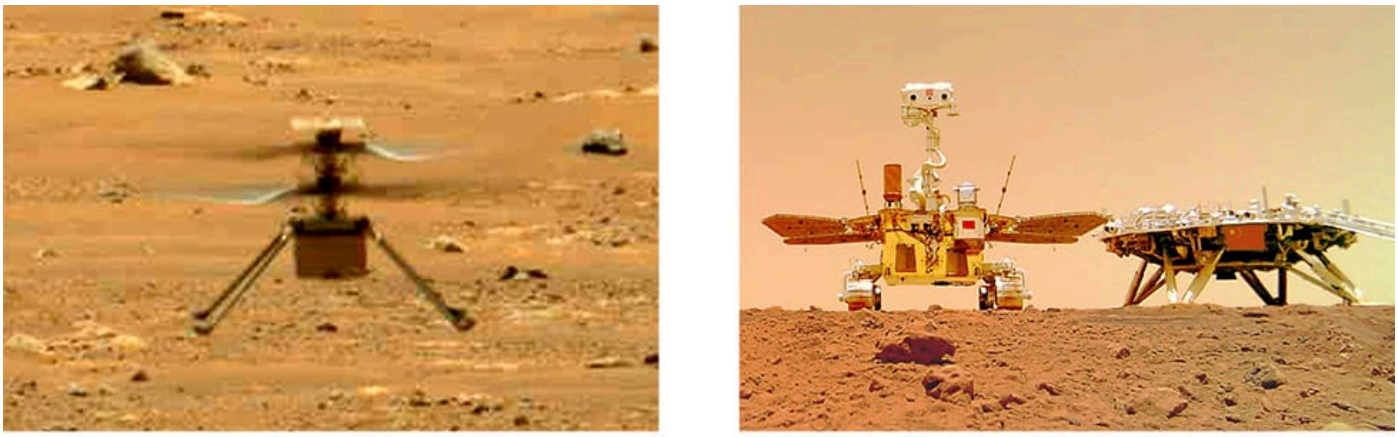


Figure 10.32 : New Vehicles on Mars. (a) The *Ingenuity* Helicopter flies above the martian surface. (b) The Chinese *Zhurong* rover, deployed in June 2021, is next to the *Tianwen-1* lander; the "group portrait" was taken by a wireless camera placed by the rover. (credit a: modification of work by NASA/JPL; credit b: modification of work by the China National Space Administration)

This page titled 15.3: Divergent Planetary Evolution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 10.7: Divergent Planetary Evolution by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

15.4: Key Terms

runaway greenhouse effect
: the process by which the greenhouse effect, rather than remaining stable or being lessened through intervention, continues to grow at an increasing rate

tectonic
: geological features that result from stresses and pressures in the crust of a planet; tectonic forces can lead to earthquakes and motion of the crust

This page titled 15.4: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

16: GIANT PLANETS

16.1: Exploring the Outer Planets
16.2: The Giant Planets
16.3: Atmosphere of the Giant Planets
16.4: Key Terms

This page titled 16: Giant Planets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

16.1: Exploring the Outer Planets

Learning Objectives

By the end of this section, you will be able to:

- Provide an overview of the composition of the giant planets
- Chronicle the robotic exploration of the outer solar system
- Summarize the missions sent to orbit the gas giants

The giant planets hold most of the mass in our planetary system. Jupiter alone exceeds the mass of all the other planets combined (Figure 11.2). The material available to build these planets can be divided into three classes by what they are made of: “gases,” “ices,” and “rocks” (see Table 11.1). The “gases” are primarily hydrogen and helium, the most abundant elements in the universe. The way it is used here, the term “ices” refers to composition only and not whether a substance is actually in a solid state. “Ices” means compounds that form from the next most abundant elements: oxygen, carbon, and nitrogen. Common ices are water, methane, and ammonia, but ices may also include carbon monoxide, carbon dioxide, and others. “Rocks” are even less abundant than ices, and include everything else: magnesium, silicon, iron, and so on.

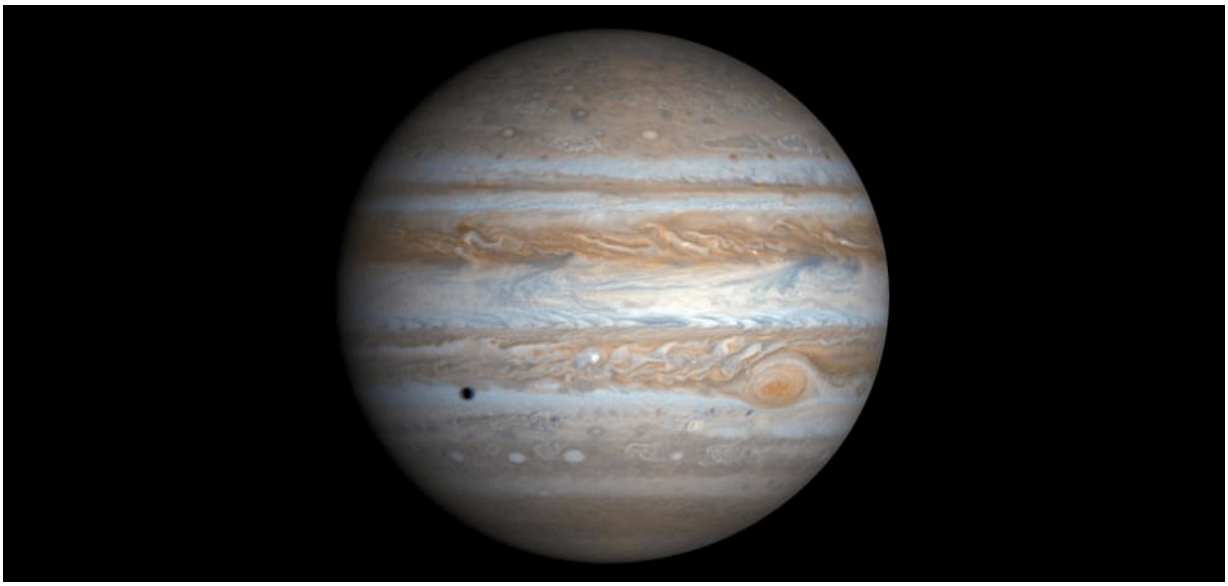


Figure 11.2 : Jupiter. The Cassini spacecraft imaged Jupiter on its way to Saturn in 2012. The giant storm system called the Great Red Spot is visible to the lower right. The dark spot to the lower left is the shadow of Jupiter’s moon Europa. (credit: modification of work by NASA/JPL)

Table 11.1: Abundances in the Outer Solar System

Type of Material	Name	Approximate % (by Mass)
Gas	Hydrogen (H_2)	75
Gas	Helium (He)	24
Ice	Water (H_2O)	0.6
Ice	Methane (CH_4)	0.4
Ice	Ammonia (NH_3)	0.1
Rock	Magnesium (Mg), iron (Fe), silicon (Si)	0.3

In the outer solar system, gases dominate the two largest planets, Jupiter and Saturn, hence their nickname “gas giants.” Uranus

and Neptune are sometimes called “ice giants” because their interiors contain far more of the “ice” component than their larger cousins. The chemistry for all four giant planet atmospheres is dominated by hydrogen. This hydrogen caused the chemistry of the outer solar system to become *reducing,* meaning that other elements tend to combine with hydrogen first. In the early solar system, most of the oxygen combined with hydrogen to make H_2O and was thus unavailable to form the kinds of oxidized compounds with other elements that are more familiar to us in the inner solar system (such as CO_2). As a result, the compounds detected in the atmosphere of the giant planets are mostly hydrogen-based gases such as methane (CH_4) and ammonia (NH_3), or more complex hydrocarbons (combinations of hydrogen and carbon) such as ethane (C_2H_6) and acetylene (C_2H_2).

Exploration of the Outer Solar System So Far

Eight spacecraft, seven from the United States and one from Europe, have penetrated beyond the asteroid belt into the realm of the giants. Table 11.2 summarizes the spacecraft missions to the outer solar system.

Table 11.2: Missions to the Giant Planets

Planet	Spacecraft[1]	Encounter Date	Type
Jupiter	Pioneer 10	December 1973	Flyby
	Pioneer 11	December 1974	Flyby
	Voyager 1	March 1979	Flyby
	Voyager 2	July 1979	Flyby
	Ulysses	February 1992	Flyby during gravity assist
	Galileo	December 1995	Orbiter and probe
	Cassini	December 2002	Flyby
	New Horizons	February 2007	Flyby during gravity assist
	Juno	July 2016	Orbiter
Saturn	Pioneer 11	September 1979	Flyby
	Voyager 1	November 1980	Flyby
	Voyager 2	August 1981	Flyby
	Cassini	July 2004	Orbiter
Uranus	Voyager 2	January 1986	Flyby
Neptune	Voyager 2	August 1989	Flyby

The challenges of exploring so far away from Earth are considerable. Flight times to the giant planets are measured in years to decades, rather than the months required to reach Venus or Mars. Even at the speed of light, messages take hours to pass between Earth and the spacecraft. If a problem develops near Saturn, for example, a wait of hours for the alarm to reach Earth and for instructions to be routed back to the spacecraft could spell disaster. Spacecraft to the outer solar system must therefore be highly reliable and capable of a greater degree of independence and autonomy. Outer solar system missions also must carry their own power sources since the Sun is too far away to provide enough energy, or else they must have very large arrays of solar cells. Heaters are required to keep instruments at proper operating temperatures, and spacecraft must have radio transmitters powerful enough to send their data to receivers on distant Earth.

The first spacecraft to investigate the regions past Mars were the NASA Pioneers 10 and 11, launched in 1972 and 1973 as pathfinders to Jupiter. One of their main objectives was simply to determine whether a spacecraft could actually navigate through the belt of asteroids that lies beyond Mars without getting destroyed by collisions with asteroidal dust. Another objective was to measure the radiation hazards in the *magnetosphere* (or zone of magnetic influence) of Jupiter. Both spacecraft passed through the asteroid belt without incident, but the energetic particles in Jupiter’s magnetic field nearly wiped out their electronics, providing information necessary for the safe design of subsequent missions.

Pioneer 10 flew past Jupiter in 1973, after which it sped outward toward the limits of the solar system. Pioneer 11 undertook a more

ambitious program, using the gravity of Jupiter to aim for Saturn, which it reached in 1979. The twin Voyager spacecraft launched the next wave of outer planet exploration in 1977. Voyagers 1 and 2 each carried 11 scientific instruments, including cameras and spectrometers, as well as devices to measure the characteristics of planetary magnetospheres. Since they kept going outward after their planetary encounters, these are now the most distant spacecraft ever launched by humanity.

Voyager 1 reached Jupiter in 1979 and used a gravity assist from that planet to take it on to Saturn in 1980. Voyager 2 arrived at Jupiter four months later, but then followed a different path to visit all the outer planets, reaching Saturn in 1981, Uranus in 1986, and Neptune in 1989. This trajectory was made possible by the approximate alignment of the four giant planets on the same side of the Sun. About once every 175 years, these planets are in such a position, and it allows a single spacecraft to visit them all by using gravity-assisted flybys to adjust course for each subsequent encounter; such a maneuver has been nicknamed a “Grand Tour” by astronomers.

Link to Learning

The Jet Propulsion Laboratory has a nice video called Voyager: The Grand Tour that describes the Voyager mission and what it found.

Making Connections: Engineering and Space Science: Teaching an Old Spacecraft New Tricks

By the time Voyager 2 arrived at Neptune in 1989, 12 years after its launch, the spacecraft was beginning to show signs of old age. The arm on which the camera and other instruments were located was “arthritic”: it could no longer move easily in all directions. The communications system was “hard of hearing”: part of its radio receiver had stopped working. The “brains” had significant “memory loss”: some of the onboard computer memory had failed. And the whole spacecraft was beginning to run out of energy: its generators had begun showing serious signs of wear.

To make things even more of a challenge, Voyager’s mission at Neptune was in many ways the most difficult of all four flybys. For example, since sunlight at Neptune is 900 times weaker than at Earth, the onboard camera had to take much longer exposures in this light-starved environment. This was a nontrivial requirement, given that the spacecraft was hurtling by Neptune at ten times the speed of a rifle bullet.

The solution was to swivel the camera backward at exactly the rate that would compensate for the forward motion of the spacecraft. Engineers had to preprogram the ship’s computer to execute an incredibly complex series of maneuvers for each image. The beautiful Voyager images of Neptune are a testament to the ingenuity of spacecraft engineers.

The sheer distance of the craft from its controllers on Earth was yet another challenge. Voyager 2 received instructions and sent back its data via on-board radio transmitter. The distance from Earth to Neptune is about 4.8 billion kilometers. Over this vast distance, the power that reached us from Voyager 2 at Neptune was approximately 10^{-16} watts, or 20 billion times less power than it takes to operate a digital watch. Thirty-eight different antennas on four continents were used by NASA to collect the faint signals from the spacecraft and decode the precious information about Neptune that they contained.

Enter the Orbiters: Galileo, Cassini, and Juno

The Pioneer and Voyager missions were flybys of the giant planets: they each produced only quick looks before the spacecraft sped onward. For more detailed studies of these worlds, we require spacecraft that can go into orbit around a planet. For Jupiter and Saturn, these orbiters were the Galileo, Cassini, and Juno spacecraft. To date, no orbiter missions have been started for Uranus and Neptune, although planetary scientists have expressed keen interest.

The Galileo spacecraft was launched toward Jupiter in 1989 and arrived in 1995. Galileo began its investigations by deploying an entry probe into Jupiter, for the first direct studies of the planet’s outer atmospheric layers.

The probe plunged at a shallow angle into Jupiter’s atmosphere, traveling at a speed of 50 kilometers *per second*—that’s fast enough to fly from New York to San Francisco in 100 seconds! This was the highest speed at which any probe has so far entered the atmosphere of a planet, and it put great demands on the heat shield protecting it. The high entry speed was a result of acceleration by the strong gravitational attraction of Jupiter.

Atmospheric friction slowed the probe within 2 minutes, producing temperatures at the front of its heat shield as high as 15,000 °C. As the probe's speed dropped to 2500 kilometers per hour, the remains of the glowing heat shield were jettisoned, and a parachute was deployed to lower the instrumented probe spacecraft more gently into the atmosphere (Figure 11.3). The data from the probe instruments were relayed to Earth via the main Galileo spacecraft.



(a) (b)

Figure 11.3 : Galileo Probe Falling into Jupiter and Juno Image of Jupiter's South Pole. (a) This artist's depiction shows the Galileo probe descending into the clouds via parachute just after the protective heat shield separated. The probe made its measurements of Jupiter's atmosphere on December 7, 1995. (b) This Juno image, taken in 2017 from about 100,000 kilometers above the cloudtops, shows the south polar region of Jupiter with its dramatic complex of storms and clouds. The enhanced-color image was processed for NASA/JPL by citizen scientist John Landino. (credit a: modification of work by NASA/Ames Research Center; credit b: modification of work by NASA/JPL-Caltech/SwRI/MSSS/John Landino)

The probe continued to operate for an hour, descending 200 kilometers into the atmosphere. A few minutes later the polyester parachute melted, and within a few hours the main aluminum and titanium structure of the probe vaporized to become a part of Jupiter itself. About 2 hours after receipt of the final probe data, the main spacecraft fired its retro-rockets so it could be captured into orbit around the planet, where its primary objectives were to study Jupiter's large and often puzzling moons.

The Cassini mission to Saturn (Figure 11.4), a cooperative venture between NASA and the European Space Agency, was similar to Galileo in its two-fold approach. Launched in 1997, Cassini arrived in 2004 and went into orbit around Saturn, beginning extensive studies of its rings and moons, as well as the planet itself. In January 2005, Cassini deployed an entry probe into the atmosphere of Saturn's large moon, Titan, where it successfully landed on the surface. (We'll discuss the probe and what it found in the chapter on Rings, Moons, and Pluto.)

The Voyager and Galileo missions to Jupiter were primarily designed to study the moons and the atmosphere of the planet. The next NASA mission, an orbiter called Juno, arrived at Jupiter in July 2016. In order to meet its objectives of studying the jovian magnetosphere, it has a very elongated (eccentric) 55-day orbit, that takes it from 4 thousand kilometers above the cloud tops out to 76 thousand kilometers. The orbit takes the craft over Jupiter's poles, giving us remarkable close-ups of the polar regions (previous spacecraft viewed the planet from lower latitudes).

Juno was originally designed without a camera, but fortunately scientists rectified this omission, adding a simple downward-looking color camera to use during close passes by Jupiter. Recognizing the value of such images, both scientific and artistic, it was decided to post the raw images and encourage "citizen scientists" to process them. The product has been many dramatic, brightly colored views of Jupiter, such as Figure 11.3.



Figure 11.4 : Earth as Seen from Saturn. This popular Cassini image shows Earth as a tiny dot (marked with an arrow) seen below Saturn's rings. It was taken in July 2013, when Saturn was 1.4 billion kilometers from Earth. (credit: modification of work by NASA/JPL-Caltech/Space Science Institute)

Footnotes

- 1Both the Ulysses and the New Horizons spacecraft (designed to study the Sun and Pluto, respectively) flew past Jupiter for a gravity boost (gaining energy by "stealing" a little bit from the giant planet's rotation).

This page titled 16.1: Exploring the Outer Planets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 11.2: Exploring the Outer Planets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

16.2: The Giant Planets

Learning Objectives

By the end of this section, you will be able to:

- Describe the basic physical characteristics, general appearance, and rotation of the giant planets
- Describe the composition and structure of Jupiter, Saturn, Uranus, and Neptune
- Compare and contrast the internal heat sources of the giant planets
- Describe the discovery and characteristics of the giant planets' magnetic fields

Let us now examine the four giant (or *jovian*) planets in more detail. Our approach is not just to catalog their characteristics, but to compare them with each other, noting their similarities and differences and attempting to relate their properties to their differing masses and distances from the Sun.

Basic Characteristics

The giant planets are very far from the Sun. Jupiter is more than five times farther from the Sun than Earth's distance (5 AU), and takes just under 12 years to circle the Sun. Saturn is about twice as far away as Jupiter (almost 10 AU) and takes nearly 30 years to complete one orbit. Uranus orbits at 19 AU with a period of 84 years, while Neptune, at 30 AU, requires 165 years for each circuit of the Sun. These long timescales make it difficult for us short-lived humans to study seasonal change on the outer planets.

Jupiter and Saturn have many similarities in composition and internal structure, although Jupiter is nearly four times more massive. Uranus and Neptune are smaller and differ in composition and internal structure from their large siblings. Some of the main properties of these four planets are summarized in Table 11.3.

Table 11.3: Basic Properties of the Jovian Planets

Planet	Distance (AU)	Period (years)	Diameter (km)	Mass (Earth = 1)	Density (g/cm^3)	Rotation (hours)
Jupiter	5.2	11.9	142,800	318	1.3	9.9
Saturn	9.5	29.5	120,540	95	0.7	10.7
Uranus	19.2	84.1	51,200	14	1.3	17.2
Neptune	30.0	164.8	49,500	17	1.6	16.1

Jupiter, the giant among giants, has enough mass to make 318 Earths. Its diameter is about 11 times that of Earth (and about one tenth that of the Sun). Jupiter's average density is 1.3 g/cm^3, much lower than that of any of the terrestrial planets. (Recall that water has a density of 1 g/cm^3.) Jupiter's material is spread out over a volume so large that about 1,300 Earths could fit within it.

Saturn's mass is 95 times that of Earth, and its average density is only 0.7 g/cm^3—the lowest of any planet. Since this is less than the density of water, Saturn would be light enough to float.

Uranus and Neptune each have a mass about 15 times that of Earth and, hence, are only 5% as massive as Jupiter. Their densities of 1.3 g/cm^3 and 1.6 g/cm^3, respectively, are much higher than that of Saturn. This is one piece of evidence that tells us that their composition must differ fundamentally from the gas giants. When astronomers began to discover other planetary systems (exoplanets), we found that planets the size of Uranus and Neptune are common, and that there are even more exoplanets intermediate in size between Earth and these ice giants, a type of planet not found in our solar system.

Appearance and Rotation

When we look at the planets, we see only their atmospheres, composed primarily of hydrogen and helium gas (see Figure 11.1). The uppermost clouds of Jupiter and Saturn, the part we see when looking down at these planets from above, are composed of ammonia crystals. On Neptune, the upper clouds are made of methane. On Uranus, we see no obvious cloud layer at all, but only a deep and featureless haze.

Seen through a telescope, Jupiter is a colorful and dynamic planet. Distinct details in its cloud patterns allow us to determine the rotation rate of its atmosphere at the cloud level, although such atmosphere rotation may have little to do with the spin of the underlying planet. Much more fundamental is the rotation of the mantle and core; these can be determined by periodic variations in radio waves coming from Jupiter, which are controlled by its magnetic field. Since the magnetic field (which we will discuss below) originates deep inside the planet, it shares the rotation of the interior. The rotation period we measure in this way is 9 hours 56 minutes, which gives Jupiter the shortest "day" of any planet. In the same way, we can measure that the underlying rotation period of Saturn is 10 hours 40 minutes. Uranus and Neptune have slightly longer rotation periods of about 17 hours, also determined from the rotation of their magnetic fields.

Link to Learning

A brief video made from Hubble Space Telescope photos shows the rotation of Jupiter with its many atmospheric features.

Remember that Earth and Mars have seasons because their spin axes, instead of "standing up straight," are tilted relative to the orbital plane of the solar system. This means that as Earth revolves around the Sun, sometimes one hemisphere and sometimes the other "leans into" the Sun.

What are the seasons like for the giant planets? The spin axis of Jupiter is tilted by only 3°, so there are no seasons to speak of. Saturn, however, does have seasons, since its spin axis is inclined at 27° to the perpendicular to its orbit. Neptune has about the same tilt as Saturn (29°); therefore, it experiences similar seasons (only more slowly). The strangest seasons of all are on Uranus, which has a spin axis tilted by 98° with respect to the north direction. Practically speaking, we can say that Uranus orbits on its side, and its ring and moon system follow along, orbiting about Uranus' equator (Figure 11.5).

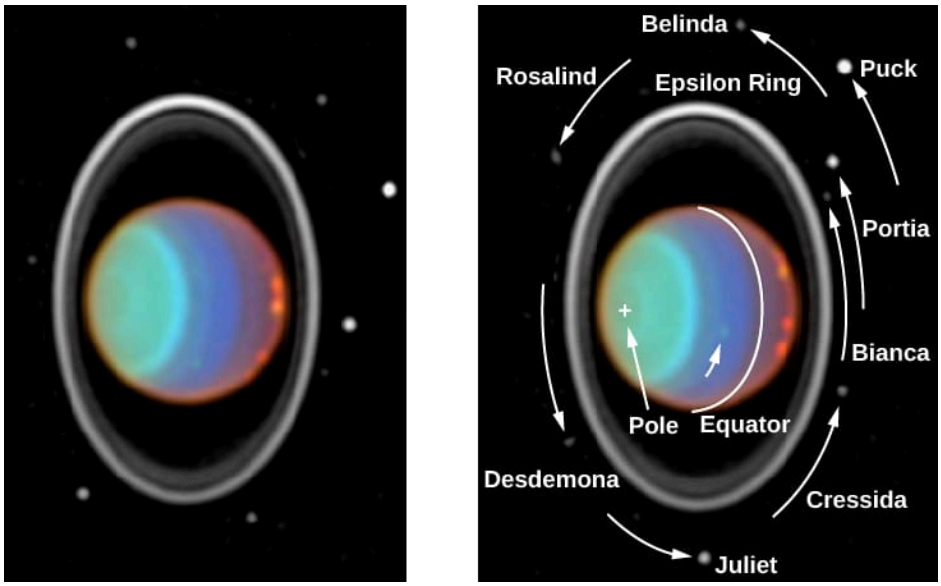


Figure 11.5 : Infrared Image of Uranus. The infrared camera on the Hubble Space Telescope took these false-color images of the planet Uranus, its ring system, and moons in 1997. The south pole of the planet (marked with a "+" on the right image) faces the Sun; its green color shows a strong local haze. The two images were taken 90 minutes apart, and during that time the five reddish clouds can be seen to rotate around the parallel to the equator. The rings (which are very faint in the visible light, but prominent in infrared) and eight moons can be seen around the equator. This was the "bull's eye" arrangement that Voyager saw as it approached Uranus in 1986. (credit: modification of work by Erich Karkoschka (University of Arizona), and NASA/ESA)

We don't know what caused Uranus to be tipped over like this, but one possibility is a collision with a large planetary body when

our system was first forming. Whatever the cause, this unusual tilt creates dramatic seasons. When Voyager 2 arrived at Uranus, its south pole was facing directly into the Sun. The southern hemisphere was experiencing a 21-year sunlit summer, while during that same period the northern hemisphere was plunged into darkness. For the next 21-year season, the Sun shines on Uranus' equator, and both hemispheres go through cycles of light and dark as the planet rotates (Figure 11.6). Then there are 21 years of an illuminated northern hemisphere and a dark southern hemisphere. After that the pattern of alternating day and night repeats.

Just as on Earth, the seasons are even more extreme at the poles. If you were to install a floating platform at the south pole of Uranus, for example, it would experience 42 years of light and 42 years of darkness. Any future astronauts crazy enough to set up camp there could spend most of their lives without ever seeing the Sun.

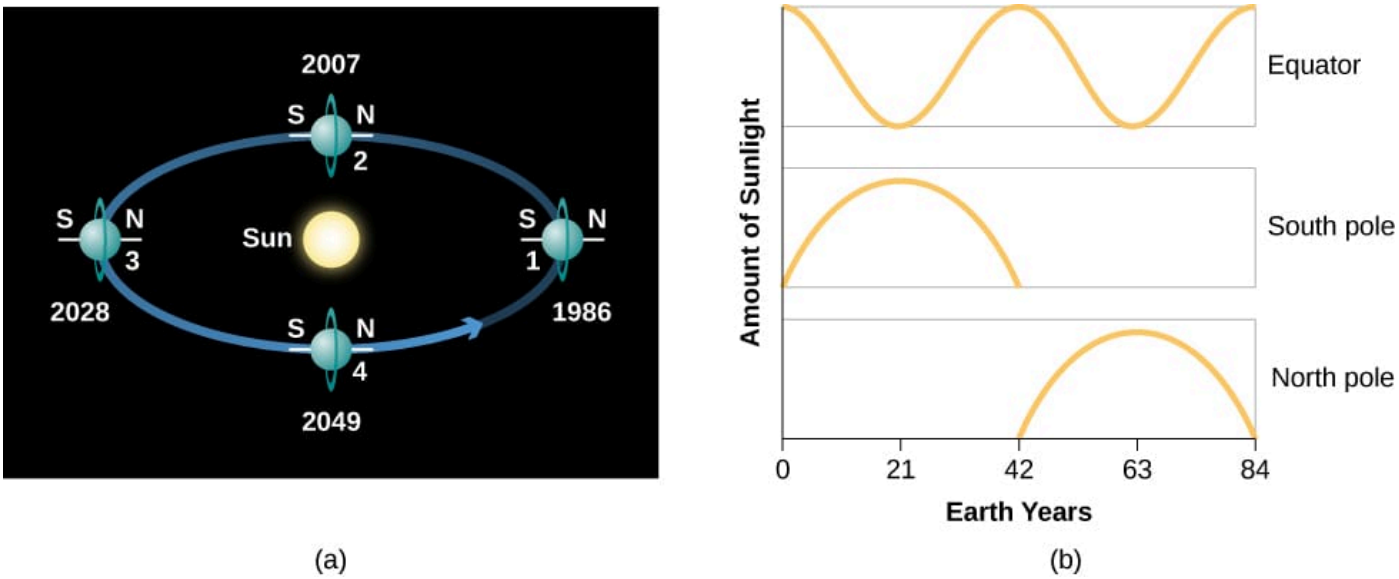


Figure 11.6 : Strange Seasons on Uranus. (a) This diagram shows the orbit of Uranus as seen from above. At the time Voyager 2 arrived (position 1), the South Pole was facing the Sun. As we move counterclockwise in the diagram, we see the planet 21 years later at each step. (b) This graph compares the amount of sunlight seen at the poles and the equator of Uranus over the course of its 84-year revolution around the Sun.

Composition and Structure

Although we cannot see into these planets, astronomers are confident that the interiors of Jupiter and Saturn are composed primarily of hydrogen and helium. Of course, these gases have been measured only in their atmosphere, but calculations first carried out more than 50 years ago showed that these two light gases are the only possible materials out of which a planet with the observed masses and densities of Jupiter and Saturn could be constructed.

The deep internal structures of these two planets are difficult to predict. This is mainly because these planets are so big that the hydrogen and helium in their centers become tremendously compressed and behave in ways that these gases can never behave on Earth. The best theoretical models we have of Jupiter's structure predict a central pressure greater than 100 million bars and a central density of about 31 g/cm^3. (By contrast, Earth's core has a central pressure of 4 million bars and a central density of 17 g/cm^3.)

At the pressures inside the giant planets, familiar materials can take on strange forms. A few thousand kilometers below the visible clouds of Jupiter and Saturn, pressures become so great that hydrogen changes from a gaseous to a liquid state. Still deeper, this liquid hydrogen is further compressed and begins to act like a metal, something it never does on Earth. (In a metal, electrons are not firmly attached to their parent nuclei but can wander around. This is why metals are such good conductors of electricity.) On Jupiter, the greater part of the interior is liquid metallic hydrogen.

Because Saturn is less massive, it has only a small volume of metallic hydrogen, but most of its interior is liquid. Uranus and Neptune are too small to reach internal pressures sufficient to liquefy hydrogen. We will return to the discussion of the metallic hydrogen layers when we examine the magnetic fields of the giant planets.

Each of these planets has a core composed of heavier materials, as demonstrated by detailed analyses of their gravitational fields.

Presumably these cores are the original rock-and-ice bodies that formed before the capture of gas from the surrounding nebula. The cores exist at pressures of tens of millions of bars. While scientists speak of the giant planet cores being composed of rock and ice, we can be sure that neither rock nor ice assumes any familiar forms at such pressures and temperatures. Remember that what is really meant by “rock” is any material made up primarily of iron, silicon, and oxygen, while the term “ice” in this chapter denotes materials composed primarily of the elements carbon, nitrogen, and oxygen in combination with hydrogen.

Figure 11.7 illustrates the likely interior structures of the four jovian planets. It appears that all four have similar cores of rock and ice. On Jupiter and Saturn, the cores constitute only a few percent of the total mass, consistent with the initial composition of raw materials shown in Table 11.1. However, most of the mass of Uranus and Neptune resides in these cores, demonstrating that the two outer planets were unable to attract massive quantities of hydrogen and helium when they were first forming.

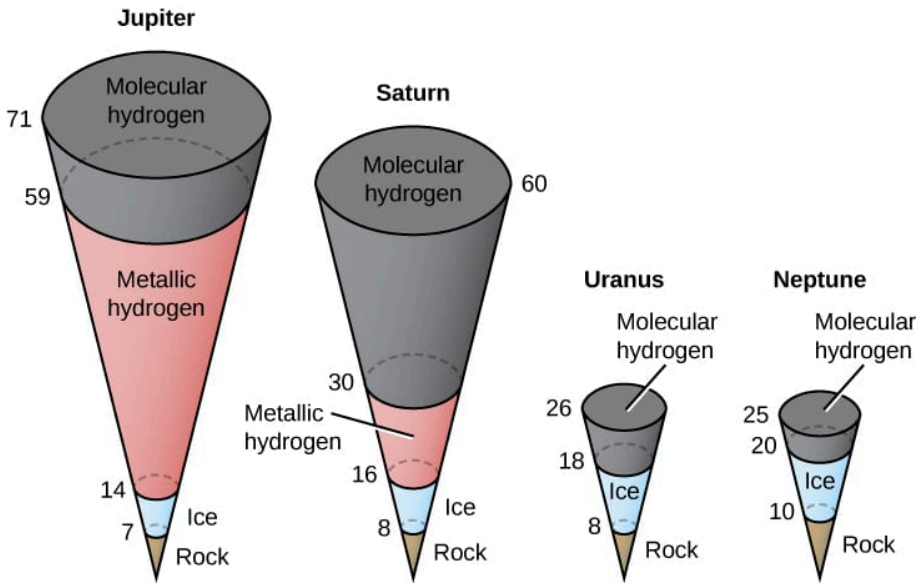


Figure 11.7 : Internal Structures of the Jovian Planets. Jupiter and Saturn are composed primarily of hydrogen and helium (but hydrogen dominates), but Uranus and Neptune consist in large part of compounds of carbon, nitrogen, and oxygen. (The diagrams are drawn to scale; numbers show radii in thousands of kilometers.)

Internal Heat Sources

Because of their large sizes, all the giant planets were strongly heated during their formation by the collapse of surrounding material onto their cores. Jupiter, being the largest, was the hottest. Some of this primordial heat can still remain inside such large planets. In addition, it is possible for giant, largely gaseous planets to generate heat after formation by slowly contracting. (With so large a mass, even a minuscule amount of shrinking can generate significant heat.) The effect of these internal energy sources is to raise the temperatures in the interiors and atmospheres of the planets higher than we would expect from the heating effect of the Sun alone.

Jupiter has the largest internal energy source, amounting to 4×10^{17} watts; that is, it is heated from inside with energy equivalent to 4 million billion 100-watt lightbulbs. This energy is about the same as the total solar energy absorbed by Jupiter. The atmosphere of Jupiter is therefore something of a cross between a normal planetary atmosphere (like Earth’s), which obtains most of its energy from the Sun, and the atmosphere of a star, which is entirely heated by an internal energy source. Most of the internal energy of Jupiter is primordial heat, left over from the formation of the planet 4.5 billion years ago.

Saturn has an internal energy source about half as large as that of Jupiter, which means (since its mass is only about one quarter as great) that it is producing twice as much energy per kilogram of material as does Jupiter. Since Saturn is expected to have much less primordial heat, there must be another source at work generating most of this 2×10^{17} watts of power. This source is the separation of helium from hydrogen in Saturn’s interior. In the liquid hydrogen mantle, the heavier helium forms droplets that sink toward the core, releasing gravitational energy. In effect, Saturn is still differentiating—letting lighter material rise and heavier material fall.

Uranus and Neptune are different. Neptune has a small internal energy source, while Uranus does not emit a measurable amount of internal heat. As a result, these two planets have almost the same atmospheric temperature, in spite of Neptune’s greater

distance from the Sun. No one knows why these two planets differ in their internal heat, but all this shows how nature can contrive to make each world a little bit different from its neighbors.

Magnetic Fields

Each of the giant planets has a strong magnetic field, generated by electric currents in its rapidly spinning interior. Associated with the magnetic fields are the planets' *magnetospheres*, which are regions around the planet within which the planet's own magnetic field dominates over the general interplanetary magnetic field. The magnetospheres of these planets are their largest features, extending millions of kilometers into space.

In the late 1950s, astronomers discovered that Jupiter was a source of radio waves that got more intense at longer rather than at shorter wavelengths—just the reverse of what is expected from thermal radiation (radiation caused by the normal vibrations of particles within all matter). Such behavior is typical, however, of the radiation emitted when high-speed electrons are accelerated by a magnetic field. We call this synchrotron radiation because it was first observed on Earth in particle accelerators, called synchrotrons. This was our first hint that Jupiter must have a strong magnetic field.

Later observations showed that the radio waves are coming from a region surrounding Jupiter with a diameter several times that of the planet itself (Figure 11.8). The evidence suggested that a vast number of charged atomic particles must be circulating around Jupiter, spiraling around the lines of force of a magnetic field associated with the planet. This is just what we observe happening, but on a smaller scale, in the Van Allen belt around Earth. The magnetic fields of Saturn, Uranus, and Neptune, discovered by the spacecraft that first passed close to these planets, work in a similar way, but are not as strong.

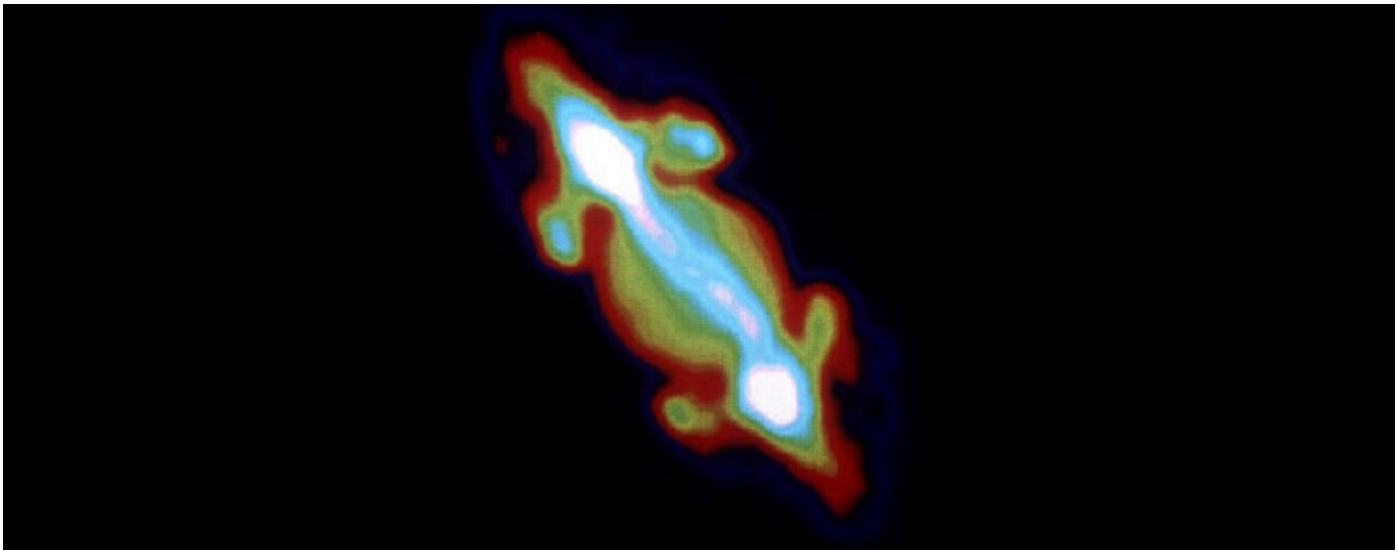


Figure 11.8 : Jupiter in Radio Waves. This false-color image of Jupiter was made with the Very Large Array (of radio telescopes) in New Mexico. We see part of the magnetosphere, brightest in the middle because the largest number of charged particles are in the equatorial zone of Jupiter. The planet itself is slightly smaller than the green oval in the center. Different colors are used to indicate different intensities of synchrotron radiation. (credit: modification of work by I. de Pater (UC Berkeley) NRAO, AUI, NSF)

Link to Learning

Learn more about the magnetosphere of Jupiter and why we continue to be interested in it from this brief NASA video.

Inside each magnetosphere, charged particles spiral around in the magnetic field; as a result, they can be accelerated to high energies. These charged particles can come from the Sun or from the neighborhood of the planet itself. In Jupiter's case, Io, one of its moons, turns out to have volcanic eruptions that blast charged particles into space and right into the jovian magnetosphere.

The axis of Jupiter's magnetic field (the line that connects the magnetic north pole with the magnetic south pole) is not aligned exactly with the axis of rotation of the planet; rather, it is tipped by about 10°. Uranus and Neptune have even greater magnetic tilts, of 60° and 55°, respectively. Saturn's field, on the other hand, is perfectly aligned with its rotation axis. Why different planets have such different magnetic tilts is not well understood.

The physical processes around the jovian planets turn out to be milder versions of what astronomers find in many distant objects, from the remnants of dead stars to the puzzling distant powerhouses we call quasars. One reason to study the magnetospheres of the giant planets and Earth is that they provide nearby accessible analogues of more energetic and challenging cosmic processes.

This page titled 16.2: The Giant Planets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 11.3: The Giant Planets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

16.3: Atmosphere of the Giant Planets

Learning Objectives

By the end of this section, you will be able to:

- Discuss the atmospheric composition of the giant planets
- Describe the cloud formation and atmospheric structure of the gas giants
- Characterize the giant planets' wind and weather patterns
- Understand the scale and longevity of storms on the giant planets

The atmospheres of the jovian planets are the parts we can observe or measure directly. Since these planets have no solid surfaces, their atmospheres are more representative of their general compositions than is the case with the terrestrial planets. These atmospheres also present us with some of the most dramatic examples of weather patterns in the solar system. As we will see, storms on these planets can grow bigger than the entire planet Earth.

Atmospheric Composition

When sunlight reflects from the atmospheres of the giant planets, the atmospheric gases leave their "fingerprints" in the spectrum of light. Spectroscopic observations of the jovian planets began in the nineteenth century, but for a long time, astronomers were not able to interpret the spectra they observed. As late as the 1930s, the most prominent features photographed in these spectra remained unidentified. Then better spectra revealed the presence of molecules of methane (CH_4) and ammonia (NH_3) in the atmospheres of Jupiter and Saturn.

At first astronomers thought that methane and ammonia might be the main constituents of these atmospheres, but now we know that hydrogen and helium are actually the dominant gases. The confusion arose because neither hydrogen nor helium possesses easily detected spectral features in the visible spectrum. It was not until the Voyager spacecraft measured the far-infrared spectra of Jupiter and Saturn that a reliable abundance for the elusive helium could be found.

The compositions of the two atmospheres are generally similar, except that on Saturn there is less helium as the result of the precipitation of helium that contributes to Saturn's internal energy source. The most precise measurements of composition were made on Jupiter by the Galileo entry probe in 1995; as a result, we know the abundances of some elements in the jovian atmosphere even better than we know those in the Sun.

Voyagers in Astronomy: James Van Allen: Several Planets under His Belt

The career of physicist James Van Allen spanned the birth and growth of the space age, and he played a major role in its development. Born in Iowa in 1914, Van Allen received his PhD from the University of Iowa. He then worked for several research institutions and served in the Navy during World War II.

After the war, Van Allen (Figure 11.9) was appointed Professor of Physics at the University of Iowa. He and his collaborators began using rockets to explore cosmic radiation in Earth's outer atmosphere. To reach extremely high altitudes, Van Allen designed a technique in which a balloon lifts and then launches a small rocket (the rocket is nicknamed "the rockoon").

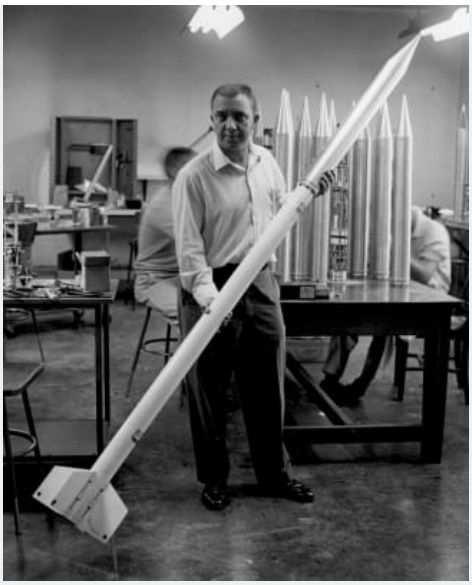


Figure 11.9 : James Van Allen (1914–2006). In this 1950s photograph, Van Allen holds a “rockoon.” (credit: modification of work by Frederick W. Kent Collection, University of Iowa Archives)

Over dinner one night in 1950, Van Allen and several colleagues came up with the idea of the International Geophysical Year (IGY), an opportunity for scientists around the world to coordinate their investigations of the physics of Earth, especially research done at high altitudes. In 1955, the United States and the Soviet Union each committed themselves to launching an Earth-orbiting satellite during IGY, a competition that began what came to be known as the space race. The IGY (stretched to 18 months) took place between July 1957 and December 1958.

The Soviet Union won the first lap of the race by launching Sputnik 1 in October 1957. The US government spurred its scientists and engineers to even greater efforts to get something into space to maintain the country’s prestige. However, the primary US satellite program, Vanguard, ran into difficulties: each of its early launches crashed or exploded. Simultaneously, a second team of rocket engineers and scientists had quietly been working on a military launch vehicle called Jupiter-C. Van Allen spearheaded the design of the instruments aboard a small satellite that this vehicle would carry. On January 31, 1958, Van Allen’s Explorer 1 became the first US satellite in space.

Unlike Sputnik, Explorer 1 was equipped to make scientific measurements of high-energy charged particles above the atmosphere. Van Allen and his team discovered a belt of highly charged particles surrounding Earth, and these belts now bear his name. This first scientific discovery of the space program made Van Allen’s name known around the world.

Van Allen and his colleagues continued to measure the magnetic and particle environment around planets with increasingly sophisticated spacecraft, including Pioneers 10 and 11, which made exploratory surveys of the environments of Jupiter and Saturn. Some scientists refer to the charged-particle zones around those planets as Van Allen belts as well. (Once, when Van Allen was giving a lecture at the University of Arizona, the graduate students in planetary science asked him if he would leave his belt at the school. It is now proudly displayed as the university’s “Van Allen belt.”)

Van Allen was a strong supporter of space science and an eloquent senior spokesperson for the American scientific community, warning NASA not to put all its efforts into human spaceflight, but to also use robotic spacecraft as productive tools for space exploration.

Clouds and Atmospheric Structure

The clouds of Jupiter (Figure 11.10) are among the most spectacular sights in the solar system, much beloved by makers of science-fiction films. They range in color from white to orange to red to brown, swirling and twisting in a constantly changing kaleidoscope of patterns. Saturn shows similar but much more subdued cloud activity; instead of vivid colors, its clouds have a nearly uniform butterscotch hue (Figure 11.11).



Figure 11.10 : Jupiter's Colorful Clouds. The vibrant colors of the clouds on Jupiter present a puzzle to astronomers: given the cool temperatures and the composition of nearly 90% hydrogen, the atmosphere should be colorless. One hypothesis suggests that perhaps colorful hydrogen compounds rise from warm areas. The actual colors are a bit more muted, as shown in Figure 11.2. (credit: modification of work by Voyager Project, JPL, and NASA)

Different gases freeze at different temperatures. At the temperatures and pressures of the upper atmospheres of Jupiter and Saturn, methane remains a gas, but ammonia can condense and freeze. (Similarly, water vapor condenses high in Earth's atmosphere to produce clouds of ice crystals.) The primary clouds that we see around these planets, whether from a spacecraft or through a telescope, are composed of frozen ammonia crystals. The ammonia clouds mark the upper edge of the planets' tropospheres; above that is the stratosphere, the coldest part of the atmosphere. (These layers were initially defined in Earth as a Planet.)

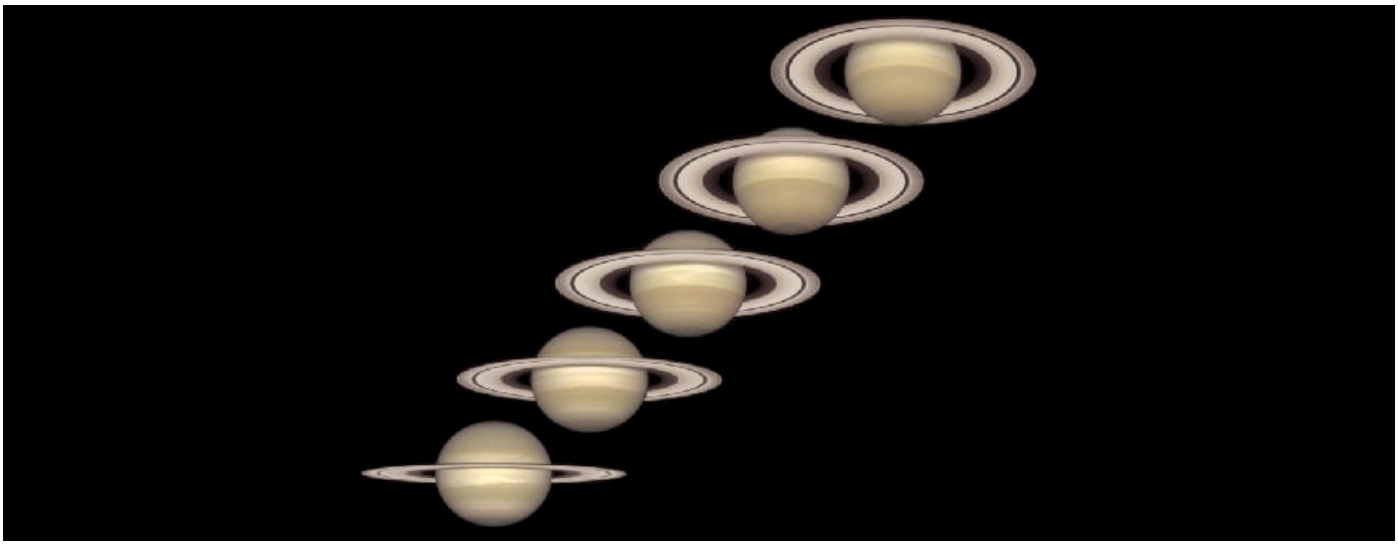


Figure 11.11 : Saturn over Five Years. These beautiful images of Saturn were recorded by the Hubble Space Telescope between 1996 and 2000. Since Saturn is tilted by 27°, we see the orientation of Saturn's rings around its equator change as the planet moves along its orbit. Note the horizontal bands in the atmosphere. (credit: modification of work by NASA and The Hubble Heritage Team (STScI/AURA))

The diagrams in Figure 11.12 show the structure and clouds in the atmospheres of all four jovian planets. On both Jupiter and Saturn, the temperature near the cloud tops is about 140 K (only a little cooler than the polar caps of Mars). On Jupiter, this cloud level is at a pressure of about 0.1 bar (one tenth the atmospheric pressure at the surface of Earth), but on Saturn it occurs lower in the atmosphere, at about 1 bar. Because the ammonia clouds lie so much deeper on Saturn, they are more difficult to see, and the overall appearance of the planet is much blander than is Jupiter's appearance.

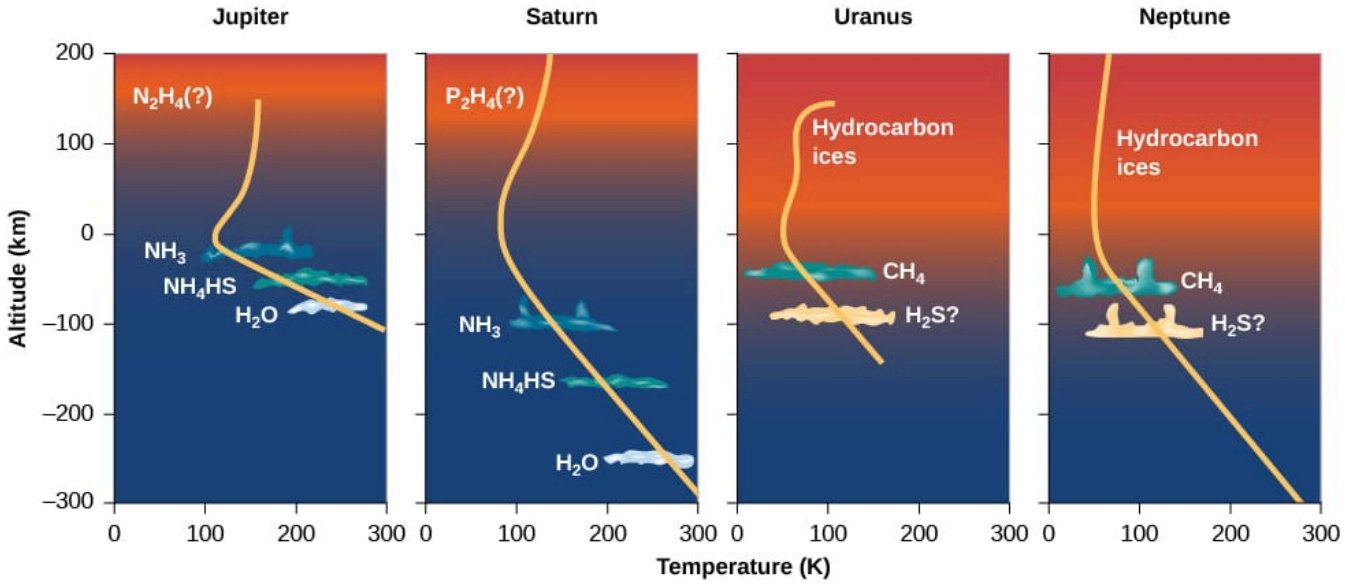


Figure 11.12 : Atmospheric Structure of the Jovian Planets. In each diagram, the yellow line shows how the temperature (see the scale on the bottom) changes with altitude (see the scale at the left). The location of the main layers on each planet is also shown.

Within the tropospheres of these planets, the temperature and pressure both increase with depth. Through breaks in the ammonia clouds, we can see tantalizing glimpses of other cloud layers that can form in these deeper regions of the atmosphere—regions that were sampled directly for Jupiter by the Galileo probe that fell into the planet.

As it descended to a pressure of 5 bars, the probe should have passed into a region of frozen water clouds, then below that into clouds of liquid water droplets, perhaps similar to the common clouds of the terrestrial troposphere. At least this is what scientists expected. But the probe saw no water clouds, and it measured a surprisingly low abundance of water vapor in the atmosphere. It soon became clear to the Galileo scientists that the probe happened to descend through an unusually dry, cloud-free region of the atmosphere—a giant downdraft of cool, dry gas. Andrew Ingersoll of Caltech, a member of the Galileo team, called this entry site the “desert” of Jupiter. It’s a pity that the probe did not enter a more representative region, but that’s the luck of the cosmic draw. The probe continued to make measurements to a pressure of 22 bars but found no other cloud layers before its instruments stopped working. It also detected lightning storms, but only at great distances, further suggesting that the probe itself was in a region of clear weather.

Above the visible ammonia clouds in Jupiter’s atmosphere, we find the clear stratosphere, which reaches a minimum temperature near 120 K. At still higher altitudes, temperatures rise again, just as they do in the upper atmosphere of Earth, because here the molecules absorb ultraviolet light from the Sun. The cloud colors are due to impurities, the product of chemical reactions among the atmospheric gases in a process we call photochemistry. In Jupiter’s upper atmosphere, photochemical reactions create a variety of fairly complex compounds of hydrogen and carbon that form a thin layer of smog far above the visible clouds. We show this smog as a fuzzy orange region in Figure 11.12; however, this thin layer does not block our view of the clouds beneath it.

The visible atmosphere of Saturn is composed of approximately 75% hydrogen and 25% helium, with trace amounts of methane, ethane, propane, and other hydrocarbons. The overall structure is similar to that of Jupiter. Temperatures are somewhat colder, however, and the atmosphere is more extended due to Saturn’s lower surface gravity. Thus, the layers are stretched out over a longer distance, as you can see in Figure 11.12. Overall, though, the same atmospheric regions, condensation cloud, and photochemical reactions that we see on Jupiter should be present on Saturn (Figure 11.13).

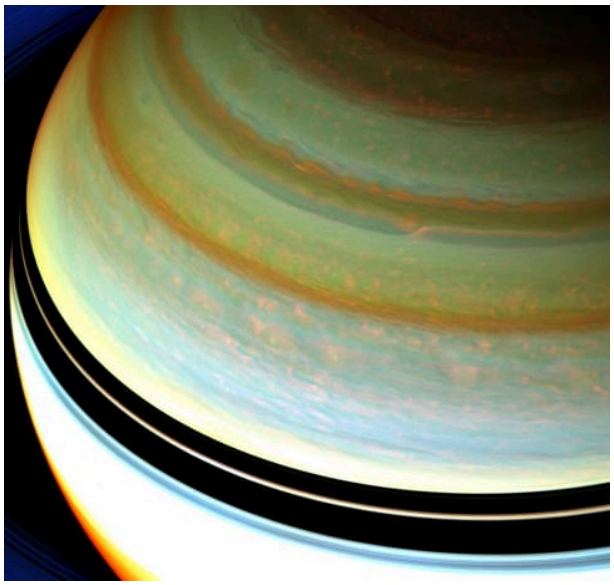


Figure 11.13 : Cloud Structure on Saturn. In this Cassini image, colors have been intensified, so we can see the bands and zones and storms in the atmosphere. The dark band is the shadow of the rings on the planet. (credit: NASA/JPL-Caltech/Space Science Institute)

Saturn has one anomalous cloud structure that has mystified scientists: a hexagonal wave pattern around the north pole, shown in Figure 11.14. The six sides of the hexagon are each longer than the diameter of Earth. Winds are also extremely high on Saturn, with speeds of up to 1800 kilometers per hour measured near the equator.

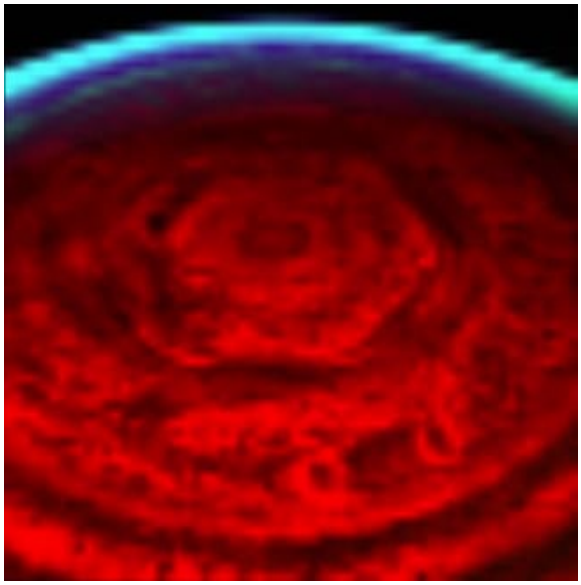


Figure 11.14 : Hexagon Pattern on Saturn's North Pole. In this infrared nighttime image from the Cassini mission, the path of Saturn's hexagonal jet stream is visible as the planet's north pole emerges from the darkness of winter. (credit: NASA/JPL/University of Arizona)

Link to Learning

See images of Saturn's hexagon with exaggerated color in this brief NASA video.

Unlike Jupiter and Saturn, Uranus is almost entirely featureless as seen at wavelengths that range from the ultraviolet to the infrared (see its rather boring image in Figure 11.1). Calculations indicate that the basic atmospheric structure of Uranus should resemble that of Jupiter and Saturn, although its upper clouds (at the 1-bar pressure level) are composed of methane rather than ammonia. However, the absence of an internal heat source suppresses up-and-down movement and leads to a very stable atmosphere with little visible structure.

Neptune differs from Uranus in its appearance, although their basic atmospheric temperatures are similar. The upper clouds are composed of methane, which forms a thin cloud layer near the top of the troposphere at a temperature of 70 K and a pressure of 1.5 bars. Most the atmosphere above this level is clear and transparent, with less haze than is found on Uranus. The scattering of sunlight by gas molecules lends Neptune a pale blue color similar to that of Earth's atmosphere (Figure 11.15). Another cloud layer, perhaps composed of hydrogen sulfide ice particles, exists below the methane clouds at a pressure of 3 bars.

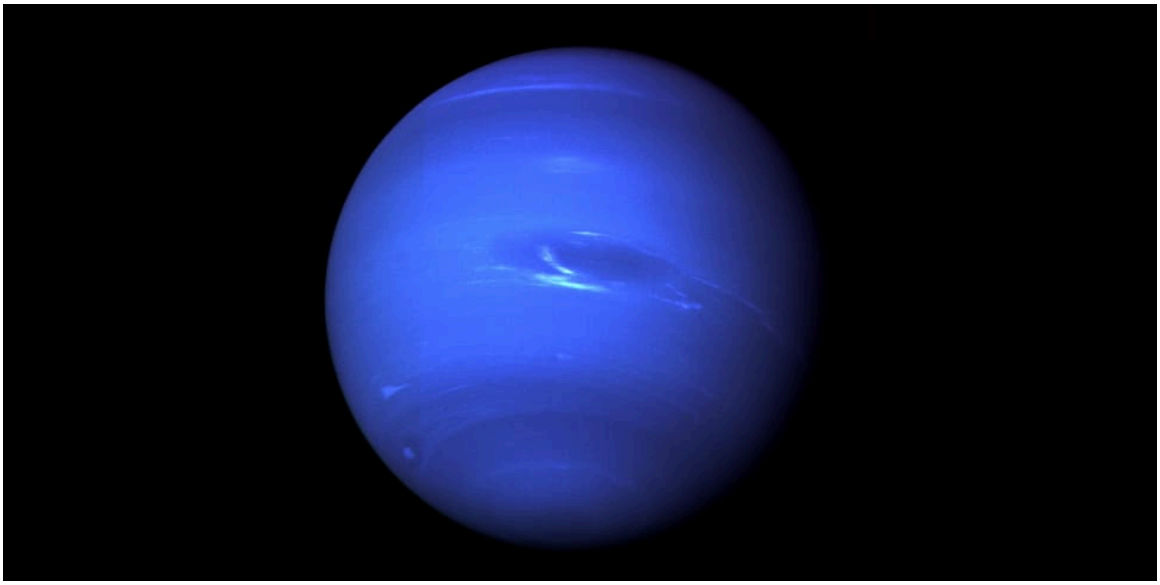


Figure 11.15 : Neptune. The planet Neptune is seen here as photographed by Voyager in 1989. The blue color, exaggerated with computer processing, is caused by the scattering of sunlight in the planet's upper atmosphere. (credit: modification of work by NASA)

Unlike Uranus, Neptune has an atmosphere in which convection currents—vertical drafts of gas—emanate from the interior, powered by the planet's internal heat source. These currents carry warm gas above the 1.5-bar cloud level, forming additional clouds at elevations about 75 kilometers higher. These high-altitude clouds form bright white patterns against the blue planet beneath. Voyager photographed distinct shadows on the methane cloud tops, permitting the altitudes of the high clouds to be calculated. Figure 11.16 is a remarkable close-up of Neptune's outer layers that could never have been obtained from Earth.

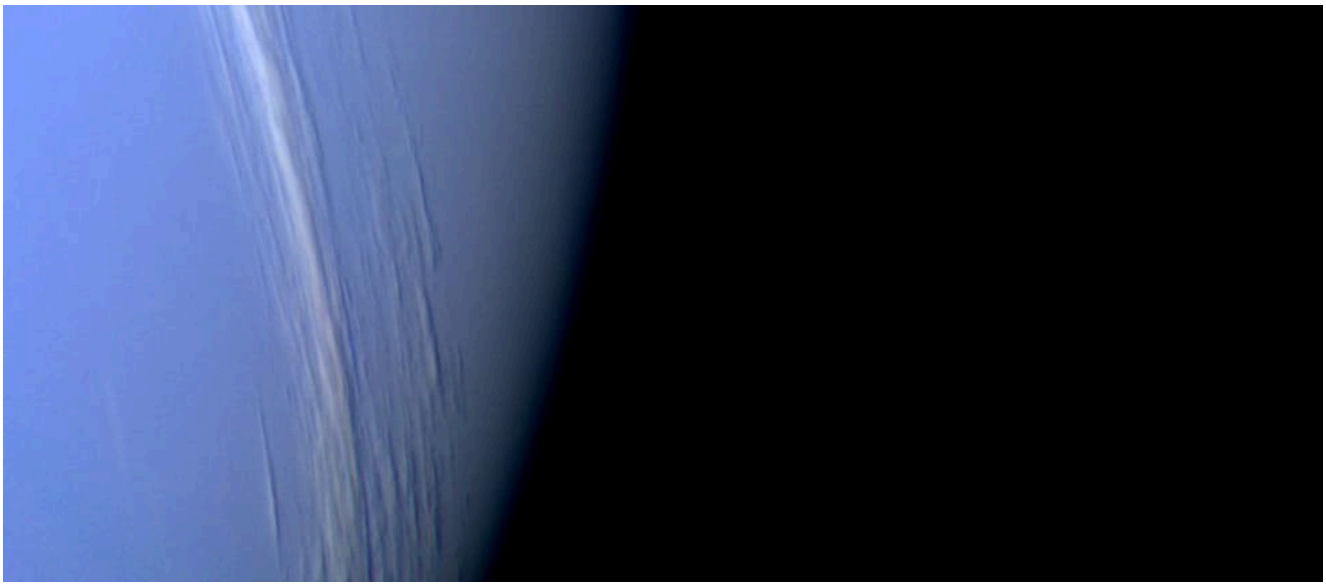


Figure 11.16 : High Clouds in the Atmosphere of Neptune. These bright, narrow cirrus clouds are made of methane ice crystals. From the shadows they cast on the thicker cloud layer below, we can measure that they are about 75 kilometers higher than the main clouds. (credit: modification of work by NASA/JPL)

Winds and Weather

The atmospheres of the jovian planets have many regions of high pressure (where there is more air) and low pressure (where there is less). Just as it does on Earth, air flows between these regions, setting up wind patterns that are then distorted by the rotation of the planet. By observing the changing cloud patterns on the jovian planets, we can measure wind speeds and track the circulation of their atmospheres.

The atmospheric motions we see on these planets are fundamentally different from those on the terrestrial planets. The giants spin faster, and their rapid rotation tends to smear out of the circulation into horizontal (east-west) patterns parallel to the equator. In addition, there is no solid surface below the atmosphere against which the circulation patterns can rub and lose energy (which is how tropical storms on Earth ultimately die out when they come over land).

As we have seen, on all the giants except Uranus, heat from the inside contributes about as much energy to the atmosphere as sunlight from the outside. This means that deep convection currents of rising hot air and falling cooler air circulate throughout the atmospheres of the planets in the vertical direction.

The main features of Jupiter's visible clouds (see Figure 11.2 and Figure 11.10, for example) are alternating dark and light bands that stretch around the planet parallel to the equator. These bands are semi-permanent features, although they shift in intensity and position from year to year. Consistent with the small tilt of Jupiter's axis, the pattern does not change with the seasons.

More fundamental than these bands are underlying east-west wind patterns in the atmosphere, which do not appear to change at all, even over many decades. These are illustrated in Figure 11.17, which indicates how strong the winds are at each latitude for the giant planets. At Jupiter's equator, a jet stream flows eastward with a speed of about 90 meters per second (300 kilometers per hour), similar to the speed of jet streams in Earth's upper atmosphere. At higher latitudes there are alternating east- and west-moving streams, with each hemisphere an almost perfect mirror image of the other. Saturn shows a similar pattern, but with a much stronger equatorial jet stream, as we noted earlier.

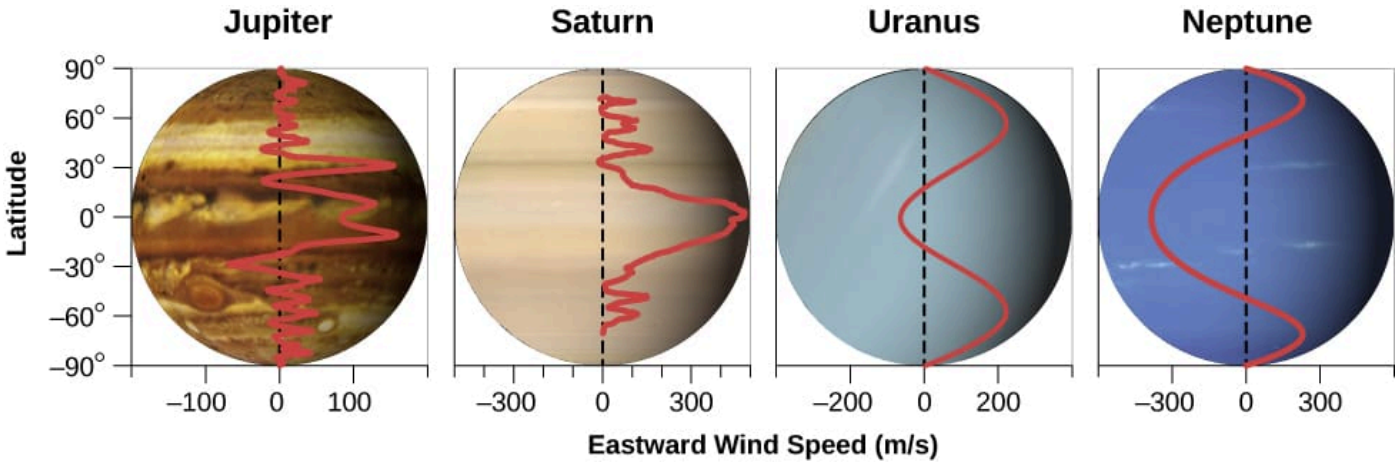


Figure 11.17 : Winds on the Giant Planets. This image compares the winds of the giant planets, illustrating that wind speed (shown on the horizontal axis) and wind direction vary with latitude (shown on the vertical axis). Winds are measured relative to a planet's internal rotation speed. A positive velocity means that the winds are blowing in the same direction as, but faster than, the planet's internal rotation. A negative velocity means that the winds are blowing more slowly than the planet's internal rotation.

The light zones on Jupiter are regions of upwelling air capped by white ammonia cirrus clouds. They apparently represent the tops of upward-moving convection currents.[2] The darker belts are regions where the cooler atmosphere moves downward, completing the convection cycle; they are darker because fewer ammonia clouds mean we can see deeper into the atmosphere, perhaps down to a region of ammonium hydrosulfide (NH_4SH) clouds. The Galileo probe sampled one of the clearest of these dry downdrafts.

In spite of the strange seasons induced by the 98° tilt of its axis, Uranus' basic circulation is parallel with its equator, as is the case on Jupiter and Saturn. The mass of the atmosphere and its capacity to store heat are so great that the alternating 42-year periods of sunlight and darkness have little effect. In fact, Voyager measurements show that the atmospheric temperature is even a few degrees higher on the dark winter side than on the hemisphere facing the Sun. This is another indication that the behavior of such giant planet atmospheres is a complex problem that we do not fully understand.

Neptune's weather is characterized by strong east-west winds generally similar to those observed on Jupiter and Saturn. The highest wind speeds near its equator reach 2100 kilometers per hour, even higher than the peak winds on Saturn. The Neptune equatorial jet stream actually approaches supersonic speeds (faster than the speed of sound in Neptune's air).

Giant Storms on Giant Planets

Superimposed on the regular atmospheric circulation patterns we have just described are many local disturbances—weather systems or *storms*, to borrow the term we use on Earth. The most prominent of these are large, oval-shaped, high-pressure regions on both Jupiter (Figure 11.18) and Neptune.

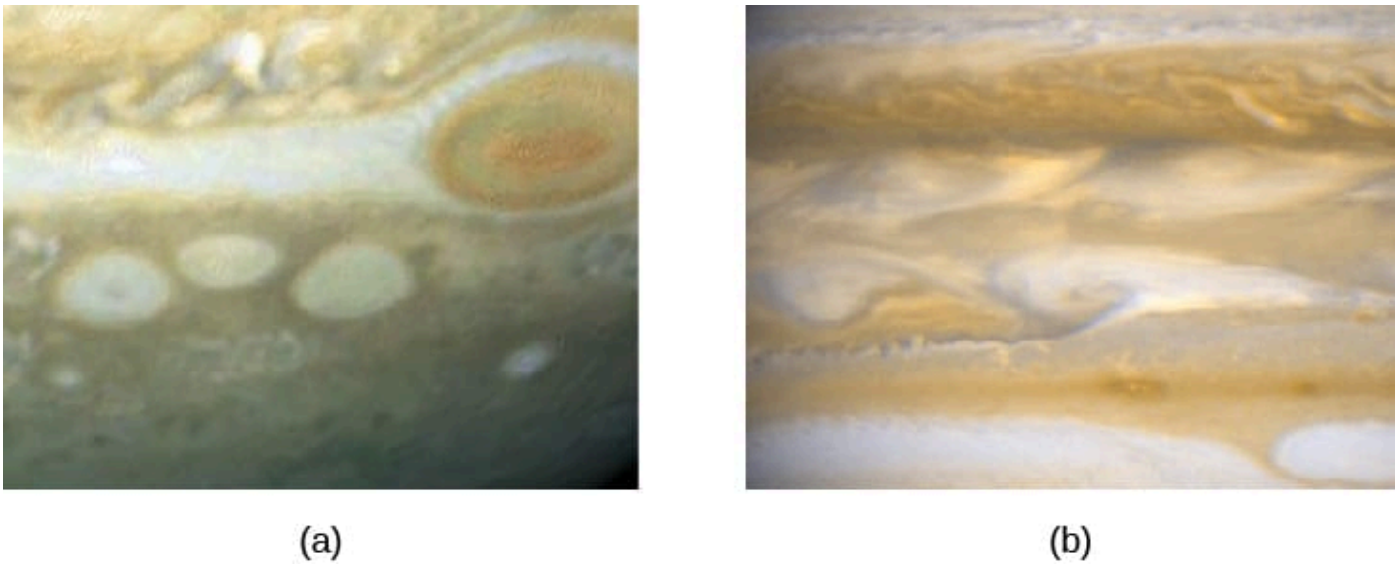


Figure 11.18 : Storms on Jupiter. Two examples of storms on Jupiter illustrate the use of enhanced color and contrast to bring out faint features. (a) The three oval-shaped white storms below and to the left of Jupiter's Great Red Spot are highly active, and moved closer together over the course of seven months between 1994 and 1995. (b) The clouds of Jupiter are turbulent and ever-changing, as shown in this Hubble Space Telescope image from 2007. (credit a: modification of work by Reta Beebe, Amy Simon (New Mexico State Univ.), and NASA; credit b: modification of work by NASA, ESA, and A. Simon-Miller (NASA Goddard Space Flight Center))

The largest and most famous of Jupiter's storms is the Great Red Spot, a reddish oval in the southern hemisphere that changes slowly; it was 25,000 kilometers long when Voyager arrived in 1979, but it had shrunk to 20,000 kilometers by the end of the Galileo mission in 2000 (Figure 11.19). The giant storm has persisted in Jupiter's atmosphere ever since astronomers were first able to observe it after the invention of the telescope, more than 300 years ago. However, it has continued to shrink and has become more nearly circular, raising speculation that we may see its end within a few decades. Measurements from the Juno spacecraft of variations in Jupiter's gravity field indicate that the depth of the Red Spot storm system is only a few hundred kilometers—a finding that challenges some of the models astronomers have made of this long-lasting weather system.



Figure 11.19 : Jupiter's Great Red Spot. This is the largest storm system on Jupiter, as seen during the Voyager spacecraft flyby. Below and to the right of the Red Spot is one of the white ovals, which are similar but smaller high-pressure features. The white oval is roughly the size of planet Earth, to give you a sense of the huge scale of the weather patterns we are seeing. The colors on the Jupiter image have been somewhat exaggerated here so astronomers (and astronomy students) can study their differences more effectively. See Figure 11.2 to get a better sense of the colors your eye would actually see near Jupiter. (credit: NASA/JPL)

In addition to its longevity, the Red Spot differs from terrestrial storms in being a high-pressure region; on our planet, such storms are regions of lower pressure. The Red Spot's counterclockwise rotation has a period of six days. Three similar but smaller disturbances (about as big as Earth) formed on Jupiter in the 1930s. They look like white ovals, and one can be seen clearly below and to the right of the Great Red Spot in Figure 11.19. In 1998, the Galileo spacecraft watched as two of these ovals collided and merged into one.

We don't know what causes the Great Red Spot or the white ovals, but we do have an idea how they can last so long once they form. On Earth, the lifetime of a large oceanic hurricane or typhoon is typically a few weeks, or even less when it moves over the continents and encounters friction with the land. Jupiter has no solid surface to slow down an atmospheric disturbance; furthermore, the sheer size of the disturbances lends them stability. We can calculate that on a planet with no solid surface, the lifetime of anything as large as the Red Spot should be measured in centuries, while lifetimes for the white ovals should be measured in decades, which is pretty much what we have observed.

Despite Neptune's smaller size and different cloud composition, Voyager showed that it had an atmospheric feature surprisingly similar to Jupiter's Great Red Spot. Neptune's Great Dark Spot was nearly 10,000 kilometers long (Figure 11.15). On both planets, the giant storms formed at latitude 20° S, had the same shape, and took up about the same fraction of the planet's diameter. The Great Dark Spot rotated with a period of 17 days, versus about 6 days for the Great Red Spot. When the Hubble Space Telescope examined Neptune in the mid-1990s, however, astronomers could find no trace of the Great Dark Spot on their images.

Although many of the details of the weather on the jovian planets are not yet understood, it is clear that if you are a fan of dramatic weather, these worlds are the place to look. We study the features in these atmospheres not only for what they have to teach us about conditions in the jovian planets, but also because we hope they can help us understand the weather on Earth just a bit better.

Example 11.1: Storms and Winds

The wind speeds in circular storm systems can be formidable on both Earth and the giant planets. Think about our big terrestrial hurricanes. If you watch their behavior in satellite images shown on weather outlets, you will see that they require about one day to rotate. If a storm has a diameter of 400 km and rotates once in 24 h, what is the wind speed?

Solution

Speed equals distance divided by time. The distance in this case is the circumference ($2\pi R$ or πd), or approximately 1250 km, and the time is 24 h, so the speed at the edge of the storm would be about 52 km/h. Toward the center of the storm, the wind speeds can be much higher.

Exercise 16.3.1

Jupiter's Great Red Spot rotates in 6 d and has a circumference equivalent to a circle with radius 10,000 km. Calculate the wind speed at the outer edge of the spot.

Answer available at the end of this section.

Footnotes

- 2Recall from earlier chapters that convection is a process in which liquids, heated from underneath, have regions where hot material rises and cooler material descends. You can see convection at work if you heat oatmeal on a stovetop or watch miso soup boil.

Answers

Exercise 16.3.1

For the Great Red Spot of Jupiter, the circumference $(2\pi R)$ is about 63,000 km. Six d equals 144 h, suggesting a speed of about 436 km/h. This is much faster than wind speeds on Earth.

This page titled 16.3: Atmosphere of the Giant Planets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 11.4: Atmosphere of the Giant Planets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

16.4: Key Terms

photochemistry
: chemical changes caused by electromagnetic radiation

synchrotron radiation
: the radiation emitted by charged particles being accelerated in magnetic fields and moving at speeds near that of light

This page titled 16.4: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

17: MOONS

17.1: Ring and Moon Systems Introduced
17.2: The Gailean Moons of Jupiter
17.3: Titan and Triton

This page titled 17: Moons is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

17.1: Ring and Moon Systems Introduced

Learning Objectives

By the end of this section, you will be able to:

- Briefly describe the system of moons around each of the jovian planets
- Describe the basic composition of each jovian planet's ring system

The rings and moons (see the moons in Figure 12.2) of the outer solar system are not composed of the same materials as the mostly rocky objects in the inner solar system. We should expect this, since they formed in regions of lower temperature, cool enough so that large quantities of water ice were available as building materials. Most of these objects also contain dark, organic compounds mixed with their ice and rock. Don't be surprised, therefore, to find that many objects in the ring and moon systems are both icy and dark.

Roughly a third of the moons in the outer solar system are in *direct* or regular orbits; that is, they revolve about their parent planet in a west-to-east direction and in the plane of the planet's equator. The majority are irregular moons that orbit in a *retrograde* (east-to-west) direction or else have orbits of high eccentricity (more elliptical than circular) or high inclination (moving in and out of the planet's equatorial plane). These irregular moons are mostly located relatively far from their planet; they were probably formed elsewhere and subsequently captured by the planet they now orbit.

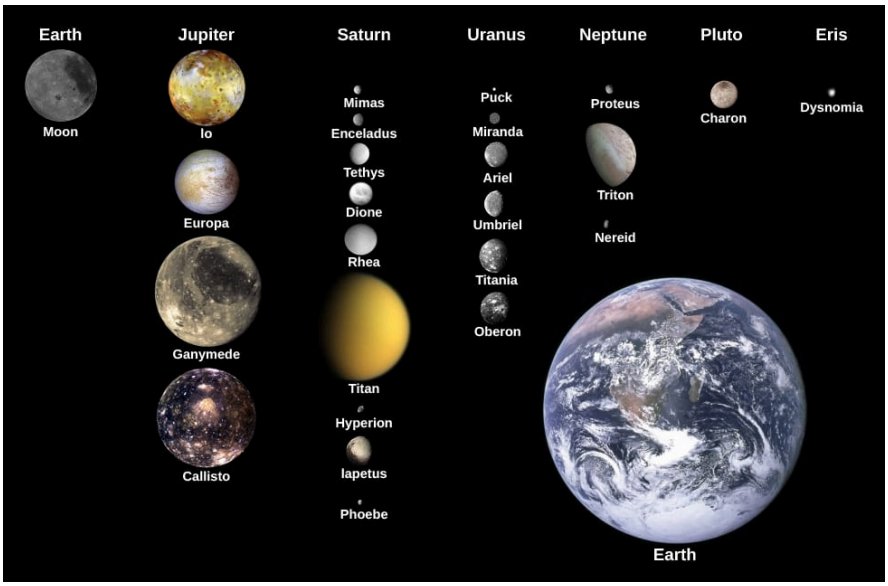


Figure 12.2 : Moons of the Solar System. This image shows some selected moons of our solar system and their comparison to the size of Earth's Moon and Earth itself. (credit: modification of work by NASA)

The Jupiter System

Jupiter has 79 known moons (that's the number as we write) and a faint ring. These include four large moons—Callisto, Ganymede, Europa, and Io (see Figure 12.1)—discovered in 1610 by Galileo and therefore often called the *Galilean moons*. The smaller of these, Europa and Io, are about the size of our Moon, while the larger, Ganymede and Callisto, are about the same size as the planet Mercury. Most of Jupiter's moons are much smaller. The majority are in retrograde orbits more than 20 million kilometers from Jupiter; these are very likely small captured asteroids.

The Saturn System

Saturn has at least 82 known moons in addition to a magnificent set of rings. The largest of the moons, Titan, is almost as big as Ganymede in Jupiter's system, and it is the only moon with a substantial atmosphere and lakes or seas of liquid hydrocarbons (such as methane and ethane) on the surface. Saturn has six other large regular moons with diameters between 400 and 1600 kilometers, a collection of small moons orbiting in or near the rings, and many captured strays similar to those of Jupiter. Mysteriously, one of Saturn's smaller moons, Enceladus, has active geysers of water being expelled into space.

The rings of Saturn, one of the most impressive sights in the solar system, are broad and flat, with a few major and many minor gaps. They are not solid, but rather a huge collection of icy fragments, all orbiting the equator of Saturn in a traffic pattern that makes rush hour in a big city look simple by comparison. Individual ring particles are composed primarily of water ice and are typically the size of ping-pong balls, tennis balls, and basketballs.

The Uranus System

The ring and moon system of Uranus is tilted at 98°, just like the planet itself. It consists of 11 rings and 27 currently known moons. The five largest moons are similar in size to the six regular moons of Saturn, with diameters of 500 to 1600 kilometers. Discovered in 1977, the rings of Uranus are narrow ribbons of dark material with broad gaps in between. Astronomers suppose that the ring particles are confined to these narrow paths by the gravitational effects of numerous small moons, many of which we have not yet glimpsed.

The Neptune System

Neptune has 14 known moons. The most interesting of these is Triton, a relatively large moon in a retrograde orbit—which is unusual. Triton has a very thin atmosphere, and active eruptions were discovered there by Voyager in its 1989 flyby. To explain its unusual characteristics, astronomers have suggested that Triton may have originated beyond the Neptune system, as a dwarf planet like Pluto. The rings of Neptune are narrow and faint. Like those of Uranus, they are composed of dark materials and are thus not easy to see.

This page titled 17.1: Ring and Moon Systems Introduced is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 12.2: Ring and Moon Systems Introduced by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

17.2: The Gailean Moons of Jupiter

Learning Objectives

By the end of this section, you will be able to:

- Describe the major features we can observe about Callisto and what we can deduce from them
- Explain the evidence for tectonic and volcanic activity on Ganymede
- Explain what may be responsible for the unusual features on the icy surface of Europa
- Describe the major distinguishing characteristic of Io
- Explain how tidal forces generate the geological activity we see on Europa and Io

From 1996 to 1999, the Galileo spacecraft careered through the jovian system on a complex but carefully planned trajectory that provided repeated close encounters with the large Galilean moons. More recently, the Juno spacecraft, also a Jupiter orbiter, has provided a few close looks at Ganymede and Europa. (Beginning in 2004, we received an even greater bonanza of information about Titan, obtained from the Cassini spacecraft and its Huygens probe, which landed on its surface. We include Titan, Saturn's one big moon, here for comparison.) Table 12.1 summarizes some basic facts about these large moons (plus our own Moon for comparison).

Table 12.1: The Largest Moons

Name	Diameter (km)	Mass (Earth's Moon = 1)	Density (g/cm^3)	Reflectivity (%)
Moon	3476	1.0	3.3	12
Callisto	4820	1.5	1.8	20
Ganymede	5270	2.0	1.9	40
Europa	3130	0.7	3.0	70
Io	3640	1.2	3.5	60
Titan	5150	1.9	1.9	20

Callisto: An Ancient, Primitive World

We begin our discussion of the Galilean moons with the outermost one, Callisto, not because it is remarkable but because it is not. This makes it a convenient object with which other, more active, worlds can be compared. Its distance from Jupiter is about 2 million kilometers, and it orbits the planet in 17 days. Like our own Moon, Callisto rotates in the same period as it revolves, so it always keeps the same face toward Jupiter. Callisto's day thus equals its month: 17 days. Its noontime surface temperature is only 130 K (about 140 °C below freezing), so that water ice is stable (it never evaporates) on its surface year round.

Callisto has a diameter of 4820 kilometers, almost the same as the planet Mercury (Figure 12.3). Yet its mass is only one-third as great, which means its density (the mass divided by the volume) must be only one-third as great as well. This tells us that Callisto has far less of the rocky and metallic materials found in the inner planets and must instead be an icy body through much of its interior. Callisto can show us how the geology of an icy object compares with those made primarily of rock.

Unlike the worlds we have studied so far, Callisto has not fully *differentiated* (separated into layers of different density materials). We can tell that it lacks a dense core from the details of its gravitational pull on the Galileo spacecraft. This surprised scientists, who expected that all the big icy moons would be differentiated. It should be easier for an icy body to differentiate than for a rocky one because the melting temperature of ice is so low. Only a little heating will soften the ice and get the process started, allowing the rock and metal to sink to the center while the slushy ice floats to the surface. Yet Callisto seems to have frozen solid before the process of differentiation was complete.

The surface of Callisto is covered with impact craters, like the lunar highlands. The survival of these craters tells us that an icy

object can retain impact craters on its surface. Callisto is unique among the planet-sized objects of the solar system in the apparent absence of interior forces to drive geological change. You might say that this moon was stillborn, and it has remained geologically dead for more than 4 billion years (Figure 12.3).

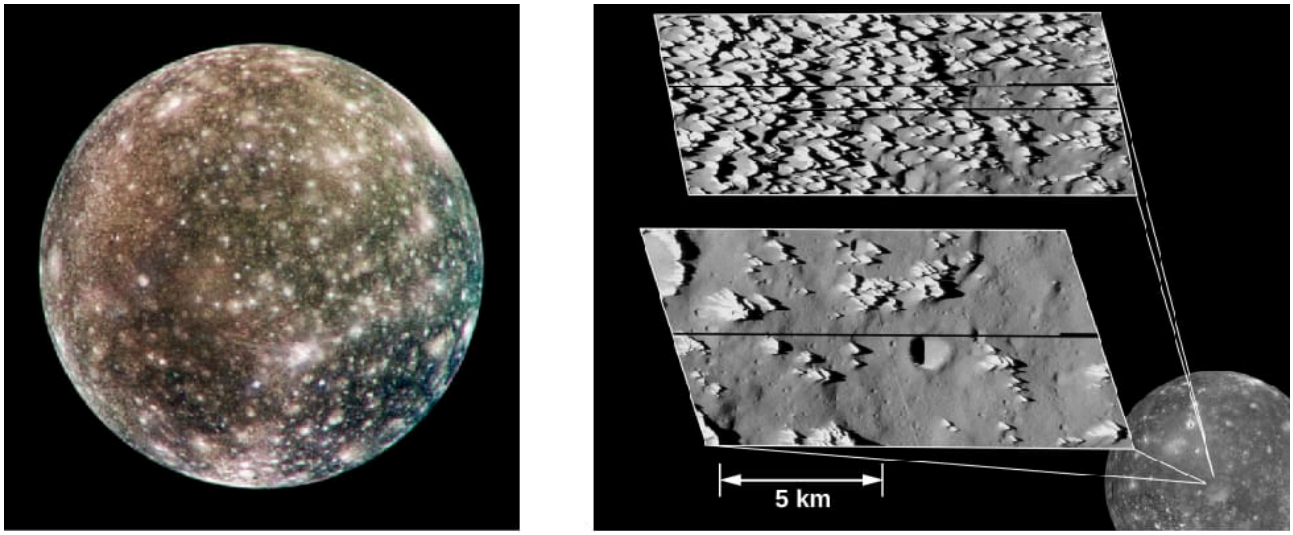


Figure 12.3 : Callisto. (a) Jupiter's outermost large moon shows a heavily cratered surface. Astronomers believe that the bright areas are mostly ice, while the darker areas are more eroded, ice-poor material. (b) These high-resolution images, taken by NASA's Galileo spacecraft in May 2001, show the icy spires (top) on Callisto's surface, with darker dust that has slid down as the ice erodes, collecting in the low-lying areas. The spires are about 80 to 100 meters tall. As the surface erodes even further, the icy spires eventually disappear, leaving impact craters exposed, as shown in the lower image. (credit a: modification of work by NASA/JPL/DLR; credit b: modification of work by NASA/JPL/Arizona State University, Academic Research Lab)

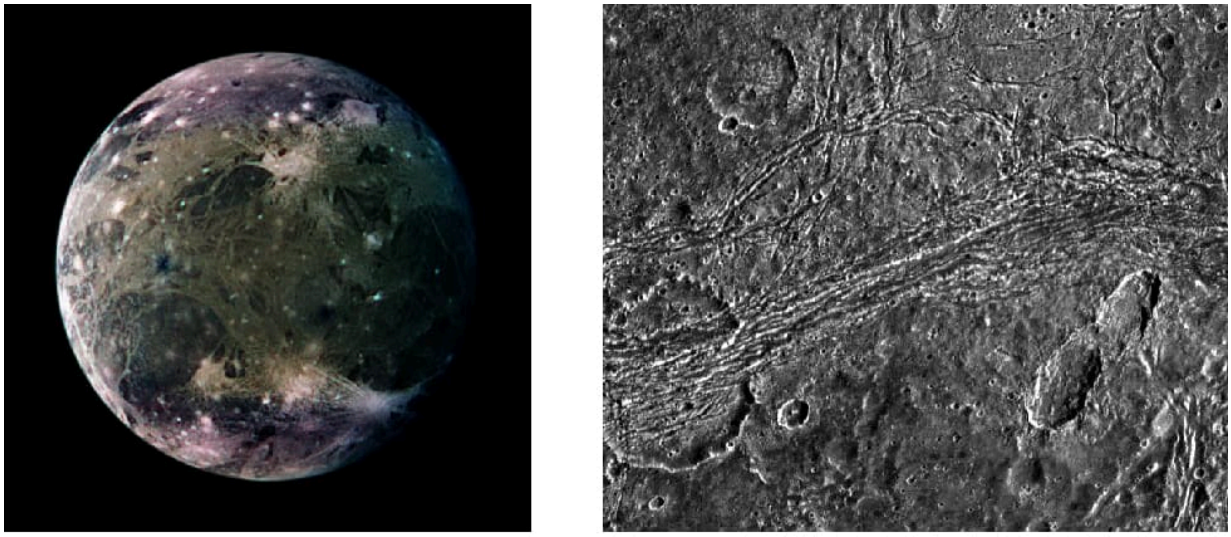
In thinking about ice so far from the Sun, we must take care not to judge its behavior from the much warmer ice we know and love on Earth. At the temperatures of the outer solar system, ice on the surface is nearly as hard as rock, and it behaves similarly. Ice on Callisto does not deform or flow like ice in glaciers on Earth.

Ganymede, the Largest Moon

Ganymede, the largest moon in the solar system, also shows a great deal of cratering (Figure 12.4). Recall from Other Worlds: An Introduction to the Solar System that we can use crater counts on solid worlds to estimate the age of the surface. The more craters, the longer the surface has been exposed to battering from space, and the older it must therefore be. About one-quarter of Ganymede's surface seems to be as old and heavily cratered as that of Callisto; the rest formed more recently, as we can tell by the sparse covering of impact craters as well as the relative freshness of those craters. If we judge from crater counts, this fresher terrain on Ganymede is somewhat younger than the lunar maria or the martian volcanic plains, perhaps 2 to 3 billion years old.

The differences between Ganymede and Callisto are more than skin deep. Ganymede is a differentiated world, like the terrestrial planets. Measurements of its gravity field tell us that the rock sank to form a core about the size of our Moon, with a mantle and crust of ice "floating" above it. In addition, the Galileo spacecraft discovered that Ganymede has a magnetic field, the sure signature of a partially molten interior. There is very likely liquid water trapped within the interior. Thus, Ganymede is not a dead world but rather a place of intermittent geological activity powered by an internal heat source. Some surface features could be as young as the surface of Venus (a few hundred million years).

The younger terrain was formed by tectonic and volcanic forces (Figure 12.4). In some places, the crust apparently cracked, flooding many of the craters with water from the interior. Extensive mountain ranges were formed from compression of the crust, forming long ridges with parallel valleys spaced a few kilometers apart. In some areas, older impact craters were split and pulled apart. There are even indications of large-scale crustal movements that are similar to the plate tectonics of Earth.



(a) (b)

Figure 12.4 : Ganymede. (a) This global view of Ganymede, the largest moon in the solar system, was taken by Voyager 2. The colors are enhanced to make spotting differences easier. Darker places are older, more heavily cratered regions; the lighter areas are younger (the reverse of our Moon). The brightest spots are sites of geologically recent impacts. (b) This close-up of Nicholson Regio on Ganymede shows an old impact crater (on the lower left-hand side) that has been split and pulled apart by tectonic forces. Against Ganymede's dark terrain, a line of grooves and ridges appears to cut through the crater, deforming its circular shape. (credit a: modification of work by NASA/JPL/DLR; credit b: modification of work by NASA/JPL/Brown University)

Why is Ganymede so different from Callisto? Possibly the small difference in size and internal heating between the two led to this divergence in their evolution. But more likely the gravity of Jupiter is to blame for Ganymede's continuing geological activity. Ganymede is close enough to Jupiter that *tidal forces* from the giant planet may have episodically heated its interior and triggered major convulsions on its crust.

A tidal force results from the unequal gravitational pull on two sides of a body. In a complex kind of modern dance, the large moons of Jupiter are caught in the varying gravity grip of both the giant planet and each other. This leads to gravitational flexing or kneading in their centers, which can heat them—an effect called tidal heating. (A fuller explanation is given in the section on Io.) We will see as we move inward to Europa and Io that the role of jovian tides becomes more important for moons close to the planet.

Europa, a Moon with an Ocean

Europa and Io, the inner two Galilean moons, are not icy worlds like most of the moons of the outer planets. With densities and sizes similar to our Moon, they appear to be predominantly rocky objects. How did they fail to acquire a majority share of the ice that must have been plentiful in the outer solar system at the time of their formation?

The most probable cause is Jupiter itself, which was hot enough to radiate a great deal of infrared energy during the first few million years after its formation. This infrared radiation would have heated the disk of material near the planet that would eventually coalesce into the closer moons. Thus, any ice near Jupiter was vaporized, leaving Europa and Io with compositions similar to planets in the inner solar system.

Despite its mainly rocky composition, Europa has an ice-covered surface, as astronomers have long known from examining spectra of sunlight reflected from it. In this it resembles Earth, which has a layer of water on its surface, but in Europa's case the water is capped by a thick crust of ice. There are very few impact craters in this ice, indicating that the surface of Europa is in a continual state of geological self-renewal. Judging from crater counts, the surface must be no more than a few million years old, and perhaps substantially less. In terms of its ability to erase impact craters, Europa is more geologically active than Earth.

When we look at close-up photos of Europa, we see a strange, complicated surface (Figure 12.5). For the most part, the icy crust is extremely smooth, but it is crisscrossed with cracks and low ridges that often stretch for thousands of kilometers. Some of these long lines are single, but most are double or multiple, looking rather like the remnants of a colossal freeway system.



Figure 12.5 : Evidence for an Ocean on Europa. (a) A close-up of an area called Conamara Chaos is shown here with enhanced color. This view is 70 kilometers wide in its long dimension. It appears that Conamara is a region where Europa's icy crust is (or recently was) relatively thin and there is easier access to the possible liquid or slushy ocean beneath. Not anchored to solid crust underneath, many of the ice blocks here seem to have slid or rotated from their original positions. In fact, the formations seen here look similar to views of floating sea-ice and icebergs in Earth's Arctic Ocean. (b) In this high-resolution view, the ice is *wrinkled* and crisscrossed by long ridges. Where these ridges intersect, we can see which ones are older and which younger; the younger ones cross over the older ones. While superficially this system of ridges resembles a giant freeway system on Europa, the ridges are much wider than our freeways and are a natural result of the flexing of the moon. (credit a: modification of work by NASA/JPL/University of Arizona; credit b: modification of work by NASA/JPL)

It is very difficult to make straight lines on a planetary surface. In discussing Mars, we explained that when Percival Lowell saw what appeared to him to be straight lines (the so-called martian "canals"), he attributed them to the engineering efforts of intelligent beings. We now know the lines on Mars were optical illusions, but the lines on Europa are real. These long cracks can form in the icy crust if it is floating without much friction on an ocean of liquid water (Figure 12.6).

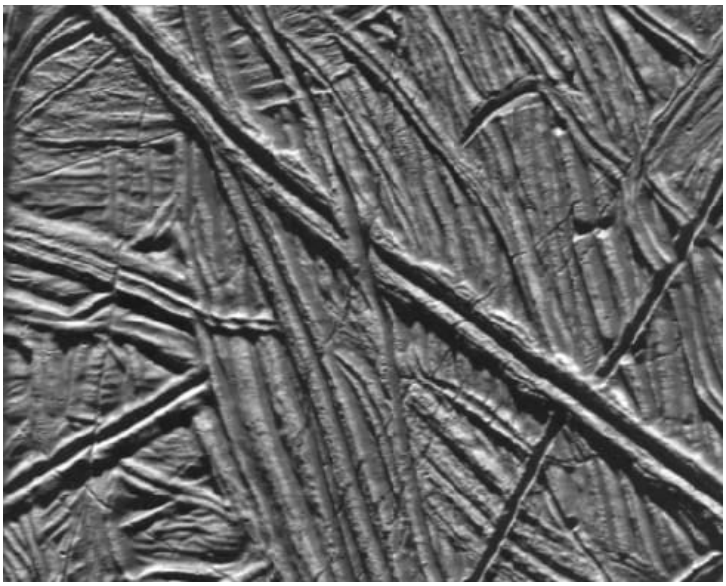


Figure 12.6 : Very High-Resolution Galileo Image of One Young Double Ridge on Europa. The area in this picture is only 15 kilometers across. It appears to have formed when viscous icy material was forced up through a long, straight crack in the crust. Note how the young ridge going from top left toward bottom right lies on top of older features, which are themselves on top of even older ones. (credit: modification of work by NASA/JPL)

The close-up Galileo images appear to confirm the existence of a global ocean. In many places, the surface of Europa looks just as we would expect for a thick layer of ice that was broken up into giant icebergs and ice floes and then refrozen in place. When the ice breaks, water or slush from below may be able to seep up through the cracks and make the ridges and multiple-line features we observe. Many episodes of ice cracking, shifting, rotating, and refreezing are required to explain the complexity we see. The icy crust might vary in thickness from a kilometer or so up to 20 kilometers. Further confirmation that a liquid ocean exists below the ice comes from measurements of the small magnetic field induced by Europa's interactions with the magnetosphere of Jupiter. The "magnetic signature" of Europa is that of a liquid water ocean, not one of ice or rock.

If Europa really has a large ocean of liquid water under its ice, then it may be the only place in the solar system, other than Earth, with really large amounts of liquid water.[1] To remain liquid, this ocean must be warmed by heat escaping from the interior of Europa. Hot (or at least warm) springs might be active there, analogous to those we have discovered in the deep oceans of Earth. The necessary internal heat is generated by tidal heating (see the discussion later in this chapter).

Link to Learning

A short film with planetary scientist Kevin Hand explains why Europa is so interesting for future exploration. Or listen to this more in-depth talk on Europa.

What makes the idea of an ocean with warm springs exciting is the discovery in Earth's oceans of large ecosystems clustered around deep ocean hot springs. Such life derives all its energy from the mineral-laden water and thrives independent of the sunlight shining on Earth's surface. Is it possible that similar ecosystems could exist today under the ice of Europa?

Many scientists now think that Europa is the most likely place beyond Earth to find life in the solar system. In response, NASA is designing a Europa mission to characterize its liquid ocean and its ice crust, and to identify locations where material from inside has risen to the surface. Such interior material might reveal direct evidence for microbial life. This "Europa Clipper" mission will not orbit Europa, where the intense energetic particles from Jupiter's magnetosphere would quickly destroy its electronics, but instead will swoop in for brief close flybys. It is planned for launch in 2024.

Io, a Volcanic Moon

Io, the innermost of Jupiter's Galilean moons, is in many ways a close twin of our Moon, with nearly the same size and density. We might therefore expect it to have experienced a similar history. Its appearance, as photographed from space, tells us another story, however (Figure 12.7). Instead of being a dead cratered world, Io turns out to have the highest level of volcanism in the solar system, greatly exceeding that of Earth.

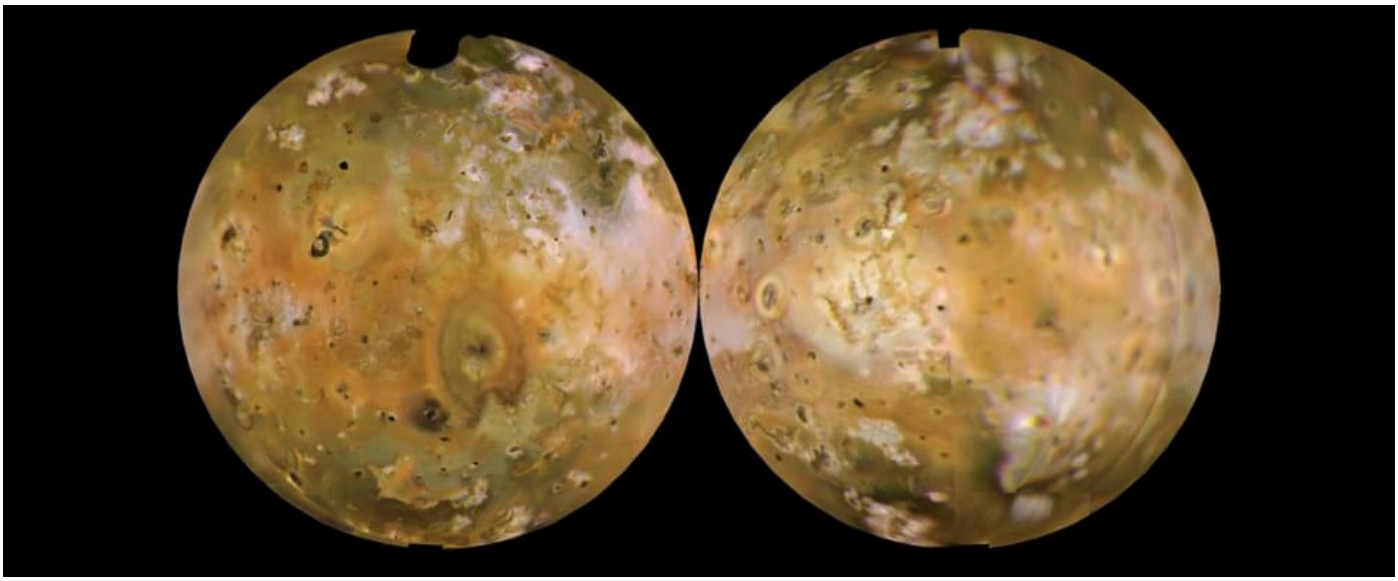


Figure 12.7 : Two Sides of Io. This composite image shows both sides of the volcanically active moon Io. The orange deposits are sulfur snow; the white is sulfur dioxide. (Carl Sagan once quipped that Io looks as if it desperately needs a shot of penicillin.) (credit: modification of work by NASA/JPL/USGS)

Io's active volcanism was discovered by the Voyager spacecraft. Eight volcanoes were seen erupting when Voyager 1 passed in March 1979, and six of these were still active four months later when Voyager 2 passed. With the improved instruments carried by the Galileo spacecraft, more than 50 eruptions were found during 1997 alone. Many of the eruptions produce graceful plumes that extend hundreds of kilometers out into space (Figure 12.8).

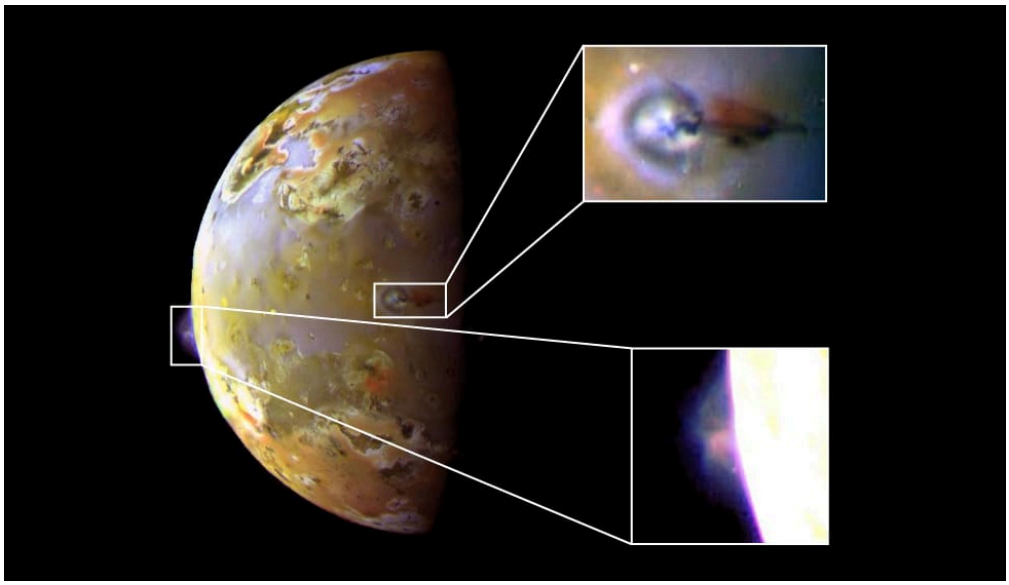


Figure 12.8 : Volcanic Eruptions on Io. This composite image from NASA's Galileo spacecraft shows close-ups (the two inset photos) of two separate volcanic eruptions on Jupiter's volcanic moon, Io. In the upper inset image, you can see a close up of a bluish plume rising about 140 kilometers above the surface of the volcano. In the lower inset image is the Prometheus plume, rising about 75 kilometers from Io's surface. The Prometheus plume is named for the Greek god of fire. (credit: modification of work by NASA/JPL)

Link to Learning

Watch a brief movie made from Voyager and Galileo data, showing a rotating Io with its dramatic surface features.

The Galileo data show that most of the volcanism on Io consists of hot silicate lava, like the volcanoes on Earth. Sometimes the hot lava encounters frozen deposits of sulfur and sulfur dioxide. When these icy deposits are suddenly heated, the result is great eruptive plumes far larger than any ejected from terrestrial volcanoes. As the rising plumes cool, the sulfur and sulfur dioxide recondense as solid particles that fall back to the surface in colorful "snowfalls" that extend as much as a thousand kilometers from the vent. Major new surface features were even seen to appear between Galileo orbits, as shown in Figure 12.9.

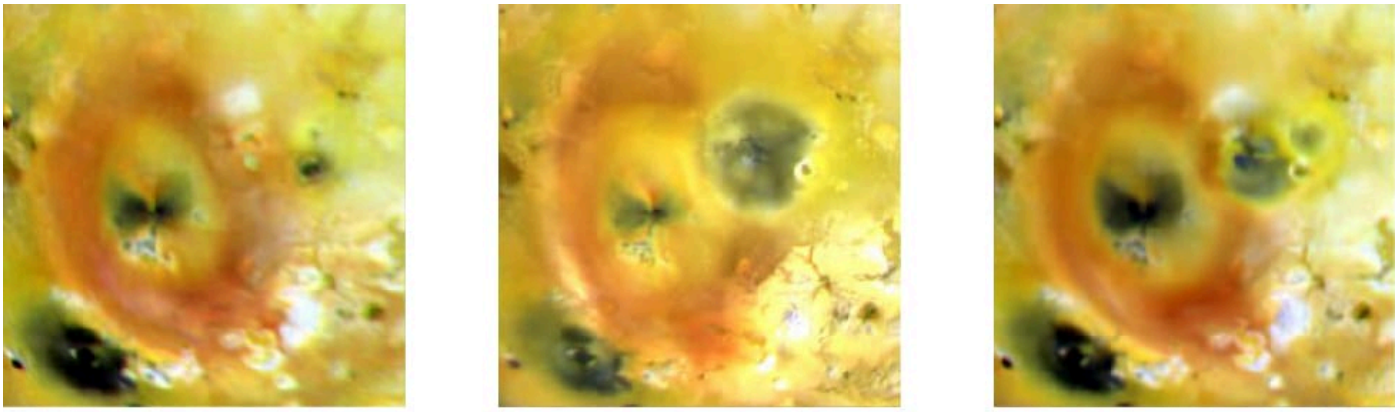


Figure 12.9 : Volcanic Changes on Io. These three images were taken of the same 1700-kilometer-square region of Io in April 1997, September 1997, and July 1999. The dark volcanic center called Pillan Patera experienced a huge eruption, producing a dark deposit some 400 kilometers across (seen as the grey area in the upper center of the middle image). In the right image, however, some of the new dark deposit is already being covered by reddish material from the volcano Pele. Also, a small unnamed volcano to the right of Pillan has erupted since 1997, and some of its dark deposit and a yellow ring around it are visible on the right image (to the right of the grey spot). The color range is exaggerated in these images. (credit: modification of work by NASA/JPL/University of Arizona)

As the Galileo mission drew to a close, controllers were willing to take risks in getting close to Io. Approaching this moon is a dangerous maneuver because the belts of atomic particles trapped in Jupiter's magnetic environment are at their most intense

near Io's orbit. Indeed, in its very first pass by Io, the spacecraft absorbed damaging radiation beyond its design levels. To keep the system working at all, controllers had to modify or disable various fault-protection software routines in the onboard computers. In spite of these difficulties, the spacecraft achieved four successful Io flybys, obtaining photos and spectra of the surface with unprecedented resolution.

Maps of Io reveal more than 100 recently active volcanoes. Huge flows spread out from many of these vents, covering about 25% of the moon's total surface with still-warm lava. From these measurements, it seems clear that the bright surface colors that first attracted attention to Io are the result of a thin veneer of sulfur compounds. The underlying volcanism is driven by eruptions of molten silicates, just like on Earth (Figure 12.10).



Figure 12.10 : Lava Fountains on Io. Galileo captured a number of eruptions along the chain of huge volcanic calderas (or pits) on Io called Tvashtar Catena in this false-color image combining infrared and visible light. The bright orange-yellow areas at left are places where fresh, hot lava is erupting from below ground. (credit: modification of work by NASA/JPL)

Tidal Heating

How can Io remain volcanically active in spite of its small size? The answer, as we hinted earlier, lies in the effect of gravity, through tidal heating. Io is about the same distance from Jupiter as our Moon is from Earth. Yet Jupiter is more than 300 times more massive than Earth, causing forces that pull Io into an elongated shape, with a several-kilometer-high bulge extending toward Jupiter.

If Io always kept exactly the same face turned toward Jupiter, this bulge would not generate heat. However, Io's orbit is not exactly circular due to gravitational perturbations (tugs) from Europa and Ganymede. In its slightly eccentric orbit, Io twists back and forth with respect to Jupiter, at the same time moving nearer and farther from the planet on each revolution. The twisting and flexing heat Io, much as repeated flexing of a wire coat hanger heats the wire.

After billions of years, this constant flexing and heating have taken their toll on Io, driving away water and carbon dioxide and other gases, so that now sulfur and sulfur compounds are the most volatile materials remaining. Its interior is entirely melted, and the crust itself is constantly recycled by volcanic activity.

In moving inward toward Jupiter from Callisto to Io, we have encountered more and more evidence of geological activity and internal heating, culminating in the violent volcanism on Io. Three of these surfaces are compared in Figure 12.11. Just as the character of the planets in our solar system depends in large measure on their distance from the Sun (and on the amount of heat they receive), so it appears that distance from a giant planet like Jupiter can play a large role in the composition and evolution of its moons (at least partly due to differences in internal heating of each moon by Jupiter's unrelenting tidal forces).

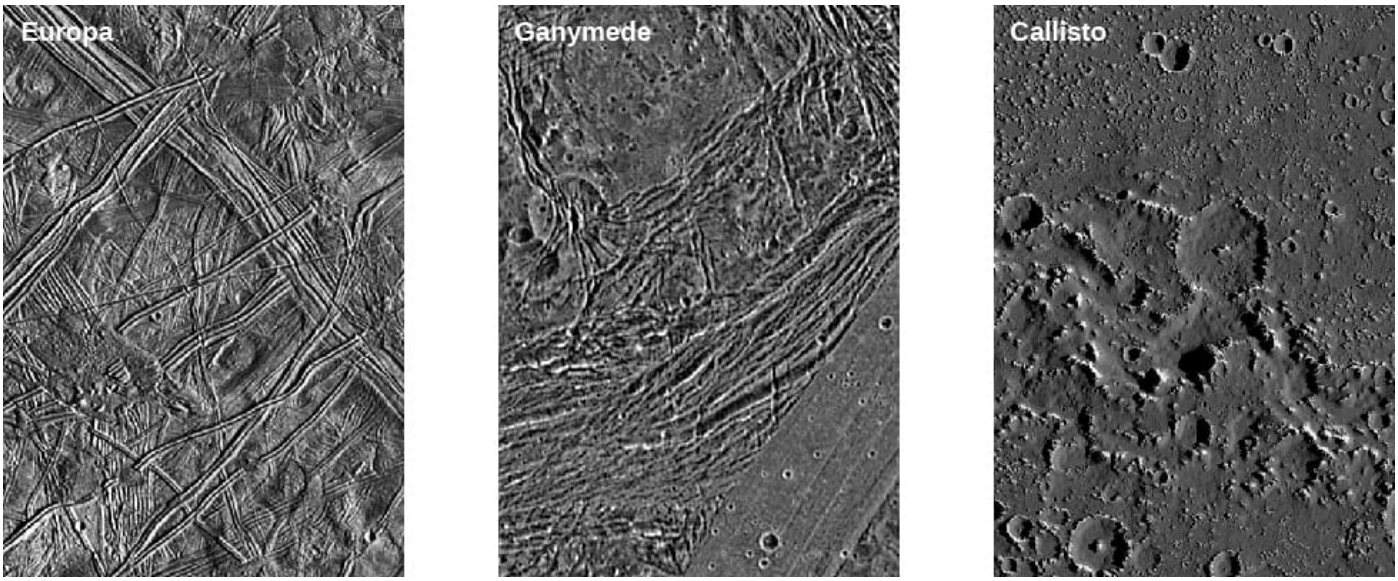


Figure 12.11 : Three Icy Moons. These Galileo images compare the surfaces of Europa, Ganymede, and Callisto at the same resolution. Note that the number of craters (and thus the age of the surface we see) increases as we go from Europa to Ganymede to Callisto. The Europa image is one of those where the system of cracks and ridges resembles a freeway system. (credit: modification of work by NASA/JPL/DLR)

Footnotes

- 1Ganymede and Saturn's moon Enceladus may have smaller amounts of liquid water under their surfaces.

This page titled 17.2: The Gailean Moons of Jupiter is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 12.3: The Gailean Moons of Jupiter by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

17.3: Titan and Triton

Learning Objectives

By the end of this section, you will be able to:

- Explain how the thick atmosphere of Titan makes bodies of liquid on its surface possible
- Describe what we learned from the landing on Titan with the Huygens probe
- Discuss the features we observed on the surface of Triton when Voyager 2 flew by

We shift our attention now to small worlds in the more distant parts of the solar system. Saturn's large moon Titan turns out to be a weird cousin of Earth, with many similarities in spite of frigid temperatures. The Cassini observations of Titan have provided some of the most exciting recent discoveries in planetary science. Neptune's moon Triton also has unusual characteristics and resembles Pluto, which we will discuss in the following section.

Titan, a Moon with Atmosphere and Hydrocarbon Lakes

Titan, first seen in 1655 by the Dutch astronomer Christiaan Huygens, was the first moon discovered after Galileo saw the four large moons of Jupiter. Titan has roughly the same diameter, mass, and density as Callisto or Ganymede. Presumably it also has a similar composition—about half ice and half rock. However, Titan is unique among moons, with a thick atmosphere and lakes and rivers and falling rain (although these are not composed of water but of hydrocarbons such as ethane and methane, which can stay liquid at the frigid temperatures on Titan). Titan is the only moon or planet other than Earth where we have found evidence of bodies of surface liquids.

The 1980 Voyager flyby of Titan determined that the surface density of its atmosphere is four times greater than that on Earth. The atmospheric pressure on this moon is 1.6 bars, higher than that on any other moon and, remarkably, even higher than that of the terrestrial planets Mars and Earth. The atmospheric composition is primarily nitrogen, an important way in which Titan's atmosphere resembles Earth's.

Also detected in Titan's atmosphere were carbon monoxide (CO), hydrocarbons (compounds of hydrogen and carbon) such as methane (CH_4), ethane (C_2H_6), and propane (C_3H_8), and nitrogen compounds such as hydrogen cyanide (HCN), cyanogen (C_2N_2), and cyanoacetylene (HC_3N). Their presence indicates an active chemistry in which sunlight interacts with atmospheric nitrogen and methane to create a rich mix of organic molecules. There are also multiple layers of hydrocarbon haze and clouds in the atmosphere, as illustrated in Figure 12.12.

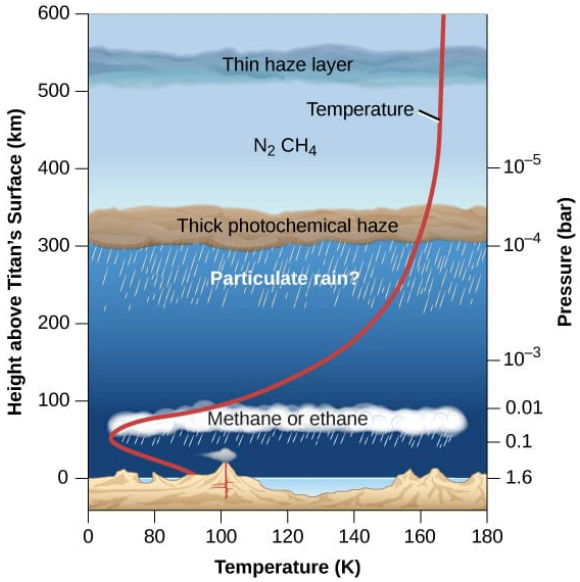


Figure 12.12 : Structure of Titan's Atmosphere. Some characteristics of Titan's atmosphere resemble those of Earth's atmosphere, although it is much colder than our planet. The red line indicates the temperature of Titan's atmosphere at different altitudes.

These Voyager discoveries motivated a much more ambitious exploration program using the NASA Cassini Saturn orbiter and a probe to land on Titan called Huygens, built by the European Space Agency. The orbiter, which included several cameras, spectrometers, and a radar imaging system, made dozens of close flybys of Titan between 2004 and 2015, each yielding radar and infrared images of portions of the surface (see Exploring the Outer Planets). The Huygens probe successfully descended by parachute through the atmosphere, photographing the surface from below the clouds, and landing on January 14, 2005. This was the first (and so far the only) spacecraft landing on a moon in the outer solar system.

At the end of its parachute descent, the 319-kilogram Huygens probe safely touched down, slid a short distance, and began sending data back to Earth, including photos and analyses of the atmosphere. It appeared to have landed on a flat, boulder-strewn plain, but both the surface and the boulders were composed of water ice, which is as hard as rock at the temperature of Titan (see Figure 12.13).

The photos taken during descent showed a variety of features, including drainage channels, suggesting that Huygens had landed on the shore of an ancient hydrocarbon lake. The sky was deep orange, and the brightness of the Sun was a thousand times less than sunlight on Earth (but still more than a hundred times brighter than under the full moon on Earth). Titan's surface temperature was 94 K (–179 °C). The warmer spacecraft heated enough of the ice where it landed for its instruments to measure released hydrocarbon gas. Measurements on the surface continued for more than an hour before the probe succumbed to the frigid temperature.

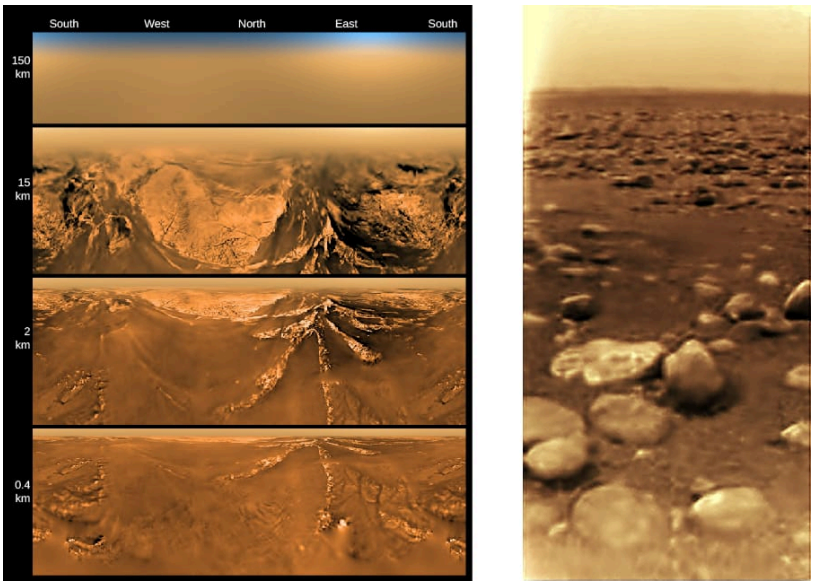


Figure 12.13 : Views of the Surface of Titan. The left image shows the views of Titan from the descent camera, in a flattened projection, at different altitudes. The right image, taken after landing, shows a boulder-strewn surface illuminated by faint reddish sunlight. The boulders are composed of water ice. (credit left: modification of work by ESA/NASA/JPL/University of Arizona; credit right: modification of work by ESA/NASA/JPL/University of Arizona; processed by Andrey Pivovarov)

Radar and infrared imaging of Titan from the Cassini orbiter gradually built up a picture of a remarkably active surface on this moon, complex and geologically young (Figure 12.14). There are large methane lakes near the polar regions that interact with the methane in the atmosphere, much as Earth's water oceans interact with the water vapor in our atmosphere. The presence of many erosional features indicates that atmospheric methane can condense and fall as rain, then flow down valleys to the big lakes. Thus, Titan has a low-temperature equivalent of the water cycle on Earth, with liquid on the surface that evaporates, forms clouds, and then condenses to fall as rain—but on Titan the liquid is a combination of methane, ethane, and a trace of other hydrocarbons. It is a weirdly familiar and yet utterly alien landscape.

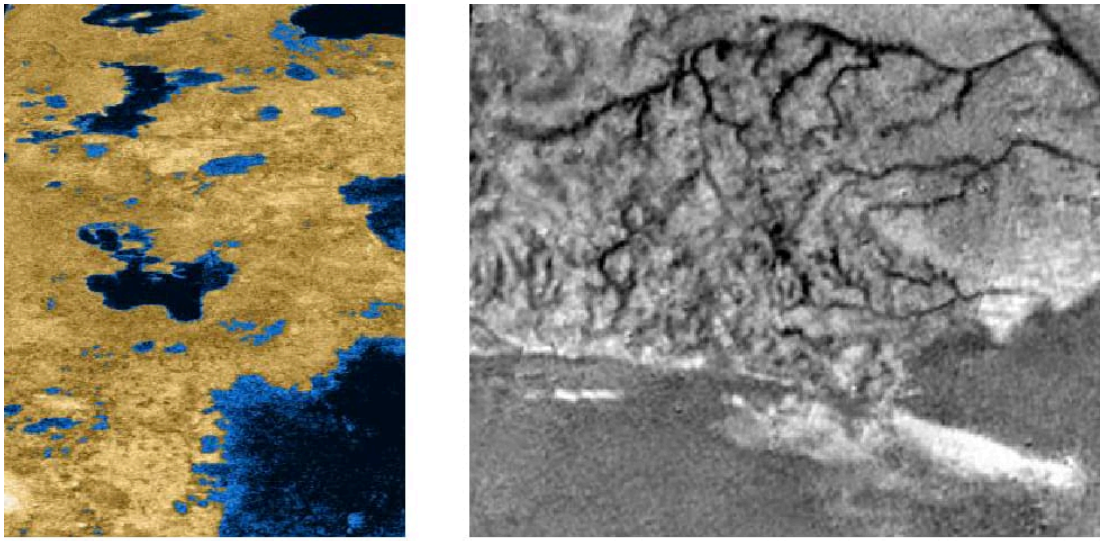


Figure 12.14 : Titan's Lakes. (a) This Cassini image from a September 2006 flyby shows the liquid lakes on Titan. Their composition is most likely a combination of methane and ethane. (Since this is a radar image, the colors are artificially added. The dark blue areas are the smooth surfaces of the liquid lakes, and yellow is the rougher solid terrain around them.) (b) This mosaic of Titan's surface from the Cassini-Huygens mission shows in detail a high ridge area and many narrow, sinuous erosion channels that appear to be part of a widespread network of "rivers" carved by flowing hydrocarbons. (credit a: modification of work by NASA/JPL-Caltech/USGS; credit b; modification of work by NASA/JPL/ESA/University of Arizona)

These discoveries raise the question of whether there could be life on Titan. Hydrocarbons are fundamental for the formation of

the large carbon molecules that are essential to life on our planet. However, the temperature on Titan is far too low for liquid water or for many of the chemical processes that are essential to life as we know it. There remains, though, an intriguing possibility that Titan might have developed a different form of low-temperature carbon-based life that could operate with liquid hydrocarbons playing the role of water. The discovery of such "life as we don't know it" could be even more exciting than finding life like ours on Mars. If such a truly alien life is present on Titan, its existence would greatly expand our understanding of the nature of life and of habitable environments.

NASA has selected a new mission to Titan for launch in 2027. Called *Dragonfly*, this mission is a drone that will fly in Titan's atmosphere, with emphasis on study of pre-biotic chemistry. Other future proposed missions include a balloon operating in the atmosphere and a "boat" floating in one of the Titan lakes.

Link to Learning

The Cassini mission scientists and the visual presentation specialists at NASA's Jet Propulsion Laboratory have put together some nice films from the images taken by Cassini and Huygens. See, for example, the Titan approach and the flyover of the Northern lakes district.

Triton and Its Volcanoes

Neptune's largest moon Triton (don't get its name confused with Titan) has a diameter of 2720 kilometers and a density of 2.1 g/cm^3, indicating that it's probably composed of about 75% rock mixed with 25% water ice. Measurements indicate that Triton's surface has the coldest temperature of any of the worlds our robot representatives have visited. Because its reflectivity is so high (about 80%), Triton reflects most of the solar energy that falls on it, resulting in a surface temperature between 35 and 40 K.

The surface material of Triton is made of frozen water, nitrogen, methane, and carbon monoxide. Methane and nitrogen exist as gas in most of the solar system, but they are frozen at Triton's temperatures. Only a small quantity of nitrogen vapor persists to form an atmosphere. Although the surface pressure of this atmosphere is only 16 millionths of a bar, this is sufficient to support thin haze or cloud layers.

Triton's surface, like that of many other moons in the outer solar system, reveals a long history of geological evolution (Figure 12.15). Although some impact craters are found, many regions have been flooded fairly recently by the local version of "lava" (perhaps water or water-ammonia mixtures). There are also mysterious regions of jumbled or mountainous terrain.



Figure 12.15 : Neptune's Moon Triton. This mosaic of Voyager 2 images of Triton shows a wide range of surface features. The pinkish area at the bottom is Triton's large southern polar cap. The south pole of Triton faces the Sun here, and the slight heating effect is driving some of the material northward, where it is colder. (credit: modification of work by NASA/JPL/USGS)

The Voyager flyby of Triton took place at a time when the moon's southern pole was tipped toward the Sun, allowing this part of the surface to enjoy a period of relative warmth. (Remember that "warm" on Triton is still outrageously colder than anything we experience on Earth.) A polar cap covers much of Triton's southern hemisphere, apparently evaporating along the northern edge. This polar cap may consist of frozen nitrogen that was deposited during the previous winter.

Remarkably, the Voyager images showed that the evaporation of Triton's polar cap generates geysers or volcanic plumes of nitrogen gas (see Figure 12.16). (Fountains of such gas rose about 10 kilometers high, visible in the thin atmosphere because dust from the surface rose with them and colored them dark.) These plumes differ from the volcanic plumes of Io in their composition and also in that they derive their energy from sunlight warming the surface rather than from internal heat.

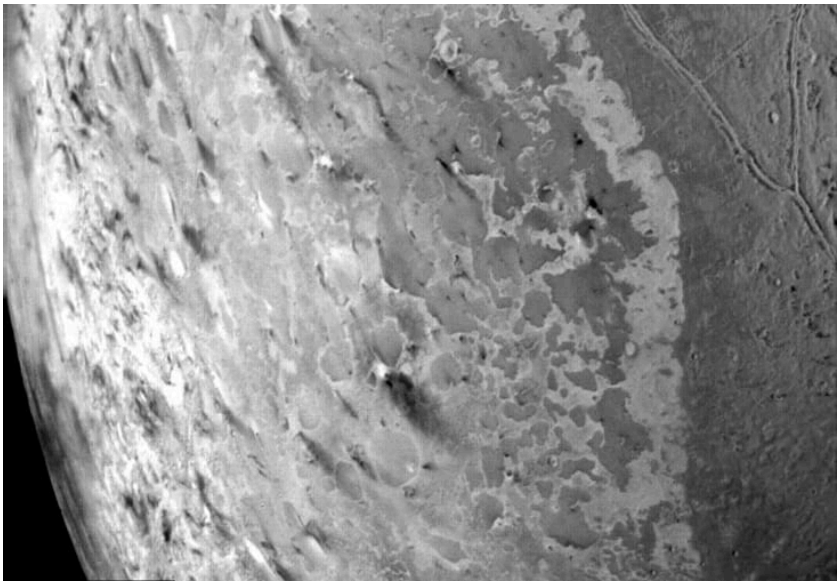


Figure 12.16 : Triton's Geysers. This close-up view shows some of the geysers on Neptune's moon Triton, with the long trains of dust pointing to the lower right in this picture. (credit: modification of work by NASA/JPL)

This page titled 17.3: Titan and Triton is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 12.4: Titan and Triton by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

18: ASTEROIDS AND COMETS

18.1: Asteroids
18.2: Asteroids and Planetary Defense
18.3: The "Long-Haired" Comets
18.4: Key Terms

This page titled 18: Asteroids and Comets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

18.1: Asteroids

Learning Objectives

By the end of this section, you will be able to:

- Outline the story of the discovery of asteroids and describe their typical orbits
- Describe the composition and classification of the various types of asteroids
- Discuss what was learned from spacecraft missions to several asteroids

The asteroids are mostly found in the broad space between Mars and Jupiter, a region of the solar system called the *asteroid belt*. Asteroids are too small to be seen without a telescope; the first of them was not discovered until the beginning of the nineteenth century.

Discovery and Orbits of the Asteroids

In the late 1700s, many astronomers were hunting for an additional planet they thought should exist in the gap between the orbits of Mars and Jupiter. The Sicilian astronomer Giovanni Piazzi thought he had found this missing planet in 1801, when he discovered the first asteroid (or as it was later called, "minor planet") orbiting at 2.8 AU from the Sun. His discovery, which he named Ceres, was quickly followed by the detection of three other little planets in similar orbits.

Clearly, there was not a single missing planet between Mars and Jupiter but rather a whole group of objects, each much smaller than our Moon. (An analogous discovery history has played out in slow motion in the outer solar system. Pluto was discovered beyond Neptune in 1930 and was initially called a planet, but early in the twenty-first century, several other similar objects were found. We now call all of them dwarf planets.)

By 1890, more than 300 of these minor planets or asteroids had been discovered by sharp-eyed observers. In that year, Max Wolf at Heidelberg introduced astronomical photography to the search for asteroids, greatly accelerating the discovery of these dim objects. In the twenty-first century, searchers use computer-driven electronic cameras, another leap in technology. Currently many asteroid searches are made with cameras in space. More than a million asteroids now have well-determined orbits.

Asteroids are given a number (corresponding to the order of discovery) and sometimes also a name. Originally, the names of asteroids were chosen from goddesses in Greek and Roman mythology. After exhausting these and other female names (including, later, those of spouses, friends, flowers, cities, and others), astronomers turned to the names of colleagues (and other people of distinction) whom they wished to honor. For example, asteroids 2410, 4859, and 68448 are named Morrison, Fraknoi, and Sidneywolff, for the three senior authors of this textbook.

The largest asteroid is Ceres (numbered 1), with a diameter just less than 1000 kilometers. As we saw, Ceres was considered a planet when it was discovered but later was called an asteroid (the first of many.) Now, it has again been reclassified and is considered one of the dwarf planets, like Pluto (see the chapter on Moons, Rings and Pluto). We still find it convenient, however, to discuss Ceres as the largest of the asteroids. Two other asteroids, Pallas and Vesta, have diameters of about 500 kilometers, and about 15 more are larger than 250 kilometers (see Table 13.1). The number of asteroids increases rapidly with decreasing size; there are about 100 times more objects 10 kilometers across than there are 100 kilometers across.

Link to Learning

The Minor Planet Center is a worldwide repository of data on asteroids. Visit it online to find out about the latest discoveries related to the small bodies in our solar system. (Note that some of the material on this site is technical; it's best to click on the menu tab for the "public" for information more at the level of this textbook.) Watch this video about asteroid size comparison for a dramatic visualization of the sizes of some better-known asteroids.

Table 13.1: The Largest Asteroids

#	Name	Year of Discovery	Orbit's Semimajor Axis (AU)	Diameter (km)	Compositional Class
1	Ceres	1801	2.77	940	C (carbonaceous)
2	Pallas	1802	2.77	540	C (carbonaceous)
3	Juno	1804	2.67	265	S (stony)
4	Vesta	1807	2.36	510	basaltic
10	Hygiea	1849	3.14	410	C (carbonaceous)
16	Psyche	1852	2.92	265	M (metallic)
31	Euphrosyne	1854	3.15	250	C (carbonaceous)
52	Europa	1858	3.10	280	C (carbonaceous)
65	Cybele	1861	3.43	280	C (carbonaceous)
87	Sylvia	1866	3.48	275	C (carbonaceous)
451	Patientia	1899	3.06	260	C (carbonaceous)
511	Davida	1903	3.16	310	C (carbonaceous)
704	Interamnia	1910	3.06	310	C (carbonaceous)

The asteroids all revolve about the Sun in the same direction as the planets, and most of their orbits lie near the plane in which Earth and other planets circle. The majority of asteroids are in the asteroid belt, the region between Mars and Jupiter that contains all asteroids with orbital periods between 3.3 to 6 years (Figure 13.2). Although more than 75% of the known asteroids are in the belt, they are not closely spaced (as they are sometimes depicted in science fiction movies). The volume of the belt is actually very large, and the typical spacing between objects (down to 1 kilometer in size) is several million kilometers. (This was fortunate for spacecraft like Galileo, Cassini, *Rosetta*, and New Horizons, which needed to travel through the asteroid belt without a collision.)

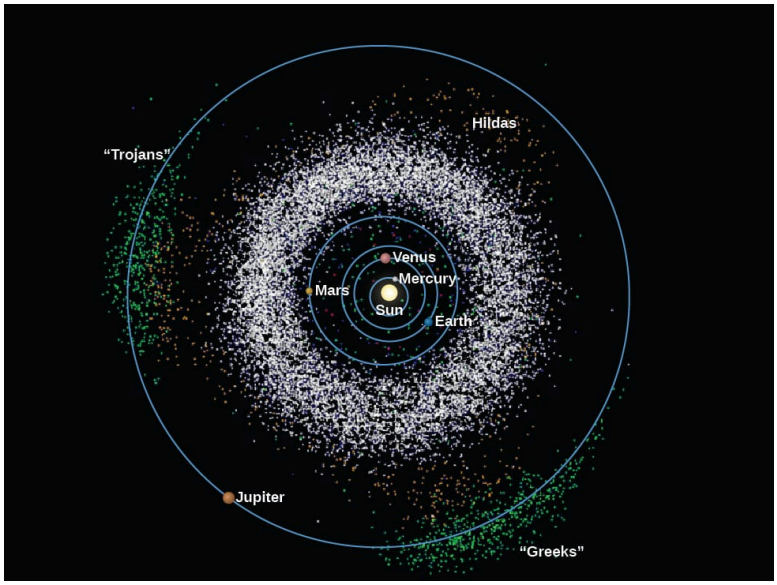


Figure 13.2 : Asteroids in the Solar System. This computer-generated diagram shows the positions of the asteroids known in 2006. If the asteroid sizes were drawn to scale, none of the dots representing an asteroid would be visible. Here, the asteroid dots are too big and give a false impression of how crowded the asteroid belt would look if you were in it. Note that in addition to those in the asteroid belt, there are also asteroids in the inner solar system and some along Jupiter's orbit (such as the Trojans and Greeks groups), controlled by the giant planet's gravity.

Still, over the long history of our solar system, there have been a good number of collisions among the asteroids themselves. In 1918, the Japanese astronomer Kiyotsugu Hirayama found that some asteroids fall into *families*, groups with similar orbital characteristics. He hypothesized that each family may have resulted from the breakup of a larger body or, more likely, from the

collision of two asteroids. Slight differences in the speeds with which the various fragments left the collision scene account for the small spread in orbits now observed for the different asteroids in a given family. Several dozen such families exist, and observations have shown that individual members of most families have similar compositions, as we would expect if they were fragments of a common parent.

Link to Learning

You can see a dramatic animated video showing the orbits of 100,000 asteroids found by one sky survey. As the 3-minute video goes on, you get to see the orbits of the planets and how the asteroids are distributed in the solar system. But note that all such videos are misleading in one sense. The asteroids themselves are really small compared to the distances covered, so they have to be depicted as larger points to be visible. If you were in the asteroid belt, there would be far more empty space than asteroids.

Composition and Classification

Asteroids are as different as black and white. The majority are very dark, with reflectivity of only 3 to 4%, like a lump of coal. However, another large group has a typical reflectivity of 15%. To understand more about these differences and how they are related to chemical composition, astronomers study the spectrum of the light reflected from asteroids for clues about their composition.

The dark asteroids are revealed from spectral studies to be *primitive* bodies (those that have changed little chemically since the beginning of the solar system) composed of silicates mixed with dark, organic carbon compounds. These are known as C-type asteroids (“C” for carbonaceous). Two of the largest asteroids, Ceres and Pallas, are primitive, as are almost all of the asteroids in the outer part of the belt. In 2020 and 2021, two of these C-type asteroids, Bennu and Ryugu, were visited by spacecraft that briefly touched their surfaces.

The second most populous group is the S-type asteroids, where “S” stands for a stony or silicate composition. Here, the dark carbon compounds are missing, resulting in higher reflectivity and clearer spectral signatures of silicate minerals. The S-type asteroids are also chemically primitive, but their different composition indicates that they were probably formed in a different location in the solar system from the C-type asteroids.

Asteroids of a third class, much less numerous than those of the first two, are composed primarily of metal and are called M-type asteroids (“M” for metallic). Spectroscopically, the identification of metal is difficult, but for at least the largest M-type asteroid, Psyche, this identification has been confirmed by radar. Since a metal asteroid, like an airplane or ship, is a much better reflector of radar than is a stony object, Psyche appears bright when we aim a radar beam at it. In 2022, a NASA mission, conveniently named Psyche, will be launched toward this asteroid.

How did such metal asteroids come to be? We suspect that each came from a parent body large enough for its molten interior to settle out or differentiate, and the heavier metals sank to the center. When this parent body shattered in a later collision, the fragments from the core were rich in metals. There is enough metal in even a 1-kilometer M-type asteroid to supply the world with iron and many other industrial metals for the foreseeable future, if we could bring one safely to Earth.

In addition to the M-type asteroids, a few other asteroids show signs of early heating and differentiation. These have basaltic surfaces like the volcanic plains of the Moon and Mars; the large asteroid Vesta (discussed in a moment) is in this last category.

The different classes of asteroids are found at different distances from the Sun (Figure 13.3). By tracing how asteroid compositions vary with distance from the Sun, we can reconstruct some of the properties of the solar nebula from which they originally formed.

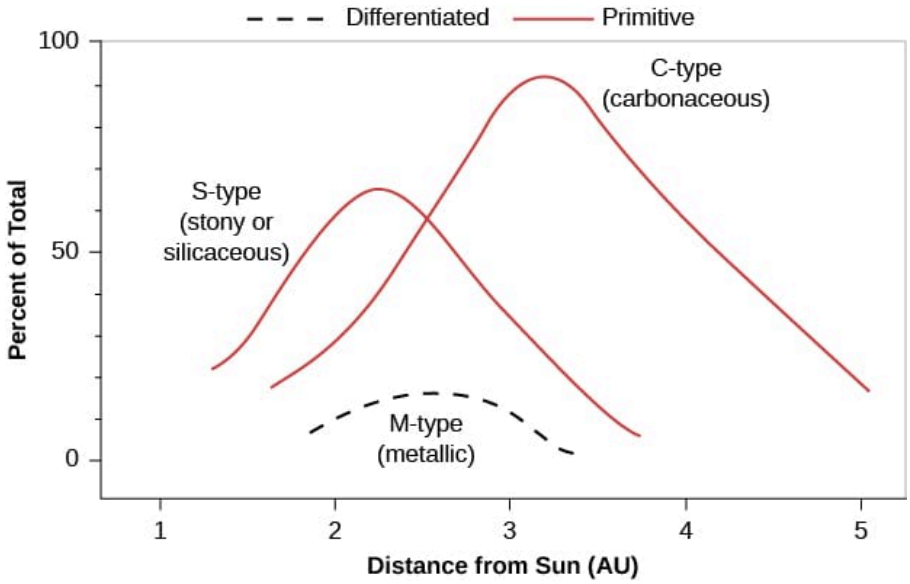


Figure 13.3 : Where Different Types of Asteroids Are Found. Asteroids of different composition are distributed at different distances from the Sun. The S-type and C-type are both primitive; the M-type consists of cores of differentiated parent bodies.

Vesta: A Differentiated Asteroid

Vesta is one of the most interesting of the asteroids. It orbits the Sun with a semi-major axis of 2.4 AU in the inner part of the asteroid belt. Its relatively high reflectivity of almost 30% makes it the brightest asteroid, so bright that it is actually visible to the unaided eye if you know just where to look. But its real claim to fame is that its surface is covered with basalt, indicating that Vesta is a differentiated object that must once have been volcanically active, in spite of its small size (about 500 kilometers in diameter).

Meteorites from Vesta's surface (Figure 13.4), identified by comparing their spectra with that of Vesta itself, have landed on Earth and are available for direct study in the laboratory. We thus know a great deal about this asteroid. The age of the lava flows from which these meteorites derived has been measured at 4.4 to 4.5 billion years, very soon after the formation of the solar system. This age is consistent with what we might expect for volcanoes on Vesta; whatever process heated such a small object was probably intense and short-lived. In 2016, a meteorite fell in Turkey that could be identified with a particular lava flow as revealed by the orbiting *Dawn* spacecraft.

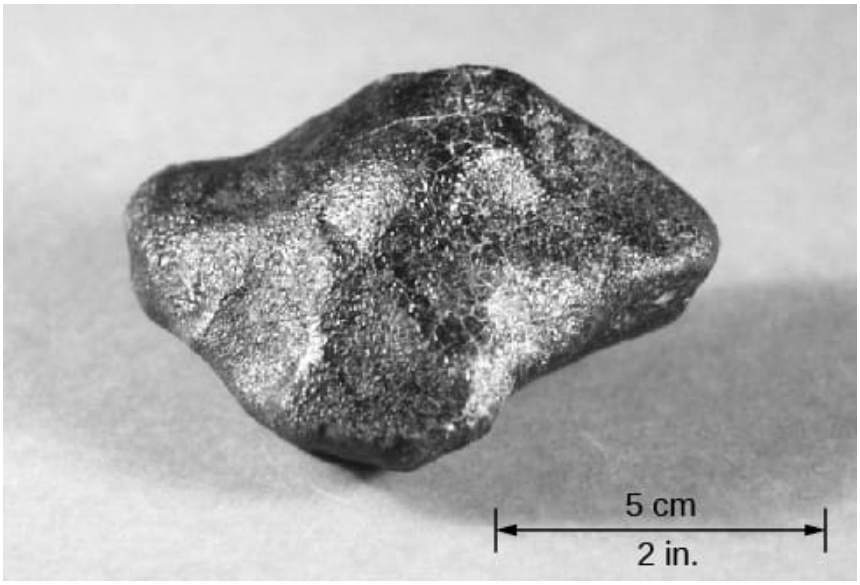


Figure 13.4 : Piece of Vesta. This meteorite (rock that fell from space) has been identified as a volcanic fragment from the crust of asteroid Vesta. (credit: modification of work by R. Kempton (New England Meteoritical Services))

Asteroids Up Close

On the way to its 1995 encounter with Jupiter, the Galileo spacecraft was targeted to fly close to two main-belt S-type asteroids called Gaspra and Ida. The Galileo camera revealed both as long and highly irregular (resembling a battered potato), as befits fragments from a catastrophic collision (Figure 13.5).

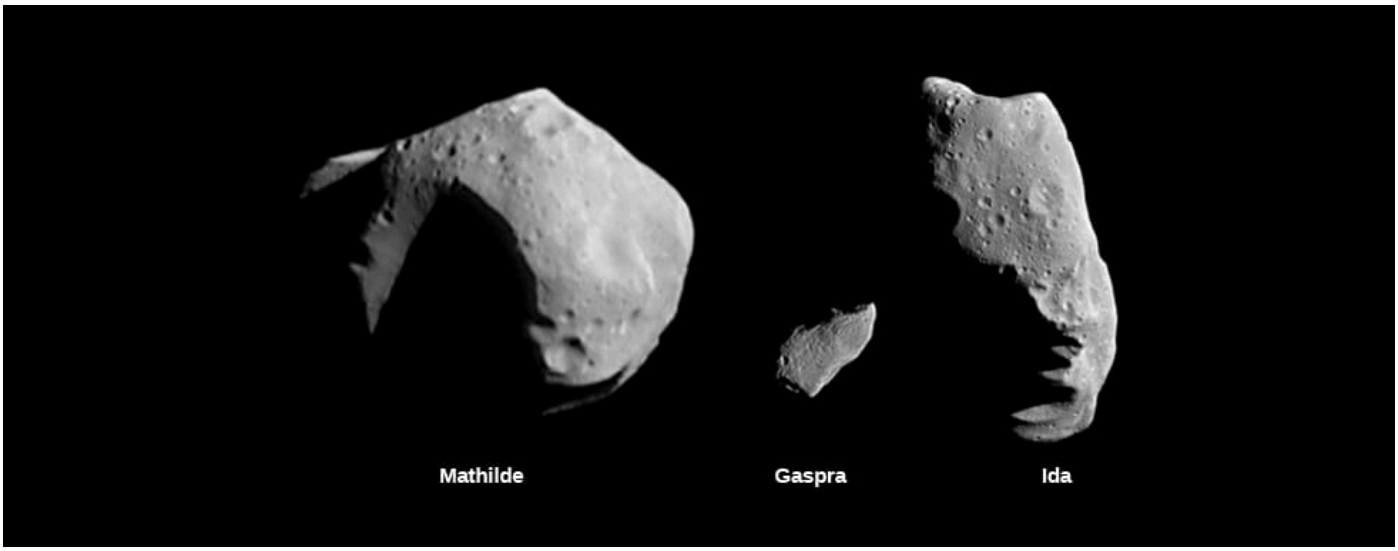


Figure 13.5 : Mathilde, Gaspra, and Ida. The first three asteroids photographed from spacecraft flybys, printed to the same scale. Gaspra and Ida are S-type and were investigated by the Galileo spacecraft; Mathilde is C-type and was a flyby target for the NEAR-Shoemaker spacecraft. (credit: modification of work by NEAR Project, Galileo Project, NASA)

The detailed images allowed us to count the craters on Gaspra and Ida, and to estimate the length of time their surfaces have been exposed to collisions. The Galileo scientists concluded that these asteroids are only about 200 million years old (that is, the collisions that formed them took place about 200 million years ago). Calculations suggest that an asteroid the size of Gaspra or Ida can expect another catastrophic collision sometime in the next billion years, at which time it will be disrupted to form another generation of still-smaller fragments.

The greatest surprise of the Galileo flyby of Ida was the discovery of a moon (which was then named Dactyl), in orbit about the asteroid (Figure 13.6). Although only 1.5 kilometers in diameter, smaller than many college campuses, Dactyl provides scientists with something otherwise beyond their reach—a measurement of the mass and density of Ida using Kepler's laws. The moon's distance of about 100 kilometers and its orbital period of about 24 hours indicate that Ida has a density of approximately 2.5 g/cm^3, which matches the density of primitive rocks. Subsequently, both large visible-light telescopes and high-powered planetary radar have discovered many other asteroid moons, so that we are now able to accumulate valuable data on asteroid masses and densities.



Figure 13.6 : Ida and Dactyl. The asteroid Ida and its tiny moon Dactyl (the small body off to its right), were photographed by the Galileo spacecraft in 1993. Irregularly shaped Ida is 56 kilometers in its longest dimension, while Dactyl is about 1.5 kilometers across. The colors have been intensified in this image; to the eye, all asteroids look basically gray. (credit: modification of work by NASA/JPL)

By the way, Phobos and Deimos, the two small moons of Mars, are probably captured asteroids (Figure 13.7). They were first studied at close range by the Viking orbiters in 1977 and later by *Mars Global Surveyor*. Both are irregular, somewhat elongated, and heavily created, resembling other smaller asteroids. Their largest dimensions are about 26 kilometers and 16 kilometers, respectively. The small outer moons of Jupiter and Saturn were probably also captured from passing asteroids, perhaps early in the history of the solar system.

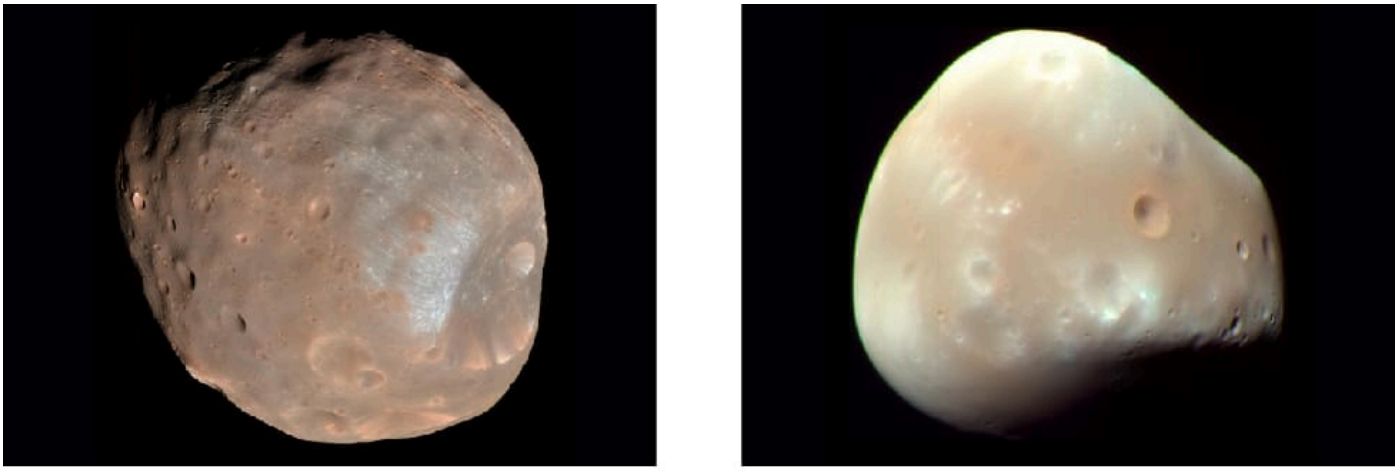


Figure 13.7 : Moons of Mars. The two small moons of Mars, (a) Phobos and (b) Deimos, were discovered in 1877 by American astronomer Asaph Hall. Their surface materials are similar to many of the asteroids in the outer asteroid belt, leading astronomers to believe that the two moons may be captured asteroids. (credit a: modification of work by NASA; credit b: modification of work by NASA/JPL-Caltech/University of Arizona)

Beginning in the 1990s, spacecraft have provided close looks at several more asteroids. The Near Earth Asteroid Rendezvous (NEAR) spacecraft went into orbit around the S-type asteroid Eros, becoming a temporary moon of this asteroid. On its way to Eros, the *NEAR* spacecraft was renamed after planetary geologist Eugene Shoemaker, a pioneer in our understanding of craters and impacts.

For a year, the NEAR-Shoemaker spacecraft orbited the little asteroid at various altitudes, measuring its surface and interior composition as well as mapping Eros from all sides (Figure 13.8). The data showed that Eros is made of some of the most chemically primitive materials in the solar system. Several other asteroids have been revealed as made of loosely bound rubble throughout, but not Eros. Its uniform density (about the same as that of Earth's crust) and extensive global-scale grooves and ridges show that it is a cracked but solid rock.

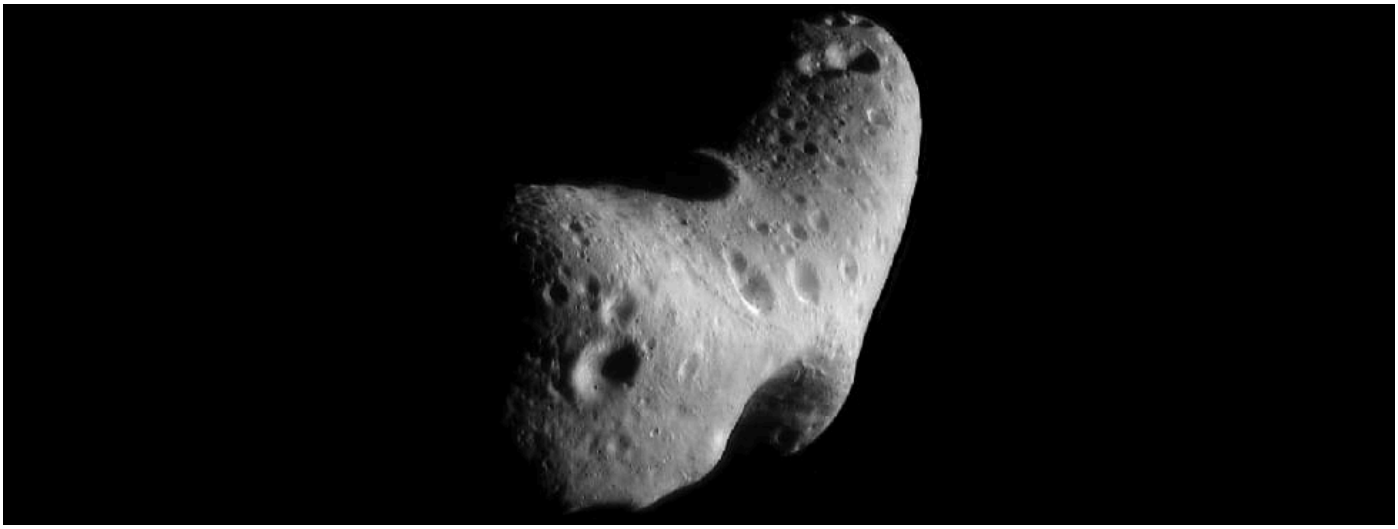


Figure 13.8 : Looking Down on the North Pole of Eros. This view was constructed from six images of the asteroid taken from an altitude of 200 kilometers. The large crater at the top has been named Psyche (after the maiden who was Eros' lover in classical mythology) and is about 5.3 kilometers wide. A saddle-shaped region can be seen directly below it. Craters of many different sizes are visible. (credit: modification of work by NASA/JHUPL)

Eros has a good deal of loose surface material that appears to have slid down toward lower elevations. In some places, the surface rubble layer is 100 meters deep. The top of loose soil is dotted with scattered, half-buried boulders. There are so many of these boulders that they are more numerous than the craters. Of course, with the gravity so low on this small world, a visiting astronaut would find loose boulders rolling toward her pretty slowly and could easily leap high enough to avoid being hit by one. Although the NEAR-Shoemaker spacecraft was not constructed as a lander, at the end of its orbital mission in 2000, it was allowed to fall gently to the surface, where it continued its chemical analysis for another week.

In 2003, Japan's Hayabusa 1 mission not only visited a small asteroid but also brought back samples to study in laboratories on Earth. The target S-type asteroid, Itokawa (shown in Figure 13.9), is much smaller than Eros, only about 500 meters long. This asteroid is elongated and appears to be the result of the collision of two separate asteroids long ago. There are almost no impact craters, but an abundance of boulders (like a pile of rubble) on the surface.

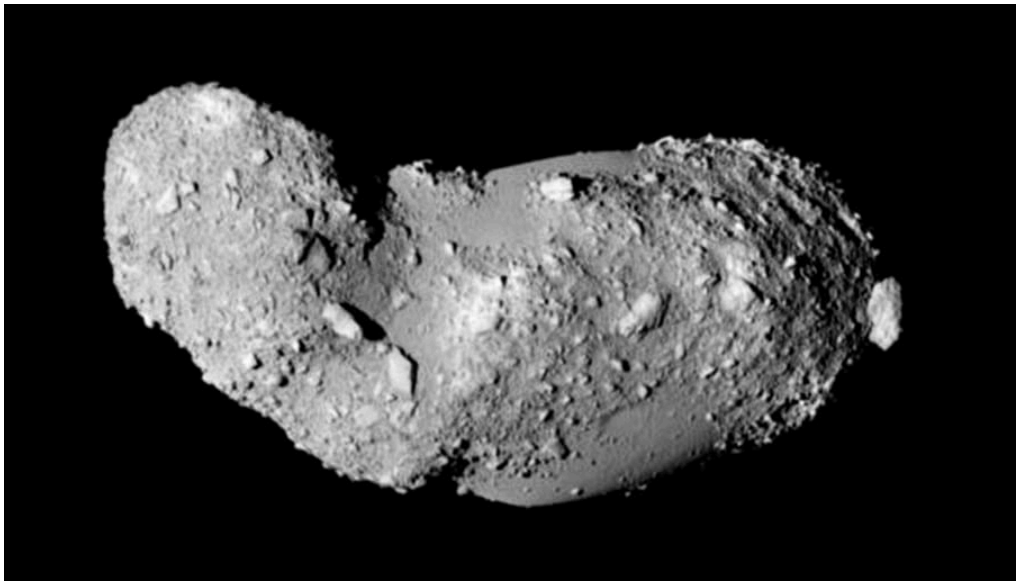


Figure 13.9 : Asteroid Itokawa. The surface of asteroid Itokawa appears to have no craters. Astronomers have hypothesized that its surface consists of rocks and ice chunks held together by a small amount of gravity, and its interior is probably also a similar rubble pile. (credit: JAXA)

The *Hayabusa* spacecraft was designed not to land, but to touch the surface just long enough to collect a small sample. This tricky maneuver failed on its first try, with the spacecraft briefly toppling over on its side. Eventually, the controllers were successful in picking up a few grains of surface material and transferring them into the return capsule (Figure 13.10). We estimate that the dust

grains *Hayabusa* picked up had been exposed on the surface of the asteroid for about 8 million years.

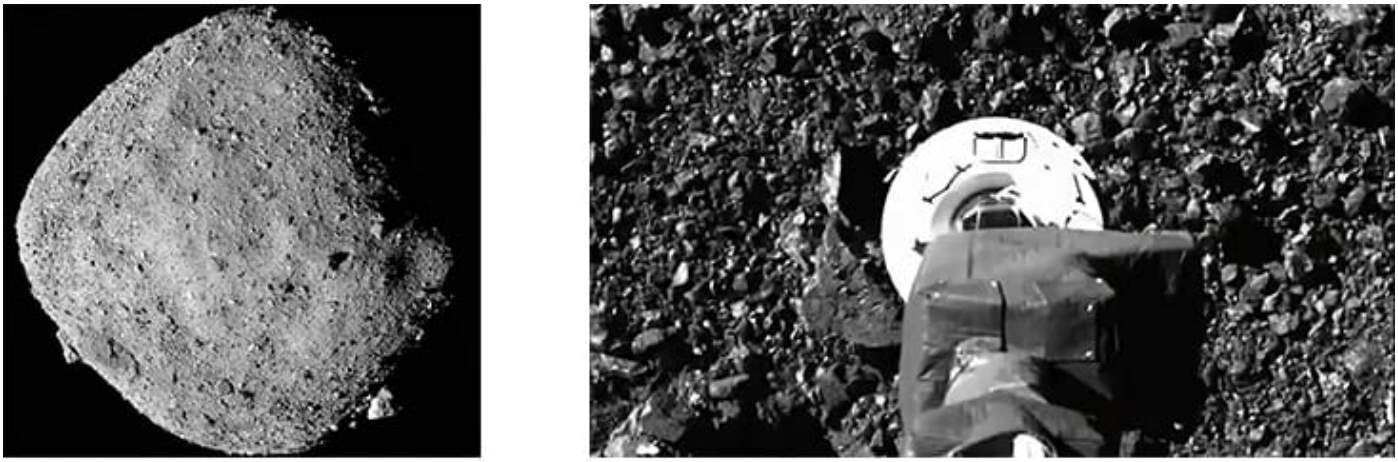


Figure 13.10 : Examining an Asteroid Up Close. (a) Near-Earth Asteroid Bennu, imaged by the Osiris-Rex Spacecraft in 2020. Bennu is a small (half-km diameter), rapidly spinning dark asteroid. Its low density suggests it is a loosely-bound rubble pile. The surface is very rough, with protruding rocks and boulders. (b) Preparing to collect a sample, the collection arm is shown just before touch-down on Bennu. For scale, the diameter of the sampling device is 30 cm. The sample of about 100 grams of surface rocks and dust will be returned to Earth in 2023. a: modification of work by NASA Goddard/University of Arizona; credit b: modification of work by NASA Goddard/University of Arizona

Following on the success of Hayabusa, both NASA and the Japanese space agency, JAXA, sent more advanced asteroid sample-return missions. Both targeted small Earth-approaching asteroids, only about 1 km in diameter. Each target is a dark object, apparently a rubble pile formed relatively recently after being broken apart in a collision. Hayabusa-2 visited the asteroid Ryugu, collected a small sample, and successfully returned it to Earth in December 2020. The U.S. mission called OSIRIS-REx visited Bennu, an asteroid that actually has a small chance of colliding with Earth in the next century. Its sample will be returned to Earth in September 2023.

The most ambitious asteroid space mission (called Dawn) has visited the two largest main belt asteroids, Ceres and Vesta, orbiting each for about a year (Figure 13.11). Their large sizes (diameters of about 1000 and 500 kilometers, respectively) make them appropriate for comparison with the planets and large moons. Both turned out to be heavily cratered, implying their surfaces are old. On Vesta, we have now actually located the large impact craters that ejected the basaltic meteorites previously identified as coming from this asteroid. These craters are so large that they sample several layers of Vesta's crustal material.

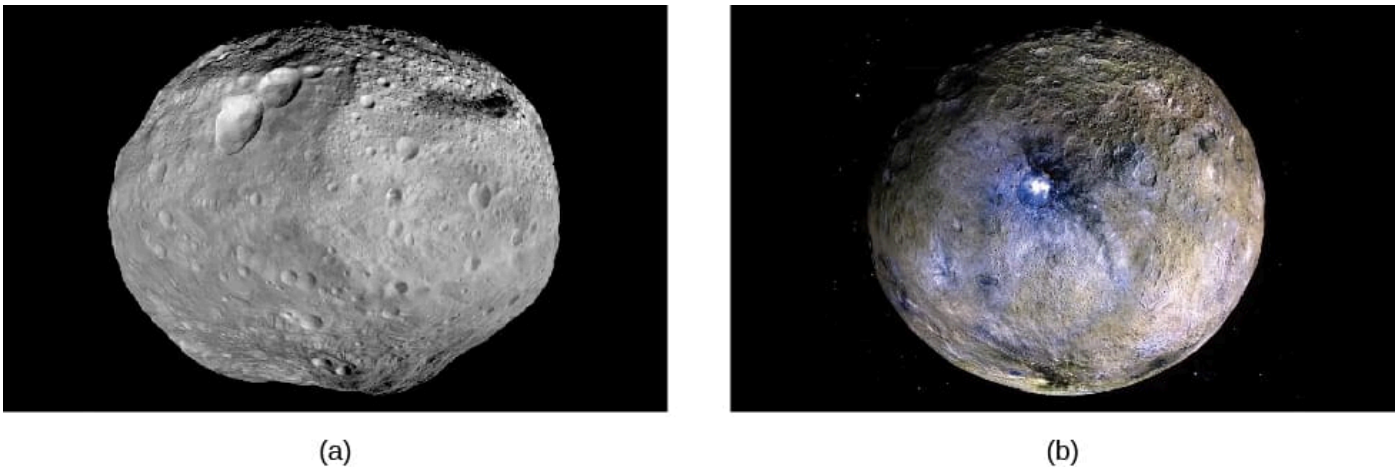


Figure 13.11 : Vesta and Ceres. The NASA *Dawn* spacecraft took these images of the large asteroids (a) Vesta and (b) Ceres. (a) Note that Vesta is not round, as Ceres (which is considered a dwarf planet) is. A mountain twice the height of Mt. Everest on Earth is visible at the very bottom of the Vesta image. (b) The image of Ceres has its colors exaggerated to bring out differences in composition. You can see a white feature in Occator crater near the center of the image. (credit a, b: modification of work by NASA/JPL-Caltech/UCLA/MPS/DLR/IDA)

Ceres has not had a comparable history of giant impacts, so its surface is covered with craters that look more like those from

the lunar highlands. One big surprise at Ceres is the presence of very bright white spots, associated primarily with the central peaks of large craters (Figure 13.12). The light-colored mineral is primarily salt, released from the interior. After repeated close flybys, data from the NASA Dawn spacecraft indicated that Ceres has (or has had) a subsurface ocean of water, with occasional eruptions on the surface. The most dramatic is the 4 kilometer tall ice volcano called Ahuna Mons (see Figure 13.12).

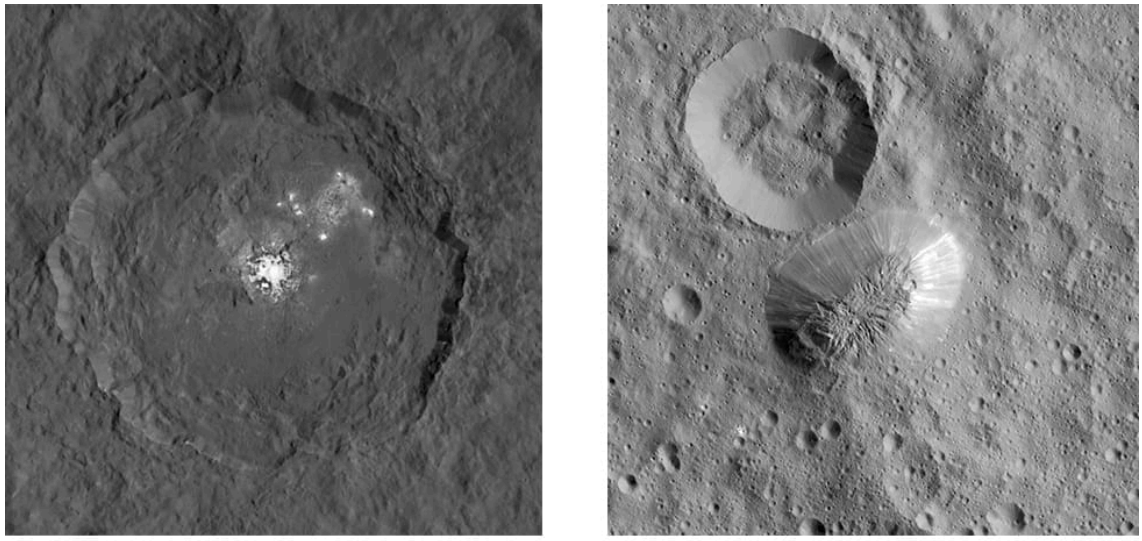


Figure 13.12 : White Spots in a Larger Crater on Ceres. White Spots in a Larger Crater on Ceres. (a) These bright features appear to be salt deposits in a Ceres crater called Occator, which is 92 kilometers across. (b) Ahuna Mons is an isolated mountain on Ceres, 4 kilometers high. It is thought to be an intrusion of ice from the interior. (credit a: modification of work by NASA/JPL-Caltech/UCLA/MPS/DLR/IDA; credit b: modification of work by NASA/JPL-Caltech/UCLA/MPS/DLR/IDA/PSI)

In late 2017, something entirely new was discovered: an interstellar asteroid. This visitor was found at a distance of 33 million kilometers with a survey telescope on Haleakala, Hawaii. As astronomers followed up on the discovery, it quickly became apparent that this asteroid was travelling far too fast to be part of the Sun's family. Its orbit is a hyperbola, and when discovered it was already rapidly leaving the inner solar system. Although it was too distant for imaging by even large telescopes, its size and shape could be estimated from its brightness and rapid light fluctuations. It is highly elongated, with an approximately cylindrical shape. The nominal dimensions are about 200 meters in length and only 35 meters across, the most extreme of any natural object. Large objects, like planets and moons, are pulled by their own gravity into roughly spherical shapes, and even small asteroids and comets (often described as "potato-shaped") rarely have irregularities of more than a factor of two.

This asteroid was named 'Oumuamua, a Hawaiian word meaning "scout" or "first to reach out." In a way, the discovery of an interstellar asteroid or comet was not unexpected. Early in solar system history, before the planet orbits sorted themselves into stable, non-intersecting paths all in the same plane, we estimate that quite a lot of mass was ejected, either whole planets or more numerous smaller fragments. Even today, an occasional comet coming in from the outer edges of the solar system can have its orbit changed by gravitational interaction with Jupiter and the Sun, and some of these escape on hyperbolic trajectories. In 2019, astronomers discovered an interstellar comet entering our solar system—more on this in the section on comets. As we have recently learned that planetary systems are common, the question became: where are similar debris objects ejected from other planetary systems? Now we have found two, and improved surveys will soon add others to this category.

Link to Learning

View an artist's rendering of the asteroid 'Oumuamua (www.openstax.org/l/30oumuamua) by the ESO. Although it was not close enough to Earth to be imaged, its long slender shape was indicated by its rapid variation in brightness as it rotated.

Link to Learning

The space agencies involved with the Dawn mission have produced nice animated "flyover" videos of Vesta and Ceres available online.

This page titled 18.1: Asteroids is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 13.2: Asteroids by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

18.2: Asteroids and Planetary Defense

Learning Objectives

By the end of this section, you will be able to:

- Recognize the threat that near-Earth objects represent for Earth
- Discuss possible defensive strategies to protect our planet

Not all asteroids are in the main asteroid belt. In this section, we consider some special groups of asteroids with orbits that approach or cross the orbit of Earth. These pose the risk of a catastrophic collision with our planet, such as the collision 65 million years ago that killed the dinosaurs.

Earth-Approaching Asteroids

Asteroids that stray far *outside* the main belt are of interest mostly to astronomers. But asteroids that come *inward*, especially those with orbits that come close to or cross the orbit of Earth, are of interest to political leaders, military planners—indeed, everyone alive on Earth. Some of these asteroids briefly become the closest celestial object to us.

In 1994, a 1-kilometer object was picked up passing closer than the Moon, causing a stir of interest in the news media. Today, it is routine to read of small asteroids coming this close to Earth. (They were always there, but only in recent years have astronomers been able to detect such faint objects.)

In 2013, a small asteroid hit our planet, streaking across the sky over the Russian city of Chelyabinsk and exploding with the energy of a nuclear bomb (Figure 13.13). The impactor was a stony object about 20 meters in diameter, exploding about 30 kilometers high with an energy of 500 kilotons (about 30 times larger than the nuclear bombs dropped on Japan in World War II). No one was hurt by the blast itself, although it briefly became as bright as the Sun, drawing many spectators to the windows in their offices and homes. When the blast wave from the explosion then reached the town, it blew out the windows. About 1500 people had to seek medical attention from injuries from the shattered glass.

A much larger atmospheric explosion took place in Russia in 1908, caused by an asteroid about 60 meters in diameter, releasing an energy of 10 megatons, as large the most powerful nuclear weapons of today. Fortunately, the area directly affected, on the Tunguska River in Siberia, was unpopulated, and no one was killed. However, the area of forest destroyed by the blast was large equal to the size of a major city (Figure 13.13).

Together with any comets that come close to our planet, such asteroids are known collectively as near-Earth objects (NEOs). As we will see (and as the dinosaurs found out 65 million years ago,) the collision of a significant-sized NEO could be a catastrophe for life on our planet.



(a)



(b)

Figure 13.13 : Impacts with Earth. (a) As the Chelyabinsk meteor passed through the atmosphere, it left a trail of smoke and briefly became as bright as the Sun. (b) Hundreds of kilometers of forest trees were knocked down and burned at the Tunguska impact site. (credit a: modification of work by Alex Alishevskikh)

Link to Learning

Visit the video compilation of the Chelyabinsk meteor streaking through the sky over the city on February 15, 2013, as taken by people who were in the area when it occurred.

Link to Learning

View this video of a non-technical talk by David Morrison to watch "The Chelyabinsk Meteor: Can We Survive a Bigger Impact?" Dr. Morrison (SETI Institute and NASA Ames Research Center) discusses the Chelyabinsk impact and how we learn about NEOs and protect ourselves; the talk is from the Silicon Valley Astronomy Lectures series.

Astronomers have urged that the first step in protecting Earth from future impacts by NEOs must be to learn what potential impactors are out there. In 1998, NASA began the Spaceguard Survey, with the goal to discover and track 90% of Earth-approaching asteroids greater than 1 kilometer in diameter. The size of 1 kilometer was selected to include all asteroids capable of causing global damage, not merely local or regional effects. At 1 kilometer or larger, the impact could blast so much dust into the atmosphere that the sunlight would be dimmed for months, causing global crop failures—an event that could threaten the survival of our civilization. The Spaceguard goal of 90% was reached in 2012 when nearly a thousand of these 1-kilometer near-Earth asteroids (NEAs) had been found, along with more than 10,000 smaller asteroids. Figure 13.14 shows how the pace of NEA discoveries has been increasing over recent years. By 2021, astronomers knew of more than 25,000.

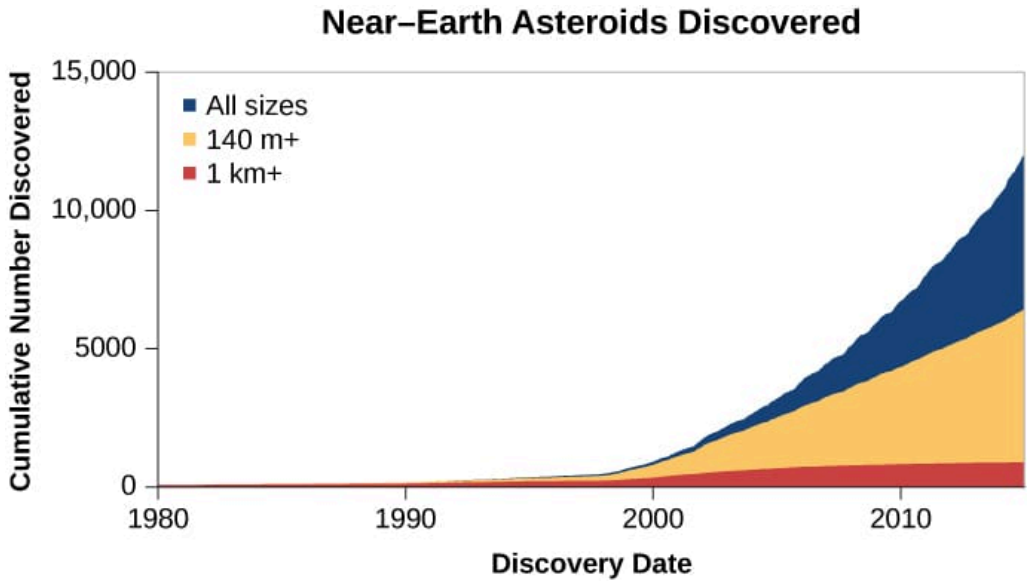


Figure 13.14 : Discovery of Near-Earth Asteroids. The accelerating rate of discovery of NEAs is illustrated in this graph, which shows the total number of known NEAs, the number over 140 meters in diameter, and the number over 1 kilometer in diameter, the size that poses the dominant impact risk on Earth.

How did astronomers know when they had discovered 90% of the larger asteroids? There are several ways to estimate the total number, even before they were individually located. One way is to look at the numbers of large craters on the dark lunar maria. Remember that these craters were made by impacts just like the ones we are considering. They are preserved on the Moon's airless surface, whereas Earth soon erases the imprints of past impacts. Thus, the number of large craters on the Moon allows us to estimate how often impacts have occurred on both the Moon and Earth over the past several billion years. The number of impacts is directly related to the number of asteroids and comets on Earth-crossing orbits.

Another approach is to see how often the surveys (which are automated searches for faint points of light that move among the stars) rediscover a previously known asteroid. At the beginning of a survey, all the NEAs it finds will be new. But as the survey becomes more complete, more and more of the moving points the survey cameras record will be *rediscoveries*. The more rediscoveries each survey experiences, the more complete our inventory of these asteroids must be.

We have been relieved to find that none of the NEAs discovered so far is on a trajectory that will impact Earth within the foreseeable future. However, we can't speak for the handful of asteroids larger than 1 kilometer that have not yet been found, or for the much more numerous smaller ones. It is estimated that there are a million NEAs capable of hitting Earth that are smaller than 1 kilometer but still large enough to destroy a city, and our surveys have found about 25% of them. Researchers who work with asteroid orbits estimate that for smaller (and therefore fainter) asteroids we are not yet tracking, we will have about a 5-second warning that one is going to hit Earth—in other words, we won't see it until it enters the atmosphere. Clearly, this estimate gives us a lot of motivation to continue these surveys to track as many asteroids as possible.

Though entirely predictable over times of a few centuries, the orbits of Earth-approaching asteroids are unstable over long time spans as they are tugged by the gravitational attractions of the planets. These objects will eventually meet one of two fates: either they will impact one of the terrestrial planets or the Sun, or they will be ejected gravitationally from the inner solar system due to a near-encounter with a planet. The probabilities of these two outcomes are about the same. The timescale for impact or ejection is only about a hundred million years, very short compared with the 4-billion-year age of the solar system. Calculations show that only approximately one quarter of the current Earth-approaching asteroids will eventually end up colliding with Earth itself.

If most of the current population of Earth-approaching asteroids will be removed by impact or ejection in a hundred million years, there must be a continuing source of new objects to replenish our supply of NEAs. Most of them come from the asteroid belt between Mars and Jupiter, where collisions between asteroids can eject fragments into Earth-crossing orbits (see Figure 13.15). Others may be "dead" comets that have exhausted their volatile materials (which we'll discuss in the next section).

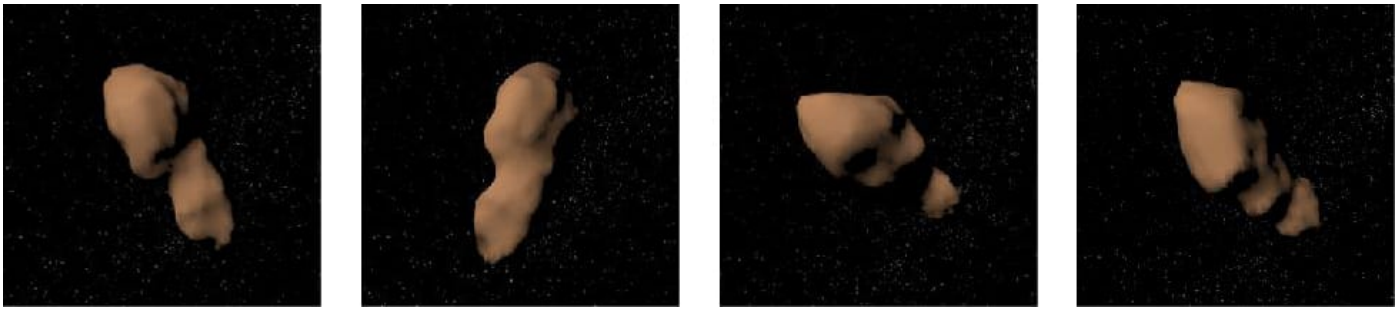


Figure 13.15 : Near-Earth Asteroid. Toutatis is a 5-kilometer long NEA that approached within 3 million kilometers of Earth in 1992. This series of images is a reconstruction its size and shape obtained from bouncing radar waves off the asteroid during its close flyby. Toutatis appears to consist of two irregular, lumpy bodies rotating in contact with each other. (Note that the color has been artificially added.) (credit: modification of work by NASA)

One reason scientists are interested in the composition and interior structure of NEAs is that humans will probably need to defend themselves against an asteroid impact someday. If we ever found one of these asteroids on a collision course with us, we would need to deflect it so it would miss Earth. The most straightforward way to deflect it would be to crash a spacecraft into it, either slowing it or speeding it up, slightly changing its orbital period. If this were done several years before the predicted collision, the asteroid would miss the planet entirely—making an asteroid impact the only natural hazard that we could eliminate completely by the application of technology. Alternatively, such deflection could be done by exploding a nuclear bomb near the asteroid to nudge it off course, although the use of nuclear explosives in space raises other concerns.

To achieve a successful deflection by either technique, we need to know more about the density and interior structure of the asteroid. A spacecraft impact or a nearby explosion would have a greater effect on a solid rocky asteroid such as Eros than on a loose rubble pile. Think of climbing a sand dune compared to climbing a rocky hill with the same slope. On the dune, much of our energy is absorbed in the slipping sand, so the climb is much more difficult and takes more energy.

The NASA Planetary Protection Office, which provides international guidance on these issues, has sponsored the first space mission to test asteroid deflection technology. The DART (Double Asteroid Redirection Test) was launched in November 2021 from the Vandenberg spaceport in California, now coincidentally operated by the military U.S. Space Force. Eleven months later, the 500-kg spacecraft will impact at a speed of about 6 km/s into Dimorphos, the 160 m diameter moon of asteroid Didymos. This will occur at a time in 2022 when the two objects are relatively close to Earth. The objective is to measure the momentum transferred to the target asteroid by using telescopes on Earth to track the change in the orbital period of Dimorphos. This mission will test both our technology to hit such a small object and the efficiency of the change in its orbit around Didymos. Note that the asteroid orbit is such that there is no possibility that the impact will lead to a later collision with Earth.

There is increasing international interest in the problem of asteroid impacts. The United Nations has formed two technical committees on planetary defense, recognizing that the entire planet is at risk from asteroid impacts. However, the fundamental problem remains one of finding NEAs in time for defensive measures to be taken. We must be able to find the next impactor before it finds us. And that's a job for the astronomers.

This page titled 18.2: Asteroids and Planetary Defense is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 13.3: Asteroids and Planetary Defense by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

18.3: The "Long-Haired" Comets

Learning Objectives

By the end of this section, you will be able to:

- Characterize the general physical appearance of comets
- Explain the range of cometary orbits
- Describe the size and composition of a typical comet's nucleus
- Discuss the atmospheres of comets
- Summarize the discoveries of the Rosetta mission

Comets differ from asteroids primarily in their icy composition, a difference that causes them to brighten dramatically as they approach the Sun, forming a temporary atmosphere. In some early cultures, these so-called "hairy stars" were considered omens of disaster. Today, we no longer fear comets, but eagerly anticipate those that come close enough to us to put on a good sky show.

Appearance of Comets

A comet is a relatively small chunk of icy material (typically a few kilometers across) that develops an atmosphere as it approaches the Sun. Later, there may be a very faint, nebulous tail, extending several million kilometers away from the main body of the comet. Comets have been observed from the earliest times: accounts of comets are found in the histories of virtually all ancient civilizations. The typical comet, however, is not spectacular in our skies, instead having the appearance of a rather faint, diffuse spot of light somewhat smaller than the Moon and many times less brilliant. (Comets seemed more spectacular to people before the invention of artificial lighting, which compromises our view of the night sky.)

Like the Moon and planets, comets appear to wander among the stars, slowly shifting their positions in the sky from night to night. Unlike the planets, however, most comets appear at unpredictable times, which perhaps explain why they frequently inspired fear and superstition in earlier times. Comets typically remain visible for periods that vary from a couple of weeks to several months. We'll say more about what they are made of and how they become visible after we discuss their motions.

Note that still images of comets give the impression that they are moving rapidly across the sky, like a bright meteor or shooting star. Looking only at such images, it is easy to confuse comets and meteors. But seen in the real sky, they are very different: the meteor burns up in our atmosphere and is gone in a few seconds, whereas the comet may be visible for weeks in nearly the same part of the sky.

Comet Orbits

The study of comets as members of the solar system dates from the time of Isaac Newton, who first suggested that they orbited the Sun on extremely elongated ellipses. Newton's colleague Edmund Halley (see the Edmund Halley: Astronomy's Renaissance Man feature box) developed these ideas, and in 1705, he published calculations of 24 comet orbits. In particular, he noted that the orbits of the bright comets that had appeared in the years 1531, 1607, and 1682 were so similar that the three could well be the same comet, returning to perihelion (closest approach to the Sun) at average intervals of 76 years. If so, he predicted that the object should next return about 1758. Although Halley had died by the time the comet appeared as he predicted, it was given the name Comet Halley (rhymes with "valley") in honor of the astronomer who first recognized it as a permanent member of our solar system, orbiting around the Sun. Its aphelion (furthest point from the Sun) is beyond the orbit of Neptune.

We now know from historical records that Comet Halley has actually been observed and recorded on every passage near the Sun since 239 BCE at intervals ranging from 74 to 79 years. The period of its return varies somewhat because of orbital changes produced by the pull of the giant planets. In 1910, Earth was brushed by the comet's tail, causing much needless public concern. Comet Halley last appeared in our skies in 1986 (Figure 13.16), when it was met by several spacecraft that gave us a wealth of information about its makeup; it will return in 2061.



Figure 13.16 : Comet Halley. This composite of three images (one in red, one in green, one in blue) shows Comet Halley as seen with a large telescope in Chile in 1986. During the time the three images were taken in sequence, the comet moved among the stars. The telescope was moved to keep the image of the comet steady, causing the stars to appear in triplicate (once in each color) in the background. (credit: modification of work by ESO)

Voyagers in Astronomy: Edmund Halley: Astronomy's Renaissance Man

Edmund Halley (Figure 13.17), a brilliant astronomer who made contributions in many fields of science and statistics, was by all accounts a generous, warm, and outgoing person. In this, he was quite the opposite of his good friend Isaac Newton, whose great work, the *Principia* (see Orbits and Gravity), Halley encouraged, edited, and helped pay to publish. Halley himself published his first scientific paper at age 20, while still in college. As a result, he was given a royal commission to go to Saint Helena (a remote island off the coast of Africa where Napoleon would later be exiled) to make the first telescopic survey of the southern sky. After returning, he received the equivalent of a master's degree and was elected to the prestigious Royal Society in England, all at the age of 22.

In addition to his work on comets, Halley was the first astronomer to recognize that the so-called "fixed" stars move relative to each other, by noting that several bright stars had changed their positions since Ptolemy's publication of the ancient Greek catalogs. He wrote a paper on the possibility of an infinite universe, proposed that some stars may be variable, and discussed the nature and size of *nebulae* (glowing cloudlike structures visible in telescopes). While in Saint Helena, Halley observed the planet Mercury going across the face of the Sun and developed the mathematics of how such transits could be used to establish the size of the solar system.

In other fields, Halley published the first table of human life expectancies (the precursor of life-insurance statistics); wrote papers on monsoons, trade winds, and tides (charting the tides in the English Channel for the first time); laid the foundations for the systematic study of Earth's magnetic field; studied evaporation and how inland waters become salty; and even designed an underwater diving bell. He served as a British diplomat, advising the emperor of Austria and squiring the future czar of Russia around England (avidly discussing, we are told, both the importance of science and the quality of local brandy).

In 1703, Halley became a professor of geometry at Oxford, and in 1720, he was appointed Astronomer Royal of England. He continued observing Earth and the sky and publishing his ideas for another 20 years, until death claimed him at age 85.

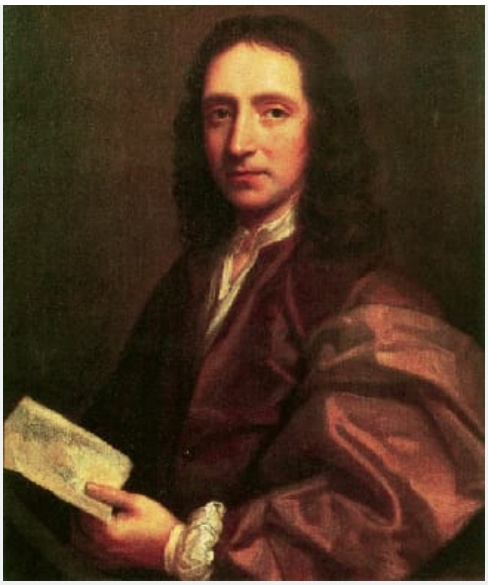


Figure 13.17 : Edmund Halley (1656–1742). Halley was a prolific contributor to the sciences. His study of comets at the turn of the eighteenth century helped predict the orbit of the comet that now bears his name.

Only a few comets return in a time measureable in human terms (shorter than a century or two), like Comet Halley does; these are called *short-period* comets. Many short-period comets have had their orbits changed by coming too close to one of the giant planets—most often Jupiter (and they are thus sometimes called Jupiter-family comets). Most comets have long periods and will take thousands of years to return, if they return at all. As we will see later in this chapter, most Jupiter-family comets come from a different source than the *long-period* comets (those with orbital periods longer than about a century).

Observational records exist for thousands of comets. We were visited by two bright comets in recent decades. First, in March 1996, came Comet Hyakutake, with a very long tail. A year later, Comet Hale-Bopp appeared; it was as bright as the brightest stars and remained visible for several weeks, even in urban areas (see the image that opens this chapter, Figure 13.1) Since then, there have been few comets visible to the naked eye, and astronomers (both professional and amateur) eagerly await such a visitor.

Table 13.2 lists some well-known comets whose history or appearance is of special interest.

Table 13.2: Some Interesting Comets

Name	Period	Significance
Great Comet of 1577	Long	Tycho Brahe showed it was beyond the Moon (a big step in our understanding)
Great Comet of 1843	Long	Brightest recorded comet; visible in daytime
Daylight Comet of 1910	Long	Brightest comet of the twentieth century
West	Long	Nucleus broke into pieces (1976)
Hyakutake	Long	Passed within 15 million km of Earth (1996)
Hale–Bopp	Long	Brightest recent comet (1997)
Swift-Tuttle	133 years	Parent comet of Perseid meteor shower
Halley	76 years	First comet found to be periodic; explored by spacecraft in 1986
Borrelly	6.8 years	Flyby by Deep Space 1 spacecraft (2000)
Biela	6.7 years	Broke up in 1846 and not seen again
Churyumov-Gerasimenko	6.5 years	Target of Rosetta mission (2014–16)
Wild 2	6.4 years	Target of Stardust sample return mission (2004)
Tempel 1	5.7 years	Target of Deep Impact mission (2005)

Name	Period	Significance
Encke	3.3 years	Shortest known period

The Comet's Nucleus

When we look at an active comet, all we normally see is its temporary atmosphere of gas and dust illuminated by sunlight. This atmosphere is called the comet's head or *coma*. Since the gravity of such small bodies is very weak, the atmosphere is rapidly escaping all the time; it must be replenished by new material, which has to come from somewhere. The source is the small, solid nucleus inside, just a few kilometers across, usually hidden by the glow from the much-larger atmosphere surrounding it. The nucleus is the real comet, the fragment of ancient icy material responsible for the atmosphere and the tail (Figure 13.18).

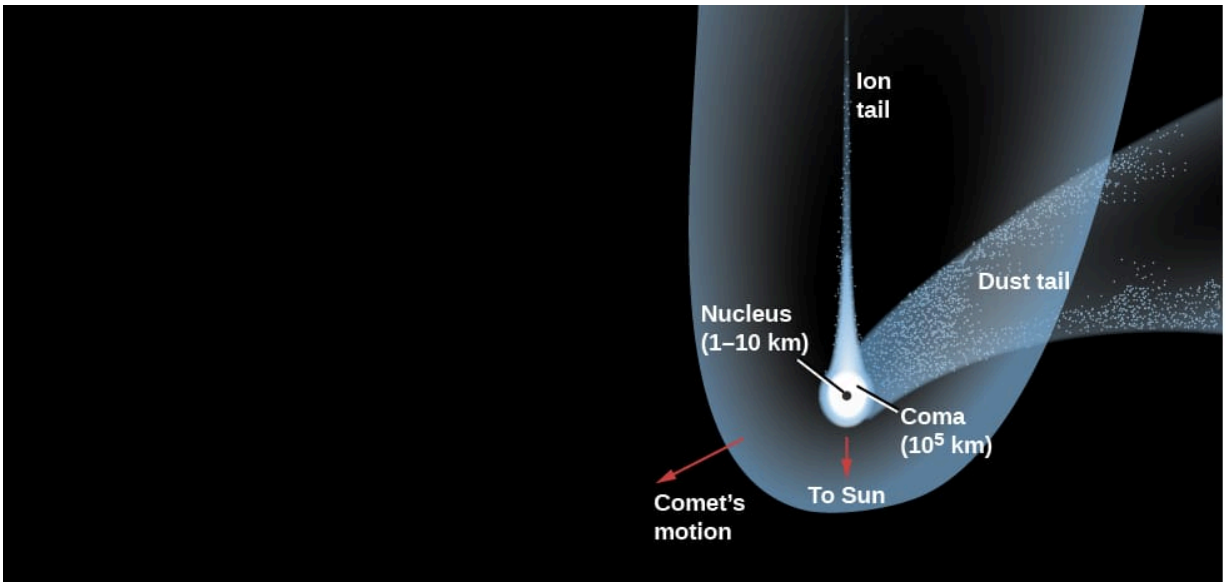


Figure 13.18 : Parts of a Comet. This schematic illustration shows the main parts of a comet. Note that the different structures are not to scale.

The modern theory of the physical and chemical nature of comets was first proposed by Harvard astronomer Fred Whipple in 1950. Before Whipple's work, many astronomers thought that a comet's nucleus might be a loose aggregation of solids, sort of an orbiting "gravel bank," Whipple proposed instead that the nucleus is a solid object a few kilometers across, composed in substantial part of water ice (but with other ices as well) mixed with silicate grains and dust. This proposal became known as the "dirty snowball" model.

The water vapor and other volatiles that escape from the nucleus when it is heated can be detected in the comet's head and tail, and therefore, we can use spectra to analyze what atoms and molecules the nucleus ice consists of. However, we are somewhat less certain of the non-icy component. We have never identified a fragment of solid matter from a comet that has survived passage through Earth's atmosphere. However, spacecraft that have approached comets have carried dust detectors, and some comet dust has even been returned to Earth (see Figure 13.19). It seems that much of the "dirt" in the dirty snowball is dark, primitive hydrocarbons and silicates, rather like the material thought to be present on the dark, primitive asteroids.

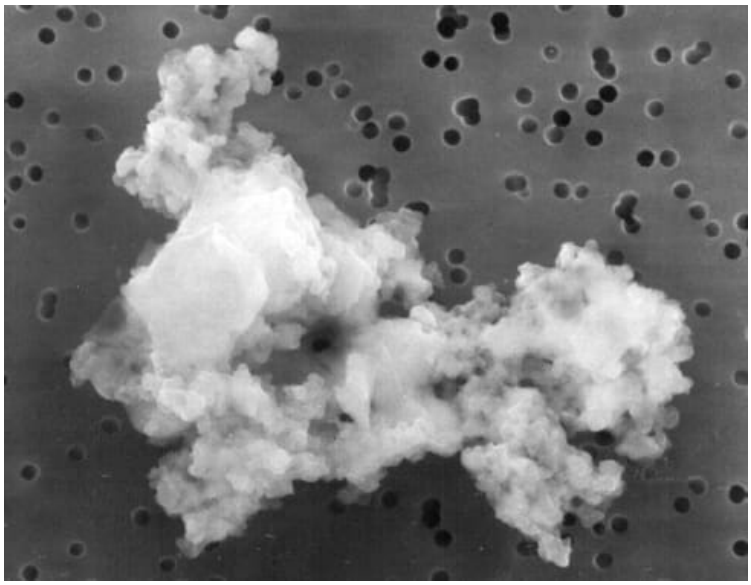


Figure 13.19 : Captured Comet Dust. This particle (seen through a microscope) is believed to be a tiny fragment of cometary dust, collected in the upper atmosphere of Earth. It measures about 10 microns, or 1/100 of a millimeter, across. (credit: NASA/JPL)

Since the nuclei of comets are small and dark, they are difficult to study from Earth. Spacecraft did obtain direct measurements of a comet nucleus, however, in 1986, when three spacecraft swept past Comet Halley at close range (see Figure 13.20). Subsequently, other spacecraft have flown close to other comets. In 2005, the NASA *Deep Impact* spacecraft even carried a probe for a high-speed impact with the nucleus of Comet Tempel 1. But by far, the most productive study of a comet has been by the 2015 Rosetta mission, which we will discuss shortly.



Figure 13.20 : Close-up of Comet Halley. This historic photograph of the black, irregularly shaped nucleus of Comet Halley was obtained by the ESA *Giotto* spacecraft from a distance of about 1000 kilometers. The bright areas are jets of material escaping from the surface. The length of the nucleus is 10 kilometers, and details as small as 1 kilometer can be made out. (credit: modification of work by ESA)

The Comet's Atmosphere

The spectacular activity that allows us to see comets is caused by the evaporation of cometary ices heated by sunlight. Beyond the asteroid belt, where comets spend most of their time, these ices are solidly frozen. But as a comet approaches the Sun, it begins to warm up. If water (H_2O) is the dominant ice, significant quantities vaporize as sunlight heats the surface above 200 K. This happens for the typical comet somewhat beyond the orbit of Mars. The evaporating H_2O in turn releases the dust that was mixed with the ice. Since the comet's nucleus is so small, its gravity cannot hold back either the gas or the dust, both of which flow away into space at speeds of about 1 kilometer per second.

The comet continues to absorb energy as it approaches the Sun. A great deal of this energy goes into the evaporation of its ice, as well as into heating the surface. However, recent observations of many comets indicate that the evaporation is not uniform and that most of the gas is released in sudden spurts, perhaps confined to a few areas of the surface. Expanding into space at a speed of about 1 kilometer per second, the comet's atmosphere can reach an enormous size. The diameter of a comet's head is often as large as Jupiter, and it can sometimes approach a diameter of a million kilometers (Figure 13.21).



Figure 13.21 : Head of Comet Halley. Here we see the cloud of gas and dust that make up the head, or coma, of Comet Halley in 1986. On this scale, the nucleus (hidden inside the cloud) would be a dot too small to see. (credit: modification of work by NASA/W. Liller)

Most comets also develop tails as they approach the Sun. A comet's tail is an extension of its atmosphere, consisting of the same gas and dust that make up its head. As early as the sixteenth century, observers realized that comet tails always point away from the Sun (Figure 13.22), not back along the comet's orbit. Newton proposed that comet tails are formed by a repulsive force of sunlight driving particles away from the head—an idea close to our modern view.

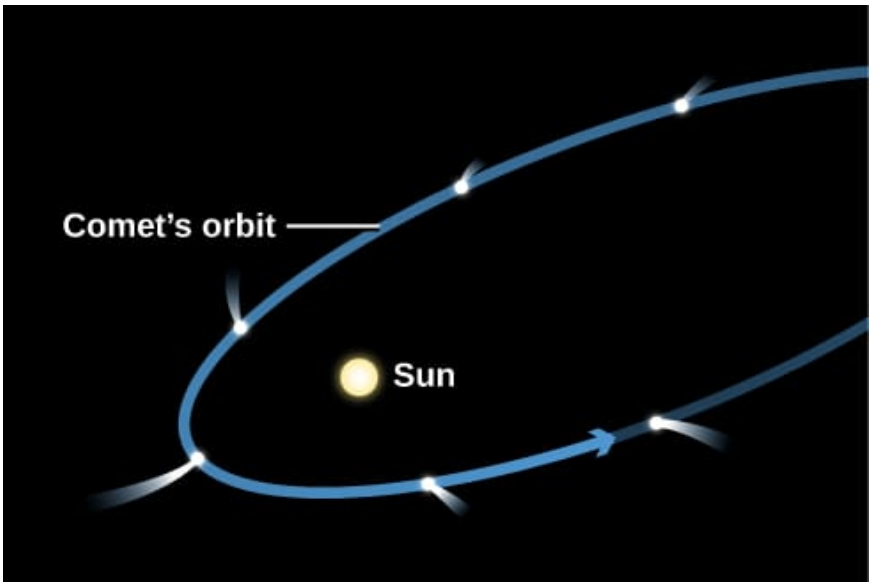


Figure 13.22 : Comet Orbit and Tail. The orientation of a typical comet tail changes as the comet passes perihelion. Approaching the Sun, the tail is behind the incoming comet head, but on the way out, the tail precedes the head.

The two different components that make up the tail (the dust and gas) act somewhat differently. The brightest part of the tail is called the *dust tail*, to differentiate it from a fainter, straight tail made of ionized gas, called the ion tail. The ion tail is carried outward by streams of ions (charged particles) emitted by the Sun. As you can see in Figure 13.23, the smoother dust tail curves a bit, as individual dust particles spread out along the comet's orbit, whereas the straight ion tail is pushed more directly outward from the Sun by our star's wind of charged particles

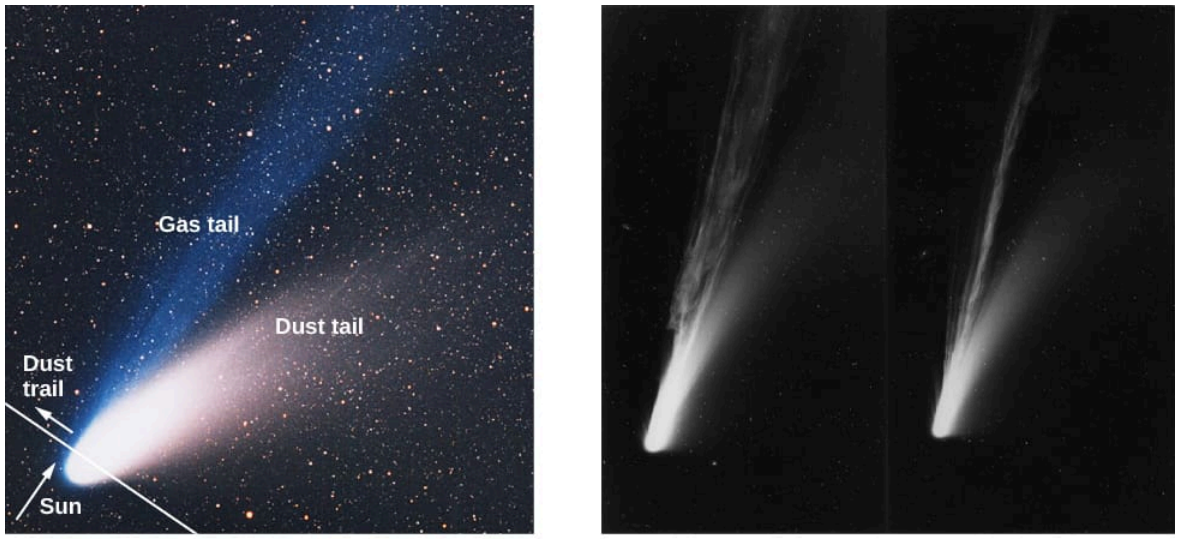


Figure 13.23 : Comet Tails. (a) As a comet nears the Sun, its features become more visible. In this illustration from NASA showing Comet Hale-Bopp, you can see a comet's two tails: the more easily visible dust tail, which can be up to 10 million kilometers long, and the fainter gas tail (or ion tail), which is up to hundreds of millions of kilometers long. The grains that make up the dust tail are the size of smoke particles. (b) Comet Mrkos was photographed in 1957 with a wide-field telescope at Palomar Observatory and also shows a clear distinction between the straight gas tail and the curving dust tail. (credit a: modification of work by ESO/E. Slawik; credit b: modification of work by Charles Kearns, George O. Abell, and Byron Hill)

Link to Learning

These days, comets close to the Sun can be found with spacecraft designed to observe our star. For example, in early July, 2011, astronomers at the ESA/NASA's Solar and Heliospheric Observatory (SOHO) witnessed a comet streaking toward the Sun, one of almost 3000 such sightings. You can also watch a brief video by NASA entitled "Why Are We Seeing So Many Sungrazing Comets?"

The Rosetta Comet Mission

The 1976 Giotto flyby of Comet Halley, which gave us the first picture of an active comet nucleus, was a watershed in our understanding of comets. In the 1990s, European scientists decided to design a much more ambitious mission that would match orbits with an incoming comet and follow it as it approached the Sun. They also proposed that a smaller spacecraft would actually try to land on the comet. The 2-ton main spacecraft was named *Rosetta*, carrying a dozen scientific instruments, and its 100-kilogram lander with nine more instruments was named *Philae*.

The Rosetta mission was launched in 2004. Delays with the launch rocket caused it to miss its original target comet, so an alternate destination was picked, Comet Churyumov-Gerasimenko (named after the two discoverers, but generally denoted 67P). This comet's period of revolution is 6.45 years, making it a Jupiter-family comet.

Since the European Space Agency did not have access to the plutonium-fueled nuclear power sources used by NASA for deep space missions, *Rosetta* had to be solar powered, requiring especially large solar panels. Even these were not enough to keep the craft operating as it matched orbits with 67P near the comet's aphelion. The only solution was to turn off all the spacecraft systems and let it coast for several years toward the Sun, out of contact with controllers on Earth until solar energy was stronger. The success of the mission depended on an automatic timer to turn the power back on as it neared the Sun. Fortunately, this strategy worked.

In August 2014, *Rosetta* began a gradual approach to the comet nucleus, which is a strangely misshapen object about 5 kilometers across, quite different from the smooth appearance of Halley's nucleus (but equally dark). Its rotation period is 12 hours. On November 12, 2014, the *Philae* lander was dropped, descending slowly for 7 hours before gently hitting the surface. It bounced and rolled, coming to rest under an overhang where there was not enough sunlight to keep its batteries charged. After operating for a few hours and sending data back to the orbiter, *Philae* went silent. The main *Rosetta* spacecraft continued

operations, however, as the level of comet activity increased, with steamers of gas jetting from the surface. As the comet approached perihelion in September 2015, the spacecraft backed off to ensure its safety.

The extent of the *Rosetta* images (and data from other instruments) far exceeds anything astronomers had seen before from a comet. The best imaging resolution was nearly a factor of 100 greater than in the best Halley images. At this scale, the comet appears surprisingly rough, with sharp angles, deep pits, and overhangs (Figure 13.24).

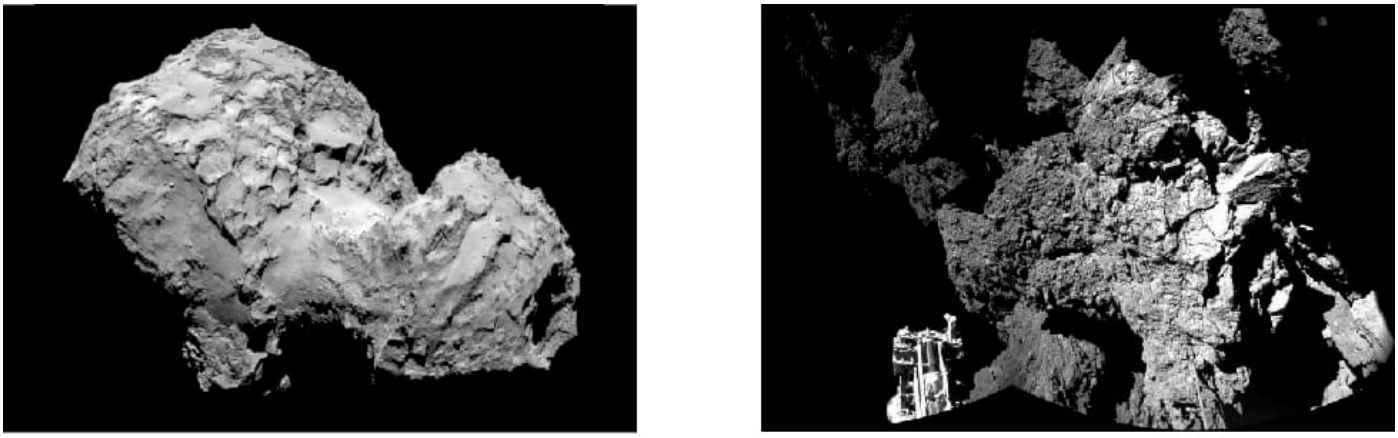


Figure 13.24 : Comet 67P's Strange Shape and Surface Features. (a) This image from the *Rosetta* camera was taken from a distance of 285 kilometers. The resolution is 5 meters. You can see that the comet consists of two sections with a connecting "neck" between them. (b) This close-up view of Comet Churyumov-Gerasimenko is from the *Philae* lander. One of the lander's three feet is visible in the foreground. The lander itself is mostly in shadow. (credit a: modification of work by ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA; credit b: modification of work by ESA/Rosetta/Philae/CIVA)

The double-lobed shape of 67P's nucleus has been tentatively attributed to the collision and merger of two independent comet nuclei long ago. The spacecraft verified that the comet's dark surface was covered with organic carbon-rich compounds, mixed with sulfides and iron-nickel grains. 67P has an average density of only 0.5 g/cm^3 (recall water in these units has a density of 1 g/cm^3.) This low density indicates that the comet is quite porous, that is, there is a large amount of empty space among its materials.

We already knew that the evaporation of comet ices was sporadic and limited to small jets, but in comet 67P, this was carried to an extreme. At any one time, more than 99% of the surface is inactive. The active vents are only a few meters across, with the material confined to narrow jets that persist for just a few minutes (Figure 13.25). The level of activity is strongly dependent on solar heating, and between July and August 2015, it increased by a factor of 10. Isotopic analysis of deuterium in the water ejected by the comet shows that it is different from the water found on Earth. Thus, apparently comets like 67P did not contribute to the origin of our oceans or the water in our bodies, as some scientists had thought.



(a) (b) (c)

Figure 13.25 : Gas Jets on Comet 67P. "(a and b) Jet activity was photographed by the *Rosetta* spacecraft near perihelion. You can see a jet suddenly appearing; it was active for only a few minutes. (c) This spectacular photo, taken near perihelion, shows the active comet surrounded by multiple jets of gas and dust. (credit a, b: modification of work by ESA/Rosetta/MPS; credit c: modification of work by ESA/Rosetta/NAVCAM)

Link to Learning

The European Space Agency is continuing to make interesting short videos illustrating the challenges and results of the Rosetta and Philae missions. For example, watch "*Rosetta*'s Moment in the Sun" to see some of the images of the comet generating plumes of gas and dust and hear about some of the dangers an active comet poses for the spacecraft.

This page titled 18.3: The "Long-Haired" Comets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 13.4: The "Long-Haired" Comets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

18.4: Key Terms

asteroid
: a stony or metallic object orbiting the Sun that is smaller than a planet but that shows no evidence of an atmosphere or of other types of activity associated with comets

asteroid belt
: the region of the solar system between the orbits of Mars and Jupiter in which most asteroids are located; the main belt, where the orbits are generally the most stable, extends from 2.2 to 3.3 AU from the Sun

comet
: a small body of icy and dusty matter that revolves about the Sun; when a comet comes near the Sun, some of its material vaporizes, forming a large head of tenuous gas and often a tail

Kuiper belt
: a region of space beyond Neptune that is dynamically stable (like the asteroid belt); the source region for most short-period comets

near-Earth asteroid (NEA)
: an Earth-approaching asteroid, one whose orbit could bring it on a collision course with our planet

near-Earth object (NEO)
: a comet or asteroid whose path intersects the orbit of Earth

nucleus (of a comet)
: the solid chunk of ice and dust in the head of a comet

Oort cloud
: the large spherical region around the Sun from which most “new” comets come; a reservoir of objects with aphelia at about 50,000 AU

tail
: (of a comet) a tail consisting of two parts: the dust tail is made of dust loosened by the sublimation of ice in a comet that is then pushed by photons from the Sun into a curved stream; the ion tail is a stream of ionized particles evaporated from a comet and then swept away from the Sun by the solar wind

This page titled 18.4: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

19: THE SUN

19.1: The Structure and Composition of the Sun
19.2: The Solar Cycle
19.3: Solar Activity above the Photosphere
19.4: Mass, Energy, and the Theory of Relativity
19.5: The Solar Interior - Theory
19.6: The Solar Interior - Observations
19.7: Key Terms

This page titled 19: The Sun is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

19.1: The Structure and Composition of the Sun

Learning Objectives

By the end of this section, you will be able to:

- Explain how the composition of the Sun differs from that of Earth
- Describe the various layers of the Sun and their functions
- Explain what happens in the different parts of the Sun's atmosphere

The Sun, like all stars, is an enormous ball of extremely hot, largely ionized gas, shining under its own power. And we do mean enormous. The Sun could fit 109 Earths side-by-side across its diameter, and it has enough volume (takes up enough space) to hold about 1.3 million Earths.

The Sun does not have a solid surface or continents like Earth, nor does it have a solid core (Figure 15.2). However, it does have a lot of structure and can be discussed as a series of layers, not unlike an onion. In this section, we describe the huge changes that occur in the Sun's extensive interior and atmosphere, and the dynamic and violent eruptions that occur daily in its outer layers.



Figure 15.2 : Earth and the Sun. Here, Earth is shown to scale with part of the Sun and a giant loop of hot gas erupting from its surface. The inset shows the entire Sun, smaller. (credit: modification of work by SOHO/EIT/ESA)

Some of the basic characteristics of the Sun are listed in Table 15.1. Although some of the terms in that table may be unfamiliar to you right now, you will get to know them as you read further.

Table 15.1: Characteristics of the Sun

Characteristic	How Found	Value
Mean distance	Radar reflection from planets	1 AU (149,597,892 km)
Maximum distance from Earth		1.521×10^8 km
Minimum distance from Earth		1.471×10^8 km
Mass	Orbit of Earth	333,400 Earth masses (1.99×10^{30} kg)
Mean angular diameter	Direct measure	31′59′′.3

Characteristic	How Found	Value
Diameter of photosphere	Angular size and distance	109.3 × Earth diameter (1.39×10^6 km)
Mean density	Mass/volume	1.41 g/cm^3 (1400 kg/m^3)
Gravitational acceleration at photosphere (surface gravity)	GM/R^2	27.9 × Earth surface gravity = 273 m/s^2
Solar constant	Instrument sensitive to radiation at all wavelengths	1370 W/m^2
Luminosity	Solar constant × area of spherical surface 1 AU in radius	3.8×10^{26} W
Spectral class	Spectrum	G2V
Effective temperature	Derived from luminosity and radius of the Sun	5800 K
Rotation period at equator	Sunspots and Doppler shift in spectra taken at the edge of the Sun	24 days 16 hours
Inclination of equator to ecliptic	Motions of sunspots	7°10′.5

Composition of the Sun's Atmosphere

Let's begin by asking what the solar atmosphere is made of. As explained in Radiation and Spectra, we can use a star's *absorption line spectrum* to determine what elements are present. It turns out that the Sun contains the same elements as Earth but *not* in the same proportions. About 73% of the Sun's mass is hydrogen, and another 25% is helium. All the other chemical elements (including those we know and love in our own bodies, such as carbon, oxygen, and nitrogen) make up only 2% of our star. The 10 most abundant gases in the Sun's visible surface layer are listed in Table 15.2. Examine that table and notice that the composition of the Sun's outer layer is very different from Earth's crust, where we live. (In our planet's crust, the three most abundant elements are oxygen, silicon, and aluminum.) Although not like our planet's, the makeup of the Sun is quite typical of stars in general.

Table 15.2: The Abundance of Elements in the Sun

Element	Percentage by Number of Atoms	Percentage By Mass
Hydrogen	92.0	73.4
Helium	7.8	25.0
Carbon	0.02	0.20
Nitrogen	0.008	0.09
Oxygen	0.06	0.80
Neon	0.01	0.16
Magnesium	0.003	0.06
Silicon	0.004	0.09
Sulfur	0.002	0.05
Iron	0.003	0.14

The fact that our Sun and the stars all have similar compositions and are made up of mostly hydrogen and helium was first shown in a brilliant thesis in 1925 by Cecilia Payne-Gaposchkin, the first woman to get a PhD in astronomy in the United States (Figure 15.3). However, the idea that the simplest light gases—hydrogen and helium—were the most abundant elements in stars was so unexpected and so shocking that she was persuaded her analysis of the data must be wrong. At the time, she wrote, "The enormous abundance derived for these elements in the stellar atmosphere is almost certainly not real." Even scientists sometimes find it hard to accept new ideas that do not agree with what everyone "knows" to be right.



Figure 15.3 : Cecilia Payne-Gaposchkin (1900–1979). Her 1925 doctoral thesis laid the foundations for understanding the composition of the Sun and the stars. Yet, being a woman, she was not given a formal appointment at Harvard, where she worked, until 1938 and was not appointed a professor until 1956. (credit: Smithsonian Institution)

Before Payne-Gaposchkin's work, everyone assumed that the composition of the Sun and stars would be much like that of Earth. It was 3 years after her thesis that other studies proved beyond a doubt that the enormous abundance of hydrogen and helium in the Sun is indeed real. (And, as we will see, the composition of the Sun and the stars is much more typical of the makeup of the universe than the odd concentration of heavier elements that characterizes our planet.)

Most of the elements found in the Sun are in the form of atoms, with a small number of molecules, all in the form of gases: the Sun is so hot that no matter can survive as a liquid or a solid. In fact, the Sun is so hot that many of the atoms in it are *ionized*, that is, stripped of one or more of their electrons. This removal of electrons from their atoms means that there is a large quantity of free electrons and positively charged ions in the Sun, making it an electrically charged environment—quite different from the neutral one in which you are reading this text. (Scientists call such a hot ionized gas a plasma.)

In the nineteenth century, scientists observed a spectral line at 530.3 nanometers in the Sun's outer atmosphere, called the corona (a layer we will discuss in a minute.) This line had never been seen before, and so it was assumed that this line was the result of a new element found in the corona, quickly named coronium. It was not until 60 years later that astronomers discovered that this emission was in fact due to highly ionized iron—iron with 13 of its electrons stripped off. This is how we first discovered that the Sun's atmosphere had a temperature of more than a million degrees.

The Layers of the Sun beneath the Visible Surface

Figure 15.4 shows what the Sun would look like if we could see all parts of it from the center to its outer atmosphere; the terms in the figure will become familiar to you as you read on.

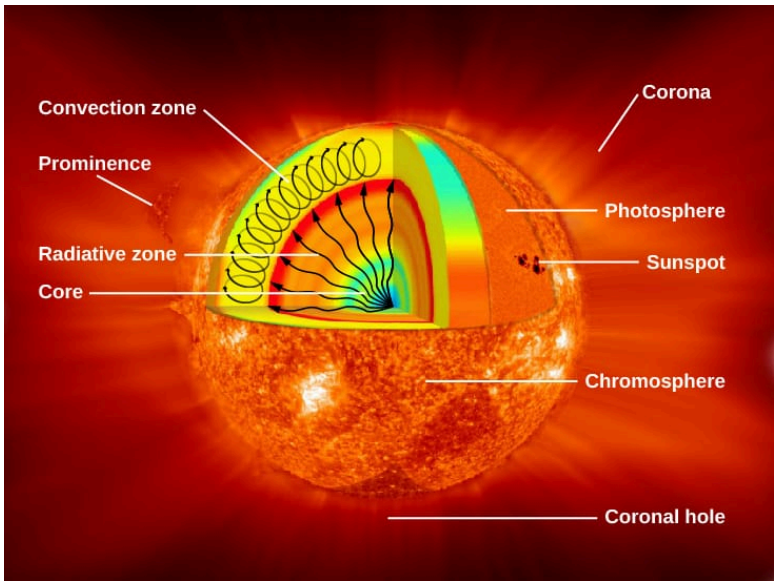


Figure 15.4 : Parts of the Sun. This illustration shows the different parts of the Sun, from the hot core where the energy is generated through regions where energy is transported outward, first by radiation, then by convection, and then out through the solar atmosphere. The parts of the atmosphere are also labeled the photosphere, chromosphere, and corona. Some typical features in the atmosphere are shown, such as coronal holes and prominences. (credit: modification of work by NASA/Goddard)

The Sun's layers are different from each other, and each plays a part in producing the energy that the Sun ultimately emits. We will begin with the core and work our way out through the layers. The Sun's *core* is extremely dense and is the source of all of its energy. Inside the core, nuclear energy is being released (in ways we will discuss in The Sun: A Nuclear Powerhouse). The core is approximately 20% of the size of the solar interior and is thought to have a temperature of approximately 15 million K, making it the hottest part of the Sun.

Above the core is a region known as the *radiative zone*—named for the primary mode of transporting energy across it. This region starts at about 25% of the distance to the solar surface and extends up to about 70% of the way to the surface. The light generated in the core is transported through the radiative zone very slowly, since the high density of matter in this region means a photon cannot travel too far without encountering a particle, causing it to change direction and lose some energy.

The *convective zone* is the outermost layer of the solar interior. It is a thick layer approximately 200,000 kilometers deep that transports energy from the edge of the radiative zone to the surface through giant convection cells, similar to a pot of boiling oatmeal. The plasma at the bottom of the convective zone is extremely hot, and it bubbles to the surface where it loses its heat to space. Once the plasma cools, it sinks back to the bottom of the convective zone.

Now that we have given a quick overview of the structure of the whole Sun, in this section, we will embark on a journey through the visible layers of the Sun, beginning with the photosphere—the visible surface.

The Solar Photosphere

Earth's air is generally transparent. But on a smoggy day in many cities, it can become opaque, which prevents us from seeing through it past a certain point. Something similar happens in the Sun. Its outer atmosphere is transparent, allowing us to look a short distance through it. But when we try to look through the atmosphere deeper into the Sun, our view is blocked. The photosphere is the layer where the Sun becomes opaque and marks the boundary past which we cannot see (Figure 15.5).

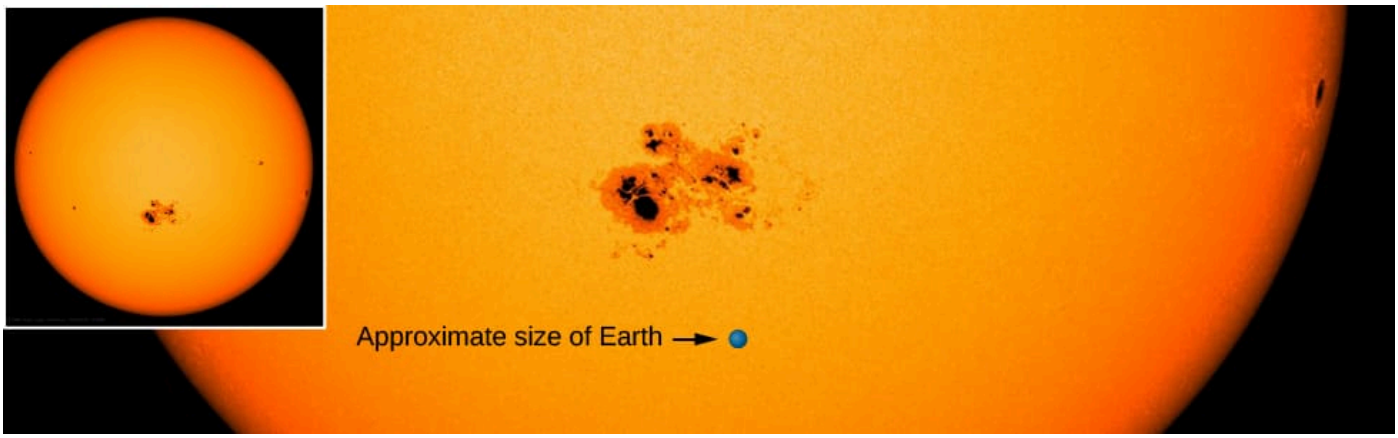


Figure 15.5 : Solar Photosphere plus Sunspots. This photograph shows the photosphere—the visible surface of the Sun. Also shown is an enlarged image of a group of sunspots; the size of Earth is shown for comparison. Sunspots appear darker because they are cooler than their surroundings. The typical temperature at the center of a large sunspot is about 3800 K, whereas the photosphere has a temperature of about 5800 K. (credit: modification of work by NASA/SDO)

As we saw, the energy that emerges from the photosphere was originally generated deep inside the Sun (more on this in The Sun: A Nuclear Powerhouse). This energy is in the form of photons, which make their way slowly toward the solar surface. Outside the Sun, we can observe *only* those photons that are emitted into the solar photosphere, where the density of atoms is sufficiently low and the photons can finally escape from the Sun without colliding with another atom or ion.

As an analogy, imagine that you are attending a big campus rally and have found a prime spot near the center of the action. Your friend arrives late and calls you on your cell phone to ask you to join her at the edge of the crowd. You decide that friendship is worth more than a prime spot, and so you work your way out through the dense crowd to meet her. You can move only a short distance before bumping into someone, changing direction, and trying again, making your way slowly to the outside edge of the crowd. All this while, your efforts are not visible to your waiting friend at the edge. Your friend can't see you until you get very close to the edge because of all the bodies in the way. So too photons making their way through the Sun are constantly bumping into atoms, changing direction, working their way slowly outward, and becoming visible only when they reach the atmosphere of the Sun where the density of atoms is too low to block their outward progress.

Astronomers have found that the solar atmosphere changes from almost perfectly transparent to almost completely opaque in a distance of just over 400 kilometers; it is this thin region that we call the *photosphere*, a word that comes from the Greek for "light sphere." When astronomers speak of the "diameter" of the Sun, they mean the size of the region surrounded by the photosphere.

The photosphere looks sharp only from a distance. If you were falling into the Sun, you would not feel any surface but would just sense a gradual increase in the density of the gas surrounding you. It is much the same as falling through a cloud while skydiving. From far away, the cloud looks as if it has a sharp surface, but you do not feel a surface as you fall into it. (One big difference between these two scenarios, however, is temperature. The Sun is so hot that you would be vaporized long before you reached the photosphere. Skydiving in Earth's atmosphere is much safer.)

We might note that the atmosphere of the Sun is not a very dense layer compared to the air in the room where you are reading this text. At a typical point in the photosphere, the pressure is less than 10% of Earth's pressure at sea level, and the density is about one ten-thousandth of Earth's atmospheric density at sea level.

Observations with telescopes show that the photosphere has a mottled appearance, resembling grains of rice spilled on a dark tablecloth or a pot of boiling oatmeal. This structure of the photosphere is called granulation (see Figure 15.6). Granules, which are typically 700 to 1000 kilometers in diameter (about the width of Texas), appear as bright areas surrounded by narrow, darker (cooler) regions. The lifetime of an individual granule is only 5 to 10 minutes. Even larger are supergranules, which are about 35,000 kilometers across (about the size of two Earths) and last about 24 hours.

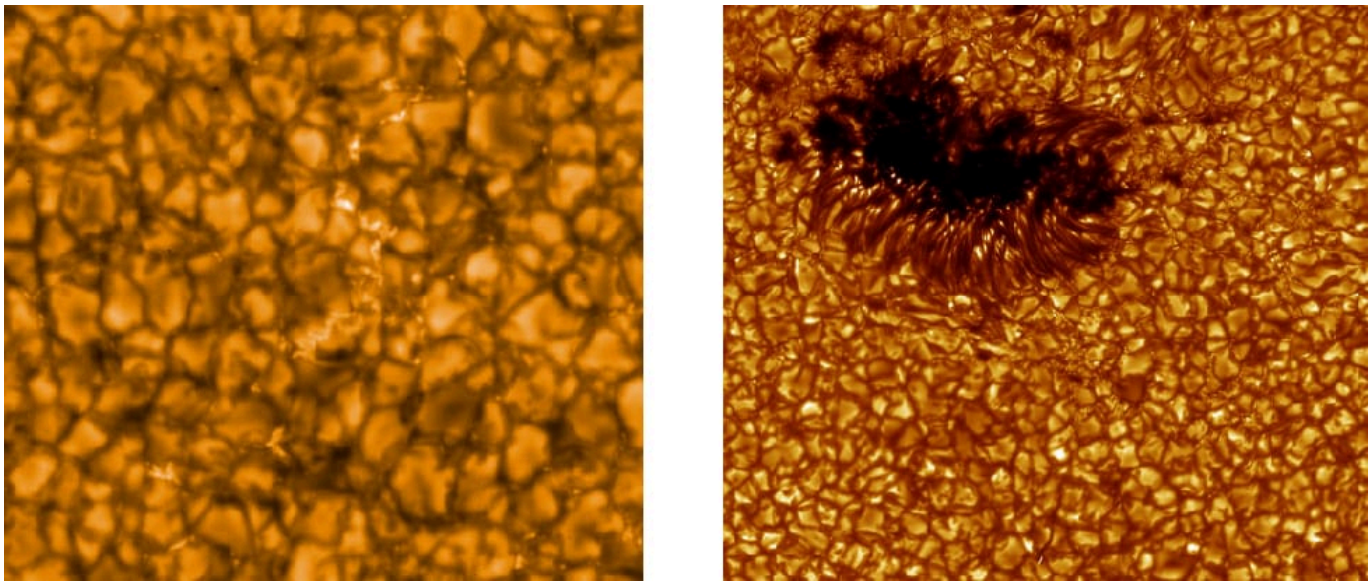


Figure 15.6 : Granulation Pattern. The surface markings of the convection cells create a granulation pattern on this dramatic image (left) taken from the Japanese Hinode spacecraft. You can see the same pattern when you heat up miso soup. The right image shows an irregular-shaped sunspot and granules on the Sun's surface, seen with the Swedish Solar Telescope on August 22, 2003. (credit left: modification of work by Hinode JAXA/NASA/PPARC; credit right: ISP/SST/Oddbjorn Engvold, Jun Elin Wiik, Luc Rouppe van der Voort)

The motions of the granules can be studied by examining the Doppler shifts in the spectra of gases just above them (see The Doppler Effect). The bright granules are columns of hotter gases rising at speeds of 2 to 3 kilometers per second from below the photosphere. As this rising gas reaches the photosphere, it spreads out, cools, and sinks down again into the darker regions between the granules. Measurements show that the centers of the granules are hotter than the intergranular regions by 50 to 100 K.

Link to Learning

See the "boiling" action of granulation in this 30-second time-lapse video from the Swedish Institute for Solar Physics.

The Chromosphere

The Sun's outer gases extend far beyond the photosphere (Figure 15.7). Because they are transparent to most visible radiation and emit only a small amount of light, these outer layers are difficult to observe. The region of the Sun's atmosphere that lies immediately above the photosphere is called the chromosphere. Until this century, the chromosphere was visible only when the photosphere was concealed by the Moon during a total solar eclipse (see the chapter on Earth, Moon, and Sky). In the seventeenth century, several observers described what appeared to them as a narrow red "streak" or "fringe" around the edge of the Moon during a brief instant after the Sun's photosphere had been covered. The name *chromosphere*, from the Greek for "colored sphere," was given to this red streak.

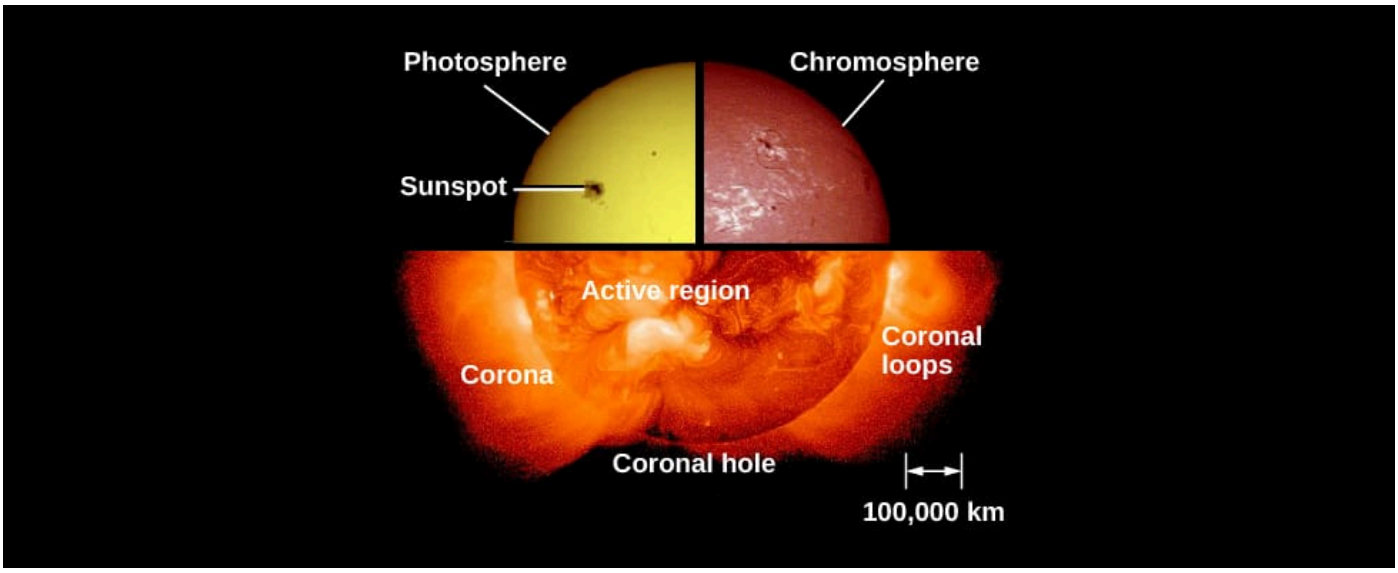


Figure 15.7 : The Sun's Atmosphere. Composite image showing the three components of the solar atmosphere: the photosphere or surface of the Sun taken in ordinary light; the chromosphere, imaged in the light of the strong red spectral line of hydrogen (H-alpha); and the corona as seen with X-rays. (credit: modification of work by NASA)

Observations made during eclipses show that the chromosphere is about 2000 to 3000 kilometers thick, and its spectrum consists of bright emission lines, indicating that this layer is composed of hot gases emitting light at discrete wavelengths. The reddish color of the chromosphere arises from one of the strongest emission lines in the visible part of its spectrum—the bright red line caused by hydrogen, the element that, as we have already seen, dominates the composition of the Sun.

In 1868, observations of the chromospheric spectrum revealed a yellow emission line that did not correspond to any previously known element on Earth. Scientists quickly realized they had found a new element and named it *helium* (after *helio*s, the Greek word for "Sun"). It took until 1895 for helium to be discovered on our planet. Today, students are probably most familiar with it as the light gas used to inflate balloons, although it turns out to be the second-most abundant element in the universe.

The temperature of the chromosphere is about 10,000 K. This means that the chromosphere is hotter than the photosphere, which should seem surprising. In all the situations we are familiar with, temperatures fall as one moves away from the source of heat, and the chromosphere is farther from the center of the Sun than the photosphere is.

The Transition Region

The increase in temperature does not stop with the chromosphere. Above it is a region in the solar atmosphere where the temperature changes from 10,000 K (typical of the chromosphere) to nearly a million degrees. The hottest part of the solar atmosphere, which has a temperature of a million degrees or more, is called the corona. Appropriately, the part of the Sun where the rapid temperature rise occurs is called the transition region. It is probably only a few tens of kilometers thick. Figure 15.8 summarizes how the temperature of the solar atmosphere changes from the photosphere outward.

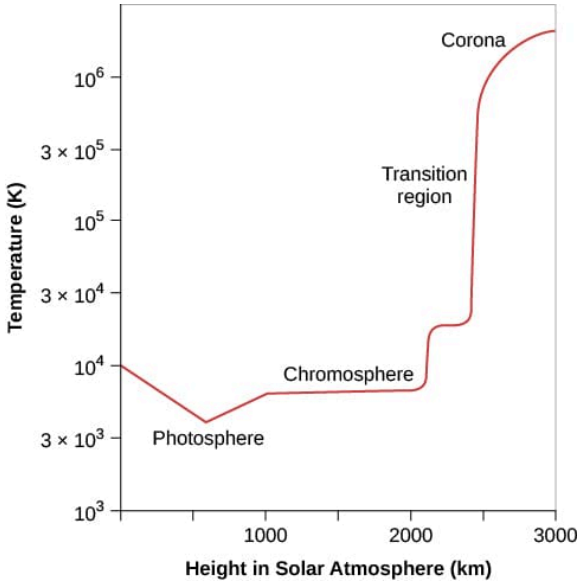


Figure 15.8 : Temperatures in the Solar Atmosphere. On this graph, temperature is shown increasing upward, and height above the photosphere is shown increasing to the right. Note the very rapid increase in temperature over a very short distance in the transition region between the chromosphere and the corona.

In 2013, NASA launched the Interface Region Imaging Spectrograph (IRIS) to study the transition region to understand better how and why this sharp temperature increase occurs. IRIS is the first space mission that is able to obtain high spatial resolution images of the different features produced over this wide temperature range and to see how they change with time and location (Figure 15.9).

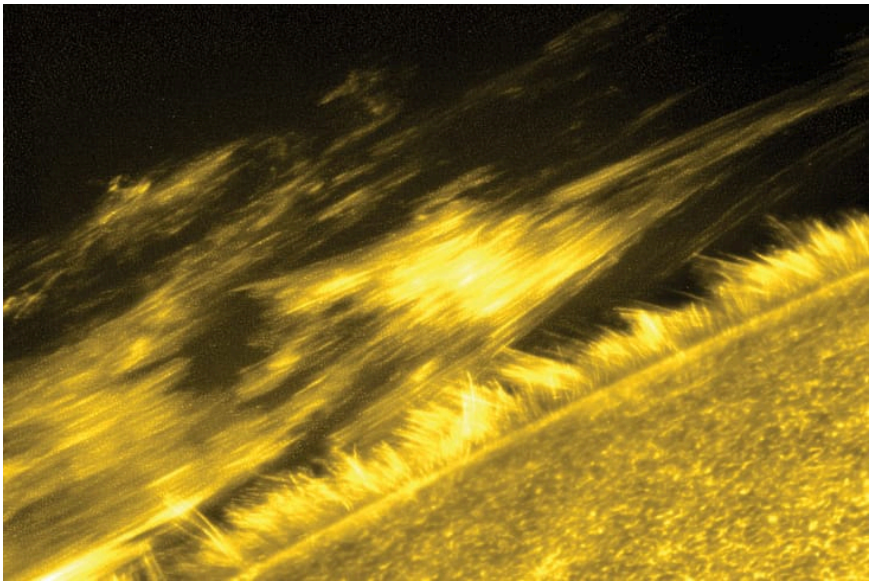


Figure 15.9 : Portion of the Transition Region. This image shows a giant ribbon of relatively cool gas threading through the lower portion of the hot corona. This ribbon (the technical term is filament) is made up of many individual threads. Time-lapse movies of this filament showed that it gradually heated as it moved through the corona. Scientists study events like this in order to try to understand what heats the chromosphere and corona to high temperatures. The “whiskers” at the edge of the Sun are spicules, jets of gas that shoot material up from the Sun’s surface and disappear after only a few minutes. This single image gives a hint of just how complicated it is to construct a model of the all the different structures and heating mechanisms in the solar atmosphere. (credit: JAXA/NASA/Hinode)

Figure 15.4 and the red graph in Figure 15.8 make the Sun seem rather like an onion, with smooth spherical shells, each one with a different temperature. For a long time, astronomers did indeed think of the Sun this way. However, we now know that while this idea of layers—photosphere, chromosphere, transition region, corona—describes the big picture fairly well, the Sun’s atmosphere is really more complicated, with hot and cool regions intermixed. For example, clouds of carbon monoxide gas with temperatures colder than 4000 K have now been found at the same height above the photosphere as the much hotter gas of the chromosphere.

The Corona

The outermost part of the Sun's atmosphere is called the *corona*. Like the chromosphere, the corona was first observed during total eclipses (Figure 15.10). Unlike the chromosphere, the corona has been known for many centuries: it was referred to by the Roman historian Plutarch and was discussed in some detail by Kepler.

The corona extends millions of kilometers above the photosphere and emits about half as much light as the full moon. The reason we don't see this light until an eclipse occurs is the overpowering brilliance of the photosphere. Just as bright city lights make it difficult to see faint starlight, so too does the intense light from the photosphere hide the faint light from the corona. While the best time to see the corona from Earth is during a total solar eclipse, it can be observed easily from orbiting spacecraft. Its brighter parts can now be photographed with a special instrument—a coronagraph—that removes the Sun's glare from the image with an occulting disk (a circular piece of material held so it is just in front of the Sun).

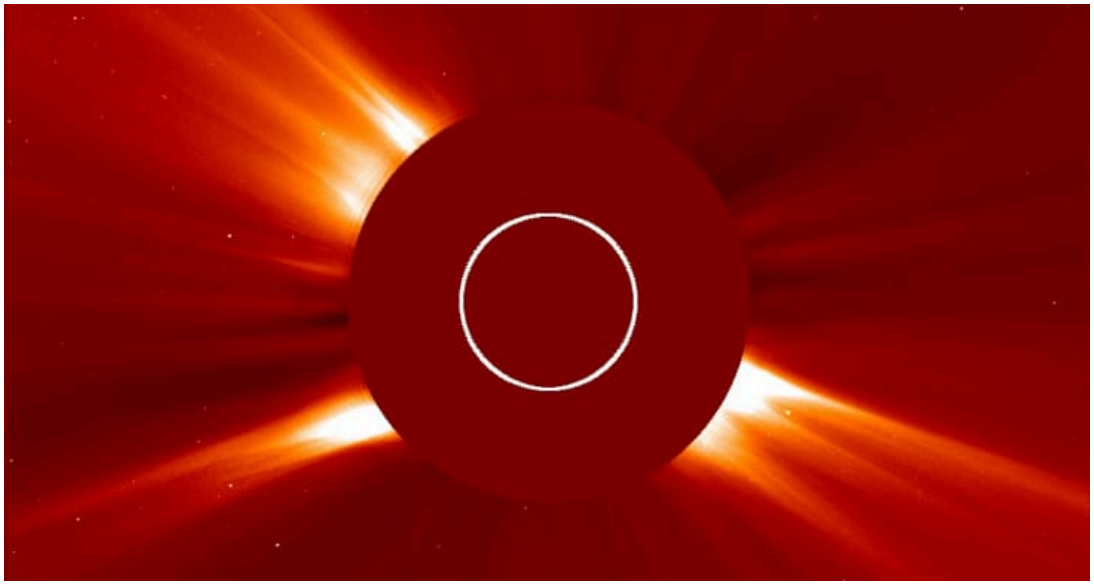


Figure 15.10 : Coronagraph. This image of the Sun was taken March 2, 2016. The larger dark circle in the center is the disk the blocks the Sun's glare, allowing us to see the corona. The smaller inner circle is where the Sun would be if it were visible in this image. (credit: modification of work by NASA/SOHO)

Studies of its spectrum show the corona to be very low in density. At the bottom of the corona, there are only about 10^9 atoms per cubic centimeter, compared with about 10^{16} atoms per cubic centimeter in the upper photosphere and 10^{19} molecules per cubic centimeter at sea level in Earth's atmosphere. The corona thins out very rapidly at greater heights, where it corresponds to a high vacuum by Earth laboratory standards. The corona extends so far into space—far past Earth—that here on our planet, we are technically living in the Sun's atmosphere.

The Solar Wind

One of the most remarkable discoveries about the Sun's atmosphere is that it produces a stream of charged particles (mainly protons and electrons) that we call the solar wind. These particles flow outward from the Sun into the solar system at a speed of about 400 kilometers per second (almost 1 million miles per hour)! The solar wind exists because the gases in the corona are so hot and moving so rapidly that they cannot be held back by solar gravity. (This wind was actually discovered by its effects on the charged tails of comets; in a sense, we can see the comet tails blow in the solar breeze the way wind socks at an airport or curtains in an open window flutter on Earth.)

Although the solar wind material is very, very rarified (i.e., *extremely* low density), the Sun has an enormous surface area. Astronomers estimate that the Sun is losing about 1–2 million tons of material each second through this wind. Although this sounds like a lot, it's so trivial compared to the enormous mass of the Sun that it can be neglected as we study the Sun.

From where in the Sun does the solar wind emerge? In visible photographs, the solar corona appears fairly uniform and smooth. X-ray and extreme ultraviolet pictures, however, show that the corona has loops, plumes, and both bright and dark regions. Large

dark regions of the corona that are relatively cool and quiet are called coronal holes (Figure 15.11). In these regions, magnetic field lines stretch far out into space away from the Sun, rather than looping back to the surface. The solar wind comes predominantly from coronal holes, where gas can stream away from the Sun into space unhindered by magnetic fields. Hot coronal gas, on the other hand, is present mainly where magnetic fields have trapped and concentrated it.

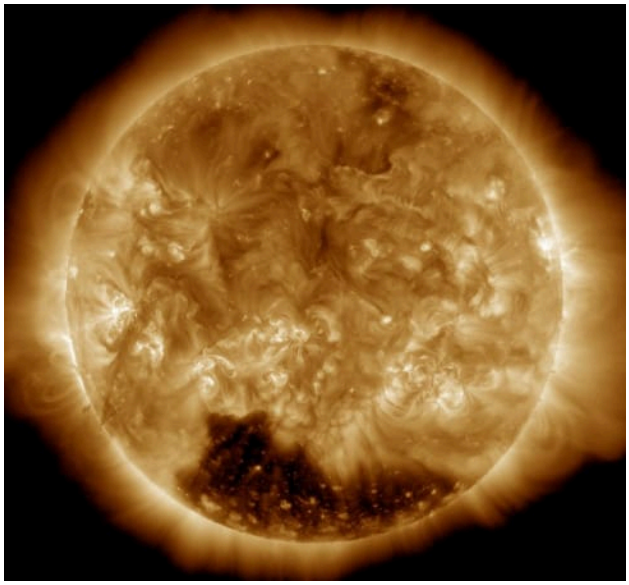


Figure 15.11 Coronal Hole. The dark area visible near the Sun's south pole on this Solar Dynamics Observer spacecraft image is a coronal hole. (credit: modification of work by NASA/SDO)

At the surface of Earth, we are protected to some degree from the solar wind by our atmosphere and Earth's magnetic field (see Earth as a Planet). However, the magnetic field lines come into Earth at the north and south magnetic poles. Here, charged particles accelerated by the solar wind can follow the field down into our atmosphere. As the particles strike molecules of air, they cause them to glow, producing beautiful curtains of light called the auroras, or the northern and southern lights (Figure 15.12).



Figure 15.12 : Aurora. The colorful glow in the sky results from charged particles in a solar wind interacting with Earth's magnetic fields. The stunning display captured here occurred over Jokulsarlon Lake in Iceland in 2013. (credit: Moyan Brenn)

Link to Learning

This NASA video explains and demonstrates the nature of the auroras and their relationship to Earth's magnetic field.

This page titled 19.1: The Structure and Composition of the Sun is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 15.2: The Structure and Composition of the Sun by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.2: The Solar Cycle

Learning Objectives

By the end of this section, you will be able to:

- Describe the sunspot cycle and, more generally, the solar cycle
- Explain how magnetism is the source of solar activity

Before the invention of the telescope, the Sun was thought to be an unchanging and perfect sphere. We now know that the Sun is in a perpetual state of change: its surface is a seething, bubbling cauldron of hot gas. Areas that are darker and cooler than the rest of the surface come and go. Vast plumes of gas erupt into the chromosphere and corona. Occasionally, there are even giant explosions on the Sun that send enormous streamers of charged particles and energy hurtling toward Earth. When they arrive, these can cause power outages and other serious effects on our planet.

Sunspots

The first evidence that the Sun changes came from studies of sunspots, which are large, dark features seen on the surface of the Sun caused by increased magnetic activity. They look darker because the spots are typically at a temperature of about 3800 K, whereas the bright regions that surround them are at about 5800 K (Figure 15.13). Occasionally, these spots are large enough to be visible to the unaided eye, and we have records going back over a thousand years from observers who noticed them when haze or mist reduced the Sun's intensity. (We emphasize what your parents have surely told you: looking at the Sun for even a brief time can cause permanent eye damage. This is the one area of astronomy where we don't encourage you to do your own observing without getting careful instructions or filters from your instructor.)

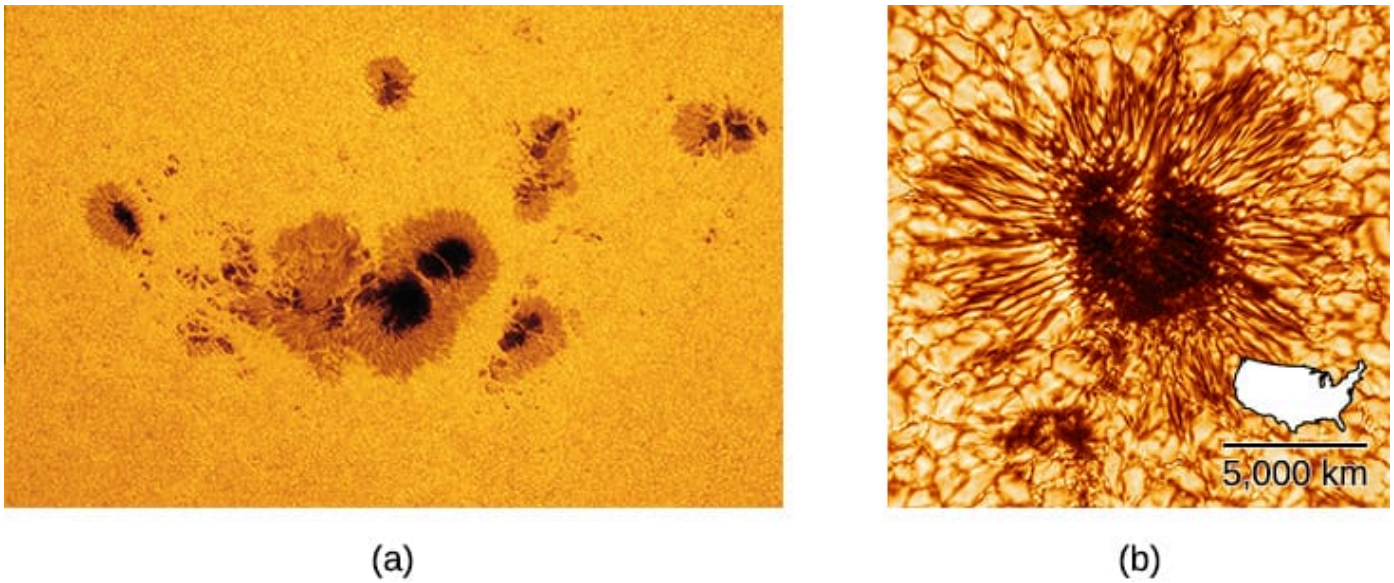


Figure 15.13 : Sunspots. a) Image of sunspots, cooler and thus darker regions on the Sun, taken in July 2012. You can see the dark, central region of each sunspot (called the umbra) surrounded by a less dark region (the penumbra). The largest spot shown here is about 11 Earths wide. b) The new Daniel Inouye Solar Telescope took the most detailed image of a sunspot ever recorded in January 2020. A map of the United States is shown to scale. Although sunspots appear dark when seen next to the hotter gases of the photosphere, an average sunspot, cut out of the solar surface and left standing in the night sky, would be about as bright as the full moon. The mottled appearance of the Sun's surface is granulation. (credit a: modification of work by NASA Goddard Space Flight Center, Alan Friedman; credit b: modification of work by NSO/AURA/NSF)

While we understand that sunspots look darker because they are cooler, they are nevertheless hotter than the surfaces of many stars. If they could be removed from the Sun, they would shine brightly. They appear dark only in contrast with the hotter, brighter

photosphere around them.

Individual sunspots come and go, with lifetimes that range from a few hours to a few months. If a spot lasts and develops, it usually consists of two parts: an inner darker core, the *umbra*, and a surrounding less dark region, the *penumbra*. Many spots become much larger than Earth, and a few, like the largest one shown in Figure 15.13, have reached diameters over 140,000 kilometers. Frequently, spots occur in groups of 2 to 20 or more. The largest groups are very complex and may have over 100 spots. Like storms on Earth, sunspots are not fixed in position, but they drift slowly compared with the Sun's rotation.

By recording the apparent motions of the sunspots as the turning Sun carried them across its disk (Figure 15.14), Galileo, in 1612, demonstrated that the Sun rotates on its axis with a rotation period of approximately 1 month. Our star turns in a west-to-east direction, like the orbital motions of the planets. The Sun, however, is a gas and does not have to rotate rigidly, the way a solid body like Earth does. Modern observations show that the speed of rotation of the Sun varies according to latitude, that is, it's different as you go north or south of the Sun's equator. The rotation period is about 25 days at the equator, 28 days at latitude 40°, and 36 days at latitude 80°. We call this behavior differential rotation.

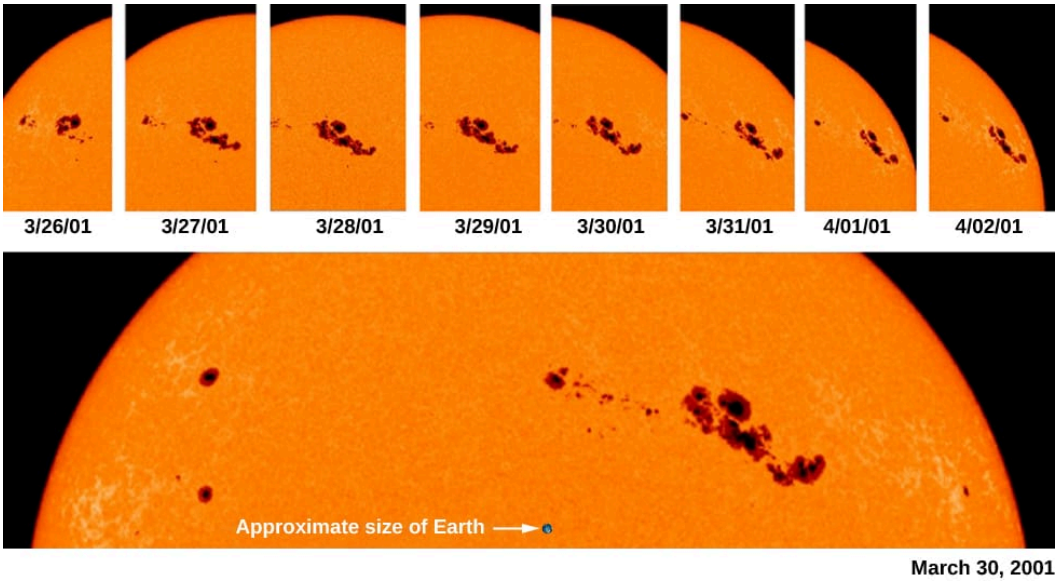


Figure 15.14 : Sunspots Rotate Across Sun's Surface. This sequence of photographs of the Sun's surface tracks the movement of sunspots across the visible hemisphere of the Sun. On March 30, 2001, this group of sunspots extended across an area about 13 times the diameter of Earth. This region produced many flares and coronal mass ejections. (credit: modification of work by SOHO/NASA/ESA)

The Sunspot Cycle

Between 1826 and 1850, Heinrich Schwabe, a German pharmacist and amateur astronomer, kept daily records of the number of sunspots. What he was really looking for was a planet inside the orbit of Mercury, which he hoped to find by observing its dark silhouette as it passed between the Sun and Earth. He failed to find the hoped-for planet, but his diligence paid off with an even-more important discovery: the sunspot cycle. He found that the number of sunspots varied systematically, in cycles about a decade long.

What Schwabe observed was that, although individual spots are short lived, the total number visible on the Sun at any one time was likely to be very much greater at certain times—the periods of *sunspot maximum*—than at other times—the periods of *sunspot minimum*. We now know that sunspot maxima occur at an *average* interval of 11 years, but the intervals between successive maxima have ranged from as short as 9 years to as long as 14 years. During sunspot maxima, more than 100 spots can often be seen at once. Even then, less than one-half of one percent of the Sun's surface is covered by spots (Figure 15.22). During sunspot minima, sometimes no spots are visible. The Sun's activity reached its most recent maximum in 2014.

Link to Learning

Watch this brief video from NASA's Goddard Space Flight Center that explains the sunspot cycle.

Explore the number of sunspots measured each year going back to 1707. Point your cursor at any year, and the diagram tells you the number of sunspots.

Magnetism and the Solar Cycle

Now that we have discussed the Sun's activity cycle, you might be asking, "Why does the Sun change in such a regular way?" Astronomers now understand that it is the Sun's changing magnetic field that drives solar activity.

The solar magnetic field is measured using a property of atoms called the *Zeeman effect*. Recall from Radiation and Spectra that an atom has many energy levels and that spectral lines are formed when electrons shift from one level to another. If each energy level is precisely defined, then the difference between them is also quite precise. As an electron changes levels, the result is a sharp, narrow spectral line (either an absorption or emission line, depending on whether the electron's energy increases or decreases in the transition).

In the presence of a strong magnetic field, however, each energy level is separated into several levels very close to one another. The separation of the levels is proportional to the strength of the field. As a result, spectral lines formed in the presence of a magnetic field are not single lines but a series of very closely spaced lines corresponding to the subdivisions of the atomic energy levels. This splitting of lines in the presence of a magnetic field is what we call the Zeeman effect (after the Dutch scientist who first discovered it in 1896).

Measurements of the Zeeman effect in the spectra of the light from sunspot regions show them to have strong magnetic fields (Figure 15.15). Bear in mind that magnets always have a north pole and a south pole. Whenever sunspots are observed in pairs, or in groups containing two principal spots, one of the spots usually has the magnetic polarity of a north-seeking magnetic pole and the other has the opposite polarity. Moreover, during a given cycle, the leading spots of pairs (or leading principle spots of groups) in the Northern Hemisphere all tend to have the same polarity, whereas those in the Southern Hemisphere all tend to have the opposite polarity.

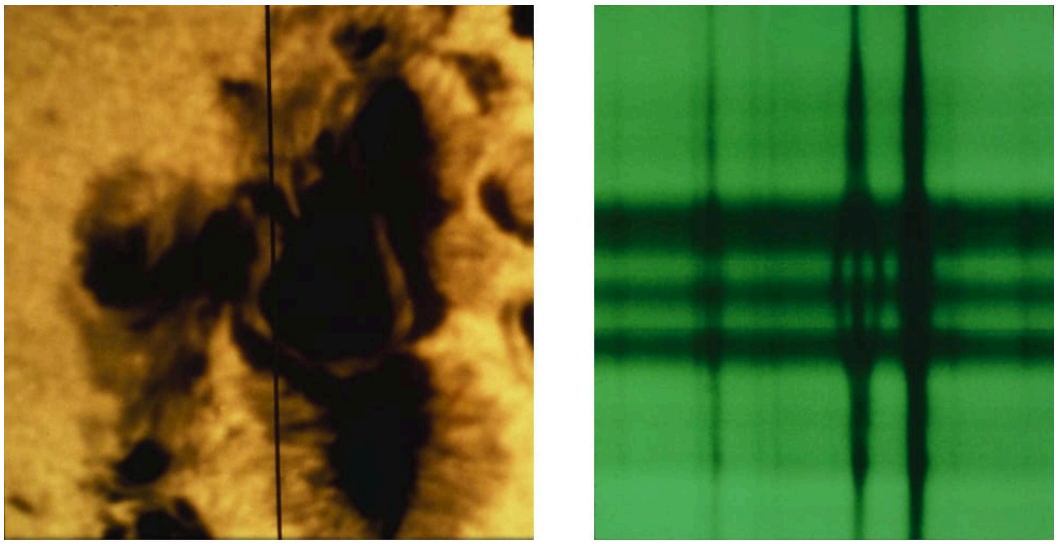


Figure 15.15 : Zeeman Effect. These photographs show how magnetic fields in sunspots are measured by means of the Zeeman effect. (left) The vertical black line indicates the position of the spectrograph slit through which light is passed to obtain the spectrum in (right). (credit: modification of work by NSO/AURA/NSF)

During the next sunspot cycle, however, the polarity of the leading spots is reversed in each hemisphere. For example, if during one cycle, the leading spots in the Northern Hemisphere all had the polarity of a north-seeking pole, then the leading spots in the Southern Hemisphere would have the polarity of a south-seeking pole. During the next cycle, the leading spots in the Northern Hemisphere would have south-seeking polarity, whereas those in the Southern Hemisphere would have north-seeking polarity. Therefore, strictly speaking, the sunspot cycle does not repeat itself in regard to magnetic polarity until two 11-year cycles have passed. A visual representation of the Sun's magnetic fields, called a *magnetogram*, can be used to see the relationship between sunspots and the Sun's magnetic field (Figure 15.16).

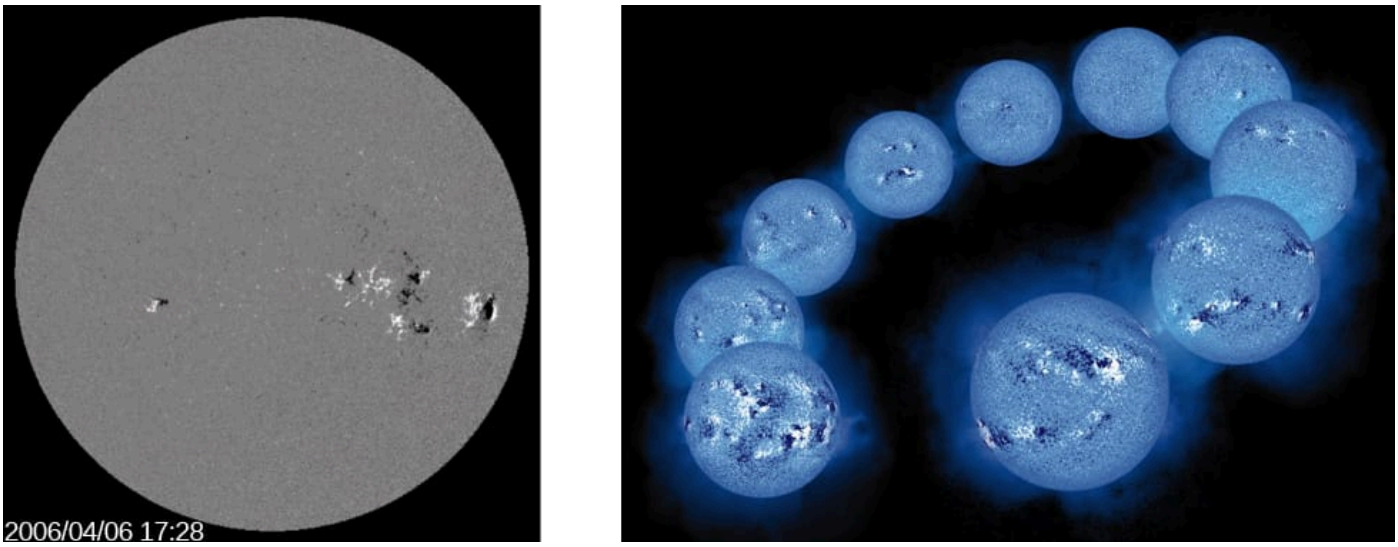


Figure 15.16 : Magnetogram and Solar Cycle. In the image on the left, called a magnetogram, we see the magnetic polarity of sunspots. The black areas are where the magnetism is pointing toward the Sun's core, whereas the white regions are where it is pointing away from the core, toward us. This dramatic sequence on the right shows the activity cycle of the Sun. The 10 maps of the magnetic field on the surface of the Sun span a period of 7.5 years. The two magnetic polarities (N and S) of the magnetic field are shown against a blue disk as dark blue to black (N) and as light blue to white (S). The earliest image, taken on January 8, 1992, is at the lower left and was taken just after solar maximum. Each image, from left to right around the arc, was taken one-half to one year after the preceding one. The last image was taken on July 25, 1999, as the Sun was approaching the next solar maximum. Note a few striking patterns in the magnetic maps: the direction from white to black polarity in the Southern Hemisphere is opposite from that in the Northern Hemisphere. (credit left: modification of work by NASA/SDO; credit right: modification of work by NASA/SOHO)

Why is the Sun such a strong and complicated magnet? Astronomers have found that it is the Sun's *dynamo* that generates the magnetic field. A dynamo is a machine that converts kinetic energy (i.e., the energy of motion) into electricity. On Earth, dynamos are found in power plants where, for example, the energy from wind or flowing water is used to cause turbines to rotate. In the Sun, the source of kinetic energy is the churning of turbulent layers of ionized gas within the Sun's interior that we mentioned earlier. These generate electric currents—moving electrons—which in turn generate magnetic fields.

Most solar researchers agree that the solar dynamo is located in the convection zone or in the interface layer between the convection zone and the radiative zone below it. As the magnetic fields from the Sun's dynamo interact, they break, reconnect, and rise through the Sun's surface.

We should say that, although we have good observations that show us *how* the Sun changes during each solar cycle, it is still very difficult to build physical models of something as complicated as the Sun that can account satisfactorily for *why* it changes. Researchers have not yet developed a generally accepted model that describes in detail the physical processes that control the solar cycle. Calculations do show that differential rotation (the idea that the Sun rotates at different rates at different latitudes) and convection just below the solar surface can twist and distort the magnetic fields. This causes them to grow and then decay, regenerating with opposite polarity approximately every 11 years. The calculations also show that as the fields grow stronger near solar maximum, they flow from the interior of the Sun toward its surface in the form of loops. When a large loop emerges from the solar surface, it creates regions of sunspot activity (Figure 15.17).

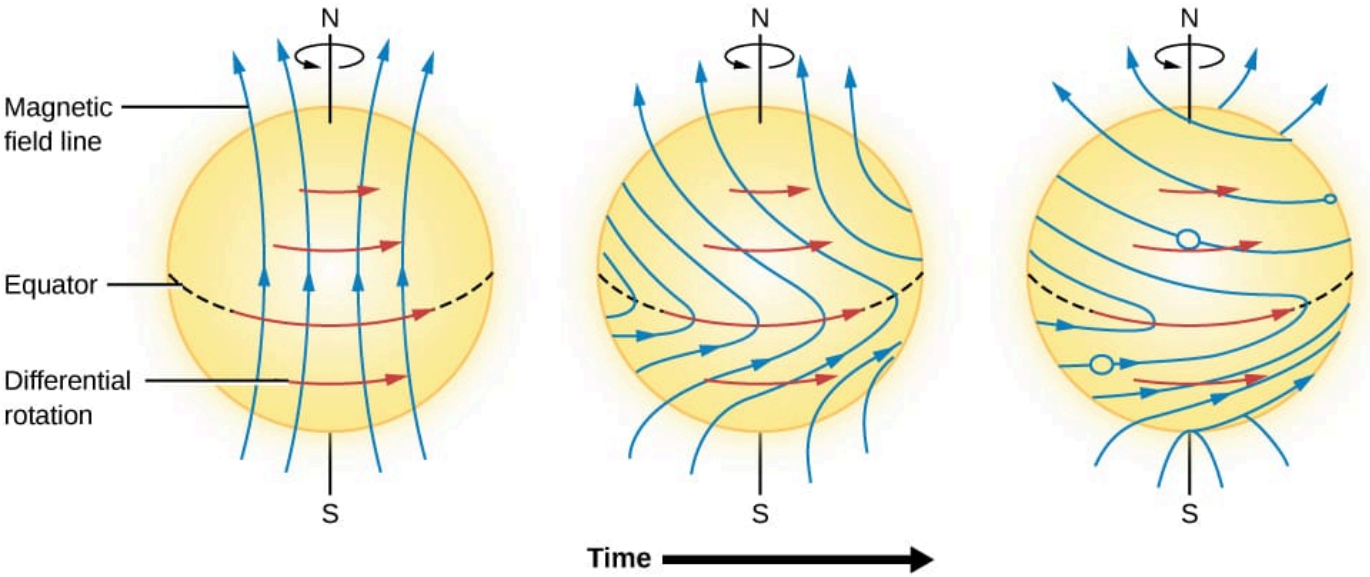


Figure 15.17 : Magnetic Field Lines Wind Up. Because the Sun spins faster at the equator than near the poles, the magnetic fields in the Sun tend to wind up as shown, and after a while make loops. This is an idealized diagram; the real situation is much more complex.

This idea of magnetic loops offers a natural explanation of why the leading and trailing sunspots in an active region have opposite polarity. The leading sunspot coincides with one end of the loop and the trailing spot with the other end. Magnetic fields also hold the key to explaining why sunspots are cooler and darker than the regions without strong magnetic fields. The forces produced by the magnetic field resist the motions of the bubbling columns of rising hot gases. Since these columns carry most of the heat from inside the Sun to the surface by means of convection, and strong magnetic fields inhibit this convection, the surface of the Sun is allowed to cool. As a result, these regions are seen as darker, cooler sunspots.

Beyond this general picture, researchers are still trying to determine why the magnetic fields are as large as they are, why the polarity of the field in each hemisphere flips from one cycle to the next, why the length of the solar cycle can vary from one cycle to the next, and why events like the Maunder Minimum occur.

Link to Learning

In this video solar scientist Holly Gilbert discusses the Sun's magnetic field.

This page titled 19.2: The Solar Cycle is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 15.3: The Solar Cycle by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.3: Solar Activity above the Photosphere

Learning Objectives

By the end of this section, you will be able to:

- Describe the various ways in which the solar activity cycle manifests itself, including flares, coronal mass ejections, prominences, and plages

Sunspots are not the only features that vary during a solar cycle. There are dramatic changes in the chromosphere and corona as well. To see what happens in the chromosphere, we must observe the emission lines from elements such as hydrogen and calcium, which emit useful spectral lines at the temperatures in that layer. The hot corona, on the other hand, can be studied by observations of X-rays and of extreme ultraviolet and other wavelengths at high energies.

Plages and Prominences

As we saw, emission lines of hydrogen and calcium are produced in the hot gases of the chromosphere. Astronomers routinely photograph the Sun through filters that transmit light only at the wavelengths that correspond to these emission lines. Pictures taken through these special filters show bright "clouds" in the chromosphere around sunspots; these bright regions are known as plages (Figure 15.18). These are regions within the chromosphere that have higher temperature and density than their surroundings. The plages actually contain all of the elements in the Sun, not just hydrogen and calcium. It just happens that the spectral lines of hydrogen and calcium produced by these clouds are bright and easy to observe.

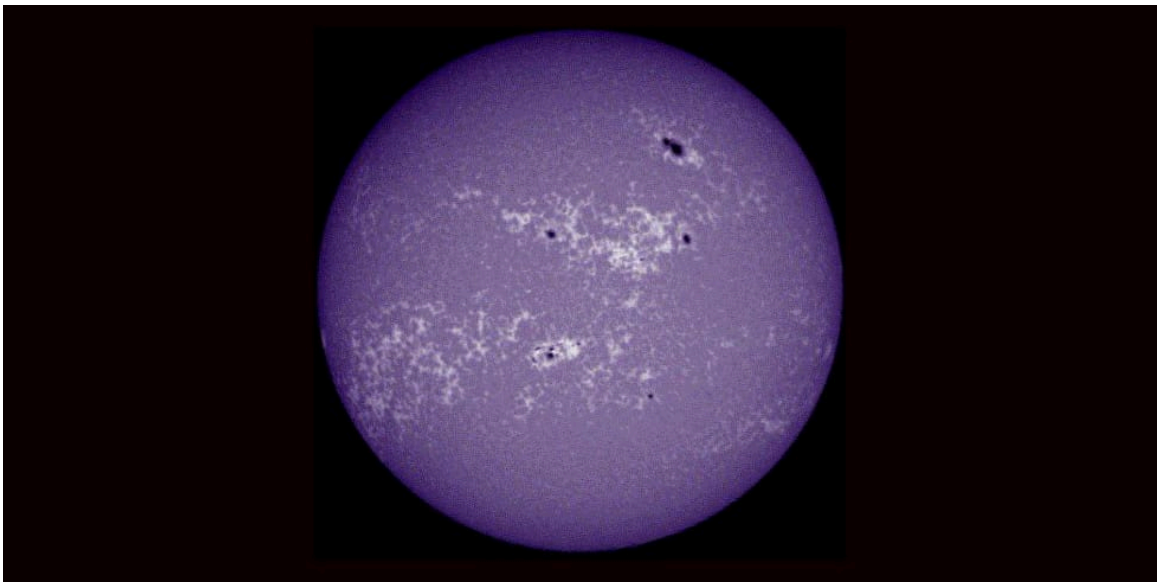


Figure 15.18 : Plages on the Sun. This image of the Sun was taken with a filter that transmits only the light of the spectral line produced by singly ionized calcium. The bright cloud-like regions are the plages. (credit: modification of work by NASA)

Moving higher into the Sun's atmosphere, we come to the spectacular phenomena called prominences (Figure 15.19), which usually originate near sunspots. Eclipse observers often see prominences as red features rising above the eclipsed Sun and reaching high into the corona. Some, the *quiescent* prominences, are graceful loops of plasma (ionized gas) that can remain nearly stable for many hours or even days. The relatively rare *eruptive* prominences appear to send matter upward into the corona at high speeds, and the most active *surge* prominences may move as fast as 1300 kilometers per second (almost 3 million miles per hour). Some eruptive prominences have reached heights of more than 1 million kilometers above the photosphere; Earth would be completely lost inside one of those awesome displays (Figure 15.19).

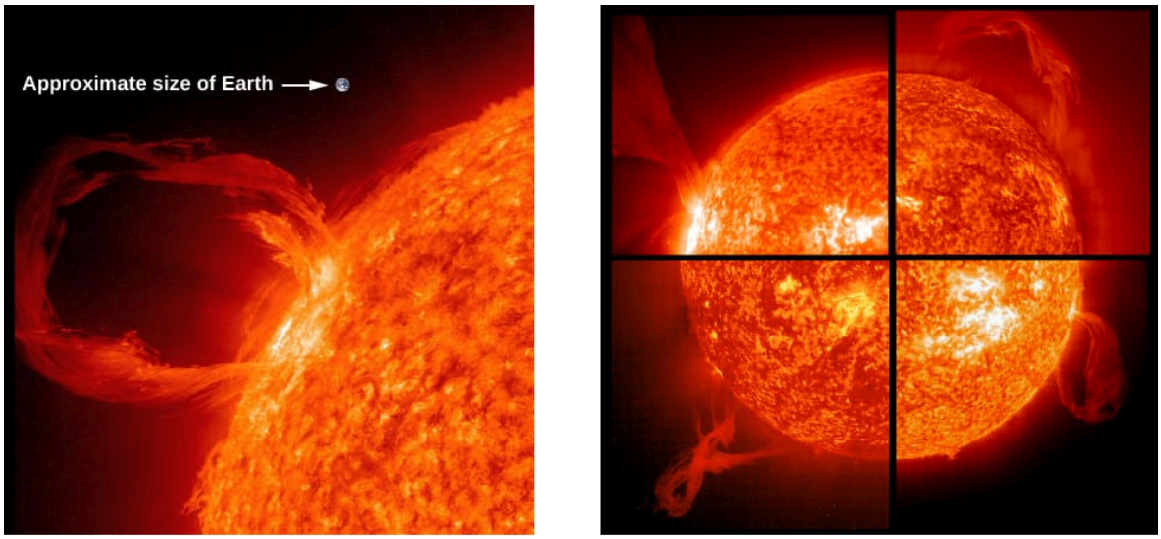


Figure 15.19 : Prominences. (a) This image of an eruptive prominence was taken in the light of singly ionized helium in the extreme ultraviolet part of the spectrum. The prominence is a particularly large one. An image of Earth is shown at the same scale for comparison. (b) A prominence is a huge cloud of relatively cool (about 60,000 K in this case), fairly dense gas suspended in the much hotter corona. These pictures, taken in ultraviolet, are color coded so that white corresponds to the hottest temperatures and dark red to cooler ones. The four images were taken, moving clockwise from the upper left, on May 15, 2001; March 28, 2000; January 18, 2000; and February 2, 2001. (credit a: modification of work by NASA/SOHO; credit b: modification of work by NASA/SDO)

Flares and Coronal Mass Ejections

The most violent event on the surface of the Sun is a rapid eruption called a solar flare (Figure 15.20). A typical flare lasts for 5 to 10 minutes and releases a total amount of energy equivalent to that of perhaps a million hydrogen bombs. The largest flares last for several hours and emit enough energy to power the entire United States at its current rate of electrical consumption for 100,000 years. Near sunspot maximum, small flares occur several times per day, and major ones may occur every few weeks.

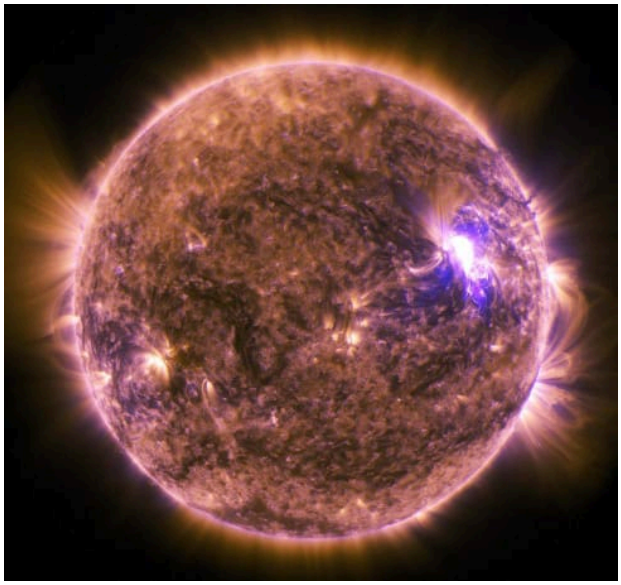


Figure 15.20 : Solar Flare. The bright white area seen on the right side of the Sun in this image from the Solar Dynamics Observer spacecraft is a solar flare that was observed on June 25, 2015. (credit: NASA/SDO)

Flares, like the one shown in Figure 15.21, are often observed in the red light of hydrogen, but the visible emission is only a tiny fraction of the energy released when a solar flare explodes. At the moment of the explosion, the matter associated with the flare is heated to temperatures as high as 10 million K. At such high temperatures, a flood of X-ray and ultraviolet radiation is emitted.

Flares seem to occur when magnetic fields pointing in opposite directions release energy by interacting with and destroying each other—much as a stretched rubber band releases energy when it breaks.

What is different about flares is that their magnetic interactions cover a large volume in the solar corona and release a tremendous amount of electromagnetic radiation. In some cases, immense quantities of coronal material—mainly protons and electrons—may also be ejected at high speeds (500–1000 kilometers per second) into interplanetary space. Such a coronal mass ejection (CME) can affect Earth in a number of ways (which we will discuss in the section on space weather).

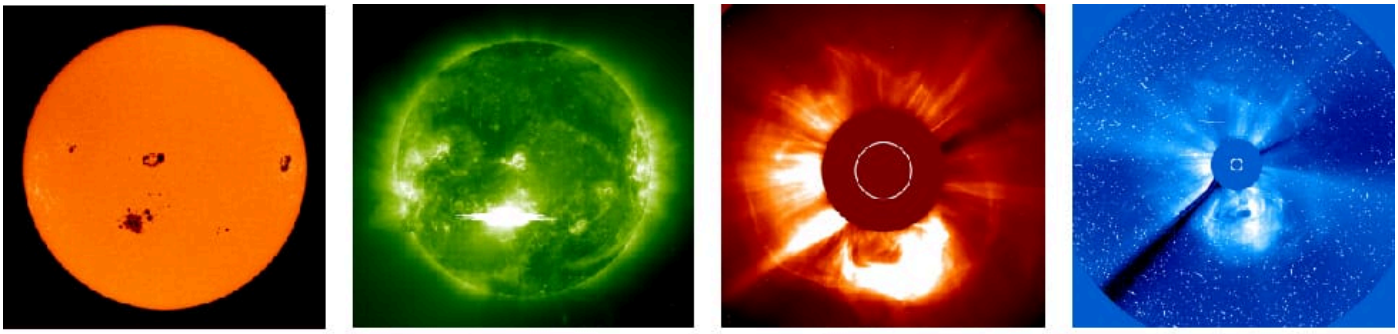


Figure 15.21 : Flare and Coronal Mass Ejection. This sequence of four images shows the evolution over time of a giant eruption on the Sun. (a) The event began at the location of a sunspot group, and (b) a flare is seen in far-ultraviolet light. (c) Fourteen hours later, a CME is seen blasting out into space. (d) Three hours later, this CME has expanded to form a giant cloud of particles escaping from the Sun and is beginning the journey out into the solar system. The white circle in (c) and (d) shows the diameter of the solar photosphere. The larger dark area shows where light from the Sun has been blocked out by a specially designed instrument to make it possible to see the faint emission from the corona. (credit a, b, c, d: modification of work by SOHO/EIT, SOHO/LASCO, SOHO/MDI (ESA & NASA))

Link to Learning

See a coronal mass ejection recorded by the Solar Dynamics Observatory.

Active Regions

To bring the discussion of the last two sections together, astronomers now realize that sunspots, flares, and bright regions in the chromosphere and corona tend to occur together on the Sun in time and space. That is, they all tend to have similar longitudes and latitudes, but they are located at different heights in the atmosphere. Because they all occur together, they vary with the sunspot cycle.

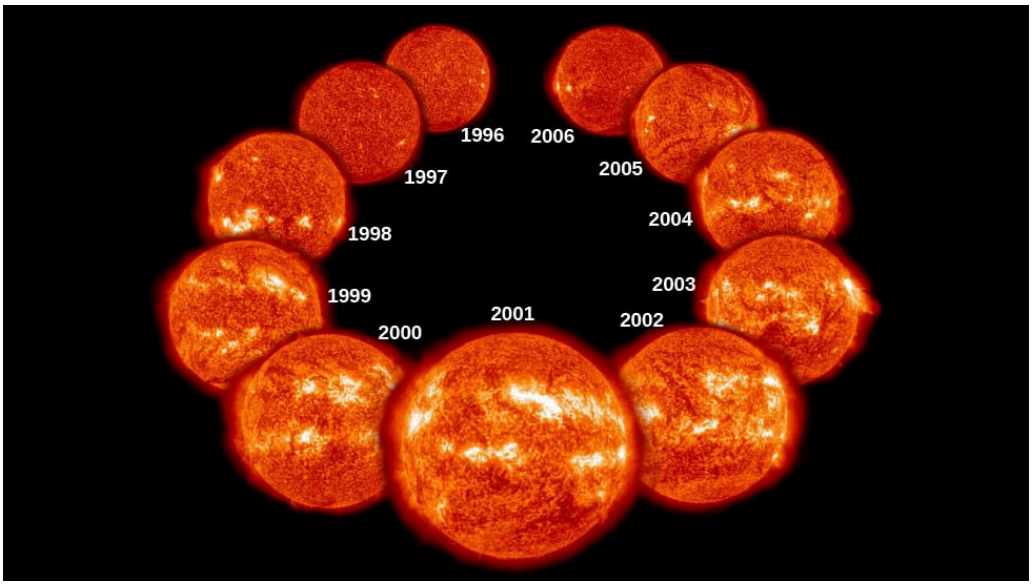


Figure 15.22 : Solar Cycle. This dramatic sequence of images taken from the SOHO satellite over a period of 11 years shows how active regions change during the solar cycle. The images were taken in the ultraviolet region of the spectrum and show that active regions on the Sun increase and decrease during the cycle. Sunspots are located in the cooler photosphere, beneath the hot gases shown in this image, and vary in phase with the emission from these hot gases—more sunspots and more emission from hot gases occur together. (credit: modification of work by ESA/NASA/SOHO)

For example, flares are more likely to occur near sunspot maximum, and the corona is much more conspicuous at that time (see Figure 15.22). A place on the Sun where a number of these phenomena are seen is called an active region (Figure 15.23). As you might deduce from our earlier discussion, active regions are always associated with strong magnetic fields.

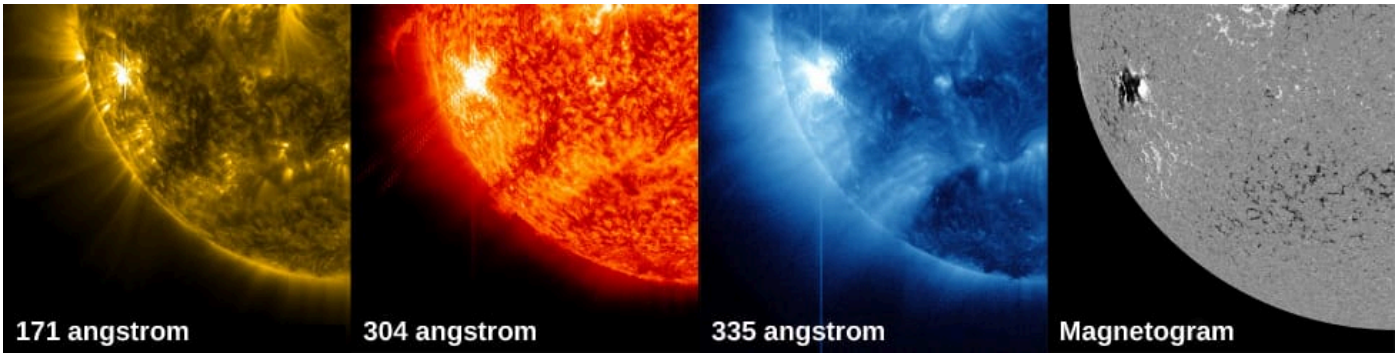


Figure 15.23 : Solar Active Region Observed at Different Heights in the Sun's Atmosphere. These four images of a solar flare on October 22, 2012, show from the left: light from the Sun at a wavelength of 171 angstroms, which shows the structure of loops of solar material in the corona; ultraviolet at 304 angstroms, which shows light from the region of the Sun's atmosphere where flares originate; light at 335 angstroms, which highlights radiation from active regions in the corona; a magnetogram, which shows magnetically active regions on the Sun. Note how these different types of activity all occur above a sunspot region with a strong magnetic field. (credit: modification of work by NASA/SDO/Goddard)

This page titled 19.3: Solar Activity above the Photosphere is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 15.4: Solar Activity above the Photosphere by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.4: Mass, Energy, and the Theory of Relativity

Learning Objectives

By the end of this section, you will be able to:

- Explain how matter can be converted into energy
- Describe the particles that make up atoms
- Describe the nucleus of an atom
- Understand the nuclear forces that hold atoms together
- Trace the nuclear reactions in the solar interior

As we have seen, energy cannot be created or destroyed, but only converted from one form to another. One of the remarkable conclusions derived by Albert Einstein (see Albert Einstein) when he developed his theory of relativity is that matter can be considered a form of energy too and can be converted into energy. Furthermore, energy can also be converted into matter. This seemed to contradict what humans had learned over thousands of years by studying nature. Matter is something we can see and touch, whereas energy is something objects have when they do things like move or heat up. The idea that matter or energy can be converted into each other seemed as outrageous as saying you could accelerate a car by turning the bumper into more speed, or that you could create a bigger front seat by slowing down your car. That would be pretty difficult to believe; yet, the universe actually works somewhat like that.

Converting Matter into Energy

The remarkable equivalence between matter and energy is given in one of the most famous equations:

$$E =$$

In this equation, E stands for energy, m stands for mass, and c, the constant that relates the two, is the speed of light $\left(3 \times 10^{8} \text{meters per second}\right)$. Note that mass is a measure of the quantity of matter, so the significance of this equation is that matter can be converted into energy and energy can be converted into matter. Let's compare this equation of converting matter and energy to some common conversion equations that have the same form:

$$\text{inches} = \text{feet} \times 12, \text{or cents} =$$

Just as each conversion formula allows you to calculate the conversion of one thing into another, when we convert matter into energy, we consider how much mass the matter has. The conversion factor in this case turns out not to be either 12 or 100, as in our examples, but another constant quantity: the speed of light squared. Note that matter does not have to travel at the speed of light (or the speed of light squared) for this conversion to occur. The factor of c^2 is just the number that Einstein showed must be used to relate mass and energy.

Notice that this formula does not tell us *how* to convert mass into energy, just as the formula for cents does not tell us where to exchange coins for a dollar bill. The formulas merely tell us what the equivalent values are if we succeed in making the conversion. When Einstein first derived his formula in 1905, no one had the faintest idea how to convert mass into energy in any practical way. Einstein himself tried to discourage speculation that the large-scale conversion of atomic mass into energy would be feasible in the near future. Today, as a result of developments in nuclear physics, we regularly convert mass into energy in power plants, nuclear weapons, and high-energy physics experiments in particle accelerators.

Because the speed of light squared (c^2) is a very large quantity, the conversion of even a small amount of mass results in a very large amount of energy. For example, the complete conversion of 1 gram of matter (about 1/28 ounce, or approximately 1 paperclip) would produce as much energy as the burning of 15,000 barrels of oil.

Scientists soon realized that the conversion of mass into energy is the source of the Sun's heat and light. With Einstein's equation of $E = mc^2$, we can calculate that the amount of energy radiated by the Sun could be produced by the complete conversion of about 4 million tons of matter into energy inside the Sun each second. Destroying 4 million tons per second sounds like a lot

when compared to earthly things, but bear in mind that the Sun is a very big reservoir of matter. In fact, we will see that the Sun contains more than enough mass to destroy such huge amounts of matter and still continue shining at its present rate for billions of years.

But knowing all that still does not tell us *how* mass can be converted into energy. To understand the process that actually occurs in the Sun, we need to explore the structure of the atom a bit further.

Voyagers in Astronomy: Albert Einstein

For a large part of his life, Albert Einstein (Figure 16.3) was one of the most recognized celebrities of his day. Strangers stopped him on the street, and people all over the world asked him for endorsements, advice, and assistance. In fact, when Einstein and the great film star Charlie Chaplin met in California, they found they shared similar feelings about the loss of privacy that came with fame. Einstein's name was a household word despite the fact that most people did not understand the ideas that had made him famous.

Einstein was born in 1879 in Ulm, Germany. Legend has it that he did not do well in school (even in arithmetic), and thousands of students have since attempted to justify a bad grade by referring to this story. Alas, like many legends, this one is not true. Records indicate that although he tended to rebel against the authoritarian teaching style in vogue in Germany at that time, Einstein was a good student.

After graduating from the Federal Polytechnic Institute in Zurich, Switzerland, Einstein at first had trouble getting a job (even as a high school teacher), but he eventually became an examiner in the Swiss Patent Office. Working in his spare time, without the benefit of a university environment but using his superb physical intuition, he wrote four papers in 1905 that would ultimately transform the way physicists looked at the world.

One of these, which earned Einstein the Nobel Prize in 1921, set part of the foundation of *quantum mechanics*—the rich, puzzling, and remarkable theory of the subatomic realm. But his most important paper presented the *special theory of relativity*, a reexamination of space, time, and motion that added a whole new level of sophistication to our understanding of those concepts. The famed equation $E = mc^2$ was actually a relatively minor part of this theory, added in a later paper.

In 1916, Einstein published his *general theory of relativity*, which was, among other things, a fundamentally new description of gravity (see Black Holes and Curved Spacetime). When this theory was confirmed by measurements of the "bending of starlight" during a 1919 eclipse (*The New York Times* headline read, "Lights All Askew in the Heavens"), Einstein became world famous.

In 1933, to escape Nazi persecution, Einstein left his professorship in Berlin and settled in the United States at the newly created Institute for Advanced Studies at Princeton. He remained there until his death in 1955, writing, lecturing, and espousing a variety of intellectual and political causes. For example, he agreed to sign a letter written by Leo Szilard and other scientists in 1939, alerting President Roosevelt to the dangers of allowing Nazi Germany to develop the atomic bomb first. And in 1952, Einstein was offered the second presidency of Israel. In declining the position, he said, "I know a little about nature and hardly anything about men."



Figure 16.3 : Albert Einstein (1879–1955). This portrait of Einstein was taken in 1912. (credit: modification of work by J. F. Langhans)

Elementary Particles

The fundamental components of atoms are the proton, neutron, and electron (see The Structure of the Atom).

Protons, neutrons, and electrons are by no means all the particles that exist. First, for each kind of particle, there is a corresponding but opposite *antiparticle*. If the particle carries a charge, its antiparticle has the opposite charge. The antielectron is the *positron*, which has the same mass as the electron but is positively charged. Similarly, the antiproton has a negative charge. The remarkable thing about such *antimatter* is that when a particle comes into contact with its antiparticle, the original particles are annihilated, and substantial amounts of energy in the form of photons are produced.

Since our world is made exclusively of ordinary particles of matter, antimatter cannot survive for very long. But individual antiparticles are found in cosmic rays (particles that arrive at the top of Earth's atmosphere from space) and can be created in particle accelerators. And, as we will see in a moment, antimatter is created in the core of the Sun and other stars.

Science fiction fans may be familiar with antimatter from the *Star Trek* television series and films. The Starship Enterprise is propelled by the careful combining of matter and antimatter in the ship's engine room. According to $E = mc^2$, the annihilation of matter and antimatter can produce a huge amount of energy, but keeping the antimatter fuel from touching the ship before it is needed must be a big problem. No wonder Scotty, the chief engineer in the original TV show, always looked worried!

In 1933, physicist Wolfgang Pauli (Figure 16.4) suggested that there might be another type of elementary particle. Energy seemed to disappear when certain types of nuclear reactions took place, violating the law of conservation of energy. Pauli was reluctant to accept the idea that one of the basic laws of physics was wrong, and he suggested a "desperate remedy." Perhaps a so-far-undetected particle, which was given the name neutrino ("little neutral one"), carried away the "missing" energy. He suggested that neutrinos were particles with zero mass, and that like photons, they moved with the speed of light.



Figure 16.4 : Wolfgang Pauli in 1945. Pauli is considered the “father” of the neutrino, having conceived of it in 1933.

The elusive neutrino was not detected until 1956. The reason it was so hard to find is that neutrinos interact very weakly with other matter and therefore are very difficult to detect. Earth is more transparent to a neutrino than the thinnest and cleanest pane of glass is to a photon of light. In fact, most neutrinos can pass completely through a star or planet without being absorbed. As we shall see, this behavior of neutrinos makes them a very important tool for studying the Sun. Since Pauli’s prediction, scientists have learned a lot more about the neutrino. We now know that there are three different types of neutrinos, and in 1998, neutrinos were discovered to have a tiny amount of mass. Indeed, it is so tiny that electrons are at least 500,000 times more massive. Ongoing research is focused on determining the mass of neutrinos more precisely, and it may still turn out that one of the three types is massless. We will return to the subject of neutrinos later in this chapter.

Some of the properties of the proton, electron, neutron, and neutrino are summarized in Table 16.1. (Other subatomic particles have been produced by experiments with particle accelerators, but they do not play a role in the generation of solar energy.)

Table 16.1: Properties of Some Common Particles

Particle	Mass (kg)	Charge
Proton	1.67265×10^{-27}	+1
Neutron	1.67495×10^{-27}	0
Electron	9.11×10^{-31}	–1
Neutrino	$< 2 \times 10^{-36}$ (uncertain)	0

The Atomic Nucleus

The nucleus of an atom is not just a loose collection of elementary particles. Inside the nucleus, particles are held together by a very powerful force called the strong nuclear force. This is short-range force, only capable of acting over distances about the size of the atomic nucleus. A quick thought experiment shows how important this force is. Take a look at your finger and consider the atoms composing it. Among them is carbon, one of the basic elements of life. Focus your imagination on the nucleus of one of your carbon atoms. It contains six protons, which have a positive charge, and six neutrons, which are neutral. Thus, the nucleus has a net charge of six positives. If only the electrical force were acting, the protons in this and every carbon atom would find each other very repulsive and fly apart.

The strong nuclear force is an attractive force, stronger than the electrical force, and it keeps the particles of the nucleus tightly bound together. We saw earlier that if under the force of gravity a star “shrinks”—bringing its atoms closer together—gravitational energy is released. In the same way, if particles come together under the strong nuclear force and unite to form an atomic nucleus, some of the nuclear energy is released. The energy given up in such a process is called the *binding energy* of the nucleus.

When such binding energy is released, the resulting nucleus has slightly less mass than the sum of the masses of the particles that came together to form it. In other words, the energy comes from the loss of mass. This slight deficit in mass is only a small fraction of the mass of one proton. But because each bit of lost mass can provide a lot of energy (remember, $E = mc^2$), this nuclear energy release can be quite substantial.

Measurements show that the binding energy is greatest for atoms with a mass near that of the iron nucleus (with a combined number of protons and neutrons equal to 56) and less for both the lighter and the heavier nuclei. Iron, therefore, is the most stable element: since it gives up the most energy when it forms, it would require the most energy to break it back down into its component particles.

What this means is that, in general, when light atomic nuclei come together to form a heavier one (up to iron), mass is lost and energy is released. This joining together of atomic nuclei is called nuclear fusion.

Energy can also be produced by breaking up heavy atomic nuclei into lighter ones (down to iron); this process is called nuclear fission. Nuclear fission was the process we learned to use first—in atomic bombs and in nuclear reactors used to generate electrical power—and it may therefore be more familiar to you. Fission also sometimes occurs spontaneously in some unstable nuclei through the process of natural radioactivity. But fission requires big, complex nuclei, whereas we know that the stars are made up predominantly of small, simple nuclei. So we must look to fusion first to explain the energy of the Sun and the stars (Figure 16.5).

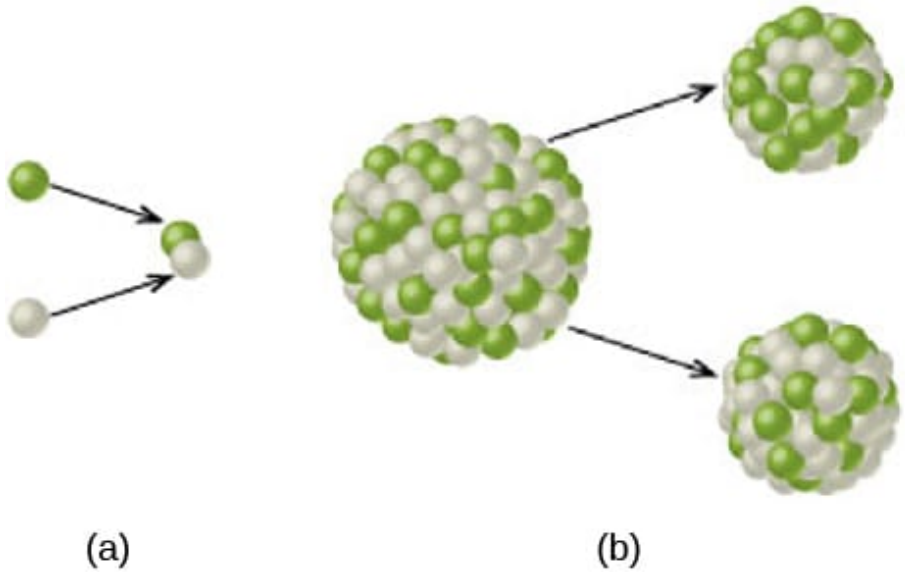


Figure 16.5 : Fusion and Fission. (a) In fusion, light atomic nuclei join together to form a heavier nuclei, releasing energy in the process. (b) In fission, energy is produced by the breaking up of heavy, complex nuclei into lighter ones.

Nuclear Attraction versus Electrical Repulsion

So far, we seem to have a very attractive prescription for producing the energy emitted by the Sun: “roll” some nuclei together and join them via nuclear fusion. This will cause them to lose some of their mass, which then turns into energy. However, every nucleus beyond simple hydrogen has two or more protons—and protons all have positive charges. Since like charges repel via the electrical force, the closer we get two nuclei to each other, the more they repel. It’s true that if we can get them within “striking distance” of the nuclear force, they will then come together with a much stronger attraction. But that striking distance is very tiny, about the size of a nucleus. How can we get nuclei close enough to participate in fusion?

The answer turns out to be heat—tremendous heat—which speeds the protons up enough to overcome the electrical forces that try to keep protons apart. Inside the Sun, as we saw, the most common element is hydrogen, whose nucleus contains only a single proton. Two protons can fuse only in regions where the temperature is greater than about 12 million K, and the speed of the protons average around 1000 kilometers per second or more. (In old-fashioned units, that’s over 2 million miles per hour!)

In our Sun, such extreme temperatures are reached only in the regions near its center, which has a temperature of 15 million K. Calculations show that nearly all of the Sun’s energy is generated within about 150,000 kilometers of its core, or within less than

10% of its total volume.

Even at these high temperatures, it is exceedingly difficult to force two protons to combine. On average, a proton will rebound from other protons in the Sun's crowded core for about 14 billion years, at the rate of *100 million collisions per second*, before it fuses with a second proton. This is, however, only the *average* waiting time. Some of the enormous numbers of protons in the Sun's inner region are "lucky" and take only a few collisions to achieve a fusion reaction: they are the protons responsible for producing the energy radiated by the Sun. Since the Sun is about 4.5 billion years old, most of its protons have not yet been involved in fusion reactions.

Nuclear Reactions in the Sun's Interior

The Sun, then, taps the energy contained in the nuclei of atoms through nuclear fusion. Let's look at what happens in more detail. Deep inside the Sun, a three-step process takes four hydrogen nuclei and fuses them together to form a single helium nucleus. The helium nucleus is slightly less massive than the four hydrogen nuclei that combine to form it, and that mass is converted into energy.

The initial step required to form one helium nucleus from four hydrogen nuclei is shown in Figure 16.6. At the high temperatures inside the Sun's core, two protons combine to make a *deuterium* nucleus, which is an isotope (or version) of hydrogen that contains one proton and one neutron. In effect, one of the original protons has been converted into a neutron in the fusion reaction. Electric charge has to be conserved in nuclear reactions, and it is conserved in this one. A positron (antimatter electron) emerges from the reaction and carries away the positive charge originally associated with one of the protons.

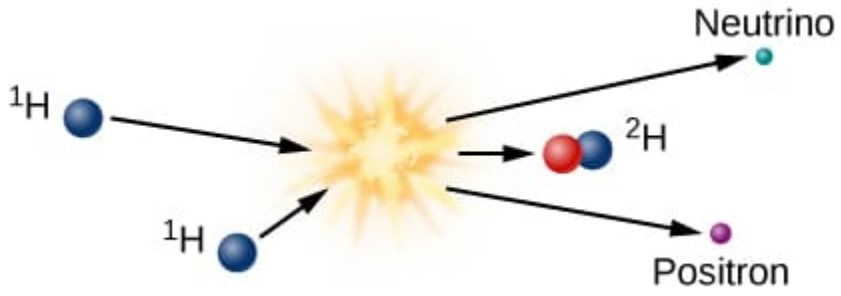


Figure 16.6 : Proton-Proton Chain, Step 1. This is the first step in the process of fusing hydrogen into helium in the Sun. High temperatures are required because this reaction starts with two hydrogen nuclei, which are protons (shown in blue at left) that must overcome electrical repulsion to combine, forming a hydrogen nucleus with a proton and a neutron (shown in red). Note that hydrogen containing one proton and one neutron is given its own name: deuterium. Also produced in this reaction are a positron, which is an antielectron, and an elusive particle named the neutrino.

Since it is antimatter, this positron will instantly collide with a nearby electron, and both will be annihilated, producing electromagnetic energy in the form of gamma-ray photons. This gamma ray, which has been created in the center of the Sun, finds itself in a world crammed full of fast-moving nuclei and electrons. The gamma ray collides with particles of matter and transfers its energy to one of them. The particle later emits another gamma-ray photon, but often the emitted photon has a bit less energy than the one that was absorbed.

Such interactions happen to gamma rays again and again and again as they make their way slowly toward the outer layers of the Sun, until their energy becomes so reduced that they are no longer gamma rays but X-rays (recall what you learned in The Electromagnetic Spectrum). Later, as the photons lose still more energy through collisions in the crowded center of the Sun, they become ultraviolet photons.

By the time they reach the Sun's surface, most of the photons have given up enough energy to be ordinary light—and they are the sunlight we see coming from our star. (To be precise, each gamma-ray photon is ultimately converted into many separate lower-energy photons of sunlight.) So, the sunlight given off by the Sun today had its origin as a gamma ray produced by nuclear reactions deep in the Sun's core. The length of time that photons require to reach the surface depends on how far a photon on average travels between collisions, and the travel time depends on what model of the complicated solar interior we accept. Estimates are somewhat uncertain but indicate that the emission of energy from the surface of the Sun can lag its production in the interior by 100,000 years to as much as 1,000,000 years.

In addition to the positron, the fusion of two hydrogen atoms to form deuterium results in the emission of a neutrino. Because

neutrinos interact so little with ordinary matter, those produced by fusion reactions near the center of the Sun travel directly to the Sun's surface and then out into space, in all directions. Neutrinos move at nearly the speed of light, and they escape the Sun about two seconds after they are created.

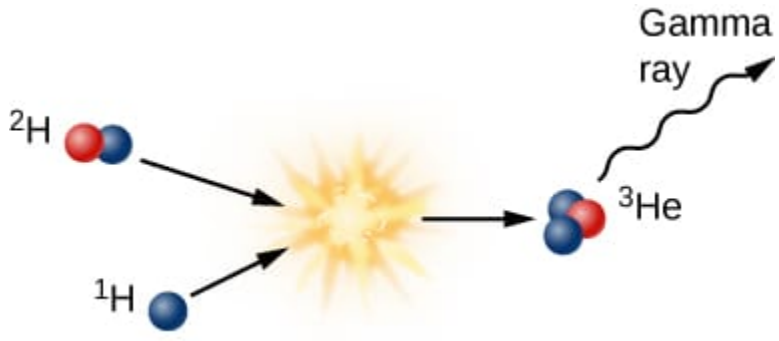


Figure 16.7 : Proton-Proton Chain, Step 2. This is the second step of the proton-proton chain, the fusion reaction that converts hydrogen into helium in the Sun. This step combines one hydrogen nucleus, which is a proton (shown in blue), with the deuterium nucleus from the previous step (shown as a red and blue particle). The product of this is an isotope of helium with two protons (blue) and one neutron (red) and energy in the form of gamma-ray radiation.

The second step in forming helium from hydrogen is to add another proton to the deuterium nucleus to create a helium nucleus that contains two protons and one neutron (Figure 16.7). In the process, some mass is again lost and more gamma radiation is emitted. Such a nucleus is helium because an element is defined by its number of protons; any nucleus with two protons is called helium. But this form of helium, which we call helium-3 (and write in shorthand as ^{3}He) is not the isotope we see in the Sun's atmosphere or on Earth. That helium has two neutrons and two protons and hence is called helium-4 (^{4}He).

To get to helium-4 in the Sun, helium-3 must combine with another helium-3 in the third step of fusion (illustrated in Figure 16.8). Note that two energetic protons are left over from this step; each of them comes out of the reaction ready to collide with other protons and to start step 1 in the chain of reactions all over again.

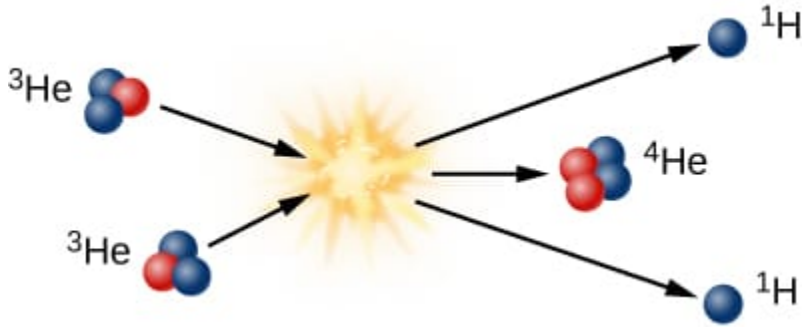


Figure 16.8 : Proton-Proton Chain, Step 3. This is the third step in the fusion of hydrogen into helium in the Sun. Note that the two helium-3 nuclei from the second step (see Figure 16.7) must combine before the third step becomes possible. The two protons that come out of this step have the energy to collide with other protons in the Sun and start step one again.

Link to Learning

Visit the Tokamak Fusion Reactor at the General Atomics Lab in San Diego, CA, for an 8-minute tour.

The Proton-Proton Chain

The nuclear reactions in the Sun that we have been discussing can be described succinctly through the following nuclear formulas:

$$^1\mathrm{H} + {}^1\mathrm{H} \longrightarrow {}^2\mathrm{H} + \mathrm{e}^+ + v$$
$$^2\mathrm{H} + {}^1\mathrm{H} \longrightarrow {}^3\mathrm{He} + \gamma$$
$$^3\mathrm{He} + {}^3\mathrm{He} \longrightarrow {}^4\mathrm{He} + {}^1\mathrm{H} + {}^1\mathrm{H}$$

Here, the superscripts indicate the total number of neutrons plus protons in the nucleus, e^+ is the positron, *v* is the neutrino, and *y* indicates that gamma rays are emitted. Note that the third step requires two helium-3 nuclei to start; the first two steps must happen twice before the third step can occur.

Although, as we discussed, the first step in this chain of reactions is very difficult and generally takes a long time, the other steps happen more quickly. After the deuterium nucleus is formed, it survives an average of only about 6 seconds before being converted into ^{3}He. About a million years after that (on average), the ^{3}He nucleus will combine with another to form ^{4}He.

We can compute the amount of energy these reactions generate by calculating the difference in the initial and final masses. The masses of hydrogen and helium atoms in the units normally used by scientists are 1.007825 *u* and 4.00268 *u*, respectively. (The unit of mass, *u*, is defined to be 1/12 the mass of an atom of carbon, or approximately the mass of a proton.) Here, we include the mass of the entire atom, not just the nucleus, because electrons are involved as well. When hydrogen is converted into helium, two positrons are created (remember, the first step happens twice), and these are annihilated with two free electrons, adding to the energy produced.

$$\begin{aligned} 4 \times 1.007825u &= 4.03130\,u\ (\text{mass of initial hydrogen atoms}) \\ &\quad - 4.00268\,u\ (\text{mass of final helium atoms}) \\ &= 0.02862\,u\ (\text{mass lost in the transformation}) \end{aligned}$$

The mass lost, 0.02862 *u*, is 0.71% of the mass of the initial hydrogen. Thus, if 1 kilogram of hydrogen is converted into helium, then the mass of the helium is only 0.9929 kilogram, and 0.0071 kilogram of material is converted into energy. The speed of light (*c*) is 3 × 10^8 meters per second, so the energy released by the conversion of just 1 kilogram of hydrogen into helium is:

$$E = mc^2$$
$$E = 0.0071\ \text{kg} \times \left(3 \times 10^8 \text{m/s}\right)^2 = 6.4 \times 10^{14}\ \text{J}$$

This amount, the energy released when a single kilogram (2.2 pounds) of hydrogen undergoes fusion, would supply all the electricity used by the typical U.S. household for roughly 17,000 years.

To produce the Sun's luminosity of 4 × 10^{26} watts, some 600 million tons of hydrogen must be converted into helium *each second*, of which about 4 million tons are converted from matter into energy. As large as these numbers are, the store of hydrogen (and thus of nuclear energy) in the Sun is still *more* enormous, and can last a long time—billions of years, in fact.

At the temperatures inside the stars with masses smaller than about 1.2 times the mass of our Sun (a category that includes the Sun itself), most of the energy is produced by the reactions we have just described, and this set of reactions is called the proton-proton chain (or sometimes, the p-p chain). In the proton-proton chain, protons collide directly with other protons to form helium nuclei.

In hotter stars, another set of reactions, called the carbon-nitrogen-oxygen (CNO) cycle, accomplishes the same net result. In the CNO cycle, carbon and hydrogen nuclei collide to initiate a series of reactions that form nitrogen, oxygen, and ultimately, helium. The nitrogen and oxygen nuclei do not survive but interact to form carbon again. Therefore, the outcome is the same as in the proton-proton chain: four hydrogen atoms disappear, and in their place, a single helium atom is created. The CNO cycle plays only a minor role in the Sun but is the main source of energy for stars with masses greater than about the mass of the Sun.

So you can see that we have solved the puzzle that so worried scientists at the end of the nineteenth century. The Sun can maintain its high temperature and energy output for billions of years through the fusion of the simplest element in the universe, hydrogen. Because most of the Sun (and the other stars) is made of hydrogen, it is an ideal "fuel" for powering a star. As will be discussed in the following chapters, we can define a star as a ball of gas capable of getting its core hot enough to initiate the fusion of hydrogen. There are balls of gas that lack the mass required to do this (Jupiter is a local example); like so many hopefuls in Hollywood, they will never be stars.

Making Connections: Fusion on Earth

Wouldn't it be wonderful if we could duplicate the Sun's energy mechanism in a controlled way on Earth? (We have already duplicated it in an uncontrolled way in hydrogen bombs, but we hope our storehouses of these will never be used.) Fusion energy would have many advantages: it would use hydrogen (or deuterium, which is heavy hydrogen) as fuel, and there is

abundant hydrogen in Earth's lakes and oceans. Water is much more evenly distributed around the world than oil or uranium, meaning that a few countries would no longer hold an energy advantage over the others. And unlike fission, which leaves dangerous byproducts, the nuclei that result from fusion are perfectly safe.

The problem is that, as we saw, it takes extremely high temperatures for nuclei to overcome their electrical repulsion and undergo fusion. When the first hydrogen bombs were exploded in tests in the 1950s, the "fuses" to get them hot enough were fission bombs. Interactions at such temperatures are difficult to sustain and control. To make fusion power on Earth, after all, we have to do what the Sun does: produce temperatures and pressures high enough to get hydrogen nuclei on intimate terms with one another.

The European Union, the United States, South Korea, Japan, China, Russia, Switzerland, and India are collaborating on the International Thermonuclear Experimental Reactor (ITER), a project to demonstrate the feasibility of controlled fusion (Figure 16.9). The facility is being built in France. Construction will require over 10,000,000 components and 2000 workers for assembly. The date for the start of operations is yet to be determined.

ITER is based on the Tokamak design, in which a large doughnut-shaped container is surrounded by superconducting magnets to confine and control the hydrogen nuclei in a strong magnetic field. Previous fusion experiments have produced about 15 million watts of energy, but only for a second or two, and they have required 100 million watts to produce the conditions necessary to achieve fusion. The goal of ITER is to build the first fusion device capable of producing 500 million watts of fusion energy for up to 1000 seconds. The challenge is keeping the deuterium and tritium—which will participate in fusion reactions—hot enough and dense enough, for a long enough time to produce energy.

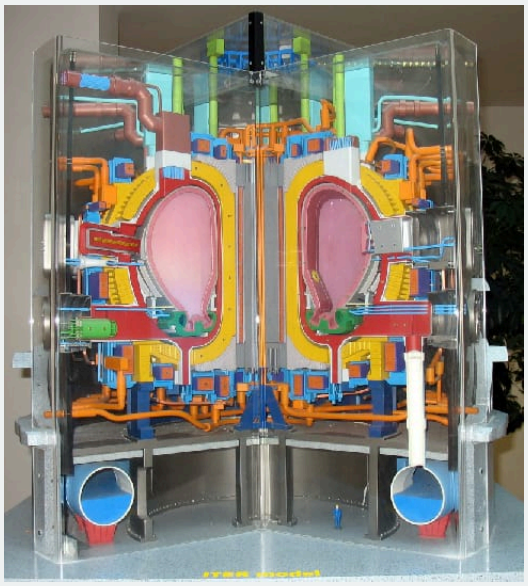


Figure 16.9 : ITER Design. The bright yellow areas in this model show where the superconducting magnets will circle the chamber within which fusion will take place. A huge magnet will keep the charged nuclei of heavy hydrogen confined. The goal is to produce 500 megawatts of energy. (credit: modification of work by Stephan Mosel)

This page titled 19.4: Mass, Energy, and the Theory of Relativity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 16.3: Mass, Energy, and the Theory of Relativity by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.5: The Solar Interior - Theory

Learning Objectives

By the end of this section, you will be able to:

- Describe the state of equilibrium of the Sun
- Understand the energy balance of the Sun
- Explain how energy moves outward through the Sun
- Describe the structure of the solar interior

Fusion of protons can occur in the center of the Sun only if the temperature exceeds 12 million K. How do we know that the Sun is actually this hot? To determine what the interior of the Sun might be like, it is necessary to resort to complex calculations. Since we can't see the interior of the Sun, we have to use our understanding of physics, combined with what we see at the surface, to construct a mathematical model of what must be happening in the interior. Astronomers use observations to build a computer program containing everything they think they know about the physical processes going on in the Sun's interior. The computer then calculates the temperature and pressure at every point inside the Sun and determines what nuclear reactions, if any, are taking place. For some calculations, we can use observations to determine whether the computer program is producing results that match what we see. In this way, the program evolves with ever-improving observations.

The computer program can also calculate how the Sun will change with time. After all, the Sun must change. In its center, the Sun is slowly depleting its supply of hydrogen and creating helium instead. Will the Sun get hotter? Cooler? Larger? Smaller? Brighter? Fainter? Ultimately, the changes in the center could be catastrophic, since eventually all the hydrogen fuel hot enough for fusion will be exhausted. Either a new source of energy must be found, or the Sun will cease to shine. We will describe the ultimate fate of the Sun in later chapters. For now, let's look at some of the things we must teach the computer about the Sun in order to carry out such calculations.

The Sun Is a Plasma

The Sun is so hot that all of the material in it is in the form of an ionized gas, called a plasma. Plasma acts much like a hot gas, which is easier to describe mathematically than either liquids or solids. The particles that constitute a gas are in rapid motion, frequently colliding with one another. This constant bombardment is the *pressure* of the gas (Figure 16.10).

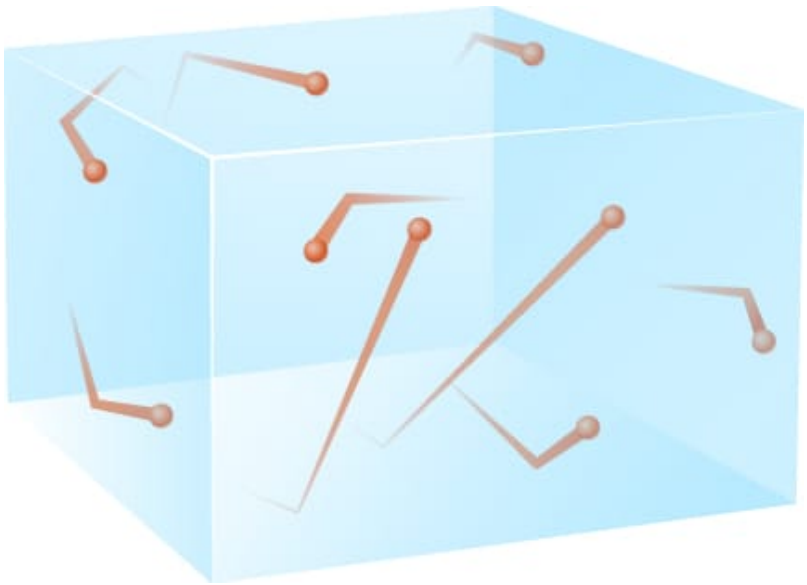


Figure 16.10 : Gas Pressure. The particles in a gas are in rapid motion and produce pressure through collisions with the surrounding material. Here, particles are shown bombarding the sides of an imaginary container.

More particles within a given volume of gas produce more pressure because the combined impact of the moving particles increases with their number. The pressure is also greater when the molecules or atoms are moving faster. Since the molecules move faster when the temperature is hotter, higher temperatures produce higher pressure.

The Sun Is Stable

The Sun, like the majority of other stars, is stable; it is neither expanding nor contracting. Such a star is said to be in a condition of *equilibrium*. All the forces within it are balanced, so that at each point within the star, the temperature, pressure, density, and so on are maintained at constant values. We will see in later chapters that even these stable stars, including the Sun, are changing as they evolve, but such evolutionary changes are so gradual that, for all intents and purposes, the stars are still in a state of equilibrium at any given time.

The mutual gravitational attraction between the masses of various regions within the Sun produces tremendous forces that tend to collapse the Sun toward its center. Yet we know from the history of Earth that the Sun has been emitting roughly the same amount of energy for billions of years, so clearly it has managed to resist collapse for a very long time. The gravitational forces must therefore be counterbalanced by some other force. That force is due to the pressure of gases within the Sun (Figure 16.11). Calculations show that, in order to exert enough pressure to prevent the Sun from collapsing due to the force of gravity, the gases at its center must be maintained at a temperature of 15 million K. Think about what this tells us. Just from the fact that the Sun is not contracting, we can conclude that its temperature must indeed be high enough at the center for protons to undergo fusion.

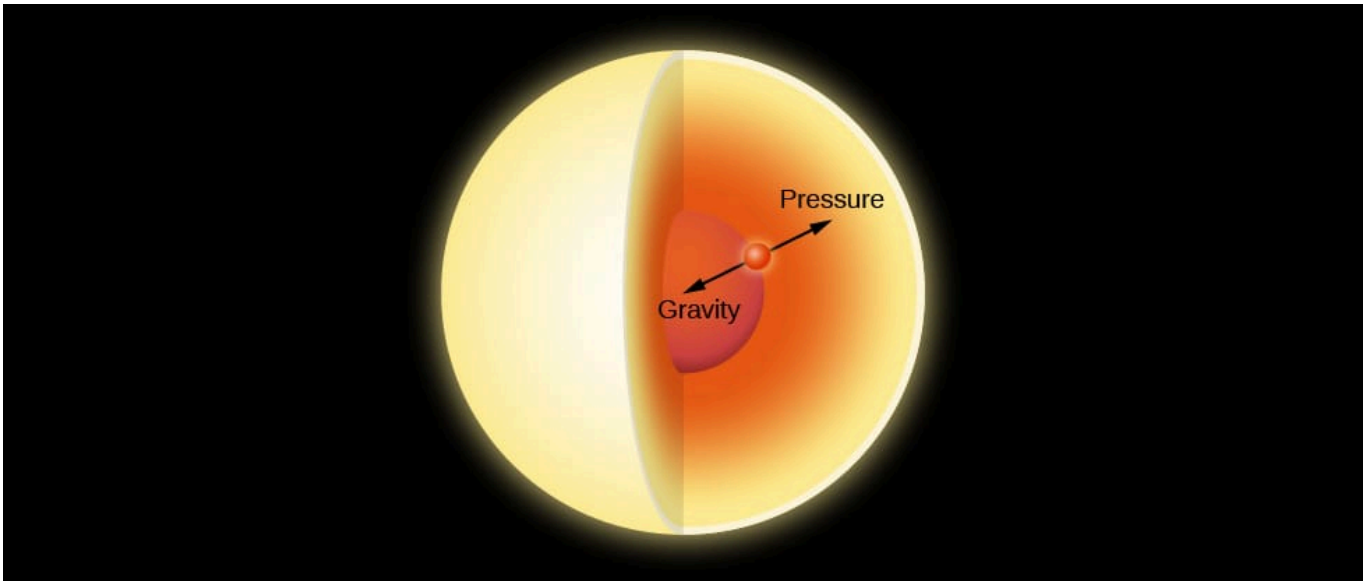


Figure 16.11 : Hydrostatic Equilibrium. In the interior of a star, the inward force of gravity is exactly balanced at each point by the outward force of gas pressure.

The Sun maintains its stability in the following way. If the internal pressure in such a star were not great enough to balance the weight of its outer parts, the star would collapse somewhat, contracting and building up the pressure inside. On the other hand, if the pressure were greater than the weight of the overlying layers, the star would expand, thus decreasing the internal pressure. Expansion would stop, and equilibrium would again be reached when the pressure at every internal point equaled the weight of the stellar layers above that point. An analogy is an inflated balloon, which will expand or contract until an equilibrium is reached between the pressure of the air inside and outside. The technical term for this condition is hydrostatic equilibrium. Stable stars are all in hydrostatic equilibrium; so are the oceans of Earth as well as Earth's atmosphere. The air's own pressure keeps it from falling to the ground.

The Sun Is Not Cooling Down

As everyone who has ever left a window open on a cold winter night knows, heat always flows from hotter to cooler regions. As energy filters outward toward the surface of a star, it must be flowing from inner, hotter regions. The temperature cannot ordinarily get cooler as we go inward in a star, or energy would flow in and heat up those regions until they were at least as hot as the outer ones. Scientists conclude that the temperature is highest at the center of a star, dropping to lower and lower values toward

the stellar surface. (The high temperature of the Sun's chromosphere and corona may therefore appear to be a paradox. But remember from The Sun: A Garden-Variety Star that these high temperatures are maintained by magnetic effects, which occur in the Sun's atmosphere.)

The outward flow of energy through a star robs it of its internal heat, and the star would cool down if that energy were not replaced. Similarly, a hot iron begins to cool as soon as it is unplugged from its source of electric energy. Therefore, a source of fresh energy must exist within each star. In the Sun's case, we have seen that this energy source is the ongoing fusion of hydrogen to form helium.

Heat Transfer in a Star

Since the nuclear reactions that generate the Sun's energy occur deep within it, the energy must be transported from the center of the Sun to its surface—where we see it in the form of both heat and light. There are three ways in which energy can be transferred from one place to another. In conduction, atoms or molecules pass on their energy by colliding with others nearby. This happens, for example, when the handle of a metal spoon heats up as you stir a cup of hot coffee. In convection, currents of warm material rise, carrying their energy with them to cooler layers. A good example is hot air rising from a fireplace. In radiation, energetic photons move away from hot material and are absorbed by some material to which they convey some or all of their energy. You can feel this when you put your hand close to the coils of an electric heater, allowing infrared photons to heat up your hand. Conduction and convection are both important in the interiors of planets. In stars, which are much more transparent, radiation and convection are important, whereas conduction can usually be ignored.

Stellar *convection* occurs as currents of hot gas flow up and down through the star (Figure 16.12). Such currents travel at moderate speeds and do not upset the overall stability of the star. They don't even result in a net transfer of mass either inward or outward because, as hot material rises, cool material falls and replaces it. This results in a convective circulation of rising and falling cells as seen in Figure 16.12. In much the same way, heat from a fireplace can stir up air currents in a room, some rising and some falling, without driving any air into or out the room. Convection currents carry heat very efficiently outward through a star. In the Sun, convection turns out to be important in the central regions and near the surface.

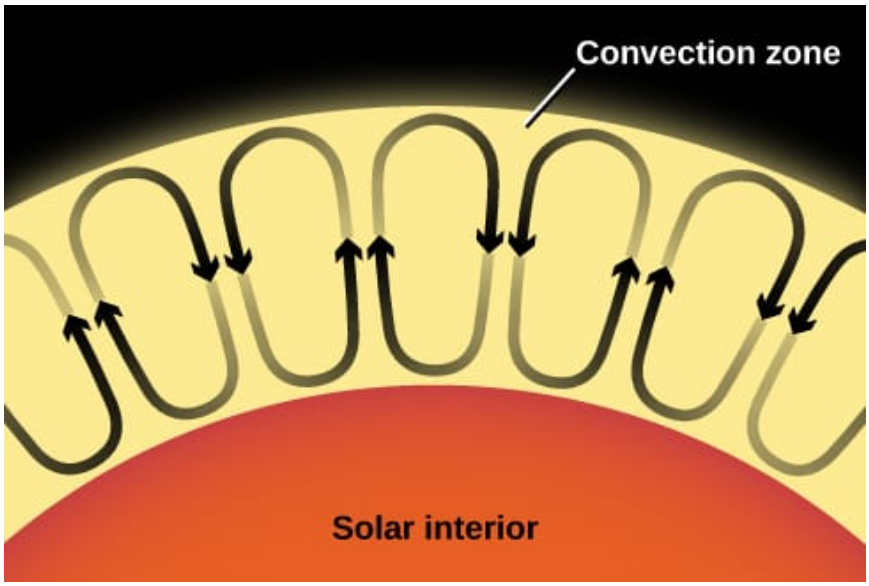


Figure 16.12 : Convection. Rising convection currents carry heat from the Sun's interior to its surface, whereas cooler material sinks downward. Of course, nothing in a real star is as simple as diagrams in textbooks suggest.

Unless convection occurs, the only significant mode of energy transport through a star is by electromagnetic radiation. Radiation is not an efficient means of energy transport in stars because gases in stellar interiors are very opaque, that is, a photon does not go far (in the Sun, typically about 0.01 meter) before it is absorbed. (The processes by which atoms and ions can interrupt the outward flow of photons—such as becoming ionized—were discussed in the section on the Formation of Spectral Lines.) The absorbed energy is always reemitted, but it can be reemitted in any direction. A photon absorbed when traveling outward in a star has almost as good a chance of being radiated back toward the center of the star as toward its surface.

A particular quantity of energy, therefore, zigzags around in an almost random manner and takes a long time to work its way from

the center of a star to its surface (Figure 16.13). Estimates are somewhat uncertain, but in the Sun, as we saw, the time required is probably between 100,000 and 1,000,000 years. If the photons were not absorbed and reemitted along the way, they would travel at the speed of light and could reach the surface in a little over 2 seconds, just as neutrinos do (Figure 16.14).

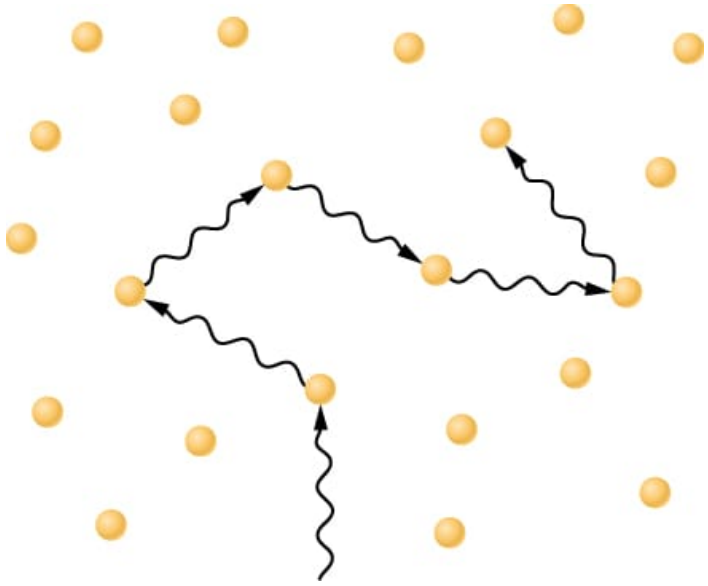


Figure 16.13 : Photons Deep in the Sun. A photon moving through the dense gases in the solar interior travels only a short distance before it interacts with one of the surrounding atoms. The resulting photon usually has a lower energy after each interaction and may then travel in any random direction.

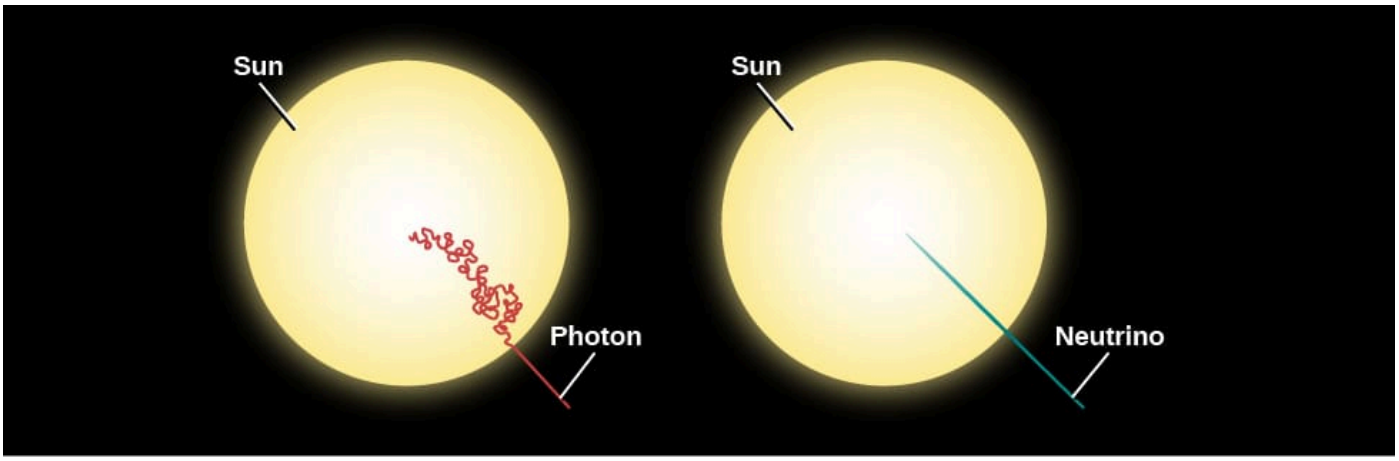


Figure 16.14 : Photon and Neutrino Paths in the Sun. (a) Because photons generated by fusion reactions in the solar interior travel only a short distance before being absorbed or scattered by atoms and sent off in random directions, estimates are that it takes between 100,000 and 1,000,000 years for energy to make its way from the center of the Sun to its surface. (b) In contrast, neutrinos do not interact with matter but traverse straight through the Sun at the speed of light, reaching the surface in only a little more than 2 seconds.

Making Connections: Heat Transfer and Cooking

The three ways that heat energy moves from higher-temperature regions to cooler regions are all used in cooking, and this is important to all of us who enjoy making or eating food.

Conduction is heat transfer by physical contact during which the energetic motion of particles in one region spread to other regions and even to adjacent objects in close contact. A tasty example of this is cooking a steak on a hot iron skillet. When a flame makes the bottom of a skillet hot, the particles in it vibrate actively and collide with neighboring particles, spreading the heat energy throughout the skillet (the ability to spread heat uniformly is a key criterion for selecting materials for cookware).

A steak sitting on the surface of the skillet picks up heat energy by the particles in the surface of the skillet colliding with particles on the surface of the steak. Many cooks will put a little oil on the pan, and this layer of oil, besides preventing sticking, increases heat transfer by filling in gaps and increasing the contact surface area.

Convection is heat transfer by the motion of matter that rises because it is hot and less dense. Heating a fluid makes it expand, which makes it less dense, so it rises. An oven is a great example of this: the fire is at the bottom of the oven and heats the air down there, causing it to expand (becoming less dense), so it rises up to where the food is. The rising hot air carries the heat from the fire to the food by convection. This is how conventional ovens work. You may also be familiar with convection ovens that use a fan to circulate hot air for more even cooking. A scientist would object to that name because normal non-fan ovens that rely on hot air rising to circulate the heat are convection ovens; technically, the ovens that use fans to help move heat are “advection” ovens. (You may not have heard about this because the scientists who complain loudly about misusing the terms convection and advection don’t get out much.)

Radiation is the transfer of heat energy by electromagnetic radiation. Although microwave ovens are an obvious example of using radiation to heat food, a simpler example is a toy oven. Toy ovens are powered by a very bright light bulb. The child-chefs prepare a mix for brownies or cookies, put it into a tray, and place it in the toy oven under the bright light bulb. The light and heat from the bulb hit the brownie mix and cook it. If you have ever put your hand near a bright light, you have undoubtedly noticed your hand getting warmed by the light.

Model Stars

Scientists use the principles we have just described to calculate what the Sun’s interior is like. These physical ideas are expressed as mathematical equations that are solved to determine the values of temperature, pressure, density, the efficiency with which photons are absorbed, and other physical quantities throughout the Sun. The solutions obtained, based on a specific set of physical assumptions, provide a theoretical model for the interior of the Sun.

Figure 16.15 schematically illustrates the predictions of a theoretical model for the Sun’s interior. Energy is generated through fusion in the core of the Sun, which extends only about one-quarter of the way to the surface but contains about one-third of the total mass of the Sun. At the center, the temperature reaches a maximum of approximately 15 million K, and the density is nearly 150 times that of water. The energy generated in the core is transported toward the surface by radiation until it reaches a point about 70% of the distance from the center to the surface. At this point, convection begins, and energy is transported the rest of the way, primarily by rising columns of hot gas.

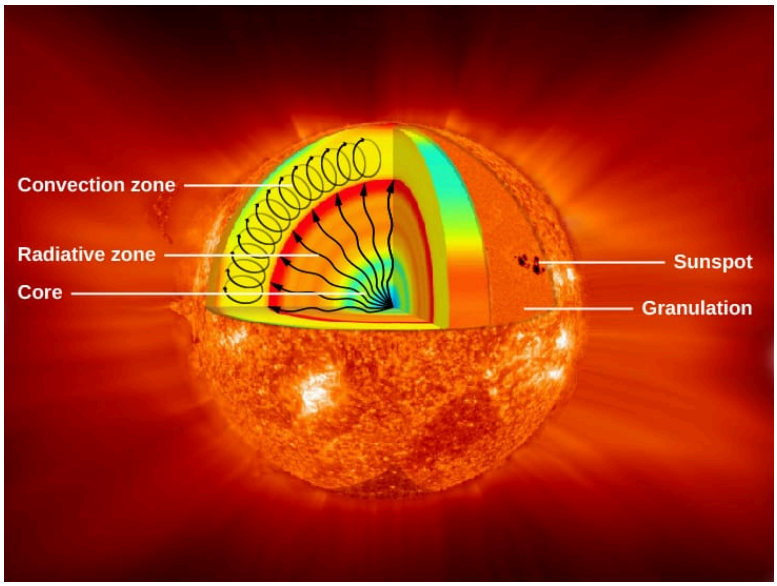


Figure 16.15 : Interior Structure of the Sun. Energy is generated in the core by the fusion of hydrogen to form helium. This energy is transmitted outward by radiation—that is, by the absorption and reemission of photons. In the outermost layers, energy is transported mainly by convection. (credit: modification of work by NASA/Goddard)

Figure 16.16 shows how the temperature, density, rate of energy generation, and composition vary from the center of the Sun to

its surface.

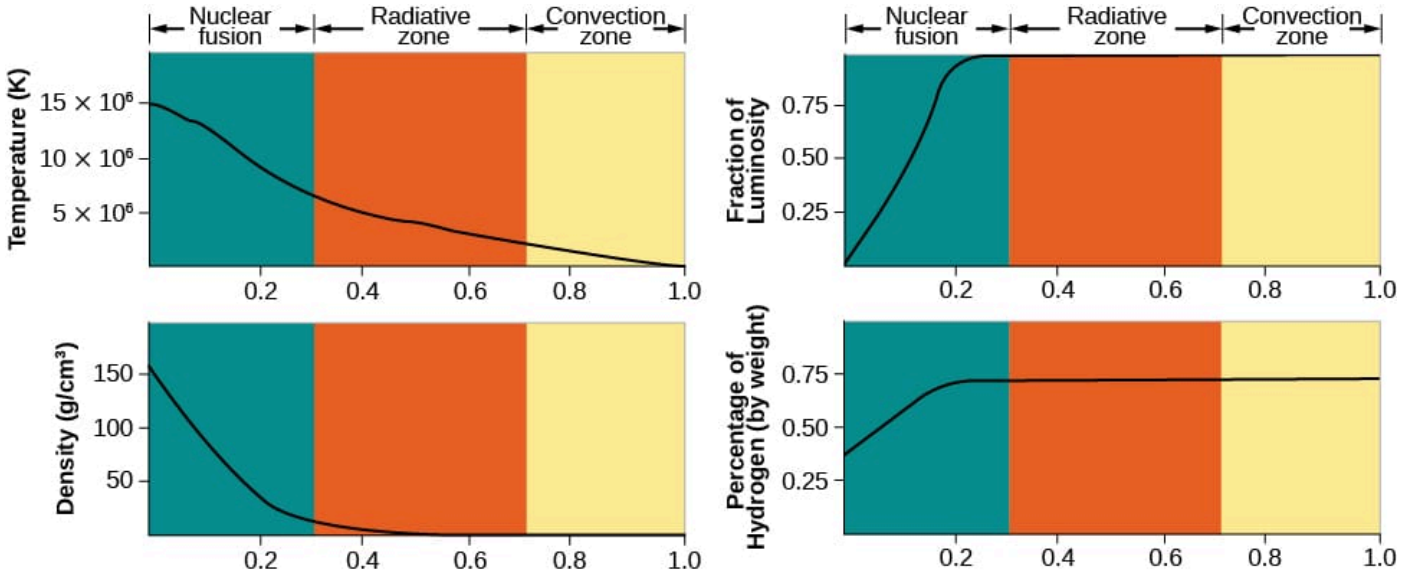


Figure 16.16 : Interior of the Sun. Diagrams showing how temperature, density, rate of energy generation, and the percentage (by mass) abundance of hydrogen vary inside the Sun. The horizontal scale shows the fraction of the Sun's radius: the left edge is the very center, and the right edge is the visible surface of the Sun, which is called the photosphere.

This page titled 19.5: The Solar Interior - Theory is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 16.4: The Solar Interior - Theory by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.6: The Solar Interior - Observations

Learning Objectives

By the end of this section, you will be able to:

- Explain how the Sun pulsates
- Explain what helioseismology is and what it can tell us about the solar interior
- Discuss how studying neutrinos from the Sun has helped understand neutrinos

Recall that when we observe the Sun's photosphere (the surface layer we see from the outside), we are not seeing very deeply into our star, certainly not into the regions where energy is generated. That's why the title of this section—observations of the solar interior—should seem very surprising. However, astronomers have indeed devised two types of measurements that can be used to obtain information about the inner parts of the Sun. One technique involves the analysis of tiny changes in the motion of small regions at the Sun's surface. The other relies on the measurement of the neutrinos emitted by the Sun.

Solar Pulsations

Astronomers discovered that the Sun pulsates—that is, it alternately expands and contracts—just as your chest expands and contracts as you breathe. This pulsation is very slight, but it can be detected by measuring the *radial velocity* of the solar surface—the speed with which it moves toward or away from us. The velocities of small regions on the Sun are observed to change in a regular way, first toward Earth, then away, then toward, and so on. It is as if the Sun were "breathing" through thousands of individual lungs, each having a size in the range of 4000 to 15,000 kilometers, each fluctuating back and forth (Figure 16.17).

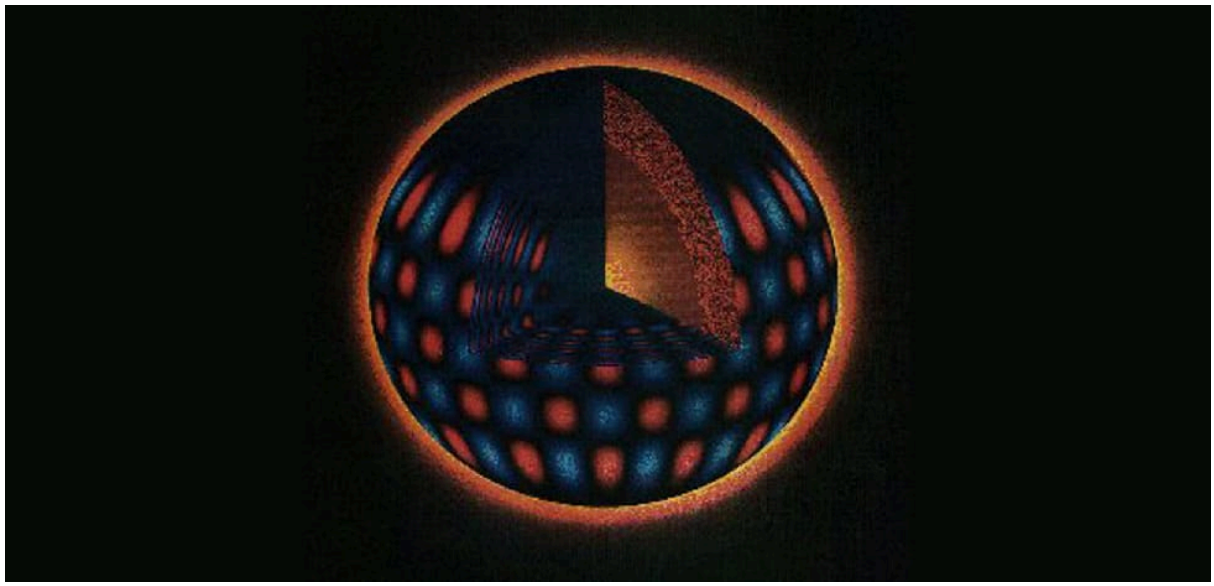


Figure 16.17 : Oscillations in the Sun. New observational techniques permit astronomers to measure small differences in velocity at the Sun's surface to infer what the deep solar interior is like. In this computer simulation, red shows surface regions that are moving away from the observer (inward motion); blue marks regions moving toward the observer (outward motion). Note that the velocity changes penetrate deep into the Sun's interior. (credit: modification of work by GONG, NOAO)

The typical velocity of one of the oscillating regions on the Sun is only a few hundred meters per second, and it takes about 5 minutes to complete a full cycle from maximum to minimum velocity and back again. The change in the size of the Sun measured at any given point is no more than a few kilometers.

The remarkable thing is that these small velocity variations can be used to determine what the interior of the Sun is like. The motion of the Sun's surface is caused by waves that reach it from deep in the interior. Study of the amplitude and cycle

length of velocity changes provides information about the temperature, density, and composition of the layers through which the waves passed before they reached the surface. The situation is somewhat analogous to the use of seismic waves generated by earthquakes to infer the properties of Earth's interior. For this reason, studies of solar oscillations (back-and-forth motions) are referred to as helioseismology.

It takes a little over an hour for waves to traverse the Sun from center to surface, so the waves, like neutrinos, provide information about what the solar interior is like at the present time. In contrast, remember that the sunlight we see today emerging from the Sun was actually generated in the core several hundred thousand years ago.

Helioseismology has shown that convection extends inward from the surface 30% of the way toward the center; we have used this information in drawing Figure 16.15. Pulsation measurements also show that the *differential rotation* that we see at the Sun's surface, with the fastest rotation occurring at the equator, persists down through the convection zone. Below the convection zone, however, the Sun, even though it is gaseous throughout, rotates as if it were a solid body like a bowling ball. Another finding from helioseismology is that the abundance of helium inside the Sun, except in the center where nuclear reactions have converted hydrogen into helium, is about the same as at its surface. That result is important to astronomers because it means we are correct when we use the abundance of the elements measured in the solar atmosphere to construct models of the solar interior.

Helioseismology also allows scientists to look beneath a sunspot and see how it works. In The Sun: A Garden-Variety Star, we said that sunspots are cool because strong magnetic fields block the outward flow of energy. Figure 16.18 shows how gas moves around underneath a sunspot. Cool material from the sunspot flows downward, and material surrounding the sunspot is pulled inward, carrying magnetic field with it and thus maintaining the strong field that is necessary to form a sunspot. As the new material enters the sunspot region, it too cools, becomes denser, and sinks, thus setting up a self-perpetuating cycle that can last for weeks.

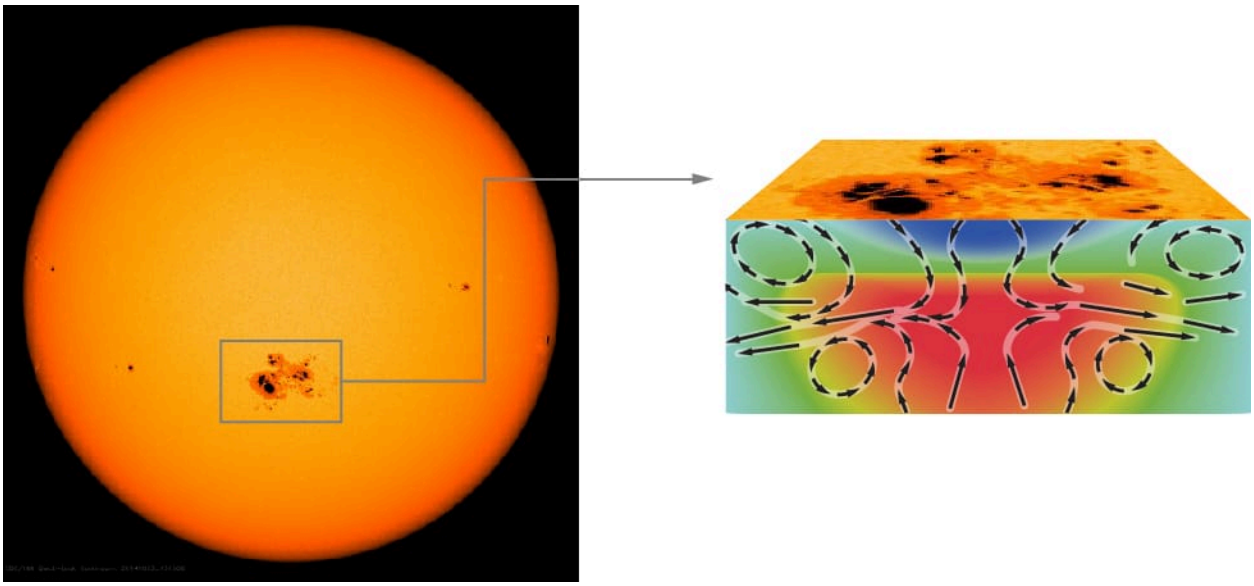


Figure 16.18 : Sunspot Structure. This drawing shows our new understanding, from helioseismology, of what lies beneath a sunspot. The black arrows show the direction of the flow of material. The intense magnetic field associated with the sunspot stops the upward flow of hot material and creates a kind of plug that blocks the hot gas. As the material above the plug cools (shown in blue), it becomes denser and plunges inward, drawing more gas and more magnetic field behind it into the spot. The concentrated magnetic field causes more cooling, thereby setting up a self-perpetuating cycle that allows a spot to survive for several weeks. Since the plug keeps hot material from flowing up into the sunspot, the region below the plug, represented by red in this picture, becomes hotter. This material flows sideways and then upward, eventually reaching the solar surface in the area surrounding the sunspot. (credit: modification of work by NASA, SDO)

The downward-flowing cool material acts as a kind of plug that block the upward flow of hot material, which is then diverted sideways and eventually reaches the solar surface in the region around the sunspot. This outward flow of hot material accounts for the paradox that we described in The Sun: A Garden-Variety Star—namely, that the Sun emits slightly more energy when more of its surface is covered by cool sunspots.

Helioseismology has become an important tool for predicting solar storms that might impact Earth. Active regions can appear and grow large in only a few days. The solar rotation period is about 28 days. Therefore, regions capable of producing solar flares and coronal mass ejections can develop on the far side of the Sun, where, for a long time, we couldn't see them directly.

Fortunately, we now have space telescopes monitoring the Sun from all angles, so we know if there are sunspots forming on the opposite side of the Sun. Moreover, sound waves travel slightly faster in regions of high magnetic field, and waves generated in active regions traverse the Sun about 6 seconds faster than waves generated in quiet regions. By detecting this subtle difference, scientists can provide warnings of a week or more to operators of electric utilities and satellites about when a potentially dangerous active region might rotate into view. With this warning, it is possible to plan for disruptions, put key instruments into safe mode, or reschedule spacewalks in order to protect astronauts.

Solar Neutrinos

The second technique for obtaining information about the Sun's interior involves the detection of a few of those elusive neutrinos created during nuclear fusion. Recall from our earlier discussion that neutrinos created in the center of the Sun make their way directly out of the Sun and travel to Earth at nearly the speed of light. As far as neutrinos are concerned, the Sun is transparent.

About 3% of the total energy generated by nuclear fusion in the Sun is carried away by neutrinos. So many protons react and form neutrinos inside the Sun's core that, scientists calculate, 35 million billion (3.5×10^{16}) solar neutrinos pass through each square meter of Earth's surface every second. If we can devise a way to detect even a few of these solar neutrinos, then we can obtain information directly about what is going on in the center of the Sun. Unfortunately for those trying to "catch" some neutrinos, Earth and everything on it are also nearly transparent to passing neutrinos, just like the Sun.

On very, very rare occasions, however, one of the billions and billions of solar neutrinos will interact with another atom. The first successful detection of solar neutrinos made use of cleaning fluid (C_2Cl_4), which is the least expensive way to get a lot of chlorine atoms together. The nucleus of a chlorine (Cl) atom in the cleaning fluid can be turned into a radioactive argon nucleus by an interaction with a neutrino. Because the argon is radioactive, its presence can be detected. However, since the interaction of a neutrino with chlorine happens so rarely, a huge amount of chlorine is needed.

Raymond Davis, Jr. (Figure 16.19) and his colleagues at Brookhaven National Laboratory, placed a tank containing nearly 400,000 liters of cleaning fluid 1.5 kilometers beneath Earth's surface in a gold mine at Lead, South Dakota. A mine was chosen so that the surrounding material of Earth would keep cosmic rays (high-energy particles from space) from reaching the cleaning fluid and creating false signals. (Cosmic-ray particles are stopped by thick layers of Earth, but neutrinos find them of no significance.) Calculations show that solar neutrinos should produce about one atom of radioactive argon in the tank each day.

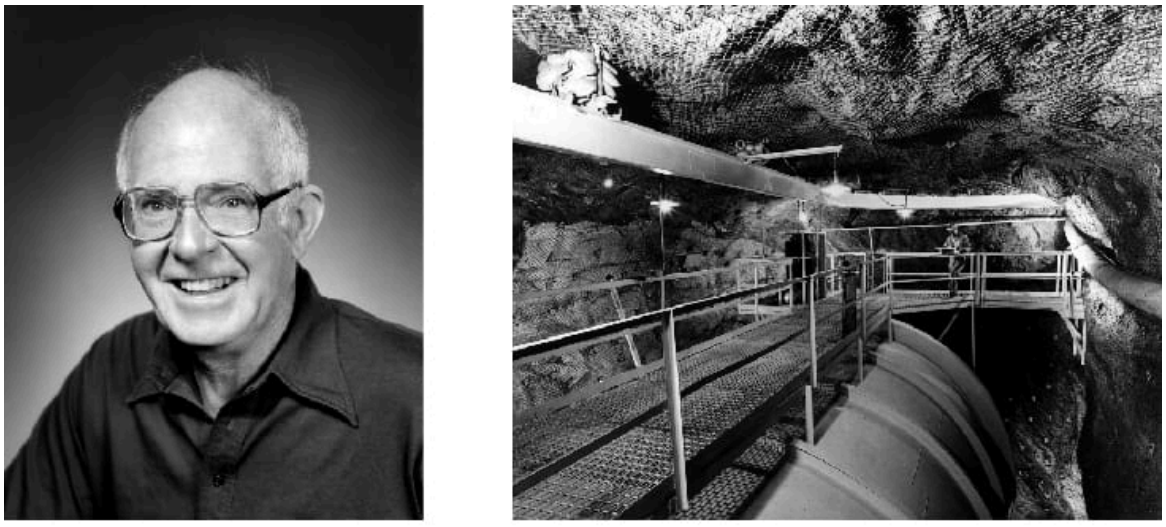


Figure 16.19 : Davis Experiment. (a) Raymond Davis received the Nobel Prize in physics in 2002. (b) Davis' experiment at the bottom of an abandoned gold mine first revealed problems with our understanding of neutrinos. (credit a: modification of work by Brookhaven National Laboratory; credit b: modification of work by the United States Department of Energy)

This was an amazing project: they counted argon atoms about once per month—and remember, they were looking for a tiny handful of argon atoms in a massive tank of chlorine atoms. When all was said and done, Davis' experiment, begun in 1970, detected only about one-third as many neutrinos as predicted by solar models! This was a shocking result because astronomers thought they had a pretty good understanding of both neutrinos and the Sun's interior. For many years, astronomers and physicists

wrestled with Davis' results, trying to find a way out of the dilemma of the "missing" neutrinos.

Eventually Davis' result was explained by the surprising discovery that there are actually three types of neutrinos. Solar fusion produces only one type of neutrino, the so-called electron neutrino, and the initial experiments to detect solar neutrinos were designed to detect this one type. Subsequent experiments showed that these neutrinos change to a different type during their journey from the center of the Sun through space to Earth in a process called *neutrino oscillation*.

An experiment, conducted at the Sudbury Neutrino Observatory in Canada, was the first one designed to capture all three types of neutrinos (Figure 16.20). The experiment was located in a mine 2 kilometers underground. The neutrino detector consisted of a 12-meter-diameter transparent acrylic plastic sphere, which contained 1000 metric tons of heavy water. Remember that an ordinary water nucleus contains two hydrogen atoms and one oxygen atom. Heavy water instead contains two deuterium atoms and one oxygen atom, and incoming neutrinos can occasionally break up the loosely bound proton and neutron that make up the deuterium nucleus. The sphere of heavy water was surrounded by a shield of 1700 metric tons of very pure water, which in turn was surrounded by 9600 photomultipliers, devices that detect flashes of light produced after neutrinos interact with the heavy water.

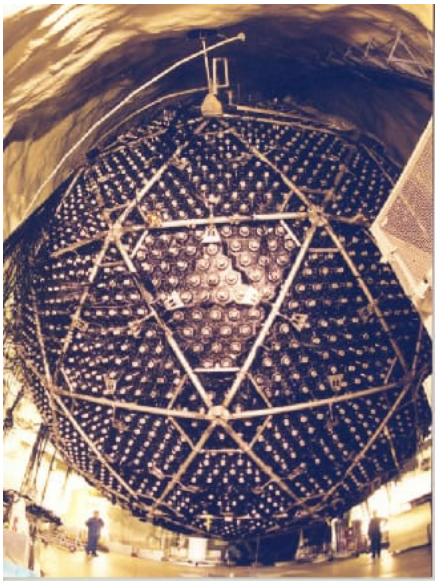


Figure 16.20 : Sudbury Neutrino Detector. The 12-meter sphere of the Sudbury Neutrino Detector lies more than 2 kilometers underground and holds 1000 metric tons of heavy water. (credit: A.B. McDonald (Queen's University) et al., The Sudbury Neutrino Observatory Institute)

To the enormous relief of astronomers who make models of the Sun, the Sudbury experiment detected about 1 neutrino per hour and has shown that the *total* number of neutrinos reaching the heavy water is just what solar models predict. Only one-third of these, however, are electron neutrinos. It appears that two-thirds of the electron neutrinos produced by the Sun transform themselves into one of the other types of neutrinos as they make their way from the core of the Sun to Earth. This is why the earlier experiments saw only one-third the number of neutrinos expected.

Although it is not intuitively obvious, such neutrino oscillations can happen only if the mass of the electron neutrino is not zero. Other experiments indicate that its mass is tiny (even compared to the electron). The 2015 Nobel Prize in physics was awarded to researchers Takaaki Kajita and Arthur B. McDonald for their work establishing the changeable nature of neutrinos. (Raymond Davis shared the 2002 Nobel Prize with Japan's Masatoshi Koshiba for the experiments that led to our understanding of the neutrino problem in the first place.) But the fact that the neutrino has mass at all has deep implications for both physics and astronomy. For example, we will look at the role that neutrinos play in the inventory of the mass of the universe in The Big Bang.

The Borexino experiment, an international experiment conducted in Italy, detected neutrinos coming from the Sun that were identified as coming from different reactions. Whereas the p-p chain is the reaction producing most of the Sun's energy, it is not the only nuclear reaction occurring in the Sun's core. There are side reactions involving nuclei of such elements as beryllium and boron. By probing the number of neutrinos that come from each reaction, the Borexino experiment has helped us confirm in detail our understanding of nuclear fusion in the Sun. In 2014, the Borexino experiment also identified neutrinos that were produced by the first step in the p-p chain, confirming the models of solar astronomers.

In 2020, scientists with the Borexino experiment were also able to identify neutrinos coming specifically from the CNO cycle of

nuclear fusion. As discussed toward the end of Telescopes Today, CNO fusion plays a very minor role in the Sun, but becomes important in stars that are more massive and hotter than the Sun.

It's amazing that a series of experiments that began with enough cleaning fluid to fill a swimming pool brought down the shafts of an old gold mine is now teaching us about the energy source of the Sun and the properties of matter! This is a good example of how experiments in astronomy and physics, coupled with the best theoretical models we can devise, continue to lead to fundamental changes in our understanding of nature.

This page titled 19.6: The Solar Interior - Observations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 16.5: The Solar Interior - Observations by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

19.7: Key Terms

conduction
process by which heat is directly transmitted through a substance when there is a difference of temperature between adjoining regions caused by atomic or molecular collisions

convection
movement caused within a gas or liquid by the tendency of hotter, and therefore less dense material, to rise and colder, denser material to sink under the influence of gravity, which consequently results in transfer of heat

fission
breaking up of heavier atomic nuclei into lighter ones

fusion
building up of heavier atomic nuclei from lighter ones

helioseismology
study of pulsations or oscillations of the Sun in order to determine the characteristics of the solar interior

hydrostatic equilibrium
balance between the weights of various layers, as in a star or Earth's atmosphere, and the pressures that support them

neutrino
fundamental particle that has no charge and a mass that is tiny relative to an electron; it rarely interacts with ordinary matter and comes in three different types

positron
particle with the same mass as an electron, but positively charged

proton-proton chain
series of thermonuclear reactions by which nuclei of hydrogen are built up into nuclei of helium

radiation
emission of energy as electromagnetic waves or photons also the transmitted energy itself

This page titled 19.7: Key Terms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

20: THE STARS

20.1: The Brightness of Stars
20.2: Colors of Stars
20.3: The Spectra of Stars (and Brown Dwarfs)
20.4: Measuring Stellar Masses
20.5: The H-R Diagram

This page titled 20: The Stars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

20.1: The Brightness of Stars

Learning Objectives

By the end of this section, you will be able to:

- Explain the difference between luminosity and apparent brightness
- Understand how astronomers specify brightness with magnitudes

Luminosity

Perhaps the most important characteristic of a star is its luminosity—the total amount of energy at all wavelengths that it emits per second. Earlier, we saw that the Sun puts out a tremendous amount of energy every second. (And there are stars far more luminous than the Sun out there.) To make the comparison among stars easy, astronomers express the luminosity of other stars in terms of the Sun's luminosity. For example, the luminosity of Sirius is about 25 times that of the Sun. We use the symbol L_{Sun} to denote the Sun's luminosity; hence, that of Sirius can be written as 25 L_{Sun}. In a later chapter, we will see that if we can measure how much energy a star emits and we also know its mass, then we can calculate how long it can continue to shine before it exhausts its nuclear energy and begins to die.

Apparent Brightness

Astronomers are careful to distinguish between the luminosity of the star (the total energy output) and the amount of energy that happens to reach our eyes or a telescope on Earth. Stars are democratic in how they produce radiation; they emit the same amount of energy in every direction in space. Consequently, only a minuscule fraction of the energy given off by a star actually reaches an observer on Earth. We call the amount of a star's energy that reaches a given area (say, one square meter) each second here on Earth its apparent brightness. If you look at the night sky, you see a wide range of apparent brightnesses among the stars. Most stars, in fact, are so dim that you need a telescope to detect them.

If all stars were the same luminosity—if they were like standard bulbs with the same light output—we could use the difference in their apparent brightnesses to tell us something we very much want to know: how far away they are. Imagine you are in a big concert hall or ballroom that is dark except for a few dozen 25-watt bulbs placed in fixtures around the walls. Since they are all 25-watt bulbs, their luminosity (energy output) is the same. But from where you are standing in one corner, they do *not* have the same apparent brightness. Those close to you appear brighter (more of their light reaches your eye), whereas those far away appear dimmer (their light has spread out more before reaching you). In this way, you can tell which bulbs are closest to you. In the same way, if all the stars had the same luminosity, we could immediately infer that the brightest-appearing stars were close by and the dimmest-appearing ones were far away.

To pin down this idea more precisely, recall from the Radiation and Spectra chapter that we know exactly how light fades with increasing distance. The energy we receive is inversely proportional to the square of the distance. If, for example, we have two stars of the same luminosity and one is twice as far away as the other, it will look four times dimmer than the closer one. If it is three times farther away, it will look nine (three squared) times dimmer, and so forth.

Alas, the stars do not all have the same luminosity. (Actually, we are pretty glad about that because having many different types of stars makes the universe a much more interesting place.) But this means that if a star looks dim in the sky, we cannot tell whether it appears dim because it has a low luminosity but is relatively nearby, or because it has a high luminosity but is very far away. To measure the luminosities of stars, we must first compensate for the dimming effects of distance on light, and to do that, we must know how far away they are. Distance is among the most difficult of all astronomical measurements. We will return to how it is determined after we have learned more about the stars. For now, we will describe how astronomers specify the apparent brightness of stars.

The Magnitude Scale

The process of measuring the apparent brightness of stars is called *photometry* (from the Greek *photo* meaning "light" and *–metry* meaning "to measure"). As we saw Observing the Sky: The Birth of Astronomy, astronomical photometry began with Hipparchus. Around 150 B.C.E., he erected an observatory on the island of Rhodes in the Mediterranean. There he prepared a catalog of nearly 1000 stars that included not only their positions but also estimates of their apparent brightnesses.

Hipparchus did not have a telescope or any instrument that could measure apparent brightness accurately, so he simply made estimates with his eyes. He sorted the stars into six brightness categories, each of which he called a magnitude. He referred to the brightest stars in his catalog as first-magnitudes stars, whereas those so faint he could barely see them were sixth-magnitude stars. During the nineteenth century, astronomers attempted to make the scale more precise by establishing exactly how much the apparent brightness of a sixth-magnitude star differs from that of a first-magnitude star. Measurements showed that we receive about 100 times more light from a first-magnitude star than from a sixth-magnitude star. Based on this measurement, astronomers then defined an accurate magnitude system in which a difference of five magnitudes corresponds exactly to a brightness ratio of 100:1. In addition, the magnitudes of stars are decimalized; for example, a star isn't just a "second-magnitude star," it has a magnitude of 2.0 (or 2.1, 2.3, and so forth). So what number is it that, when multiplied together five times, gives you this factor of 100? Play on your calculator and see if you can get it. The answer turns out to be about 2.5, which is the fifth root of 100. This means that a magnitude 1.0 star and a magnitude 2.0 star differ in brightness by a factor of about 2.5. Likewise, we receive about 2.5 times as much light from a magnitude 2.0 star as from a magnitude 3.0 star. What about the difference between a magnitude 1.0 star and a magnitude 3.0 star? Since the difference is 2.5 times for each "step" of magnitude, the total difference in brightness is 2.5 × 2.5 = 6.25 times.

Here are a few rules of thumb that might help those new to this system. If two stars differ by 0.75 magnitudes, they differ by a factor of about 2 in brightness. If they are 2.5 magnitudes apart, they differ in brightness by a factor of 10, and a 4-magnitude difference corresponds to a difference in brightness of a factor of 40.You might be saying to yourself at this point, "Why do astronomers continue to use this complicated system from more than 2000 years ago?" That's an excellent question and, as we shall discuss, astronomers today can use other ways of expressing how bright a star looks. But because this system is still used in many books, star charts, and computer apps, we felt we had to introduce students to it (even though we were very tempted to leave it out.)

The brightest stars, those that were traditionally referred to as first-magnitude stars, actually turned out (when measured accurately) not to be identical in brightness. For example, the brightest star in the sky, Sirius, sends us about 10 times as much light as the average first-magnitude star. On the modern magnitude scale, Sirius, the star with the brightest apparent magnitude, has been assigned a magnitude of –1.5. Other objects in the sky can appear even brighter. Venus at its brightest is of magnitude –4.4, while the Sun has a magnitude of –26.8. Figure 17.2 shows the range of observed magnitudes from the brightest to the faintest, along with the actual magnitudes of several well-known objects. The important fact to remember when using magnitude is that the system goes backward: the *larger* the magnitude, the *fainter* the object you are observing.

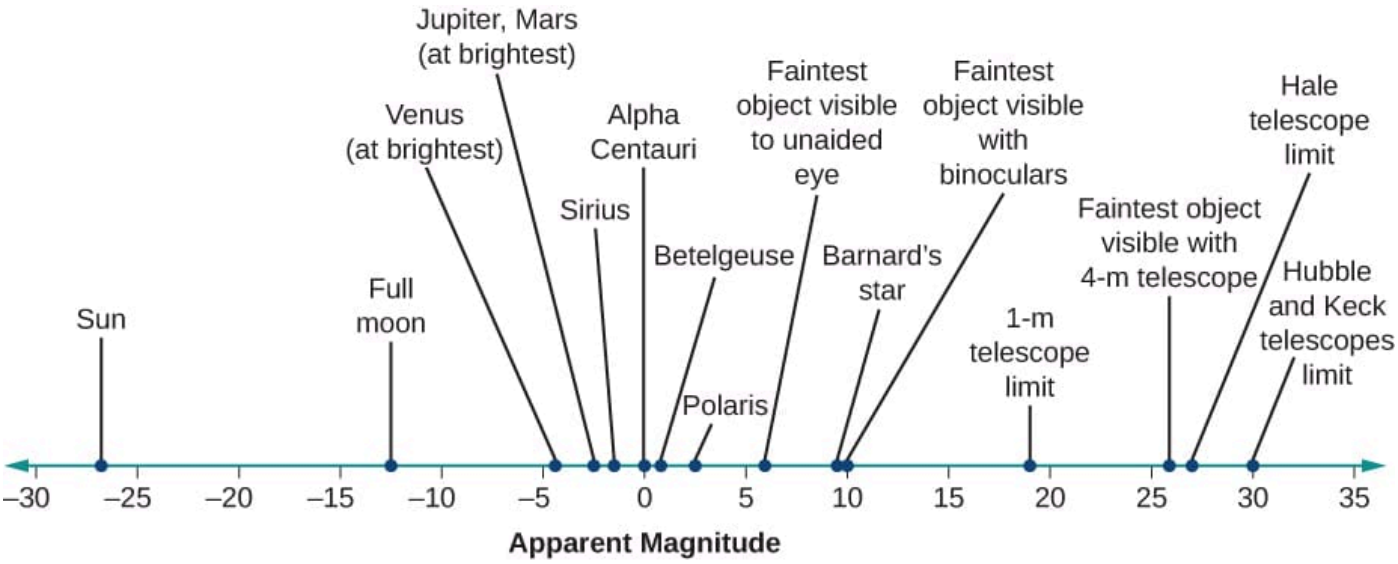


Figure 17.2 : Apparent Magnitudes of Well-Known Objects. The faintest magnitudes that can be detected by the unaided eye, binoculars, and large telescopes are also shown.

Example 17.1: The Magnitude Equation

Even scientists can't calculate fifth roots in their heads, so astronomers have summarized the above discussion in an equation to help calculate the difference in brightness for stars with different magnitudes. If m_1 and m_2 are the magnitudes of two stars, then we can calculate the ratio of their brightness $\left(\frac{b_2}{b_1}\right)$ using this equation:

$$m_1 - m_2 = 2.5 \log\left(\frac{b_2}{b_1}\right) \qquad \text{or} \qquad \frac{b_2}{b_1} = 2.5^{m_1-}$$

Here is another way to write this equation:

$$\frac{b_2}{b_1} = \left(100^{0.2}\right)^{m_1-}$$

Let's do a real example, just to show how this works. Imagine that an astronomer has discovered something special about a dim star (magnitude 8.5), and she wants to tell her students how much dimmer the star is than Sirius. Star 1 in the equation will be our dim star and star 2 will be Sirius.

Solution

Remember, Sirius has a magnitude of -1.5. In that case:

$$\begin{aligned} \frac{b_2}{b_1} &= \left(100^{0.2}\right)^{8.5-(-1.5)} = \left(100^{0.2}\right)^{10} \\ &= (100)^2 = 100 \times 100 = 10, \end{aligned}$$

Exercise 20.1.1

It is a common misconception that Polaris (magnitude 2.0) is the brightest star in the sky, but, as we saw, that distinction actually belongs to Sirius (magnitude –1.5). How does Sirius' apparent brightness compare to that of Polaris?

Answer available at the end of this section.

Other Units of Brightness

Although the magnitude scale is still used for visual astronomy, it is not used at all in newer branches of the field. In radio astronomy, for example, no equivalent of the magnitude system has been defined. Rather, radio astronomers measure the amount of energy being collected each second by each square meter of a radio telescope and express the brightness of each source in terms of, for example, watts per square meter.

Similarly, most researchers in the fields of infrared, X-ray, and gamma-ray astronomy use energy per area per second rather than magnitudes to express the results of their measurements. Nevertheless, astronomers in all fields are careful to distinguish between the *luminosity* of the source (even when that luminosity is all in X-rays) and the amount of energy that happens to reach us on Earth. After all, the luminosity is a really important characteristic that tells us a lot about the object in question, whereas the energy that reaches Earth is an accident of cosmic geography.

To make the comparison among stars easy, in this text, we avoid the use of magnitudes as much as possible and will express the luminosity of other stars in terms of the Sun's luminosity. For example, the luminosity of Sirius is 25 times that of the Sun. We use the symbol L_{Sun} to denote the Sun's luminosity; hence, that of Sirius can be written as 25 L_{Sun}.

Answers

Exercise 20.1.1

$\frac{b_{\text{Sirius}}}{b_{\text{Polaris}}} = \left(100^{0.2}\right)^{2.0-(-1.5)} = \left(100^{0.2}\right)^{3.5} = 100^{0.7} =$

(Hint: If you only have a basic calculator, you may wonder how to take 100 to the 0.7th power. But this is something you can ask Google to do. Google now accepts mathematical questions and will answer them. So try it for yourself. Ask Google, "What is 100 to the 0.7th power?")

Our calculation shows that Sirius' apparent brightness is 25 times greater than Polaris' apparent brightness.

This page titled 20.1: The Brightness of Stars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 17.2: The Brightness of Stars by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

20.2: Colors of Stars

Learning Objectives

By the end of this section, you will be able to:

- Compare the relative temperatures of stars based on their colors
- Understand how astronomers use color indexes to measure the temperatures of stars

Look at the beautiful picture of the stars in the Sagittarius Star Cloud shown in Figure 17.3. The stars show a multitude of colors, including red, orange, yellow, white, and blue. As we have seen, stars are not all the same color because they do not all have identical temperatures. To define *color* precisely, astronomers have devised quantitative methods for characterizing the color of a star and then using those colors to determine stellar temperatures. In the chapters that follow, we will provide the temperature of the stars we are describing, and this section tells you how those temperatures are determined from the colors of light the stars give off.

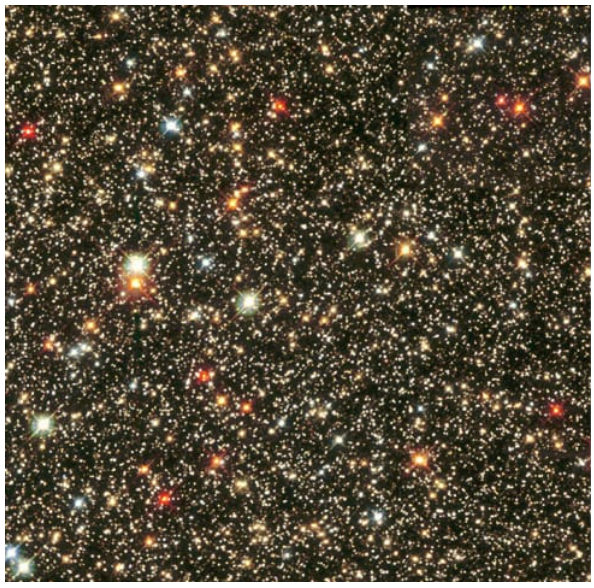


Figure 17.3 : Sagittarius Star Cloud. This image, which was taken by the Hubble Space Telescope, shows stars in the direction toward the center of the Milky Way Galaxy. The bright stars glitter like colored jewels on a black velvet background. The color of a star indicates its temperature. Blue-white stars are much hotter than the Sun, whereas red stars are cooler. On average, the stars in this field are at a distance of about 25,000 light-years (which means it takes light 25,000 years to traverse the distance from them to us) and the width of the field is about 13.3 light-years. (credit: Hubble Heritage Team (AURA/STScI/NASA))

Color and Temperature

As we learned in The Electromagnetic Spectrum section, Wien's law relates stellar color to stellar temperature. Blue colors dominate the visible light output of very hot stars (with much additional radiation in the ultraviolet). On the other hand, cool stars emit most of their visible light energy at red wavelengths (with more radiation coming off in the infrared) (Table 17.1). The color of a star therefore provides a measure of its intrinsic or true surface temperature (apart from the effects of reddening by interstellar dust, which will be discussed in Between the Stars: Gas and Dust in Space). Color does not depend on the distance to the object. This should be familiar to you from everyday experience. The color of a traffic signal, for example, appears the same no matter how far away it is. If we could somehow take a star, observe it, and then move it much farther away, its apparent brightness (magnitude) would change. But this change in brightness is the same for all wavelengths, and so its color would remain the same.

Table 17.1: Example Star Colors and Corresponding Approximate Temperatures

Star Color	Approximate Temperature	Example
Blue	25,000 K	Spica
White	10,000 K	Vega
Yellow	6000 K	Sun
Orange	4000 K	Aldebaran
Red	3000 K	Betelgeuse

Link to Learning

Go to this interactive simulation from the University of Colorado to see the color of a star changing as the temperature is changed.

The hottest stars have temperatures of over 40,000 K, and the coolest stars have temperatures of about 2000 K. Our Sun's surface temperature is about 6000 K; its peak wavelength color is a slightly greenish-yellow. In space, the Sun would look white, shining with about equal amounts of reddish and bluish wavelengths of light. It looks somewhat yellow as seen from Earth's surface because our planet's nitrogen molecules scatter some of the shorter (i.e., blue) wavelengths out of the beams of sunlight that reach us, leaving more long wavelength light behind. This also explains why the sky is blue: the blue sky is sunlight scattered by Earth's atmosphere.

Color Indices

In order to specify the exact color of a star, astronomers normally measure a star's apparent brightness through filters, each of which transmits only the light from a particular narrow band of wavelengths (colors). A crude example of a filter in everyday life is a green-colored, plastic, soft drink bottle, which, when held in front of your eyes, lets only the green colors of light through.

One commonly used set of filters in astronomy measures stellar brightness at three wavelengths corresponding to ultraviolet, blue, and yellow light. The filters are named: U (ultraviolet), B (blue), and V (visual, for yellow). These filters transmit light near the wavelengths of 360 nanometers (nm), 420 nm, and 540 nm, respectively. The brightness measured through each filter is usually expressed in magnitudes. The difference between any two of these magnitudes—say, between the blue and the visual magnitudes (B–V)—is called a color index.

By agreement among astronomers, the ultraviolet, blue, and visual magnitudes of the UBV system are adjusted to give a color index of 0 to a star with a surface temperature of about 10,000 K, such as Vega. The B–V color indexes of stars range from –0.4 for the bluest stars, with temperatures of about 40,000 K, to +2.0 for the reddest stars, with temperatures of about 2000 K. The B–V index for the Sun is about +0.65. Note that, by convention, the B–V index is always the "bluer" minus the "redder" color.

Why use a color index if it ultimately implies temperature? Because the brightness of a star through a filter is what astronomers actually measure, and we are always more comfortable when our statements have to do with measurable quantities.

This page titled 20.2: Colors of Stars is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 17.3: Colors of Stars by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

20.3: The Spectra of Stars (and Brown Dwarfs)

Learning Objectives

By the end of this section, you will be able to:

- Describe how astronomers use spectral classes to characterize stars
- Explain the difference between a star and a brown dwarf

Measuring colors is only one way of analyzing starlight. Another way is to use a spectrograph to spread out the light into a spectrum (see the Radiation and Spectra and the Astronomical Instruments chapters). In 1814, the German physicist Joseph Fraunhofer observed that the spectrum of the Sun shows dark lines crossing a continuous band of colors. In the 1860s, English astronomers Sir William Huggins and Lady Margaret Huggins (Figure 17.4) succeeded in identifying some of the lines in stellar spectra as those of known elements on Earth, showing that the same chemical elements found in the Sun and planets exist in the stars. Since then, astronomers have worked hard to perfect experimental techniques for obtaining and measuring spectra, and they have developed a theoretical understanding of what can be learned from spectra. Today, spectroscopic analysis is one of the cornerstones of astronomical research.

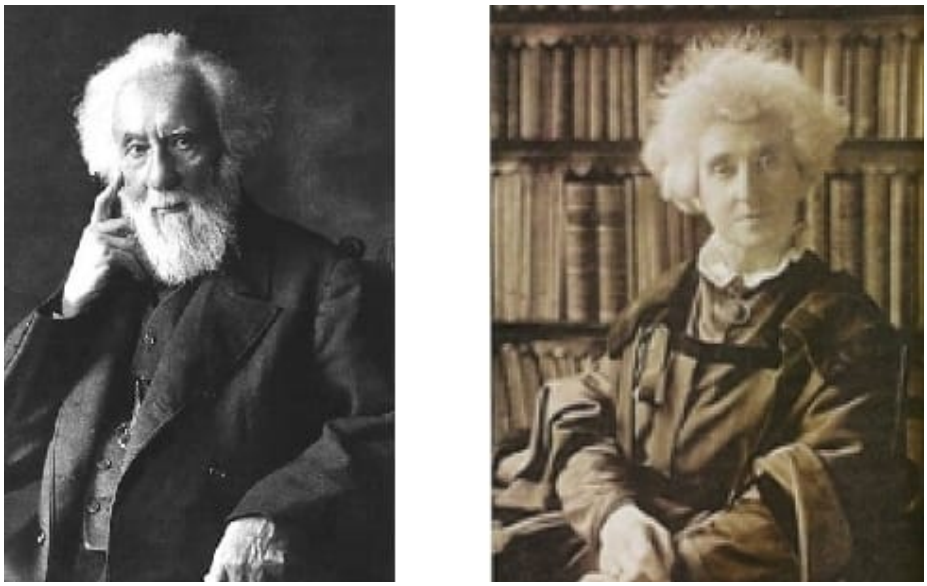


Figure 17.4 : William Huggins (1824–1910) and Margaret Huggins (1848–1915). William and Margaret Huggins were the first to identify the lines in the spectrum of a star other than the Sun; they also took the first spectrogram, or photograph of a stellar spectrum.

Formation of Stellar Spectra

When the spectra of different stars were first observed, astronomers found that they were not all identical. Since the dark lines are produced by the chemical elements present in the stars, astronomers first thought that the spectra differ from one another because stars are not all made of the same chemical elements. This hypothesis turned out to be wrong. *The primary reason that stellar spectra look different is because the stars have different temperatures*. Most stars have nearly the same composition as the Sun, with only a few exceptions.

Hydrogen, for example, is by far the most abundant element in most stars. However, lines of hydrogen are not seen in the spectra of the hottest and the coolest stars. In the atmospheres of the hottest stars, hydrogen atoms are completely ionized. Because the electron and the proton are separated, ionized hydrogen cannot produce absorption lines. (Recall from the Formation of Spectral Lines section, the lines are the result of electrons in orbit around a nucleus changing energy levels.)

In the atmospheres of the coolest stars, hydrogen atoms have their electrons attached and can switch energy levels to produce

lines. However, practically all of the hydrogen atoms are in the lowest energy state (unexcited) in these stars and thus can absorb only those photons able to lift an electron from that first energy level to a higher level. Photons with enough energy to do this lie in the ultraviolet part of the electromagnetic spectrum, and there are very few ultraviolet photons in the radiation from a cool star. What this means is that if you observe the spectrum of a very hot or very cool star with a typical telescope on the surface of Earth, the most common element in that star, hydrogen, will show very weak spectral lines or none at all.

The hydrogen lines in the visible part of the spectrum (called *Balmer lines*) are strongest in stars with intermediate temperatures—not too hot and not too cold. Calculations show that the optimum temperature for producing visible hydrogen lines is about 10,000 K. At this temperature, an appreciable number of hydrogen atoms are excited to the second energy level. They can then absorb additional photons, rise to still-higher levels of excitation, and produce a dark absorption line. Similarly, every other chemical element, in each of its possible stages of ionization, has a characteristic temperature at which it is most effective in producing absorption lines in any particular part of the spectrum.

Classification of Stellar Spectra

Astronomers use the patterns of lines observed in stellar spectra to sort stars into a spectral class. Because a star's temperature determines which absorption lines are present in its spectrum, these spectral classes are a measure of its surface temperature. There are seven standard spectral classes. From hottest to coldest, these seven spectral classes are designated O, B, A, F, G, K, and M. Recently, astronomers have added three additional classes for even cooler objects—L, T, and Y.

At this point, you may be looking at these letters with wonder and asking yourself why astronomers didn't call the spectral types A, B, C, and so on. You will see, as we tell you the history, that it's an instance where tradition won out over common sense.

In the 1880s, Williamina Fleming devised a system to classify stars based on the strength of hydrogen absorption lines. Spectra with the strongest lines were classified as "A" stars, the next strongest "B," and so on down the alphabet to "O" stars, in which the hydrogen lines were very weak. But we saw above that hydrogen lines alone are not a good indicator for classifying stars, since their lines disappear from the visible light spectrum when the stars get too hot or too cold.

In the 1890s, Annie Jump Cannon revised this classification system, focusing on just a few letters from the original system: A, B, F, G, K, M, and O. Instead of starting over, Cannon also rearranged the existing classes—in order of decreasing temperature—into the sequence we have learned: O, B, A, F, G, K, M. As you can read in the feature on Annie Cannon: Classifier of the Stars in this chapter, she classified around 500,000 stars over her lifetime, classifying up to three stars per minute by looking at the stellar spectra.

Link to Learning

For a deep dive into spectral types, explore the interactive project at the Sloan Digital Sky Survey in which you can practice classifying stars yourself.

To help astronomers remember this crazy order of letters, Cannon created a mnemonic, "Oh Be A Fine Girl, Kiss Me." (If you prefer, you can easily substitute "Guy" for "Girl.") Other mnemonics, which we hope will not be relevant for you, include "Oh Brother, Astronomers Frequently Give Killer Midterms" and "Oh Boy, An F Grade Kills Me!" With the new L, T, and Y spectral classes, the mnemonic might be expanded to "Oh Be A Fine Girl (Guy), Kiss Me Like That, Yo!"

Each of these spectral classes, except possibly for the Y class which is still being defined, is further subdivided into 10 subclasses designated by the numbers 0 through 9. A B0 star is the hottest type of B star; a B9 star is the coolest type of B star and is only slightly hotter than an A0 star.

And just one more item of vocabulary: for historical reasons, astronomers call all the elements heavier than helium *metals*, even though most of them do not show metallic properties. (If you are getting annoyed at the peculiar jargon that astronomers use, just bear in mind that every field of human activity tends to develop its own specialized vocabulary. Just try reading a credit card or social media agreement form these days without training in law!)

Let's take a look at some of the details of how the spectra of the stars change with temperature. (It is these details that allowed Annie Cannon to identify the spectral types of stars as quickly as three per minute!) As Figure 17.5 shows, in the hottest O stars (those with temperatures over 28,000 K), only lines of ionized helium and highly ionized atoms of other elements are conspicuous. Hydrogen lines are strongest in A stars with atmospheric temperatures of about 10,000 K. Ionized metals provide

the most conspicuous lines in stars with temperatures from 6000 to 7500 K (spectral type F). In the coolest M stars (below 3500 K), absorption bands of titanium oxide and other molecules are very strong. By the way, the spectral class assigned to the Sun is G2. The sequence of spectral classes is summarized in Table 17.2.

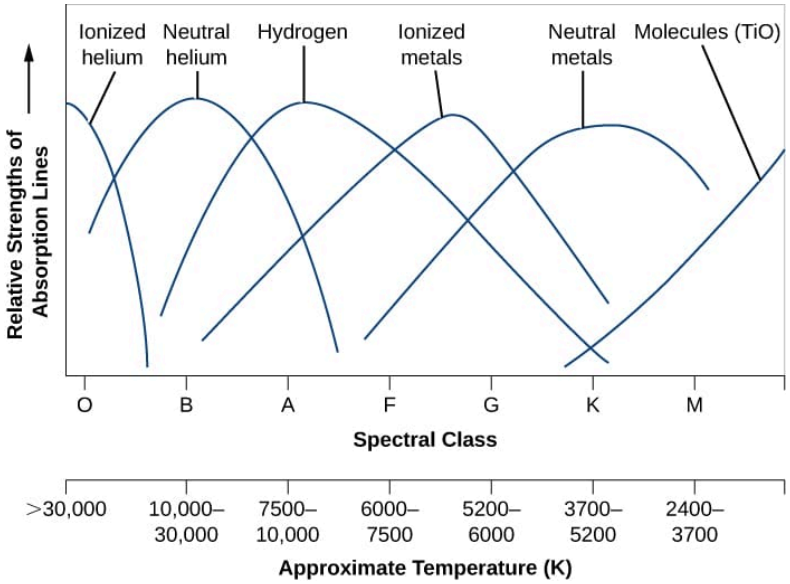


Figure 17.5 : Absorption Lines in Stars of Different Temperatures. This graph shows the strengths of absorption lines of different chemical species (atoms, ions, molecules) as we move from hot (left) to cool (right) stars. The sequence of spectral types is also shown.

Link to Learning

Use the Spectrum Explorer to see how each element's absorption lines change with temperature.

Table 17.2: Spectral Classes for Stars

Spectral Class	Color	Approximate Temperature (K)	Principal Features	Examples
O	Blue	> 30,000	Neutral and ionized helium lines, weak hydrogen lines	10 Lacertae
B	Blue-white	10,000–30,000	Neutral helium lines, strong hydrogen lines	Rigel, Spica
A	White	7500–10,000	Strongest hydrogen lines, weak ionized calcium lines, weak ionized metal (e.g., iron, magnesium) lines	Sirius, Vega
F	Yellow-white	6000–7500	Strong hydrogen lines, strong ionized calcium lines, weak sodium lines, many ionized metal lines	Canopus, Procyon
G	Yellow	5200–6000	Weaker hydrogen lines, strong ionized calcium lines, strong sodium lines, many lines of ionized and neutral metals	Sun, Capella
K	Orange	3700–5200	Very weak hydrogen lines, strong ionized calcium lines, strong sodium lines, many lines of neutral metals	Arcturus, Aldebaran
M	Red	2400–3700	Strong lines of neutral metals and molecular bands of titanium oxide dominate	Betelgeuse, Antares
L	Red	1300–2400	Metal hydride lines, alkali metal lines (e.g., sodium, potassium, rubidium)	Teide 1
T	Magenta	700–1300	Methane lines	Gliese 229B

Spectral Class	Color	Approximate Temperature (K)	Principal Features	Examples
Y	Infrared[1]	< 700	Ammonia lines	WISE 1828+2650

To see how spectral classification works, let's use Figure 17.5. Suppose you have a spectrum in which the hydrogen lines are about half as strong as those seen in an A star. Looking at the lines in our figure, you see that the star could be either a B star or a G star. But if the spectrum also contains helium lines, then it is a B star, whereas if it contains lines of ionized iron and other metals, it must be a G star.

If you look at Figure 17.6, you can see that you, too, could assign a spectral class to a star whose type was not already known. All you have to do is match the pattern of spectral lines to a standard star (like the ones shown in the figure) whose type has already been determined.

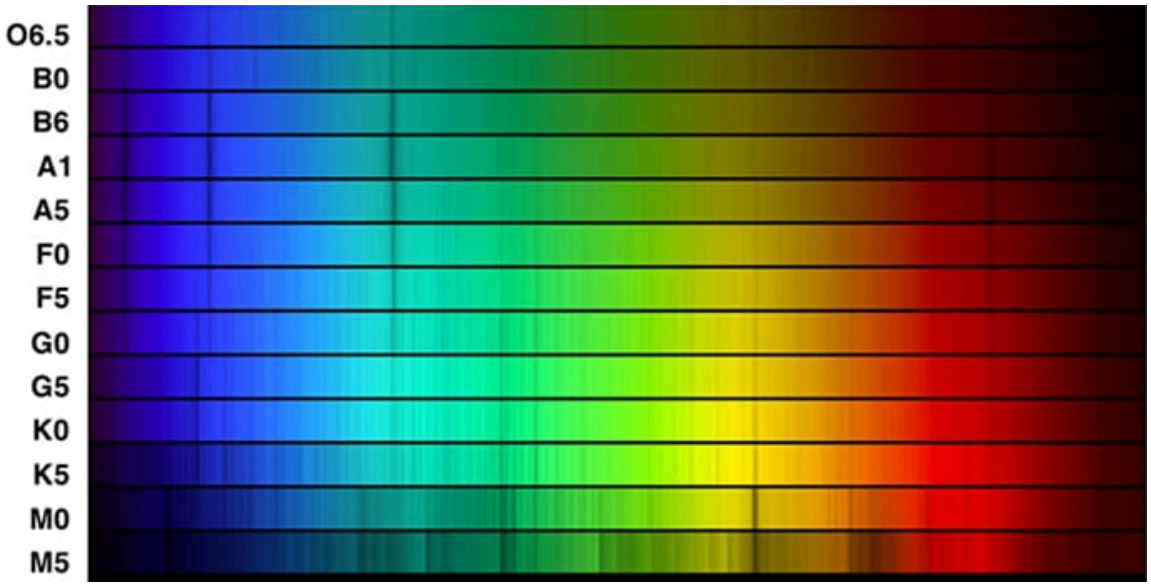


Figure 17.6 : Spectra of Stars with Different Spectral Classes. This image compares the spectra of the different spectral classes. The spectral class assigned to each of these stellar spectra is listed at the left of the picture. The strongest four lines seen at spectral type A1 (one in the red, one in the blue-green, and two in the blue) are Balmer lines of hydrogen. Note how these lines weaken at both higher and lower temperatures, as Figure 17.5 also indicates. The strong pair of closely spaced lines in the yellow in the cool stars is due to neutral sodium (one of the neutral metals in Figure 17.5). (Credit: modification of work by NOAO/AURA/NSF)

Both colors and spectral classes can be used to estimate the temperature of a star. Spectra are harder to measure because the light has to be bright enough to be spread out into all colors of the rainbow, and detectors must be sensitive enough to respond to individual wavelengths. In order to measure colors, the detectors need only respond to the many wavelengths that pass simultaneously through the colored filters that have been chosen—that is, to *all* the blue light or *all* the yellow-green light.

Voyagers in Astronomy: Annie Cannon: Classifier of the Stars

Annie Jump Cannon was born in Delaware in 1863 (Figure 17.7). In 1880, she went to Wellesley College, one of the new breed of US colleges opening up to educate young women. Wellesley, only 5 years old at the time, had the second student physics lab in the country and provided excellent training in basic science. After college, Cannon spent a decade with her parents but was very dissatisfied, longing to do scientific work. After her mother's death in 1893, she returned to Wellesley as a teaching assistant and also to take courses at Radcliffe, the women's college associated with Harvard.



Figure 17.7 : Annie Jump Cannon (1863–1941). Cannon is well-known for her classifications of stellar spectra. (credit: modification of work by Smithsonian Institution)

In the late 1800s, the director of the Harvard Observatory, Edward C. Pickering, needed lots of help with his ambitious program of classifying stellar spectra. The basis for these studies was a monumental collection of nearly a million photographic spectra of stars, obtained from many years of observations made at Harvard College Observatory in Massachusetts as well as at its remote observing stations in South America and South Africa. Pickering quickly discovered that educated young women could be hired as assistants for one-third or one-fourth the salary paid to men, and they would often put up with working conditions and repetitive tasks that men with the same education would not tolerate. These women became known as the Harvard Computers. (We should emphasize that astronomers were not alone in reaching such conclusions about the relatively new idea of upper-class, educated women working outside the home: women were exploited and undervalued in many fields. This is a legacy from which our society is just beginning to emerge.)

Cannon was hired by Pickering as one of the "computers" to help with the classification of spectra. She became so good at it that she could visually examine and determine the spectral types of several hundred stars per hour (dictating her conclusions to an assistant). She made many discoveries while investigating the Harvard photographic plates, including 300 variable stars (stars whose luminosity changes periodically). But her main legacy is a marvelous catalog of spectral types for hundreds of thousands of stars, which served as a foundation for much of twentieth-century astronomy.

In 1911, a visiting committee of astronomers reported that "she is the one person in the world who can do this work quickly and accurately" and urged Harvard to give Cannon an official appointment in keeping with her skill and renown. Not until 1938, however, did Harvard appoint her an astronomer at the university; she was then 75 years old.

Cannon received the first honorary degree Oxford awarded to a woman, and she became the first woman to be elected an officer of the American Astronomical Society, the main professional organization of astronomers in the US. She generously donated the money from one of the major prizes she had won to found a special award for women in astronomy, now known as the Annie Jump Cannon Prize. True to form, she continued classifying stellar spectra almost to the very end of her life in 1941.

Spectral Classes L, T, and Y

The scheme devised by Cannon worked well until 1988, when astronomers began to discover objects even cooler than M9-type stars. We use the word *object* because many of the new discoveries are not true stars. A star is defined as an object that during some part of its lifetime derives 100% of its energy from the same process that makes the Sun shine—the fusion of hydrogen nuclei (protons) into helium. Objects with masses less than about 7.5% of the mass of our Sun (about 0.075 M_{Sun}) do not become hot enough for hydrogen fusion to take place. Even before the first such "failed star" was found, this class of objects, with masses intermediate between stars and planets, was given the name brown dwarfs.

Brown dwarfs are very difficult to observe because they are extremely faint and cool, and they put out most of their light in the infrared part of the spectrum. It was only after the construction of very large telescopes, like the Keck telescopes in Hawaii, and the development of very sensitive infrared detectors, that the search for brown dwarfs succeeded. The first brown dwarf was discovered in 1988, and, as of the summer of 2015, there are more than 2200 known brown dwarfs.

Initially, brown dwarfs were given spectral classes like $M10^{+}$ or "much cooler than M9," but so many are now known that it

is possible to begin assigning spectral types. The hottest brown dwarfs are given types L0–L9 (temperatures in the range 2400–1300 K), whereas still cooler (1300–700 K) objects are given types T0–T9 (see Figure 17.8). In class L brown dwarfs, the lines of titanium oxide, which are strong in M stars, have disappeared. This is because the L dwarfs are so cool that atoms and molecules can gather together into dust particles in their atmospheres; the titanium is locked up in the dust grains rather than being available to form molecules of titanium oxide. Lines of steam (hot water vapor) are present, along with lines of carbon monoxide and neutral sodium, potassium, cesium, and rubidium. Methane (CH_4) lines are strong in class-T brown dwarfs, as methane exists in the atmosphere of the giant planets in our own solar system.

In 2009, astronomers discovered ultra-cool brown dwarfs with temperatures of 500–600 K. These objects exhibited absorption lines due to ammonia (NH_3), which are not seen in T dwarfs. A new spectral class, Y, was created for these objects. As of 2021, over 50 brown dwarfs belonging to spectral class Y have been discovered, some with temperatures comparable to that of the human body (about 300 K).[2]

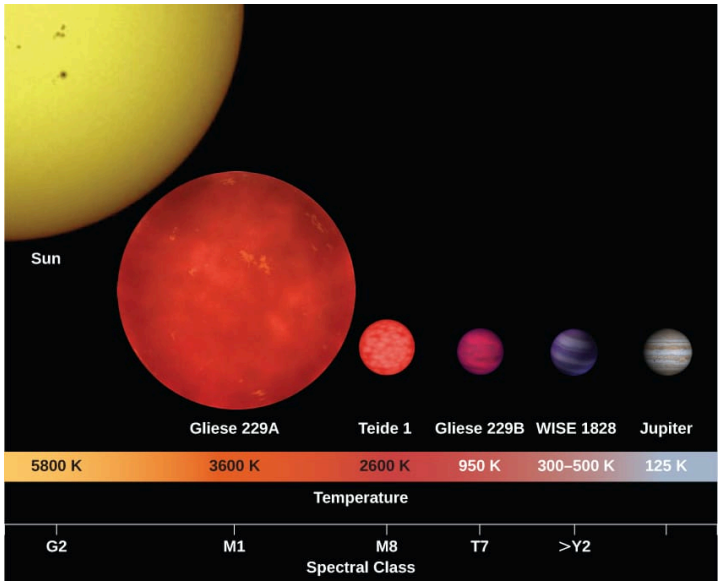


Figure 17.8 : Brown Dwarfs. This illustration shows the sizes and surface temperatures of brown dwarfs Teide 1, Gliese 229B, and WISE1828 in relation to the Sun, a red dwarf star (Gliese 229A), and Jupiter. (credit: modification of work by MPIA/V. Joergens)

Most brown dwarfs start out with atmospheric temperatures and spectra like those of true stars with spectral classes of M6.5 and later, even though the brown dwarfs are not hot and dense enough in their interiors to fuse hydrogen. In fact, the spectra of brown dwarfs and true stars are so similar from spectral types late M through L that it is not possible to distinguish the two types of objects based on spectra alone. An independent measure of mass is required to determine whether a specific object is a brown dwarf or a very low mass star. Since brown dwarfs cool steadily throughout their lifetimes, the spectral type of a given brown dwarf changes with time over a billion years or more from late M through L, T, and Y spectral types.

Low-Mass Brown Dwarfs vs. High-Mass Planets

An interesting property of brown dwarfs is that they are all about the same radius as Jupiter, regardless of their masses. Amazingly, this covers a range of masses from about 13 to 80 times the mass of Jupiter (M_J). This can make distinguishing a low-mass brown dwarf from a high-mass planet very difficult.

So, what is the difference between a low-mass brown dwarf and a high-mass planet? The International Astronomical Union considers the distinctive feature to be *deuterium fusion*. Although brown dwarfs do not sustain regular (proton-proton) hydrogen fusion, they are capable of fusing deuterium (a rare form of hydrogen with one proton and one neutron in its nucleus). The fusion of deuterium can happen at a lower temperature than the fusion of hydrogen. If an object has enough mass to fuse deuterium (about 13 M_J or 0.012 M_{Sun}), it is a brown dwarf. Objects with less than 13 M_J do not fuse deuterium and are usually considered planets.

Footnotes

- 1Absorption by sodium and potassium atoms makes Y dwarfs appear a bit less red than L dwarfs.
- 2The record-holder for the coldest Type Y dwarf is WISE 0855-0714, at around 250 K.

This page titled 20.3: The Spectra of Stars (and Brown Dwarfs) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 17.4: The Spectra of Stars (and Brown Dwarfs) by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

20.4: Measuring Stellar Masses

Learning Objectives

By the end of this section, you will be able to:

- Distinguish the different types of binary star systems
- Understand how we can apply Newton's version of Kepler's third law to derive the sum of star masses in a binary star system
- Apply the relationship between stellar mass and stellar luminosity to determine the physical characteristics of a star

The mass of a star—how much material it contains—is one of its most important characteristics. If we know a star's mass, as we shall see, we can estimate how long it will shine and what its ultimate fate will be. Yet the mass of a star is very difficult to measure directly. Somehow, we need to put a star on the cosmic equivalent of a scale.

Luckily, not all stars live like the Sun, in isolation from other stars. About half the stars are binary stars—two stars that orbit each other, bound together by gravity. Masses of binary stars can be calculated from measurements of their orbits, just as the mass of the Sun can be derived by measuring the orbits of the planets around it (see Orbits and Gravity).

Binary Stars

Before we discuss in more detail how mass can be measured, we will take a closer look at stars that come in pairs. The first binary star was discovered in 1650, less than half a century after Galileo began to observe the sky with a telescope. John Baptiste Riccioli (1598–1671), an Italian astronomer, noted that the star Mizar, in the middle of the Big Dipper's handle, appeared through his telescope as two stars. Since that discovery, thousands of binary stars have been cataloged. (Astronomers call any pair of stars that appear to be close to each other in the sky *double stars*, but not all of these form a true binary, that is, not all of them are physically associated. Some are just chance alignments of stars that are actually at different distances from us.) Although stars most commonly come in pairs, there are also triple and quadruple systems.

One well-known binary star is Castor, located in the constellation of Gemini. By 1804, astronomer William Herschel, who also discovered the planet Uranus, had noted that the fainter component of Castor had slightly changed its position relative to the brighter component. (We use the term "component" to mean a member of a star system.) Here was evidence that one star was moving around another. It was actually the first evidence that gravitational influences exist outside the solar system. The orbital motion of a binary star is shown in Figure 18.4. A binary star system in which both of the stars can be seen with a telescope is called a visual binary.

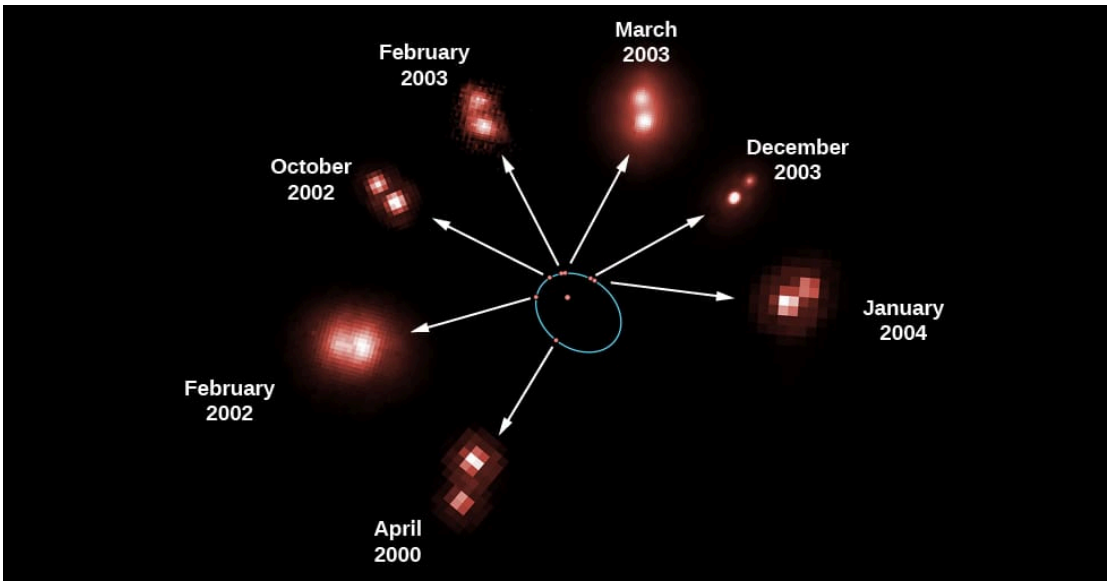


Figure 18.4 : Revolution of a Binary Star. This figure shows seven observations of the mutual revolution of two stars, one a brown dwarf and one an ultra-cool L dwarf. Each red dot on the orbit, which is shown by the blue ellipse, corresponds to the position of one of the dwarfs relative to the other. The reason that the pair of stars looks different on the different dates is that some images were taken with the Hubble Space Telescope and others were taken from the ground. The arrows point to the actual observations that correspond to the positions of each red dot. From these observations, an international team of astronomers directly measured the mass of an ultra-cool brown dwarf star for the first time. Barely the size of the planet Jupiter, the dwarf star weighs in at just 8.5% of the mass of our Sun. (credit: modification of work by ESA/ NASA and Herve Bouy (Max-Planck-Institut für Extraterrestrische Physik/ESO, Germany))

Edward C. Pickering (1846–1919), at Harvard, discovered a second class of binary stars in 1889—a class in which only one of the stars is actually seen directly. He was examining the spectrum of Mizar and found that the dark absorption lines in the brighter star's spectrum were usually double. Not only were there two lines where astronomers normally saw only one, but the spacing of the lines was constantly changing. At times, the lines even became single. Pickering correctly deduced that the brighter component of Mizar, called Mizar A, is itself really two stars that revolve about each other in a period of 104 days. A star like Mizar A, which appears as a single star when photographed or observed visually through the telescope, but which spectroscopy shows really to be a double star, is called a spectroscopic binary.

Mizar, by the way, is a good example of just how complex such star systems can be. Mizar has been known for centuries to have a faint companion called Alcor, which can be seen without a telescope. Mizar and Alcor form an *optical double*—a pair of stars that appear close together in the sky but do not orbit each other. Through a telescope, as Riccioli discovered in 1650, Mizar can be seen to have another, closer companion that does orbit it; Mizar is thus a visual binary. The two components that make up this visual binary, known as Mizar A and Mizar B, are both spectroscopic binaries. So, Mizar is really a quadruple system of stars.

Strictly speaking, it is not correct to describe the motion of a binary star system by saying that one star orbits the other. Gravity is a *mutual* attraction. Each star exerts a gravitational force on the other, with the result that both stars orbit a point between them called the *center of mass*. Imagine that the two stars are seated at either end of a seesaw. The point at which the fulcrum would have to be located in order for the seesaw to balance is the center of mass, and it is always closer to the more massive star (Figure 18.5).

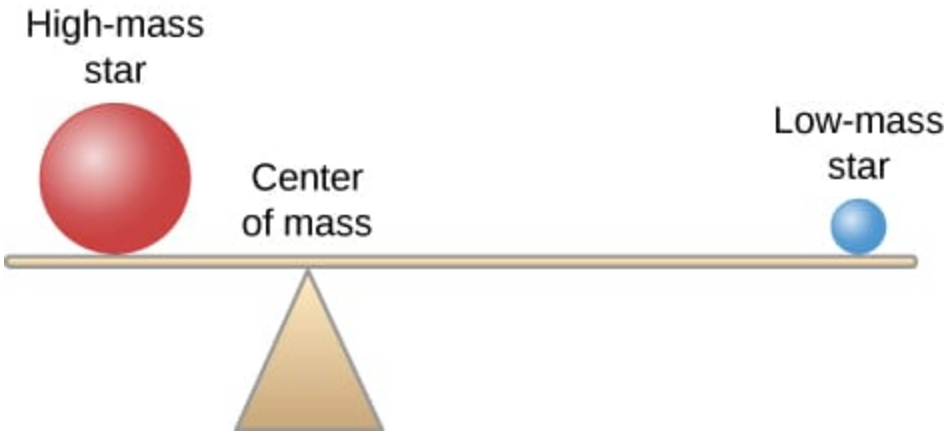


Figure 18.5 : Binary Star System. In a binary star system, both stars orbit their center of mass. The image shows the relative positions of two, different-mass stars from their center of mass, similar to how two masses would have to be located on a seesaw in order to keep it level. The star with the higher mass will be found closer to the center of mass, while the star with the lower mass will be farther from it.

Figure 18.6 shows two stars (A and B) moving around their center of mass, along with one line in the spectrum of each star that we observe from the system at different times. When one star is approaching us relative to the center of mass, the other star is receding from us. In the top left illustration, star A is moving toward us, so the line in its spectrum is Doppler-shifted toward the blue end of the spectrum. Star B is moving away from us, so its line shows a redshift. When we observe the composite spectrum of the two stars, the line appears double. When the two stars are both moving across our line of sight (neither away from nor toward us), they both have the same radial velocity (that of the pair's center of mass); hence, the spectral lines of the two stars come together. This is shown in the two bottom illustrations in Figure 18.6.

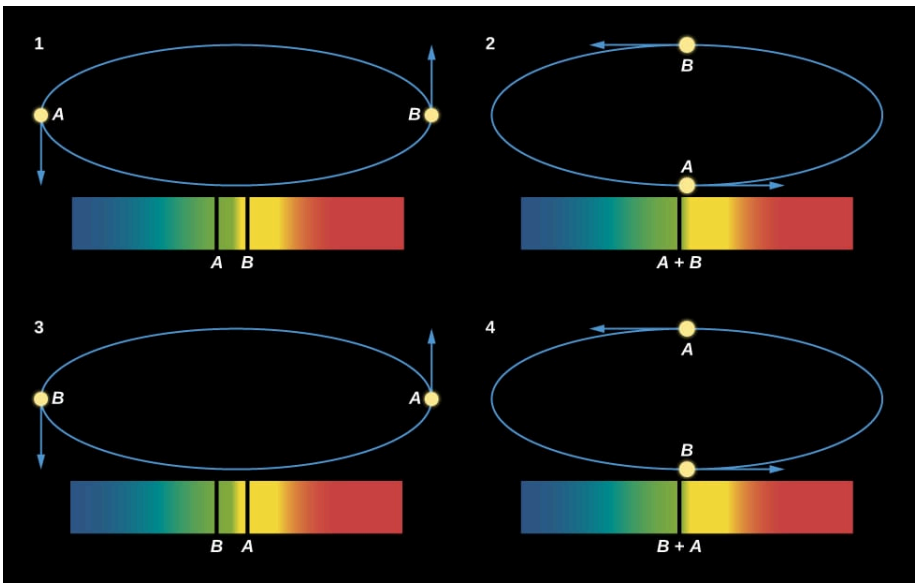


Figure 18.6 : Motions of Two Stars Orbiting Each Other and What the Spectrum Shows. We see changes in velocity because when one star is moving toward Earth, the other is moving away; half a cycle later, the situation is reversed. Doppler shifts cause the spectral lines to move back and forth. In diagrams 1 and 3, lines from both stars can be seen well separated from each other. When the two stars are moving perpendicular to our line of sight (that is, they are not moving either toward or away from us), the two lines are exactly superimposed, and so in diagrams 2 and 4, we see only a single spectral line. Note that in the diagrams, the orbit of the star pair is tipped slightly with respect to the viewer (or if the viewer were looking at it in the sky, the orbit would be tilted with respect to the viewer's line of sight). If the orbit were exactly in the plane of the page or screen (or the sky), then it would look nearly circular, but we would see no change in radial velocity (no part of the motion would be toward us or away from us.) If the orbit were perpendicular to the plane of the page or screen, then the stars would appear to move back and forth in a straight line, and we would see the largest-possible radial velocity variations. This diagram has been simplified for clarity. For more realistic situations, see the animation in the following Link to Learning feature box.

A plot showing how the velocities of the stars change with time is called a *radial velocity curve*; the curve for the binary system in Figure 18.6 is shown in Figure 18.7.

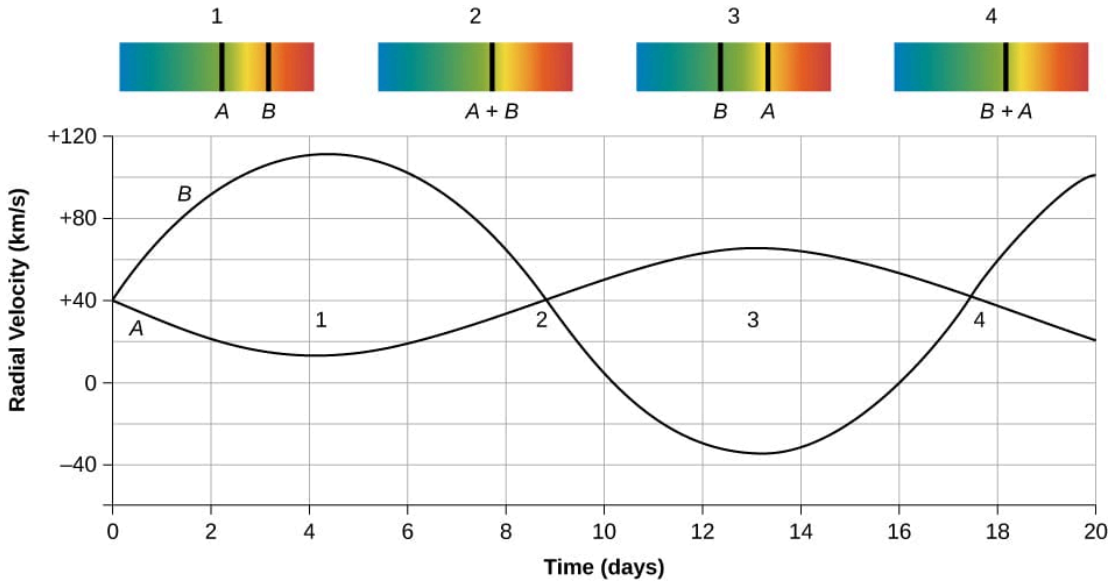


Figure 18.7 : Radial Velocities in a Spectroscopic Binary System. These curves plot the radial velocities of two stars in a spectroscopic binary system, showing how the stars alternately approach and recede from Earth. Note that positive velocity means the star is moving away from us, and negative velocity means the star is moving toward us. The center of mass of the system itself is also moving away from us, indicated by the positive velocity of 40 kilometers per second. The positions on the curve corresponding to the illustrations in Figure 18.6 are marked with the diagram number (1–4).

Link to Learning

This animation lets you follow the orbits of a binary star system in various combinations of the masses of the two stars.

Masses from the Orbits of Binary Stars

We can estimate the masses of binary star systems using Newton's reformulation of Kepler's third law (discussed in Newton's Universal Law of Gravitation). Kepler found that the time a planet takes to go around the Sun is related by a specific mathematical formula to its distance from the Sun. In our binary star situation, if two objects are in mutual revolution, then the period (*P*) with which they go around each other is related to the semimajor axis (*D*) of the orbit of one with respect to the other, according to this equation

$$D^3 = (M_1 + M_2)P^2$$

where *D* is in astronomical units, *P* is measured in years, and $M_1 + M_2$ is the sum of the masses of the two stars in units of the Sun's mass. This is a very useful formula for astronomers; it says that if we can observe the size of the orbit and the period of mutual revolution of the stars in a binary system, we can calculate the sum of their masses.

Most spectroscopic binaries have periods ranging from a few days to a few months, with separations of usually less than 1 AU between their member stars. Recall that an AU is the distance from Earth to the Sun, so this is a small separation and very hard to see at the distances of stars. This is why many of these systems are known to be double only through careful study of their spectra.

We can analyze a radial velocity curve (such as the one in Figure 18.7) to determine the masses of the stars in a spectroscopic binary. This is complex in practice but not hard in principle. We measure the speeds of the stars from the Doppler effect. We then determine the period—how long the stars take to go through an orbital cycle—from the velocity curve. Knowing how fast the stars are moving and how long they take to go around tells us the circumference of the orbit and, hence, the separation of the stars in kilometers or astronomical units. From Kepler's law, the period and the separation allow us to calculate the sum of the stars' masses.

Of course, knowing the sum of the masses is not as useful as knowing the mass of each star separately. But the relative orbital speeds of the two stars can tell us how much of the total mass each star has. As we saw in our seesaw analogy, the more massive

star is closer to the center of mass and therefore has a smaller orbit. Therefore, it moves more slowly to get around in the same time compared to the more distant, lower-mass star. If we sort out the speeds relative to each other, we can sort out the masses relative to each other. In practice, we also need to know how the binary system is oriented in the sky to our line of sight, but if we do, and the just-described steps are carried out carefully, the result is a calculation of the masses of each of the two stars in the system.

To summarize, a good measurement of the motion of two stars around a common center of mass, combined with the laws of gravity, allows us to determine the masses of stars in such systems. These mass measurements are absolutely crucial to developing a theory of how stars evolve. One of the best things about this method is that it is independent of the location of the binary system. It works as well for stars 100 light-years away from us as for those in our immediate neighborhood.

To take a specific example, Sirius is one of the few binary stars in Appendix J for which we have enough information to apply Kepler's third law:

$$D^3 = (M_1 + M_2)P^2$$

In this case, the two stars, the one we usually call Sirius and its very faint companion, are separated by about 20 AU and have an orbital period of about 50 years. If we place these values in the formula we would have

$$(20)^3 = (M_1 + M_2)(50)^2$$
$$8000 = (M_1 + M_2)(2500)$$

This can be solved for the sum of the masses:

$$M_1 + M_2 = \frac{8000}{2500} = 3.2$$

Therefore, the sum of masses of the two stars in the Sirius binary system is 3.2 times the Sun's mass. In order to determine the individual mass of each star, we would need the velocities of the two stars and the orientation of the orbit relative to our line of sight.

The Range of Stellar Masses

How large can the mass of a star be? Stars more massive than the Sun are rare. None of the stars within 30 light-years of the Sun has a mass greater than four times that of the Sun. Searches at large distances from the Sun have led to the discovery of a few stars with masses up to about 100 times that of the Sun, and a handful of stars (a few out of several billion) may have masses as large as 250 solar masses. However, most stars have less mass than the Sun.

According to theoretical calculations, the smallest mass that a true star can have is about 1/12 that of the Sun. By a "true" star, astronomers mean one that becomes hot enough to fuse protons to form helium (as discussed in The Sun: A Nuclear Powerhouse). Objects with masses between roughly 1/100 and 1/12 that of the Sun may produce energy for a brief time by means of nuclear reactions involving deuterium, but they do not become hot enough to fuse protons. Such objects are intermediate in mass between stars and planets and have been given the name brown dwarfs (Figure 18.8). Brown dwarfs are similar to Jupiter in radius but have masses from approximately 13 to 80 times larger than the mass of Jupiter.[2]



(a) (b)

Figure 18.8 : Brown Dwarfs in Orion. These images, taken with the Hubble Space Telescope, show the region surrounding the Trapezium star cluster inside the star-forming region called the Orion Nebula. (a) No brown dwarfs are seen in the visible light image, both because they put out very little light in the visible and because they are hidden within the clouds of dust in this region. (b) This image was taken in infrared light, which can make its way to us through the dust. The faintest objects in this image are brown dwarfs with masses between 13 and 80 times the mass of Jupiter. (credit a: NASA, C.R. O'Dell and S.K. Wong (Rice University); credit b: NASA; K.L. Luhman (Harvard-Smithsonian Center for Astrophysics) and G. Schneider, E. Young, G. Rieke, A. Cotera, H. Chen, M. Rieke, R. Thompson (Steward Observatory))

Still-smaller objects with masses less than about 1/100 the mass of the Sun (or 10 Jupiter masses) are called planets. They may radiate energy produced by the radioactive elements that they contain, and they may also radiate heat generated by slowly compressing under their own weight (a process called gravitational contraction). However, their interiors will never reach temperatures high enough for any nuclear reactions, to take place. Jupiter, whose mass is about 1/1000 the mass of the Sun, is unquestionably a planet, for example. Until the 1990s, we could only detect planets in our own solar system, but now we have thousands of them elsewhere as well. (We will discuss these exciting observations in The Birth of Stars and the Discovery of Planets outside the Solar System.)

The Mass-Luminosity Relation

Now that we have measurements of the characteristics of many different types of stars, we can search for relationships among the characteristics. For example, we can ask whether the mass and luminosity of a star are related. It turns out that for most stars, they are: The more massive stars are generally also the more luminous. This relationship, known as the mass-luminosity relation, is shown graphically in Figure 18.9. Each point represents a star whose mass and luminosity are both known. The horizontal position on the graph shows the star's mass, given in units of the Sun's mass, and the vertical position shows its luminosity in units of the Sun's luminosity.

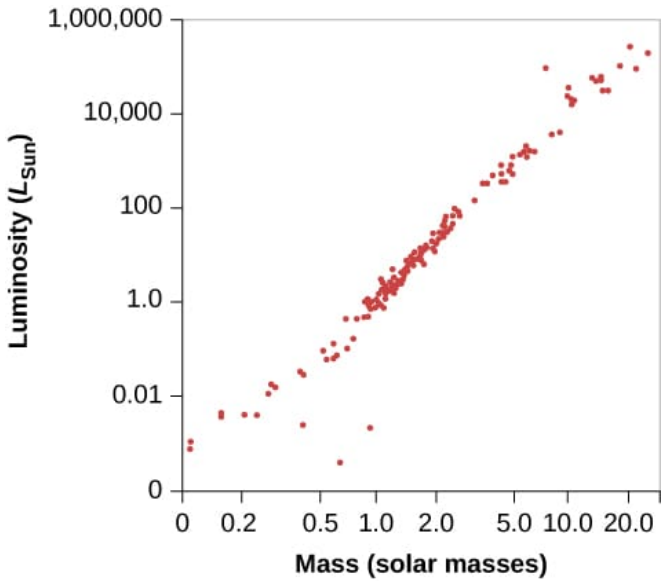


Figure 18.9 : Mass-Luminosity Relation. The plotted points show the masses and luminosities of stars. The three points lying below the sequence of points are all white dwarf stars.

We can also say this in mathematical terms.

$$L \sim M^{3.9}$$

It's a reasonably good approximation to say that luminosity (expressed in units of the Sun's luminosity) varies as the fourth power of the mass (in units of the Sun's mass). (The symbol ~ means the two quantities are proportional.) If two stars differ in mass by a factor of 2, then the more massive one will be 2^4, or about 16 times brighter; if one star is 1/3 the mass of another, it will be approximately 81 times less luminous.

Example 18.1: Calculating the Mass from the Luminosity of a Star

The mass-luminosity formula can be rewritten so that a value of mass can be determined if the luminosity is known.

Solution

First, we must get our units right by expressing both the mass and the luminosity of a star in units of the Sun's mass and luminosity:

$$L/L_{\text{Sun}} = (M/M_{\text{Sun}})^4$$

Now we can take the 4th root of both sides, which is equivalent to taking both sides to the 1/4 = 0.25 power. The formula in this case would be:

$$M/M_{\text{Sun}} = (L/L_{\text{Sun}})^{0.25} = (L/L_{\text{Sun}})^{0.25}$$

Exercise 20.4.1

In the previous section, we determined the sum of the masses of the two stars in the Sirius binary system (Sirius and its faint companion) using Kepler's third law to be 3.2 solar masses. Using the mass-luminosity relationship, calculate the mass of each individual star.

Answer available at the end of this section.

Notice how good this mass-luminosity relationship is. Most stars (see Figure 18.9) fall along a line running from the lower-left (low

mass, low luminosity) corner of the diagram to the upper-right (high mass, high luminosity) corner. About 90% of all stars obey the mass-luminosity relation. Later, we will explore why such a relationship exists and what we can learn from the roughly 10% of stars that “disobey” it.

Link to Learning

This visualization illustrates how all the properties of a star are determined by the mass that the star happened to form with and how these properties change when you change the mass.

Footnotes

- 2Exactly where to put the dividing line between planets and brown dwarfs is a subject of some debate among astronomers as we write this book (as is, in fact, the exact definition of each of these objects). Even those who accept deuterium fusion (see The Birth of Stars and the Discovery of Planets outside the Solar System) as the crucial issue for brown dwarfs concede that, depending on the composition of the star and other factors, the lowest mass for such a dwarf could be anywhere from 11 to 16 Jupiter masses.

Answers

Exercise $20.4.1$

In Appendix J, Sirius is listed with a luminosity 23 times that of the Sun. This value can be inserted into the mass-luminosity relationship to get the mass of Sirius: $M/M_{\text{Sun}} = 23^{0.25} = 2.2$
The mass of the companion star to Sirius is then $3.2 - 2.2 = 1.0$ solar mass.

This page titled 20.4: Measuring Stellar Masses is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 18.3: Measuring Stellar Masses by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

20.5: The H-R Diagram

Learning Objectives

By the end of this section, you will be able to:

- Identify the physical characteristics of stars that are used to create an H–R diagram, and describe how those characteristics vary among groups of stars
- Discuss the physical properties of most stars found at different locations on the H–R diagram, such as radius, and for main sequence stars, mass

In this chapter and Analyzing Starlight, we described some of the characteristics by which we might classify stars and how those are measured. These ideas are summarized in Table 18.2. We have also given an example of a relationship between two of these characteristics in the mass-luminosity relation. When the characteristics of large numbers of stars were measured at the beginning of the twentieth century, astronomers were able to begin a deeper search for patterns and relationships in these data.

Table 18.2: Measuring the Characteristics of Stars

Characteristic	Technique
Surface temperature	1. Determine the color (very rough). 2. Measure the spectrum and get the spectral type.
Chemical composition	Determine which lines are present in the spectrum.
Luminosity	Measure the apparent brightness and compensate for distance.
Radial velocity	Measure the Doppler shift in the spectrum.
Rotation	Measure the width of spectral lines.
Mass	Measure the period and radial velocity curves of spectroscopic binary stars.
Diameter	1. Measure the way a star's light is blocked by the Moon. 2. Measure the light curves and Doppler shifts for eclipsing binary stars.

To help understand what sorts of relationships might be found, let's look briefly at a range of data about human beings. If you want to understand humans by comparing and contrasting their characteristics—without assuming any previous knowledge of these strange creatures—you could try to determine which characteristics lead you in a fruitful direction. For example, you might plot the heights of a large sample of humans against their weights (which is a measure of their mass). Such a plot is shown in Figure 18.12 and it has some interesting features. In the way we have chosen to present our data, height increases upward, whereas weight increases to the left. Notice that humans are not randomly distributed in the graph. Most points fall along a sequence that goes from the upper left to the lower right.

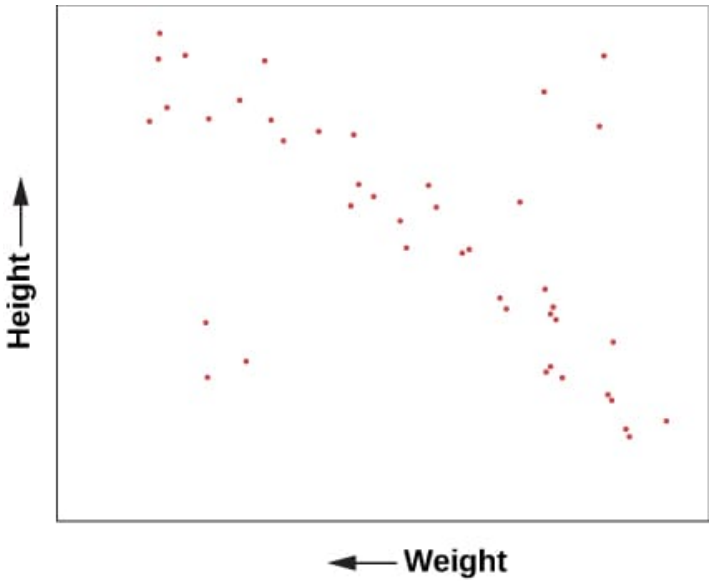


Figure 18.12 : Height versus Weight. The plot of the heights and weights of a representative group of human beings. Most points lie along a "main sequence" representing most people, but there are a few exceptions.

We can conclude from this graph that human height and weight are related. Generally speaking, taller human beings weigh more, whereas shorter ones weigh less. This makes sense if you are familiar with the structure of human beings. Typically, if we have bigger bones, we have more flesh to fill out our larger frame. It's not mathematically exact—there is a wide range of variation—but it's not a bad overall rule. And, of course, there are some dramatic exceptions. You occasionally see a short human who is very overweight and would thus be more to the bottom left of our diagram than the average sequence of people. Or you might have a very tall, skinny fashion model with great height but relatively small weight, who would be found near the upper right.

A similar diagram has been found extremely useful for understanding the lives of stars. In 1913, American astronomer Henry Norris Russell plotted the luminosities of stars against their spectral classes (a way of denoting their surface temperatures). This investigation, and a similar independent study in 1911 by Danish astronomer Ejnar Hertzsprung, led to the extremely important discovery that the temperature and luminosity of stars are related (Figure 18.13).

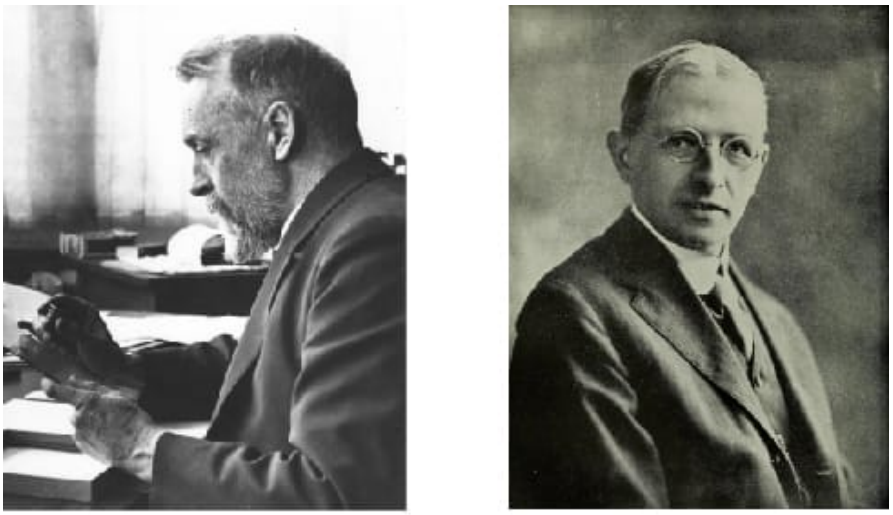


Figure 18.13 : Hertzsprung (1873–1967) and Russell (1877–1957). (a) Ejnar Hertzsprung and (b) Henry Norris Russell independently discovered the relationship between the luminosity and surface temperature of stars that is summarized in what is now called the H–R diagram.

Voyagers in Astronomy

Henry Norris Russell

When Henry Norris Russell graduated from Princeton University, his work had been so brilliant that the faculty decided to create a new level of honors degree beyond “summa cum laude” for him. His students later remembered him as a man whose thinking was three times faster than just about anybody else’s. His memory was so phenomenal, he could correctly quote an enormous number of poems and limericks, the entire Bible, tables of mathematical functions, and almost anything he had learned about astronomy. He was nervous, active, competitive, critical, and very articulate; he tended to dominate every meeting he attended. In outward appearance, he was an old-fashioned product of the nineteenth century who wore high-top black shoes and high starched collars, and carried an umbrella every day of his life. His 264 papers were enormously influential in many areas of astronomy.

Born in 1877, the son of a Presbyterian minister, Russell showed early promise. When he was 12, his family sent him to live with an aunt in Princeton so he could attend a top preparatory school. He lived in the same house in that town until his death in 1957 (interrupted only by a brief stay in Europe for graduate work). He was fond of recounting that both his mother and his maternal grandmother had won prizes in mathematics, and that he probably inherited his talents in that field from their side of the family.

Before Russell, American astronomers devoted themselves mainly to surveying the stars and making impressive catalogs of their properties, especially their spectra (as described in Analyzing Starlight. Russell began to see that interpreting the spectra of stars required a much more sophisticated understanding of the physics of the atom, a subject that was being developed by European physicists in the 1910s and 1920s. Russell embarked on a lifelong quest to ascertain the physical conditions inside stars from the clues in their spectra; his work inspired, and was continued by, a generation of astronomers, many trained by Russell and his collaborators.

Russell also made important contributions in the study of binary stars and the measurement of star masses, the origin of the solar system, the atmospheres of planets, and the measurement of distances in astronomy, among other fields. He was an influential teacher and popularizer of astronomy, writing a column on astronomical topics for *Scientific American* magazine for more than 40 years. He and two colleagues wrote a textbook for college astronomy classes that helped train astronomers and astronomy enthusiasts over several decades. That book set the scene for the kind of textbook you are now reading, which not only lays out the facts of astronomy but also explains how they fit together. Russell gave lectures around the country, often emphasizing the importance of understanding modern physics in order to grasp what was happening in astronomy.

Harlow Shapley, director of the Harvard College Observatory, called Russell “the dean of American astronomers.” Russell was certainly regarded as the leader of the field for many years and was consulted on many astronomical problems by colleagues from around the world. Today, one of the highest recognitions that an astronomer can receive is an award from the American Astronomical Society called the Russell Prize, set up in his memory.

Features of the H–R Diagram

Following Hertzsprung and Russell, let us plot the temperature (or spectral class) of a selected group of nearby stars against their luminosity and see what we find (Figure 18.14). Such a plot is frequently called the *Hertzsprung–Russell diagram*, abbreviated H–R diagram. It is one of the most important and widely used diagrams in astronomy, with applications that extend far beyond the purposes for which it was originally developed more than a century ago.

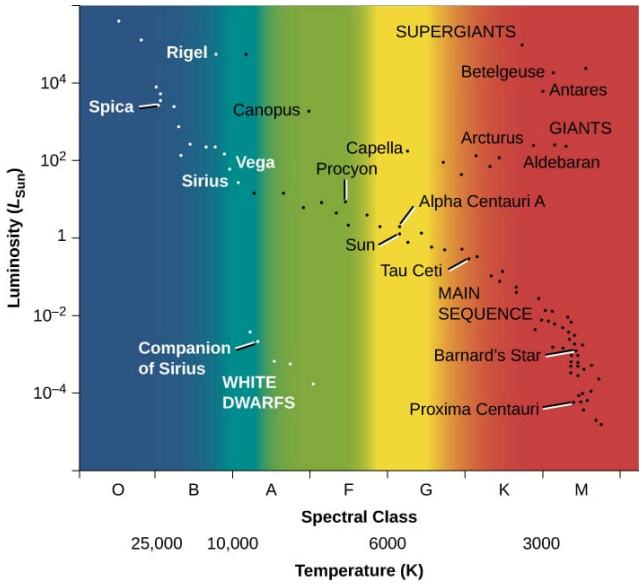


Figure 18.14 : H–R Diagram for a Selected Sample of Stars. In such diagrams, luminosity is plotted along the vertical axis. Along the horizontal axis, we can plot either temperature or spectral type (also sometimes called spectral class). Several of the brightest stars are identified by name. Most stars fall on the main sequence.

It is customary to plot H–R diagrams in such a way that temperature increases toward the left and luminosity toward the top. Notice the similarity to our plot of height and weight for people (Figure 18.12). Stars, like people, are not distributed over the diagram at random, as they would be if they exhibited all combinations of luminosity and temperature. Instead, we see that the stars cluster into certain parts of the H–R diagram. The great majority are aligned along a narrow sequence running from the upper left (hot, highly luminous) to the lower right (cool, less luminous). This band of points is called the main sequence. It represents a relationship between *temperature* and *luminosity* that is followed by most stars. We can summarize this relationship by saying that hotter stars are more luminous than cooler ones.

A number of stars, however, lie above the main sequence on the H–R diagram, in the upper-right region, where stars have low temperature and high luminosity. How can a star be at once cool, meaning each square meter on the star does not put out all that much energy, and yet very luminous? The only way is for the star to be enormous—to have so many square meters on its surface that the *total* energy output is still large. These stars must be *giants* or *supergiants*, the stars of huge diameter we discussed earlier.

There are also some stars in the lower-left corner of the diagram, which have high temperature and low luminosity. If they have high surface temperatures, each square meter on that star puts out a lot of energy. How then can the overall star be dim? It must be that it has a very small total surface area; such stars are known as white dwarfs (white because, at these high temperatures, the colors of the electromagnetic radiation that they emit blend together to make them look bluish-white). We will say more about these puzzling objects in a moment. Figure 18.15 is a schematic H–R diagram for a large sample of stars, drawn to make the different types more apparent.

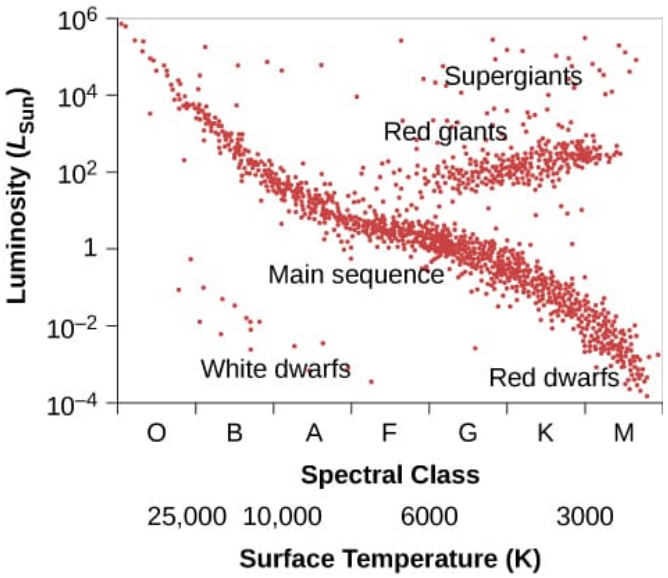


Figure 18.15 : Schematic H–R Diagram for Many Stars. Ninety percent of all stars on such a diagram fall along a narrow band called the main sequence. A minority of stars are found in the upper right; they are both cool (and hence red) and bright, and must be giants. Some stars fall in the lower left of the diagram; they are both hot and dim, and must be white dwarfs.

Now, think back to our discussion of star surveys. It is difficult to plot an H–R diagram that is truly representative of all stars because most stars are so faint that we cannot see those outside our immediate neighborhood. The stars plotted in Figure 18.14 were selected because their distances are known. This sample omits many intrinsically faint stars that are nearby but have not had their distances measured, so it shows fewer faint main-sequence stars than a "fair" diagram would. To be truly representative of the stellar population, an H–R diagram should be plotted for all stars within a certain distance. Unfortunately, our knowledge is reasonably complete only for stars within 10 to 20 light-years of the Sun, among which there are no giants or supergiants. Still, from many surveys (and more can now be done with new, more powerful telescopes), we estimate that about 90% of the true stars overall (excluding brown dwarfs) in our part of space are main-sequence stars, about 10% are white dwarfs, and fewer than 1% are giants or supergiants.

These estimates can be used directly to understand the lives of stars. Permit us another quick analogy with people. Suppose we survey people just like astronomers survey stars, but we want to focus our attention on the location of young people, ages 6 to 18 years. Survey teams fan out and take data about where such youngsters are found at all times during a 24-hour day. Some are found in the local pizza parlor, others are asleep at home, some are at the movies, and many are in school. After surveying a very large number of young people, one of the things that the teams determine is that, averaged over the course of the 24 hours, one-third of all youngsters are found in school.

How can they interpret this result? Does it mean that two-thirds of students are truants and the remaining one-third spend all their time in school? No, we must bear in mind that the survey teams counted youngsters throughout the full 24-hour day. Some survey teams worked at night, when most youngsters were at home asleep, and others worked in the late afternoon, when most youngsters were on their way home from school (and more likely to be enjoying a pizza). If the survey was truly representative, we *can* conclude, however, that if an average of one-third of all youngsters are found in school, then humans ages 6 to 18 years must spend about one-third of *their time* in school.

We can do something similar for stars. We find that, on average, 90% of all stars are located on the main sequence of the H–R diagram. If we can identify some activity or life stage with the main sequence, then it follows that stars must spend 90% of their lives in that activity or life stage.

Understanding the Main Sequence

In The Sun: A Nuclear Powerhouse, we discussed the Sun as a representative star. We saw that what stars such as the Sun "do for a living" is to convert protons into helium deep in their interiors via the process of nuclear fusion, thus producing energy. The fusion of protons to helium is an excellent, long-lasting source of energy for a star because the bulk of every star consists of hydrogen atoms, whose nuclei are protons.

Our computer models of how stars evolve over time show us that a typical star will spend about 90% of its life fusing the abundant hydrogen in its core into helium. This then is a good explanation of why 90% of all stars are found on the main sequence in the H–R diagram. But if all the stars on the main sequence are doing the same thing (fusing hydrogen), why are they distributed along a sequence of points? That is, why do they differ in luminosity and surface temperature (which is what we are plotting on the H–R diagram)?

To help us understand how main-sequence stars differ, we can use one of the most important results from our studies of model stars. Astrophysicists have been able to show that the structure of stars that are in equilibrium and derive all their energy from nuclear fusion is completely and uniquely determined by just two quantities: the *total mass* and the *composition* of the star. This fact provides an interpretation of many features of the H–R diagram.

Imagine a cluster of stars forming from a cloud of interstellar "raw material" whose chemical composition is similar to the Sun's. (We'll describe this process in more detail in The Birth of Stars and Discovery of Planets outside the Solar System, but for now, the details will not concern us.) In such a cloud, all the clumps of gas and dust that become stars begin with the same chemical composition and differ from one another only in mass. Now suppose that we compute a model of each of these stars for the time at which it becomes stable and derives its energy from nuclear reactions, but before it has time to alter its composition appreciably as a result of these reactions.

The models calculated for these stars allow us to determine their luminosities, temperatures, and sizes. If we plot the results from the models—one point for each model star—on the H–R diagram, we get something that looks just like the main sequence we saw for real stars.

And here is what we find when we do this. The model stars with the largest masses are the hottest and most luminous, and they are located at the upper left of the diagram.

The least-massive model stars are the coolest and least luminous, and they are placed at the lower right of the plot. The other model stars all lie along a line running diagonally across the diagram. In other words, *the main sequence turns out to be a sequence of stellar masses*.

This makes sense if you think about it. The most massive stars have the most gravity and can thus compress their centers to the greatest degree. This means they are the hottest inside and the best at generating energy from nuclear reactions deep within. As a result, they shine with the greatest luminosity and have the hottest surface temperatures. The stars with lowest mass, in turn, are the coolest inside and least effective in generating energy. Thus, they are the least luminous and wind up being the coolest on the surface. Our Sun lies somewhere in the middle of these extremes (as you can see in Figure 18.14). The characteristics of representative main-sequence stars (excluding brown dwarfs, which are not true stars) are listed in Table 18.3.

Table 18.3: Characteristics of Main-Sequence Stars

Spectral Type	Mass (Sun = 1)	Luminosity (Sun = 1)	Temperature	Radius (Sun = 1)
O5	40	7×10^5	40,000 K	18
B0	16	2.7×10^5	28,000 K	7
A0	3.3	55	10,000 K	2.5
F0	1.7	5	7500 K	1.4
G0	1.1	1.4	6000 K	1.1
K0	0.8	0.35	5000 K	0.8
M0	0.4	0.05	3500 K	0.6

Note that we've seen this 90% figure come up before. This is exactly what we found earlier when we examined the mass-luminosity relation (Figure 18.9). We observed that 90% of all stars seem to follow the relationship; these are the 90% of all stars that lie on the main sequence in our H–R diagram. Our models and our observations agree.

What about the other stars on the H–R diagram—the giants and supergiants, and the white dwarfs? As we will see in the next few chapters, these are what main-sequence stars turn into as they age: They are the later stages in a star's life. As a star consumes its nuclear fuel, its source of energy changes, as do its chemical composition and interior structure. These changes cause the star to alter its luminosity and surface temperature so that it no longer lies on the main sequence on our diagram. Because stars

spend much less time in these later stages of their lives, we see fewer stars in those regions of the H–R diagram.

Extremes of Stellar Luminosities, Diameters, and Densities

We can use the H–R diagram to explore the extremes in size, luminosity, and density found among the stars. Such extreme stars are not only interesting to fans of the *Guinness Book of World Records*; they can teach us a lot about how stars work. For example, we saw that the most massive main-sequence stars are the most luminous ones. We know of a few extreme stars that are a million times more luminous than the Sun, with masses that exceed 100 times the Sun's mass. These superluminous stars, which are at the upper left of the H–R diagram, are exceedingly hot, very blue stars of spectral type O. These are the stars that would be the most conspicuous at vast distances in space.

The cool supergiants in the upper corner of the H–R diagram are as much as 10,000 times as luminous as the Sun. In addition, these stars have diameters very much larger than that of the Sun. As discussed above, some supergiants are so large that if the solar system could be centered in one, the star's surface would lie beyond the orbit of Mars (see Figure 18.16). We will have to ask, in coming chapters, what process can make a star swell up to such an enormous size, and how long these "swollen" stars can last in their distended state.

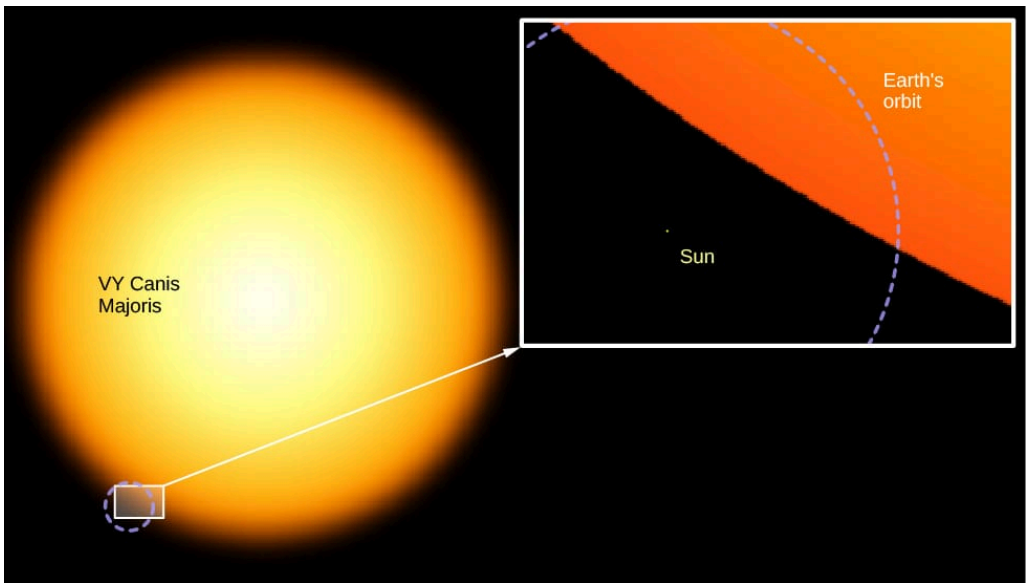


Figure 18.16 : The Sun and a Supergiant. Here you see how small the Sun looks in comparison to one of the largest known stars: VY Canis Majoris, a supergiant.

In contrast, the very common red, cool, low-luminosity stars at the lower end of the main sequence are much smaller and more compact than the Sun. An example of such a red dwarf is Ross 614B, with a surface temperature of 2700 K and only 1/2000 of the Sun's luminosity. We call such a star a dwarf because its diameter is only 1/10 that of the Sun. A star with such a low luminosity also has a low mass (about 1/12 that of the Sun). This combination of mass and diameter means that it is so compressed that the star has an average density about 80 times that of the Sun. Its density must be higher, in fact, than that of any known solid found on the surface of Earth. (Despite this, the star is made of gas throughout because its center is so hot.)

The faint, red, main-sequence stars are not the stars of the most extreme densities, however. The white dwarfs, at the lower-left corner of the H–R diagram, have densities many times greater still.

The White Dwarfs

The first white dwarf star was detected in 1862. Called Sirius B, it forms a binary system with Sirius A, the brightest-appearing star in the sky. It eluded discovery and analysis for a long time because its faint light tends to be lost in the glare of nearby Sirius A (Figure 18.17). (Since Sirius is often called the Dog Star—being the brightest star in the constellation of Canis Major, the big dog—Sirius B is sometimes nicknamed the Pup.)

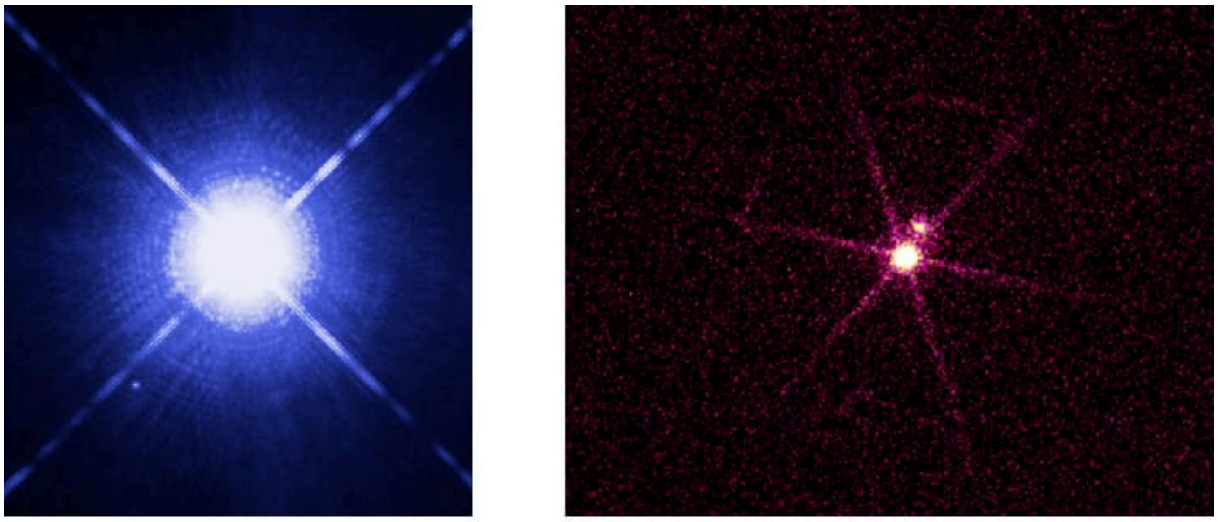


Figure 18.17 : Two Views of Sirius and Its White Dwarf Companion. (a) The (visible light) image, taken with the Hubble Space Telescope, shows bright Sirius A, and, below it and off to its left, faint Sirius B. (b) This image of the Sirius star system was taken with the Chandra X-Ray Telescope. Now, the bright object is the white dwarf companion, Sirius B. Sirius A is the faint object above it; what we are seeing from Sirius is probably not actually X-ray radiation but rather ultraviolet light that has leaked into the detector. Note that the ultraviolet intensities of these two objects are completely reversed from the situation in visible light because Sirius B is hotter and emits more higher-frequency radiation. (credit a: modification of work by NASA, H.E. Bond and E. Nelan (Space Telescope Science Institute), M. Barstow and M. Burleigh (University of Leicester) and J.B. Holberg (University of Arizona); credit b: modification of work by NASA/SAO/CXC)

We have now found thousands of white dwarfs. Table 18.1 shows that about 7% of the true stars (spectral types O–M) in our local neighborhood are white dwarfs. A good example of a typical white dwarf is the nearby star 40 Eridani B. Its surface temperature is a relatively hot 12,000 K, but its luminosity is only 1/275 L_{Sun}. Calculations show that its radius is only 1.4% of the Sun's, or about the same as that of Earth, and its volume is 2.5×10^{-6} that of the Sun. Its mass, however, is 0.57 times the Sun's mass, just a little more than half. To fit such a substantial mass into so tiny a volume, the star's density must be about 210,000 times the density of the Sun, or more than 300,000 g/cm^3. A teaspoonful of this material would have a mass of some 1.6 tons! At such enormous densities, matter cannot exist in its usual state; we will examine the particular behavior of this type of matter in The Death of Stars. For now, we just note that white dwarfs are dying stars, reaching the end of their productive lives and ready for their stories to be over.

The British astrophysicist (and science popularizer) Arthur Eddington (1882–1944) described the first known white dwarf this way:

> The message of the companion of Sirius, when decoded, ran: "I am composed of material three thousand times denser than anything you've ever come across. A ton of my material would be a little nugget you could put in a matchbox." What reply could one make to something like that? Well, the reply most of us made in 1914 was, "Shut up; don't talk nonsense."

Today, however, astronomers not only accept that stars as dense as white dwarfs exist but (as we will see) have found even denser and stranger objects in their quest to understand the evolution of different types of stars.

Link to Learning

Use this interactive H-R diagram to explore the sizes of stars based on their placement. Overlay the closest stars and the brightest stars, and consider where they lie on the figure. How many stars are among both the closest and brightest stars?

This page titled 20.5: The H-R Diagram is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 18.5: The H-R Diagram by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

21: EXOPLANETS

21.1: Planets beyond the Solar System- Search and Discovery
21.2: Exoplanets Everywhere - What we are Learning

This page titled 21: Exoplanets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

21.1: Planets beyond the Solar System- Search and Discovery

Learning Objectives

By the end of this section, you will be able to:

- Describe the orbital motion of planets in our solar system using Kepler's laws
- Compare the indirect and direct observational techniques for exoplanet detection

For centuries, astronomers have dreamed of finding planets around other stars, including other planets like Earth. Direct observations of such distant planets are very difficult, however. You might compare a planet orbiting a star to a mosquito flying around one of those giant spotlights at a shopping center opening. From close up, you might spot the mosquito. But imagine viewing the scene from some distance away—say, from an airplane. You could see the spotlight just fine, but what are your chances of catching the mosquito in that light? Instead of making direct images, astronomers have relied on indirect observations and have now succeeded in detecting a multitude of planets around other stars.

In 1995, after decades of effort, we found the first such exoplanet (a planet outside our solar system) orbiting a main-sequence star, and today we know that most stars form with planets. This is an example of how persistence and new methods of observation advance the knowledge of humanity. By studying exoplanets, astronomers hope to better understand our solar system in context of the rest of the universe. For instance, how does the arrangement of our solar system compare to planetary systems in the rest of the universe? What do exoplanets tell us about the process of planet formation? And how does knowing the frequency of exoplanets influence our estimates of whether there is life elsewhere?

Searching for Orbital Motion

Most exoplanet detections are made using techniques where we observe the *effect* that the planet exerts on the host star. For example, the gravitational tug of an unseen planet will cause a small wobble in the host star. Or, if its orbit is properly aligned, a planet will periodically cross in front of the star, causing the brightness of the star to dim.

To understand how a planet can move its host star, consider a single Jupiter-like planet. Both the planet and the star actually revolve about their *common center of mass*. Remember that gravity is a mutual attraction. The star and the planet each exert a force on the other, and we can find a stable point, the center of mass, between them about which both objects move. The smaller the mass of a body in such a system, the larger its orbit. A massive star barely swings around the center of mass, while a low-mass planet makes a much larger "tour."

Suppose the planet is like Jupiter and has a mass about one-thousandth that of its star; in this case, the size of the star's orbit is one-thousandth the size of the planet's. To get a sense of how difficult observing such motion might be, let's see how hard Jupiter would be to detect in this way from the distance of a nearby star. Consider an alien astronomer trying to observe our own system from Alpha Centauri, the closest star system to our own (about 4.3 light-years away). There are two ways this astronomer could try to detect the orbital motion of the Sun. One way would be to look for changes in the Sun's position on the sky. The second would be to use the Doppler effect to look for changes in its velocity. Let's discuss each of these in turn.

The diameter of Jupiter's apparent orbit viewed from Alpha Centauri is 1$0$ seconds of arc, and that of the Sun's orbit is $0.0$1$0$ seconds of arc. (Remember, 1 second of arc is 1/3600 degree.) If they could measure the apparent position of the Sun (which is bright and easy to detect) to sufficient precision, they would describe an orbit of diameter $0.0$1$0$ seconds of arc with a period equal to that of Jupiter, which is 12 years.

In other words, if they watched the Sun for 12 years, they would see it wiggle back and forth in the sky by this minuscule fraction of a degree. From the observed motion and the period of the "wiggle," they could deduce the mass of Jupiter and its distance using Kepler's laws. (To refresh your memory about these laws, see the chapter on Orbits and Gravity.)

Measuring positions in the sky this accurately is extremely difficult, and so far, astronomers have not made any confirmed detections of planets using this technique. However, we have been successful in using spectrometers to measure the changing velocity of stars with planets around them.

As the star and planet orbit each other, part of their motion will be in our line of sight (toward us or away from us). Such motion can be measured using the *Doppler effect* and the star's spectrum. As the star moves back and forth in orbit around the system's center of mass in response to the gravitational tug of an orbiting planet, the lines in its spectrum will shift back and forth.

Let's again consider the example of the Sun. Its *radial velocity* (motion toward or away from us) changes by about 13 meters per second with a period of 12 years because of the gravitational pull of Jupiter. This corresponds to about 30 miles per hour, roughly the speed at which many of us drive around town. Detecting motion at this level in a star's spectrum presents an enormous technical challenge, but several groups of astronomers around the world, using specialized spectrographs designed for this purpose, have succeeded. Note that the change in speed does not depend on the distance of the star from the observer. Using the Doppler effect to detect planets will work at any distance, as long as the star is bright enough to provide a good spectrum and a large telescope is available to make the observations (Figure 21.16).

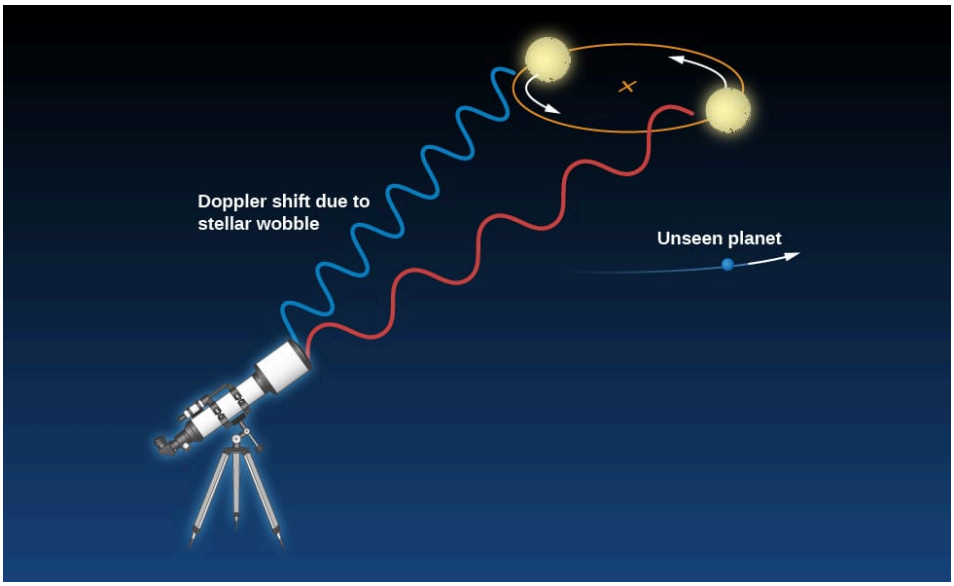


Figure 21.16 : Doppler Method of Detecting Planets. The motion of a star around a common center of mass with an orbiting planet can be detected by measuring the changing speed of the star. When the star is moving away from us, the lines in its spectrum show a tiny redshift; when it is moving toward us, they show a tiny blueshift. The change in color (wavelength) has been exaggerated here for illustrative purposes. In reality, the Doppler shifts we measure are extremely small and require sophisticated equipment to be detected.

The first successful use of the Doppler effect to find a planet around another star was in 1995. Michel Mayor and Didier Queloz of the Geneva Observatory (Figure 21.17) used this technique to find a planet orbiting a star resembling our Sun called 51 Pegasi, about 40 light-years away. (The star can be found in the sky near the great square of Pegasus, the flying horse of Greek mythology, one of the easiest-to-find star patterns.) To everyone's surprise, the planet takes a mere 4.2 days to orbit around the star. (Remember that Mercury, the innermost planet in our solar system, takes 88 days to go once around the Sun, so 4.2 days seems fantastically short.)

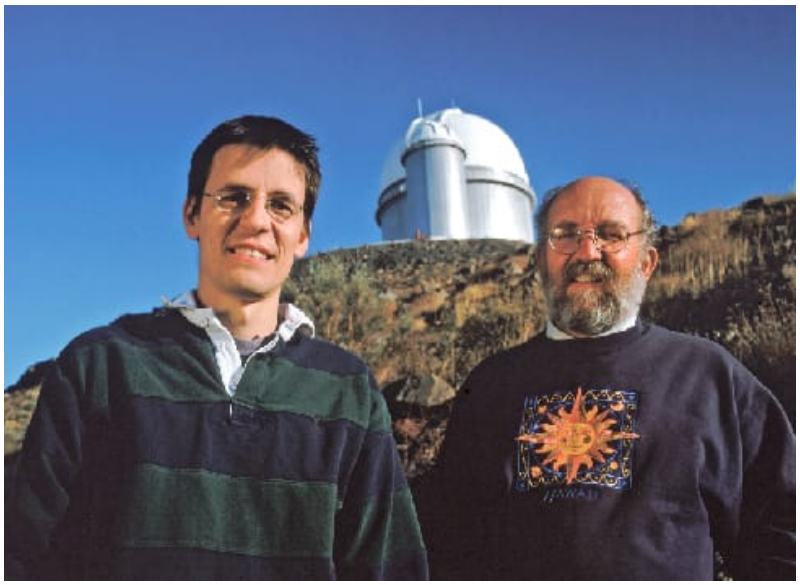


Figure 21.17 : Planet Discoverers. In 1995, Didier Queloz and Michel Mayor of the Geneva Observatory were the first to discover a planet around a regular star (51 Pegasi), for which they received the 2019 Nobel Prize in physics. They are seen here at an observatory in Chile where they are continuing their planet hunting. (credit: Weinstein/Ciel et Espace Photos)

Mayor and Queloz's findings mean the planet must be very close to 51 Pegasi, circling it about 7 million kilometers away (Figure 21.18). At that distance, the energy of the star should heat the planet's surface to a temperature of a few thousand degrees Celsius (a bit hot for future tourism). From its motion, astronomers calculate that it has at least half the mass of Jupiter[1], making it clearly a jovian and not a terrestrial-type planet.

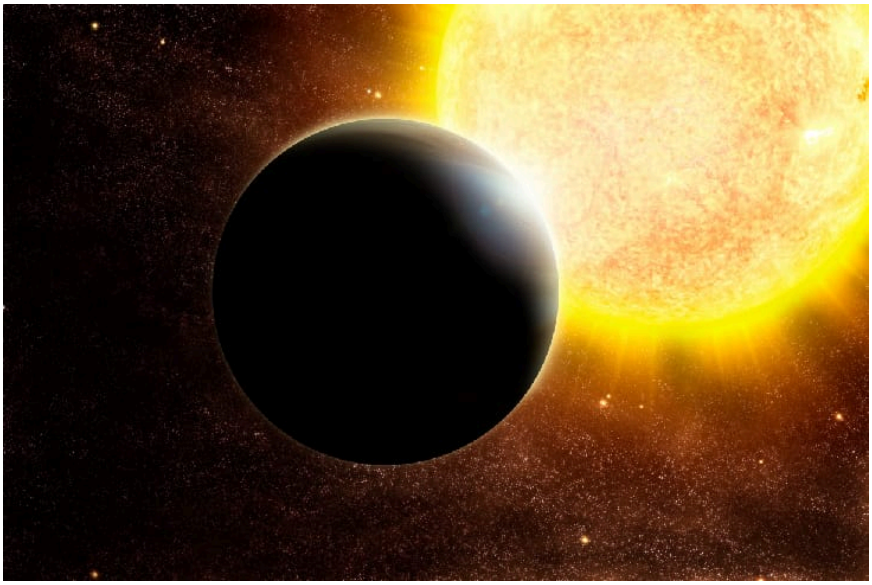


Figure 21.18 : Hot Jupiter. Artist Greg Bacon painted this impression of a hot, Jupiter-type planet orbiting close to a sunlike star. The artist shows bands on the planet like Jupiter, but we only estimate the mass of most hot, Jupiter-type planets from the Doppler method and don't know what conditions on the planet are like. (credit: ESO)

Since that initial planet discovery, the rate of progress has been breathtaking. Almost a thousand giant planets have been discovered using the Doppler technique. Many of these giant planets are orbiting close to their stars—astronomers have called these *hot Jupiters*.

The existence of giant planets so close to their stars was a surprise, and these discoveries have forced us to rethink our ideas about how planetary systems form. But for now, bear in mind that the Doppler-shift method—which relies on the pull of a planet making its star "wiggle" back and forth around the center of mass—is most effective at finding planets that are both close to their stars and massive. These planets cause the biggest "wiggles" in the motion of their stars and the biggest Doppler shifts in the

spectrum. Plus, they will be found sooner, since astronomers like to monitor the star for at least one full orbit (and perhaps more) and hot Jupiters take the shortest time to complete their orbit.

So if such planets exist, we would expect to be finding this type first. Scientists call this a *selection effect*—where our technique of discovery selects certain kinds of objects as "easy finds." As an example of a selection effect in everyday life, imagine you decide you are ready for a new romantic relationship in your life. To begin with, you only attend social events on campus, all of which require a student ID to get in. Your selection of possible partners will then be limited to students at your college. That may not give you as diverse a group to choose from as you want. In the same way, when we first used the Doppler technique, it selected massive planets close to their stars as the most likely discoveries. As we spend longer times watching target stars and as our ability to measure smaller Doppler shifts improves, this technique can reveal more distant and less massive planets too.

Link to Learning

View a series of animations demonstrating solar system motion and Kepler's laws, and select animation 1 (Kepler's laws) from the dropdown playlist. To view an animation demonstrating the radial velocity curve for an exoplanet, select animation 29 (radial velocity curve for an exoplanet) and animation 30 (radial velocity curve for an exoplanet—elliptical orbit) from the dropdown playlist.

Transiting Planets

The second method for indirect detection of exoplanets is based not on the motion of the star but on its brightness. When the orbital plane of the planet is tilted or inclined so that it is viewed edge-on, we will see the planet cross in front of the star once per orbit, causing the star to dim slightly; this event is known as transit. Figure 21.19 shows a sketch of the transit at three time steps: (1) out of transit, (2) the start of transit, and (3) full transit, along with a sketch of the light curve, which shows the drop in the brightness of the host star. The amount of light blocked—the depth of the transit—depends on the area of the planet (its size) compared to the star. If we can determine the size of the star, the transit method tells us the size of the planet.

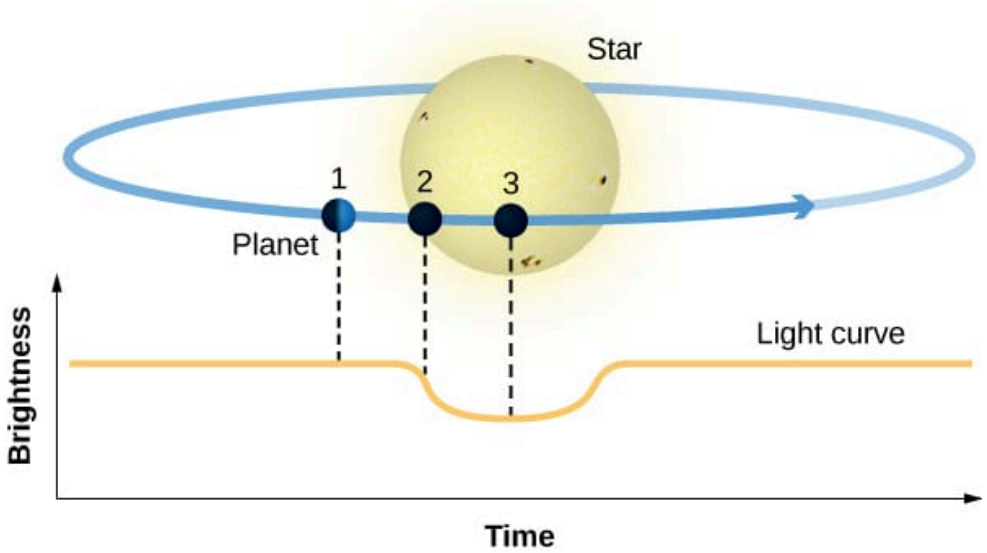


Figure 21.19 : Planet Transits. As the planet transits, it blocks out some of the light from the star, causing a temporary dimming in the brightness of the star. The top figure shows three moments during the transit event and the bottom panel shows the corresponding light curve: (1) out of transit, (2) transit ingress, and (3) the full drop in brightness.

The interval between successive transits is the length of the year for that planet, which can be used (again using Kepler's laws) to find its distance from the star. Larger planets like Jupiter block out more starlight than small earthlike planets, making transits by giant planets easier to detect, even from ground-based observatories. But by going into space, above the distorting effects of Earth's atmosphere, the transit technique has been extended to exoplanets as small as Mars.

Example 21.1: Transit Depth

In a transit, the planet's circular disk blocks the light of the star's circular disk. The area of a circle is πR^2. The amount of light the planet blocks, called the transit depth, is then given by

$$\frac{\pi R^2_{\text{planet}}}{\pi R^2_{\text{star}}} = \frac{R^2_{\text{planet}}}{R^2_{\text{star}}} = \left(\frac{R_{\text{planet}}}{R_{\text{star}}}\right)^2$$

Now calculate the transit depth for a star the size of the Sun with a gas giant planet the size of Jupiter.

Solution

The radius of Jupiter is 71,400 km, while the radius of the Sun is 695,700 km. Substituting into the equation, we get $\left(\frac{R_{\text{planet}}}{R_{\text{star}}}\right)^2 = \left(\frac{71{,}400 \text{ km}}{695{,}700 \text{ km}}\right)^2 = 0.01$ or 1%, which can easily be detected with the instruments on board the Kepler spacecraft.

Exercise 21.1.1

What is the transit depth for a star half the size of the Sun with a much smaller planet, like the size of Earth?

Answer available at the end of this section.

The Doppler method allows us to estimate the mass of a planet. If the same object can be studied by both the Doppler and transit techniques, we can measure both the mass and the size of the exoplanet. This is a powerful combination that can be used to derive the average density (mass/volume) of the planet. In 1999, using measurements from ground-based telescopes, the first transiting planet was detected orbiting the star HD 209458. The planet transits its parent star for about 3 hours every 3.5 days as we view it from Earth. Doppler measurements showed that the planet around HD 209458 has about 70% the mass of Jupiter, but its radius is about 35% larger than Jupiter's. This was the first case where we could determine what an exoplanet was made of—with that mass and radius, HD 209458 must be a gas and liquid world like Jupiter or Saturn.

It is even possible to learn something about the planet's atmosphere. When the planet passes in front of HD 209458, the atoms in the planet's atmosphere absorb starlight. Observations of this absorption were first made at the wavelengths of yellow sodium lines and showed that the atmosphere of the planet contains sodium; now, other elements can be measured as well.

Link to Learning

Try a transit simulator that demonstrates how a planet passing in front of its parent star can lead to the planet's detection. Follow the instructions to run the animation on your computer.

Transiting planets reveal such a wealth of information that the French Space Agency (CNES) and the European Space Agency (ESA) launched the CoRoT space telescope in 2007 to detect transiting exoplanets. CoRoT discovered 32 transiting exoplanets, including the first transiting planet with a size and density similar to Earth. In 2012, the spacecraft suffered an onboard computer failure, ending the mission. Meanwhile, NASA built a much more powerful transit observatory called Kepler.

In 2009, NASA launched the Kepler space telescope, dedicated to the discovery of transiting exoplanets. This spacecraft stared continuously at more than 150,000 stars in a small patch of sky near the constellation of Cygnus—just above the plane of our Milky Way Galaxy (Figure 21.20). Kepler's cameras and ability to measure small changes in brightness very precisely enabled the discovery of thousands of exoplanets, including many multi-planet systems. The spacecraft required three reaction wheels—a type of wheel used to help control slight rotation of the spacecraft—to stabilize the pointing of the telescope and monitor the brightness of the same group of stars over and over again. Kepler was launched with four reaction wheels (one a spare), but by May 2013, two wheels had failed and the telescope could no longer be accurately pointed toward the target area. Kepler had been designed to operate for 4 years, and ironically, the pointing failure occurred exactly 4 years and 1 day after it began observing.

However, this failure did not end the mission. The Kepler telescope continued to observe for two more years, looking for short-period transits in different parts of the sky. We discuss the Kepler results in the next section Exoplanets Everywhere: What We

Are Learning. A new NASA mission called TESS (Transiting Exoplanet Survey Satellite) is now carrying out a transit survey of the nearer (and therefore brighter) stars all over the sky. TESS completed its first planned survey of the sky in August 2020 and reported 66 exoplanet discoveries and almost 2100 candidates that need further observations to confirm.

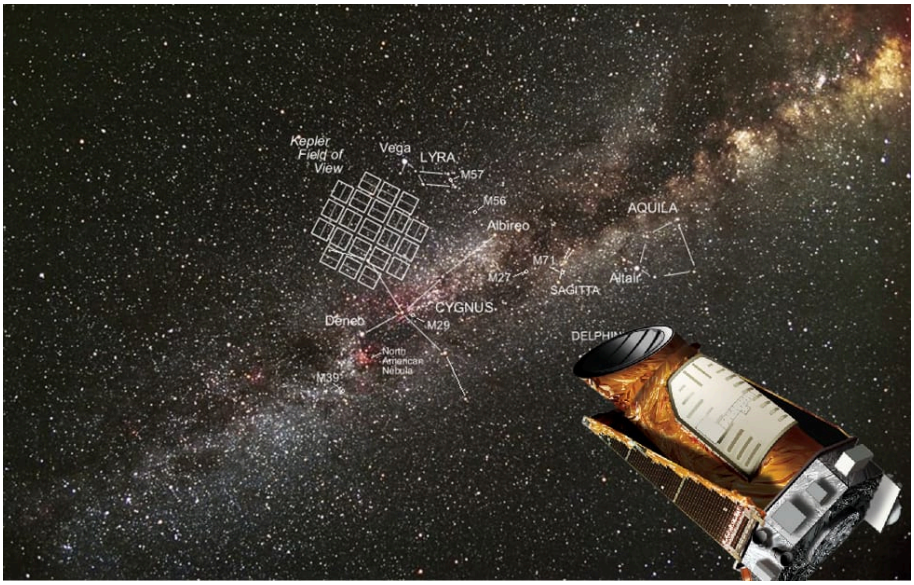


Figure 21.20 : Kepler's Field of View. The boxes show the region where the Kepler spacecraft cameras took images of over 150,000 stars regularly, to find transiting planets. (credit "field of view": modification of work by NASA/Kepler mission; credit "spacecraft": modification of work by NASA/Kepler mission/Wendy Stenzel)

What do we mean, exactly, by "discovery" of transiting exoplanets? A single transit shows up as a very slight drop in the brightness of the star, lasting several hours. However, astronomers must be on guard against other factors that might produce a false transit, especially when working at the limit of precision of the telescope. We must wait for a second transit of similar depth. But when another transit is observed, we don't initially know whether it might be due to another planet in a different orbit. The "discovery" occurs only when a third transit is found with similar depth and the same spacing in time as the first pair.

Computers normally conduct the analysis, which involves searching for tiny, periodic dips in the light from each star, extending over 4 years of observation. But the Kepler mission also has a program in which non-astronomers—citizen scientists—can examine the data. These dedicated volunteers have found several transits that were missed by the computer analyses, showing that the human eye and brain sometimes recognize unusual events that a computer was not programmed to look for.

Measuring three or four evenly spaced transits is normally enough to "discover" an exoplanet. But in a new field like exoplanet research, we would like to find further independent verification. The strongest confirmation happens when ground-based telescopes are also able to detect a Doppler shift with the same period as the transits. However, this is generally not possible for Earth-size planets. One of the most convincing ways to verify that a dip in brightness is due to a planet is to find more planets orbiting the same star—a *planetary system*. Multi-planet systems also provide alternative ways to estimate the masses of the planets, as we will discuss in the next section.

The selection effects (or biases) in the Kepler data are similar to those in Doppler observations. Large planets are easier to find than small ones, and short-period planets are easier than long-period planets. If we require three transits to establish the presence of a planet, we are of course limited to discovering planets with orbital periods less than one-third of the observing interval. Thus, it was only in its fourth and final year of operation that Kepler was able to find planets with orbits like Earth's that require 1 year to go around their star.

Direct Detection

The best possible evidence for an earthlike planet elsewhere would be an image. After all, "seeing is believing" is a very human prejudice. But imaging a distant planet is a formidable challenge indeed. Suppose, for example, you were a great distance away and wished to detect reflected light from Earth. Earth intercepts and reflects less than one billionth of the Sun's radiation, so its apparent brightness in visible light is less than one billionth that of the Sun. Compounding the challenge of detecting such a faint speck of light, the planet is swamped by the blaze of radiation from its parent star.

Even today, the best telescope mirrors' optics have slight imperfections that prevent the star's light from coming into focus in a completely sharp point.

Direct imaging works best for young gas giant planets that emit infrared light and reside at large separations from their host stars. Young giant planets emit more infrared light because they have more internal energy, stored from the process of planet formation. Even then, clever techniques must be employed to subtract out the light from the host star. In 2008, three such young planets were discovered orbiting HR 8799, a star in the constellation of Pegasus (Figure 21.21). Two years later, a fourth planet was detected closer to the star. Additional planets may reside even closer to HR 8799, but if they exist, they are currently lost in the glare of the star.

Since then, a number of planets around other stars have been found using direct imaging. However, one challenge is to tell whether the objects we are seeing are indeed planets or if they are brown dwarfs (failed stars) in orbit around a star.

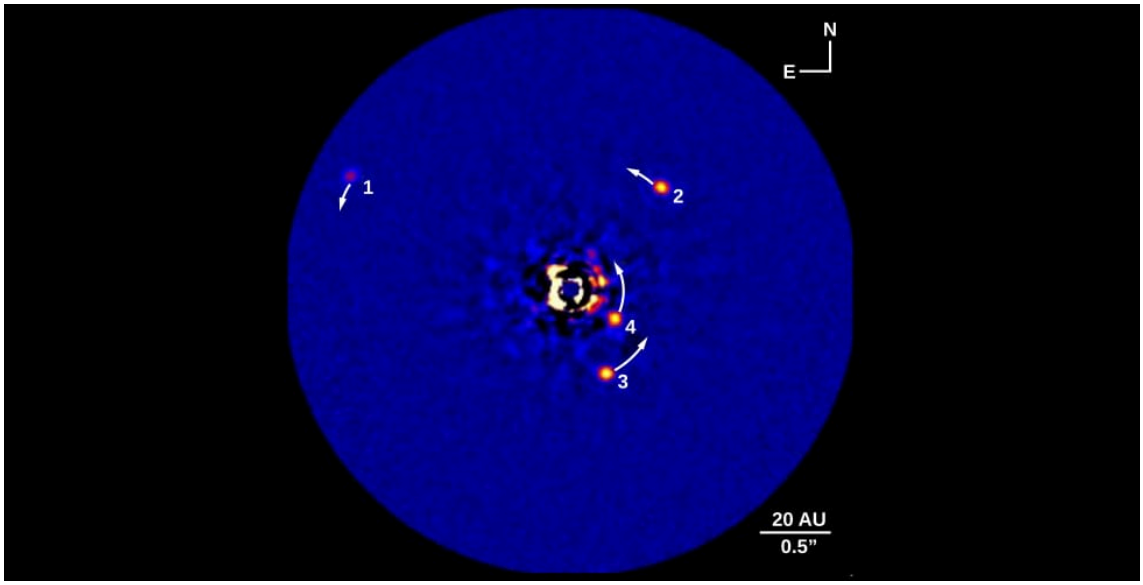


Figure 21.21 : Exoplanets around HR 8799. This image shows Keck telescope observations of four directly imaged planets orbiting HR 8799. A size scale for the system gives the distance in AU (remember that one astronomical unit is the distance between Earth and the Sun.) (credit: modification of work by Ben Zuckerman)

Direct imaging is an important technique for characterizing an exoplanet. The brightness of the planet can be measured at different wavelengths. These observations provide an estimate for the temperature of the planet's atmosphere; in the case of HR 8799 planet 1, the color suggests the presence of thick clouds. Spectra can also be obtained from the faint light to analyze the atmospheric constituents. A spectrum of HR 8799 planet 1 indicates a hydrogen-rich atmosphere, while the closer planet 4 shows evidence for methane in the atmosphere.

Another way to overcome the blurring effect of Earth's atmosphere is to observe from space. Infrared may be the optimal wavelength range in which to observe because planets get brighter in the infrared while stars like our Sun get fainter, thereby making it easier to detect a planet against the glare of its star. Special optical techniques can be used to suppress the light from the central star and make it easier to see the planet itself. However, even if we go into space, it will be difficult to obtain images of Earth-size planets.

Footnotes

- 1The Doppler method only allows us to find the minimum mass of a planet. To determine the exact mass using the Doppler shift and Kepler's laws, we must also have the angle at which the planet's orbit is oriented to our view—something we don't have any independent way of knowing in most cases. Still, if the minimum mass is half of Jupiter's, the actual mass can only be larger than that, and we are sure that we are dealing with a jovian planet.

Answers

Exercise 21.1.1

The radius of Earth is 6371 km. Therefore,
$\left(\frac{R_{\text{planet}}}{R_{\text{star}}}\right)^2 = \left(\frac{6371\text{ km}}{695{,}700/2\text{ km}}\right)^2 = \left(\frac{6371\text{ km}}{347{,}850\text{ km}}\right)^2 = 0.0003$, or significantly less than 1%.

This page titled 21.1: Planets beyond the Solar System- Search and Discovery is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 21.5: Planets beyond the Solar System- Search and Discovery by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

21.2: Exoplanets Everywhere - What we are Learning

Learning Objectives

By the end of this section, you will be able to:

- Explain what we have learned from our discovery of exoplanets
- Identify which kind of exoplanets appear to be the most common in the Galaxy
- Discuss the kinds of planetary systems we are finding around other stars

Before the discovery of exoplanets, most astronomers expected that other planetary systems would be much like our own—planets following roughly circular orbits, with the most massive planets several AU from their parent star. Such systems do exist in large numbers, but many exoplanets and planetary systems are very different from those in our solar system. Another surprise is the existence of whole classes of exoplanets that we simply don't have in our solar system: planets with masses between the mass of Earth and Neptune, and planets that are several times more massive than Jupiter.

Kepler Results

The Kepler telescope has been responsible for the discovery of most exoplanets, especially at smaller sizes, as illustrated in Figure 21.22, where the Kepler discoveries are plotted in yellow. You can see the wide range of sizes, including planets substantially larger than Jupiter and smaller than Earth. The absence of Kepler-discovered exoplanets with orbital periods longer than a few hundred days is a consequence of the 4-year lifetime of the mission. (Remember that three evenly spaced transits must be observed to register a discovery.) At the smaller sizes, the absence of planets much smaller than one earth radius is due to the difficulty of detecting transits by very small planets. In effect, the "discovery space" for Kepler was limited to planets with orbital periods less than 400 days and sizes larger than Mars.

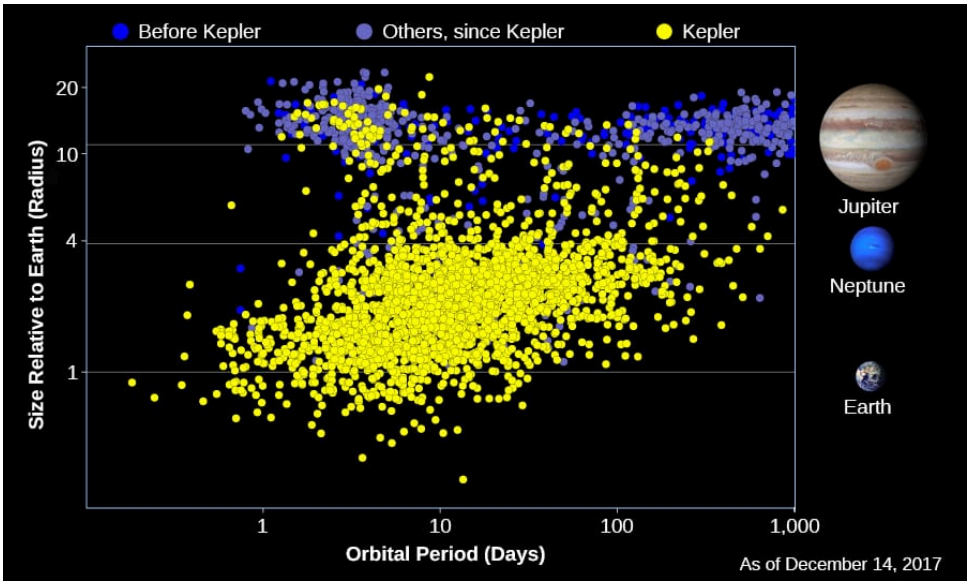


Figure 21.22 : Exoplanet Discoveries through 2017. The vertical axis shows the radius of each planet compared to Earth. Horizontal lines show the size of Earth, Neptune, and Jupiter. The horizontal axis shows the time each planet takes to make one orbit (and is given in Earth days). Recall that Mercury takes 88 days and Earth takes a little more than 365 days to orbit the Sun. The yellow dots show planets discovered by the Kepler mission using transits, and the blue dots are the discoveries by other projects and techniques. When this graph was made, astronomers had discovered just over 3,500 exoplanets; as of the start of 2022, that number has risen to roughly 5,000. (credit: modification of work by NASA/Ames Research Center/Jessie Dotson and Wendy Stenzel)

One of the primary objectives of the Kepler mission was to find out how many stars hosted planets and especially to estimate the frequency of earthlike planets. Although Kepler looked at only a very tiny fraction of the stars in the Galaxy, the sample size was

large enough to draw some interesting conclusions. While the observations apply only to the stars observed by Kepler, those stars are reasonably representative, and so astronomers can extrapolate to the entire Galaxy.

Figure 21.23 shows that the Kepler discoveries include many rocky, Earth-size planets, far more than Jupiter-size gas planets. This immediately tells us that the initial Doppler discovery of many hot Jupiters was a biased sample, in effect, finding the odd planetary systems because they were the easiest to detect. However, there is one huge difference between this observed size distribution and that of planets in our solar system. The most common planets have radii between 1.4 and 2.8 that of Earth, sizes for which we have no examples in the solar system. These have been nicknamed super-Earths, while the other large group with sizes between 2.8 and 4 that of Earth are often called mini-Neptunes.

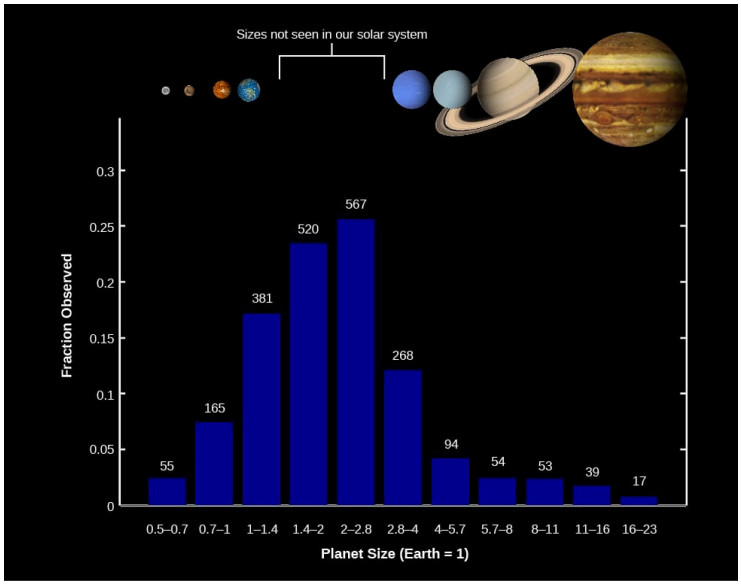


Figure 21.23 : Kepler Discoveries. This bar graph shows the number of planets of each size range found among the first 2213 Kepler planet discoveries. Sizes range from half the size of Earth to 20 times that of Earth. On the vertical axis, you can see the fraction that each size range makes up of the total. Note that planets that are between 1.4 and 4 times the size of Earth make up the largest fractions, yet this size range is not represented among the planets in our solar system. (credit: modification of work by NASA/Kepler mission)

What a remarkable discovery it is that the most common types of planets in the Galaxy are completely absent from our solar system and were unknown until Kepler's survey. However, recall that really small planets were difficult for the Kepler instruments to find. So, to estimate the frequency of Earth-size exoplanets, we need to correct for this sampling bias. The result is the corrected size distribution shown in Figure 21.24. Notice that in this graph, we have also taken the step of showing not the number of Kepler detections but the average number of planets per star for solar-type stars (spectral types F, G, and K).

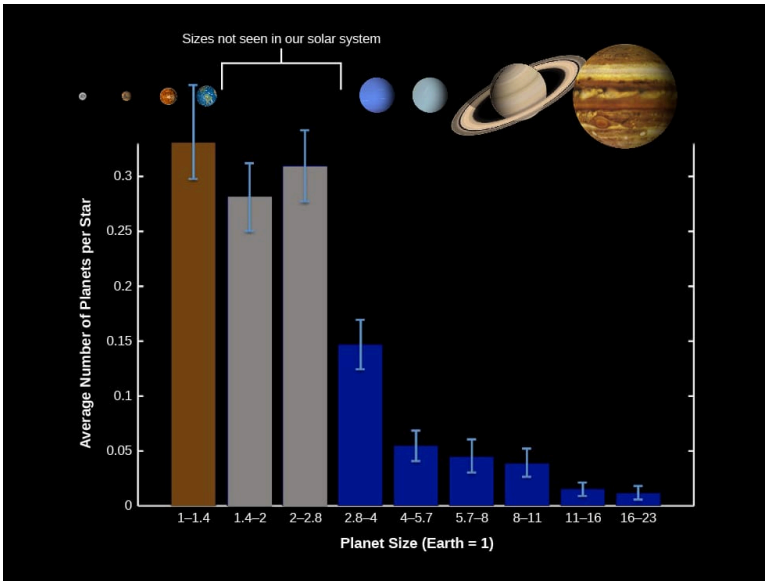


Figure 21.24 : Size Distribution of Planets for Stars Similar to the Sun. We show the average number of planets per star in each planet size range. (The average is less than one because some stars will have zero planets of that size range.) This distribution, corrected for biases in the Kepler data, shows that Earth-size planets may actually be the most common type of exoplanets. (credit: modification of work by NASA/Kepler mission)

We see that the most common planet sizes of are those with radii from 1 to 3 times that of Earth—what we have called “Earths” and “super-Earths.” Each group occurs in about one-third to one-quarter of stars. In other words, if we group these sizes together, we can conclude there is nearly one such planet per star! And remember, this census includes primarily planets with orbital periods less than 2 years. We do not yet know how many undiscovered planets might exist at larger distances from their star.

To estimate the number of Earth-size planets in our Galaxy, we need to remember that there are approximately 100 billion stars of spectral types F, G, and K. Therefore, we estimate that there are about 30 billion Earth-size planets in our Galaxy. If we include the super-Earths too, then there could be one hundred billion in the whole Galaxy. This idea—that planets of roughly Earth’s size are so numerous—is surely one of the most important discoveries of modern astronomy.

Planets with Known Densities

For several hundred exoplanets, we have been able to measure both the size of the planet from transit data and its mass from Doppler data, yielding an estimate of its density. Comparing the average density of exoplanets to the density of planets in our solar system helps us understand whether they are rocky or gaseous in nature. This has been particularly important for understanding the structure of the new categories of super-Earths and mini-Neptunes with masses between 3–10 times the mass of Earth. A key observation so far is that planets that are more than 10 times the mass of Earth have substantial gaseous envelopes (like Uranus and Neptune) whereas lower-mass planets are predominately rocky in nature (like the terrestrial planets).

Figure 21.25 compares all the exoplanets that have both mass and radius measurements. The dependence of the radius on planet mass is also shown for a few illustrative cases—hypothetical planets made of pure iron, rock, water, or hydrogen.

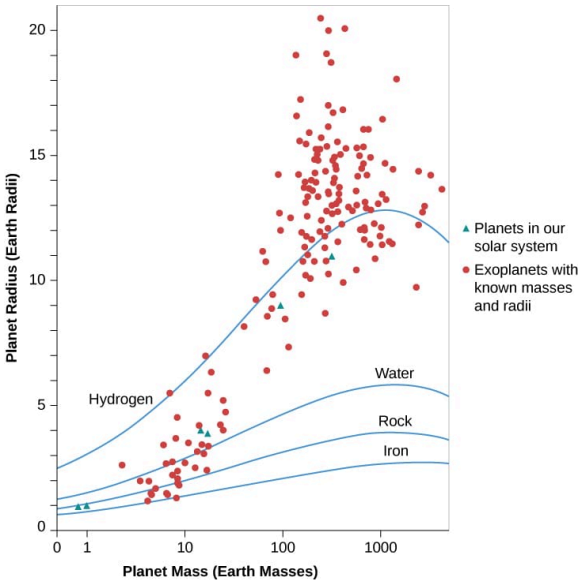


Figure 21.25 : Exoplanets with Known Densities. Exoplanets with known masses and radii (red circles) are plotted along with solid lines that show the theoretical size of pure iron, rock, water, and hydrogen planets with increasing mass. Masses are given in multiples of Earth's mass. (For comparison, Jupiter contains enough mass to make 320 Earths.) The green triangles indicate planets in our solar system.

At lower masses, notice that as the mass of these hypothetical planets increases, the radius also increases. That makes sense—if you were building a model of a planet out of clay, your toy planet would increase in size as you added more clay. However, for the highest mass planets ($M > 1000\ M_{\text{Earth}}$) in Figure 21.25, notice that the radius stops increasing and the planets with greater mass are actually smaller. This occurs because increasing the mass also increases the gravity of the planet, so that compressible materials (even rock is compressible) will become more tightly packed, shrinking the size of the more massive planet.

In reality, planets are not pure compositions like the hypothetical water or iron planet. Earth is composed of a solid iron core, an outer liquid-iron core, a rocky mantle and crust, and a relatively thin atmospheric layer. Exoplanets are similarly likely to be differentiated into compositional layers. The theoretical lines in Figure 21.25 are simply guides that suggest a range of possible compositions.

Astronomers who work on the complex modeling of the interiors of rocky planets make the simplifying assumption that the planet consists of two or three layers. This is not perfect, but it is a reasonable approximation and another good example of how science works. Often, the first step in understanding something new is to narrow down the range of possibilities. This sets the stage for refining and deepening our knowledge. In Figure 21.25, the two green triangles with roughly 1 M_{Earth} and 1 R_{Earth} represent Venus and Earth. Notice that these planets fall between the models for a pure iron and a pure rock planet, consistent with what we would expect for the known mixed-chemical composition of Venus and Earth.

In the case of gaseous planets, the situation is more complex. Hydrogen is the lightest element in the periodic table, yet many of the detected exoplanets in Figure 21.25 with masses greater than $100\ M_{\text{Earth}}$ have radii that suggest they are lower in density than a pure hydrogen planet. Hydrogen is the lightest element, so what is happening here? Why do some gas giant planets have inflated radii that are larger than the fictitious pure hydrogen planet? Many of these planets reside in short-period orbits close to the host star where they intercept a significant amount of radiated energy. If this energy is trapped deep in the planet atmosphere, it can cause the planet to expand.

Planets that orbit close to their host stars in slightly eccentric orbits have another source of energy: the star will raise tides in these planets that tend to circularize the orbits. This process also results in tidal dissipation of energy that can inflate the atmosphere. It would be interesting to measure the size of gas giant planets in wider orbits where the planets should be cooler—the expectation is that unless they are very young, these cooler gas giant exoplanets (sometimes called "cold Jupiters") should not be inflated. But we don't yet have data on these more distant exoplanets.

Exoplanetary Systems

As we search for exoplanets, we don't expect to find only one planet per star. Our solar system has eight major planets, half a dozen dwarf planets, and millions of smaller objects orbiting the Sun. The evidence we have of planetary systems in formation

also suggest that they are likely to produce multi-planet systems.

The first planetary system was found around the star Upsilon Andromedae in 1999 using the Doppler method, and many others have been found since then. If such exoplanetary system are common, let's consider which systems we expect to find in the Kepler transit data.

A planet will transit its star only if Earth lies in the plane of the planet's orbit. If the planets in other systems do not have orbits in the same plane, we are unlikely to see multiple transiting objects. Also, as we have noted before, Kepler was sensitive only to planets with orbital periods less than about 4 years. What we expect from Kepler data, then, is evidence of coplanar planetary systems confined to what would be the realm of the terrestrial planets in our solar system.

By 2022, astronomers gathered data on over 800 such exoplanet systems. Many have only two known planets, but a few have as many as five, and one has eight (the same number of planets as our own solar system). For the most part, these are very compact systems with most of their planets closer to their star than Mercury is to the Sun. The figure below shows one of the largest exoplanet systems: that of the star called Kepler-62 (Figure 21.26). Our solar system is shown to the same scale, for comparison (note that the Kepler-62 planets are drawn with artistic license; we have no detailed images of any exoplanets).

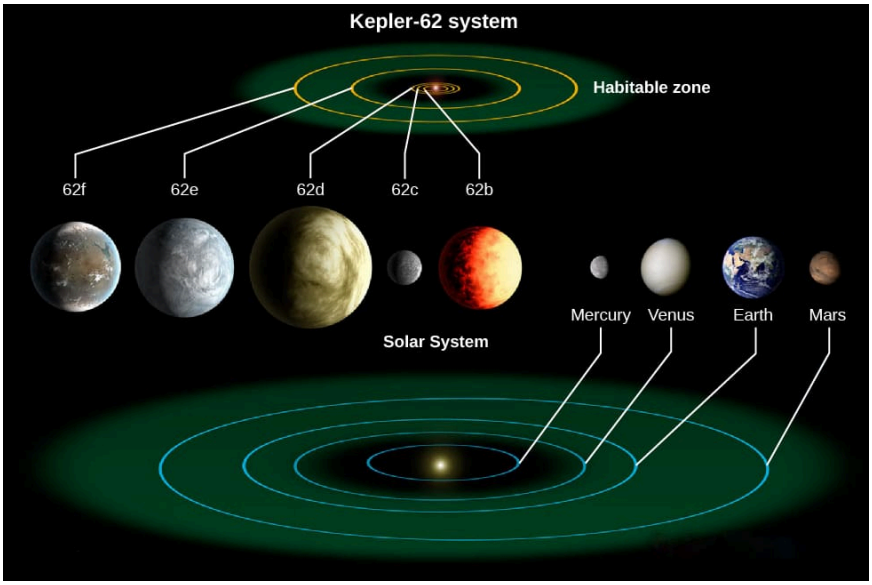


Figure 21.26 : Exoplanet System Kepler-62, with the Solar System Shown to the Same Scale. The green areas are the "habitable zones," the range of distance from the star where surface temperatures are likely to be consistent with liquid water. (credit: modification of work by NASA/ Ames/JPL-Caltech)

All but one of the planets in the K-62 system are larger than Earth. These are super-Earths, and one of them (62d) is in the size range of a mini-Neptune, where it is likely to be largely gaseous. The smallest planet in this system is about the size of Mars. The three inner planets orbit very close to their star, and only the outer two have orbits larger than Mercury in our system. The green areas represent each star's "habitable zone," which is the distance from the star where we calculate that surface temperatures would be consistent with liquid water. The Kepler-62 habitable zone is much smaller than that of the Sun because the star is intrinsically fainter.

With closely spaced systems like this, the planets can interact gravitationally with each other. The result is that the observed transits occur a few minutes earlier or later than would be predicted from simple orbits. These gravitational interactions have allowed the Kepler scientists to calculate masses for the planets, providing another way to learn about exoplanets.

Kepler has discovered some interesting and unusual planetary systems. For example, most astronomers expected planets to be limited to single stars. But we have found planets orbiting close double stars, so that the planet would see two suns in its sky, like those of the fictional planet Tatooine in the *Star Wars* films. At the opposite extreme, planets can orbit one star of a wide, double-star system without major interference from the second star.

This page titled 21.2: Exoplanets Everywhere - What we are Learning is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa via source content that was edited to the style and standards of the LibreTexts platform.

- 21.6: Exoplanets Everywhere - What we are Learning by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/astronomy.

INDEX

A

absorption spectrum
9.3: Spectroscopy in Astronomy
accelerate
4.1: The Birth of Modern Astronomy
active region
19.3: Solar Activity above the Photosphere
adaptive optics
10.2: Telescopes Today
angular momentum
5.1: Newton's Great Synthesis
aperture
10.1: Telescopes
aphelion
5.3: Orbits in the Solar System
apogee
5.3: Orbits in the Solar System
apparent brightness
20.1: The Brightness of Stars
apparent magnitude
3.2: Ancient Astronomy
apparent solar time
6.3: Keeping Time
Arctic Circle
6.2: The Seasons
Ascension
6.1: Earth and Sky
asteroid
11.1: Overview of Our Planetary System
18.1: Asteroids
asteroid belt
18.1: Asteroids
5.3: Orbits in the Solar System
astronomical
3.1: The Sky Above
astronomical unit
4.2: The Laws of Planetary Motion
astronomical unit (AU)
2.1: A Tour of the Universe
4.2: The Laws of Planetary Motion
AU
4.2: The Laws of Planetary Motion
Aurora
19.1: The Structure and Composition of the Sun

B

Balmer lines
20.3: The Spectra of Stars (and Brown Dwarfs)
bar
12.3: Earth's Atmosphere
basalt
12.1: The Global Perspective
binary stars
20.4: Measuring Stellar Masses
blackbody
9.2: The Electromagnetic Spectrum
brown dwarf
20.3: The Spectra of Stars (and Brown Dwarfs)
20.4: Measuring Stellar Masses

C

Callisto
17.2: The Gailean Moons of Jupiter
CCD
10.3: Visible-Light Detectors and Instruments
celestial equator
3.1: The Sky Above
celestial poles
3.1: The Sky Above
celestial sphere
3.1: The Sky Above
chromatic aberration
10.1: Telescopes
chromosphere
19.1: The Structure and Composition of the Sun
circumpolar zone
3.1: The Sky Above
CME
19.3: Solar Activity above the Photosphere
color index
20.2: Colors of Stars
coma (comet)
18.3: The "Long-Haired" Comets
comet
11.1: Overview of Our Planetary System
18.3: The "Long-Haired" Comets
conduction
19.5: The Solar Interior - Theory
constellation
2.2: The Universe on the Large Scale
continuous spectrum
9.3: Spectroscopy in Astronomy
convection
12.2: Earth's Crust
19.5: The Solar Interior - Theory
Copernicus
4.1: The Birth of Modern Astronomy
core
12.1: The Global Perspective
corona
19.1: The Structure and Composition of the Sun
coronal hole
19.1: The Structure and Composition of the Sun
coronal mass ejection
19.3: Solar Activity above the Photosphere
coronal mass ejection (CME)
19.3: Solar Activity above the Photosphere
Cosmology
3.2: Ancient Astronomy
crust
12.1: The Global Perspective
cycle of the seasons
6.2: The Seasons

D

declination
6.1: Earth and Sky
density
5.1: Newton's Great Synthesis
detector
10.1: Telescopes
differential force
8.1: Ocean Tides and the Moon
differential forces
8.1: Ocean Tides and the Moon
differential rotation
19.2: The Solar Cycle
differentiation
11.2: Compostion and Structure of Planets
dispersion
9.3: Spectroscopy in Astronomy
Doppler effect
9.6: The Doppler Effect

E

Early Calendars
6.4: The Calendar
Eccentricity
4.2: The Laws of Planetary Motion
Ecliptic
3.1: The Sky Above
electromagnetic radiation
9.1: The Behavior of Light
electromagnetic spectrum
9.2: The Electromagnetic Spectrum
ellipse
4.2: The Laws of Planetary Motion
emission spectrum
9.3: Spectroscopy in Astronomy
energy flux
9.2: The Electromagnetic Spectrum
energy level
9.4: The Structure of the Atom
epicycle
3.2: Ancient Astronomy
escape speed
5.4: Motions of Satellites and Spacecraft
Europa
17.2: The Gailean Moons of Jupiter
excitation
9.5: Formation of Spectral Lines
exoplanet
21.1: Planets beyond the Solar System- Search and Discovery
eyepiece
10.1: Telescopes

F

fault
12.2: Earth's Crust
fission
19.4: Mass, Energy, and the Theory of Relativity
focus
10.1: Telescopes
4.2: The Laws of Planetary Motion
frequency
9.1: The Behavior of Light

fusion
19.4: Mass, Energy, and the Theory of Relativity

G

gamma rays
9.2: The Electromagnetic Spectrum
Ganymede
17.2: The Gailean Moons of Jupiter
geocentric
3.1: The Sky Above
Geology of Mars
15.1: The Geology of Mars
giant planet
11.1: Overview of Our Planetary System
granite
12.1: The Global Perspective
granulation
19.1: The Structure and Composition of the Sun
gravity
5.2: Newton's Universal Law of Gravitation
great circle
6.1: Earth and Sky
greenhouse effect
12.4: Life, Chemical Evolution, and Climate Change
greenhouse gas
12.4: Life, Chemical Evolution, and Climate Change
Gregorian Calendar
6.4: The Calendar
ground state
9.5: Formation of Spectral Lines

H

H–R diagram
20.5: The H-R Diagram
heliocentric
4.1: The Birth of Modern Astronomy
helioseismology
19.6: The Solar Interior - Observations
highlands
13.2: The Lunar Surface
horizon
3.1: The Sky Above
hubble telescope
1.6: Consequences of Light Travel Time
hydrostatic equilibrium
19.5: The Solar Interior - Theory
hypothesis
1.3: The Nature of Science

I

igneous rock
12.2: Earth's Crust
infrared
9.2: The Electromagnetic Spectrum
interference
10.4: Radio Telescopes
interferometer
10.4: Radio Telescopes
interferometer array
10.4: Radio Telescopes
International Date Line
6.3: Keeping Time
inverse square law
9.1: The Behavior of Light
ion
9.5: Formation of Spectral Lines
ionization
9.5: Formation of Spectral Lines
island universes
2.2: The Universe on the Large Scale
isotope
9.4: The Structure of the Atom

K

Kepler's first law
4.2: The Laws of Planetary Motion
Kepler's second law
4.2: The Laws of Planetary Motion
Kepler's third law
4.2: The Laws of Planetary Motion

L

lo
17.2: The Gailean Moons of Jupiter
luminosity
20.1: The Brightness of Stars
lunar eclipse
8.2: Eclipses of the Sun and Moon

M

Magellanic Clouds
2.2: The Universe on the Large Scale
magnetosphere
12.1: The Global Perspective
magnitude
20.1: The Brightness of Stars
main sequence
20.5: The H-R Diagram
major axis
4.2: The Laws of Planetary Motion
mantle
12.1: The Global Perspective
mare
13.2: The Lunar Surface
Margaret Huggins
20.3: The Spectra of Stars (and Brown Dwarfs)
mass extinction
12.5: Cosmic Influences on the Evolution of Earth
Maunder Minimum
19.2: The Solar Cycle
Mean Solar Time
6.3: Keeping Time
mercury
13.5: Mercury
meridian
6.1: Earth and Sky
metamorphic rock
12.2: Earth's Crust
meteor
11.1: Overview of Our Planetary System
meteorite
11.1: Overview of Our Planetary System
15.1: The Geology of Mars
microwave
9.2: The Electromagnetic Spectrum
momentum
5.1: Newton's Great Synthesis

N

NEA
18.2: Asteroids and Planetary Defense
NEO
18.2: Asteroids and Planetary Defense
neutrino
19.4: Mass, Energy, and the Theory of Relativity
Newton's first law
5.1: Newton's Great Synthesis
Newton's second law
5.1: Newton's Great Synthesis
Newton's third law
5.1: Newton's Great Synthesis
north pole
6.2: The Seasons
Northern Hemisphere
6.2: The Seasons
nucleus
18.3: The "Long-Haired" Comets
9.4: The Structure of the Atom
nucleus (of a comet)
18.3: The "Long-Haired" Comets

O

orbit
4.2: The Laws of Planetary Motion
orbital period
4.2: The Laws of Planetary Motion
orbital speed
4.2: The Laws of Planetary Motion
ozone
12.3: Earth's Atmosphere

P

parallax
3.2: Ancient Astronomy
perigee
5.3: Orbits in the Solar System
perihelion
5.3: Orbits in the Solar System
perturbation
5.5: Gravity with More Than Two Bodies
phases of the Moon
7.1: Phases and Motions of the Moon
photochemistry
16.3: Atmosphere of the Giant Planets
photon
9.1: The Behavior of Light
photosphere
19.1: The Structure and Composition of the Sun
photosynthesis
12.4: Life, Chemical Evolution, and Climate Change
plage
19.3: Solar Activity above the Photosphere
planet
3.1: The Sky Above
planetesimals
11.4: Origin of the Solar System
plasma
19.1: The Structure and Composition of the Sun

plate tectonics
12.2: Earth's Crust
Positron
19.4: Mass, Energy, and the Theory of Relativity
precession
3.2: Ancient Astronomy
precession of Earth
3.2: Ancient Astronomy
prime focus
10.1: Telescopes
primitive rock
12.2: Earth's Crust
prolate spheroid
8.1: Ocean Tides and the Moon
prominence
19.3: Solar Activity above the Photosphere

R

radar
10.4: Radio Telescopes
radial velocity
9.6: The Doppler Effect
radiation
19.5: The Solar Interior - Theory
radio waves
9.2: The Electromagnetic Spectrum
radioactivity
11.3: Dating Planetary Surfaces
Reflecting Telescope
10.1: Telescopes
refracting telescope
10.1: Telescopes
refraction
6.2: The Seasons
resolution
10.2: Telescopes Today
retrograde motion
3.2: Ancient Astronomy
rift zone
12.2: Earth's Crust
right ascension
6.1: Earth and Sky
runaway greenhouse effect
14.3: The Massive Atmosphere of Venus

S

satellite
5.3: Orbits in the Solar System
scientific laws
1.4: The Laws of Nature
scientific notation
1.5: Numbers in Astronomy
sedimentary rock
12.2: Earth's Crust
seeing
10.2: Telescopes Today
seismic wave
12.1: The Global Perspective
semimajor axis
4.2: The Laws of Planetary Motion
sidereal day
6.3: Keeping Time
sidereal month
7.1: Phases and Motions of the Moon
solar day
6.3: Keeping Time
solar flare
19.3: Solar Activity above the Photosphere
solar month
7.1: Phases and Motions of the Moon
solar nebula
11.4: Origin of the Solar System
solar wind
19.1: The Structure and Composition of the Sun
Southern Hemisphere
6.2: The Seasons
space
2.3: The Universe of the Very Small
spectral class
20.3: The Spectra of Stars (and Brown Dwarfs)
spectrometer
9.3: Spectroscopy in Astronomy
spectroscopic binary
20.4: Measuring Stellar Masses
stratosphere
12.3: Earth's Atmosphere
subduction
12.2: Earth's Crust
sunspot
19.2: The Solar Cycle
sunspot cycle
19.2: The Solar Cycle
synchronous rotation
7.1: Phases and Motions of the Moon
synchrotron radiation
16.2: The Giant Planets

T

tail
18.3: The "Long-Haired" Comets
tectonic
14.2: The Geology of Venus
telescope
10.1: Telescopes
terrestrial planet
11.1: Overview of Our Planetary System
tidal force
17.2: The Gailean Moons of Jupiter
tidal heating
17.2: The Gailean Moons of Jupiter
tides
8.1: Ocean Tides and the Moon
Titan
17.3: Titan and Triton
transit
21.1: Planets beyond the Solar System- Search and Discovery
transit depth
21.1: Planets beyond the Solar System- Search and Discovery
transition region
19.1: The Structure and Composition of the Sun
triton
17.3: Titan and Triton
Tropic of Cancer
6.2: The Seasons
Tropic of Capricorn
6.2: The Seasons
Troposphere
12.3: Earth's Atmosphere
Twin Mars Exploration Rovers (MER)
15.1: The Geology of Mars

U

ultraviolet
9.2: The Electromagnetic Spectrum

V

velocity
5.1: Newton's Great Synthesis
visible light
9.2: The Electromagnetic Spectrum
visual binary
20.4: Measuring Stellar Masses
volcano
12.2: Earth's Crust

W

wavelength
9.1: The Behavior of Light
white dwarf
20.5: The H-R Diagram
Wien's law
9.2: The Electromagnetic Spectrum
William Huggins
20.3: The Spectra of Stars (and Brown Dwarfs)

Z

Zodiac
3.1: The Sky Above

This page titled Index is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

GLOSSARY

S

Sample Word 1 | Sample Definition 1

This page titled Glossary is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

DETAILED LICENSING

Overview

Title: Astronomy (Mastrapa)

Webpages: 124

All licenses found:

- CC BY 4.0: 99.2% (123 pages)
- Undeclared: 0.8% (1 page)

By Page

- Astronomy (Mastrapa) - *CC BY 4.0*
  - Front Matter - *CC BY 4.0*
    - TitlePage - *CC BY 4.0*
    - InfoPage - *CC BY 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY 4.0*
  - 1: Introduction to Astronomy - *CC BY 4.0*
    - 1.1: Introduction - *CC BY 4.0*
    - 1.2: The Nature of Astronomy - *CC BY 4.0*
    - 1.3: The Nature of Science - *CC BY 4.0*
    - 1.4: The Laws of Nature - *CC BY 4.0*
    - 1.5: Numbers in Astronomy - *CC BY 4.0*
    - 1.6: Consequences of Light Travel Time - *CC BY 4.0*
  - 2: The Universe - *CC BY 4.0*
    - 2.1: A Tour of the Universe - *CC BY 4.0*
    - 2.2: The Universe on the Large Scale - *CC BY 4.0*
    - 2.3: The Universe of the Very Small - *CC BY 4.0*
    - 2.4: A Conclusion and a Beginning - *CC BY 4.0*
  - 3: Ancient Astronomy - *CC BY 4.0*
    - 3.1: The Sky Above - *CC BY 4.0*
    - 3.2: Ancient Astronomy - *CC BY 4.0*
    - 3.3: Key Terms - *CC BY 4.0*
  - 4: Modern Astronomy - *CC BY 4.0*
    - 4.1: The Birth of Modern Astronomy - *CC BY 4.0*
    - 4.2: The Laws of Planetary Motion - *CC BY 4.0*
  - 5: Physics and Gravity - *CC BY 4.0*
    - 5.1: Newton's Great Synthesis - *CC BY 4.0*
    - 5.2: Newton's Universal Law of Gravitation - *CC BY 4.0*
    - 5.3: Orbits in the Solar System - *CC BY 4.0*
    - 5.4: Motions of Satellites and Spacecraft - *CC BY 4.0*
    - 5.5: Gravity with More Than Two Bodies - *CC BY 4.0*
    - 5.6: Key Terms - *CC BY 4.0*
  - 6: Time - *CC BY 4.0*
    - 6.1: Earth and Sky - *CC BY 4.0*
    - 6.2: The Seasons - *CC BY 4.0*
    - 6.3: Keeping Time - *CC BY 4.0*
    - 6.4: The Calendar - *CC BY 4.0*
  - 7: The Moon - *CC BY 4.0*
    - 7.1: Phases and Motions of the Moon - *CC BY 4.0*
  - 8: The Moon 2 - *CC BY 4.0*
    - 8.1: Ocean Tides and the Moon - *CC BY 4.0*
    - 8.2: Eclipses of the Sun and Moon - *CC BY 4.0*
    - 8.3: Key Terms - *CC BY 4.0*
  - 9: Light - *CC BY 4.0*
    - 9.1: The Behavior of Light - *CC BY 4.0*
    - 9.2: The Electromagnetic Spectrum - *CC BY 4.0*
    - 9.3: Spectroscopy in Astronomy - *CC BY 4.0*
    - 9.4: The Structure of the Atom - *CC BY 4.0*
    - 9.5: Formation of Spectral Lines - *CC BY 4.0*
    - 9.6: The Doppler Effect - *CC BY 4.0*
    - 9.7: Key Terms - *CC BY 4.0*
  - 10: Observing - *CC BY 4.0*
    - 10.1: Telescopes - *CC BY 4.0*
    - 10.2: Telescopes Today - *CC BY 4.0*
    - 10.3: Visible-Light Detectors and Instruments - *CC BY 4.0*
    - 10.4: Radio Telescopes - *CC BY 4.0*
    - 10.5: Observations outside Earth's Atmosphere - *CC BY 4.0*
    - 10.6: The Future of Large Telescopes - *CC BY 4.0*
    - 10.7: Key Terms - *CC BY 4.0*
  - 11: Solar System - *CC BY 4.0*
    - 11.1: Overview of Our Planetary System - *CC BY 4.0*
    - 11.2: Compostion and Structure of Planets - *CC BY 4.0*
    - 11.3: Dating Planetary Surfaces - *CC BY 4.0*
    - 11.4: Origin of the Solar System - *CC BY 4.0*
    - 11.5: Summary - *CC BY 4.0*
  - 12: Earth - *CC BY 4.0*
    - 12.1: The Global Perspective - *CC BY 4.0*
    - 12.2: Earth's Crust - *CC BY 4.0*
    - 12.3: Earth's Atmosphere - *CC BY 4.0*

- 12.4: Life, Chemical Evolution, and Climate Change - *CC BY 4.0*
- 12.5: Cosmic Influences on the Evolution of Earth - *CC BY 4.0*
- 12.6: Key Terms - *CC BY 4.0*

◦ 13: The Moon and Mercury - *CC BY 4.0*
- 13.1: General Properties of the Moon - *CC BY 4.0*
- 13.2: The Lunar Surface - *CC BY 4.0*
- 13.3: Impact Craters - *CC BY 4.0*
- 13.4: The Origin of the Moon - *CC BY 4.0*
- 13.5: Mercury - *CC BY 4.0*
- 13.6: Key Terms - *CC BY 4.0*

◦ 14: Venus - *CC BY 4.0*
- 14.1: The Nearest Planets - An Overview - *CC BY 4.0*
- 14.2: The Geology of Venus - *CC BY 4.0*
- 14.3: The Massive Atmosphere of Venus - *CC BY 4.0*

◦ 15: Mars - *CC BY 4.0*
- 15.1: The Geology of Mars - *CC BY 4.0*
- 15.2: Water and Life on Mars - *CC BY 4.0*
- 15.3: Divergent Planetary Evolution - *CC BY 4.0*
- 15.4: Key Terms - *CC BY 4.0*

◦ 16: Giant Planets - *CC BY 4.0*
- 16.1: Exploring the Outer Planets - *CC BY 4.0*
- 16.2: The Giant Planets - *CC BY 4.0*
- 16.3: Atmosphere of the Giant Planets - *CC BY 4.0*
- 16.4: Key Terms - *CC BY 4.0*

◦ 17: Moons - *CC BY 4.0*
- 17.1: Ring and Moon Systems Introduced - *CC BY 4.0*
- 17.2: The Gailean Moons of Jupiter - *CC BY 4.0*
- 17.3: Titan and Triton - *CC BY 4.0*

◦ 18: Asteroids and Comets - *CC BY 4.0*
- 18.1: Asteroids - *CC BY 4.0*
- 18.2: Asteroids and Planetary Defense - *CC BY 4.0*
- 18.3: The "Long-Haired" Comets - *CC BY 4.0*
- 18.4: Key Terms - *CC BY 4.0*

◦ 19: The Sun - *CC BY 4.0*
- 19.1: The Structure and Composition of the Sun - *CC BY 4.0*
- 19.2: The Solar Cycle - *CC BY 4.0*
- 19.3: Solar Activity above the Photosphere - *CC BY 4.0*
- 19.4: Mass, Energy, and the Theory of Relativity - *CC BY 4.0*
- 19.5: The Solar Interior - Theory - *CC BY 4.0*
- 19.6: The Solar Interior - Observations - *CC BY 4.0*
- 19.7: Key Terms - *CC BY 4.0*

◦ 20: The Stars - *CC BY 4.0*
- 20.1: The Brightness of Stars - *CC BY 4.0*
- 20.2: Colors of Stars - *CC BY 4.0*
- 20.3: The Spectra of Stars (and Brown Dwarfs) - *CC BY 4.0*
- 20.4: Measuring Stellar Masses - *CC BY 4.0*
- 20.5: The H-R Diagram - *CC BY 4.0*

◦ 21: Exoplanets - *CC BY 4.0*
- 21.1: Planets beyond the Solar System- Search and Discovery - *CC BY 4.0*
- 21.2: Exoplanets Everywhere - What we are Learning - *CC BY 4.0*

◦ Back Matter - *CC BY 4.0*
- Index - *CC BY 4.0*
- Glossary - *CC BY 4.0*
- Detailed Licensing - *CC BY 4.0*
- Detailed Licensing - *CC BY 4.0*

This page titled Astronomy (Mastrapa) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.

DETAILED LICENSING

Overview

Title: Astronomy (Mastrapa)

Webpages: 124

All licenses found:

- CC BY 4.0: 99.2% (123 pages)
- Undeclared: 0.8% (1 page)

By Page

- Astronomy (Mastrapa) - *CC BY 4.0*
  - Front Matter - *CC BY 4.0*
    - TitlePage - *CC BY 4.0*
    - InfoPage - *CC BY 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY 4.0*
  - 1: Introduction to Astronomy - *CC BY 4.0*
    - 1.1: Introduction - *CC BY 4.0*
    - 1.2: The Nature of Astronomy - *CC BY 4.0*
    - 1.3: The Nature of Science - *CC BY 4.0*
    - 1.4: The Laws of Nature - *CC BY 4.0*
    - 1.5: Numbers in Astronomy - *CC BY 4.0*
    - 1.6: Consequences of Light Travel Time - *CC BY 4.0*
  - 2: The Universe - *CC BY 4.0*
    - 2.1: A Tour of the Universe - *CC BY 4.0*
    - 2.2: The Universe on the Large Scale - *CC BY 4.0*
    - 2.3: The Universe of the Very Small - *CC BY 4.0*
    - 2.4: A Conclusion and a Beginning - *CC BY 4.0*
  - 3: Ancient Astronomy - *CC BY 4.0*
    - 3.1: The Sky Above - *CC BY 4.0*
    - 3.2: Ancient Astronomy - *CC BY 4.0*
    - 3.3: Key Terms - *CC BY 4.0*
  - 4: Modern Astronomy - *CC BY 4.0*
    - 4.1: The Birth of Modern Astronomy - *CC BY 4.0*
    - 4.2: The Laws of Planetary Motion - *CC BY 4.0*
  - 5: Physics and Gravity - *CC BY 4.0*
    - 5.1: Newton's Great Synthesis - *CC BY 4.0*
    - 5.2: Newton's Universal Law of Gravitation - *CC BY 4.0*
    - 5.3: Orbits in the Solar System - *CC BY 4.0*
    - 5.4: Motions of Satellites and Spacecraft - *CC BY 4.0*
    - 5.5: Gravity with More Than Two Bodies - *CC BY 4.0*
    - 5.6: Key Terms - *CC BY 4.0*
  - 6: Time - *CC BY 4.0*
    - 6.1: Earth and Sky - *CC BY 4.0*
    - 6.2: The Seasons - *CC BY 4.0*
    - 6.3: Keeping Time - *CC BY 4.0*
    - 6.4: The Calendar - *CC BY 4.0*
  - 7: The Moon - *CC BY 4.0*
    - 7.1: Phases and Motions of the Moon - *CC BY 4.0*
  - 8: The Moon 2 - *CC BY 4.0*
    - 8.1: Ocean Tides and the Moon - *CC BY 4.0*
    - 8.2: Eclipses of the Sun and Moon - *CC BY 4.0*
    - 8.3: Key Terms - *CC BY 4.0*
  - 9: Light - *CC BY 4.0*
    - 9.1: The Behavior of Light - *CC BY 4.0*
    - 9.2: The Electromagnetic Spectrum - *CC BY 4.0*
    - 9.3: Spectroscopy in Astronomy - *CC BY 4.0*
    - 9.4: The Structure of the Atom - *CC BY 4.0*
    - 9.5: Formation of Spectral Lines - *CC BY 4.0*
    - 9.6: The Doppler Effect - *CC BY 4.0*
    - 9.7: Key Terms - *CC BY 4.0*
  - 10: Observing - *CC BY 4.0*
    - 10.1: Telescopes - *CC BY 4.0*
    - 10.2: Telescopes Today - *CC BY 4.0*
    - 10.3: Visible-Light Detectors and Instruments - *CC BY 4.0*
    - 10.4: Radio Telescopes - *CC BY 4.0*
    - 10.5: Observations outside Earth's Atmosphere - *CC BY 4.0*
    - 10.6: The Future of Large Telescopes - *CC BY 4.0*
    - 10.7: Key Terms - *CC BY 4.0*
  - 11: Solar System - *CC BY 4.0*
    - 11.1: Overview of Our Planetary System - *CC BY 4.0*
    - 11.2: Compostion and Structure of Planets - *CC BY 4.0*
    - 11.3: Dating Planetary Surfaces - *CC BY 4.0*
    - 11.4: Origin of the Solar System - *CC BY 4.0*
    - 11.5: Summary - *CC BY 4.0*
  - 12: Earth - *CC BY 4.0*
    - 12.1: The Global Perspective - *CC BY 4.0*
    - 12.2: Earth's Crust - *CC BY 4.0*
    - 12.3: Earth's Atmosphere - *CC BY 4.0*

- 12.4: Life, Chemical Evolution, and Climate Change - *CC BY 4.0*
- 12.5: Cosmic Influences on the Evolution of Earth - *CC BY 4.0*
- 12.6: Key Terms - *CC BY 4.0*

◦ 13: The Moon and Mercury - *CC BY 4.0*
- 13.1: General Properties of the Moon - *CC BY 4.0*
- 13.2: The Lunar Surface - *CC BY 4.0*
- 13.3: Impact Craters - *CC BY 4.0*
- 13.4: The Origin of the Moon - *CC BY 4.0*
- 13.5: Mercury - *CC BY 4.0*
- 13.6: Key Terms - *CC BY 4.0*

◦ 14: Venus - *CC BY 4.0*
- 14.1: The Nearest Planets - An Overview - *CC BY 4.0*
- 14.2: The Geology of Venus - *CC BY 4.0*
- 14.3: The Massive Atmosphere of Venus - *CC BY 4.0*

◦ 15: Mars - *CC BY 4.0*
- 15.1: The Geology of Mars - *CC BY 4.0*
- 15.2: Water and Life on Mars - *CC BY 4.0*
- 15.3: Divergent Planetary Evolution - *CC BY 4.0*
- 15.4: Key Terms - *CC BY 4.0*

◦ 16: Giant Planets - *CC BY 4.0*
- 16.1: Exploring the Outer Planets - *CC BY 4.0*
- 16.2: The Giant Planets - *CC BY 4.0*
- 16.3: Atmosphere of the Giant Planets - *CC BY 4.0*
- 16.4: Key Terms - *CC BY 4.0*

◦ 17: Moons - *CC BY 4.0*
- 17.1: Ring and Moon Systems Introduced - *CC BY 4.0*
- 17.2: The Gailean Moons of Jupiter - *CC BY 4.0*
- 17.3: Titan and Triton - *CC BY 4.0*

◦ 18: Asteroids and Comets - *CC BY 4.0*
- 18.1: Asteroids - *CC BY 4.0*
- 18.2: Asteroids and Planetary Defense - *CC BY 4.0*
- 18.3: The "Long-Haired" Comets - *CC BY 4.0*
- 18.4: Key Terms - *CC BY 4.0*

◦ 19: The Sun - *CC BY 4.0*
- 19.1: The Structure and Composition of the Sun - *CC BY 4.0*
- 19.2: The Solar Cycle - *CC BY 4.0*
- 19.3: Solar Activity above the Photosphere - *CC BY 4.0*
- 19.4: Mass, Energy, and the Theory of Relativity - *CC BY 4.0*
- 19.5: The Solar Interior - Theory - *CC BY 4.0*
- 19.6: The Solar Interior - Observations - *CC BY 4.0*
- 19.7: Key Terms - *CC BY 4.0*

◦ 20: The Stars - *CC BY 4.0*
- 20.1: The Brightness of Stars - *CC BY 4.0*
- 20.2: Colors of Stars - *CC BY 4.0*
- 20.3: The Spectra of Stars (and Brown Dwarfs) - *CC BY 4.0*
- 20.4: Measuring Stellar Masses - *CC BY 4.0*
- 20.5: The H-R Diagram - *CC BY 4.0*

◦ 21: Exoplanets - *CC BY 4.0*
- 21.1: Planets beyond the Solar System- Search and Discovery - *CC BY 4.0*
- 21.2: Exoplanets Everywhere - What we are Learning - *CC BY 4.0*

◦ Back Matter - *CC BY 4.0*
- Index - *CC BY 4.0*
- Glossary - *CC BY 4.0*
- Detailed Licensing - *CC BY 4.0*
- Detailed Licensing - *CC BY 4.0*

This page titled Astronomy (Mastrapa) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rachel Mastrapa.