# INTRODUCTION TO INTERDISCIPLINARY PHYSICAL SCIENCE



*Arunava Roy*
University of Oklahoma



# Introduction to Interdisciplinary Physical Science

Arunava Roy
University of Oklahoma

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 10/14/2026

# TABLE OF CONTENTS


**Licensing**

**1: Module 1a-Scientific Method**

1.1: Prelude to Science and the Realm of Physics, Physical Quantities, and Units

1.2: Observing the World

**1.3: Developing the Scientific Method**

1.3.1: Careful Observations- Documentation, Measurements, and Patterns

1.3.2: Experiments

1.3.3: A Scientific Method

**1.4: Processes of Science**

1.4.1: Laws and Theories in Science

1.4.2: The Scientific Process

**2: Module 1b-Motion**

2.1: Physics- An Introduction

2.2: Prelude to One-Dimensional Kinematics

2.3: Displacement

2.4: Vectors, Scalars, and Coordinate Systems

2.5: Time, Velocity, and Speed

2.6: Acceleration

2.7: Falling Objects

2.8: Prelude to Two-Dimensional Kinematics

2.9: Kinematics in Two Dimensions - An Introduction

2.10: Projectile Motion

2.11: Prelude to Uniform Circular Motion and Gravitation

2.12: Rotation Angle and Angular Velocity

2.13: Centripetal Acceleration

**3: Module 2-Motion and Forces**

3.1: Prelude to Dynamics- Newton's Laws of Motion

3.2: Development of Force Concept

3.3: Newton's First Law of Motion - Inertia

3.4: Newton's Second Law of Motion- Concept of a System

3.5: Newton's Third Law of Motion- Symmetry in Forces

3.6: Summary of Newton's Laws

3.7: Newton's Universal Law of Gravitation

**4: Module 3-Energy and Momentum**

4.1: Prelude to Work, Energy, and Energy Resources

4.2: Work- The Scientific Definition

4.3: Kinetic Energy and the Work-Energy Theorem
4.4: Gravitational Potential Energy
4.5: Conservative Forces and Potential Energy
4.6: Conservation of Energy
4.7: Power
4.8: Work, Energy, and Power in Humans
4.9: World Energy Use
4.10: Prelude to Linear Momentum and Collisions
4.11: Linear Momentum and Force
4.12: Conservation of Momentum
4.13: Energy and Momentum

**5: Module 4-Thermodynamics and Waves**

5.1: Prelude to Temperature, Kinetic Theory, and the Gas Laws
5.2: Temperature
5.3: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature
5.4: Phase Changes
5.5: Prelude to Heat and Heat Transfer Methods
5.6: Heat
5.7: Temperature Change and Heat Capacity
5.8: Phase Change and Latent Heat
5.9: Heat Transfer Methods
5.10: Conduction
5.11: Convection
5.12: Radiation
5.13: Prelude to Thermodynamics
5.14: The First Law of Thermodynamics
5.15: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency
5.16: Waves
5.17: Superposition and Interference
5.18: Prelude to the Physics of Hearing
5.19: Sound
5.20: Speed of Sound, Frequency, and Wavelength
5.21: Doppler Effect and Sonic Booms
5.22: Prelude to Electromagnetic Waves
5.23: The Electromagnetic Spectrum

**6: Module 5- Optics**

6.1: Prelude to Geometric Optics
6.2: The Ray Aspect of Light
6.3: The Law of Reflection
6.4: The Law of Refraction

6.5: Dispersion - Rainbows and Prisms
6.6: Image Formation by Lenses
6.7: Image Formation by Mirrors

**7: Module 6a- Electricity, Magnetism**

7.1: Prelude to Electric Charge and Electric Field
7.2: Static Electricity and Charge - Conservation of Charge
7.3: Conductors and Insulators
7.4: Coulomb's Law
7.5: Electric Field- Concept of a Field Revisited
7.7: Electric Potential Energy- Potential Difference
7.8: Prelude to Electric Current, Resistance, and Ohm's Law
7.9: Current
7.10: Ohm’s Law - Resistance and Simple Circuits
7.11: Resistance and Resistivity
7.12: 20.4 Electric Power and Energy
7.13: Alternating Current versus Direct Current
7.14: Electric Hazards and the Human Body
7.15: Prelude to Circuits and DC Instruments
7.16: Resistors in Series and Parallel
7.17: Prelude to Magnetism
7.18: Magnets
7.19: Ferromagnets and Electromagnets
7.20: Magnetic Fields and Magnetic Field Lines
7.21: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies
7.06: Introduction to Electric Potential and Electric Energy

**8: Module 6b-Introduction to Chemistry**

8.1: Chemistry in Context
8.4: Prelude to Atoms
8.5: Early Ideas in Atomic Theory
8.6: Evolution of Atomic Theory
8.7: Atomic Structure and Symbolism
8.8: Chemical Formulas
8.9: The Periodic Table
**8.10: Organizing the Elements**
8.10.1: Mendeleev's Periodic Table
8.10.2: Periodic Law
8.10.3: Modern Periodic Table- Periods and Groups
8.11: Chemical Nomenclature
8.02: Phases and Classification of Matter
8.03: Physical and Chemical Properties

## 9: Module 7-Atomic Structures and Periodicity

9.1: Introduction to Atoms and Elements

9.2: Modern Understandings of Elements

**9.3: The Development of Atomic Theory**

9.3.1: Law of Conservation of Mass

9.3.2: Law of Definite Proportions

9.3.3: Law of Multiple Proportions

9.3.4: Dalton's Atomic Theory

9.4: Prelude to Chemical Bonding and Molecular Geometry

9.5: Ionic Bonding

9.6: Ionic Bonding (1)

9.7: Covalent Bonds and Molecules

9.8: Covalent Bonding

9.9: Prelude to Liquids and Solids

9.10: Intermolecular Forces

9.11: Intermolecular Forces (1)

## 10: Module 8- Chemical Reactions and Organic Chemistry

10.1: Introduction to Chemical Changes

10.2: Aspects of a Chemical Change

10.3: Chemical Equations- Representing Chemical Changes

10.4: Formula Mass and the Mole Concept

10.5: Balancing Equations

10.6: Energy and Chemical Reactions

10.7: Classifying Reactions

10.8: Acid Base Reactions

10.9: Polymerization Reactions

10.10: Determining Empirical and Molecular Formulas

10.11: Prelude to Stoichiometry

10.12: Balancing Equations (1)

10.13: Writing and Balancing Chemical Equations

10.14: Prelude to Kinetics

10.15: Chemical Reaction Rates

10.16: Brønsted-Lowry Acids and Bases

10.17: pH and pOH

**10.18: Organic Chemistry**

10.18.1: Introduction

10.18.2: Hydrocarbons

10.18.3: Alcohols and Ethers

10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters

10.18.5: Amines and Amides

Index
Glossary
Detailed Licensing

This page titled Introduction to Interdisciplinary Physical Science is shared under a license and was authored, remixed, and/or curated by LibreTexts.

# 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 Public Domain license and was authored, remixed, and/or curated by LibreTexts.

# 1: MODULE 1A-SCIENTIFIC METHOD


1.1: Prelude to Science and the Realm of Physics, Physical Quantities, and Units

1.2: Observing the World

**1.3: Developing the Scientific Method**

1.3.1: Careful Observations- Documentation, Measurements, and Patterns

1.3.2: Experiments

1.3.3: A Scientific Method

**1.4: Processes of Science**

1.4.1: Laws and Theories in Science

1.4.2: The Scientific Process



This page titled 1: Module 1a-Scientific Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

# 1.1: Prelude to Science and the Realm of Physics, Physical Quantities, and Units

What is your first reaction when you hear the word “physics”? Did you imagine working through difficult equations or memorizing formulas that seem to have no real use in life outside the physics classroom? Many people come to the subject of physics with a bit of fear. But as you begin your exploration of this broad-ranging subject, you may soon come to realize that physics plays a much larger role in your life than you first thought, no matter your life goals or career choice.



Figure 1.1.1: Galaxies are as immense as atoms are small. Yet the same laws of physics describe both, and all the rest of nature—an indication of the underlying unity in the universe. The laws of physics are surprisingly few in number, implying an underlying simplicity to nature’s apparent complexity. (credit: NASA, JPL-Caltech, P. Barmby, Harvard-Smithsonian Center for Astrophysics)

For example, take a look at the image above. This image is of the Andromeda Galaxy, which contains billions of individual stars, huge clouds of gas, and dust. Two smaller galaxies are also visible as bright blue spots in the background. At a staggering 2.5 million light years from the Earth, this galaxy is the nearest one to our own galaxy (which is called the Milky Way). The stars and planets that make up Andromeda might seem to be the furthest thing from most people’s regular, everyday lives. But Andromeda is a great starting point to think about the forces that hold together the universe. The forces that cause Andromeda to act as it does are the same forces we contend with here on Earth, whether we are planning to send a rocket into space or simply raise the walls for a new home. The same gravity that causes the stars of Andromeda to rotate and revolve also causes water to flow over hydroelectric dams here on Earth. Tonight, take a moment to look up at the stars. The forces out there are the same as the ones here on Earth. Through a study of physics, you may gain a greater understanding of the interconnectedness of everything we can see and know in this universe.

Think now about all of the technological devices that you use on a regular basis. Computers, smart phones, GPS systems, MP3 players, and satellite radio might come to mind. Next, think about the most exciting modern technologies that you have heard about in the news, such as trains that levitate above tracks, “invisibility cloaks” that bend light around them, and microscopic robots that fight cancer cells in our bodies. All of these groundbreaking advancements, commonplace or unbelievable, rely on the principles of physics. Aside from playing a significant role in technology, professionals such as engineers, pilots, physicians, physical therapists, electricians, and computer programmers apply physics concepts in their daily work. For example, a pilot must understand how wind forces affect a flight path and a physical therapist must understand how the muscles in the body experience forces as they move and bend. As you will learn in this text, physics principles are propelling new, exciting technologies, and these principles are applied in a wide range of careers.

In this text, you will begin to explore the history of the formal study of physics, beginning with natural philosophy and the ancient Greeks, and leading up through a review of Sir Isaac Newton and the laws of physics that bear his name. You will also be introduced to the standards scientists use when they study physical quantities and the interrelated system of measurements most of the scientific community uses to communicate in a single mathematical language. Finally, you will study the limits of our ability to be accurate and precise, and the reasons scientists go to painstaking lengths to be as clear as possible regarding their own limitations.

This page titled 1.1: Prelude to Science and the Realm of Physics, Physical Quantities, and Units is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 1.0: Prelude to Science and the Realm of Physics, Physical Quantities, and Units by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 1.2: Observing the World

**Learning Objectives**

- Explain the importance of observation in science and some of the obstacles that have historical limited observations.

What is the role of observation in our lives? We can use our senses, not just our vision, but our hearing, smell, taste, and touch to learn more about the world. We might do this for simple enjoyment, or we might do it because we want a better understanding of how the world works. If we observe something because we want to understand the world, we engage in an act of science, even if we do not realize it yet.

Observations were not necessarily always held in such high regard, however. During the time of Greek philosophy, around 400 BCE, Plato suggested that we cannot truly trust our senses. Plato claimed that to truly understand the world we must apply reason and believe that we can only see the shadows of some more true reality. But reason without observation is merely conjecture. Plato's student Aristotle wrote that men have more teeth than women. This is a statement that is obviously not based on true observations, but on reasoning alone. (And faulty reasoning at that!) When our reasoning is not based on observations about the natural world it can go in incorrect, and sometimes dangerous, directions. Some credit Aristotle with the beginning of natural philosophy, the precursor to science. However, the approach did not give as much primacy to observations as we now understand them to have in science.

During the medieval period (1100-1700 CE) in Europe, scholasticism became a dominant mode of thought. Like Greek philosophy before it, this mode of thought was also based more on reasoning than on observations of nature. In this case, the reasoning was related to trying to reconcile these same ideas from Greek philosophy with religious texts and writings. Scholastics debated about things such as the bodily functions of angels among other ideas we would now find silly (possibly even another debate about teeth, in this case those of a horse, though it is unclear if this story is factual.) The scholastic tradition did give rise to universities in Europe, and it was from these universities that scientific thought would eventually emerge. However the tradition itself was about appealing to an authority and particular methods of debate rather than about the work of observation and inquiry that would later develop in science.

Throughout time and across many cultures, observations about the world have led to understanding how it works. Sometimes these observations led to useful information such as the knowledge of which plants to eat and which to avoid. Sometimes these observations were paired with ideas about what might have caused a natural effect, perhaps a god throwing around lightning bolts because he is upset about something. Observation itself is the first step of science. But for the field of science to fully develop as it is today, additional steps were also necessary.

## Section Summary

- Observation is the first step of science.

This page titled 1.2: Observing the World is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.3: DEVELOPING THE SCIENTIFIC METHOD

1.3.1: Careful Observations- Documentation, Measurements, and Patterns
1.3.2: Experiments
1.3.3: A Scientific Method

This page titled 1.3: Developing the Scientific Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.3.1: Careful Observations- Documentation, Measurements, and Patterns

**Learning Objectives**

- Explain the importance of documentation, measurement and pattern recognition in science.
- Correctly document measurements.
- Express patterns qualitatively, graphically, and mathematically.

It takes better observations to have a better understanding of the natural world. There are several ways that we can do this, many of which pre-date what we now refer to as science.

## Documentation

It's one thing to say that observations do not matter, but it's quite another thing to say that maybe you made an observation but do not remember everything about it correctly. Careful documentation can aid in this process. Maybe you have an interest in starting a regular routine that you think will improve the quality of your life. Documenting the start of this routine and the quality of your life that you experience afterwards might be a way of learning whether the routine is helpful (and of course there are many apps for things like fitness, budgeting, time management, and cooking that have been created to help with this tracking process). The same process can be applied to any observation about the natural world. What time of year did you plant something, and did it grow the way you hoped it would? This information can be helpful to you the next year to make plans for planting again.

If enough people are able to document their observations of the world, this information can be compiled and shared for others to use. Note the difference between compiling observations and compiling ideas (as the scholastics did). If your goal is to understand the world that we live in, the written documentation of what someone else has observed can be quite helpful. Perhaps it was observed in a time or place that you don't have access to. Perhaps the records of someone else's plant growing process will help you to get the results that you want when you plant your own. This sort of documentation is one of the key aspects of science. Documentation is helpful in recognizing patterns in the natural world. As we shall soon see, this recognition of patterns is an important part of science.

Around 1400 CE, during the Ming dynasty, the Yongle Encyclopedia was created. This was one of the largest encyclopedias ever created, and included many subject areas beyond just science. The ability to share information in this way is one of the keys to scientific progress. Scientific knowledge is not built by the works of a single man (as one might assume based on the way Scientific Laws have often been named), but by the collective work of a community of diverse scholars.




Figure 1.3.1.1: A page from the Yongle Encyclopedia from the Chester Beatty Library. Public Domain

## Measurements

Another way to learn more about the world is to make better observations. If you observe something as large, or tall, or tiny those are subjective observations which do not allow for finding patterns in nature. If you can quantify the observation it allows for an increased ability to recognize patterns. Maybe large things that are 8 feet tall follow different patterns than large things that are 10 feet tall. You will not know this if you are not able to measure (and then document) that difference.

People learned how to make repeatable measurements before documented history. This was helpful for economic reasons. If you wanted to sell and/or meal plan the food you had harvested, it was helpful to know how many ephas of grain you had and how many shekels of coin you could receive in return for them. The epha was a measure of volume and the shekel was a measure of mass. In fact, the process of measuring a harvest each year could be considered a scientific observation.

Measurements have become a very key part of science. As we make or report a measurement, it is important that we include both a number and a unit. Think of it like the difference between 100 dollars and 100 cents. If you do not know the unit, you do not actually know anything about the results of the measurement.

One of the things that has allowed science to progress so much is the ability to make better measurements. Sometimes new devices allow us to make better observations that unlock some understanding of the natural world that we did not have before.

> **Key Concept:** Every measurement must include both a number and a unit.

## Patterns

As more observations are documented, and the quality of those observations is improved by the use of measurements, patterns might begin to emerge. These patterns are often the key to understanding the next step in the process of science. As we begin to notice patterns, we can then use our understanding of those patterns to make predictions about what might happen next.

Some of the earliest scientists lived before recorded history and they used this sort of pattern recognition of the natural world for survival. Keeping track of the patterns of both weather and astronomy helped ancient nomadic people decide when it was time to move towards geographies that had better resources for them. By around 700 BCE, Babylonian astronomers had developed an observation based predictive model for the movement of the planets. This may have been the very first scientific revolution.




Figure 1.3.1.2: A Babylonian tablet from 164 BCE recording an observation of what has since been called Halley's comet (British Museum, public domain)

We have many skills to aid with pattern recognition in science that we will introduce throughout this text. There are often trends where two aspects of an object change together. For example, as a child gets older they get taller. One way to measure this trend would be to mark their height on a wall on each of their birthdays. You would likely find that children grow at variable rates (hence the term growth spurt). In this case, a purely qualitative description is sufficient, a sentence such as “children grow as they get older.”

But for some patterns, the trend is more pronounced. For example, if you were to measure the mass and volume of a single aluminum pellet and then repeat this process for 5, 10, and 20 aluminum pellets you would definitely see a trend: as the mass goes up, so does the volume. But you might notice this trend is not quite as variable as that for growing children. At this point we could make a line graph of the data to see what kind of variability exists. The graph would show a straight line connecting the data points. We call this sort of relationship a direct relationship. We can express it as

$$m = kV$$

where m is the mass, V is the volume, and k is some constant that relates these two quantities together. (It turns out the constant k is a property called density, and this is a relationship we will explore in more detail later in this text.)

This practice of changing some aspect of nature and measuring another aspect of it is a key part of scientific experiments that we will explore in more detail in the next subsection. It is also a very powerful tool. If you wanted a certain mass of aluminum pellets but were only able to measure the volume of those pellets, you could use the direct relationship developed here in order to determine what volume of pellets to use. These sorts of relationships in the field of science are what has allowed us to create the high tech society in which we now live.

## Section Summary

- Documentation is a useful aspect of the process of science, and access to information previously documented by others can be just as useful.
- More accurate observations can occur via measurements, and new measuring devices continually expand the scope of science.
- Every measurement contains both a number and a unit.
- Pattern recognition is very important in science, and leads to important descriptions of the relationship between measurements.
- In a direct relationship, two aspects of nature increase together which can be shown graphically or in an equation.

This page titled 1.3.1: Careful Observations- Documentation, Measurements, and Patterns is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.3.2: Experiments

**Learning Objectives**

- Explain the key aspects of an experiment.

Careful experiments such as the one we discussed with the aluminum pellets are a hallmark of science. But science was not always done in this way. Two researchers in the 1600s proposed new ideas for exploring the natural world that turned the focus of science towards experimentation (also referred to as empiricism).

Francis Bacon (1561-1626) advocated for inductive reasoning in contrast to the deductive reasoning that was practiced by the scholastics at the time. Deductive reasoning was based on ideas, whereas inductive reasoning was based on observations. The challenge for Bacon was that if there was some flaw in the ideas that were used to perform the deduction, it would carry over into what was deduced (as we have already seen in this text). Bacon advocated for making several observations and from those observations a cause might be observed to emerge. This was the initial the basis for what we might now call a controlled experiment. In a controlled experiment, we change one of the variables and see what might happen while leaving other variables unexamined.

The year after Bacon died, Robert Boyle was born (1627-1691). There were still many scholastics at the universities who wanted to understand nature via deductive reasoning from ancient texts, but Boyle continued the focus on inductive reasoning and experimentation. Boyle expanded on these ideas to include the need to replicate an experiment, and even to record experiments that failed to achieve the expected result. Boyle's innovations were owing to both his curiosity and his considerable wealth, which allowed him to conduct such costly experiments over and over again. (Imagine how much more science could have been done at the time if everyone had access to the time and resources necessary to do such work?)

Some of what made this idea unique at the time was probably Boyle's considerable wealth which allowed him to pay for these expensive experiments over and over again.

Boyle was credited for a gas law that we will discuss later in this text. The work of discovering it was largely done by others, but Boyle's prestige and wealth resulted in him getting credit for its discovery. On one hand we can see how it might be unfair for Boyle to have this natural law named after him when the work was mostly done by others. On the other hand, it can be thought of as an example of how modern science works: a community of experimentalist arriving at a result through collaboration and sometimes competition.

Through the work of these and other scholars of the time, empiricism began to emerge as the most powerful form of scientific thought. The general process of experimentation involves observing nature, manipulating nature in some way, observing nature after this manipulation, comparing these results and then trying to explain them. A more formalized experimental process will be introduced in the next subsection.

## Section Summary

- In a controlled experiment, one aspect of nature is changed and then observations are made.
- Replication of experiments was one of the initial aspects of the scientific method.
- In an experiment nature is observed, some change is made to it, additional observations are made, and then the results are compared.

This page titled 1.3.2: Experiments is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.3.3: A Scientific Method

**Learning Objectives**

- Explain the steps of the scientific method.
- Correctly apply the following terms to descriptions of a scientific process: observations, hypothesis, experiment, independent variable, dependent variable, experimental group, control group, repetition.

As we shall see soon enough, there are many ways of doing science in the modern world. For much of science there is a very formal way of approaching questions about the natural world referred to as the scientific method. This is the method often taught in public schools. It is convenient at that level in that it provides a finished product at the end and follows a procedure which provides the sort of checklist that makes the lives of students and teachers at least somewhat predictable. It is also a good way of introducing students to the type of thinking and observing of the natural world which is necessary in science.

## The Four Steps of the Scientific Method

The four steps of the scientific method are listed below. To help illustrate the scientific method, an example that an entomologist (a biologist who specializes in insects) might use is given in italics below each step.

### Step 1: Observations & Questions

Observe something in the natural world and ask a question about how it works. The part of the natural world that is observed and investigated is usually the area that the scientist specializes in. An entomologist for example, would ask questions about how insects function.

*"The life cycle of a fruit fly is about 30 days (at 29 degrees Celsius). How do changes in temperature affect the life cycle of a fruit fly?"*

### Step 2: Hypothesis

Make a hypothesis (an educated guess) which attempts to answer the question. A useful hypothesis is a **testable** statement.

*"Decreasing the temperature of a fruit fly's environment will increase the time it takes the fruit fly to complete its life cycle."*

### Step 3: Experiment

Design and carry out an experiment that is capable of testing the hypothesis. In other words, the experiment must be designed so that it will produce results that either clearly support or clearly falsify (disprove) the hypothesis. It helps to use "If-Then" predictions based on your hypothesis.

*"Place 100 fruit flies at 18 degrees Celsius for one generation. Also place 100 fruit flies at 29 degrees Celsius for one generation. If the hypothesis is correct, then the fruit flies that develop at 18 degrees Celsius will complete their life cycle after those fruit flies that are placed at 29 degrees Celsius."*

### Step 4: Analyze Results and State Conclusions

Reject the hypothesis if the results are not consistent with the hypothesis or accept the hypothesis as possibly true if the results are consistent with the hypothesis. Notice that the hypothesis is not "proven to be true" even if the results do support it. This is because there may be explanations other than the hypothesis for the experimental result.

*For example, if the fruit flies placed at 18 degrees Celsius do develop slower, it may be that their food is not as soft making it more difficult for the fruit flies to eat at the lower temperature, causing them to eat less food and thus grow slower.*

If the experimental results do not support the hypothesis, the hypothesis may be modified and additional experiments may be done to test the new or revised hypothesis.

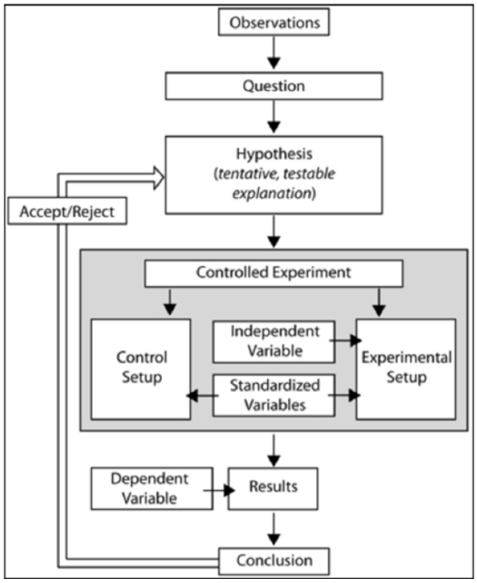



Figure 1.3.3.1: A schematic of the scientific method. Credit: Katherine Harris (Hartnell College) (CC BY-NC-SA)

## Characteristics of a Well-Designed Experiment

The most challenging part of the scientific method is usually the third step, designing and carrying out an experiment to test the hypothesis. A well-designed experiment should include all of the following characteristics:

**1. An independent variable**. The independent variable is the part of the experiment that the scientist changes or manipulates to see what effect occurs.
*"The temperature is the independent variable, since that is what the experiment changes to see its effect."*

**2. A dependent variable**. The dependent variable is the part of the experiment that changes because of the change in the independent variable. In other words, the dependent variable is the effect while the independent variable is the cause.

*"The length of the fruit flies' life cycle is the dependent variable, since the time of development is expected to change because of the temperature."*

**3. An experimental group**. The experimental group is the group of subjects where the independent variable is set to an unusual or test level.

*"The fruit flies placed at 18 degrees Celsius are the experimental group, since the effect of lower temperature on the life cycle of fruit flies is what is being tested."*

**4. A control group**. A control group is the group of subjects in the experiment that the experimental group is compared to. For the control group, the independent variable is set to a normal or usual level (which may be zero, if that is considered a normal level).

*"The fruit flies kept at 29 degrees are the control group, since this is the optimal temperature for fruit fly development."*

Some experiments include things called positive controls and negative controls. These are slightly different than control groups. Positive and negative controls serve to show that the experiment is working correctly. A **positive control** is a part of the experiment that is deliberately designed to give a positive result. It shows that the experiment is capable of producing a positive result when it is supposed to. A **negative control** is a part of the experiment that is designed to give a negative result. It shows that the experiment is capable of producing a negative result when it is supposed to.

**5. The experiment should contain repetition**. This means that there should be more than one subject in the experimental group and the control group. Why? In general, the more repetition, the less likely that your results are due to random chance.

*"The experimental group and the control group each contained 100 fruit flies."*

**6. The experiment should be well-defined**. One aspect of "well-defined" is that the procedure (the steps) must be written down and clearly described. The true test of a well-written experimental procedure is that another scientist could duplicate it exactly using just the written directions. Another aspect of "well-defined" is that everything in the experiment, such as the materials, chemicals, equipment, environmental conditions, and the subjects (the organisms involved in the experiment), should be described as exactly as possible. All of the factors that are kept equal in the experiment and control groups are called **standardized variables**. Another aspect of "well-defined" is that all parts of the experiment should be quantified. This means they should be measured by numbers.

*a) All fruit flies in this experiment are the same species (Drosophila melanogaster). All fruit flies were placed in plastic vials with standard cornmeal media and a cotton stopper. All fruit flies received artificial sunlight for 16 hours per day.*

*b) One group of 100 fruit flies (the experimental group) were placed in an incubator set to 18 degrees Celsius.*

*c) One group of 100 fruit flies (the control group) were placed in an incubator set to 29 degrees Celsius.*

*d) In order to accurately determine the length of the life cycle of the fruit flies, 100 adult flies were allowed to lay eggs for 2 days and then removed from vials. Vials were monitored daily to determine the stage of the life cycle of the offspring of the original adult fruit flies. The end of the life cycle was recorded as the time when half of the fruit flies had undergone eclosion (adults emerge from the pupa case).*

## Section Summary

- In the scientific method, the research makes observations, formulates a testable hypothesis, performs an experiment, and then analyzes the results of the experiment.
- A valid hypothesis is testable, and can be used to rule out a possible explanation for a phenomenon.
- In a controlled experiment, an independent variable is manipulated while a dependent variable is observed. The manipulation is done on an experimental group and then compared to a control group.
- A good experiment should contain repetition and the steps should be clearly defined.

## Contributions and Attributions

This page was remixed by Jamie MacArthur (Madera Community College) from Scientific Method Tutorial by Dr. Katherine Harris

This page titled 1.3.3: A Scientific Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.4: PROCESSES OF SCIENCE

1.4.1: Laws and Theories in Science
1.4.2: The Scientific Process

This page titled 1.4: Processes of Science is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

# 1.4.1: Laws and Theories in Science

**Learning Objectives**

- Identify examples of laws and theories in science.

We have already discussed some of the methods of science, but what is the end result of science? For a single scientist, the end result would be a collection of all of the work they have done over their career, which may have variable results. For the scientific community as a whole, the end result is often going to be laws, theories, and principles. Although there are many scientific laws which bear the name of a single researcher, it is never really a single person who discovers anything, but the scientific community as a whole. (In some fields this is recognized by including multiple contributors in the naming of a phenomena. For example: the Soave-Redlich-Kwong equation of state is used in thermodynamics.) This can be no better epitomized than by considering the words of Isaac Newton, for whom three of the most important laws of physics have been named: "if I have seen further....it is by standing on the shoulders of giants."

Scientists search for answers to questions and solutions to problems by using a procedure called the scientific method, as discussed previously. This procedure consists of making observations, formulating hypotheses, and designing experiments, which in turn lead to additional observations, hypotheses, and experiments in repeated cycles (Figure 1.4.1.1). Eventually, after the work of many scientists, those cycles might result in laws or theories.

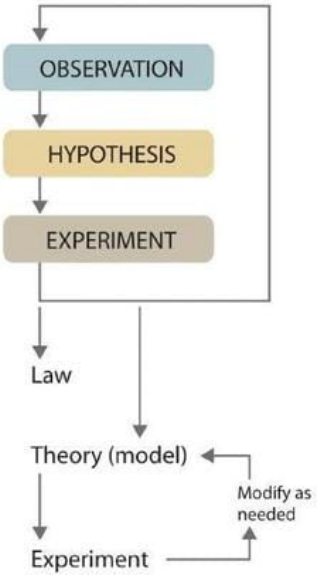



Figure 1.4.1.1: The Scientific Method. As depicted in this flowchart, the scientific method consists of making observations, formulating hypotheses, and designing experiments.

## Scientific Laws

After many experimental data have been collected and analyzed, the scientific community may begin to think that the results are sufficiently reproducible (i.e., dependable) to merit being summarized in a **law**, a verbal or mathematical description of a phenomenon that allows for general predictions. A law simply says what happens; it does not address the question of why. Notice that this is very different from our use of the word law in our every day lives. We might drive a certain speed because it is the law. In this sense we can consider the law to be prescriptive: we do something because the law tells us to. Natural laws are not prescriptive but descriptive. An apple does not fall from a tree because it is written on a sign somewhere. Instead, someone writes down the law of gravity because they observed that the movement of apples and other objects followed a pattern. Laws are unlikely to change greatly over time unless later laws are found which apply to contexts which were not considered at the time.

**Definition: Scientific Law**

A verbal or mathematical description of a phenomenon that allows for general predictions.

One example of a law, the Law of Definite Proportions, was discovered by the French scientist Joseph Proust (1754–1826). This law states that a chemical substance always contains the same proportions of elements by mass. Thus sodium chloride (table salt) always contains the same proportion by mass of sodium to chlorine, in this case 39.34% sodium and 60.66% chlorine by mass, and sucrose (table sugar) is always 42.11% carbon, 6.48% hydrogen, and 51.41% oxygen by mass. The law does not say why the elements exist in the same proportion every time, simply that they do.

## Scientific Theories

Whereas a law states only what happens, a **theory** explains why nature behaves as it does. How is a theory different than a hypothesis? When one makes a hypothesis, there is generally a reason to believe something about what will happen, but it is not necessarily generalizable beyond the particular situation which is being investigated. To be accepted as a scientific theory, there needs to be a preponderance of evidence to support a particular explanation for observed phenomena. Theories are not considered to be more or less valid than laws by those in the scientific community, they are simply different types of information. This is in contrast to how we use these terms in our regular lives. We often use the word theory to represent something that is at best a hypothesis, and more likely a conjecture. For example "My theory is that at the end of this mystery novel, Moriarty will be the culprit." This use of the word theory is merely conjecture about a fictional character, not a combination of explanations coming from multiple experiments performed by several research groups. So when we use the term "it was just a theory" we should be careful not to apply that to scientific theories. Theories don't "graduate" into a law when they have sufficient evidence. They are an explanation as opposed to a description. In a few cases, technological advancements have allowed us to see things that we couldn't before. For example, the atomic theory was considered a well established fact long before the first atoms were observed with electron microscopes. However, we still refer to it as the atomic theory because it has explanatory power for a few scientific laws, such as the law of definite proportions that we just mentioned.

**Definition: Scientific Theory**

When a preponderance of evidence supports a particular explanation for observed phenomena.

The theory developed to explain the extinction of the dinosaurs, for example, is that Earth occasionally encounters small- to medium-sized asteroids, and these encounters may have unfortunate implications for the continued existence of most species. The hypothesis that this might be the case was proposed by Luis and Walter Alvarez based on their discovery of a rare element in a geological site in Italy in the 1980s. Over the course of a few decades, this hypothesis became recognized as a theory as additional data emerged to support it. This theory is consistent with the bulk of evidence amassed to date by multiple researchers over many different disciplines. Figure $1.4.1.2$ summarizes the application of the scientific method in this case.

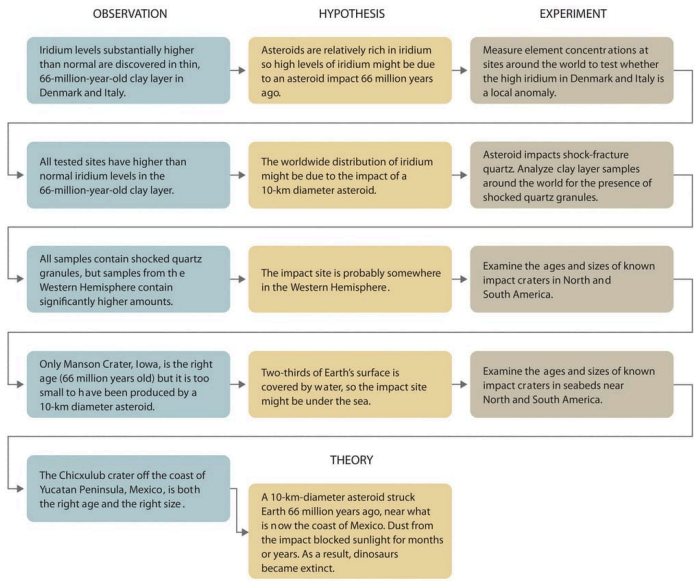



Figure 1.4.1.2: A Summary of How the Scientific Method Was Used in Developing the Asteroid Impact Theory to Explain the Disappearance of the Dinosaurs from Earth

Since scientists can enter the cycle shown in Figure 1.4.1.1 at any point, the actual application of the scientific method to different topics can take many different forms, as we shall see in the next subsection.

It is important to remember that scientists have a tendency to formulate hypotheses in familiar terms simply because it is difficult to propose something that has never been encountered or imagined before. As a result, scientists sometimes discount or overlook unexpected findings that disagree with the basic assumptions behind the hypothesis or theory being tested. Fortunately, truly important findings are immediately subject to independent verification by scientists in other laboratories, which science is a self-correcting discipline. When the Alvarezes originally suggested that an extraterrestrial impact caused the extinction of the dinosaurs, the response was almost universal skepticism and scorn. In only 20 years, however, the persuasive nature of the evidence overcame the skepticism of many scientists, and their initial hypothesis has now evolved into a theory that has revolutionized paleontology and geology.

## Section Summary

- Scientific laws are verbal or mathematical descriptions of a phenomenon that allow for general predictions.
- A scientific theory exists when a preponderance of evidence supports a particular explanation for observed phenomena.

## Contributors

Curated from resources found in this text by Libretexts.

This page titled 1.4.1: Laws and Theories in Science is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

- Section 4: has no license indicated.

# 1.4.2: The Scientific Process

## Learning Objectives

- Explain different steps within the process of science.

The scientific method alone is not enough to make real progress in accumulation of scientific knowledge, but using it as the hub of a cyclic process has led to the massive rate of scientific and technological advancement we have seen over the last century. Science can be thought of as a continuous process guided by the scientific method, as discussed earlier in this text. Modern science is done according to a complex process of checks and balances, such as replication and peer-review. This complexity emerged to help ensure the integrity of scientific results, but the process remains rooted in the basic scientific method. Notice in **Figure** 1.4.2.1: the large amount of work done within the boxes. The results of that work must be combined with the work of many other scientists to result in either a law or theory.

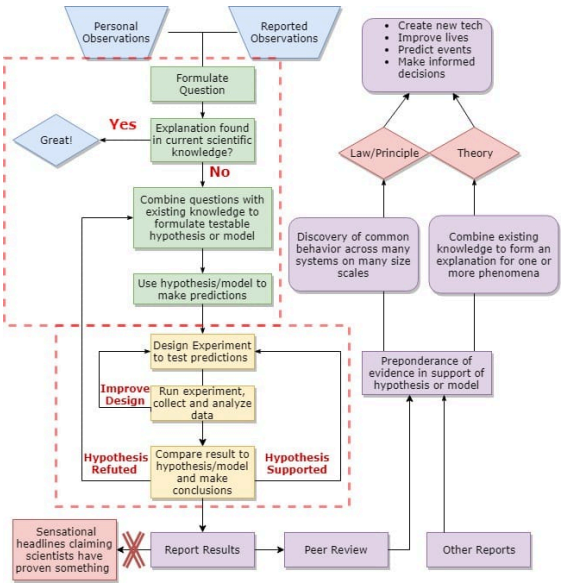



**Figure** 1.4.2.1: The complex modern scientific process built around the basic scientific method (within dashed lines).

In some fields of research, no single scientist can even keep up with the work inside of the dashed lines. The process of science involves many people working together to better understand the world. Some scientists are very good at making observations (often involving complex or expensive instrumentation). Others are great at finding models (usually mathematical, and sometimes requiring expensive computer programs) that explain the data that have already been collected. In some cases it is not possible to design an experiment that will test everything you would like to. For example, we cannot create another Earth populated with dinosaurs in order to test the impact theory we discussed in the previous subsection. And oftentimes the most interesting results are not at all related to the hypothesis but hints at some aspect of reality that nobody would have imagined otherwise. Some scientists have even won Nobel Prizes for discovering things they were not even looking for! Two examples of this are the Big Bang and buckyballs.

The previous diagram illustrates the complex scientific process, but also highlights the basic scientific method on which the whole process is built. After observation, the basic scientific method follows the green and yellow boxes within the dotted line in the diagram below. Generally speaking, the green boxes comprise theoretical science and the yellow boxes comprise experimental science. Because of the degrees of specialization in each of these areas, the need to collaborate with other scientists to complete the process is quite important.

The uncertainty associated with all measurements means that science cannot *prove* anything, despite what the media often claims. Instead, the scientific process produces reviewed and reproduced conclusions that account for uncertainty. Scientific conclusions provide evidence for or against hypotheses.

The following video includes some additional ways of thinking about the process of science:

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- The scientific method guides the progress of much of the progress of science.
- There is often specialization among researchers which focuses their work on a limited number of the steps within the process.
- Not all scientific questions can be answered with experiments, and not all experiments lead to the expected results.
- Science does not prove something is true, but can provide evidence in support of or opposed to a given hypothesis.

## Contributions and Attributions

- Remixed and edited by Jamie MacArthur (Madera Community College)
- Adapted from The Scientific Process by Lawrence Davis (Umpqua Community College)

This page titled 1.4.2: The Scientific Process is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.2: The Scientific Process by Lawrence Davis is licensed CC BY-NC-SA 4.0. Original source: https://openoregon.pressbooks.pub/bpsupmat/.

# 2: MODULE 1B-MOTION


2.1: Physics- An Introduction

2.2: Prelude to One-Dimensional Kinematics

2.3: Displacement

2.4: Vectors, Scalars, and Coordinate Systems

2.5: Time, Velocity, and Speed

2.6: Acceleration

2.7: Falling Objects

2.8: Prelude to Two-Dimensional Kinematics

2.9: Kinematics in Two Dimensions - An Introduction

2.10: Projectile Motion

2.11: Prelude to Uniform Circular Motion and Gravitation

2.12: Rotation Angle and Angular Velocity

2.13: Centripetal Acceleration



This page titled 2: Module 1b-Motion is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Arunava Roy.

# 2.1: Physics- An Introduction

**Learning Objectives**

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

- Explain the difference between a principle and a law.
- Explain the difference between a model and a theory.

The physical universe is enormously complex in its detail. Every day, each of us observes a great variety of objects and phenomena. Over the centuries, the curiosity of the human race has led us collectively to explore and catalog a tremendous wealth of information. From the flight of birds to the colors of flowers, from lightning to gravity, from quarks to clusters of galaxies, from the flow of time to the mystery of the creation of the universe, we have asked questions and assembled huge arrays of facts. In the face of all these details, we have discovered that a surprisingly small and unified set of physical laws can explain what we observe. As humans, we make generalizations and seek order. We have found that nature is remarkably cooperative—it exhibits the *underlying order and simplicity* we so value.



Figure 2.1.1: The flight formations of migratory birds such as Canada geese are governed by the laws of physics. (credit: David Merrett)

It is the underlying order of nature that makes science in general, and physics in particular, so enjoyable to study. For example, what do a bag of chips and a car battery have in common? Both contain energy that can be converted to other forms. The law of conservation of energy (which says that energy can change form but is never lost) ties together such topics as food calories, batteries, heat, light, and watch springs. Understanding this law makes it easier to learn about the various forms energy takes and how they relate to one another. Apparently unrelated topics are connected through broadly applicable physical laws, permitting an understanding beyond just the memorization of lists of facts.

The unifying aspect of physical laws and the basic simplicity of nature form the underlying themes of this text. In learning to apply these laws, you will, of course, study the most important topics in physics. More importantly, you will gain analytical abilities that will enable you to apply these laws far beyond the scope of what can be included in a single book. These analytical skills will help you to excel academically, and they will also help you to think critically in any professional career you choose to pursue. This module discusses the realm of physics (to define what physics is), some applications of physics (to illustrate its relevance to other disciplines), and more precisely what constitutes a physical law (to illuminate the importance of experimentation to theory).

## Science and the Realm of Physics

Science consists of the theories and laws that are the general truths of nature as well as the body of knowledge they encompass. Scientists are continually trying to expand this body of knowledge and to perfect the expression of the laws that describe it. **Physics** is concerned with describing the interactions of energy, matter, space, and time, and it is especially interested in what fundamental mechanisms underlie every phenomenon. The concern for describing the basic phenomena in nature essentially

defines the *realm of physics*.

Physics aims to describe the function of everything around us, from the movement of tiny charged particles to the motion of people, cars, and spaceships. In fact, almost everything around you can be described quite accurately by the laws of physics. Consider a smart phone (Figure 2.1.2). Physics describes how electricity interacts with the various circuits inside the device. This knowledge helps engineers select the appropriate materials and circuit layout when building the smart phone. Next, consider a GPS system. Physics describes the relationship between the speed of an object, the distance over which it travels, and the time it takes to travel that distance. When you use a GPS device in a vehicle, it utilizes these physics equations to determine the travel time from one location to another.



Figure 2.1.2: The Apple "iPhone" is a common smart phone with a GPS function. Physics describes the way that electricity flows through the circuits of this device. Engineers use their knowledge of physics to construct an iPhone with features that consumers will enjoy. One specific feature of an iPhone is the GPS function. GPS uses physics equations to determine the driving time between two locations on a map. (credit: @gletham GIS, Social, Mobile Tech Images)

## Applications of Physics

You need not be a scientist to use physics. On the contrary, knowledge of physics is useful in everyday situations as well as in nonscientific professions. It can help you understand how microwave ovens work, why metals should not be put into them, and why they might affect pacemakers (Figures 2.1.3 and 2.1.4). Physics allows you to understand the hazards of radiation and rationally evaluate these hazards more easily. Physics also explains the reason why a black car radiator helps remove heat in a car engine, and it explains why a white roof helps keep the inside of a house cool. Similarly, the operation of a car's ignition system as well as the transmission of electrical signals through our body's nervous system are much easier to understand when you think about them in terms of basic physics.




Figure 2.1.3: The laws of physics help us understand how common appliances work. For example, the laws of physics can help explain how microwave ovens heat up food, and they also help us understand why it is dangerous to place metal objects in a microwave oven. (credit: MoneyBlogNewz)

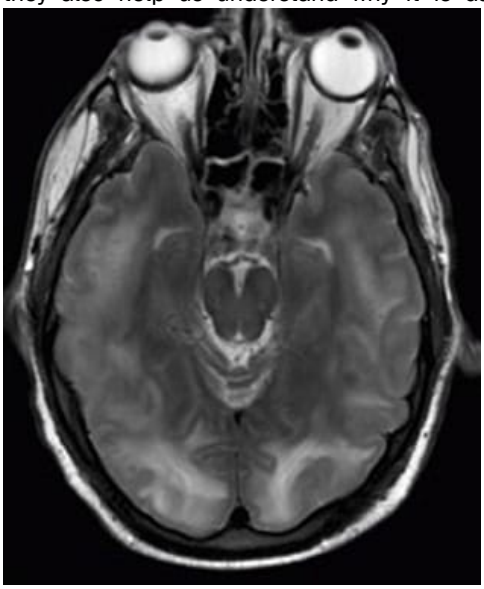


Figure 2.1.4: These two applications of physics have more in common than meets the eye. Microwave ovens use electromagnetic waves to heat food. Magnetic resonance imaging (MRI) also uses electromagnetic waves to yield an image of the brain, from which the exact location of tumors can be determined. (credit: Rashmi Chawla, Daniel Smith, and Paul E. Marik)

Physics is the foundation of many important disciplines and contributes directly to others. Chemistry, for example—since it deals with the interactions of atoms and molecules—is rooted in atomic and molecular physics. Most branches of engineering are applied physics. In architecture, physics is at the heart of structural stability, and is involved in the acoustics, heating, lighting, and cooling of buildings. Parts of geology rely heavily on physics, such as radioactive dating of rocks, earthquake analysis, and heat transfer in the Earth. Some disciplines, such as biophysics and geophysics, are hybrids of physics and other disciplines.

Physics has many applications in the biological sciences. On the microscopic level, it helps describe the properties of cell walls and cell membranes (Figures 2.1.5 and 2.1.6). On the macroscopic level, it can explain the heat, work, and power associated with the human body. Physics is involved in medical diagnostics, such as x-rays, magnetic resonance imaging (MRI), and ultrasonic blood flow measurements. Medical therapy sometimes directly involves physics; for example, cancer radiotherapy uses ionizing radiation. Physics can also explain sensory phenomena, such as how musical instruments make sound, how the eye detects color, and how lasers can transmit information.

It is not necessary to formally study all applications of physics. What is most useful is knowledge of the basic laws of physics and a skill in the analytical methods for applying them. The study of physics also can improve your problem-solving skills. Furthermore,

physics has retained the most basic aspects of science, so it is used by all of the sciences, and the study of physics makes other sciences easier to understand.

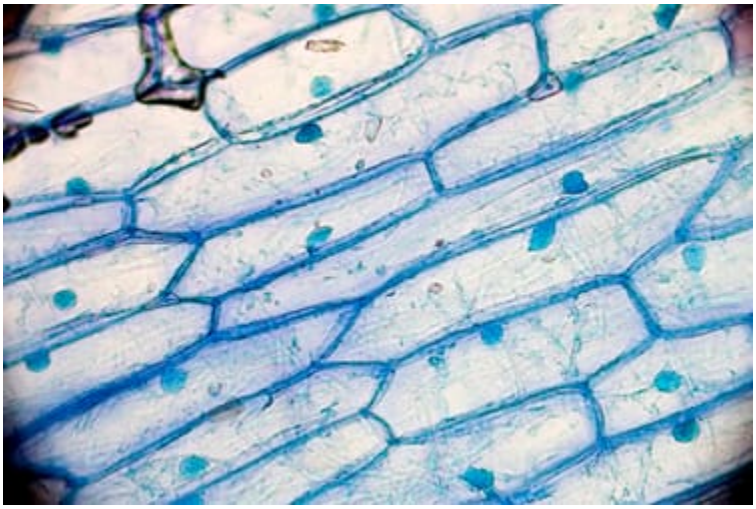


Figure 2.1.5: Physics, chemistry, and biology help describe the properties of cell walls in plant cells, such as the onion cells seen here. (credit: UmbertoSalvagnin)

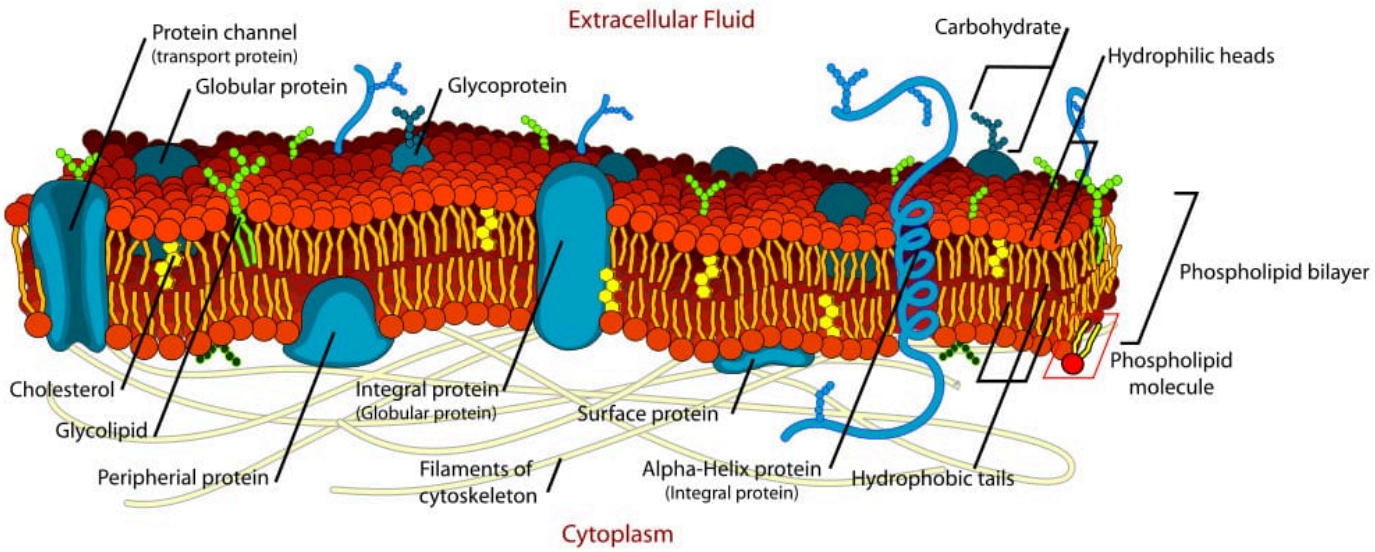



Figure 2.1.6: An artist's rendition of the the structure of a cell membrane. Membranes form the boundaries of animal cells and are complex in structure and function. Many of the most fundamental properties of life, such as the firing of nerve cells, are related to membranes. The disciplines of biology, chemistry, and physics all help us understand the membranes of animal cells. (credit: Mariana Ruiz)

## Models, Theories, and Laws; The Role of Experimentation

The laws of nature are concise descriptions of the universe around us; they are human statements of the underlying laws or rules that all natural processes follow. Such laws are intrinsic to the universe; humans did not create them and so cannot change them. We can only discover and understand them. Their discovery is a very human endeavor, with all the elements of mystery, imagination, struggle, triumph, and disappointment inherent in any creative effort (Figures 2.1.7 and 2.1.8). The cornerstone of discovering natural laws is observation; science must describe the universe as it is, not as we may imagine it to be.

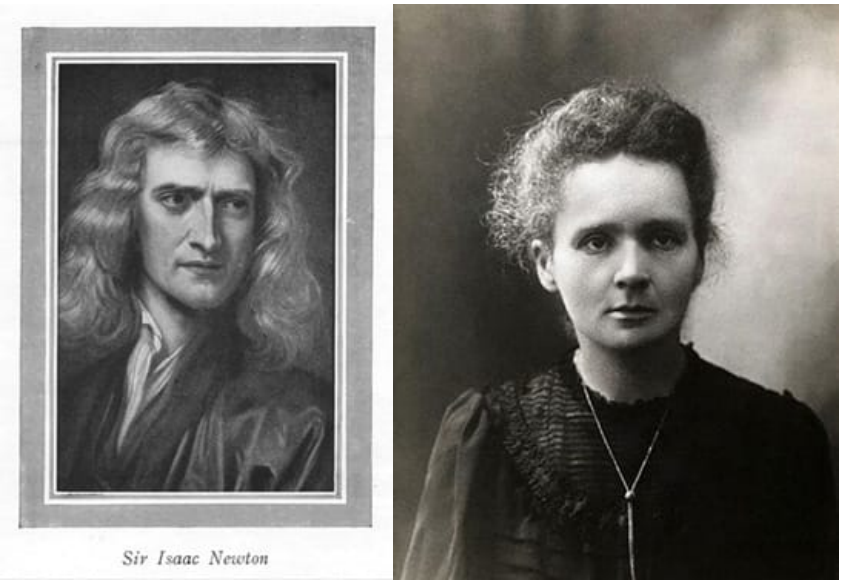



Figure 2.1.8: Marie Curie (1867–1934) sacrificed monetary assets to help finance her early research and damaged her physical well-being with radiation exposure. She is the only person to win Nobel prizes in both physics and chemistry. One of her daughters also won a Nobel Prize. (credit: Wikimedia Commons)

We all are curious to some extent. We look around, make generalizations, and try to understand what we see—for example, we look up and wonder whether one type of cloud signals an oncoming storm. As we become serious about exploring nature, we become more organized and formal in collecting and analyzing data. We attempt greater precision, perform controlled experiments (if we can), and write down ideas about how the data may be organized and unified. We then formulate models, theories, and laws based on the data we have collected and analyzed to generalize and communicate the results of these experiments.

A **model** is a representation of something that is often too difficult (or impossible) to display directly. While a model is justified with experimental proof, it is only accurate under limited situations. An example is the planetary model of the atom in which electrons are pictured as orbiting the nucleus, analogous to the way planets orbit the Sun. (See Figure 9.) We cannot observe electron orbits directly, but the mental image helps explain the observations we can make, such as the emission of light from hot gases (atomic spectra). Physicists use models for a variety of purposes. For example, models can help physicists analyze a scenario and perform a calculation, or they can be used to represent a situation in the form of a computer simulation. A theory is an explanation for patterns in nature that is supported by scientific evidence and verified multiple times by various groups of researchers. Some theories include models to help visualize phenomena, whereas others do not. Newton's theory of gravity, for example, does not require a model or mental image, because we can observe the objects directly with our own senses. The kinetic theory of gases, on the other hand, is a model in which a gas is viewed as being composed of atoms and molecules. Atoms and molecules are too small to be observed directly with our senses—thus, we picture them mentally to understand what our instruments tell us about the behavior of gases.

A **law** uses concise language to describe a generalized pattern in nature that is supported by scientific evidence and repeated experiments. Often, a law can be expressed in the form of a single mathematical equation. Laws and theories are similar in that they are both scientific statements that result from a tested hypothesis and are supported by scientific evidence. However, the designation*law* is reserved for a concise and very general statement that describes phenomena in nature, such as the law that energy is conserved during any process, or Newton's second law of motion, which relates force, mass, and acceleration by the simple equation

$$F = ma. \tag{2.1.1}$$

A theory, in contrast, is a less concise statement of observed phenomena. For example, the Theory of Evolution and the Theory of Relativity cannot be expressed concisely enough to be considered a law. The biggest difference between a law and a theory is that a theory is much more complex and dynamic. A law describes a single action, whereas a theory explains an entire group of related phenomena. And, whereas a law is a postulate that forms the foundation of the scientific method, a theory is the end result of that process.

Less broadly applicable statements are usually called principles (such as Pascal's principle, which is applicable only in fluids), but the distinction between laws and principles often is not carefully made.

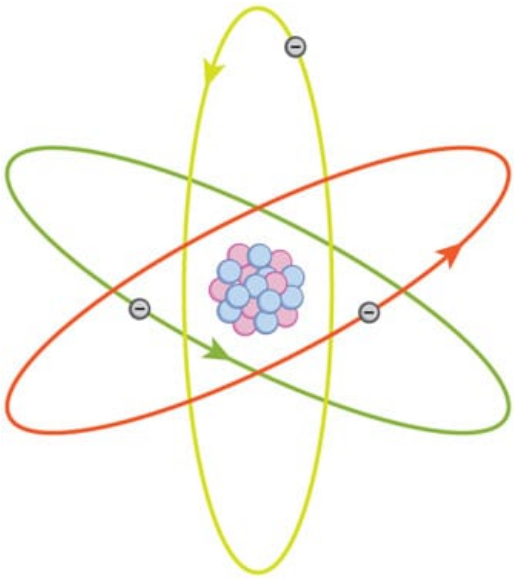


Figure 2.1.9: What is a model? This planetary model of the atom shows electrons orbiting the nucleus. It is a drawing that we use to form a mental image of the atom that we cannot see directly with our eyes because it is too small.

## MODELS, THEORIES, AND LAWS

Models, theories, and laws are used to help scientists analyze the data they have already collected. However, often after a model, theory, or law has been developed, it points scientists toward new discoveries they would not otherwise have made.

The models, theories, and laws we devise sometimes *imply the existence of objects or phenomena as yet unobserved*. These predictions are remarkable triumphs and tributes to the power of science. It is the underlying order in the universe that enables scientists to make such spectacular predictions. However, if *experiment* does not verify our predictions, then the theory or law is wrong, no matter how elegant or convenient it is. Laws can never be known with absolute certainty because it is impossible to perform every imaginable experiment in order to confirm a law in every possible scenario. Physicists operate under the assumption that all scientific laws and theories are valid until a counterexample is observed. If a good-quality, verifiable experiment contradicts a well-established law, then the law must be modified or overthrown completely.

The study of science in general and physics in particular is an adventure much like the exploration of uncharted ocean. Discoveries are made; models, theories, and laws are formulated; and the beauty of the physical universe is made more sublime for the insights gained.

## THE SCIENTIFIC METHOD

As scientists inquire and gather information about the world, they follow a process called the scientific method. This process typically begins with an observation and question that the scientist will research. Next, the scientist typically performs some research about the topic and then devises a hypothesis. Then, the scientist will test the hypothesis by performing an experiment. Finally, the scientist analyzes the results of the experiment and draws a conclusion. Note that the scientific method can be applied to many situations that are not limited to science, and this method can be modified to suit the situation.

Consider an example. Let us say that you try to turn on your car, but it will not start. You undoubtedly wonder: Why will the car not start? You can follow a scientific method to answer this question. First off, you may perform some research to determine a variety of reasons why the car will not start. Next, you will state a hypothesis. For example, you may believe that the car is not starting because it has no engine oil. To test this, you open the hood of the car and examine the oil level. You observe that the oil is at an acceptable level, and you thus conclude that the oil level is not contributing to your car issue. To troubleshoot the issue further, you may devise a new hypothesis to test and then repeat the process again.

# The Evolution of Natural Philosophy into Modern Physics

Physics was not always a separate and distinct discipline. It remains connected to other sciences to this day. The word *physics*comes from Greek, meaning nature. The study of nature came to be called “natural philosophy.” From ancient times through the Renaissance, natural philosophy encompassed many fields, including astronomy, biology, chemistry, physics, mathematics, and medicine. Over the last few centuries, the growth of knowledge has resulted in ever-increasing specialization and branching of natural philosophy into separate fields, with physics retaining the most basic facets (Figures 2.1.10-Figure 2.1.12 ). Physics as it developed from the Renaissance to the end of the 19th century is called **classical physics**. It was transformed into modern physics by revolutionary discoveries made starting at the beginning of the 20th century.




Figure 2.1.12: Niels Bohr (1885–1962) made fundamental contributions to the development of quantum mechanics, one part of modern physics. (credit: United States Library of Congress Prints and Photographs Division)

Classical physics is not an exact description of the universe, but it is an excellent approximation under the following conditions: Matter must be moving at speeds less than about 1% of the speed of light, the objects dealt with must be large enough to be seen with a microscope, and only weak gravitational fields, such as the field generated by the Earth, can be involved. Because humans live under such circumstances, classical physics seems intuitively reasonable, while many aspects of modern physics seem bizarre. This is why models are so useful in modern physics—they let us conceptualize phenomena we do not ordinarily experience. We can relate to models in human terms and visualize what happens when objects move at high speeds or imagine what objects too small to observe with our senses might be like. For example, we can understand an atom’s properties because we can picture it in our minds, although we have never seen an atom with our eyes. New tools, of course, allow us to better picture phenomena we cannot see. In fact, new instrumentation has allowed us in recent years to actually “picture” the atom.

## LIMITS ON THE LAWS OF CLASSICAL PHYSICS

For the laws of classical physics to apply, the following criteria must be met: Matter must be moving at speeds less than about 1% of the speed of light, the objects dealt with must be large enough to be seen with a microscope, and only weak gravitational fields (such as the field generated by the Earth) can be involved.

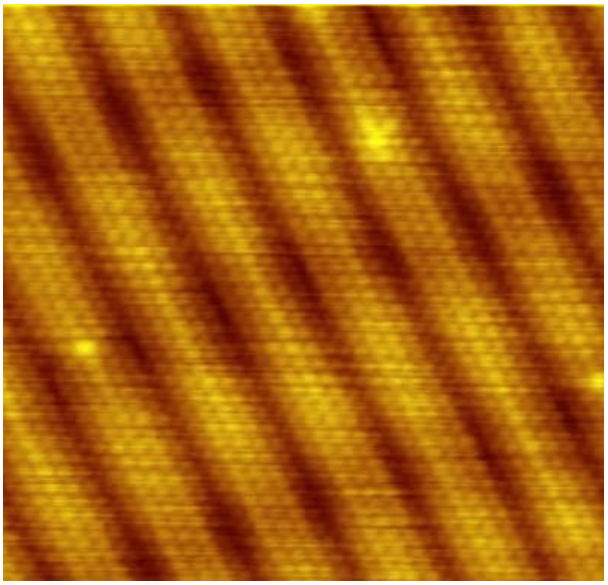


Figure 2.1.13: Using a scanning tunneling microscope (STM), scientists can see the individual atoms that compose this sheet of gold. (credit: Erwinrossen)

Some of the most spectacular advances in science have been made in modern physics. Many of the laws of classical physics have been modified or rejected, and revolutionary changes in technology, society, and our view of the universe have resulted. Like science fiction, modern physics is filled with fascinating objects beyond our normal experiences, but it has the advantage over science fiction of being very real. Why, then, is the majority of this text devoted to topics of classical physics? There are two main reasons: Classical physics gives an extremely accurate description of the universe under a wide range of everyday circumstances, and knowledge of classical physics is necessary to understand modern physics.

**Modern physics** itself consists of the two revolutionary theories, relativity and quantum mechanics. These theories deal with the very fast and the very small, respectively. **Relativity** must be used whenever an object is traveling at greater than about 1% of the speed of light or experiences a strong gravitational field such as that near the Sun. **Quantum mechanics** must be used for objects smaller than can be seen with a microscope. The combination of these two theories is *relativistic quantum mechanics,* and it describes the behavior of small objects traveling at high speeds or experiencing a strong gravitational field. Relativistic quantum mechanics is the best universally applicable theory we have. Because of its mathematical complexity, it is used only when necessary, and the other theories are used whenever they will produce sufficiently accurate results. We will find, however, that we can do a great deal of modern physics with the algebra and trigonometry used in this text.

### Exercise 2.1.1

A friend tells you he has learned about a new law of nature. What can you know about the information even before your friend describes the law? How would the information be different if your friend told you he had learned about a scientific theory rather than a law?

*Answer available at the end of this section.*

### PHET EXPLORATIONS: EQUATION GRAPHER

Learn about graphing polynomials. The shape of the curve changes as the constants are adjusted. View the curves for the individual terms (e.g. $y = bx$) to see how they add to generate the polynomial curve.

## Summary

- Science seeks to discover and describe the underlying order and simplicity in nature.
- Physics is the most basic of the sciences, concerning itself with energy, matter, space and time, and their interactions.
- Scientific laws and theories express the general truths of nature and the body of knowledge they encompass. These laws of

nature are rules that all natural processes appear to follow.

## Glossary

**classical physics**
: physics that was developed from the Renaissance to the end of the 19th century

**physics**
: the science concerned with describing the interactions of energy, matter, space, and time; it is especially interested in what fundamental mechanisms underlie every phenomenon

**model**
: representation of something that is often too difficult (or impossible) to display directly

**theory**
: an explanation for patterns in nature that is supported by scientific evidence and verified multiple times by various groups of researchers

**law**
: a description, using concise language or a mathematical formula, a generalized pattern in nature that is supported by scientific evidence and repeated experiments

**scientific method**
: a method that typically begins with an observation and question that the scientist will research; next, the scientist typically performs some research about the topic and then devises a hypothesis; then, the scientist will test the hypothesis by performing an experiment; finally, the scientist analyzes the results of the experiment and draws a conclusion

**modern physics**
: the study of relativity, quantum mechanics, or both

**relativity**
: the study of objects moving at speeds greater than about 1% of the speed of light, or of objects being affected by a strong gravitational field

**quantum mechanics**
: the study of objects smaller than can be seen with a microscope

## Answers

**Exercise** 2.1.1

Without knowing the details of the law, you can still infer that the information your friend has learned conforms to the requirements of all laws of nature: it will be a concise description of the universe around us; a statement of the underlying rules that all natural processes follow. If the information had been a theory, you would be able to infer that the information will be a large-scale, broadly applicable generalization.

This page titled 2.1: Physics- An Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 1.1: Physics- An Introduction by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.2: Prelude to One-Dimensional Kinematics

Objects are in motion everywhere we look. Everything from a tennis game to a space-probe flyby of the planet Neptune involves motion. When you are resting, your heart moves blood through your veins. And even in inanimate objects, there is continuous motion in the vibrations of atoms and molecules. Questions about motion are interesting in and of themselves: *How long will it take for a space probe to get to Mars? Where will a football land if it is thrown at a certain angle?* But an understanding of motion is also key to understanding other concepts in physics. An understanding of acceleration, for example, is crucial to the study of force.



Figure 2.2.1:The motion of an American kestrel through the air can be described by the bird's displacement, speed, velocity, and acceleration. When it flies in a straight line without any change in direction, its motion is said to be one dimensional. (credit: Vince Maidens, Wikimedia Commons)

Our formal study of physics begins with kinematics which is defined as the *study of motion without considering its causes*. The word "kinematics" comes from a Greek term meaning motion and is related to other English words such as "cinema" (movies) and "kinesiology" (the study of human motion). In one-dimensional kinematics and Two-Dimensional Kinematics we will study only the *motion* of a football, for example, without worrying about what forces cause or change its motion. Such considerations come in other chapters. In this chapter, we examine the simplest type of motion—namely, motion along a straight line, or one-dimensional motion. In Two-Dimensional Kinematics, we apply concepts developed here to study motion along curved paths (two- and three-dimensional motion); for example, that of a car rounding a curve.

This page titled 2.2: Prelude to One-Dimensional Kinematics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 2.0: Prelude to One-Dimensional Kinematics by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.3: Displacement

**Learning Objectives**

- Define position, displacement, distance, and distance traveled.
- Explain the relationship between position and displacement.
- Distinguish between displacement and distance traveled.
- Calculate displacement and distance given initial position, final position, and the path between the two.

These cyclists in Vietnam can be described by their position relative to buildings and a canal. Their motion can be described by their change in position, or displacement, in the frame of reference.



Figure 2.3.1: These cyclists in Vietnam can be described by their position relative to buildings and a canal. Their motion can be described by their change in position, or displacement, in the frame of reference. (credit: Suzan Black, Fotopedia)

## Position

In order to describe the motion of an object, you must first be able to describe its position—where it is at any particular time. More precisely, you need to specify its position relative to a convenient reference frame. Earth is often used as a reference frame, and we often describe the position of an object as it relates to stationary objects in that reference frame. For example, a rocket launch would be described in terms of the position of the rocket with respect to the Earth as a whole, while a professor's position could be described in terms of where she is in relation to the nearby white board. In other cases, we use reference frames that are not stationary but are in motion relative to the Earth. To describe the position of a person in an airplane, for example, we use the airplane, not the Earth, as the reference frame.

## Displacement

If an object moves relative to a reference frame (for example, if a professor moves to the right relative to a white board or a passenger moves toward the rear of an airplane), then the object's position changes. This change in position is known as displacement. The word "displacement" implies that an object has moved, or has been displaced.

## Definition: Displacement

Displacement is the change in position of an object:

$$\Delta x = x_f - x_0,$$

where $\Delta x$ is displacement, $x_f$ is the final position, and $x_0$ is the initial position.

In this text the upper case Greek letter size $\Delta$(delta) always means "change in" whatever quantity follows it; thus, size $\Delta x$ means change in position. Always solve for displacement by subtracting initial position size $x_0$ from final position $x_f$.

Note that the SI unit for displacement is the meter ($m$) (see Section on Physical Quantities and Units), but sometimes kilometers, miles, feet, and other units of length are used. Keep in mind that when units other than the meter are used in a problem, you may need to convert them into meters to complete the calculation.

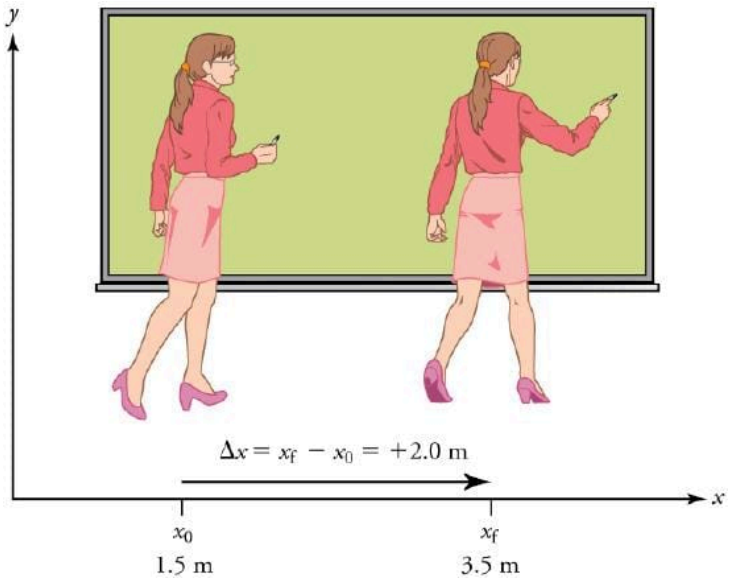



Figure 2.3.2: A professor paces left and right while lecturing. Her position relative to Earth is given by x. The +2 m displacement of the professor relative to Earth is represented by an arrow pointing to the right.

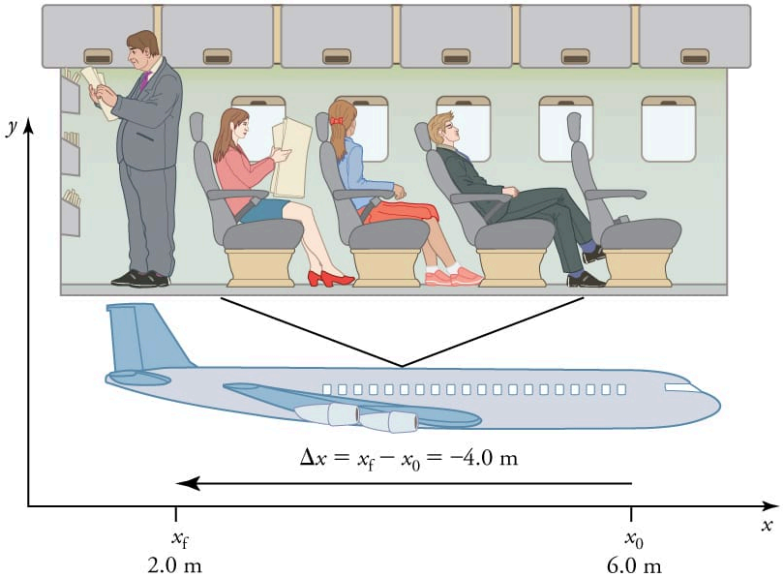



Figure 2.3.3: A passenger moves from his seat to the back of the plane. His location relative to the airplane is given by x. The -4 m displacement of the passenger relative to the plane is represented by an arrow toward the rear of the plane. Notice that the arrow representing his displacement is twice as long as the arrow representing the displacement of the professor (he moves twice as far)

Note that displacement has a direction as well as a magnitude. The professor's displacement in Figure 2.3.2 is 2.0 m to the right, and the airline passenger's displacement is 4.0 m toward the rear in Figure 2.3.3. In one-dimensional motion, direction can be specified with a plus or minus sign. When you begin a problem, you should select which direction is positive (usually that will be to the right or up, but you are free to select positive as being any direction). The professor's initial position is size $x_0$ = 1.5 m and her final position is $x_f$ = 3.5 m. Thus her displacement is

$$\begin{aligned}\Delta x &= x_f - x_0 \\ &= 3.5\,m - 1.5\,m \\ &= +2.0\,m.\end{aligned}$$

In this coordinate system, motion to the right is positive, whereas motion to the left is negative. Similarly, the airplane passenger's initial position is $x_0 = 6.0m$ and his final position is $x_f = 2.0m$, so his displacement is

$$\begin{aligned}\Delta x &= x_f - x_0 \\ &= 2.0\,m - 6.0\,m \\ &= -4.0\,m.\end{aligned}$$

His displacement is negative because his motion is toward the rear of the plane, or in the negative size 12{x} {} direction in our coordinate system.

## Distance

Although displacement is described in terms of direction, distance is not. **Distance** is defined to be the *magnitude or size of displacement between two positions*. Note that the distance between two positions is not the same as the distance traveled between them. **Distance traveled** is *the total length of the path traveled between two positions*. Distance has no direction and, thus, no sign. For example, the distance the professor walks is 2.0 m. The distance the airplane passenger walks is 4.0 m.

### Misconception Alert: Distance Traveled vs. Magnitude of Displacement

It is important to note that the distance traveled, however, can be greater than the magnitude of the displacement (by magnitude, we mean just the size of the displacement without regard to its direction; that is, just a number with a unit). For example, the professor could pace back and forth many times, perhaps walking a distance of 150 m during a lecture, yet still end up only 2.0 m to the right of her starting point. In this case her displacement would be +2.0 m, the magnitude of her displacement would be 2.0 m, but the distance she traveled would be 150 m. In kinematics we nearly always deal with displacement and magnitude of displacement, and almost never with distance traveled. One way to think about this is to assume you marked the start of the motion and the end of the motion. The displacement is simply the difference in the position of the two marks and is independent of the path taken in traveling between the two marks. The distance traveled, however, is the total length of the path taken between the two marks.

### Exercise 2.3.1

A cyclist rides 3 km west and then turns around and rides 2 km east.

a. What is her displacement?
b. What distance does she ride?
c. What is the magnitude of her displacement?

*Answer available at the end of this section.*

## Summary

- Kinematics is the study of motion without considering its causes. In this chapter, it is limited to motion along a straight line, called one-dimensional motion.

- Displacement is the change in position of an object.
- In symbols, displacement $\Delta x$ is defined to be

$$\Delta x = x_f - x_0,$$

where $x_0$ is the initial position and xf is the final position. In this text, the Greek letter $\Delta$ (delta) always means "change in" whatever quantity follows it. The SI unit for displacement is the meter (m). Displacement has a direction as well as a magnitude.

- When you start a problem, assign which direction will be positive.
- Distance is the magnitude of displacement between two positions.
- Distance traveled is the total length of the path traveled between two positions.

# Glossary

**kinematics**
the study of motion without considering its causes

**position**
the location of an object at a particular time

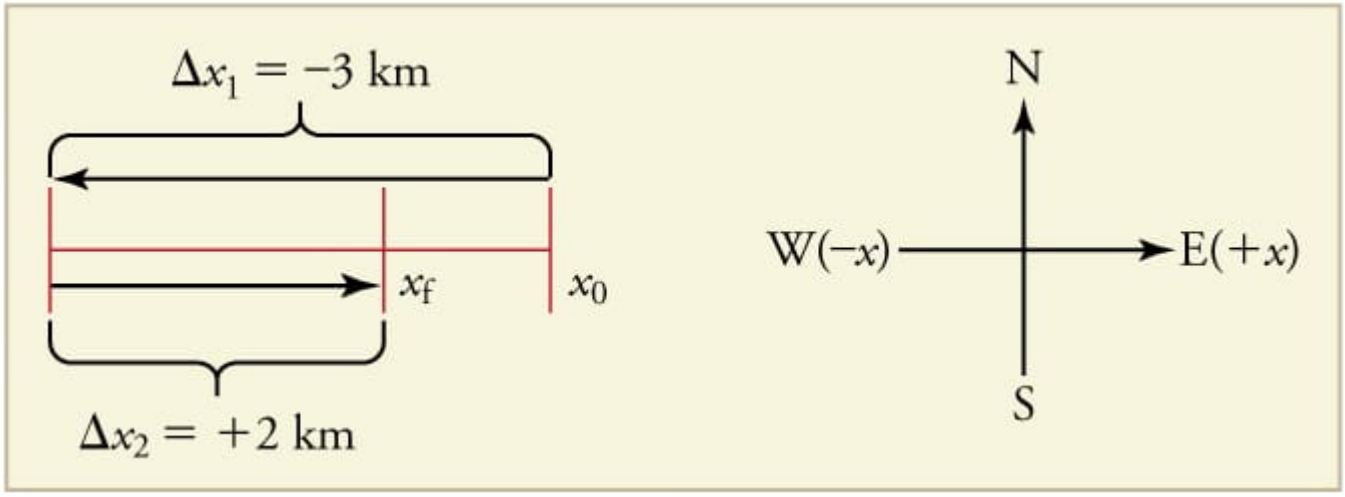
**displacement**
the change in position of an object

**distance**
the magnitude of displacement between two positions

**distance traveled**
the total length of the path traveled between two positions

## Answers

**Exercise** 2.3.1




The rider's displacement is $\Delta_x = x_f - x_0 = -1km$. (The displacement is negative because we take east to be positive and west to be negative.)
The distance traveled is $3km + 2km = 5km$.
The magnitude of the displacement is $1km$.

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

- 2.1: Displacement by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.4: Vectors, Scalars, and Coordinate Systems

**Learning Objectives**

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

- Define and distinguish between scalar and vector quantities.
- Assign a coordinate system for a scenario involving one-dimensional motion.

What is the difference between distance and displacement? Whereas displacement is defined by both direction and magnitude, distance is defined only by magnitude. Displacement is an example of a vector quantity. Distance is an example of a scalar quantity. A vector is any quantity with both *magnitude and direction*. Other examples of vectors include a velocity of 90 km/h east and a force of 500 newtons straight down.




Figure 2.4.1: The motion of this Eclipse Concept jet can be described in terms of the distance it has traveled (a scalar quantity) or its displacement in a specific direction (a vector quantity). In order to specify the direction of motion, its displacement must be described based on a coordinate system. In this case, it may be convenient to choose motion toward the left as positive motion (it is the forward direction for the plane), although in many cases, the x -coordinate runs from left to right, with motion to the right as positive and motion to the left as negative. (credit: Armchair Aviator, Flickr)

The direction of a vector in one-dimensional motion is given simply by a plus (+) or minus (–) sign. Vectors are represented graphically by arrows. An arrow used to represent a vector has a length proportional to the vector's magnitude (e.g., the larger the magnitude, the longer the length of the vector) and points in the same direction as the vector.

Some physical quantities, like distance, either have no direction or none is specified. A scalar is any quantity that has a magnitude, but no direction. For example, a 20°C temperature, the 250 kilocalories (250 Calories) of energy in a candy bar, a 90 km/h speed limit, a person's 1.8 m height, and a distance of 2.0 m are all scalars—quantities with no specified direction. Note, however, that a scalar can be negative, such as a –20°C temperature. In this case, the minus sign indicates a point on a scale rather than a direction. Scalars are never represented by arrows.

## Coordinate Systems for One-Dimensional Motion

In order to describe the direction of a vector quantity, you must designate a coordinate system within the reference frame. For one-dimensional motion, this is a simple coordinate system consisting of a one-dimensional coordinate line. In general, when describing horizontal motion, motion to the right is usually considered positive, and motion to the left is considered negative. With vertical motion, motion up is usually positive and motion down is negative. In some cases, however, as with the jet in Figure {2{.}4{.}\eatSpaces{2} \endSpaces}\), it can be more convenient to switch the positive and negative directions. For example, if you are analyzing the motion of falling objects, it can be useful to define downwards as the positive direction. If people in a race are running to the left, it is useful to define left as the positive direction. It does not matter as long as the system is clear and consistent. Once you assign a positive direction and start solving a problem, you cannot change it.

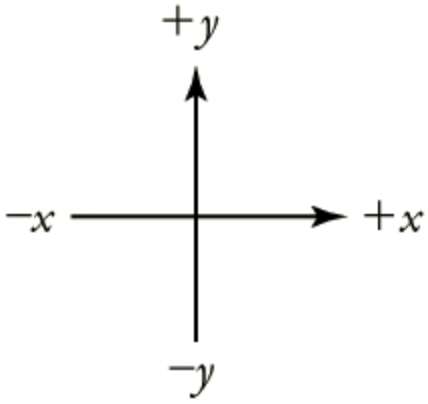



Figure 2.4.2: It is usually convenient to consider motion upward or to the right as positive (+) and motion downward or to the left as negative (–) .

**Exercise 2.4.1**

A person's speed can stay the same as he or she rounds a corner and changes direction. Given this information, is speed a scalar or a vector quantity? Explain.

*Answer available at the end of this section.*

## Summary

- A vector is any quantity that has magnitude and direction.
- A scalar is any quantity that has magnitude but no direction.
- Displacement and velocity are vectors, whereas distance and speed are scalars.
- In one-dimensional motion, direction is specified by a plus or minus sign to signify left or right, up or down, and the like.

## Glossary

**scalar**
: a quantity that is described by magnitude, but not direction

**vector**
: a quantity that is described by both magnitude and direction

## Answers

**Exercise 2.4.1**

Speed is a scalar quantity. It does not change at all with direction changes; therefore, it has magnitude only. If it were a vector quantity, it would change as direction changes (even if its magnitude remained constant).

This page titled 2.4: Vectors, Scalars, and Coordinate Systems is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 2.2: Vectors, Scalars, and Coordinate Systems by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.5: Time, Velocity, and Speed

**Learning Objectives**

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

- Explain the relationships between instantaneous velocity, average velocity, instantaneous speed, average speed, displacement, and time.
- Calculate velocity and speed given initial position, initial time, final position, and final time.
- Derive a graph of velocity vs. time given a graph of position vs. time.
- Interpret a graph of velocity vs. time.

There is more to motion than distance and displacement. Questions such as, “How long does a foot race take?” and “What was the runner’s speed?” cannot be answered without an understanding of other concepts. In this section we add definitions of time, velocity, and speed to expand our description of motion.



Figure 2.5.1: The motion of these racing snails can be described by their speeds and their velocities. (credit: tobitasflickr, Flickr)

## Time

As discussed in Physical Quantities and Units, the most fundamental physical quantities are defined by how they are measured. This is the case with time. Every measurement of time involves measuring a change in some physical quantity. It may be a number on a digital clock, a heartbeat, or the position of the Sun in the sky. In physics, the definition of **time** is simple—time is *change*, or the interval over which change occurs. It is impossible to know that time has passed unless something changes.

The amount of time or change is calibrated by comparison with a standard. The SI unit for time is the second, abbreviated s. We might, for example, observe that a certain pendulum makes one full swing every 0.75 s. We could then use the pendulum to measure time by counting its swings or, of course, by connecting the pendulum to a clock mechanism that registers time on a dial. This allows us to not only measure the amount of time, but also to determine a sequence of events.

How does time relate to motion? We are usually interested in elapsed time for a particular motion, such as how long it takes an airplane passenger to get from his seat to the back of the plane. To find elapsed time, we note the time at the beginning and end of the motion and subtract the two. For example, a lecture may start at 11:00 A.M. and end at 11:50 A.M., so that the elapsed time would be 50 min. **Elapsed time** $\Delta t$ is the difference between the ending time and beginning time,

$$\Delta t = t_f - t_0,$$

where $\Delta t$ is the change in time or elapsed time, $t_f$ is the time at the end of the motion, and $t_0$ is the time at the beginning of the motion. (As usual, the delta symbol, $\Delta$, means the change in the quantity that follows it.)

Life is simpler if the beginning time $t_0$ is taken to be zero, as when we use a stopwatch. If we were using a stopwatch, it would simply read zero at the start of the lecture and 50 min at the end. If $t_0 = 0$, then

$$\Delta t = t_f \equiv t. \tag{2.5.1}$$

In this text, for simplicity's sake,

- motion starts at time equal to zero ($t_0 = 0$)
- the symbol $t$ is used for elapsed time unless otherwise specified ($\Delta t = t_f \equiv t$)

## Velocity

Your notion of velocity is probably the same as its scientific definition. You know that if you have a large displacement in a small amount of time you have a large velocity, and that velocity has units of distance divided by time, such as miles per hour or kilometers per hour.

### Definition: AVERAGE VELOCITY

**Average velocity** is *displacement (change in position) divided by the time of travel,*

$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{x_f - x_0}{t_f - t_0}. \tag{2.5.2}$$

where $\bar{v}$ is the average (indicated by the bar over the $v$) velocity, $\Delta x$ is the change in position (or displacement), and $x_f$ and $x_0$ are the final and beginning positions at times $t_f$ and $t_0$, respectively. If the starting time $t_0$ is taken to be zero, then the average velocity is simply

$$\bar{v} = \frac{\Delta x}{t}. \tag{2.5.3}$$

Notice that this definition indicates that *velocity is a vector because displacement is a vector*. It has both magnitude and direction. The SI unit for velocity is meters per second or m/s, but many other units, such as km/h, mi/h (also written as mph), and cm/s, are in common use. Suppose, for example, an airplane passenger took 5 seconds to move –4 m (the negative sign indicates that displacement is toward the back of the plane). His average velocity would be

$$\bar{v} = \frac{\Delta x}{t} = \frac{-4\,m}{5\,s} = -0.8\,m/s. \tag{2.5.4}$$

The minus sign indicates the average velocity is also toward the rear of the plane.

The average velocity of an object does not tell us anything about what happens to it between the starting point and ending point, however. For example, we cannot tell from average velocity whether the airplane passenger stops momentarily or backs up before he goes to the back of the plane. To get more details, we must consider smaller segments of the trip over smaller time intervals.

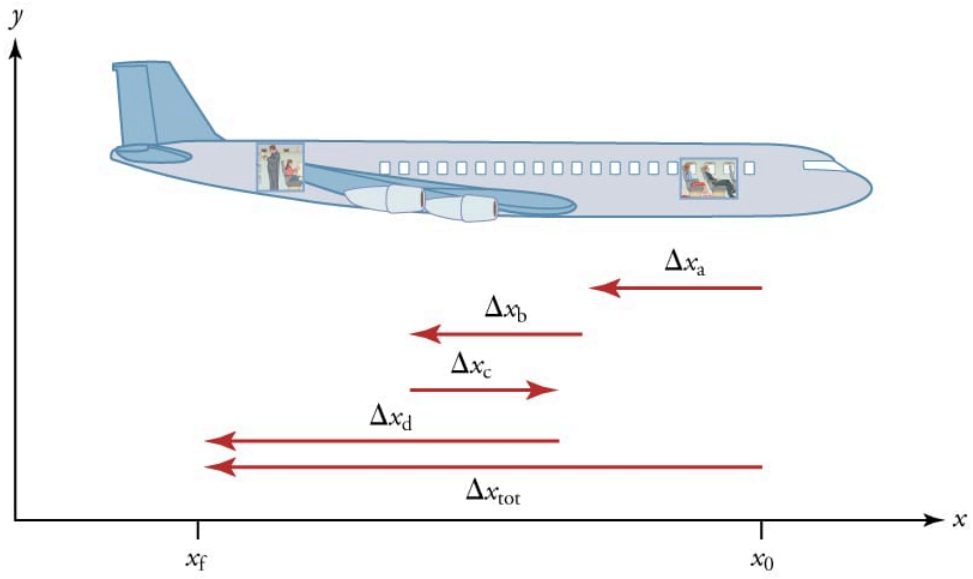



Figure 2.5.2: A more detailed record of an airplane passenger heading toward the back of the plane, showing smaller segments of his trip.

The smaller the time intervals considered in a motion, the more detailed the information. When we carry this process to its logical conclusion, we are left with an infinitesimally small interval. Over such an interval, the average velocity becomes the *instantaneous velocity* or the *velocity at a specific instant*. A car's speedometer, for example, shows the magnitude (but not the direction) of the instantaneous velocity of the car. (Police give tickets based on instantaneous velocity, but when calculating how long it will take to get from one place to another on a road trip, you need to use average velocity.) **Instantaneous velocity** $v$ is the average velocity at a specific instant in time (or over an infinitesimally small time interval).

Mathematically, finding instantaneous velocity, $v$, at a precise instant $t$ can involve taking a limit, a calculus operation beyond the scope of this text. However, under many circumstances, we can find precise values for instantaneous velocity without calculus.

## Speed

In everyday language, most people use the terms "speed" and "velocity" interchangeably. In physics, however, they do not have the same meaning and they are distinct concepts. One major difference is that speed has no direction. Thus *speed is a scalar*. Just as we need to distinguish between instantaneous velocity and average velocity, we also need to distinguish between instantaneous speed and average speed.

**Instantaneous speed** is the magnitude of instantaneous velocity. For example, suppose the airplane passenger at one instant had an instantaneous velocity of –3.0 m/s (the minus meaning toward the rear of the plane). At that same time his instantaneous speed was 3.0 m/s. Or suppose that at one time during a shopping trip your instantaneous velocity is 40 km/h due north. Your instantaneous speed at that instant would be 40 km/h—the same magnitude but without a direction. Average speed, however, is very different from average velocity. **Average speed** is the distance traveled divided by elapsed time.

We have noted that distance traveled can be greater than displacement. So average speed can be greater than average velocity, which is displacement divided by time. For example, if you drive to a store and return home in half an hour, and your car's odometer shows the total distance traveled was 6 km, then your average speed was 12 km/h. Your average velocity, however, was zero, because your displacement for the round trip is zero. (Displacement is change in position and, thus, is zero for a round trip.) Thus average speed is *not* simply the magnitude of average velocity.

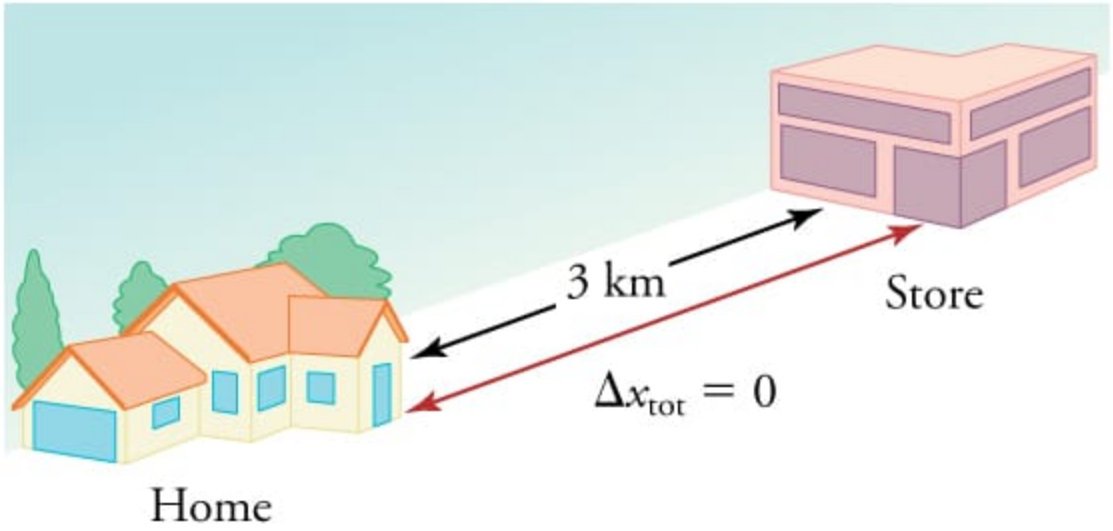



Figure 2.5.3: During a 30-minute round trip to the store, the total distance traveled is 6 km. The average speed is 12 km/h. The displacement for the round trip is zero, since there was no net change in position. Thus the average velocity is zero.

Another way of visualizing the motion of an object is to use a graph. A plot of position or of velocity as a function of time can be very useful. For example, for this trip to the store, the position, velocity, and speed-vs.-time graphs are displayed in Figure 2.5.4. (Note that these graphs depict a very simplified **model** of the trip. We are assuming that speed is constant during the trip, which is unrealistic given that we'll probably stop at the store. But for simplicity's sake, we will model it with no stops or changes in speed. We are also assuming that the route between the store and the house is a perfectly straight line.)

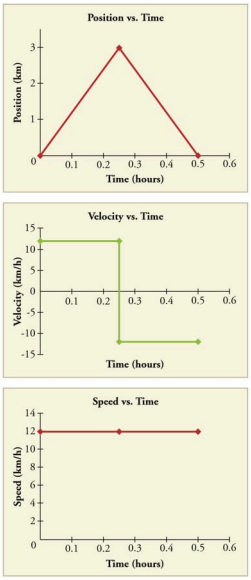



Figure 2.5.4: Position vs. time, velocity vs. time, and speed vs. time on a trip. Note that the velocity for the return trip is negative.

## MAKING CONNECTIONS: TAKE-HOME INVESTIGATION - GETTING A SENSE OF SPEED

If you have spent much time driving, you probably have a good sense of speeds between about 10 and 70 miles per hour. But what are these in meters per second? What do we mean when we say that something is moving at 10 m/s? To get a better sense of what these values really mean, do some observations and calculations on your own:

- calculate typical car speeds in meters per second
- estimate jogging and walking speed by timing yourself; convert the measurements into both m/s and mi/h
- determine the speed of an ant, snail, or falling leaf

**Exercise 2.5.1**

A commuter train travels from Baltimore to Washington, DC, and back in 1 hour and 45 minutes. The distance between the two stations is approximately 40 miles. What is

a. the average velocity of the train, and
b. the average speed of the train in m/s?

*Answer available at the end of this section.*

$$\frac{80miles}{105minutes} \times \frac{5280feet}{1mile} \times \frac{1meter}{3.28feet} \times \frac{1minute}{60seconds} = 20m/s$$

## Summary

- Time is measured in terms of change, and its SI unit is the second (s). Elapsed time for an event is
$$\Delta t = t_f - t_0,$$
where $t_f$ is the final time and $t_0$ is the initial time. The initial time is often taken to be zero, as if measured with a stopwatch; the elapsed time is then just $t$.
- Average velocity $\bar{v}$ is defined as displacement divided by the travel time. In symbols, average velocity is
$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{x_f - x_0}{t_f - t_0}.$$
- The SI unit for velocity is m/s.
- Velocity is a vector and thus has a direction.
- Instantaneous velocity $v$ is the velocity at a specific instant or the average velocity for an infinitesimal interval.
- Instantaneous speed is the magnitude of the instantaneous velocity.
- Instantaneous speed is a scalar quantity, as it has no direction specified.
- Average speed is the total distance traveled divided by the elapsed time. (Average speed is *not* the magnitude of the average velocity.) Speed is a scalar quantity; it has no direction associated with it.

## Glossary

**average speed**
distance traveled divided by time during which motion occurs

**average velocity**
displacement divided by time over which displacement occurs

**instantaneous velocity**
velocity at a specific instant, or the average velocity over an infinitesimal time interval

**instantaneous speed**
magnitude of the instantaneous velocity

**time**
change, or the interval over which change occurs

**model**
simplified description that contains only those elements necessary to describe the physics of a physical situation

**elapsed time**
the difference between the ending time and beginning time

## Answers

**Exercise** 2.5.1

(a) The average velocity of the train is zero because $x_f = x_0$; the train ends up at the same place it starts.

(b) The average speed of the train is calculated below. Note that the train travels 40 miles one way and 40 miles back, for a total distance of 80 miles.

$$\frac{distance}{time} = \frac{80 miles}{105 minutes}$$

This page titled 2.5: Time, Velocity, and Speed is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 2.3: Time, Velocity, and Speed by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.6: Acceleration

## Learning Objectives

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

- Define and distinguish between instantaneous acceleration, average acceleration, and deceleration.
- Calculate acceleration given initial time, initial velocity, final time, and final velocity.

In everyday conversation, to accelerate means to speed up. The accelerator in a car can in fact cause it to speed up. The greater the **acceleration**, the greater the change in velocity over a given time. The formal definition of acceleration is consistent with these notions, but more inclusive.



Figure 2.6.1: A plane decelerates, or slows down, as it comes in for landing in St. Maarten. Its acceleration is opposite in direction to its velocity. (credit: Steve Conry, Flickr)

## Definition: Average Acceleration

The *average acceleration* is the rate at which velocity changes,

$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{v_f - v_0}{t_f - t_0} \tag{2.6.1}$$

where $\bar{a}$ is average acceleration, $v$ is velocity, and $t$ is time. (The bar over the $a$ means *average*.)

Because acceleration is velocity in m/s divided by time in s, the SI units for acceleration are $m/s^2$, meters per second squared or meters per second per second, which literally means by how many meters per second the velocity changes every second.

Recall that velocity is a vector—it has both magnitude and direction. This means that a change in velocity can be a change in magnitude (or speed), but it can also be a change in direction. For example, if a car turns a corner at constant speed, it is accelerating because its direction is changing. The quicker you turn, the greater the acceleration. So there is an acceleration when velocity changes either in magnitude (an increase or decrease in speed) or in direction, or both.

## ACCELERATION AS A VECTOR

Acceleration is a vector in the same direction as the change in velocity, $\Delta v$. Since velocity is a vector, it can change either in magnitude or in direction. Acceleration is therefore a change in either speed or direction, or both.

Keep in mind that although acceleration is in the direction of the *change* in velocity, it is not always in the direction of *motion*. When an object slows down, its acceleration is opposite to the direction of its motion. This is known as **deceleration.**




Figure 2.6.2: A subway train in Sao Paulo, Brazil, decelerates as it comes into a station. It is accelerating in a direction opposite to its direction of motion. (credit: Yusuke Kawasaki, Flickr)

## MISCONCEPTION ALERT: DECELERATION VS. NEGATIVE ACCELERATION

Deceleration always refers to acceleration in the direction opposite to the direction of the velocity. Deceleration always reduces speed. Negative acceleration, however, is acceleration in the negative direction in the chosen coordinate system. Negative acceleration may or may not be deceleration, and deceleration may or may not be considered negative acceleration. If acceleration has the same sign as the velocity, the object is speeding up. If acceleration has the opposite sign as the velocity, the object is slowing down. For example, consider Figure 2.6.2.

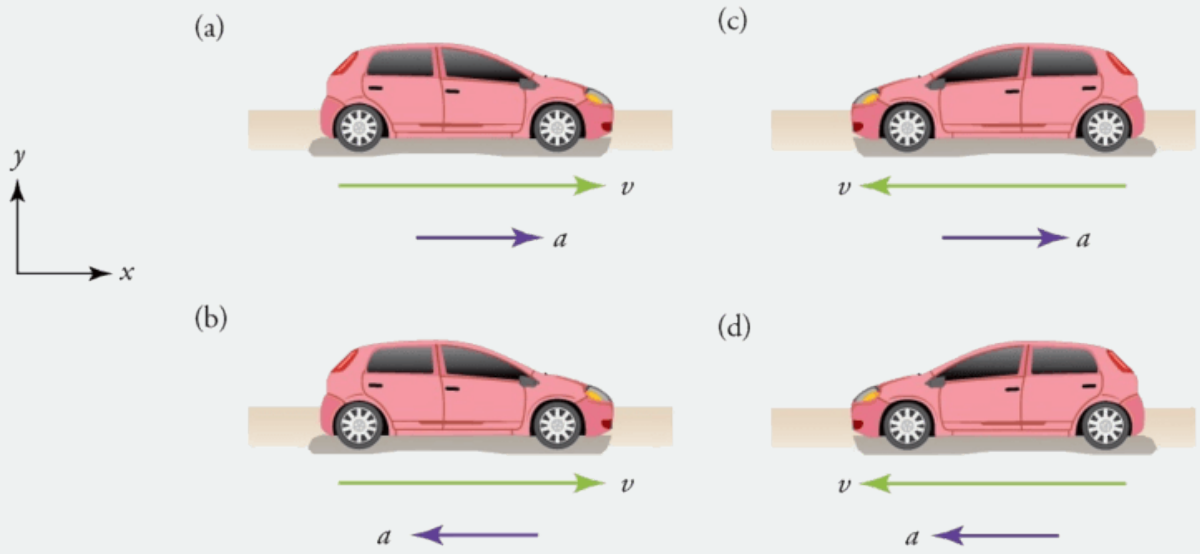



Figure 2.6.3: (a) This car is speeding up as it moves toward the right. It therefore has positive acceleration in our coordinate system. (b) This car is slowing down as it moves toward the right. Therefore, it has negative acceleration in our coordinate system, because its acceleration is toward the left. The car is also decelerating: the direction of its acceleration is opposite to its direction of motion. (c) This car is moving toward the left, but slowing down over time. Therefore, its acceleration is positive in our coordinate system because it is toward the right. However, the car is decelerating because its acceleration is opposite to its motion. (d) This car is speeding up as it moves toward the left. It has negative acceleration because it is accelerating toward the left. However, because its acceleration is in the same direction as its motion, it is speeding up (not decelerating).

## Example 2.6.1: Calculating Acceleration: A Racehorse Leaves the Gate

A racehorse coming out of the gate accelerates from rest to a velocity of 15.0 m/s due west in 1.80 s. What is its average acceleration?

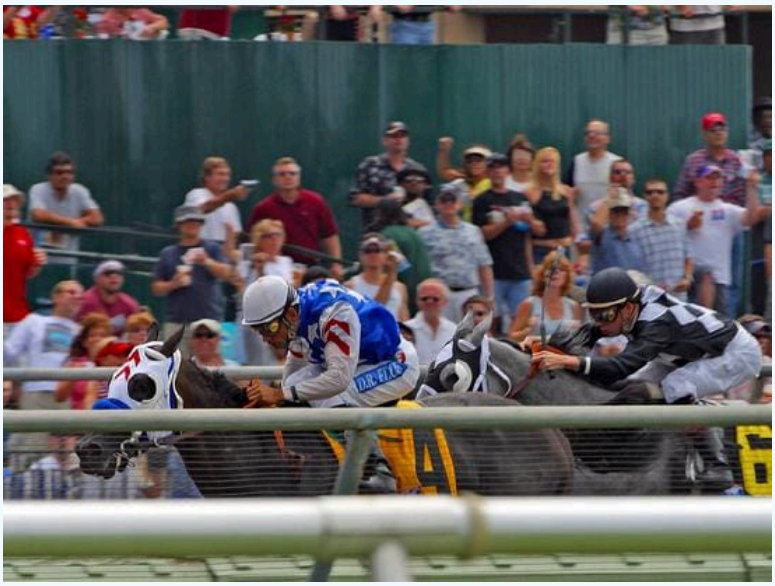


Figure 2.6.4: Two racehorses running toward the left. (credit: Jon Sullivan, PD Photo.org)

**Strategy**

First we draw a sketch and assign a coordinate system to the problem. This is a simple problem, but it always helps to visualize it. Notice that we assign east as positive and west as negative. Thus, in this case, we have negative velocity.

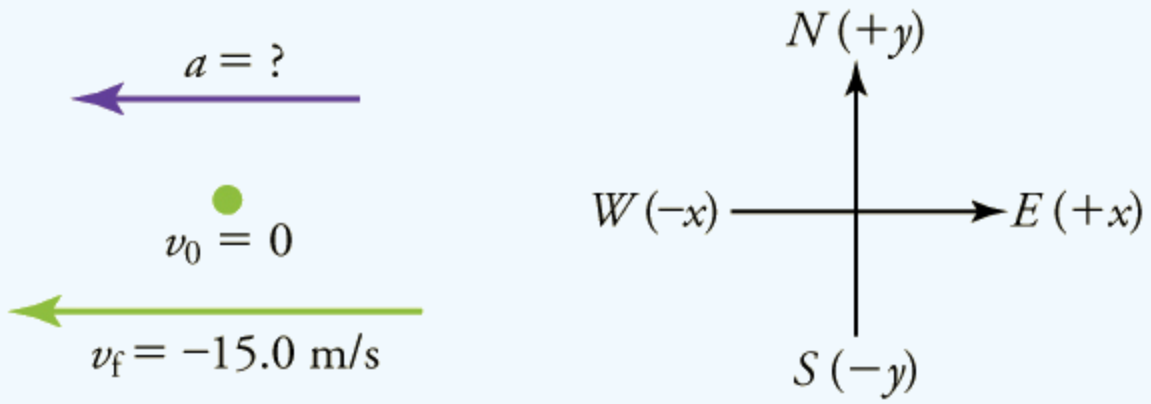



Figure 2.6.5.

We can solve this problem by identifying $\Delta v$ and $\Delta t$ from the given information and then calculating the average acceleration directly from the Equation ???:

$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{v_f - v_0}{t_f - t_0}.$$

**Solution**

1. Identify the knowns. $v_0 = 0, v_f = -15.0 m/s$ (the negative sign indicates direction toward the west), $\Delta t = 1.80 s$.
2. Find the change in velocity. Since the horse is going from zero to $-15.0 m/s$, its change in velocity equals its final velocity:
$$\Delta v = v_f = -15.0\, m/s.$$
3. Plug in the known values ( $\Delta v$ and $\Delta t$ ) and solve for the unknown $\bar{a}$.
$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{-15.0 m/s}{1.80 s} = -8.33 m/s^2.$$

**Discussion**

The negative sign for acceleration indicates that acceleration is toward the west. An acceleration of $8.33\ m/s^2$ due west means that the horse increases its velocity by 8.33 m/s due west each second, that is, 8.33 meters per second per second, which we write as $8.33\ m/s^2$. This is truly an average acceleration, because the ride is not smooth. We shall see later that an acceleration of this magnitude would require the rider to hang on with a force nearly equal to his weight.

## Instantaneous Acceleration

Instantaneous acceleration $a$, or the acceleration at a specific instant in time, is obtained by the same process as discussed for instantaneous velocity in Time, Velocity, and Speed—that is, by considering an infinitesimally small interval of time. How do we find instantaneous acceleration using only algebra? The answer is that we choose an average acceleration that is representative of the motion. Figure $2.6.6$ shows graphs of instantaneous acceleration versus time for two very different motions. In Figure $2.6.6a$, the acceleration varies slightly and the average over the entire interval is nearly the same as the instantaneous acceleration at any time. In this case, we should treat this motion as if it had a constant acceleration equal to the average (in this case about $1.8 m/s^2$ ). In Figure $2.6.6b$, the acceleration varies drastically over time. In such situations it is best to consider smaller time intervals and choose an average acceleration for each. For example, we could consider motion over the time intervals from 0 to 1.0 s and from 1.0 to 3.0 s as separate motions with accelerations of $+3.0\ m/s^2$ and $-2.0\ m/s^2$, respectively.

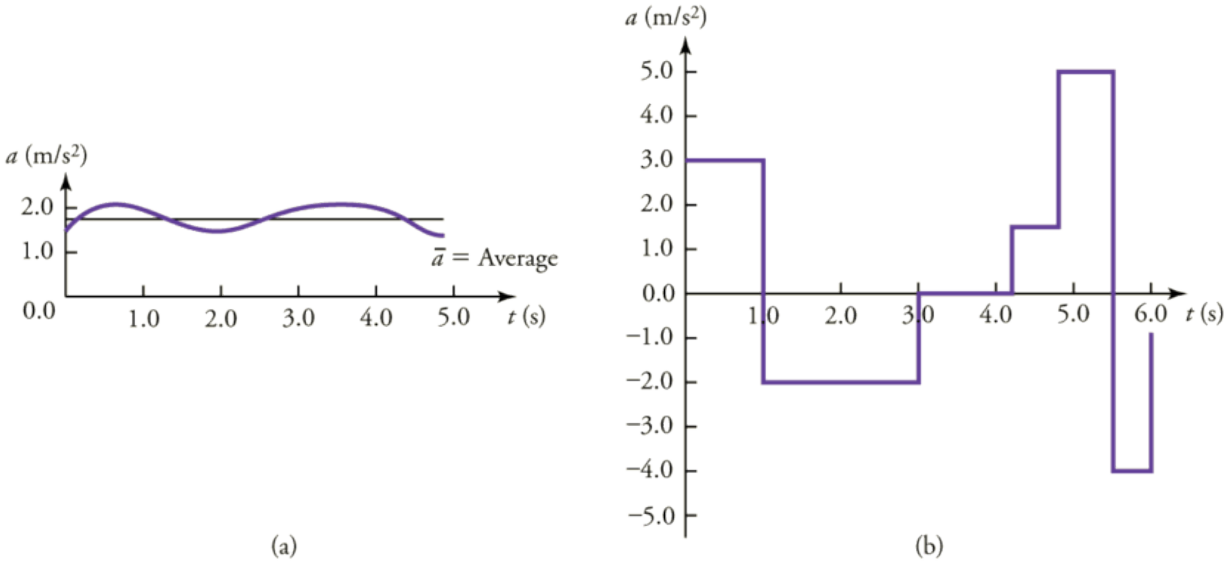



Figure 2.6.6: Graphs of instantaneous acceleration versus time for two different one-dimensional motions. (a) Here acceleration varies only slightly and is always in the same direction, since it is positive. The average over the interval is nearly the same as the acceleration at any given time. (b) Here the acceleration varies greatly, perhaps representing a package on a post office conveyor belt that is accelerated forward and backward as it bumps along. It is necessary to consider small time intervals (such as from 0 to 1.0 s) with constant or nearly constant acceleration in such a situation.

The next several examples consider the motion of the subway train shown in Figure 2.6.7. In Figure 2.6.7$a$ the shuttle moves to the right, and in Figure 2.6.7$b$ it moves to the left. The examples are designed to further illustrate aspects of motion and to illustrate some of the reasoning that goes into solving problems.

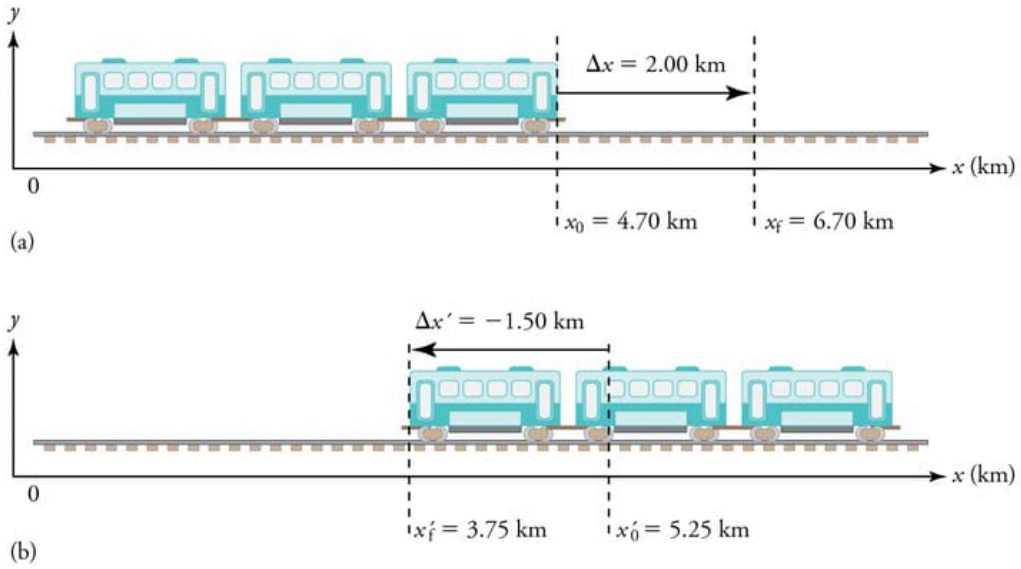



Figure 2.6.7: One-dimensional motion of a subway train considered in Examples 2.6.2 - 2.6.5. Here we have chosen the x -axis so that + means to the right and – means to the left for displacements, velocities, and accelerations. (a) The subway train moves to the right from $x_0$ to $x_f$. Its displacement $\Delta x$ is +2.0 km. (b) The train moves to the left from $x_0$ to $x_f$. Its displacement $\Delta x$ is $-1.5km$. (Note that the prime symbol (′) is used simply to distinguish between displacement in the two different situations. The distances of travel and the size of the cars are on different scales to fit everything into the diagram.)

## Example 2.6.2: Calculating Displacement - A Subway Train

What are the magnitude and sign of displacements for the motions of the subway train shown in parts (a) and (b) of Figure 2.6.7?

**Strategy**

A drawing with a coordinate system is already provided, so we don't need to make a sketch, but we should analyze it to make sure we understand what it is showing. Pay particular attention to the coordinate system. To find displacement, we use the equation $\Delta x = x_f - x_0$. This is straightforward since the initial and final positions are given.

**Solution**

1. Identify the knowns. In the figure we see that $x_f = 6.70\ km$ and $x_0 = 4.70\ km$ for part (a), and $x'_f = 3.75\ km$ and $x'_0 = 5.25\ km$ for part (b).
2. Solve for displacement in part (a).

$$\begin{aligned}\Delta x &= x_f - x_0 \\ &= 6.70\,km - 4.70\,km \\ &= +2.00\,km\end{aligned}$$

3. Solve for displacement in part (b).

$$\begin{aligned}\Delta x' &= x'_f - x'_0 \\ &= 3.75\,km - 5.25\,km \\ &= -1.50\,km\end{aligned}$$

**Discussion**

The direction of the motion in (a) is to the right and therefore its displacement has a positive sign, whereas motion in (b) is to the left and thus has a negative sign.

## Example 2.6.3: Comparing Distance Traveled with Displacement - A Subway Train

What are the distances traveled for the motions shown in parts (a) and (b) of the subway train in Figure 2.6.7?

**Strategy**

To answer this question, think about the definitions of distance and distance traveled, and how they are related to displacement. Distance between two positions is defined to be the magnitude of displacement, which was found in Example 2.6.2. Distance traveled is the total length of the path traveled between the two positions (see Section on Displacement). In the case of the subway train shown in Figure 2.6.7, the distance traveled is the same as the distance between the initial and final positions of the train.

**Solution**

1. The displacement for part (a) was +2.00 km. Therefore, the distance between the initial and final positions was 2.00 km, and the distance traveled was 2.00 km.

2. The displacement for part (b) was $-1.5km$. Therefore, the distance between the initial and final positions was 1.50 km, and the distance traveled was 1.50 km.

**Discussion**

Distance is a scalar. It has magnitude but no sign to indicate direction.

## Example 2.6.4: Calculating Acceleration: A Subway Train Speeding Up

Suppose the train in Figure 2.6.7$a$ accelerates from rest to 30.0 km/h in the first 20.0 s of its motion. What is its average acceleration during that time interval?

**Strategy**

It is worth it at this point to make a simple sketch:

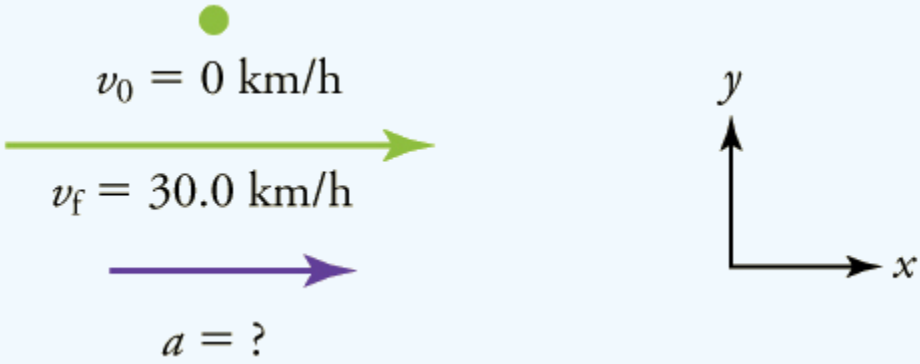



Figure 2.6.8: This problem involves three steps. First we must determine the change in velocity, then we must determine the change in time, and finally we use these values to calculate the acceleration.

**Solution**

1. Identify the knowns. $v_0 = 0$ (the trains starts at rest), $v_f = 30.0km/h$, and $\Delta t = 20.0s$.
2. Calculate $\Delta v$. Since the train starts from rest, its change in velocity is $\Delta v = +30.0km/h$, where the plus sign means velocity to the right.
3. Plug in known values and solve for the unknown, $\bar{a}$.

$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{+30.0km/h}{20.0s}$$

4. Since the units are mixed (we have both hours and seconds for time), we need to convert everything into SI units of meters and seconds. (See Physical Quantities and Units for more guidance.)

$$\bar{a} = \left(\frac{+30km/h}{20.0s}\right)\left(\frac{10^3m}{1km}\right)\left(\frac{1h}{3600s}\right) = 0.417m/s^2$$

**Discussion**

The plus sign means that acceleration is to the right. This is reasonable because the train starts from rest and ends up with a velocity to the right (also positive). So acceleration is in the same direction as the change in velocity, as is always the case.

## Example 2.6.5: Calculate Acceleration

A Subway Train Slowing Down: Now suppose that at the end of its trip, the train in Figure 2.6.7$a$ slows to a stop from a speed of 30.0 km/h in 8.00 s. What is its average acceleration while stopping?

**Strategy**

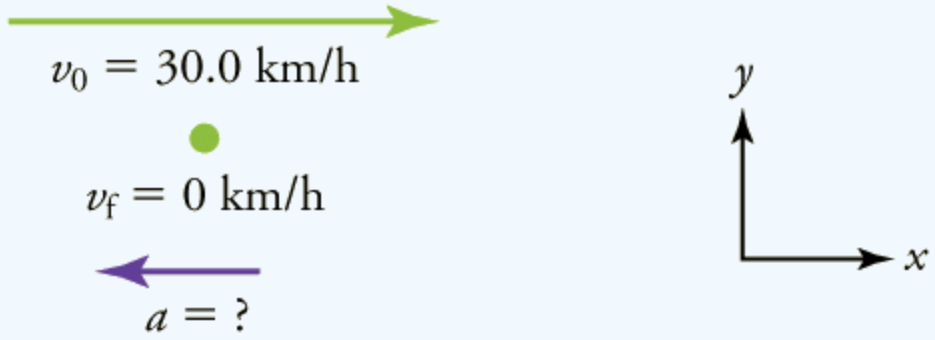



Figure 2.6.9:.

In this case, the train is decelerating and its acceleration is negative because it is toward the left. As in the previous example, we must find the change in velocity and the change in time and then solve for acceleration.

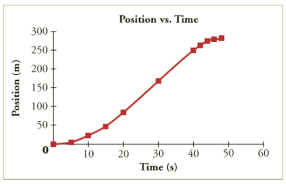
**Solution**

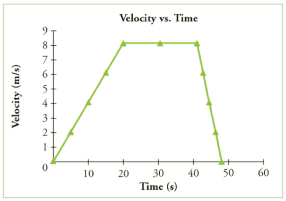
1. Identify the knowns. $v_0 = 30.0km/h, v_f = 0km/h$ (the train is stopped, so its velocity is 0), and $\Delta t = 8.00s$.
2. Solve for the change in velocity, $\Delta v$.
$$\Delta v = v_f - v_0 = 0 - 30.0km/h = -30.0km/h$$
3. Plug in the knowns, $\Delta v$ and $\Delta t$, and solve for $\bar{a}$.
$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{-30.0km/h}{8.00s}$$
4. Convert the units to meters and seconds.
$$\bar{a} = \frac{\Delta v}{\Delta t} = \left(\frac{-30.0km/h}{8.00s}\right)\left(\frac{10^3m}{1km}\right)\left(\frac{1h}{3600s}\right) = -1.04m/s^2.$$

**Discussion**

The minus sign indicates that acceleration is to the left. This sign is reasonable because the train initially has a positive velocity in this problem, and a negative acceleration would oppose the motion. Again, acceleration is in the same direction as the change in velocity, which is negative here. This acceleration can be called a deceleration because it has a direction opposite to the velocity.

The graphs of position, velocity, and acceleration vs. time for the trains in Example 2.6.4 and 2.6.5 are displayed in Figure 2.6.10. (We have taken the velocity to remain constant from 20 to 40 s, after which the train decelerates.)



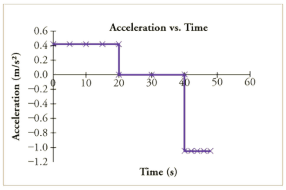




Figure 2.6.10: (a) Position of the train over time. Notice that the train's position changes slowly at the beginning of the journey, then more and more quickly as it picks up speed. Its position then changes more slowly as it slows down at the end of the journey. In the middle of the journey, while the velocity remains constant, the position changes at a constant rate. (b) Velocity of the train over time. The train's velocity increases as it accelerates at the beginning of the journey. It remains the same in the middle of the journey (where there is no acceleration). It decreases as the train decelerates at the end of the journey. (c) The acceleration of the train over time. The train has positive acceleration as it speeds up at the beginning of the journey. It has no acceleration as it travels at constant velocity in the middle of the journey. Its acceleration is negative as it slows down at the end of the journey.

## Example 2.6.6: Calculating Average Velocity: The Subway Train

What is the average velocity of the train in part b of Example 2.6.2, and shown again below, if it takes 5.00 min to make its trip?

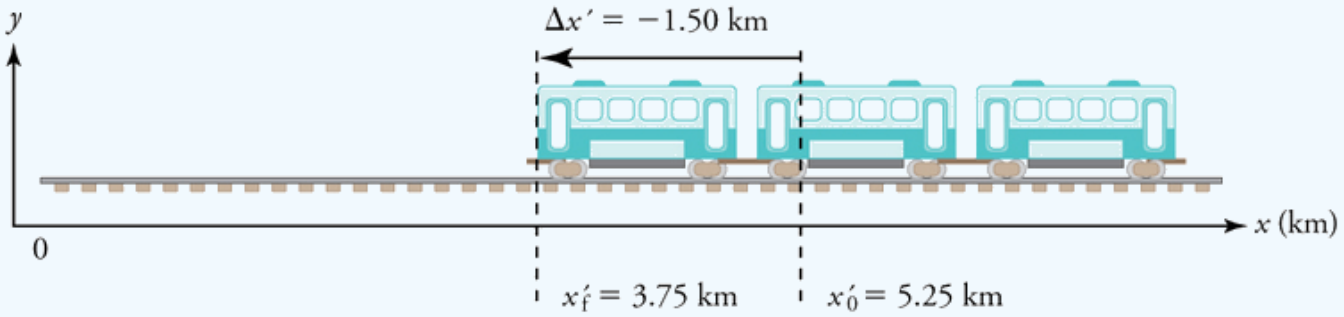



Figure 2.6.11

**Strategy**

Average velocity is displacement divided by time. It will be negative here, since the train moves to the left and has a negative displacement.

**Solution**

1. Identify the knowns.

$$x'_f = 3.75km,\ x'_0 = 5.25km,\ \Delta t = 5.00min.$$

2. Determine displacement, $\Delta x'$. We found $\Delta x'$ to be $-1.5km$ in Example 2.6.7.
3. Solve for average velocity.

$$\bar{v} = \frac{\Delta x'}{\Delta t} = \frac{-1.50km}{5.00min}$$

4. Convert units.

$$\bar{v} = \frac{\Delta x'}{\Delta t} = (\frac{-1.50km}{5.00min})(\frac{60min}{1}h) = -18.0km/h$$

**Discussion**

The negative velocity indicates motion to the left.

## Example 2.6.7: Calculating Deceleration: The Subway Train

Finally, suppose the train in Figure 2.6.7 slows to a stop from a velocity of 20.0 km/h in 10.0 s. What is its average acceleration?

**Strategy**

Once again, let's draw a sketch:

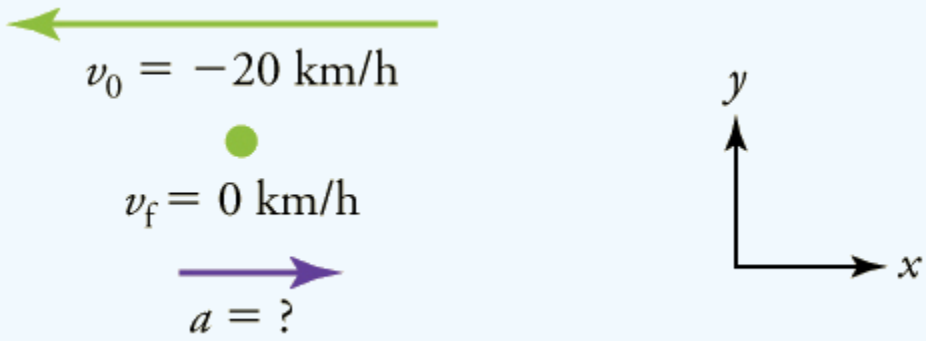



Figure 2.6.12: As before, we must find the change in velocity and the change in time to calculate average acceleration.

**Solution**

1. Identify the knowns. $v_0 = -20\ km/h$, $v_f = 0\ km/h$, $\Delta t = 10.0\ s$.
2. Calculate $\Delta v$. The change in velocity here is actually positive, since

$$\Delta v = v_f - v_0 = 0 - (-20\,km/h) = +20\,km/h.$$

3. Solve for $\bar{a}$.

$$\bar{a} = \frac{\Delta v}{\Delta t} = \frac{+20.0\,km/h}{10.0\,s}$$

4. Convert units.

$$\bar{a} = \left(\frac{+20.0\,km/h}{10.0\,s}\right)\left(\frac{10^3\,m}{1\,km}\right)\left(\frac{1\,h}{3600\,s}\right) = +0.556\,m/^s2$$

**Discussion**

The plus sign means that acceleration is to the right. This is reasonable because the train initially has a negative velocity (to the left) in this problem and a positive acceleration opposes the motion (and so it is to the right). Again, acceleration is in the same direction as the change in velocity, which is positive here. As in Example 2.6.5, this acceleration can be called a deceleration since it is in the direction opposite to the velocity.

## Sign and Direction

Perhaps the most important thing to note about these examples is the signs of the answers. In our chosen coordinate system, plus means the quantity is to the right and minus means it is to the left. This is easy to imagine for displacement and velocity. But it is a little less obvious for acceleration. Most people interpret negative acceleration as the slowing of an object. This was not the case in Example 2.6.5, where a positive acceleration slowed a negative velocity. The crucial distinction was that the acceleration was in the opposite direction from the velocity. In fact, a negative acceleration will increase a negative velocity. For example, the train moving to the left in Figure Figure 2.6.11 is sped up by an acceleration to the left. In that case, both v and a are negative. The plus and minus signs give the directions of the accelerations. If acceleration has the same sign as the velocity, the object is speeding up. If acceleration has the opposite sign as the velocity, the object is slowing down.

**Exercise 2.6.1**

An airplane lands on a runway traveling east. Describe its acceleration.

*Answer available at the end of this section.*

**PHET EXPLORATIONS: MOVING MAN SIMULATION**

Learn about position, velocity, and acceleration graphs with the PhET Moving Man simulation. Move the little man back and forth with the mouse and plot his motion. Set the position, velocity, or acceleration and let the simulation move the man for you.

## Summary

- Acceleration is the rate at which velocity changes. In symbols, **average acceleration** $\bar{a}$ is $\bar{a} = \frac{\Delta v}{\Delta t} = \frac{v_f - v_0}{t_f - t_0}$.
- The SI unit for acceleration is 2 .
- Acceleration is a vector, and thus has a both a magnitude and direction.
- Acceleration can be caused by either a change in the magnitude or the direction of the velocity.
- Instantaneous acceleration is the acceleration at a specific instant in time.
- Deceleration is an acceleration with a direction opposite to that of the velocity.

## Glossary

**acceleration**
: the rate of change in velocity; the change in velocity over time

**average acceleration**
: the change in velocity divided by the time over which it changes

**instantaneous acceleration**
: acceleration at a specific point in time

**deceleration**
: acceleration in the direction opposite to velocity; acceleration that results in a decrease in velocity

### Answers

**Exercise** 2.6.1

If we take east to be positive, then the airplane has negative acceleration, as it is accelerating toward the west. It is also decelerating: its acceleration is opposite in direction to its velocity.

This page titled 2.6: Acceleration is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 2.4: Acceleration by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.7: Falling Objects

**Learning Objectives**

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

- Describe the effects of gravity on objects in motion.
- Describe the motion of objects that are in free fall.
- Calculate the position and velocity of objects in free fall.

## Falling Objects

Falling objects form an interesting class of motion problems. For example, we can estimate the depth of a vertical mine shaft by dropping a rock into it and listening for the rock to hit the bottom. By applying the kinematics developed so far to falling objects, we can examine some interesting situations and learn much about gravity in the process.

## Gravity

The most remarkable and unexpected fact about falling objects is that, if air resistance and friction are negligible, then in a given location all objects fall toward the center of Earth with the *same constant acceleration*, *independent of their mass*. This experimentally determined fact is unexpected, because we are so accustomed to the effects of air resistance and friction that we expect light objects to fall slower than heavy ones.

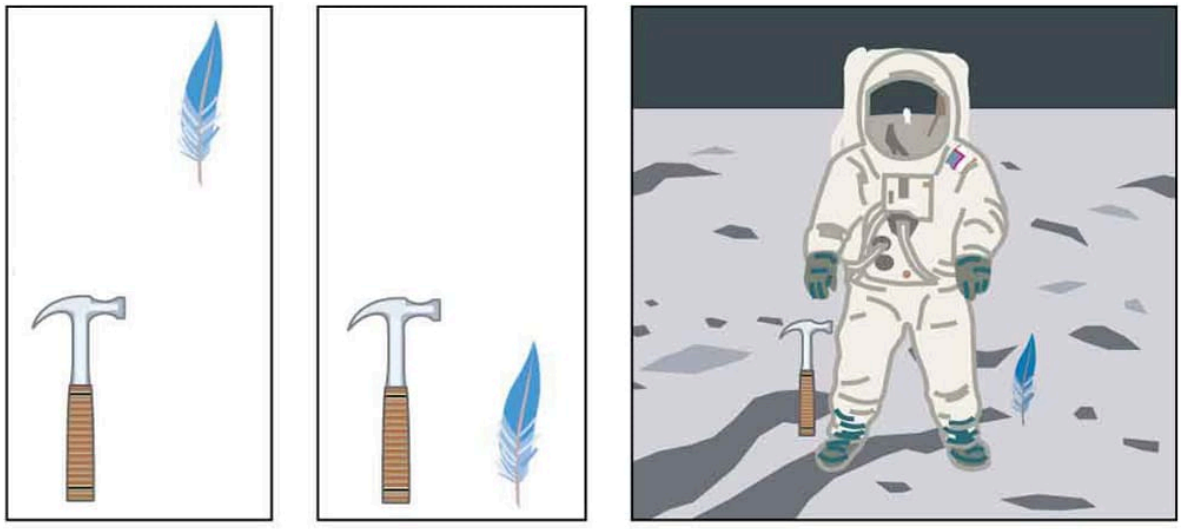



Figure 2.7.1: A hammer and a feather will fall with the same constant acceleration if air resistance is considered negligible. This is a general characteristic of gravity not unique to Earth, as astronaut David R. Scott demonstrated on the Moon in 1971, where the acceleration due to gravity is only .67 m/s2.

In the real world, air resistance can cause a lighter object to fall slower than a heavier object of the same size. A tennis ball will reach the ground after a hard baseball dropped at the same time. (It might be difficult to observe the difference if the height is not large.) Air resistance opposes the motion of an object through the air, while friction between objects—such as between clothes and a laundry chute or between a stone and a pool into which it is dropped—also opposes motion between them. For the ideal situations of these first few chapters, an object *falling without air resistance or friction* is defined to be in free-fall.

The force of gravity causes objects to fall toward the center of Earth. The acceleration of free-falling objects is therefore called

the acceleration due to gravity. The acceleration due to gravity is *constant*, which means we can apply the kinematics equations to any falling object where air resistance and friction are negligible. This opens a broad class of interesting situations to us. The acceleration due to gravity is so important that its magnitude is given its own symbol, . It is constant at any given location on Earth and has the average value

$$g = 9.80 m/s^2. \tag{2.7.1}$$

Although varies from .78 m/s2 to 9.83 m/s2, depending on latitude, altitude, underlying geological formations, and local topography, the average value of 9.80 m/s2 will be used in this text unless otherwise specified. The direction of the acceleration due to gravity is downward (towards the center of Earth). In fact, its direction defines what we call vertical. Note that whether the acceleration a in the kinematic equations has the value +g or –g depends on how we define our coordinate system. If we define the upward direction as negative, then a=–g=–9.80 m/s2, and if we define the downward direction as positive, then =g=9.80 m/s2.

## One-Dimensional Motion Involving Gravity

The best way to see the basic features of motion involving gravity is to start with the simplest situations and then progress toward more complex ones. So we start by considering straight up and down motion with no air resistance or friction. These assumptions mean that the velocity (if there is any) is vertical. If the object is dropped, we know the initial velocity is zero. Once the object has left contact with whatever held or threw it, the object is in free-fall. Under these circumstances, the motion is one-dimensional and has constant acceleration of magnitude g. We will also represent vertical displacement with the symbol y and use x for horizontal displacement.

### KINEMATIC EQUATIONS FOR OBJECTS IN FREE-FALL WHERE ACCELERATION = -G

$$v = v_0 - gt$$

$$y = y_0 + v_0 t - \tfrac{1}{2} g t^2$$

$$v^2 = v_0^2 - 2g(y - y_0)$$

### Example 2.7.1: Calculating Position and Velocity of a Falling Object: A Rock Thrown Upward

A person standing on the edge of a high cliff throws a rock straight up with an initial velocity of 13.0 m/s. The rock misses the edge of the cliff as it falls back to earth. Calculate the position and velocity of the rock 1.00 s, 2.00 s, and 3.00 s after it is thrown, neglecting the effects of air resistance.

**Strategy**

Draw a sketch.

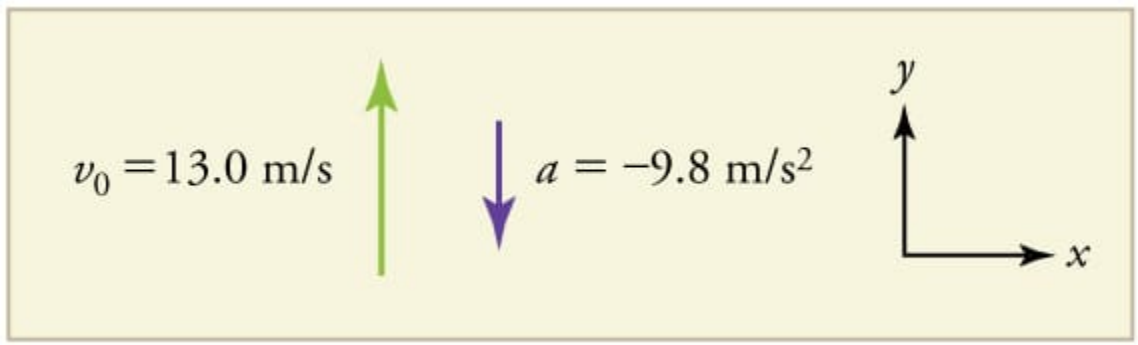



Figure 2.7.2

We are asked to determine the position $y$ at various times. It is reasonable to take the initial position $y_0$ to be zero. This problem involves one-dimensional motion in the vertical direction. We use plus and minus signs to indicate direction, with up being positive and down negative. Since up is positive, and the rock is thrown upward, the initial velocity must be positive too. The acceleration due to gravity is downward, so a is negative. It is crucial that the initial velocity and the acceleration due to gravity have opposite signs. Opposite signs indicate that the acceleration due to gravity opposes the initial motion and will slow and eventually reverse it.

Since we are asked for values of position and velocity at three times, we will refer to these as $y_1$ and $v_1$; $y_2$ and $v_2$; and $y_3$ and $v_3$.

**Solution for Position $y_1$**

1. Identify the knowns. We know that $y_0 = 0; v_0 = 13.0m/s; a = -g = -9.80m/s^2$; and $t = 1.00s$.

2. Identify the best equation to use. We will use $y = y_0 + v_0t + \frac{1}{2}at^2$ because it includes only one unknown, $y$ (or $y_1$, here), which is the value we want to find.

3. Plug in the known values and solve for $y_1$.

$$y_1 = 0 + (13.0m/s)(1.00s) + \tfrac{1}{2}(-9.80m/s^2)(1.00s)^2 = 8.10m$$

**Discussion**

The rock is 8.10 m above its starting point at $t = 1.00s$, since $y_1 > y_0$. It could be moving up or down; the only way to tell is to calculate $v_1$ and find out if it is positive or negative.

**Solution for Velocity $v_1$**

1. Identify the knowns. We know that $y_0 = 0; v_0 = 13.0m/s; a = -g = -9.80m/s^2$; and $t = 1.00s$. We also know from the solution above that $y_1 = 8.10m$.

2. Identify the best equation to use. The most straightforward is $v = v_0 - gt$ (from $v = v_0 + at$, where a=gravitational acceleration=–g).

3. Plug in the knowns and solve.

$$v_1 = v_0 - gt = 13.0m/s - (9.80m/s^2)(1.00s) = 3.20m/s$$

**Discussion**

The positive value for $v_1$ means that the rock is still heading upward at $t = 1.00s$. However, it has slowed from its original 13.0 m/s, as expected.

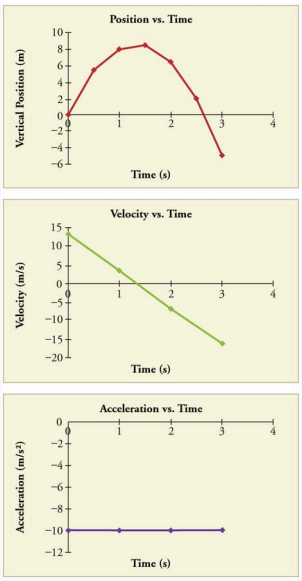
**Solution for Remaining Times**

The procedures for calculating the position and velocity at $t = 2.00s$ and $3.00s$ are the same as those above. The results are summarized in Table and illustrated in Figure.

***Results***

| Time, *t* | Position, *y* | Velocity, *v* | Acceleration, *a* |
|---|---|---|---|
| .00 s | .10 m | .20 m/s\) | 9.80 m/s^2\) |
| .00 s | .40 m | 6.60 m/s\) | 9.80 m/s^2\) |
| .00 s | 5.10 m | 16.4 m/s\) | 9.80 m/s^2\) |

Graphing the data helps us understand it more clearly.




*$PageIndex3$: Vertical position, vertical velocity, and vertical acceleration vs. time for a rock thrown vertically up at the edge of a cliff. Notice that velocity changes linearly with time and that acceleration is constant. Misconception Alert! Notice that the position vs. time graph shows vertical position only. It is easy to get the impression that the graph shows some horizontal motion—the shape of the graph looks like the path of a projectile. But this is not the case; the horizontal axis is time, not space. The actual path of the rock in space is straight up, and straight down.*

**Discussion**

The interpretation of these results is important. At 1.00 s the rock is above its starting point and heading upward, since $y_1$ and $v_1$ are both positive. At 2.00 s, the rock is still above its starting point, but the negative velocity means it is moving downward. At 3.00 s, both $y_3$ and $v_3$ are negative, meaning the rock is below its starting point and continuing to move downward. Notice that when the rock is at its highest point (at 1.5 s), its velocity is zero, but its acceleration is still $-9.80 m/s^2$. Its acceleration is $-9.80 m/s^2$ for the whole trip—while it is moving up and while it is moving down. Note that the values for y are the positions (or displacements) of the rock, not the total distances traveled. Finally, note that free-fall applies to upward motion as well as downward. Both have the same acceleration—the acceleration due to gravity, which remains constant the entire time. Astronauts training in the famous Vomit Comet, for example, experience free-fall while arcing up as well as down, as we will discuss in more detail later.The interpretation of these results is important. At 1.00 s the rock is above its starting point and heading upward, since y1 and v1 are both positive. At 2.00 s, the rock is still above its starting point, but the negative velocity means it is moving downward. At 3.00 s, both y3 and v3 are negative, meaning the rock is below its starting point and continuing to move downward. Notice that when the rock is at its highest point (at 1.5 s), its velocity is zero, but its acceleration is still –9.80 m/s2. Its acceleration is –9.80 m/s2 for the whole trip—while it is moving up and while it is moving down. Note that the values for y are the positions (or displacements) of the rock, not the total distances traveled. Finally, note that free-fall applies to upward motion as well as downward. Both have the same acceleration—the acceleration due to gravity, which remains constant the entire time. Astronauts training in the famous Vomit Comet, for example, experience free-fall while arcing up as well as down, as we will discuss in more detail later.

## MAKING CONNECTIONS: TAKE-HOME EXPERIMENT—REACTION TIME

A simple experiment can be done to determine your reaction time. Have a friend hold a ruler between your thumb and index finger, separated by about 1 cm. Note the mark on the ruler that is right between your fingers. Have your friend drop the ruler unexpectedly, and try to catch it between your two fingers. Note the new reading on the ruler. Assuming acceleration is that due to gravity, calculate your reaction time. How far would you travel in a car (moving at 30 m/s) if the time it took your foot to go from the gas pedal to the brake was twice this reaction time?

## Example 2.7.2:Calculating Velocity of a Falling Object: A Rock Thrown Down

What happens if the person on the cliff throws the rock straight down, instead of straight up? To explore this question, calculate the velocity of the rock when it is 5.10 m below the starting point, and has been thrown downward with an initial speed of 13.0 m/s.

**Strategy**

Draw a sketch.

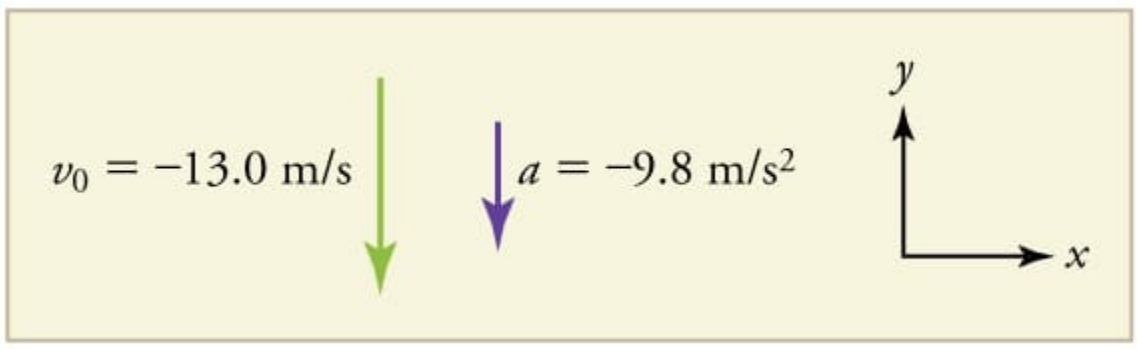



Figure 2.7.4

Since up is positive, the final position of the rock will be negative because it finishes below the starting point at $y_0 = 0$. Similarly, the initial velocity is downward and therefore negative, as is the acceleration due to gravity. We expect the final velocity to be negative since the rock will continue to move downward.

**Solution**

1. Identify the knowns. $y_0 = 0; y_1 = -5.10m; v_0 = -13.0m/s; a = -g = -9.80m/s^2$.

2. Choose the kinematic equation that makes it easiest to solve the problem. The equation $v^2 = v_0^2 + 2a(y - y_0)$ works well because the only unknown in it is $v$. (We will plug $y_1$ in for $y$.)

3. Enter the known values

$$v^2 = (-13.0m/s)^2 + 2(-9.80m/s^2)(-5.10m - 0m) = 268.96m^2/s^2,$$

where we have retained extra significant figures because this is an intermediate result.

Taking the square root, and noting that a square root can be positive or negative, gives

$$v = \pm 16.4m/s.$$

The negative root is chosen to indicate that the rock is still heading down. Thus,

$$v = -16.4m/s.$$

**Discussion**

Note that *this is exactly the same velocity the rock had at this position when it was thrown straight upward with the same initial speed.* (See Example and Figure(a).) This is not a coincidental result. Because we only consider the acceleration due to gravity in this problem, the speed of a falling object depends only on its initial speed and its vertical position relative to the starting point. For example, if the velocity of the rock is calculated at a height of 8.10 m above the starting point (using the method from Example) when the initial velocity is 13.0 m/s straight up, a result of $\pm 3.20m/s$ is obtained. Here both signs are meaningful; the positive value occurs when the rock is at 8.10 m and heading up, and the negative value occurs when the rock is at 8.10 m and heading back down. It has the same speed but the opposite direction.

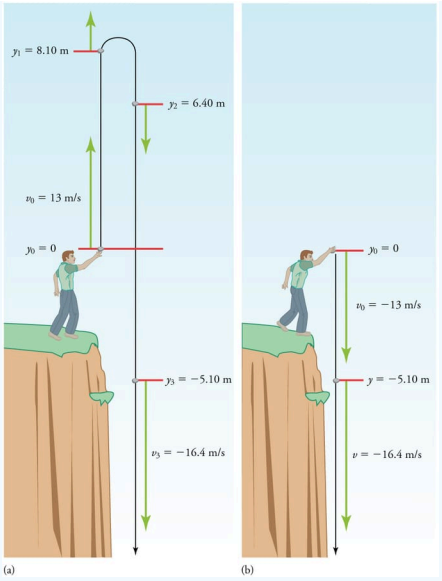



Figure $PageIndex5$: (a) A person throws a rock straight up, as explored in Example. The arrows are velocity vectors at 0, 1.00, 2.00, and 3.00 s. (b) A person throws a rock straight down from a cliff with the same initial speed as before, as in Example. Note that at the same distance below the point of release, the rock has the same velocity in both cases.

Another way to look at it is this: In Example, the rock is thrown up with an initial velocity of $13.0m/s$. It rises and then falls back down. When its position is $y = 0$ on its way back down, its velocity is $-13.0m/s$. That is, it has the same speed on its way down as on its way up. We would then expect its velocity at a position of $y = -5.10$ m to be the same whether we have thrown it upwards at $+13.0m/s$ or thrown it downwards at $-13.0m/s$. The velocity of the rock on its way down from $y = 0$ is the same whether we have thrown it up or down to start with, as long as the speed with which it was initially thrown is the same.

## Example 2.7.3:Find g from Data on a Falling Object

The acceleration due to gravity on Earth differs slightly from place to place, depending on topography (e.g., whether you are on a hill or in a valley) and subsurface geology (whether there is dense rock like iron ore as opposed to light rock like salt beneath you.) The precise acceleration due to gravity can be calculated from data taken in an introductory physics laboratory course. An object, usually a metal ball for which air resistance is negligible, is dropped and the time it takes to fall a known distance is measured. See, for example, Figure. Very precise results can be produced with this method if sufficient care is taken in measuring the distance fallen and the elapsed time.

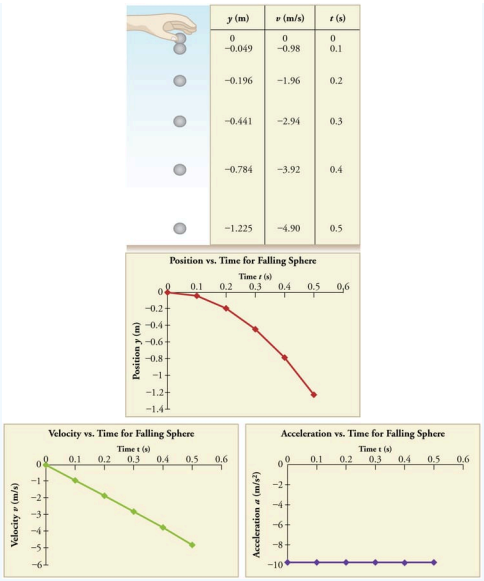



Figure 2.7.6: Positions and velocities of a metal ball released from rest when air resistance is negligible. Velocity is seen to increase linearly with time while displacement increases with time squared. Acceleration is a constant and is equal to gravitational acceleration.

Suppose the ball falls 1.0000 m in 0.45173 s. Assuming the ball is not affected by air resistance, what is the precise acceleration due to gravity at this location?

**Strategy**

Draw a sketch.

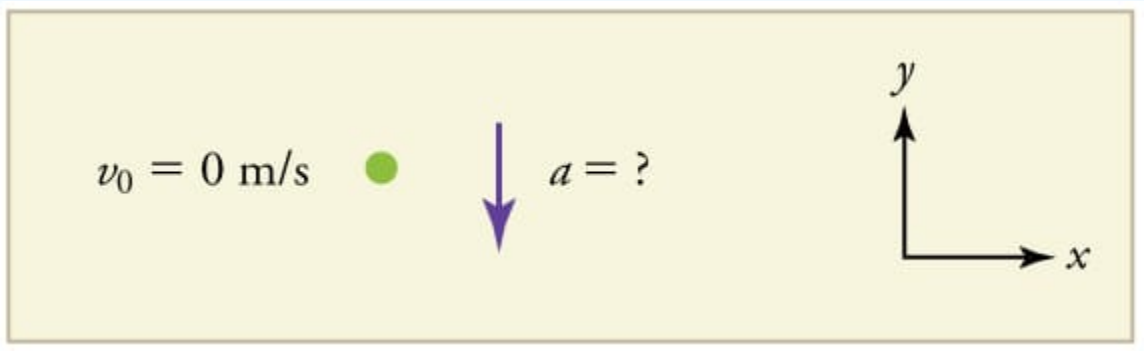



Figure 2.7.7:

We need to solve for acceleration $a$. Note that in this case, displacement is downward and therefore negative, as is acceleration.

**Solution**

1. Identify the knowns. $y_0 = 0; y = -1.0000m; t = 0.45173; v_0 = 0.$

2. Choose the equation that allows you to solve for $a$ using the known values.

$$y = y_0 + v_0 t + \tfrac{1}{2}at^2$$

3. Substitute 0 for $v_0$ and rearrange the equation to solve for $a$. Substituting 0 for $v_0$yields

$$y = y_0 + \tfrac{1}{2}at^2.$$

Solving for $a$ gives

$$a = \frac{2(y-y_0)}{t^2}.$$

4. Substitute known values yields

$$a = \frac{2(-1.0000m-0)}{(0.45173s)^2} = -9.8010m/s^2,$$

so, because $a = -g$ with the directions we have chosen,

$$g = 9.8010m/s^2.$$

**Discussion**

The negative value for a indicates that the gravitational acceleration is downward, as expected. We expect the value to be somewhere around the average value of $9.80m/s^2$, so $9.8010m/s^2$ makes sense. Since the data going into the calculation are relatively precise, this value for g is more precise than the average value of $9.80m/s$; it represents the local value for the acceleration due to gravity.

## Exercise 2.7.1

A chunk of ice breaks off a glacier and falls 30.0 meters before it hits the water. Assuming it falls freely (there is no air resistance), how long does it take to hit the water?

*Answer available at the end of this section.*

## PHET EXPLORATIONS: EQUATION GRAPHER

Learn about graphing polynomials. The shape of the curve changes as the constants are adjusted. View the curves for the individual terms (e.g. $y = bx$) to see how they add to generate the polynomial curve.




Figure 2.7.8: Equation Grapher

# Summary

- An object in free-fall experiences constant acceleration if air resistance is negligible.
- On Earth, all free-falling objects have an acceleration due to gravity $g$, which averages

$$g = 9.80m/s^2.$$

- Whether the acceleration a should be taken as $+g$ or $-g$ is determined by your choice of coordinate system. If you choose the upward direction as positive, $a = -g = -9.80m/s^2$ is negative. In the opposite case, $a = +g = 9.80m/s^2$ is positive. Since acceleration is constant, the kinematic equations above can be applied with the appropriate $+g$ or $-g$ substituted for a.
- For objects in free-fall, up is normally taken as positive for displacement, velocity, and acceleration.

# Glossary

**free-fall**
the state of movement that results from gravitational force only

**acceleration due to gravity**
acceleration of an object as a result of gravity

## Answers

**Exercise** 2.7.1

We know that initial position $y_0 = 0$, final position $y = -30.0m$, and $a = -g = -9.80m/s^2$. We can then use the equation $y = y_0 + v_0t + \frac{1}{2}at^2$ to solve for $t$. Inserting $a = -g$, we obtain

$y = 0 + 0 - \frac{1}{2}gt^2$

$t^2 = \frac{2y}{-g}$

$t = \pm\sqrt{\frac{2y}{-g}} = \pm\sqrt{\frac{2(-30.0m)}{-9.80m/s^2}} = \pm\sqrt{6.12s^2} = 2.47s \approx 2.5s$

where we take the positive value as the physically relevant answer. Thus, it takes about 2.5 seconds for the piece of ice to hit the water.

This page titled 2.7: Falling Objects is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 2.7: Falling Objects by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.8: Prelude to Two-Dimensional Kinematics

The arc of a basketball, the orbit of a satellite, a bicycle rounding a curve, a swimmer diving into a pool, blood gushing out of a wound, and a puppy chasing its tail are but a few examples of motions along curved paths. In fact, most motions in nature follow curved paths rather than straight lines. Motion along a curved path on a flat surface or a plane (such as that of a ball on a pool table or a skater on an ice rink) is two-dimensional, and thus described by two-dimensional kinematics. Motion not confined to a plane, such as a car following a winding mountain road, is described by three-dimensional kinematics. Both two- and three-dimensional kinematics are simple extensions of the one-dimensional kinematics developed for straight-line motion in the previous chapter. This simple extension will allow us to apply physics to many more situations, and it will also yield unexpected insights about nature.



Figure 2.8.1: Everyday motion that we experience is, thankfully, rarely as tortuous as a rollercoaster ride like this—the Dragon Khan in Spain's Universal Port Aventura Amusement Park. However, most motion is in curved, rather than straight-line, paths. Motion along a curved path is two- or three-dimensional motion, and can be described in a similar fashion to one-dimensional motion. (credit: Boris23/Wikimedia Commons)

This page titled 2.8: Prelude to Two-Dimensional Kinematics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 3.0: Prelude to Two-Dimensional Kinematics by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.9: Kinematics in Two Dimensions - An Introduction

**Learning Objectives**

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

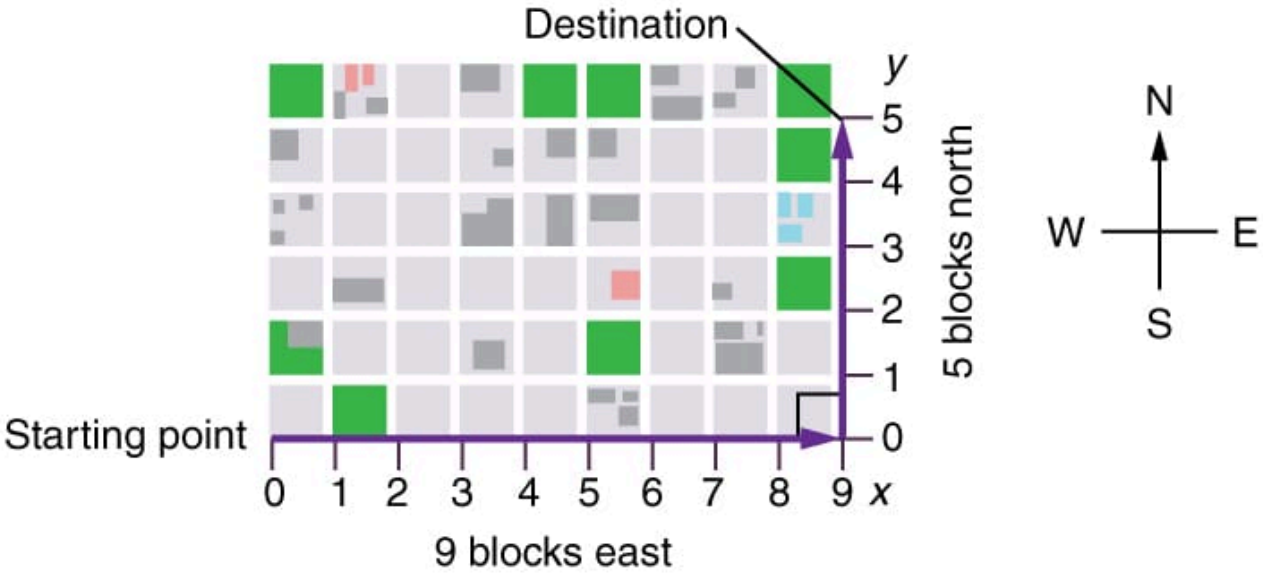
- Observe that motion in two dimensions consists of horizontal and vertical components.
- Understand the independence of horizontal and vertical vectors in two-dimensional motion.



Figure 2.9.1: Walkers and drivers in a city like New York are rarely able to travel in straight lines to reach their destinations. Instead, they must follow roads and sidewalks, making two-dimensional, zigzagged paths. (credit: Margaret W. Carruthers)

## Two-Dimensional Motion: Walking in a City

Suppose you want to walk from one point to another in a city with uniform square blocks, as pictured in Figure 2.9.2.




Figure

2.9.2: A pedestrian walks a two-dimensional path between two points in a city. In this scene, all blocks are square and are the same size.

The straight-line path that a helicopter might fly is blocked to you as a pedestrian, and so you are forced to take a two-dimensional path, such as the one shown. You walk 14 blocks in all, 9 east followed by 5 north. What is the straight-line distance?

An old adage states that the shortest distance between two points is a straight line. The two legs of the trip and the straight-line path form a right triangle, and so the Pythagorean theorem,

$$a^2 + b^2 = c^2 \tag{2.9.1}$$

can be used to find the straight-line distance.

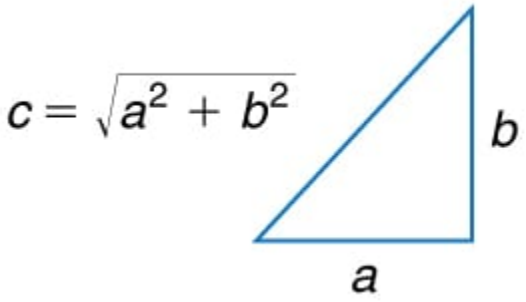



Figure 2.9.3: The Pythagorean theorem relates the length of the legs of a right triangle, labeled $a$ and $b$, with the hypotenuse, labeled $c$. The relationship is given by: $a^2 + b^2 = c^2$. This can be rewritten, solving for $c$ : $c = \sqrt{a^2 + b^2}$.

The hypotenuse of the triangle is the straight-line path, and so in this case its length in units of city blocks is $\sqrt{(9blocks)^2 + (5blocks)^2} = 10.3blocks$, considerably shorter than the 14 blocks you walked. (Note that we are using three significant figures in the answer. Although it appears that "9" and "5" have only one significant digit, they are discrete numbers. In this case "9 blocks" is the same as "9.0 or 9.00 blocks." We have decided to use three significant figures in the answer in order to show the result more precisely.)

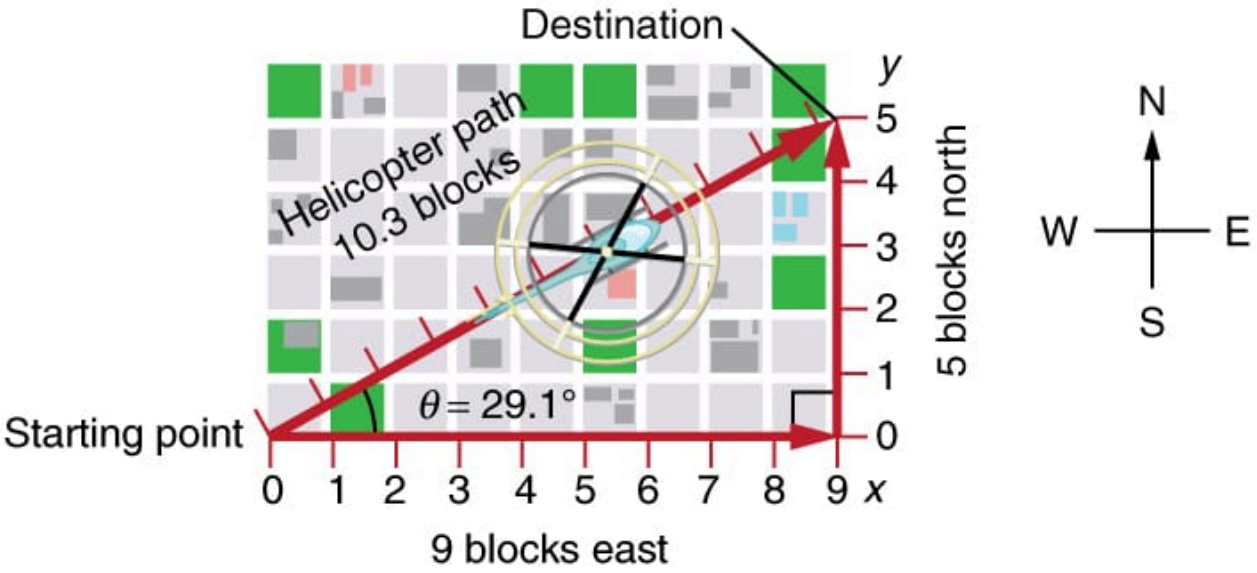



Figure 2.9.4 : The straight-line path followed by a helicopter between the two points is shorter than the 14 blocks walked by the pedestrian. All blocks are square and the same size.

The fact that the straight-line distance (10.3 blocks) in Figure is less than the total distance walked (14 blocks) is one example of a general characteristic of vectors. (Recall that vectors are quantities that have both magnitude and direction.)

As for one-dimensional kinematics, we use arrows to represent vectors. The length of the arrow is proportional to the vector's magnitude. The arrow's length is indicated by hash marks in Figure and Figure. The arrow points in the same direction as the

vector. For two-dimensional motion, the path of an object can be represented with three vectors: one vector shows the straight-line path between the initial and final points of the motion, one vector shows the horizontal component of the motion, and one vector shows the vertical component of the motion. The horizontal and vertical components of the motion add together to give the straight-line path. For example, observe the three vectors in Figure. The first represents a 9-block displacement east. The second represents a 5-block displacement north. These vectors are added to give the third vector, with a 10.3-block total displacement. The third vector is the straight-line path between the two points. Note that in this example, the vectors that we are adding are perpendicular to each other and thus form a right triangle. This means that we can use the Pythagorean theorem to calculate the magnitude of the total displacement. (Note that we cannot use the Pythagorean theorem to add vectors that are not perpendicular. We will develop techniques for adding vectors having any direction, not just those perpendicular to one another, in Vector Addition and Subtraction: Graphical Methods and Vector Addition and Subtraction: Analytical Methods.)

## The Independence of Perpendicular Motions

The person taking the path shown in Figure walks east and then north (two perpendicular directions). How far he or she walks east is only affected by his or her motion eastward. Similarly, how far he or she walks north is only affected by his or her motion northward.

**INDEPENDENCE OF MOTION**

The horizontal and vertical components of two-dimensional motion are independent of each other. Any motion in the horizontal direction does not affect motion in the vertical direction, and vice versa.

This is true in a simple scenario like that of walking in one direction first, followed by another. It is also true of more complicated motion involving movement in two directions at once. For example, let's compare the motions of two baseballs. One baseball is dropped from rest. At the same instant, another is thrown horizontally from the same height and follows a curved path. A stroboscope has captured the positions of the balls at fixed time intervals as they fall.

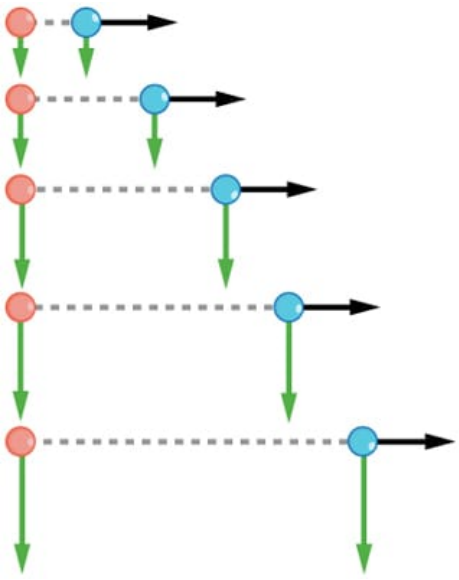


Figure 2.9.5: This shows the motions of two identical balls—one falls from rest, the other has an initial horizontal velocity. Each subsequent position is an equal time interval. Arrows represent horizontal and vertical velocities at each position. The ball on the right has an initial horizontal velocity, while the ball on the left has no horizontal velocity. Despite the difference in horizontal velocities, the vertical velocities and positions are identical for both balls. This shows that the vertical and horizontal motions are independent.

It is remarkable that for each flash of the strobe, the vertical positions of the two balls are the same. This similarity implies that the vertical motion is independent of whether or not the ball is moving horizontally. (Assuming no air resistance, the vertical motion of a falling object is influenced by gravity only, and not by any horizontal forces.) Careful examination of the ball thrown horizontally shows that it travels the same horizontal distance between flashes. This is due to the fact that there are no additional forces on the ball in the horizontal direction after it is thrown. This result means that the horizontal velocity is constant, and

affected neither by vertical motion nor by gravity (which is vertical). Note that this case is true only for ideal conditions. In the real world, air resistance will affect the speed of the balls in both directions.

The two-dimensional curved path of the horizontally thrown ball is composed of two independent one-dimensional motions (horizontal and vertical). The key to analyzing such motion, called *projectile motion*, is to *resolve* (break) it into motions along perpendicular directions. Resolving two-dimensional motion into perpendicular components is possible because the components are independent. We shall see how to resolve vectors in Vector Addition and Subtraction: Graphical Methods and Vector Addition and Subtraction: Analytical Methods. We will find such techniques to be useful in many areas of physics.

PHET EXPLORATIONS: LADYBUG MOTION 2D

Learn about position, velocity and acceleration vectors. Move the ladybug by setting the position, velocity or acceleration, and see how the vectors change. Choose linear, circular or elliptical motion, and record and playback the motion to analyze the behavior.




Figure 2.9.5: Ladybug Motion 2D

## Summary

- The shortest path between any two points is a straight line. In two dimensions, this path can be represented by a vector with horizontal and vertical components.
- The horizontal and vertical components of a vector are independent of one another. Motion in the horizontal direction does not affect motion in the vertical direction, and vice versa.

## Glossary

**vector**
: a quantity that has both magnitude and direction; an arrow used to represent quantities with both magnitude and direction

This page titled 2.9: Kinematics in Two Dimensions - An Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 3.1: Kinematics in Two Dimensions - An Introduction by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.10: Projectile Motion

## Learning Objectives

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

- Identify and explain the properties of a projectile, such as acceleration due to gravity, range, maximum height, and trajectory.
- Determine the location and velocity of a projectile at different points in its trajectory.
- Apply the principle of independence of motion to solve projectile motion problems.

**Projectile motion** is the **motion** of an object thrown or projected into the air, subject to only the acceleration of gravity. The object is called a **projectile**, and its path is called its **trajectory.** The motion of falling objects, as covered in Problem-Solving Basics for One-Dimensional Kinematics, is a simple one-dimensional type of projectile motion in which there is no horizontal movement. In this section, we consider two-dimensional projectile motion, such as that of a football or other object for which **air resistance** *is negligible*.

The most important fact to remember here is that motions along perpendicular axes are independent and thus can be analyzed separately. This fact was discussed in Kinematics in Two Dimensions: An Introduction, where vertical and horizontal motions were seen to be independent. The key to analyzing two-dimensional projectile motion is to break it into two motions, one along the horizontal axis and the other along the vertical. (This choice of axes is the most sensible, because acceleration due to gravity is vertical—thus, there will be no acceleration along the horizontal axis when air resistance is negligible.) As is customary, we call the horizontal axis the *x*-axis and the vertical axis the *y*-axis. Figure illustrates the notation for displacement, where $s$ is defined to be the total displacement and $x$ and $y$ are its components along the horizontal and vertical axes, respectively. The magnitudes of these vectors are s, x, and y. (Note that in the last section we used the notation $A$ to represent a vector with components $A_x$ and $A_y$. If we continued this format, we would call displacement $s$ with components $s_x$ and $s_y$. However, to simplify the notation, we will simply represent the component vectors as $x$ and $y$.)

Of course, to describe motion we must deal with velocity and acceleration, as well as with displacement. We must find their components along the x- and *y*-axes, too. We will assume all forces except gravity (such as air resistance and friction, for example) are negligible. The components of acceleration are then very simple: $a_y = -g = -9.80 m/s^2$(. (Note that this definition assumes that the upwards direction is defined as the positive direction. If you arrange the coordinate system instead such that the downwards direction is positive, then acceleration due to gravity takes a positive value.) Because gravity is vertical, $a_x = 0$. Both accelerations are constant, so the kinematic equations can be used.

### REVIEW OF KINEMATIC EQUATIONS (CONSTANT )

$$x = x_0 + \bar{v}t$$
$$\bar{v} = \frac{v_0 + v}{2}$$
$$v = v_0 + at$$
$$x = x_0 + v_0 t + \frac{1}{2}at^2$$
$$v^2 = v_0^2 + 2a(x - x_0).$$

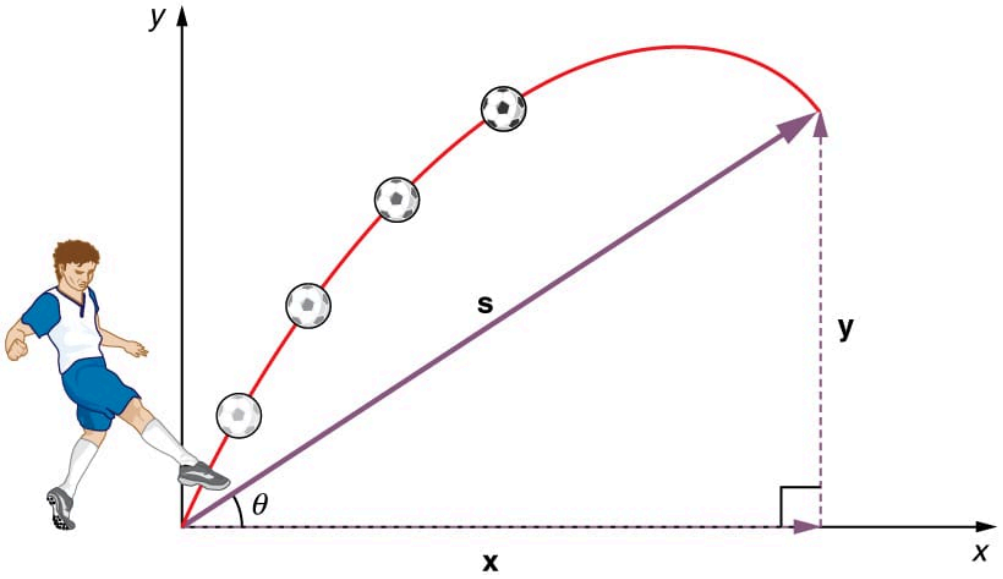



Figure 2.10.1: The total displacement $s$ of a soccer ball at a point along its path. The vector $s$ has components $x$ and $y$ along the horizontal and vertical axes. Its magnitude is $s$, and it makes an angle $\theta$ with the horizontal.

Given these assumptions, the following steps are then used to analyze projectile motion:

**Step 1**. *Resolve or break the motion into horizontal and vertical components along the x- and y-axes.* These axes are perpendicular, so $A_x = Acos\theta$ and $A_y = Asin\theta$ are used. The magnitude of the components of displacement $s$ along these axes are $x$ and $y$. The magnitudes of the components of the velocity $v$ are $v_x = vcos\theta$ and $v_y = vsin\theta$, where $v$ is the magnitude of the velocity and θ is its direction, as shown in Figure. Initial values are denoted with a subscript 0, as usual.

**Step 2.** *Treat the motion as two independent one-dimensional motions, one horizontal and the other vertical.* The kinematic equations for horizontal and vertical motion take the following forms:

$$\text{Horizontal Motion}(a_x = 0)$$
$$x = x_0 + v_x t$$
$$v_x = v_{0x} = v_x = \text{velocity is a constant.}$$
$$\text{Vertical Motion}(\text{assuming positive is up} a_y = -g = -9.80 m/s^2)$$
$$y = y_0 + \frac{1}{2}(v_{0y} + v_y)t$$
$$v_y = v_{0y} - gt$$
$$y = y_0 + v_{0y}t - \frac{1}{2}gt^2$$
$$v_y^2 = v_{0y}^2 - 2g(y - y_0).$$

**Step 3.** *Solve for the unknowns in the two separate motions—one horizontal and one vertical.* Note that the only common variable between the motions is time *t*. The problem solving procedures here are the same as for one-dimensional **kinematics** and are illustrated in the solved examples below.

**Step 4.** *Recombine the two motions to find the total displacement $s$ and velocity $v$.* Because the *x* - and *y* -motions are perpendicular, we determine these vectors by using the techniques outlined in the Vector Addition and Subtraction: Analytical Methods and employing $A = \sqrt{A_x^2 + A_y^2}$ and $\theta = tan^{-1}(A_y/A_x)$ in the following form, where $\theta$ is the direction of the displacement $s$ and $\theta_v$ is the direction of the velocity $v$:

**Total displacement and velocity**

$$s = \sqrt{x^2 + y^2}$$
$$\theta = tan^{-1}(y/x)$$
$$v = \sqrt{v_x^2 + v_y^2}$$
$$\theta_v = tan^{-1}(v_y/v_x).$$

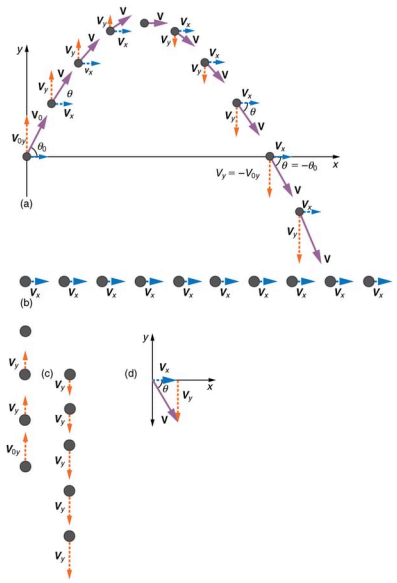


Figure 2.10.2: (a) We analyze two-dimensional projectile motion by breaking it into two independent one-dimensional motions along the vertical and horizontal axes. (b) The horizontal motion is simple, because $a_x = 0$ and $v_x$ is thus constant. (c) The velocity in the vertical direction begins to decrease as the object rises; at its highest point, the vertical velocity is zero. As the object falls towards the Earth again, the vertical velocity increases again in magnitude but points in the opposite direction to the initial vertical velocity. (d) The x - and y -motions are recombined to give the total velocity at any given point on the trajectory.

## Example 2.10.1: A Fireworks Projectile Explodes High and Away

During a fireworks display, a shell is shot into the air with an initial speed of 70.0 m/s at an angle of **75.0°** above the horizontal, as illustrated in Figure. The fuse is timed to ignite the shell just as it reaches its highest point above the ground.

a. Calculate the height at which the shell explodes.
b. How much time passed between the launch of the shell and the explosion?
c. What is the horizontal displacement of the shell when it explodes?

**Strategy**

Because air resistance is negligible for the unexploded shell, the analysis method outlined above can be used. The motion can be broken into horizontal and vertical motions in which $a_x = 0$ and $a_y = -g$. We can then define $x_0$ and $y_0$ to be zero and solve for the desired quantities.

**Solution for (a)**

By “height” we mean the altitude or vertical position y above the starting point. The highest point in any trajectory, called the apex, is reached when $v_y = 0$. Since we know the initial and final velocities as well as the initial position, we use the following equation to find $y$:

$$v_y^2 = v_{0y}^2 - 2g(y - y_0).$$

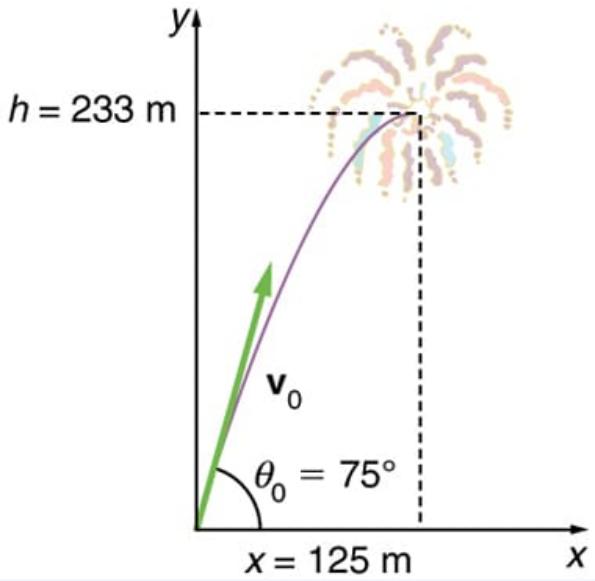



Figure 2.10.3: The trajectory of a fireworks shell. The fuse is set to explode the shell at the highest point in its trajectory, which is found to be at a height of 233 m and 125 m away horizontally.

Because $y_0$ and $v_y$ are both zero, the equation simplifies to

$$0 = v_{0y}^2 - 2gy.$$

Solving for $y$ gives

$$y = \frac{v_{0y}^2}{2g}.$$

Now we must find $v_{0y}$, the component of the initial velocity in the *y*-direction. It is given by $v_{0y} = v_0 sin\theta$, where $v_{0y}$ is the initial velocity of 70.0 m/s, and $\theta_0 = 75.0\,^{\circ}$ is the initial angle. Thus,

$$v_{0y} = v_0 sin\theta_0 = (70.0 m/s)(sin 75\,^{\circ}) = 67.6 m/s.$$

and $y$ is

$$y = \frac{(67.6 m/s)^2}{2(9.80 m/s^2)},$$

so that

$$y = 233 m.$$

**Discussion for (a)**

Note that because up is positive, the initial velocity is positive, as is the maximum height, but the acceleration due to gravity is negative. Note also that the maximum height depends only on the vertical component of the initial velocity, so that any projectile with a 67.6 m/s initial vertical component of velocity will reach a maximum height of 233 m (neglecting air resistance). The numbers in this example are reasonable for large fireworks displays, the shells of which do reach such heights before exploding. In practice, air resistance is not completely negligible, and so the initial velocity would have to be somewhat larger than that given to reach the same height.

**Solution for (b)**

As in many physics problems, there is more than one way to solve for the time to the highest point. In this case, the easiest method is to use $y = y_0 + \frac{1}{2}(v_{0y} + v_y)t$. Because y0 is zero, this equation reduces to simply

$$y = \frac{1}{2}(v_{0y} + v_y)t.$$

Note that the final vertical velocity, $v_y$, at the highest point is zero. Thus,

$$t = \frac{2y}{(v_{0y} + v_y)} = \frac{2(233m)}{(67.6m/s)} = 6.90s.$$

**Discussion for (b)**

This time is also reasonable for large fireworks. When you are able to see the launch of fireworks, you will notice several seconds pass before the shell explodes. (Another way of finding the time is by using $y = y_0 + v_{0y}t - \frac{1}{2}gt^2$, and solving the quadratic equation for $t$.)

**Solution for (c)**

Because air resistance is negligible, $a_x = 0$ and the horizontal velocity is constant, as discussed above. The horizontal displacement is horizontal velocity multiplied by time as given by $x = x_0 + v_x t$, where $x_0$ is equal to zero:

$$x = v_x t,$$

where $v_x$ is the *x*-component of the velocity, which is given by $v_x = v_0 cos\theta_0$. Now,

$$v_x = v_0 cos\theta_0 = (70.0m/s)(cos75.0\,^{o}) = 18.1m/s.$$

The time $t$ for both motions is the same, and so $x$ is

$$x = (18.1m/s)(6.90s) = 125m.$$

**Discussion for (c)**

The horizontal motion is a constant velocity in the absence of air resistance. The horizontal displacement found here could be useful in keeping the fireworks fragments from falling on spectators. Once the shell explodes, air resistance has a major effect, and many fragments will land directly below.

In solving part (a) of the preceding example, the expression we found for y is valid for any projectile motion where air resistance is negligible. Call the maximum height $y = h$; then,

$$h = \frac{v_{0y}^2}{2g}.$$

This equation defines the *maximum height* of a projectile and depends only on the vertical component of the initial velocity.

## DEFINING A COORDINATE SYSTEM

It is important to set up a coordinate system when analyzing projectile motion. One part of defining the coordinate system is to define an origin for the $x$ and $y$ positions. Often, it is convenient to choose the initial position of the object as the origin such that $x_0 = 0$ and $y_0 = 0$. It is also important to define the positive and negative directions in the $x$ and $y$ directions. Typically, we define the positive vertical direction as upwards, and the positive horizontal direction is usually the direction of the object's motion. When this is the case, the vertical acceleration, $g$, takes a negative value (since it is directed downwards towards the Earth). However, it is occasionally useful to define the coordinates differently. For example, if you are analyzing the motion of a ball thrown downwards from the top of a cliff, it may make sense to define the positive direction downwards since the motion of the ball is solely in the downwards direction. If this is the case, g takes a positive value.

### Example 2.10.2: Calculating Projectile Motion: Hot Rock Projectile

Kilauea in Hawaii is the world's most continuously active volcano. Very active volcanoes characteristically eject red-hot rocks and lava rather than smoke and ash. Suppose a large rock is ejected from the volcano with a speed of 25.0 m/s and at an angle above the horizontal, as shown in Figure. The rock strikes the side of the volcano at an altitude 20.0 m lower than its starting point. (a) Calculate the time it takes the rock to follow this path. (b) What are the magnitude and direction of the rock's velocity at impact?

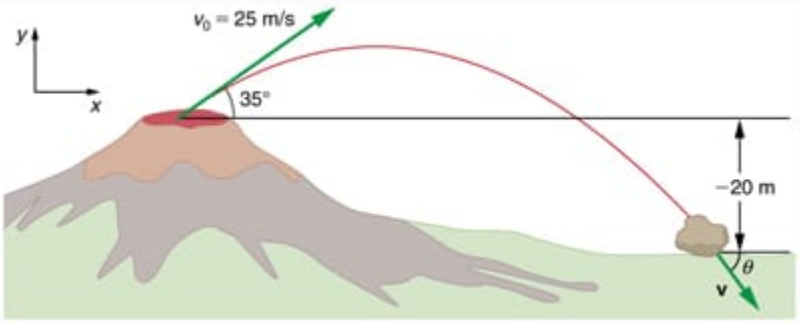



Figure 2.10.4: The trajectory of a rock ejected from the Kilauea volcano.

**Strategy**

Again, resolving this two-dimensional motion into two independent one-dimensional motions will allow us to solve for the desired quantities. The time a projectile is in the air is governed by its vertical motion alone. We will solve for $t$ first. While the rock is rising and falling vertically, the horizontal motion continues at a constant velocity. This example asks for the final velocity. Thus, the vertical and horizontal results will be recombined to obtain $v$ and $\theta_v$ at the final time t determined in the first part of the example.

**Solution for (a)**

While the rock is in the air, it rises and then falls to a final position 20.0 m lower than its starting altitude. We can find the time for this by using

$$y = y_0 + v_{0y}t - \frac{1}{2}gt^2.$$

If we take the initial position $y_0$ to be zero, then the final position is $y = -20.0m$. Now the initial vertical velocity is the vertical component of the initial velocity, found from $v_{0y} = v_0 sin\theta_0 = (25.0m/s)(sin35.0^{\circ}) = 14.3m/s$. Substituting known values yields

$$-20.0m = (14.3m/s)t - (4.90m/s^2)t^2.$$

Rearranging terms gives a quadratic equation in $t$:

$$(4.90m/s^2)t^2 - (14.3m/s)t - (20.0m) = 0.$$

This expression is a quadratic equation of the form $at^2 + bt + c = 0$, where the constants are $a = 4.90, b = -14.3$, and $c = -20.0$. Its solutions are given by the quadratic formula:

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

This equation yields two solutions: $t = 3.96$ and $t = -1.03$. (It is left as an exercise for the reader to verify these solutions.) The time is $t = 3.96s$ or $-1.03s$. The negative value of time implies an event before the start of motion, and so we discard it. Thus,

$$t = 3.96s.$$

**Discussion for (a)**

The time for projectile motion is completely determined by the vertical motion. So any projectile that has an initial vertical velocity of 14.3 m/s and lands 20.0 m below its starting altitude will spend 3.96 s in the air.

**Solution for (b)**

From the information now in hand, we can find the final horizontal and vertical velocities $v_x$ and $v_y$ and combine them to find the total velocity v and the angle $\theta_0$ it makes with the horizontal. Of course, $v_x$ is constant so we can solve for it at any horizontal location. In this case, we chose the starting point since we know both the initial velocity and initial angle. Therefore:

$$v_x = v_0 cos\theta_0 = (25.0m/s)(cos35^{\circ}) = 20.5m/s.$$

The final vertical velocity is given by the following equation:

$$v_y = v_{0y} - gt,$$

where $v_{0y}$ was found in part (a) to be **14.3 m/s**. Thus,

$$v_y = 14.3m/s - (9.80m/s^2)(3.96s)$$

so that

$$v_y = -24.5m/s.$$

To find the magnitude of the final velocity v we combine its perpendicular components, using the following equation:

$$v = \sqrt{v_x^2 + v_y^2} = \sqrt{(20.5m/s)^2 + (-24.5m/s)^2},$$

which gives

$$v = 31.9m/s.$$

The direction θv is found from the equation:

$$\theta_v = tan^{-1}(v_y/v_x)$$

so that

$$\theta_v = tan^{-1}(-24.5/20.5) = tan^{-1}(-1.19).$$

Thus,

$$\theta_v = -50.1^{\circ}.$$

**Discussion for (b)**

The negative angle means that the velocity is **50.1°** below the horizontal. This result is consistent with the fact that the final vertical velocity is negative and hence downward—as you would expect because the final altitude is 20.0 m lower than the initial altitude. (See Figure.)

One of the most important things illustrated by projectile motion is that vertical and horizontal motions are independent of each other. Galileo was the first person to fully comprehend this characteristic. He used it to predict the range of a projectile. On level ground, we define **range** to be the horizontal distance traveled by a projectile. Galileo and many others were interested in the range of projectiles primarily for military purposes—such as aiming cannons. However, investigating the range of projectiles can shed light on other interesting phenomena, such as the orbits of satellites around the Earth. Let us consider projectile range further.

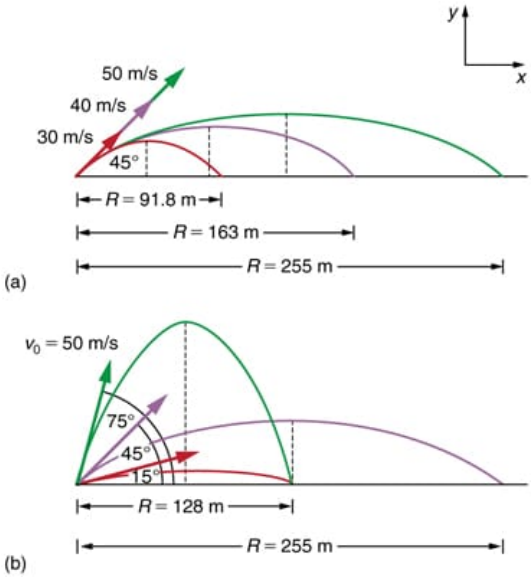



Figure $2.10.5$: Trajectories of projectiles on level ground. (a) The greater the initial speed $v_0$, the greater the range for a given initial angle. (b) The effect of initial angle $\theta_0$ on the range of a projectile with a given initial speed. Note that the range is the same for 15° and 75°, although the maximum heights of those paths are different.

How does the initial velocity of a projectile affect its range? Obviously, the greater the initial speed $v_0$, the greater the range, as shown in Figure(a). The initial angle $\theta_0$ also has a dramatic effect on the range, as illustrated in Figure(b). For a fixed initial speed, such as might be produced by a cannon, the maximum range is obtained with $\theta_0 = 45^{\circ}$. This is true only for conditions neglecting air resistance. If air resistance is considered, the maximum angle is approximately $38^{\circ}$. Interestingly, for every initial angle except $45^{\circ}$, there are two angles that give the same range—the sum of those angles is $90^{\circ}$. The range also depends on the value of the acceleration of gravity $g$. The lunar astronaut Alan Shepherd was able to drive a golf ball a great distance on the Moon because gravity is weaker there. The range $R$ of a projectile on level ground for which air resistance is negligible is given by

$$R = \frac{v_0^2 sin2\theta_0}{g},$$

where $v_0$ is the initial speed and $\theta_0$ is the initial angle relative to the horizontal. The proof of this equation is left as an end-of-chapter problem (hints are given), but it does fit the major features of projectile range as described.

When we speak of the range of a projectile on level ground, we assume that $R$ is very small compared with the circumference of the Earth. If, however, the range is large, the Earth curves away below the projectile and acceleration of gravity changes direction along the path. The range is larger than predicted by the range equation given above because the projectile has farther to fall than it would on level ground. (See Figure.) If the initial speed is great enough, the projectile goes into orbit. This is called exit velocity. This possibility was recognized centuries before it could be accomplished. When an object is in orbit, the Earth curves away from underneath the object at the same rate as it falls. The object thus falls continuously but never hits the surface. These and other aspects of orbital motion, such as the rotation of the Earth, will be covered analytically and in greater depth later in this text.

Once again we see that thinking about one topic, such as the range of a projectile, can lead us to others, such as the Earth orbits. In Addition of Velocities, we will examine the addition of velocities, which is another important aspect of two-dimensional kinematics and will also yield insights beyond the immediate topic.

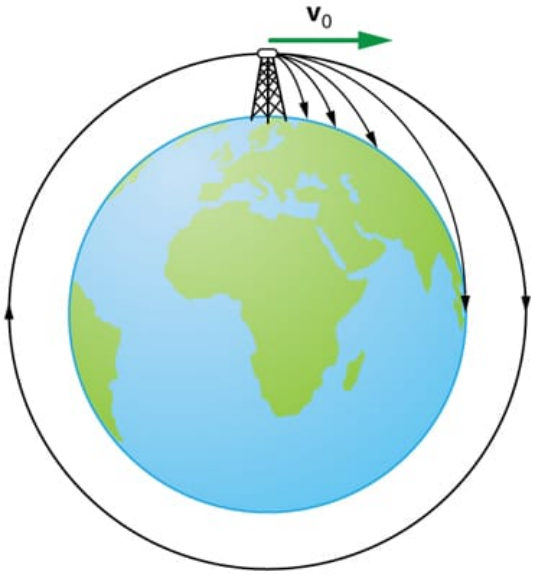



Figure 2.10.6: Projectile to satellite. In each case shown here, a projectile is launched from a very high tower to avoid air resistance. With increasing initial speed, the range increases and becomes longer than it would be on level ground because the Earth curves away underneath its path. With a large enough initial speed, orbit is achieved.

PHET EXPLORATIONS: PROJECTILE MOTION

Blast a Buick out of a cannon! Learn about projectile motion by firing various objects. Set the angle, initial speed, and mass. Add air resistance. Make a game out of this simulation by trying to hit a target.




Figure 2.10.7: Projectile Motion

## Summary

- Projectile motion is the motion of an object through the air that is subject only to the acceleration of gravity.
- To solve projectile motion problems, perform the following steps:

1. Determine a coordinate system. Then, resolve the position and/or velocity of the object in the horizontal and vertical components. The components of position s are given by the quantities x and y, and the components of the velocity v are given by vx=vcosθ and vy=vsinθ, where v is the magnitude of the velocity and θ is its direction.

2. Analyze the motion of the projectile in the horizontal direction using the following equations:

$$\text{Horizontal motion}(a_x = 0)$$

$$x = x_0 + v_x t$$

$$v_x = v_{0x} = v_x = \text{velocity is a constant.}$$

3. Analyze the motion of the projectile in the vertical direction using the following equations:

$$\text{Vertical motion(Assuming positive direction is up;} a_y = -g = -9.80 m/s^2)$$

$$y = y_0 + \frac{1}{2}(v_{0y} + v_y)t$$

$$v_y = v_{0y} - gt$$

$$y = y_0 + v_{0y}t - \frac{1}{2}gt^2$$

$$v_y^2 = v_{0y}^2 - 2g(y - y_0).$$

4. Recombine the horizontal and vertical components of location and/or velocity using the following equations:

$$s = \sqrt{x^2 + y^2}$$

$$\theta = tan^{-1}(y/x)$$

$$v = \sqrt{v_x^2 + v_y^2}$$

$$\theta_v = tan^{-1}(v_y/v_x).$$

- The maximum height $h$ of a projectile launched with initial vertical velocity $v_{0y}$ is given by

$$h = \frac{v_{0y}^2}{2g}.$$

- The maximum horizontal distance traveled by a projectile is called the **range**. The range $R$ of a projectile on level ground launched at an angle $\theta_0$ above the horizontal with initial speed $v_0$ is given by

$$R = \frac{v_0^2 sin 2\theta_0}{g}.$$

# Glossary

**air resistance**
: a frictional force that slows the motion of objects as they travel through the air; when solving basic physics problems, air resistance is assumed to be zero

**kinematics**
: the study of motion without regard to mass or force

**motion**
: displacement of an object as a function of time

**projectile**
: an object that travels through the air and experiences only acceleration due to gravity

**projectile motion**
: the motion of an object that is subject only to the acceleration of gravity

**range**
: the maximum horizontal distance that a projectile travels

**trajectory**
: the path of a projectile through the air

This page titled 2.10: Projectile Motion is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 3.4: Projectile Motion by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.11: Prelude to Uniform Circular Motion and Gravitation

Many motions, such as the arc of a bird's flight or Earth's path around the Sun, are curved. Recall that Newton's first law tells us that motion is along a straight line at constant speed unless there is a net external force. We will therefore study not only motion along curves, but also the forces that cause it, including gravitational forces. In some ways, this chapter is a continuation of Dynamics: Newton's Laws of Motion as we study more applications of Newton's laws of motion.



Figure 2.11.1: This Australian Grand Prix Formula 1 race car moves in a circular path as it makes the turn. Its wheels also spin rapidly—the latter completing many revolutions, the former only part of one (a circular arc). The same physical principles are involved in each. (credit: Richard Munckton)

This chapter deals with the simplest form of curved motion, **uniform circular motion**, motion in a circular path at constant speed. Studying this topic illustrates most concepts associated with rotational motion and leads to the study of many new topics we group under the name *rotation*. Pure *rotational motion* occurs when points in an object move in circular paths centered on one point. Pure *translational motion* is motion with no rotation. Some motion combines both types, such as a rotating hockey puck moving along ice.

## Glossary

**uniform circular motion**
: the motion of an object in a circular path at constant speed

This page titled 2.11: Prelude to Uniform Circular Motion and Gravitation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 6.0: Prelude to Uniform Circular Motion and Gravitation by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.12: Rotation Angle and Angular Velocity

**Learning Objectives**

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

- Define arc length, rotation angle, radius of curvature and angular velocity.
- Calculate the angular velocity of a car wheel spin.

In Kinematics, we studied motion along a straight line and introduced such concepts as displacement, velocity, and acceleration. Two-Dimensional Kinematics dealt with motion in two dimensions. Projectile motion is a special case of two-dimensional kinematics in which the object is projected into the air, while being subject to the gravitational force, and lands a distance away. In this chapter, we consider situations where the object does not land but moves in a curve. We begin the study of uniform circular motion by defining two angular quantities needed to describe rotational motion.

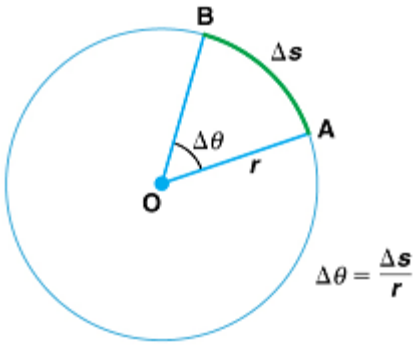
## Rotation Angle

When objects rotate about some axis—for example, when the CD (compact disc) in Figure rotates about its center—each point in the object follows a circular arc. Consider a line from the center of the CD to its edge. Each pit used to record sound along this line moves through the same angle in the same amount of time. The rotation angle is the amount of rotation and is analogous to linear distance. We define the rotation angle $\Delta\theta$ to be the ratio of the arc length to the radius of curvature:

$$\Delta\theta = \frac{\Delta s}{r}. \tag{2.12.1}$$




Figure 2.12.1: All points on a CD travel in circular arcs. The pits along a line




from the center to the edge all move through the same angle $\Delta\theta$ in a time $\Delta t$. Figure 2.12.2 : The radius of a circle is rotated through an angle $\Delta\theta$. The arc length $\delta s$ is described on the circumference.

The arc length $\Delta s$ is the distance traveled along a circular path as shown in Figure. Note that is the radius of curvature of the circular path.We know that for one complete revolution, the arc length is the circumference of a circle of radius $r$. The circumference of a circle is $2\pi r$.

Thus for one complete revolution the rotation angle is

$$\Delta\theta = \frac{2\pi r}{r} = 2\pi. \tag{2.12.2}$$

This result is the basis for defining the units used to measure rotation angles, $\Delta\theta$ to be **radians** (rad), defined so that

$$2\pi\, radians = 1\ revolution. \tag{2.12.3}$$

A comparison of some useful angles expressed in both degrees and radians is shown in Table 2.12.1.

*Table 2.12.1: Comparison in Angular Units*

| Degree Measure | Radian Measure |
|---|---|
| $30^o$ | $\frac{\pi}{6}$ |
| $60^o$ | $\frac{\pi}{3}$ |
| $90^o$ | $\frac{\pi}{2}$ |
| $120^o$ | $\frac{2\pi}{3}$ |
| $135^o$ | $\frac{3\pi}{4}$ |
| $180^o$ | $\pi$ |

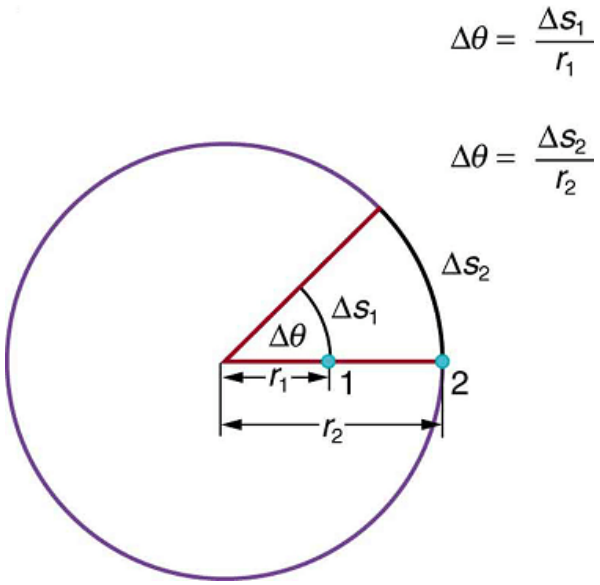



Figure 2.12.3: Points 1 and 2 rotate through the same angle $(\Delta\theta)$, but point 2 moves through a greater arc length $(\Delta s)$ because it is at a greater distance from the center of rotation $(r)$.

If $\Delta\theta = 2\pi\ rad$, then the CD has made one complete revolution, and every point on the CD is back at its original position. Because there are $360^o$ in a circle or one revolution, the relationship between radians and degrees is thus

$$2\pi\, rad = 360^o \tag{2.12.4}$$

so that

$$1\, rad = \frac{360^o}{2\pi} \approx 57.3^o \tag{2.12.5}$$

## Angular Velocity

How fast is an object rotating? We define angular velocity $\omega$ as the rate of change of an angle. In symbols, this is

$$\omega = \frac{\Delta\theta}{\Delta t}, \tag{2.12.6}$$

where an angular rotation $\Delta\theta$ takes place in a time $\Delta t$. The greater the rotation angle in a given amount of time, the greater the angular velocity. The units for angular velocity are radians per second (rad/s).Angular velocity $\omega$ is analogous to linear velocity $v$. To get the precise relationship between angular and linear velocity, we again consider a pit on the rotating CD. This pit moves an arc length $\Delta s$ in a time $\Delta t$, and so it has a linear velocity

$$v = \frac{\Delta s}{\Delta t}. \tag{2.12.7}$$

From $\Delta\theta = \frac{\Delta s}{r}$ we see that $\Delta s = r\Delta\theta$. Substituting this into the expression for $v$ gives

$$v = \frac{r\Delta\theta}{\Delta t} = r\omega. \tag{2.12.8}$$

We write this relationship in two different ways and gain two different insights:

$$v = r\omega,\ or\ \omega = \frac{v}{r}. \tag{2.12.9}$$

The first relationship in $v = r\omega,\ \ or\ \ \omega = \frac{v}{r}$ states that the linear velocity $v$ is proportional to the distance from the center of rotation, thus, it is largest for a point on the rim (largest $r$), as you might expect. We can also call this linear speed $v$ of a point on the rim the ***tangential speed***. The second relationship in $v = r\omega,\ \ or\ \ \omega = \frac{v}{r}$ can be illustrated by considering the tire of a moving car. Note that the speed of a point on the rim of the tire is the same as the speed $v$ of the car. See Figure So the faster the

car moves, the faster the tire spins—large $v$ means a large $\omega$, because $v = r\omega$. Similarly, a larger-radius tire rotating at the same angular velocity $(\omega)$ will produce a greater linear speed $(v)$ for the car.

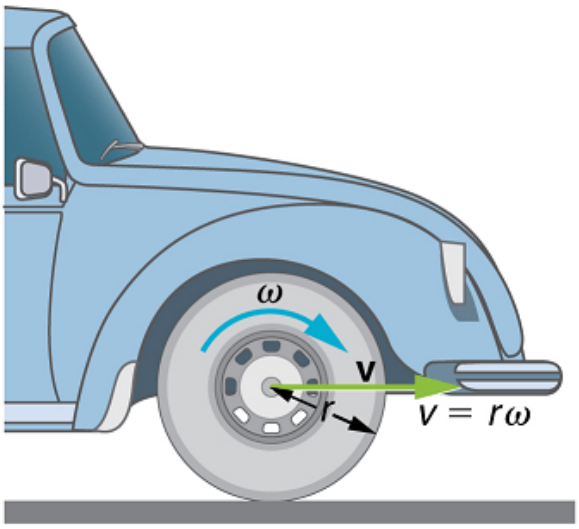



Figure 2.12.4: A car moving at a velocity $v$ to the right has a tire rotating with an angular velocity $\omega$. The speed of the tread of the tire relative to the axle is $v$, the same as if the car were jacked up. Thus the car moves forward at linear velocity $v = r\omega$, where $r$ is the tire radius. A larger angular velocity for the tire means a greater velocity for the car.

## Example 2.12.1: How Fast Does the Car Tire Spin?

Calculate the angular velocity of a 0.300 m radius car tire when the car travels at $15.0 m/s$ (about $54\ km/h$). See Figure.

**Strategy**

Because the linear speed of the tire rim is the same as the speed of the car, we have $v = 15.0 m/s$. The radius of the tire is given to be $r = 0.300\ m$. Knowing $v$ and $r$, we can use the second relationship in $v = \omega r$, $\omega = \frac{v}{r}$ to calculate the angular velocity.

**Solution**

To calculate the angular velocity, we will use the following relationship:

$$\omega = \frac{v}{r}. \tag{2.12.10}$$

Substituting the knowns,

$$\omega = \frac{15.0\, m/s}{0.300\, m} = 50.0\, rad/s. \tag{2.12.11}$$

**Discussion**

When we cancel units in the above calculation, we get 50.0/s. But the angular velocity must have units of rad/s. Because radians are actually unitless (radians are defined as a ratio of distance), we can simply insert them into the answer for the angular velocity. Also note that if an earth mover with much larger tires, say 1.20 m in radius, were moving at the same speed of 15.0 m/s, its tires would rotate more slowly. They would have an angular velocity

$$\omega = (15.0\, m/s)/(1.20\, m) = 12.5\, rad/s. \tag{2.12.12}$$

Both $\omega$ and $v$ have directions (hence they are angular and linear *velocities*, respectively). Angular velocity has only two directions with respect to the axis of rotation—it is either clockwise or counterclockwise. Linear velocity is tangent to the path, as illustrated in Figure.

## TAKE-HOME EXPERIMENT

Tie an object to the end of a string and swing it around in a horizontal circle above your head (swing at your wrist). Maintain uniform speed as the object swings and measure the angular velocity of the motion. What is the approximate speed of the object? Identify a point close to your hand and take appropriate measurements to calculate the linear speed at this point. Identify other circular motions and measure their angular velocities.

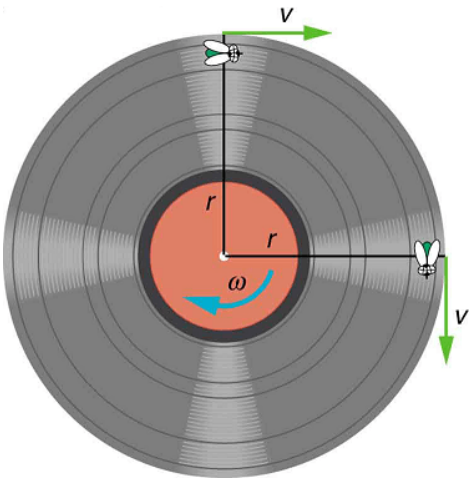



Figure 2.12.5: As an object moves in a circle, here a fly on the edge of an old-fashioned vinyl record, its instantaneous velocity is always tangent to the circle. The direction of the angular velocity is clockwise in this case.

## PHET EXPLORATIONS: LADYBUG REVOLUTION

Join the ladybug in an exploration of rotational motion. Rotate the merry-go-round to change its angle, or choose a constant angular velocity or angular acceleration. Explore how circular motion relates to the bug's x,y position, velocity, and acceleration using vectors or graphs.

# Section Summary

- Uniform circular motion is motion in a circle at constant speed. The rotation angle $\delta\theta$ is defined as the ratio of the arc length to the radius of curvature:

$$\Delta\theta = \frac{\Delta s}{r} \tag{2.12.13}$$

  where arc length $\delta s$ is distance traveled along a circular path and $r$ is the radius of curvature of the circular path. The quantity $\Delta\theta$ is measured in units of radians (rad), for which

$$2\pi\, rad = 360^o = 1\, revolution. \tag{2.12.14}$$

- The conversion between radians and degrees is

$$1\, rad = 57.3^o. \tag{2.12.15}$$

- Angular velocity $\omega$ is the rate of change of an angle,

$$\omega = \frac{\Delta\theta}{\Delta t}, \tag{2.12.16}$$

where a rotation $\Delta\theta$ takes place in a time $\Delta t$. The units of angular velocity are radians per second (rad/s). Linear velocity $v$ and angular velocity $\omega$ are related by

$$v = r\omega,\ or\ \omega = \frac{v}{r}. \tag{2.12.17}$$

# Glossary

**arc length**
: $\Delta s$, the distance traveled by an object along a circular path

**pit**
: a tiny indentation on the spiral track moulded into the top of the polycarbonate layer of CD

**rotation angle**
: the ratio of the arc length to the radius of curvature on a circular path: $\Delta\theta=\frac{\Delta s}{r}$

**radius of curvature**
: radius of a circular path

**radians**
: a unit of angle measurement

**angular velocity**
: $\omega$, the rate of change of the angle with which an object moves on a circular path

This page titled 2.12: Rotation Angle and Angular Velocity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 6.1: Rotation Angle and Angular Velocity by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 2.13: Centripetal Acceleration

## Learning Objectives

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

- Establish the expression for centripetal acceleration.
- Explain the centrifuge.

We know from kinematics that acceleration is a change in velocity, either in its magnitude or in its direction, or both. In uniform circular motion, the direction of the velocity changes constantly, so there is always an associated acceleration, even though the magnitude of the velocity might be constant. You experience this acceleration yourself when you turn a corner in your car. (If you hold the wheel steady during a turn and move at constant speed, you are in uniform circular motion.) What you notice is a sideways acceleration because you and the car are changing direction. The sharper the curve and the greater your speed, the more noticeable this acceleration will become. In this section we examine the direction and magnitude of that acceleration.

Figure $2.13.1$ shows an object moving in a circular path at constant speed. The direction of the instantaneous velocity is shown at two points along the path. Acceleration is in the direction of the change in velocity, which points directly toward the center of rotation (the center of the circular path). This pointing is shown with the vector diagram in the figure. We call the acceleration of an object moving in uniform circular motion (resulting from a net external force) the **centripetal acceleration** $a_c$; centripetal means “toward the center” or “center seeking.”

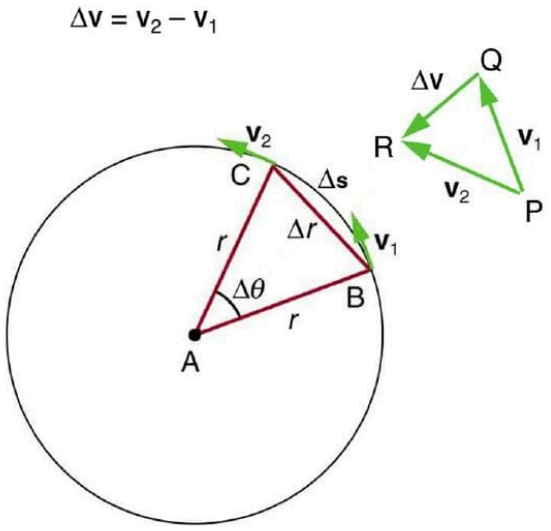



Figure $2.13.1$:The directions of the velocity of an object at two different points are shown, and the change in velocity $\Delta v$ is seen to point directly toward the center of curvature. (See small inset.) Because $a_c = \Delta v/\Delta t$, the acceleration is also toward the center; $a_c$ is called centripetal acceleration. (Because $\delta\theta$ is very small, the arc length $\Delta s$ is equal to the chord length $\Delta r$ for small time differences.)

The direction of centripetal acceleration is toward the center of curvature, but what is its magnitude? Note that the triangle formed by the velocity vectors and the one formed by the radii $r$ and $\Delta s$ are similar. Both the triangles ABC and PQR are isosceles triangles (two equal sides). The two equal sides of the velocity vector triangle are the speeds $v_1 = v_2 = v$ Using the properties of two similar triangles, we obtain

$$\frac{\Delta v}{v} = \frac{\Delta s}{r}. \tag{2.13.1}$$

Acceleration is $\Delta v/\Delta t$ and so we first solve this expression for $\delta v$:

$$\delta v = \frac{v}{r}\Delta s. \tag{2.13.2}$$

Then we divide this by $\Delta t$, yielding

$$\frac{\Delta v}{\Delta t} = \frac{v}{r} \times \frac{\Delta s}{\Delta t}. \tag{2.13.3}$$

Finally, noting that $\Delta v/\Delta t = a_c$ and that $\delta s/\Delta t = v$ the linear or tangential speed, we see that the magnitude of the centripetal acceleration is

$$a_c = \frac{v^2}{r}, \tag{2.13.4}$$

which is the acceleration of an object in a circle of radius $r$ at a speed $v$.

So, centripetal acceleration is greater at high speeds and in sharp curves (smaller radius), as you have noticed when driving a car. But it is a bit surprising that $a_c$ is proportional to speed squared, implying, for example, that it is four times as hard to take a curve at 100 km/h than at 50 km/h. A sharp corner has a small radius, so that $a_c$ is greater for tighter turns, as you have probably noticed.It is also useful to express $a_c$ in terms of angular velocity. Substituting $v = r\omega$ into the above expression, we find $a_c = (r\omega^2)/r = r\omega^2$. We can express the magnitude of centripetal acceleration using either of two equations:

$$a_c = \frac{v^2}{r};\ a_c = r\omega^2 \tag{2.13.5}$$

Recall that the direction of $a_c$ is toward the center. You may use whichever expression is more convenient, as illustrated in examples below.

A **centrifuge** (Figure $2.13.2b$) is a rotating device used to separate specimens of different densities. High centripetal acceleration significantly decreases the time it takes for separation to occur, and makes separation possible with small samples. Centrifuges are used in a variety of applications in science and medicine, including the separation of single cell suspensions such as bacteria, viruses, and blood cells from a liquid medium and the separation of macromolecules, such as DNA and protein, from a solution. Centrifuges are often rated in terms of their centripetal acceleration relative to acceleration due to gravity $(g)$ maximum centripetal acceleration of several hundred thousand $g$ is possible in a vacuum. Human centrifuges, extremely large centrifuges, have been used to test the tolerance of astronauts to the effects of accelerations larger than that of Earth's gravity.

### Example$2.13.1$: Centripetal Acceleration vs. Gravity?

What is the magnitude of the centripetal acceleration of a car following a curve of radius 500 m at a speed of 25.0 m/s (about 90 km/h)? Compare the acceleration with that due to gravity for this fairly gentle curve taken at highway speed. See Figure $2.13.2a$.

**Strategy**

Because $v$ and $r$ are given, the first expression in $a_c = \frac{v^2}{r}$ : $a_c = r\omega^2$ is the most convenient to use.

**Solution**

Entering the given values of $v = 25.0\ m/s and r = 500\ m$ into the first expression for $a_c$ gives

$$a_c = \frac{v^2}{r} = \frac{(25.0 m/s)^2}{500\, m} = 1.25\, m/s^2.$$

**Discussion**

To compare this with the acceleration due to gravity $(g = 9.80\ m/s^2)$, we take the ratio of $a_c/g = (1.25\ m/s^2)/(9.80\ m/s^2) = 0.128$. Thus, $a_c = 0.128g$ and is noticeable especially if you were not wearing a seat belt.

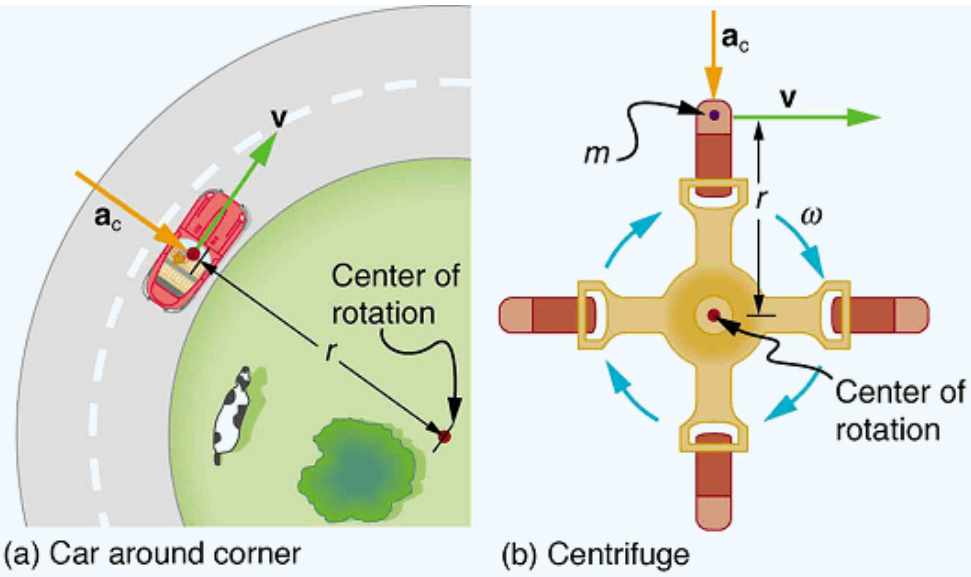



Figure 2.13.2: (a) The car following a circular path at constant speed is accelerated perpendicular to its velocity, as shown. The magnitude of this centripetal acceleration is found in Example. (b) A particle of mass in a centrifuge is rotating at constant angular velocity . It must be accelerated perpendicular to its velocity or it would continue in a straight line. The magnitude of the necessary acceleration is found in Example.

## Example2.13.2:How Big is the Centripetal Acceleration in an Ultracentrifuge?

Calculate the centripetal acceleration of a point 7.50 cm from the axis of an **ultracentrifuge** spinning at $7.4 \times 10^7\ rev/min$. Determine the ratio of this acceleration to that due to gravity. See Figure Figure 2.13.2*b*.

**Strategy**

The term rev/min stands for revolutions per minute. By converting this to radians per second, we obtain the angular velocity $\omega$. Because $r$ is given, we can use the second expression in the equation $a_c = \frac{v^2}{r};\ a_c = r\omega^2$ to calculate the centripetal acceleration.

**Solution**

To convert $7.40 \times 10^4\ rev/min$ to radians per second, we use the facts that one revolution is $2\pi\ rad$ and one minute is 60.0 s. Thus,

$$\omega = 7.40 \times 10^4 \frac{rev}{min} \times \frac{2\pi\, rad}{1\, rev} \times \frac{1\, min}{60\, sec} = 7745\, rad/sec.$$

Now the centripetal acceleration is given by the second expression in $a_c = \frac{v^2}{r};\ a_c = r\omega^2$ as

$$a_c = r\omega^2.$$

Converting 7.50 cm to meters and substituting known values gives

$$a_c = (0.0750\, m)(7854\, rad/sec)^2 = 4.50 \times 10^6\, m/s^2.$$

Note that the unitless radians are discarded in order to get the correct units for centripetal acceleration. Taking the ratio of $a_c$ to $g$ yields

$$\frac{a_c}{g} = \frac{4.63 \times 10^6}{9.80} = 4.59 \times 10^5.$$

**Discussion**

This last result means that the centripetal acceleration is 472,000 times as strong as $g$. It is no wonder that such high $\omega$ centrifuges are called ultracentrifuges. The extremely large accelerations involved greatly decrease the time needed to cause the sedimentation of blood cells or other materials.

Of course, a net external force is needed to cause any acceleration, just as Newton proposed in his second law of motion. So a net external force is needed to cause a centripetal acceleration. In the section on Centripetal Force, we will consider the forces involved in circular motion.

PHET EXPLORATIONS: LADYBUG MOTION 2D

Learn about position, velocity and acceleration vectors. Move the ladybug by setting the position, velocity or acceleration, and see how the vectors change. Choose linear, circular or elliptical motion, and record and playback the motion to analyze the behavior.

## Summary

- Centripetal acceleration $a_c$ is the acceleration experienced while in uniform circular motion. It always points toward the center of rotation. It is perpendicular to the linear velocity $v$ and has the magnitude

$$a_c = \frac{v^2}{r};\ a_c = r\omega^2.$$

- The unit of centripetal acceleration is $m/s^2$.

## Glossary

**centripetal acceleration**
: the acceleration of an object moving in a circle, directed toward the center

**ultracentrifuge**
: a centrifuge optimized for spinning a rotor at very high speeds

This page titled 2.13: Centripetal Acceleration is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 6.2: Centripetal Acceleration by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3: MODULE 2-MOTION AND FORCES


3.1: Prelude to Dynamics- Newton’s Laws of Motion

3.2: Development of Force Concept

3.3: Newton’s First Law of Motion - Inertia

3.4: Newton’s Second Law of Motion- Concept of a System

3.5: Newton’s Third Law of Motion- Symmetry in Forces

3.6: Summary of Newton's Laws

3.7: Newton’s Universal Law of Gravitation



This page titled 3: Module 2-Motion and Forces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

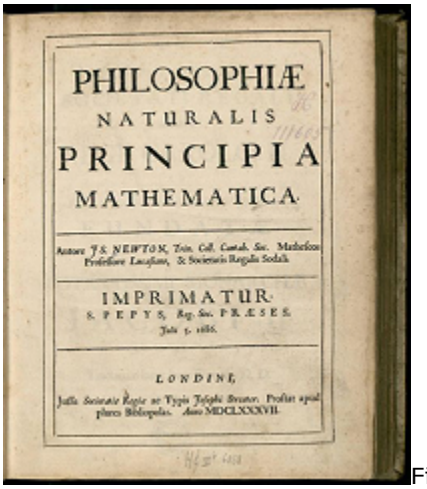
# 3.1: Prelude to Dynamics- Newton's Laws of Motion

Motion draws our attention. Motion itself can be beautiful, causing us to marvel at the forces needed to achieve spectacular motion, such as that of a dolphin jumping out of the water, or a pole vaulter, or the flight of a bird, or the orbit of a satellite. The study of motion is kinematics, but kinematics only *describes* the way objects move—their velocity and their acceleration. **Dynamics** considers the forces that affect the motion of moving objects and systems. Newton's laws of motion are the foundation of dynamics. These laws provide an example of the breadth and simplicity of principles under which nature functions. They are also universal laws in that they apply to similar situations on Earth as well as in space.



Figure 4.1.1. Newton's laws of motion describe the motion of the dolphin's path. (credit: Jin Jang)

Isaac Newton's (1642–1727) laws of motion were just one part of the monumental work that has made him legendary. The development of Newton's laws marks the transition from the Renaissance into the modern era. This transition was characterized by a revolutionary change in the way people thought about the physical universe. For many centuries natural philosophers had debated the nature of the universe based largely on certain rules of logic with great weight given to the thoughts of earlier classical philosophers such as Aristotle (384–322 BC). Among the many great thinkers who contributed to this change were Newton and Galileo.


PHILOSOPHIÆ
NATURALIS
PRINCIPIA
MATHEMATICA

IMPRIMATUR
S. PEPYS, Reg. Soc. PRÆSES.

LONDINI

Figure 4.1.2. Issac Newton's monumental work, Philosophiae Naturalis Principia Mathematica, was published in 1687. It proposed scientific laws that are still used today to describe the motion of objects. (credit: Service commun de la documentation de l'Université de Strasbourg)

Galileo was instrumental in establishing *observation* as the absolute determinant of truth, rather than “logical” argument. Galileo’s use of the telescope was his most notable achievement in demonstrating the importance of observation. He discovered moons orbiting Jupiter and made other observations that were inconsistent with certain ancient ideas and religious dogma. For this reason, and because of the manner in which he dealt with those in authority, Galileo was tried by the Inquisition and punished. He spent the final years of his life under a form of house arrest. Because others before Galileo had also made discoveries by *observing* the nature of the universe, and because repeated observations verified those of Galileo, his work could not be suppressed or denied. After his death, his work was verified by others, and his ideas were eventually accepted by the church and scientific communities.

Galileo also contributed to the formation of what is now called Newton’s first law of motion. Newton made use of the work of his predecessors, which enabled him to develop laws of motion, discover the law of gravity, invent calculus, and make great contributions to the theories of light and color. It is amazing that many of these developments were made with Newton working alone, without the benefit of the usual interactions that take place among scientists today.

It was not until the advent of modern physics early in the 20th century that it was discovered that Newton’s laws of motion produce a good approximation to motion only when the objects are moving at speeds much, much less than the speed of light and when those objects are larger than the size of most molecules (about m in diameter). These constraints define the realm of classical mechanics, as discussed in Introduction to the Nature of Science and Physics. At the beginning of the 20th century, Albert Einstein (1879–1955) developed the theory of relativity and, along with many other scientists, developed quantum theory. This theory does not have the constraints present in classical physics. All of the situations we consider in this chapter, and all those preceding the introduction of relativity in Special Relativity, are in the realm of classical physics.

## MAKING CONNECTIONS: PAST AND PRESENT PHILOSOPHY

*The importance of observation* and the concept of *cause and effect* were not always so entrenched in human thinking. This realization was a part of the evolution of modern physics from natural philosophy. The achievements of Galileo, Newton, Einstein, and others were key milestones in the history of scientific thought. Most of the scientific theories that are described in this book descended from the work of these scientists.*The importance of observation* and the concept of *cause and effect* were not always so entrenched in human thinking. This realization was a part of the evolution of modern physics from natural philosophy. The achievements of Galileo, Newton, Einstein, and others were key milestones in the history of scientific thought. Most of the scientific theories that are described in this book descended from the work of these scientists.

This page titled 3.1: Prelude to Dynamics- Newton’s Laws of Motion is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 4.0: Prelude to Dynamics- Newton’s Laws of Motion by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3.2: Development of Force Concept

## Learning Objectives

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

- Understand the definition of force.

**Dynamics** is the study of the forces that cause objects and systems to move. To understand this, we need a working definition of force. Our intuitive definition of **force**—that is, a push or a pull—is a good place to start. We know that a push or pull has both magnitude and direction (therefore, it is a vector quantity) and can vary considerably in each regard. For example, a cannon exerts a strong force on a cannonball that is launched into the air. In contrast, Earth exerts only a tiny downward pull on a flea. Our everyday experiences also give us a good idea of how multiple forces add. If two people push in different directions on a third person, as illustrated in Figure 3.2.1, we might expect the total force to be in the direction shown. Since force is a vector, it adds just like other vectors, as illustrated in Figure 3.2.1$a$ for two ice skaters. Forces, like other vectors, are represented by arrows and can be added using the familiar head-to-tail method or by trigonometric methods. These ideas were developed in Two-Dimensional Kinematics.

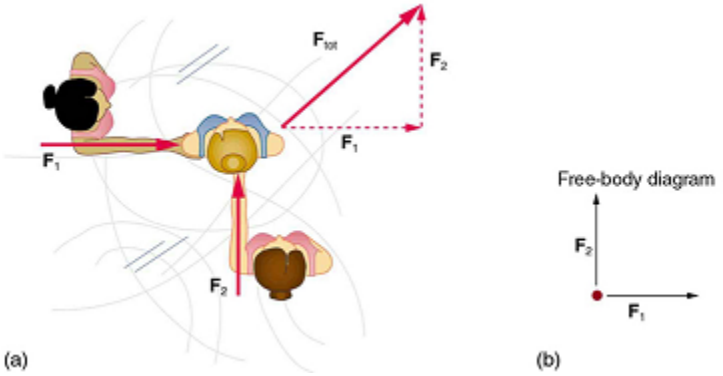



Figure 3.2.1: Part (a) shows an overhead view of two ice skaters pushing on a third. Forces are vectors and add like other vectors, so the total force on the third skater is in the direction shown. In part (b), we see a free-body diagram representing the forces acting on the third skater.

Figure 3.2.1$b$ is our first example of a free-body diagram, which is a technique used to illustrate all the external forces acting on a body. The body is represented by a single isolated point (or free body), and only those forces acting *on* the body from the outside (external forces) are shown. (These forces are the only ones shown, because only external forces acting on the body affect its motion. We can ignore any internal forces within the body.) Free-body diagrams are very useful in analyzing forces acting on a system and are employed extensively in the study and application of Newton's laws of motion.

A more quantitative definition of force can be based on some standard force, just as distance is measured in units relative to a standard distance. One possibility is to stretch a spring a certain fixed distance, as illustrated in Figure 3.2.2, and use the force it exerts to pull itself back to its relaxed shape—called a *restoring force*—as a standard. The magnitude of all other forces can be stated as multiples of this standard unit of force. Many other possibilities exist for standard forces. (One that we will encounter in Magnetism is the magnetic force between two wires carrying electric current.) Some alternative definitions of force will be given later in this chapter.

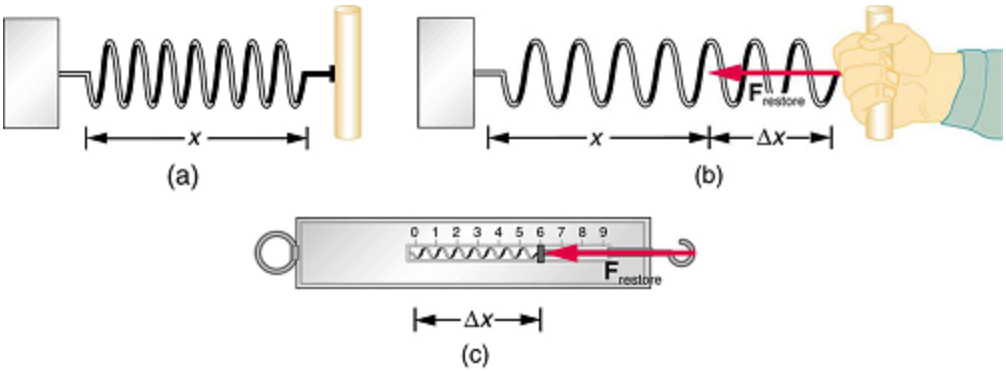



Figure 3.2.2:The force exerted by a stretched spring can be used as a standard unit of force. (a) This spring has a length when undistorted. (b) When stretched a distance $\Delta x$ the spring exerts a restoring force, $F_{restore}$ *which is reproducible. (c) A spring scale is one device that uses a spring to measure force. The force* $F_{restore}$ *is exerted on whatever is attached to the hook. Here* $F_{restore}$ *has a magnitude of 6 units in the force standard being employed.*

### TAKE HOME EXPERIMENT: FORCE STANDARDS

To investigate force standards and cause and effect, get two identical rubber bands. Hang one rubber band vertically on a hook. Find a small household item that could be attached to the rubber band using a paper clip, and use this item as a weight to investigate the stretch of the rubber band. Measure the amount of stretch produced in the rubber band with one, two, and four of these (identical) items suspended from the rubber band. What is the relationship between the number of items and the amount of stretch? How large a stretch would you expect for the same number of items suspended from two rubber bands? What happens to the amount of stretch of the rubber band (with the weights attached) if the weights are also pushed to the side with a pencil?

## Summary

- **Dynamics** is the study of how forces affect the motion of objects.
- **Force** is a push or pull that can be defined in terms of various standards, and it is a vector having both magnitude and direction.
- **External forces** are any outside forces that act on a body. A **free-body diagram** is a drawing of all external forces acting on a body.

## Glossary

**dynamics**
: the study of how forces affect the motion of objects and systems

**external force**
: a force acting on an object or system that originates outside of the object or system

**free-body diagram**
: a sketch showing all of the external forces acting on an object or system; the system is represented by a dot, and the forces are represented by vectors extending outward from the dot

**force**
: a push or pull on an object with a specific magnitude and direction; can be represented by vectors; can be expressed as a multiple of a standard force

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

- 4.1: Development of Force Concept by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3.3: Newton’s First Law of Motion - Inertia

## Learning Objectives

- Define mass and inertia.
- Understand Newton's first law of motion.

Experience suggests that an object at rest will remain at rest if left alone, and that an object in motion tends to slow down and stop unless some effort is made to keep it moving. What **Newton’s first law of motion** states, however, is the following:

## Newton’s First Law of Motion

A body at rest remains at rest, or, if in motion, remains in motion at a constant velocity unless acted on by a net external force.

Note the repeated use of the verb “remains.” We can think of this law as preserving the status quo of motion.

Rather than contradicting our experience, **Newton’s first law of motion** states that there must be a *cause* (which is a net external force) *for there to be any change in velocity (either a change in magnitude or direction)*. We will define *net external force* in the next section. An object sliding across a table or floor slows down due to the net force of friction acting on the object. If friction disappeared, would the object still slow down?

The idea of cause and effect is crucial in accurately describing what happens in various situations. For example, consider what happens to an object sliding along a rough horizontal surface. The object quickly grinds to a halt. If we spray the surface with talcum powder to make the surface smoother, the object slides farther. If we make the surface even smoother by rubbing lubricating oil on it, the object slides farther yet. Extrapolating to a frictionless surface, we can imagine the object sliding in a straight line indefinitely. Friction is thus the *cause* of the slowing (consistent with Newton’s first law). The object would not slow down at all if friction were completely eliminated. Consider an air hockey table. When the air is turned off, the puck slides only a short distance before friction slows it to a stop. However, when the air is turned on, it creates a nearly frictionless surface, and the puck glides long distances without slowing down. Additionally, if we know enough about the friction, we can accurately predict how quickly the object will slow down. Friction is an external force.

Newton’s first law is completely general and can be applied to anything from an object sliding on a table to a satellite in orbit to blood pumped from the heart. Experiments have thoroughly verified that any change in velocity (speed or direction) must be caused by an external force. The idea of *generally applicable or universal laws* is important not only here—it is a basic feature of all laws of physics. Identifying these laws is like recognizing patterns in nature from which further patterns can be discovered. The genius of Galileo, who first developed the idea for the first law, and Newton, who clarified it, was to ask the fundamental question, “What is the cause?” Thinking in terms of cause and effect is a worldview fundamentally different from the typical ancient Greek approach when questions such as “Why does a tiger have stripes?” would have been answered in Aristotelian fashion, “That is the nature of the beast.” True perhaps, but not a useful insight.

## Mass

The property of a body to remain at rest or to remain in motion with constant velocity is called **inertia**. Newton’s first law is often called the law of inertia. As we know from experience, some objects have more inertia than others. It is obviously more difficult to change the motion of a large boulder than that of a basketball, for example. The inertia of an object is measured by its **mass**. Roughly speaking, mass is a measure of the amount of “stuff” (or matter) in something. The quantity or amount of matter in an object is determined by the numbers of atoms and molecules of various types it contains. Unlike weight, mass does not vary with location. The mass of an object is the same on Earth, in orbit, or on the surface of the Moon. In practice, it is very difficult to count and identify all of the atoms and molecules in an object, so masses are not often determined in this manner. Operationally, the masses of objects are determined by comparison with the standard kilogram.

**Exercise 3.3.1**

Which has more mass: a kilogram of cotton balls or a kilogram of gold?

*Answer available at the end of this section.*

## Summary

- **Newton's first law of motion** states that a body at rest remains at rest, or, if in motion, remains in motion at a constant velocity unless acted on by a net external force. This is also known as the **law of inertia**.
- **Inertia** is the tendency of an object to remain at rest or remain in motion. Inertia is related to an object's mass.
- **Mass** is the quantity of matter in a substance.

# Glossary

**inertia**
: the tendency of an object to remain at rest or remain in motion

**law of inertia**
: see Newton's first law of motion

**mass**
: the quantity of matter in a substance; measured in kilograms

**Newton's first law of motion**
: a body at rest remains at rest, or, if in motion, remains in motion at a constant velocity unless acted on by a net external force; also known as the law of inertia

## Answers

**Exercise** 3.3.1

They are equal. A kilogram of one substance is equal in mass to a kilogram of another substance. The quantities that might differ between them are volume and density.

This page titled 3.3: Newton's First Law of Motion - Inertia is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 4.2: Newton's First Law of Motion - Inertia by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3.4: Newton's Second Law of Motion- Concept of a System

## Learning Objectives

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

- Define net force, external force, and system.
- Understand Newton's second law of motion.
- Apply Newton's second law to determine the weight of an object.

**Newton's second law of motion** is closely related to Newton's first law of motion. It mathematically states the cause and effect relationship between force and changes in motion. Newton's second law of motion is more quantitative and is used extensively to calculate what happens in situations involving a force. Before we can write down Newton's second law as a simple equation giving the exact relationship of force, mass, and acceleration, we need to sharpen some ideas that have already been mentioned.

First, what do we mean by a change in motion? The answer is that a change in motion is equivalent to a change in velocity. A change in velocity means, by definition, that there is an **acceleration**. Newton's first law says that a net external force causes a change in motion; thus, we see that a *net external force causes acceleration*.

Another question immediately arises. What do we mean by an external force? An intuitive notion of external is correct — an **external force** acts from outside the **system** of interest. For example, in Figure $3.4.1a$ the system of interest is the wagon plus the child in it. The two forces exerted by the other children are external forces. An internal force acts between elements of the system. Again looking at Figure $3.4.1a$, the force the child in the wagon exerts to hang onto the wagon is an internal force between elements of the system of interest. Only external forces affect the motion of a system, according to Newton's first law. (The internal forces actually cancel, as we shall see in the next section.) *You must define the boundaries of the system before you can determine which forces are external*. Sometimes the system is obvious, whereas other times identifying the boundaries of a system is more subtle. The concept of a system is fundamental to many areas of physics, as is the correct application of Newton's laws. This concept will be revisited many times on our journey through physics.

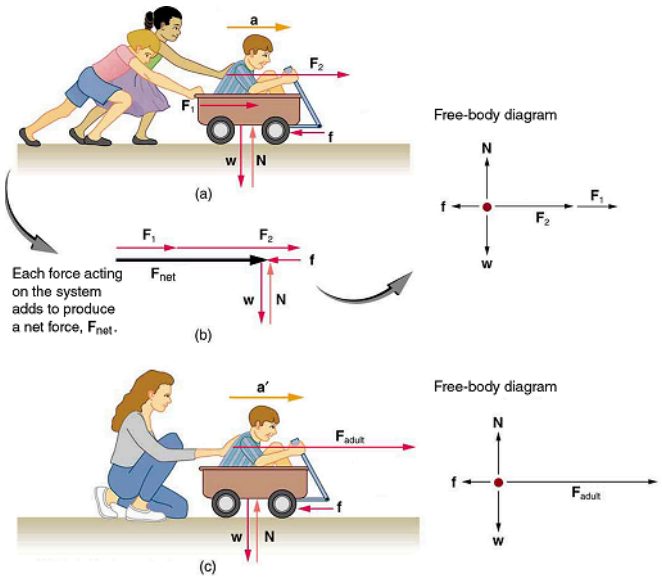



Figure $3.4.1$: Different forces exerted on the same mass produce different accelerations. (a) Two children push a wagon with a child in it. Arrows representing all external forces are shown. The system of interest is the wagon and its rider. The weight $w$ of the system and the support of the ground $N$ are also shown for completeness and are assumed to cancel. The vector $f$ represents the friction acting on the wagon, and it acts to the left, opposing the motion of the wagon. (b) All of the external forces acting on the system add together to produce a net force, $F_{net}$ . The free-body diagram shows all of the forces acting on the system of interest. The dot represents the center of mass of the system. Each force vector extends from this dot. Because there are two forces acting to the right, we draw the vectors collinearly.

(c) A larger net external force produces a larger acceleration $(al > a)$ when an adult pushes the child.

Now, it seems reasonable that acceleration should be directly proportional to and in the same direction as the net (total) external force acting on a system. This assumption has been verified experimentally and is illustrated in Figure. In part (a), a smaller force causes a smaller acceleration than the larger force illustrated in part (c). For completeness, the vertical forces are also shown; they are assumed to cancel since there is no acceleration in the vertical direction. The vertical forces are the weight $w$ and the support of the ground $N$, and the horizontal force $f$ represents the force of friction. These will be discussed in more detail in later sections. For now, we will define **friction** as a force that opposes the motion past each other of objects that are touching. Figure $3.4.1b$ shows how vectors representing the external forces add together to produce a net force, $F_{net}$.

To obtain an equation for Newton's second law, we first write the relationship of acceleration and net external force as the proportionality

$$a \propto F_{net} \tag{3.4.1}$$

where the symbol $\propto$ means "proportional to," and $F_{net}$ is the **net external force**. (The net external force is the vector sum of all external forces and can be determined graphically, using the head-to-tail method, or analytically, using components. The techniques are the same as for the addition of other vectors, and are covered in the chapter section on Two-Dimensional Kinematics.) This proportionality states what we have said in words—*acceleration is directly proportional to the net external force*. Once the system of interest is chosen, it is important to identify the external forces and ignore the internal ones. It is a tremendous simplification not to have to consider the numerous internal forces acting between objects within the system, such as muscular forces within the child's body, let alone the myriad of forces between atoms in the objects, but by doing so, we can easily solve some very complex problems with only minimal error due to our simplification

Now, it also seems reasonable that acceleration should be inversely proportional to the mass of the system. In other words, the larger the mass (the inertia), the smaller the acceleration produced by a given force. And indeed, as illustrated in Figure, the same net external force applied to a car produces a much smaller acceleration than when applied to a basketball. The proportionality is written as

$$a \propto \frac{1}{m}, \tag{3.4.2}$$

where $m$ is the mass of the system. Experiments have shown that acceleration is exactly inversely proportional to mass, just as it is exactly linearly proportional to the net external force.

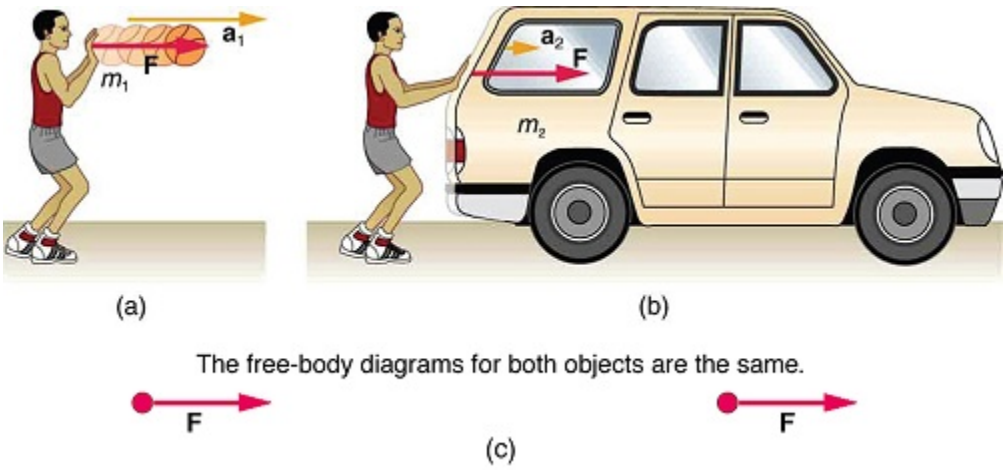



Figure $3.4.2$:The same force exerted on systems of different masses produces different accelerations. (a) A basketball player pushes on a basketball to make a pass. (The effect of gravity on the ball is ignored.) (b) The same player exerts an identical force on a stalled SUV and produces a far smaller acceleration (even if friction is negligible). (c) The free-body diagrams are identical, permitting direct comparison of the two situations. A series of patterns for the free-body diagram will emerge as you do more problems.

It has been found that the acceleration of an object depends *only* on the net external force and the mass of the object. Combining the two proportionalities just given yields Newton's second law of motion.

### Newton's Second Law of Motion

The acceleration of a system is directly proportional to and in the same direction as the net external force acting on the system, and inversely proportional to its mass. In equation form, Newton's second law of motion is

$$a = \frac{F_{net}}{m} \tag{3.4.3}$$

This is often written in the more familiar form

$$F_{net} = ma. \tag{3.4.4}$$

When only the magnitude of force and acceleration are considered, this equation is simply

$$F_{net} = ma. \tag{3.4.5}$$

Although these last two equations are really the same, the first gives more insight into what Newton's second law means. The law is a *cause and effect relationship* among three quantities that is not simply based on their definitions. The validity of the second law is completely based on experimental verification.

## Units of Force

$F_{net} = ma$ is used to define the units of force in terms of the three basic units for mass, length, and time. The SI unit of force is called the newton (abbreviated N) and is the force needed to accelerate a 1-kg system at the rate of $1m/s^2$ That is, since $F_{net} = ma$,

$$1N = 1kg \cdot s^2 \tag{3.4.6}$$

While almost the entire world uses the newton for the unit of force, in the United States the most familiar unit of force is the pound (lb), where 1 N = 0.225 lb.

## Weight and the Gravitational Force

When an object is dropped, it accelerates toward the center of Earth. Newton's second law states that a net force on an object is responsible for its acceleration. If air resistance is negligible, the net force on a falling object is the gravitational force, commonly called its **weight** $w$. Weight can be denoted as a vector $w$ because it has a direction; down is, by definition, the direction of gravity, and hence weight is a downward force. The magnitude of weight is denoted as $w$. Galileo was instrumental in showing that, in the absence of air resistance, all objects fall with the same acceleration $w$. Using Galileo's result and Newton's second law, we can derive an equation for weight.

Consider an object with mass $m$ falling downward toward Earth. It experiences only the downward force of gravity, which has magnitude $w$. Newton's second law states that the magnitude of the net external force on an object is $F_{net} = ma$.

Since the object experiences only the downward force of gravity, $F_{net} = w$. We know that the acceleration of an object due to gravity is $g$, or $a = g$. Substituting these into Newton's second law gives

### WEIGHT

This is the equation for weight - the gravitational force on mass $m$:

$$w = mg \tag{3.4.7}$$

Since weight $g = 9.80m/s^2$ on Earth, the weight of a 1.0 kg object on Earth is 9.8 N, as we see:

$$w = mg = (1.0kg)(9.8m/s^2) = 9.8N. \tag{3.4.8}$$

Recall that $g$ can take a positive or negative value, depending on the positive direction in the coordinate system. Be sure to take this into consideration when solving problems with weight.

When the net external force on an object is its weight, we say that it is in **free-fall**. That is, the only force acting on the object is the force of gravity. In the real world, when objects fall downward toward Earth, they are never truly in free-fall because there is always some upward force from the air acting on the object.

The acceleration due to gravity $g$ varies slightly over the surface of Earth, so that the weight of an object depends on location and is not an intrinsic property of the object. Weight varies dramatically if one leaves Earth's surface. On the Moon, for example, the acceleration due to gravity is only $1.67 m/s^2$. A 1.0-kg mass thus has a weight of 9.8 N on Earth and only about 1.7 N on the Moon.

The broadest definition of weight in this sense is that *the weight of an object is the gravitational force on it from the nearest large body*, such as Earth, the Moon, the Sun, and so on. This is the most common and useful definition of weight in physics. It differs dramatically, however, from the definition of weight used by NASA and the popular media in relation to space travel and exploration. When they speak of "weightlessness" and "microgravity," they are really referring to the phenomenon we call "free-fall" in physics. We shall use the above definition of weight, and we will make careful distinctions between free-fall and actual weightlessness.

It is important to be aware that weight and mass are very different physical quantities, although they are closely related. Mass is the quantity of matter (how much "stuff") and does not vary in classical physics, whereas weight is the gravitational force and does vary depending on gravity. It is tempting to equate the two, since most of our examples take place on Earth, where the weight of an object only varies a little with the location of the object. Furthermore, the terms *mass* and *weight* are used interchangeably in everyday language; for example, our medical records often show our "weight" in kilograms, but never in the correct units of newtons.

### COMMON MISCONCEPTIONS: MASS VS. WEIGHT

Mass and weight are often used interchangeably in everyday language. However, in science, these terms are distinctly different from one another. Mass is a measure of how much matter is in an object. The typical measure of mass is the kilogram (or the "slug" in English units). Weight, on the other hand, is a measure of the force of gravity acting on an object. Weight is equal to the mass of an object $(m)$ multiplied by the acceleration due to gravity $(g)$. Like any other force, weight is measured in terms of newtons (or pounds in English units).

Assuming the mass of an object is kept intact, it will remain the same, regardless of its location. However, because weight depends on the acceleration due to gravity, the weight of an object can change when the object enters into a region with stronger or weaker gravity. For example, the acceleration due to gravity on the Moon is $1.67 m/s^2$ (which is much less than the acceleration due to gravity on Earth, $9.80 m/s^2$ ). If you measured your weight on Earth and then measured your weight on the Moon, you would find that you "weigh" much less, even though you do not look any skinnier. This is because the force of gravity is weaker on the Moon. In fact, when people say that they are "losing weight," they really mean that they are losing "mass" (which in turn causes them to weigh less)

### TAKE-HOME EXPERIMENT: MASS AND WEIGHT

What do bathroom scales measure? When you stand on a bathroom scale, what happens to the scale? It depresses slightly. The scale contains springs that compress in proportion to your weight—similar to rubber bands expanding when pulled. The springs provide a measure of your weight (for an object which is not accelerating). This is a force in newtons (or pounds). In most countries, the measurement is divided by 9.80 to give a reading in mass units of kilograms. The scale measures weight but is calibrated to provide information about mass. While standing on a bathroom scale, push down on a table next to you. What happens to the reading? Why? Would your scale measure the same "mass" on Earth as on the Moon?

### Example 3.4.1:What Acceleration Can a Person Produce when pushing a Lawn Mower?

Suppose that the net external force (push minus friction) exerted on a lawn mower is 51 N (about 11 lb) parallel to the ground. The mass of the mower is 24 kg. What is its acceleration?




Figure 3.4.3: The net force on a lawn mower is 51 N to the right. At what rate does the lawn mower accelerate to the right?

**Strategy**

Since $F_{net}$ and $m$ are given, the acceleration can be calculated directly from Newton's second law as stated in $F_{net} = ma$ .

**Solution**

The magnitude of the acceleration $a$ is $a = \frac{F_{net}}{m}$. Entering known values gives

$$a = \frac{51\ N}{24\ kg} \tag{3.4.9}$$

Substituting the units $kg \cdot m/s^2$ for N yields

$$a = \frac{51\ kg/s^2}{24\ kg} = 2.1\ m/s^2 \tag{3.4.10}$$

**Discussion**

The direction of the acceleration is the same direction as that of the net force, which is parallel to the ground. There is no information given in this example about the individual external forces acting on the system, but we can say something about their relative magnitudes. For example, the force exerted by the person pushing the mower must be greater than the friction opposing the motion (since we know the mower moves forward), and the vertical forces must cancel if there is to be no acceleration in the vertical direction (the mower is moving only horizontally). The acceleration found is small enough to be reasonable for a person pushing a mower. Such an effort would not last too long because the person's top speed would soon be reached.

## Example 3.4.2:What Rocket Thrust Accelerates This Sled?

Prior to manned space flights, rocket sleds were used to test aircraft, missile equipment, and physiological effects on human subjects at high speeds. They consisted of a platform that was mounted on one or two rails and propelled by several rockets. Calculate the magnitude of force exerted by each rocket, called its thrust $T$ for the four-rocket propulsion system shown in Figure. The sled's initial acceleration is $49m/s^2$ the mass of the system is 2100 kg, and the force of friction opposing the motion is known to be 650 N.

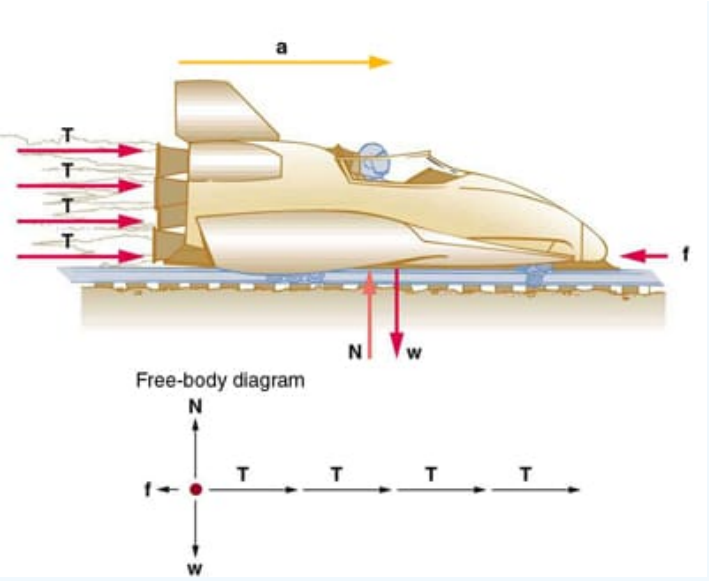



Figure 4.4.4. A sled experiences a rocket thrust that accelerates it to the right. Each rocket creates an identical thrust $T$. As in other situations where there is only horizontal acceleration, the vertical forces cancel. The ground exerts an upward force $N$ on the system that is equal in magnitude and opposite in direction to its weight, $w$. The system here is the sled, its rockets, and rider, so none of the forces between these objects are considered. The arrow representing friction $(f)$ is drawn larger than scale.

**Strategy**

Although there are forces acting vertically and horizontally, we assume the vertical forces cancel since there is no vertical acceleration. This leaves us with only horizontal forces and a simpler one-dimensional problem. Directions are indicated with plus or minus signs, with right taken as the positive direction. See the free-body diagram in the figure.

**Solution**

Since acceleration, mass, and the force of friction are given, we start with Newton's second law and look for ways to find the thrust of the engines. Since we have defined the direction of the force and acceleration as acting "to the right," we need to consider only the magnitudes of these quantities in the calculations. Hence we begin with

$$F_{net} = ma. \tag{3.4.11}$$

, where $F_{net}$ is the net force along the horizontal direction. We can see from Figure that the engine thrusts add, while friction opposes the thrust. In equation form, the net external force is

$$F_{net} = 4T - f. \tag{3.4.12}$$

Substituting this into Newton's second law gives

$$F_{net} = ma = 4T - f. \tag{3.4.13}$$

Using a little algebra, we solve for the total thrust 4T:

$$4T = ma + f. \tag{3.4.14}$$

Substituting known values yields

$$4T = ma + f = (2100\,kg)(49\,m/s^2) + 650\,N \tag{3.4.15}$$

So the total thrust is

$$1 \times 10^5 N, \tag{3.4.16}$$

and the individual thrusts are

$$T = \frac{1 \times 10^5}{4} = 2.6 \times 10^4\,N \tag{3.4.17}$$

**Discussion**

The numbers are quite large, so the result might surprise you. Experiments such as this were performed in the early 1960s to test the limits of human endurance and the setup designed to protect human subjects in jet fighter emergency ejections. Speeds of 1000 km/h were obtained, with accelerations of 45 $g$-s. (Recall that $g$, the acceleration due gravity is $9.80\ m/s^2$. When we say that an acceleration is 45 $g$-s, it is $45 \times 9.80 m/s^2$, which is approximately $440 m/s^2$ ). While living subjects are not used any more, land speeds of 10,000 km/h have been obtained with rocket sleds. In this example, as in the preceding one, the system of interest is obvious. We will see in later examples that choosing the system of interest is crucial—and the choice is not always obvious.

Newton's second law of motion is more than a definition; it is a relationship among acceleration, force, and mass. It can help us make predictions. Each of those physical quantities can be defined independently, so the second law tells us something basic and universal about nature. The next section introduces the third and final law of motion.

Summary

- Acceleration, $a$, is defined as a change in velocity, meaning a change in its magnitude or direction, or both.
- An external force is one acting on a system from outside the system, as opposed to internal forces, which act between components within the system.
- Newton's second law of motion states that the acceleration of a system is directly proportional to and in the same direction as the net external force acting on the system, and inversely proportional to its mass.
- In equation form, Newton's second law of motion is $a = \frac{F_{net}}{m}$
- This is often written in the more familiar form: $F_{net} = ma$.
- The weight $w$ of an object is defined as the force of gravity acting on an object of mass $m$. The object experiences an acceleration due to gravity $g$:

$$w = mg. \tag{3.4.18}$$

- If the only force acting on an object is due to gravity, the object is in free fall.
- Friction is a force that opposes the motion past each other of objects that are touching.

This page titled 3.4: Newton's Second Law of Motion- Concept of a System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 4.3: Newton's Second Law of Motion- Concept of a System by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3.5: Newton's Third Law of Motion- Symmetry in Forces

## Learning Objectives

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

- Understand Newton's third law of motion.
- Apply Newton's third law to define systems and solve problems of motion.

There is a passage in the musical *Man of la Mancha* that relates to Newton's third law of motion. Sancho, in describing a fight with his wife to Don Quixote, says, "Of course I hit her back, Your Grace, but she's a lot harder than me and you know what they say, 'Whether the stone hits the pitcher or the pitcher hits the stone, it's going to be bad for the pitcher.'" This is exactly what happens whenever one body exerts a force on another—the first also experiences a force (equal in magnitude and opposite in direction). Numerous common experiences, such as stubbing a toe or throwing a ball, confirm this. It is precisely stated in **Newton's third law of motion.**

## NEWTON'S THIRD LAW OF MOTION

Whenever one body exerts a force on a second body, the first body experiences a force that is equal in magnitude and opposite in direction to the force that it exerts.

This law represents a certain *symmetry in nature*: Forces always occur in pairs, and one body cannot exert a force on another without experiencing a force itself. We sometimes refer to this law loosely as "action-reaction," where the force exerted is the action and the force experienced as a consequence is the reaction. Newton's third law has practical uses in analyzing the origin of forces and understanding which forces are external to a system.

We can readily see Newton's third law at work by taking a look at how people move about. Consider a swimmer pushing off from the side of a pool, as illustrated in Figure. She pushes against the pool wall with her feet and accelerates in the direction *opposite* to that of her push. The wall has exerted an equal and opposite force back on the swimmer. You might think that two equal and opposite forces would cancel, but they do not *because they act on different systems*. In this case, there are two systems that we could investigate: the swimmer or the wall. If we select the swimmer to be the system of interest, as in the figure, then $F_{wall\ on\ feet}$ is an external force on this system and affects its motion. The swimmer moves in the direction of $F_{wall\ on\ feet}$. In contrast, the force $F_{feet\ on\ wall}$ acts on the wall and not on our system of interest. Thus $F_{feet\ on\ wall}$ does not directly affect the motion of the system and does not cancel $F_{wall\ on\ feet}$. Note that the swimmer pushes in the direction opposite to that in which she wishes to move. The reaction to her push is thus in the desired direction.

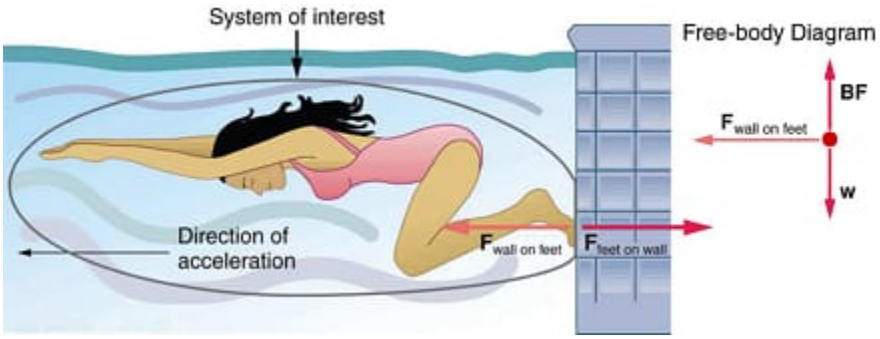



Figure 3.5.1: When the swimmer exerts a force $F_{feet\ on\ wall}$ on the wall, she accelerates in the direction opposite to that of her push. This means the net external force on her is in the direction opposite to $F_{feet\ on\ wall}$ This opposition occurs because, in accordance with Newton's third law of motion, the wall exerts a force $F_{wall\ on\ feet}$ on her, equal in magnitude but in the direction opposite to the one she exerts on it. The line around the swimmer indicates the system of interest. Note that $F_{feet\ on\ wall}$ does not act on this system (the swimmer) and, thus, does not cancel $F_{wall\ on\ feet}$. Thus the free-body diagram shows only $F_{wall\ on\ feet}$, $w$, the gravitational force, and $BF$, the buoyant force of the water supporting the swimmer's weight. The vertical forces $w$ and $BF$ cancel since there is no vertical motion.

Other examples of Newton's third law are easy to find. As a professor paces in front of a whiteboard, she exerts a force backward on the floor. The floor exerts a reaction force forward on the professor that causes her to accelerate forward. Similarly, a car accelerates because the ground pushes forward on the drive wheels in reaction to the drive wheels pushing backward on the ground. You can see evidence of the wheels pushing backward when tires spin on a gravel road and throw rocks backward. In another example, rockets move forward by expelling gas backward at high velocity. This means the rocket exerts a large backward force on the gas in the rocket combustion chamber, and the gas therefore exerts a large reaction force forward on the rocket. This reaction force is called thrust. It is a common misconception that rockets propel themselves by pushing on the ground or on the air behind them. They actually work better in a vacuum, where they can more readily expel the exhaust gases. Helicopters similarly create lift by pushing air down, thereby experiencing an upward reaction force. Birds and airplanes also fly by exerting force on air in a direction opposite to that of whatever force they need. For example, the wings of a bird force air downward and backward in order to get lift and move forward. An octopus propels itself in the water by ejecting water through a funnel from its body, similar to a jet ski. In a situation similar to Sancho's, professional cage fighters experience reaction forces when they punch, sometimes breaking their hand by hitting an opponent's body.

## Exercise 3.5.1: Getting up to speed: Choosing the Correct System

A physics professor pushes a cart of demonstration equipment to a lecture hall, as seen in Figure. Her mass is 65.0 kg, the cart's is 12.0 kg, and the equipment's is 7.0 kg. Calculate the acceleration produced when the professor exerts a backward force of 150 N on the floor. All forces opposing the motion, such as friction on the cart's wheels and air resistance, total 24.0 N.

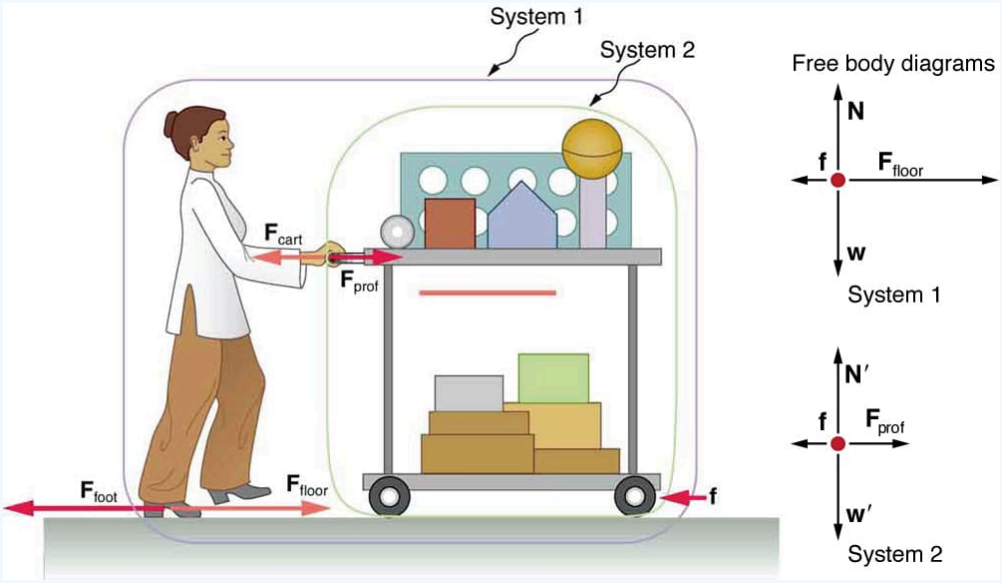



Figure 3.5.2: A professor pushes a cart of demonstration equipment. The lengths of the arrows are proportional to the magnitudes of the forces (except for $f$, since it is too small to draw to scale). Different questions are asked in each example; thus, the system of interest must be defined differently for each. System 1 is appropriate for Example, since it asks for the acceleration of the entire group of objects. Only $F_{floor}$ and $f$ are external forces acting on System 1 along the line of motion. All other forces either cancel or act on the outside world. System 2 is chosen for this example so that $F_{prof}$ will be an external force and enter into Newton's second law. Note that the free-body diagrams, which allow us to apply Newton's second law, vary with the system chosen.

**Strategy**

Since they accelerate as a unit, we define the system to be the professor, cart, and equipment. This is System 1 in Figure. The professor pushes backward with a force $F_{foot}$ of 150 N. According to Newton's third law, the floor exerts a forward reaction force $F_{floor}$ of 150 N on System 1. Because all motion is horizontal, we can assume there is no net force in the vertical direction. The problem is therefore one-dimensional along the horizontal direction. As noted, $f$ opposes the motion and is thus in the opposite direction of $F_{floor}$. Note that we do not include the forces $F_{prof}$ or $F_{cart}$ because these are internal forces, and we do not include $F_{foot}$ because it acts on the floor, not on the system. There are no other significant forces acting on System 1. If the net external force can be found from all this information, we can use Newton's second law to find the acceleration as requested. See the free-body diagram in the figure.

**Solution**

Newton's second law is given by

$$a = \frac{F_{net}}{m}.$$

The net external force on System 1 is deduced from Figure and the discussion above to be

$$F_{net} = F_{floor} - f = 150\,N - 24.0\,N = 126\,N.$$

The mass of System 1 is

$$m = (65.0 + 12.0 + 7.0) = 84\,kg.$$

These values of $F_{net}$ and $m$ produce an acceleration of

$$a = \frac{F_{net}}{m}$$

$$a = \frac{126\,N}{84\,kg} = 1.5\,m/s^2.$$

**Discussion**

None of the forces between components of System 1, such as between the professor's hands and the cart, contribute to the net external force because they are internal to System 1. Another way to look at this is to note that forces between components of a system cancel because they are equal in magnitude and opposite in direction. For example, the force exerted by the professor on the cart results in an equal and opposite force back on her. In this case both forces act on the same system and, therefore, cancel. Thus internal forces (between components of a system) cancel. Choosing System 1 was crucial to solving this problem.

## Example 3.5.2: Force on the Cart: Choosing a New System

Calculate the force the professor exerts on the cart in Figure using data from the previous example if needed.

**Strategy**

If we now define the system of interest to be the cart plus equipment (System 2 in Figure), then the net external force on System 2 is the force the professor exerts on the cart minus friction. The force she exerts on the cart, $F_{prof}$ is an external force acting on System 2. $F_{prof}$ was internal to System 1, but it is external to System 2 and will enter Newton's second law for System 2.

**Solution**

Newton's second law can be used to find $F_{prof}$ Starting with

$$a = \frac{F_{net}}{m}$$

and noting that the magnitude of the net external force on System 2 is

$$F_{net} = F_{prof} - f,$$

we solve for $F_{prof}$, the desired quantity

$$F_{net} + f.$$

The value of $f$ is given, so we must calculate net $F_{net}$. That can be done since both the acceleration and mass of System 2 are known. Using Newton's second law we see that

$$F_{net} = ma,$$

where the mass of System 2 is 19.0 kg ( m = 12.0 kg + 7.0 kg) and its acceleration was found to be $a = 1.5\ m/s^2$ in the previous example. Thus,

$$F_{net} = ma$$

$$F_{net} = (19.0\,kg)(1.5\,m/s^2) = 29\,N.$$

Now we can find the desired force:

$$F_{prof} = F_{net} + f,$$

$$F_{prof} = 29\,N + 24.0\,N = 53\,N.$$

**Discussion**

It is interesting that this force is significantly less than the 150-N force the professor exerted backward on the floor. Not all of that 150-N force is transmitted to the cart; some of it accelerates the professor.

The choice of a system is an important analytical step both in solving problems and in thoroughly understanding the physics of the situation (which is not necessarily the same thing).

PHET EXPLORATIONS: GRAVITY FORCE LAB

Visualize the gravitational force that two objects exert on each other. Change properties of the objects in order to see how it changes the gravity force.

# Section Summary

- **Newton's third law of motion** represents a basic symmetry in nature. It states: Whenever one body exerts a force on a second body, the first body experiences a force that is equal in magnitude and opposite in direction to the force that the first body exerts.
- A **thrust** is a reaction force that pushes a body forward in response to a backward force. Rockets, airplanes, and cars are pushed forward by a thrust reaction force.

## Glossary

**Newton's third law of motion**
: whenever one body exerts a force on a second body, the first body experiences a force that is equal in magnitude and opposite in direction to the force that the first body exerts

**thrust**
: a reaction force that pushes a body forward in response to a backward force; rockets, airplanes, and cars are pushed forward by a thrust reaction force

This page titled 3.5: Newton's Third Law of Motion- Symmetry in Forces is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 4.4: Newton's Third Law of Motion- Symmetry in Forces by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 3.6: Summary of Newton's Laws

Understanding Newton's laws of motion has been very important for our society in that this understanding has led to numerous technological advances. Think of all of the human constructed moving objects in this world. Each of these were designed and created by scientists and engineers with a deep understanding of Newton's Laws: airplanes, boats, cars, cannons, elevators, rockets, satellites, space shuttles, trains, trucks, and many more. Some of these vehicles we even trust with our lives! And it is not just the moving objects that require an understanding of Newton's laws of motion, Newton's laws are also important in order to understand how to build stationary objects that will stand the test of time. Bridges, buildings, fences, roads, swimming pools, towers and many more constructions can only be created to be reliable with a proper understanding of Newton's laws. Newton's laws are fundamental for creating the world in which we live. We will take a minute now to review these laws. Every example we have seen so far in this chapter can be solved with a proper understanding of Newton's laws of motion.

**NEWTON'S FIRST LAW OF MOTION**

A body at rest remains at rest, or, if in motion, remains in motion at a constant velocity unless acted on by a net external force.

Rather than contradicting our experience, **Newton's first law of motion** states that there must be a *cause* (which is a net external force) *for there to be any change in velocity (either a change in magnitude or direction)*. Please review earlier portions of this chapter for more on Newton's first law of motion.

As it turns out, the acceleration of an object depends only on the net external force and the mass of the object. Combining the two proportionalities just given yields Newton's second law of motion.

NEWTON'S SECOND LAW OF MOTION

$$\boldsymbol{a} = \frac{\boldsymbol{F}_{\text{net}}}{m}$$

This is often written in the more familiar form

$$\boldsymbol{F}_{\text{net}} = m\boldsymbol{a},$$

with the vector notation indicating that the net external force is in the same direction as acceleration. When only the magnitude of force and acceleration are considered, this equation is simply (note the lack of vector notations)

$$F_{\text{net}} = ma.$$

The acceleration of a system is directly proportional to and in the same direction as the net external force acting on the system, and inversely proportional to its mass. This was also discussed previously in this chapter.

Whenever one body exerts a force on another—the first also experiences a force (equal in magnitude and opposite in direction). Numerous common experiences, such as stubbing a toe or throwing a ball, confirm this. It is precisely stated in **Newton's third law of motion**.

NEWTON'S THIRD LAW OF MOTION

Whenever one body exerts a force on a second body, the first body experiences a force by the second body that is equal in magnitude and opposite in direction to the force that it itself exerts.

This law represents a certain *symmetry in nature*: Forces always occur in pairs, and one body cannot exert a force on another without experiencing a force itself. We sometimes refer to this law loosely as "action-reaction," where the force exerted is the action and the force experienced as a consequence is the reaction. Newton's third law has practical uses in analyzing the origin

of forces and understanding which forces are external to a system. Again, this is review from material previously covered.

This page titled 3.6: Summary of Newton's Laws is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Jamie MacArthur.

- 3.3: Newton's First Law of Motion- Inertia by Andrew Park is licensed CC BY 4.0.
- 3.4: Newton's Second Law of Motion- Force and Acceleration by Andrew Park is licensed CC BY 4.0.
- 3.5: Newton's Third Law of Motion- Symmetry in Forces by Andrew Park is licensed CC BY 4.0.

# 3.7: Newton's Universal Law of Gravitation

## Learning Objectives

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

- Explain Earth's gravitational force.
- Describe the gravitational effect of the Moon on Earth.
- Discuss weightlessness in space.
- Examine the Cavendish experiment

What do aching feet, a falling apple, and the orbit of the Moon have in common? Each is caused by the gravitational force. Our feet are strained by supporting our weight—the force of Earth's gravity on us. An apple falls from a tree because of the same force acting a few meters above Earth's surface. And the Moon orbits Earth because gravity is able to supply the necessary centripetal force at a distance of hundreds of millions of meters. In fact, the same force causes planets to orbit the Sun, stars to orbit the center of the galaxy, and galaxies to cluster together. Gravity is another example of underlying simplicity in nature. It is the weakest of the four basic forces found in nature, and in some ways the least understood. It is a force that acts at a distance, without physical contact, and is expressed by a formula that is valid everywhere in the universe, for masses and distances that vary from the tiny to the immense.

Sir Isaac Newton was the first scientist to precisely define the gravitational force, and to show that it could explain both falling bodies and astronomical motions. See Figure. But Newton was not the first to suspect that the same force caused both our weight and the motion of planets. His forerunner Galileo Galilei had contended that falling bodies and planetary motions had the same cause. Some of Newton's contemporaries, such as Robert Hooke, Christopher Wren, and Edmund Halley, had also made some progress toward understanding gravitation. But Newton was the first to propose an exact mathematical form and to use that form to show that the motion of heavenly bodies should be conic sections—circles, ellipses, parabolas, and hyperbolas. This theoretical prediction was a major triumph—it had been known for some time that moons, planets, and comets follow such paths, but no one had been able to propose a mechanism that caused them to follow these paths and not others.



Figure 3.7.1: According to early accounts, Newton was inspired to make the connection between falling bodies and astronomical motions when he saw an apple fall from a tree and realized that if the gravitational force could extend above the ground to a tree, it might also reach the Sun. The inspiration of Newton's apple is a part of worldwide folklore and may even be based in fact. Great importance is attached to it because Newton's universal law of gravitation and his laws of motion answered very old questions about nature and gave tremendous support to the notion of underlying simplicity and unity in nature. Scientists still expect underlying simplicity to emerge from their ongoing inquiries into nature.

The gravitational force is relatively simple. It is always attractive, and it depends only on the masses involved and the distance

between them. Stated in modern language, Newton's universal law of gravitation states that every particle in the universe attracts every other particle with a force along a line joining them. The force is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

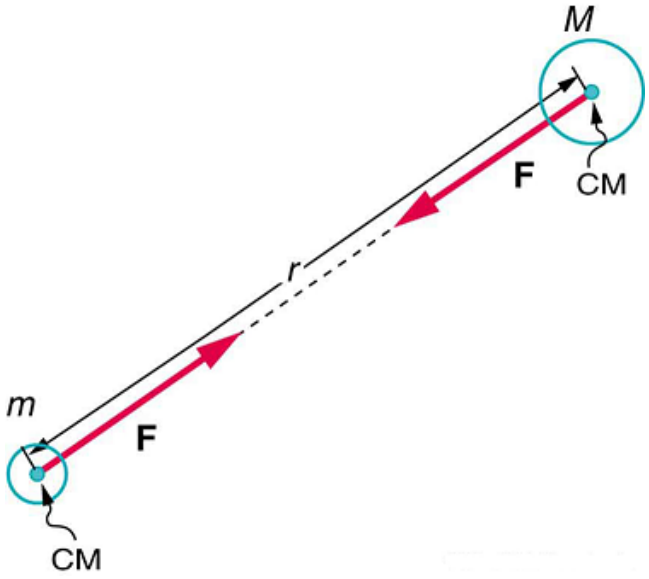



Figure 3.7.2: Gravitational attraction is along a line joining the centers of mass of these two bodies. The magnitude of the force is the same on each, consistent with Newton's third law.

MISCONCEPT ALERT

The magnitude of the force on each object (one has larger mass than the other) is the same, consistent with Newton's third law.

The bodies we are dealing with tend to be large. To simplify the situation we assume that the body acts as if its entire mass is concentrated at one specific point called the **center of mass** (CM), which will be further explored in Linear Momentum and Collisions. For two bodies having masses $m$ and $M$ with a distance $r$ between their centers of mass, the equation for Newton's universal law of gravitation is

$$F = G\frac{mM}{r^2}, \tag{3.7.1}$$

where $F$ is the magnitude of the gravitational force and $G$ is a proportionality factor called the gravitational constant. $G$ is a universal gravitational constant—that is, it is thought to be the same everywhere in the universe. It has been measured experimentally to be

$$G = 6.673 \times 10^{-11}\frac{N \cdot m^2}{kg^2} \tag{3.7.2}$$

in SI units. Note that the units of $G$ are such that a force in newtons is obtained from $F = G\frac{mM}{r^2}$, when considering masses in kilograms and distance in meters. For example, two 1.000 kg masses separated by 1.000 m will experience a gravitational attraction of $6.673 \times 10^{-11}\ N$.

This is an extraordinarily small force. The small magnitude of the gravitational force is consistent with everyday experience. We are unaware that even large objects like mountains exert gravitational forces on us. In fact, our body weight is the force of attraction of the *entire Earth* on us with a mass of $6 \times 10^{24}\ kg$.

Recall that the acceleration due to gravity $g$ is about $9.80\ m/s^2$ on Earth. We can now determine why this is so. The weight of an object *mg* is the gravitational force between it and Earth. Substituting *mg* for $F$ in Newton's universal law of gravitation gives

$$mg = G\frac{mM}{r^2}, \tag{3.7.3}$$

where $m$ is the mass of the object, $M$ is the mass of Earth, and $r$ is the distance to the center of Earth (the distance between the

centers of mass of the object and Earth). See Figure. The mass $m$ of the object cancels, leaving an equation for $g$:

$$g = G\frac{M}{r^2}. \tag{3.7.4}$$

Substituting known values for Earth's mass and radius (to three significant figures),

$$g = \left(6.673 \times 10^{-11}\,\frac{N \cdot m^2}{kg^2}\right) \times \frac{5.98 \times 10^{24}\,kg}{(6.38 \times 10^{6}\,m)^2}, \tag{3.7.5}$$

and we obtain a value for the acceleration of a falling body:

$$g = 9.80\,m/s^2. \tag{3.7.6}$$

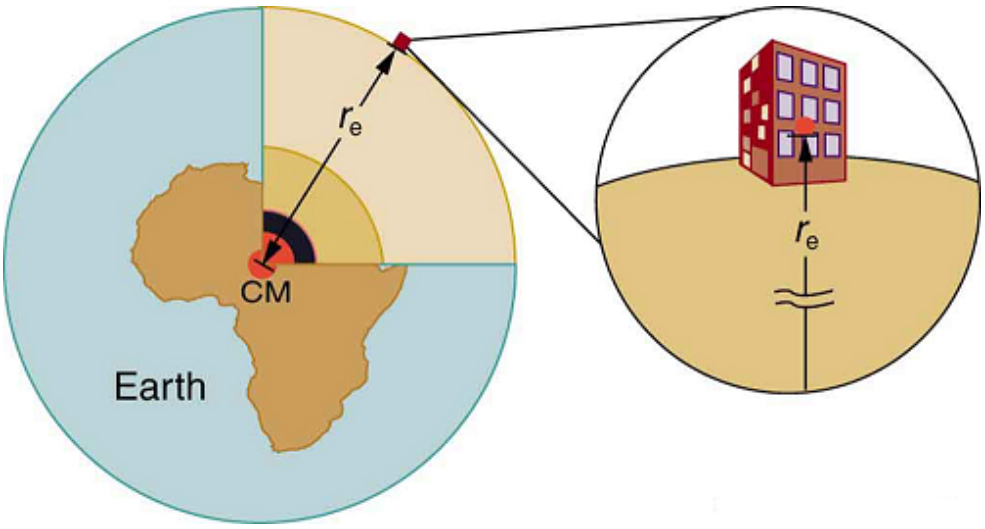



Figure 3.7.3: The distance between the centers of mass of Earth and an object on its surface is very nearly the same as the radius of Earth, because Earth is so much larger than the object.

This is the expected value *and is independent of the body's mass*. Newton's law of gravitation takes Galileo's observation that all masses fall with the same acceleration a step further, explaining the observation in terms of a force that causes objects to fall—in fact, in terms of a universally existing force of attraction between masses.

## TAKE HOME EXPERIMENT

Take a marble, a ball, and a spoon and drop them from the same height. Do they hit the floor at the same time? If you drop a piece of paper as well, does it behave like the other objects? Explain your observations.

## MAKING CONNECTIONS

Attempts are still being made to understand the gravitational force. As we shall see in Particle Physics, modern physics is exploring the connections of gravity to other forces, space, and time. General relativity alters our view of gravitation, leading us to think of gravitation as bending space and time.

In the following example, we make a comparison similar to one made by Newton himself. He noted that if the gravitational force caused the Moon to orbit Earth, then the acceleration due to gravity should equal the centripetal acceleration of the Moon in its orbit. Newton found that the two accelerations agreed "pretty nearly."

## Example 3.7.1: Earth's Gravitational Force Is the Centripetal Force Making the Moon Move in a Curved Path

a. Find the acceleration due to Earth's gravity at the distance of the Moon.
b. Calculate the centripetal acceleration needed to keep the Moon in its orbit (assuming a circular orbit about a fixed Earth), and compare it with the value of the acceleration due to Earth's gravity that you have just found.

**Strategy for (a)**

This calculation is the same as the one finding the acceleration due to gravity at Earth's surface, except that $r$ is the distance from the center of Earth to the center of the Moon. The radius of the Moon's nearly circular orbit is $3.84 \times 10^8\ m$.

**Solution for (a)**

Substituting known values into the expression for $g$ found above, remembering that $M$ is the mass of Earth not the Moon, yields

$$g = G\frac{M}{r^2} = \left(6.67 \times 10^{-11}\frac{N \cdot m^2}{kg^2}\right) \times \frac{5.98 \times 10^{24}\,kg}{3.84 \times 10^8\,m)^2} \tag{3.7.7}$$

$$= 2.70 \times 10^{-3}\,m/s^2. \tag{3.7.8}$$

**Strategy for (b)**

Centripetal acceleration can be calculated using either form of

$$a_c = \frac{v^2}{r} \tag{3.7.9}$$

$$a_c = r\omega^2 \tag{3.7.10}$$

We choose to use the second form:

$$a_c = r\omega^2, \tag{3.7.11}$$

where $\omega$ is the angular velocity of the Moon about Earth.

**Solution for (b)**

Given that the period (the time it takes to make one complete rotation) of the Moon's orbit is 27.3 days, (d) and using

$$1\,d \times 24\frac{hr}{d} \times 60\frac{min}{hr} \times 60\frac{s}{min} = 86,400\,s \tag{3.7.12}$$

we see that

$$a_c = r\omega^2 = (3.84 \times 10^8\,m)(2.66 \times 10^{-6}\,rad/s^2) \tag{3.7.13}$$

$$= 2.72 \times 10^{-3}\,m/s^2. \tag{3.7.14}$$

The direction of the acceleration is toward the center of the Earth.

**Discussion**

The centripetal acceleration of the Moon found in (b) differs by less than 1% from the acceleration due to Earth's gravity found in (a). This agreement is approximate because the Moon's orbit is slightly elliptical, and Earth is not stationary (rather the Earth-Moon system rotates about its center of mass, which is located some 1700 km below Earth's surface). The clear implication is that Earth's gravitational force causes the Moon to orbit Earth.

Why does Earth not remain stationary as the Moon orbits it? This is because, as expected from Newton's third law, if Earth

exerts a force on the Moon, then the Moon should exert an equal and opposite force on Earth (see Figure). We do not sense the Moon's effect on Earth's motion, because the Moon's gravity moves our bodies right along with Earth but there are other signs on Earth that clearly show the effect of the Moon's gravitational force as discussed in Satellites and Kepler's Laws: An Argument for Simplicity.

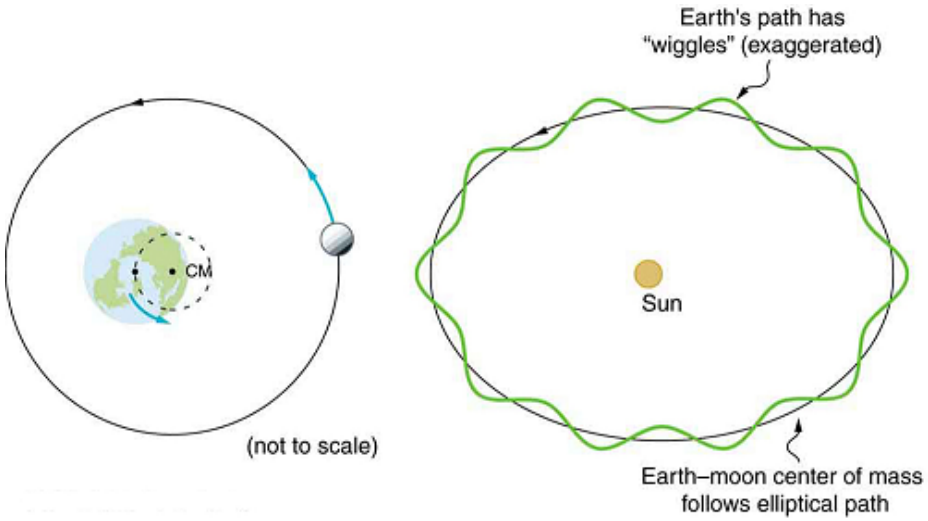



Figure 3.7.4: (a) Earth and the Moon rotate approximately once a month around their common center of mass. (b) Their center of mass orbits the Sun in an elliptical orbit, but Earth's path around the Sun has "wiggles" in it. Similar wiggles in the paths of stars have been observed and are considered direct evidence of planets orbiting those stars. This is important because the planets' reflected light is often too dim to be observed.

## Tides

Ocean tides are one very observable result of the Moon's gravity acting on Earth. Figure is a simplified drawing of the Moon's position relative to the tides. Because water easily flows on Earth's surface, a high tide is created on the side of Earth nearest to the Moon, where the Moon's gravitational pull is strongest. Why is there also a high tide on the opposite side of Earth? The answer is that Earth is pulled toward the Moon more than the water on the far side, because Earth is closer to the Moon. So the water on the side of Earth closest to the Moon is pulled away from Earth, and Earth is pulled away from water on the far side. As Earth rotates, the tidal bulge (an effect of the tidal forces between an orbiting natural satellite and the primary planet that it orbits) keeps its orientation with the Moon. Thus there are two tides per day (the actual tidal period is about 12 hours and 25.2 minutes), because the Moon moves in its orbit each day as well).

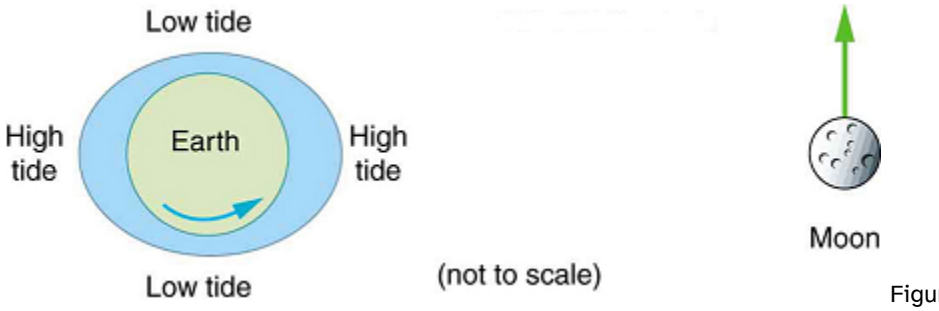



Figure 3.7.5: The Moon causes ocean tides by attracting the water on the near side more than Earth, and by attracting Earth more than the water on the far side. The distances and sizes are not to scale. For this simplified representation of the Earth-Moon system, there are two high and two low tides per day at any location, because Earth rotates under the tidal bulge.

The Sun also affects tides, although it has about half the effect of the Moon. However, the largest tides, called spring tides, occur when Earth, the Moon, and the Sun are aligned. The smallest tides, called neap tides, occur when the Sun is at a $90^{o}$ angle to the Earth-Moon alignment.

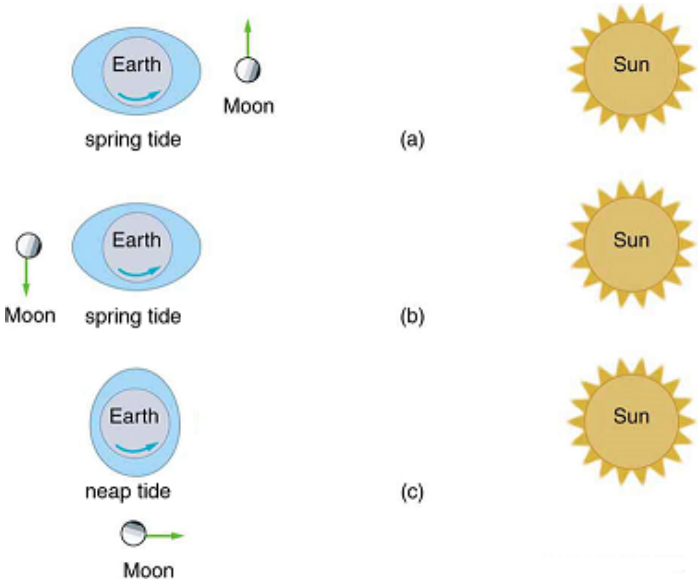



Figure 3.7.6: (a, b) Spring tides: The highest tides occur when Earth, the Moon, and the Sun are aligned. (c) Neap tide: The lowest tides occur when the Sun lies at $90^o$ to the Earth-Moon alignment. Note that this figure is not drawn to scale.

Tides are not unique to Earth but occur in many astronomical systems. The most extreme tides occur where the gravitational force is the strongest and varies most rapidly, such as near black holes (see Figure). A few likely candidates for black holes have been observed in our galaxy. These have masses greater than the Sun but have diameters only a few kilometers across. The tidal forces near them are so great that they can actually tear matter from a companion star.

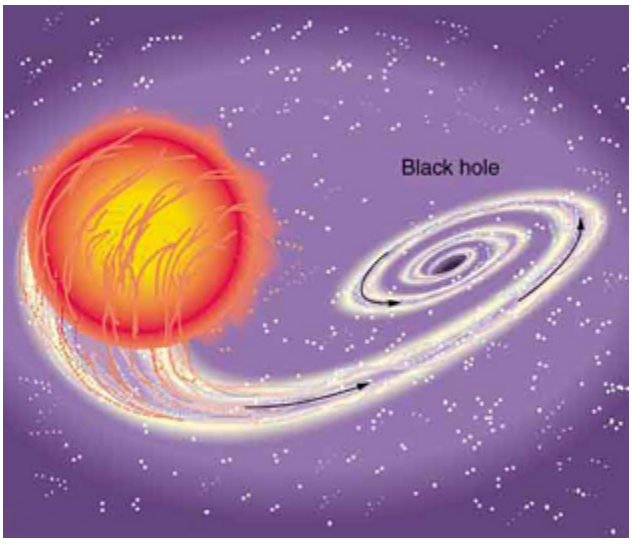



Figure 3.7.7: A black hole is an object with such strong gravity that not even light can escape it. This black hole was created by the supernova of one star in a two-star system. The tidal forces created by the black hole are so great that it tears matter from the companion star. This matter is compressed and heated as it is sucked into the black hole, creating light and X-rays observable from Earth.

## ”Weightlessness” and Microgravity

In contrast to the tremendous gravitational force near black holes is the apparent gravitational field experienced by astronauts orbiting Earth. What is the effect of “weightlessness” upon an astronaut who is in orbit for months? Or what about the effect of weightlessness upon plant growth? Weightlessness doesn’t mean that an astronaut is not being acted upon by the gravitational force. There is no “zero gravity” in an astronaut’s orbit. The term just means that the astronaut is in free-fall, accelerating with the acceleration due to gravity. If an elevator cable breaks, the passengers inside will be in free fall and will experience weightlessness. You can experience short periods of weightlessness in some rides in amusement parks.



Figure 3.7.8: Astronauts experiencing weightlessness on board the International Space Station. (credit: NASA).

Microgravity refers to an environment in which the apparent net acceleration of a body is small compared with that produced by Earth at its surface. Many interesting biology and physics topics have been studied over the past three decades in the presence of microgravity. Of immediate concern is the effect on astronauts of extended times in outer space, such as at the International Space Station. Researchers have observed that muscles will atrophy (waste away) in this environment. There is also a corresponding loss of bone mass. Study continues on cardiovascular adaptation to space flight. On Earth, blood pressure is usually higher in the feet than in the head, because the higher column of blood exerts a downward force on it, due to gravity. When standing, 70% of your blood is below the level of the heart, while in a horizontal position, just the opposite occurs. What difference does the absence of this pressure differential have upon the heart?

Some findings in human physiology in space can be clinically important to the management of diseases back on Earth. On a somewhat negative note, spaceflight is known to affect the human immune system, possibly making the crew members more vulnerable to infectious diseases. Experiments flown in space also have shown that some bacteria grow faster in microgravity than they do on Earth. However, on a positive note, studies indicate that microbial antibiotic production can increase by a factor of two in space-grown cultures. One hopes to be able to understand these mechanisms so that similar successes can be achieved on the ground. In another area of physics space research, inorganic crystals and protein crystals have been grown in outer space that have much higher quality than any grown on Earth, so crystallography studies on their structure can yield much better results.

Plants have evolved with the stimulus of gravity and with gravity sensors. Roots grow downward and shoots grow upward. Plants might be able to provide a life support system for long duration space missions by regenerating the atmosphere, purifying water, and producing food. Some studies have indicated that plant growth and development are not affected by gravity, but there is still uncertainty about structural changes in plants grown in a microgravity environment.

## The Cavendish Experiment: Then and Now

As previously noted, the universal gravitational constant $G$ is determined experimentally. This definition was first done accurately by Henry Cavendish (1731–1810), an English scientist, in 1798, more than 100 years after Newton published his universal law of gravitation. The measurement of $G$ is very basic and important because it determines the strength of one of the four forces in nature. Cavendish's experiment was very difficult because he measured the tiny gravitational attraction between two ordinary-sized masses (tens of kilograms at most), using apparatus like that in Figure. Remarkably, his value for $G$ differs by less than 1% from the best modern value.One important consequence of knowing $G$ was that an accurate value for Earth's mass could finally be obtained. This was done by measuring the acceleration due to gravity as accurately as possible and then calculating the mass of Earth $M$ from the relationship Newton's universal law of gravitation gives

$$mg = G\frac{mM}{r^2}, \tag{3.7.15}$$

where $m$ is the mass of the object, $M$ is the mass of Earth, and $r$ is the distance to the center of Earth (the distance between the

centers of mass of the object and Earth). See Figure. The mass $m$ of the object cancels, leaving an equation for $g$:

$$g = G\frac{M}{r^2}. \tag{3.7.16}$$

Rearranging to solve for $M$ yields

$$M = \frac{gr^2}{G}. \tag{3.7.17}$$

so $M$ can be calculated because all quantities on the right, including the radius of Earth $r$, are known from direct measurements. We shall see in Satellites and Kepler's Laws: An Argument for Simplicity that knowing $G$ also allows for the determination of astronomical masses. Interestingly, of all the fundamental constants in physics, $G$ is by far the least well determined.

The Cavendish experiment is also used to explore other aspects of gravity. One of the most interesting questions is whether the gravitational force depends on substance as well as mass—for example, whether one kilogram of lead exerts the same gravitational pull as one kilogram of water. A Hungarian scientist named Roland von Eötvös pioneered this inquiry early in the 20th century. He found, with an accuracy of five parts per billion, that the gravitational force does not depend on the substance. Such experiments continue today, and have improved upon Eötvös' measurements. Cavendish-type experiments such as those of Eric Adelberger and others at the University of Washington, have also put severe limits on the possibility of a fifth force and have verified a major prediction of general relativity—that gravitational energy contributes to rest mass. Ongoing measurements there use a torsion balance and a parallel plate (not spheres, as Cavendish used) to examine how Newton's law of gravitation works over sub-millimeter distances. On this small-scale, do gravitational effects depart from the inverse square law? So far, no deviation has been observed.

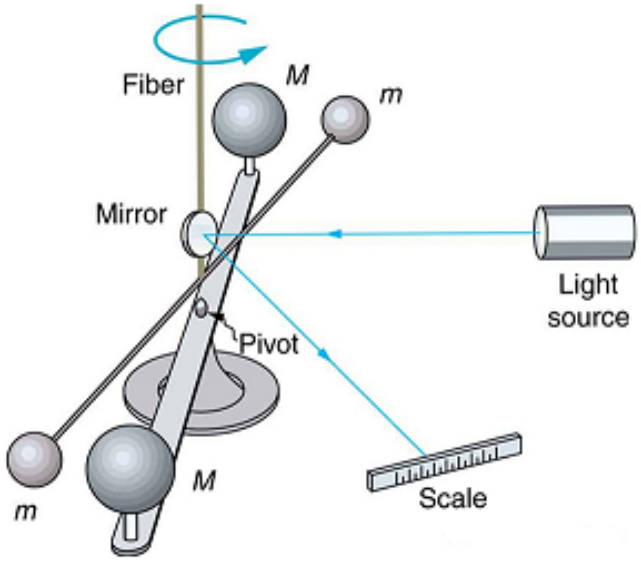



Figure 3.7.9: Cavendish used an apparatus like this to measure the gravitational attraction between the two suspended spheres $(m)$ and the two on the stand $(M)$ by observing the amount of torsion (twisting) created in the fiber. Distance between the masses can be varied to check the dependence of the force on distance. Modern experiments of this type continue to explore gravity.

## Summary

- Newton's universal law of gravitation: Every particle in the universe attracts every other particle with a force along a line joining them. The force is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. In equation form, this is

$$F = G\frac{mM}{r^2} \tag{3.7.18}$$

  where F is the magnitude of the gravitational force. $G$ is the gravitational constant, given by $G = 6.63 \times 10^{-11}\ N \cdot m^2/kg^2$.

- Newton's law of gravitation applies universally.

# Glossary

**gravitational constant, *G***
: a proportionality factor used in the equation for Newton's universal law of gravitation; it is a universal constant—that is, it is thought to be the same everywhere in the universe

**center of mass**
: the point where the entire mass of an object can be thought to be concentrated

**microgravity**
: an environment in which the apparent net acceleration of a body is small compared with that produced by Earth at its surface

**Newton's universal law of gravitation**
: every particle in the universe attracts every other particle with a force along a line joining them; the force is directly proportional to the product of their masses and inversely proportional to the square of the distance between them

This page titled 3.7: Newton's Universal Law of Gravitation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

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

# 4: MODULE 3-ENERGY AND MOMENTUM


4.1: Prelude to Work, Energy, and Energy Resources

4.2: Work- The Scientific Definition

4.3: Kinetic Energy and the Work-Energy Theorem

4.4: Gravitational Potential Energy

4.5: Conservative Forces and Potential Energy

4.6: Conservation of Energy

4.7: Power

4.8: Work, Energy, and Power in Humans

4.9: World Energy Use

4.10: Prelude to Linear Momentum and Collisions

4.11: Linear Momentum and Force

4.12: Conservation of Momentum

4.13: Energy and Momentum



This page titled 4: Module 3-Energy and Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

# 4.1: Prelude to Work, Energy, and Energy Resources

*Energy* plays an essential role both in everyday events and in scientific phenomena. You can no doubt name many forms of energy, from that provided by our foods, to the energy we use to run our cars, to the sunlight that warms us on the beach. You can also cite examples of what people call energy that may not be scientific, such as someone having an energetic personality. Not only does energy have many interesting forms, it is involved in almost all phenomena, and is one of the most important concepts of physics. What makes it even more important is that the total amount of energy in the universe is constant. Energy can change forms, but it cannot appear from nothing or disappear without a trace. Energy is thus one of a handful of physical quantities that we say is *conserved*.



Figure 4.1.1: How many forms of energy can you identify in this photograph of a wind farm in Iowa? (credit: Jürgen from Sandesneben, Germany, Wikimedia Commons)

**Conservation of energy** (as physicists like to call the principle that energy can neither be created nor destroyed) is based on experiment. Even as scientists discovered new forms of energy, conservation of energy has always been found to apply. Perhaps the most dramatic example of this was supplied by Einstein when he suggested that mass is equivalent to energy (his famous equation $E = mc^2$.

From a societal viewpoint, energy is one of the major building blocks of modern civilization. Energy resources are key limiting factors to economic growth. The world use of energy resources, especially oil, continues to grow, with ominous consequences economically, socially, politically, and environmentally. We will briefly examine the world's energy use patterns at the end of this chapter.

There is no simple, yet accurate, scientific definition for energy. Energy is characterized by its many forms and the fact that it is conserved. We can loosely define energy as the ability to do work, admitting that in some circumstances not all **energy** is available to do work. Because of the association of energy with work, we begin the chapter with a discussion of work. Work is intimately related to energy and how energy moves from one system to another or changes form.

This page titled 4.1: Prelude to Work, Energy, and Energy Resources is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.0: Prelude to Work, Energy, and Energy Resources by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.2: Work- The Scientific Definition

**Learning Objectives**

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

- Explain how an object must be displaced for a force on it to do work.
- Explain how relative directions of force and displacement determine whether the work done is positive, negative, or zero.

## What It Means to Do Work

The scientific definition of work differs in some ways from its everyday meaning. Certain things we think of as hard work, such as writing an exam or carrying a heavy load on level ground, are not work as defined by a scientist. The scientific definition of work reveals its relationship to energy—whenever work is done, energy is transferred. For work, in the scientific sense, to be done, a force must be exerted and there must be motion or displacement in the direction of the force.

Formally, the work done on a system by a constant force is defined to be *the product of the component of the force in the direction of motion times the distance through which the force acts*. For one-way motion in one dimension, this is expressed in equation form as

$$W = |\vec{F}| \cos\theta |\vec{d}| \tag{4.2.1}$$

where $W$ is work, $d$ is the displacement of the system, and $\theta$ is the angle between the force vector $\vec{F}$ and the displacement vector $\vec{d}$, as in Figure 4.2.1. We can also write Equation 4.2.1 as

$$W = F\,d\,\cos\theta \tag{4.2.2}$$

To find the work done on a system that undergoes motion that is not one-way or that is in two or three dimensions, we divide the motion into one-way one-dimensional segments and add up the work done over each segment.

**What is Work?**

The work done on a system by a constant force is *the product of the component of the force in the direction of motion times the distance through which the force acts*. For one-way motion in one dimension, this is expressed in equation form as

$$W = F\,d\,\cos\theta \tag{4.2.3}$$

where $W$ is work, $F$ is the magnitude of the force on the system, $d$ is the magnitude of the displacement of the system, and $\theta$ is the angle between the force vector $F$ d the displacement vector $d$.

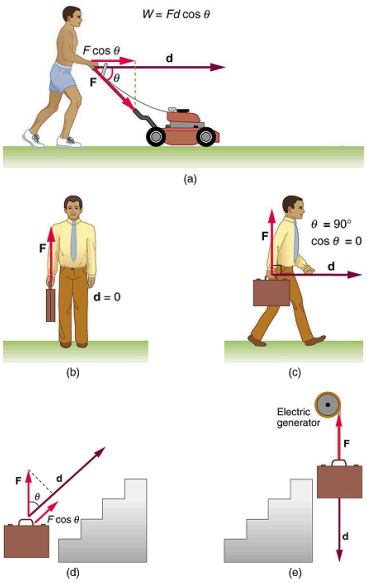



Figure $4.2.1$: Examples of work. (a) The work done by the force $F$ on this lawn mower is $Fd\ cos\ \theta$. Note that $F\ cos\ \theta$ is the component of the force in the direction of motion. (b) A person holding a briefcase does no work on it, because there is no motion. No energy is transferred to or from the briefcase. (c) The person moving the briefcase horizontally at a constant speed does no work on it, and transfers no energy to it. (d) Work is done on the briefcase by carrying it upstairs at constant speed, because there is necessarily a component of force $F$ in the direction of the motion. Energy is transferred to the briefcase and could in turn be used to do work. (e) When the briefcase is lowered, energy is transferred out of the briefcase and into an electric generator. Here the work done on the briefcase by the generator is negative, removing energy from the briefcase, because $F$ and $d$ are in opposite directions.

To examine what the definition of work means, let us consider the other situations shown in Figure. The person holding the briefcase in Figure $4.2.1b$does no work, for example. Here $d = 0$, so $W = 0$. Why is it you get tired just holding a load? The answer is that your muscles are doing work against one another, *but they are doing no work on the system of interest* (the "briefcase-Earth system" - see Gravitational Potential Energy for more details). There must be motion for work to be done, and there must be a component of the force in the direction of the motion. For example, the person carrying the briefcase on level ground in Figure $4.2.1c$ does no work on it, because the force is perpendicular to the motion. That is, $\cos\ 90^o = 0$, so $W = 0$.

In contrast, when a force exerted on the system has a component in the direction of motion, such as in Figure $4.2.1d$, work *is* done—energy is transferred to the briefcase. Finally, in Figure $4.2.1e$, energy is transferred from the briefcase to a generator. There are two good ways to interpret this energy transfer. One interpretation is that the briefcase's weight does work on the generator, giving it energy. The other interpretation is that the generator does negative work on the briefcase, thus removing energy from it. The drawing shows the latter, with the force from the generator upward on the briefcase, and the displacement downward. This makes $\theta = 180^o$, and $\cos\ 180^o = -1$, therefore $W$ is negative.

## Calculating Work

Work and energy have the same units. From the definition of work, we see that those units are force times distance. Thus, in SI units, work and energy are measured in newton-meters. A newton-meter is given the special name joule (J), and $1\ J = 1\ N \cdot m = 1\ kg\ m^2/s^2$. One joule is not a large amount of energy; it would lift a small 100-gram apple a distance of about 1 meter.

### Example $4.2.1$: Calculating the Work You Do to Push a Lawn Mower Across a Large Lawn

How much work is done on the lawn mower by the person in Figure (a) if he exerts a constant force of 75.0 N at an angle $35^o$ below the horizontal and pushes the mower 25 m. on level ground? Convert the amount of work from joules to kilocalories and compare it with this person's average daily intake of **10,000 kJ** (about **2400 kcal**) of food energy. One *calorie* (1 cal) of heat is the amount required to warm 1 g of water by $1^oC$ and is equivalent to **4,184 J**, while one *food calorie* (**1 kcal**) is equivalent to **4,184 J**.

**Strategy**

We can solve this problem by substituting the given values into the definition of work done on a system, stated in the equation $W = Fd\ cos\ \theta$. The force, angle, and displacement are given, so that only the work $W$ is unknown.

**Solution**

The equation for the work is (Equation 4.2.2):

$$W = Fd\ \cos\ \theta$$

Substituting the known values gives

$$\begin{aligned} W &= (75\,N)(25.0\,m)(cos\,35^o) \\ &= 1536\,J \\ &= 1.54 \times 10^3\,J \end{aligned}$$

Converting the work in joules to kilocalories yields $W = (1536\ \ J)(1\ \ kcal/4184\ \ J) = 0.367kcal$. The ratio of the work done to the daily consumption is

$$\frac{W}{2400\,kcal} = 1.53 \times 10^{-4}.$$

**Discussion**

This ratio is a tiny fraction of what the person consumes, but it is typical. Very little of the energy released in the consumption of food is used to do work. Even when we "work" all day long, less than 10% of our food energy intake is used to do work and more than 90% is converted to thermal energy or stored as chemical energy in fat.

## Summary

- Work is the transfer of energy by a force acting on an object as it is displaced.
- The work $W$ that a force $F$ does on an object is the product of the magnitude $F$ of the force, times the magnitude $d$ of the displacement, times the cosine of the angle $\theta$ between them. In symbols,

$$W = Fd\ \cos\ \theta. \tag{4.2.4}$$

- The SI unit for work and energy is the joule (J), where $1\ \ J = 1\ \ N \cdot m = 1\ \ kg\ \ m^2/s^2$.
- The work done by a force is zero if the displacement is either zero or perpendicular to the force.
- The work done is positive if the force and displacement have the same direction, and negative if they have opposite direction.

## Glossary

**energy**
: the ability to do work

**work**
: the transfer of energy by a force that causes an object to be displaced; the product of the component of the force in the direction of the displacement and the magnitude of the displacement

**joule**
: SI unit of work and energy, equal to one newton-meter

This page titled 4.2: Work- The Scientific Definition is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.1: Work- The Scientific Definition by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.3: Kinetic Energy and the Work-Energy Theorem

**Learning Objectives**

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

- Explain work as a transfer of energy and net work as the work done by the net force.
- Explain and apply the work-energy theorem.

## Work Transfers Energy

What happens to the work done on a system? Energy is transferred into the system, but in what form? Does it remain in the system or move on? The answers depend on the situation. For example, if the lawn mower in [link](a) is pushed just hard enough to keep it going at a constant speed, then energy put into the mower by the person is removed continuously by friction, and eventually leaves the system in the form of heat transfer. In contrast, work done on the briefcase by the person carrying it up stairs in [link](d) is stored in the briefcase-Earth system and can be recovered at any time, as shown in [link](e). In fact, the building of the pyramids in ancient Egypt is an example of storing energy in a system by doing work on the system. Some of the energy imparted to the stone blocks in lifting them during construction of the pyramids remains in the stone-Earth system and has the potential to do work.

In this section we begin the study of various types of work and forms of energy. We will find that some types of work leave the energy of a system constant, for example, whereas others change the system in some way, such as making it move. We will also develop definitions of important forms of energy, such as the energy of motion.

## Net Work and the Work-Energy Theorem

We know from the study of Newton's laws in Dynamics: Force and Newton's Laws of Motion that net force causes acceleration. We will see in this section that work done by the net force gives a system energy of motion, and in the process we will also find an expression for the energy of motion.

Let us start by considering the total, or net, work done on a system. **Net work** is defined to be the sum of work done by all external forces—that is, net work is the work done by the net external force $F_{net}$. In equation form, this is $W_{net} = F_{net} d\ cos\ \theta$, where $\theta$ is the angle between the force vector and the displacement vector. Figure (a) shows a graph of force versus displacement for the component of the force in the direction of the displacement—that is, an $F\ cos\ \theta$ vs. $d$ graph. In this case, $F\ cos\ \theta$ is constant. You can see that the area under the graph is $F\ cos\ \theta$, or the work done. Figure (b) shows a more general process where the force varies. The area under the curve is divided into strips, each having an average force $(F\,cos\,\theta)_{i(ave)}$. The work done is e $(F\,cos\,\theta)_{i(ave)} d_i$ for each strip, and the total work done is the sum of the $W_i$. Thus the total work done is the total area under the curve, a useful property to which we shall refer later.

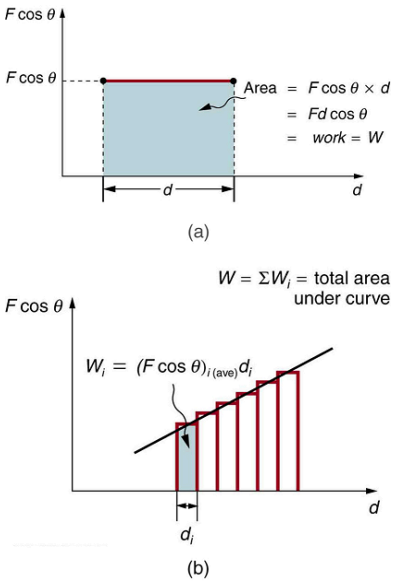



Figure 4.3.1: (a) A graph of $F\ cos\ \theta$ vs. $d$ when $F\ cos\ \theta$ is constant. The area under the curve represents the work done by the force. (b) A graph of $F\ cos\ \theta$ vs. $d$ in which the force varies. The work done for each interval is the area of each strip; thus, the total area under the curve equals the total work done.

Net work will be simpler to examine if we consider a one-dimensional situation where a force is used to accelerate an object in a direction parallel to its initial velocity. Such a situation occurs for the package on the roller belt conveyor system shown in Figure.

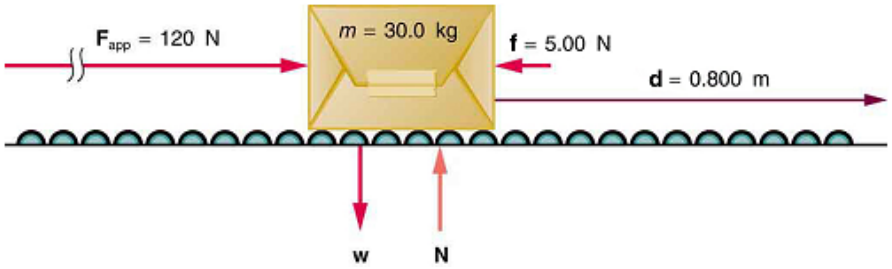



Figure 4.3.2: A package on a roller belt is pushed horizontally through a distance $d$.

The force of gravity and the normal force acting on the package are perpendicular to the displacement and do no work. Moreover, they are also equal in magnitude and opposite in direction so they cancel in calculating the net force. The net force arises solely from the horizontal applied force $F_{app}$ and the horizontal friction force $f$. Thus, as expected, the net force is parallel to the displacement, so that $\theta = 0$ and $cos\ \theta = 1$, and the net work is given by

$$W_{net} = F_{net}d. \tag{4.3.1}$$

The effect of the net force $F_{net}$ is to accelerate the package from $v_0$ to $v$ The kinetic energy of the package increases, indicating that the net work done on the system is positive. (See Example.) By using Newton's second law, and doing some algebra, we can reach an interesting conclusion. Substituting $F = ma$ from Newton's second law gives

$$W_{net} = mad. \tag{4.3.2}$$

To get a relationship between net work and the speed given to a system by the net force acting on it, we take $d = x - x_0$ and use the equation studied in Motion Equations for Constant Acceleration in One Dimension for the change in speed over a distance $d$ if the acceleration has the constant value $a$, namely $v^2 = v_0^2 + 2ad$. (note that $a$ appears in the expression for the net work). Solving for acceleration gives $a = \frac{v^2 - v_0^2}{2d}$. When $a$ is substituted into the preceding expression for $W_{net}$ we obtain

$$W_{net} = m\left(\frac{v^2 - v_0^2}{2d}\right)d. \tag{4.3.3}$$

The $d$ cancels, and we rearrange this to obtain

$$W_{net} = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2. \tag{4.3.4}$$

This expression is called the **work-energy theorem,** and it actually applies *in general* (even for forces that vary in direction and magnitude), although we have derived it for the special case of a constant force parallel to the displacement. The theorem implies that the net work on a system equals the change in the quantity $\frac{1}{2}mv^2$. This quantity is our first example of a form of energy.

### Work-Energy Theorem

The net work on a system equals the change in the quantity $\frac{1}{2}mv^2$.

$$W_{net} = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2. \tag{4.3.5}$$

The quantity $\frac{1}{2}mv^2$ in the work-energy theorem is defined to be the translational **kinetic energy** (KE) of a mass $m$ moving at a speed $v$. (*Translational* kinetic energy is distinct from *rotational* kinetic energy, which is considered later.) In equation form, the translational kinetic energy,

$$KE = \frac{1}{2}mv^2, \tag{4.3.6}$$

is the energy associated with translational motion. Kinetic energy is a form of energy associated with the motion of a particle, single body, or system of objects moving together.

We are aware that it takes energy to get an object, like a car or the package in Figure, up to speed, but it may be a bit surprising that kinetic energy is proportional to speed squared. This proportionality means, for example, that a car traveling at 100 km/h has four times the kinetic energy it has at 50 km/h, helping to explain why high-speed collisions are so devastating. We will now consider a series of examples to illustrate various aspects of work and energy.

### Example4.3.1: Calculating the Kinetic Energy of a Package

Suppose a 30.0-kg package on the roller belt conveyor system in Figure 7.03.2 is moving at 0.500 m/s. What is its kinetic energy?

**Strategy**

Because the mass $m$ and the speed $v$ are given, the kinetic energy can be calculated from its definition as given in the equation $KE = \frac{1}{2}mv^2$.

**Solution**

The kinetic energy is given by

$$KE = \frac{1}{2}mv^2. \tag{4.3.7}$$

Entering known values gives

$$KE = 0.5(30.0\,kg)(0.500\,m/s)^2, \tag{4.3.8}$$

which yields

$$KE = 3.75\,kg \cdot m^2/s^2 = 3.75\,J \tag{4.3.9}$$

**Discussion**

Note that the unit of kinetic energy is the joule, the same as the unit of work, as mentioned when work was first defined. It is also interesting that, although this is a fairly massive package, its kinetic energy is not large at this relatively low speed. This fact is consistent with the observation that people can move packages like this without exhausting themselves.

## Example4.3.2: Determining the Work to Accelerate a Package

Suppose that you push on the 30.0-kg package in Figure 7.03.2. with a constant force of 120 N through a distance of 0.800 m, and that the opposing friction force averages 5.00 N.

(a) Calculate the net work done on the package. (b) Solve the same problem as in part (a), this time by finding the work done by each force that contributes to the net force.

**Strategy and Concept for (a)**

This is a motion in one dimension problem, because the downward force (from the weight of the package) and the normal force have equal magnitude and opposite direction, so that they cancel in calculating the net force, while the applied force, friction, and the displacement are all horizontal. (See Figure 7.03.2.) As expected, the net work is the net force times distance.

**Solution for (a)**

The net force is the push force minus friction, or $F_{net} = 120\ N - 5.00\ N = 115\ N$. Thus the net work is

$$W_{net} = F_{net}d = (115\,N)(0.800\,m) \tag{4.3.10}$$

$$= 92.0\,N \cdot m = 92.0\,J \tag{4.3.11}$$

**Discussion for (a)**

This value is the net work done on the package. The person actually does more work than this, because friction opposes the motion. Friction does negative work and removes some of the energy the person expends and converts it to thermal energy. The net work equals the sum of the work done by each individual force.

**Strategy and Concept for (b)**

The forces acting on the package are gravity, the normal force, the force of friction, and the applied force. The normal force and force of gravity are each perpendicular to the displacement, and therefore do no work.

**Solution for (b)**

The applied force does work.

$$W_{app} = F_{app}d\,cos\,(0^o) = F_{app}d \tag{4.3.12}$$

$$= (120\,N)(0.800\,m) \tag{4.3.13}$$

$$= 96.0\,J \tag{4.3.14}$$

The friction force and displacement are in opposite directions, so that $\theta = 180^o$, and the work done by friction is

$$W_{fr} = F_{fr}d\,cos\,(180^o) \tag{4.3.15}$$

$$= -(5.00\,N)(0.800\,m) \tag{4.3.16}$$

$$= -4.00\,J \tag{4.3.17}$$

So the amounts of work done by gravity, by the normal force, by the applied force, and by friction are, respectively,

$$W_{gr} = 0, \tag{4.3.18}$$

$$W_N = 0, \tag{4.3.19}$$

$$W_{app} = 96.0\,J, \tag{4.3.20}$$

$$W_{fr} = -4\,J. \tag{4.3.21}$$

The total work done as the sum of the work done by each force is then seen to be

$$W_{total} = W_{gr} + W_N + W_{app} + W_{fr} = 92.0\,J. \tag{4.3.22}$$

**Discussion for (b)**

The calculated total work $W_{total}$ as the sum of the work by each force agrees, as expected, with the work $W_{net}$ done by the net force. The work done by a collection of forces acting on an object can be calculated by either approach.

## Example 4.3.3: Determining Speed from Work and Energy

Find the speed of the package in Figure 7.03.2. at the end of the push, using work and energy concepts.

**Strategy**

Here the work-energy theorem can be used, because we have just calculated the net work $W_{net}$ and the initial kinetic energy, $\frac{1}{2}mv_0^2$ These calculations allow us to find the final kinetic energy, $\frac{1}{2}mv^2$ and thus the final speed $v$.

**Solution**

The work-energy theorem in equation form is

$$W_{net} = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2. \tag{4.3.23}$$

Solving for $\frac{1}{2}mv^2$ gives

$$\frac{1}{2}mv^2 = W_{net} + \frac{1}{2}mv_0^2 \tag{4.3.24}$$

Thus,

$$\frac{1}{2}mv^2 = 92.0\,J + 3.75\,J = 95.75\,J. \tag{4.3.25}$$

Solving for the final speed as requested and entering known values gives

$$v = \sqrt{\frac{2(95.75\,J)}{m}} = \sqrt{\frac{191.5\,kg \cdot m^2/s^2}{30.0\,kg}} \tag{4.3.26}$$

$$= 2.53\,m/s \tag{4.3.27}$$

**Discussion**

Using work and energy, we not only arrive at an answer, we see that the final kinetic energy is the sum of the initial kinetic energy and the net work done on the package. This means that the work indeed adds to the energy of the package.

## Example 4.3.4: Work and Energy Can Reveal Distance, Too

How far does the package in Figure 7.03.2. coast after the push, assuming friction remains constant? Use work and energy considerations.

**Strategy**

We know that once the person stops pushing, friction will bring the package to rest. In terms of energy, friction does negative work until it has removed all of the package's kinetic energy. The work done by friction is the force of friction times the distance traveled times the cosine of the angle between the friction force and displacement; hence, this gives us a way of finding the distance traveled after the person stops pushing.

**Solution**

The normal force and force of gravity cancel in calculating the net force. The horizontal friction force is then the net force, and it acts opposite to the displacement, so $\theta = 180^o$. To reduce the kinetic energy of the package to zero, the work $W_{fr}$ by friction must be minus the kinetic energy that the package started with plus what the package accumulated due to the pushing. Thus $W_{fr} = -95.75\ J$. Furthermore, $W_{fr} = df'\ cos\ \theta = -Fd'$, where $d'$ is the distance it takes to stop. Thus,

$$d' = -\frac{W_{fr}}{f} = \frac{-95.75\,J}{5.00\,N}, \tag{4.3.28}$$

and so

$$d' = 19.2\,m \tag{4.3.29}$$

**Discussion**

This is a reasonable distance for a package to coast on a relatively friction-free conveyor system. Note that the work done by friction is negative (the force is in the opposite direction of motion), so it removes the kinetic energy.

Some of the examples in this section can be solved without considering energy, but at the expense of missing out on gaining insights about what work and energy are doing in this situation. On the whole, solutions involving energy are generally shorter and easier than those using kinematics and dynamics alone.

## Summary

- The net work $W_{net}$ is the work done by the net force acting on an object.
- Work done on an object transfers energy to the object.
- The translational kinetic energy of an object of mass $m$ moving at speed $v$ is $KE = \frac{1}{2}mv^2$.
- The work-energy theorem states that the net work $W_{net}$ on a system changes its kinetic energy, $W_{net} = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2$.

## Glossary

**net work**
: work done by the net force, or vector sum of all the forces, acting on an object

**work-energy theorem**
: the result, based on Newton's laws, that the net work done on an object is equal to its change in kinetic energy

**kinetic energy**
: the energy an object has by reason of its motion, equal to $\frac{1}{2}mv^2$ for the translational (i.e., non-rotational) motion of an object of mass $m$ moving at speed $v$

This page titled 4.3: Kinetic Energy and the Work-Energy Theorem is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.2: Kinetic Energy and the Work-Energy Theorem by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.4: Gravitational Potential Energy

**Learning Objectives**

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

- Explain gravitational potential energy in terms of work done against gravity.
- Show that the gravitational potential energy of an object of mass *m* at height *h* on Earth is given by $PE_g = mgh$
- Show how knowledge of the potential energy as a function of position can be used to simplify calculations and explain physical phenomena.

## Work Done Against Gravity

Climbing stairs and lifting objects is work in both the scientific and everyday sense—it is work done against the gravitational force. When there is work, there is a transformation of energy. The work done against the gravitational force goes into an important form of stored energy that we will explore in this section.

Let us calculate the work done in lifting an object of mass $m$ through a height $h$ such as in Figure. If the object is lifted straight up at constant speed, then the force needed to lift it is equal to its weight $mg$. The work done on the mass is then $W = Fd = mgh$. We define this to be the **gravitational potential energy** $(PE_g)$ put into (or gained by) the object-Earth system. This energy is associated with the state of separation between two objects that attract each other by the gravitational force. For convenience, we refer to this as the $PE_g$ gained by the object, recognizing that this is energy stored in the gravitational field of Earth. Why do we use the word “system”? Potential energy is a property of a system rather than of a single object—due to its physical position. An object’s gravitational potential is due to its position relative to the surroundings within the Earth-object system. The force applied to the object is an external force, from outside the system. When it does positive work it increases the gravitational potential energy of the system. Because gravitational potential energy depends on relative position, we need a reference level at which to set the potential energy equal to 0. We usually choose this point to be Earth’s surface, but this point is arbitrary; what is important is the *difference* in gravitational potential energy, because this difference is what relates to the work done. The difference in gravitational potential energy of an object (in the Earth-object system) between two rungs of a ladder will be the same for the first two rungs as for the last two rungs.

## Converting Between Potential Energy and Kinetic Energy

Gravitational potential energy may be converted to other forms of energy, such as kinetic energy. If we release the mass, gravitational force will do an amount of work equal to $mgh$ on it, thereby increasing its kinetic energy by that same amount (by the work-energy theorem). We will find it more useful to consider just the conversion of $PE_g$ to $KE$ without explicitly considering the intermediate step of work. (See Example $4.4.2$.) This shortcut makes it is easier to solve problems using energy (if possible) rather than explicitly using forces.

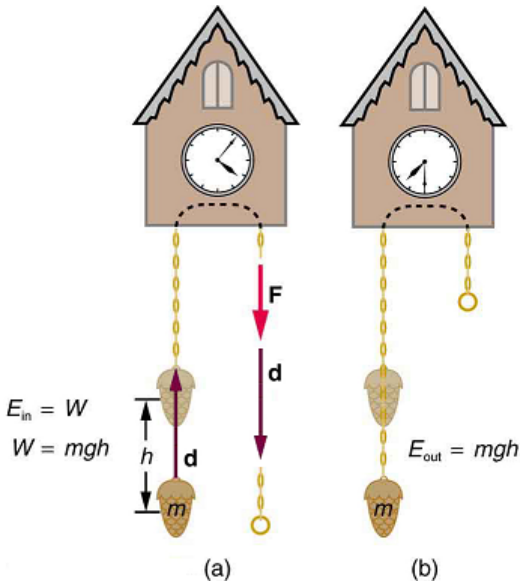



Figure 4.4.1: (a) The work done to lift the weight is stored in the mass-Earth system as gravitational potential energy. (b) As the weight moves downward, this gravitational potential energy is transferred to the cuckoo clock.

More precisely, we define the *change* in gravitational potential energy $\Delta PE_g$ to be

$$\Delta PE_g = mgh, \tag{4.4.1}$$

where, for simplicity, we denote the change in height by $h$ rather than the usual $\Delta h$. Note that $h$ is positive when the final height is greater than the initial height, and vice versa. For example, if a 0.500-kg mass hung from a cuckoo clock is raised 1.00 m, then its change in gravitational potential energy is

$$mgh = (0.500\,kg)(9.80\,m/s^2)(1.00\,m) \tag{4.4.2}$$

$$= 4.90\,kg \cdot m^2/s^2 = 4.90\,J. \tag{4.4.3}$$

Note that the units of gravitational potential energy turn out to be joules, the same as for work and other forms of energy. As the clock runs, the mass is lowered. We can think of the mass as gradually giving up its 4.90 J of gravitational potential energy, *without directly considering the force of gravity that does the work*.

## Using Potential Energy to Simplify Calculations

The equation $\Delta PE_g = mgh$ applies for any path that has a change in height of $h$, not just when the mass is lifted straight up. (See Figure.) It is much easier to calculate $mgh$ (a simple multiplication) than it is to calculate the work done along a complicated path. The idea of gravitational potential energy has the double advantage that it is very broadly applicable and it makes calculations easier. From now on, we will consider that any change in vertical position $h$ of a mass $m$ is accompanied by a change in gravitational potential energy $mgh$, and we will avoid the equivalent but more difficult task of calculating work done by or against the gravitational force.

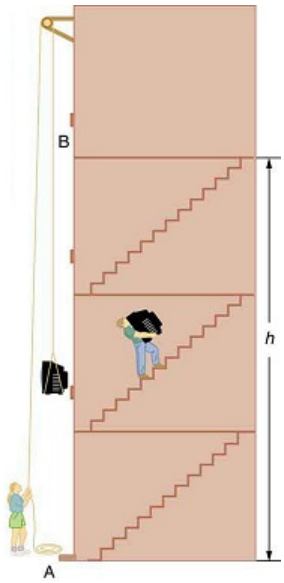



Figure 4.4.2: The change in gravitational potential energy $(\Delta PE_g)$ between points A and B is independent of the path $\Delta PE_g = mgh$ for any path between the two points. Gravity is one of a small class of forces where the work done by or against the force depends only on the starting and ending points, not on the path between them.

## Example 4.4.1: The Force to Stop Falling

A 60.0-kg person jumps onto the floor from a height of 3.00 m. If he lands stiffly (with his knee joints compressing by 0.500 cm), calculate the force on the knee joints.

**Strategy**

This person's energy is brought to zero in this situation by the work done on him by the floor as he stops. The initial $PE_g$ is transformed into $KE$ as he falls. The work done by the floor reduces this kinetic energy to zero.

**Solution**

The work done on the person by the floor as he stops is given by

$$W = Fd\cos\theta = -Fd \tag{4.4.4}$$

with a minus sign because the displacement while stopping and the force from floor are in opposite directions $(cos\ \theta = cos\ 180^o = -1.)$ The floor removes energy from the system, so it does negative work.

The kinetic energy the person has upon reaching the floor is the amount of potential energy lost by falling through height $h$:

$$KE = -\Delta PE_g = -mgh \tag{4.4.5}$$

The distance $d$ that the person's knees bend is much smaller than the height $h$ of the fall, so the additional change in gravitational potential energy during the knee bend is ignored. The work $W$ done by the floor on the person stops the person and brings the person's kinetic energy to zero:

$$W = -KE = mgh \tag{4.4.6}$$

Combining this equation with the expression for $W$ gives

$$-Fd = mgh. \tag{4.4.7}$$

Recalling that $h$ is negative because the person fell *down*, the force on the knee joints is given by

$$F = -\frac{mgh}{d} = -\frac{(60.0\,kg)(9.80\,m/s^2)(-3.00\,m)}{5.00\times10^{-3}\,m} = 3.53\times10^5\,N. \tag{4.4.8}$$

**Discussion**

Such a large force (500 times more than the person's weight) over the short impact time is enough to break bones. A much better way to cushion the shock is by bending the legs or rolling on the ground, increasing the time over which the force acts. A bending motion of 0.5 m this way yields a force 100 times smaller than in the example. A kangaroo's hopping shows this method in action. The kangaroo is the only large animal to use hopping for locomotion, but the shock in hopping is cushioned by the bending of its hind legs in each jump.(See Figure.)



Figure 4.4.3: The work done by the ground upon the kangaroo reduces its kinetic energy to zero as it lands. However, by applying the force of the ground on the hind legs over a longer distance, the impact on the bones is reduced. (credit: Chris Samuel, Flickr)

## Example 4.4.2: Finding the Speed of a Roller Coaster from its Height

(a) What is the final speed of the roller coaster shown in Figure, if it starts from rest at the top of the 20.0 m hill and work done by frictional forces is negligible? (b) What is its final speed (again assuming negligible friction) if its initial speed is 5.00 m/s?

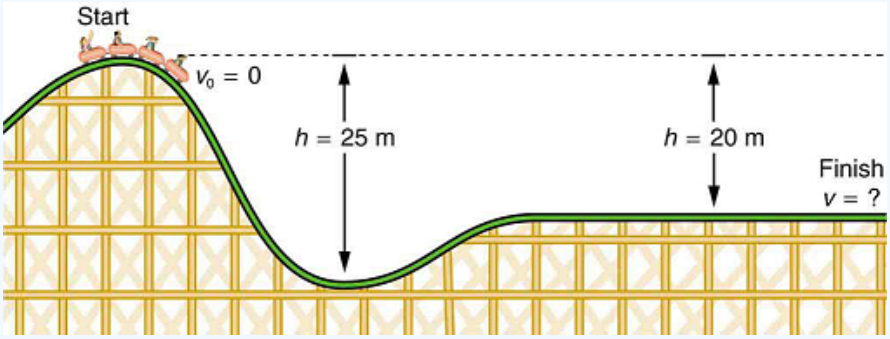



Figure 4.4.4: The speed of a roller coaster increases as gravity pulls it downhill and is greatest at its lowest point. Viewed in terms of energy, the roller-coaster-Earth system's gravitational potential energy is converted to kinetic energy. If work done by friction is negligible, all $\Delta PE_g$ is converted to $KE$.

**Strategy**

The roller coaster loses potential energy as it goes downhill. We neglect friction, so that the remaining force exerted by the track is the normal force, which is perpendicular to the direction of motion and does no work. The net work on the roller coaster is then done by gravity alone. The *loss* of gravitational potential energy from moving *downward* through a distance $h$ equals the *gain* in kinetic energy. This can be written in equation form as $-\Delta PE = \Delta KE$. Using the equations for $PE_g$ and $KE$ we can solve for the final speed $v$, which is the desired quantity.

**Solution for (a)**

Here the initial kinetic energy is zero, so that $\Delta KE = \frac{1}{2}mv^2$. The equation for change in potential energy states that $\Delta PE_g = mgh$. Since $h$ is negative in this case, we will rewrite this as $\Delta PE_g = -mg|h|$ to show the minus sign clearly. Thus,

$$-\Delta PE_g = \Delta KE \tag{4.4.9}$$

becomes

$$mg|h| = \frac{1}{2}mv^2 \tag{4.4.10}$$

Solving for $v$ we find that mass cancels and that

$$v = \sqrt{2g|h|}. \tag{4.4.11}$$

Substituting known values,

$$v = \sqrt{2(9.80\,m/s^2)(20.0\,m)} \tag{4.4.12}$$

$$= 19.8\,m/s \tag{4.4.13}$$

**Solution for (b)**

Again $-\Delta PE_g = \Delta KE$. In this case there is initial kinetic energy, so $\Delta KE = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2$. Thus,

$$mgh = \frac{1}{2}mv^2 - \frac{1}{2}m_0^2. \tag{4.4.14}$$

Rearranging gives

$$\frac{1}{2}mv^2 = mg|h| + \frac{1}{2}mv_0^2. \tag{4.4.15}$$

This means that the final kinetic energy is the sum of the initial kinetic energy and the gravitational potential energy. Mass again cancels, and

$$v = \sqrt{2g|h| + v_0^2}. \tag{4.4.16}$$

This equation is very similar to the kinematics equation $v = \sqrt{v_0^2 + 2ad}$, but it is more general—the kinematics equation is valid only for constant acceleration, whereas our equation above is valid for any path regardless of whether the object moves with a constant acceleration. Now, substituting known values gives

$$v = \sqrt{2(9.80\,m/s^2)(20.0\,m) + (5.00)^2} \tag{4.4.17}$$

$$= 20.4\,m/s. \tag{4.4.18}$$

**Discussion and Implications**

First, note that mass cancels. This is quite consistent with observations made in Falling Objects that all objects fall at the same rate if friction is negligible. Second, only the speed of the roller coaster is considered; there is no information about its direction at any point. This reveals another general truth. When friction is negligible, the speed of a falling body depends only on its initial speed and height, and not on its mass or the path taken. For example, the roller coaster will have the same final

speed whether it falls 20.0 m straight down or takes a more complicated path like the one in the figure. Third, and perhaps unexpectedly, the final speed in part (b) is greater than in part (a), but by far less than 5.00 m/s. Finally, note that speed can be found at *any* height along the way by simply using the appropriate value of $h$ at the point of interest.

We have seen that work done by or against the gravitational force depends only on the starting and ending points, and not on the path between, allowing us to define the simplifying concept of gravitational potential energy. We can do the same thing for a few other forces, and we will see that this leads to a formal definition of the law of conservation of energy.

### Making Connections: Take-Home Investigation—Converting Potential to Kinetic Energy

One can study the conversion of gravitational potential energy into kinetic energy in this experiment. On a smooth, level surface, use a ruler of the kind that has a groove running along its length and a book to make an incline (see Figure). Place a marble at the 10-cm position on the ruler and let it roll down the ruler. When it hits the level surface, measure the time it takes to roll one meter. Now place the marble at the 20-cm and the 30-cm positions and again measure the times it takes to roll 1 m on the level surface. Find the velocity of the marble on the level surface for all three positions. Plot velocity squared versus the distance traveled by the marble. What is the shape of each plot? If the shape is a straight line, the plot shows that the marble's kinetic energy at the bottom is proportional to its potential energy at the release point.

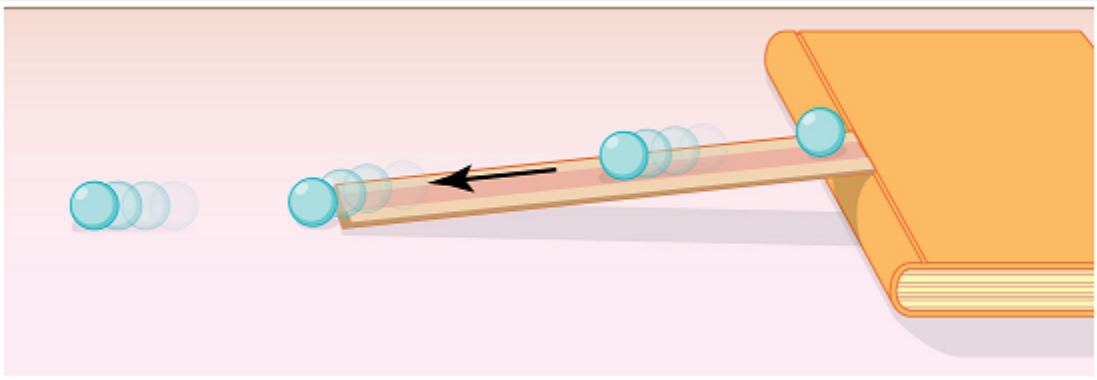


Figure 4.4.5: A marble rolls down a ruler, and its speed on the level surface is measured.

## Summary

- Work done against gravity in lifting an object becomes potential energy of the object-Earth system.
- The change in gravitational potential energy $\Delta PE_g$, is $\Delta PE_g = mgh$, with $h$ being the increase in height and $g$ the acceleration due to gravity.
- The gravitational potential energy of an object near Earth's surface is due to its position in the mass-Earth system. Only differences in gravitational potential energy, $\Delta PE_g$, have physical significance.
- As an object descends without friction, its gravitational potential energy changes into kinetic energy corresponding to increasing speed, so that $\Delta KE = -\Delta PE_g$.

## Glossary

**gravitational potential energy**
the energy an object has due to its position in a gravitational field

This page titled 4.4: Gravitational Potential Energy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.3: Gravitational Potential Energy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.5: Conservative Forces and Potential Energy

## Learning Objectives

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

- Define conservative force, potential energy, and mechanical energy.
- Explain the potential energy of a spring in terms of its compression when Hooke's law applies.
- Use the work-energy theorem to show how having only conservative forces implies conservation of mechanical energy.

Work is done by a force, and some forces, such as weight, have special characteristics. A *conservative force* is one, like the gravitational force, for which work done by or against it depends only on the starting and ending points of a motion and not on the path taken. We can define a *potential energy (PE)* for any conservative force, just as we did for the gravitational force. For example, when you wind up a toy, an egg timer, or an old-fashioned watch, you do work against its spring and store energy in it. (We treat these springs as ideal, in that we assume there is no friction and no production of thermal energy.) This stored energy is recoverable as work, and it is useful to think of it as potential energy contained in the spring. Indeed, the reason that the spring has this characteristic is that its force is *conservative*. That is, a conservative force results in stored or potential energy. Gravitational potential energy is one example, as is the energy stored in a spring. We will also see how conservative forces are related to the conservation of energy.

## Potential Energy and Conservative Forces

Potential energy is the energy a system has due to position, shape, or configuration. It is stored energy that is completely recoverable.

A conservative force is one for which work done by or against it depends only on the starting and ending points of a motion and not on the path taken.

We can define a potential energy (PE) for any conservative force. The work done against a conservative force to reach a final configuration depends on the configuration, not the path followed, and is the potential energy added.

## Potential Energy of a Spring

First, let us obtain an expression for the potential energy stored in a spring $(PE_s$. We calculate the work done to stretch or compress a spring that obeys Hooke's law. (Hooke's law was examined in Elasticity: Stress and Strain, and states that the magnitude of force $F$ on the spring and the resulting deformation $\Delta L$ are proportional, $F = k\Delta L$). (See Figure.) For our spring, we will replace $\Delta L$ (the amount of deformation produced by a force $F$) by the distance $x$ that the spring is stretched or compressed along its length. So the force needed to stretch the spring has magnitude $F = kx$, where $k$ is the spring's force constant. The force increases linearly from 0 at the start to $kx$ in the fully stretched position. The average force is $kx/2$. Thus the work done in stretching or compressing the spring is

$$W_s = Fd = (\frac{kx}{2})x = \frac{1}{2}kx^2. \tag{4.5.1}$$

Alternatively, we noted in Kinetic Energy and the Work-Energy Theorem that the area under a graph of $F$ vs. $x$ is the work done by the force. In Figure (c) we see that this area is also $\frac{1}{2}kx^2$. We therefore define the **potential energy of a spring**, \(PE_s\) to be

$$PE_s = \frac{1}{2}kx^2, \tag{4.5.2}$$

where $k$ is the spring's force constant and $x$ is the displacement from its undeformed position. The potential energy represents the work done *on* the spring and the energy stored in it as a result of stretching or compressing it a distance $x$. The potential energy of the spring $PE_s$ does not depend on the path taken; it depends only on the stretch or squeeze $x$ in the final configuration.

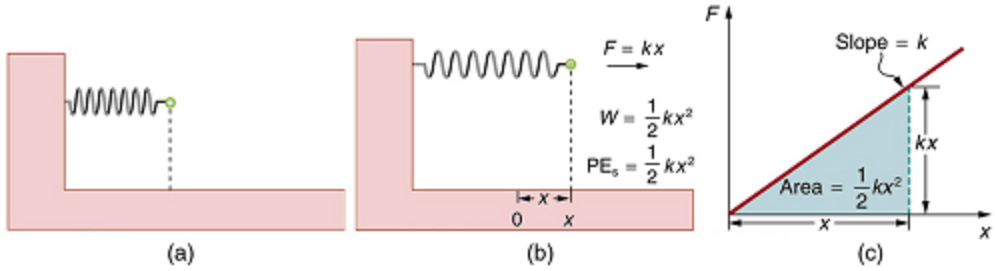



Figure 4.5.1: (a) An undeformed spring has no $PE_s$ stored in it. (b) The force needed to stretch (or compress) the spring a distance $x$ has a magnitude $F = kx$, and the work done to stretch (or compress) it is $\frac{1}{2}kx^2$. Because the force is conservative, this work is stored as potential energy $(PE_s)$ in the spring, and it can be fully recovered. (c) A graph of $F$ vs $x$ has a slope of $k$, and the area under the graph is $\frac{1}{2}kx^2$. Thus the work done or potential energy stored is $\frac{1}{2}kx^2$.

The equation $PE_s = \frac{1}{2}kx^2$ has general validity beyond the special case for which it was derived. Potential energy can be stored in any elastic medium by deforming it. Indeed, the general definition of potential energy is energy due to position, shape, or configuration. For shape or position deformations, stored energy is $PE_s = \frac{1}{2}kx^2$, where $k$ is the force constant of the particular system and $x$ is its deformation. Another example is seen in Figure for a guitar string.

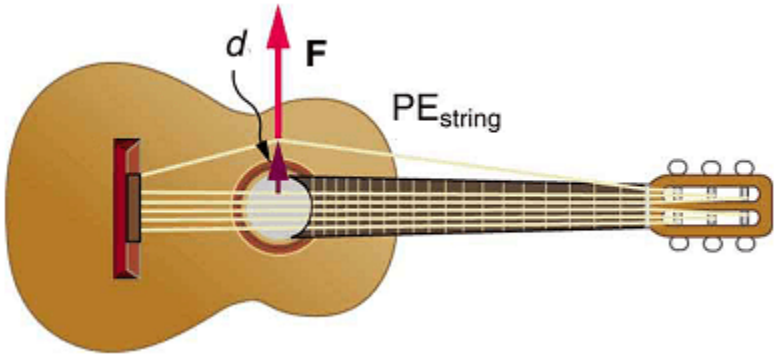



Figure 4.5.2: Work is done to deform the guitar string, giving it potential energy. When released, the potential energy is converted to kinetic energy and back to potential as the string oscillates back and forth. A very small fraction is dissipated as sound energy, slowly removing energy from the string.

## Conservation of Mechanical Energy

Let us now consider what form the work-energy theorem takes when only conservative forces are involved. This will lead us to the conservation of energy principle. The work-energy theorem states that the net work done by all forces acting on a system equals its change in kinetic energy. In equation form, this is

$$W_{net} = \frac{1}{2}mv^2 - \frac{1}{2}mv_0^2 = \Delta KE. \tag{4.5.3}$$

If only conservative forces act, then

$$W_{net} = W_c \tag{4.5.4}$$

where $W_c$ is the total work done by all conservative forces. Thus,

$$W_c = \Delta KE. \tag{4.5.5}$$

Now, if the conservative force, such as the gravitational force or a spring force, does work, the system loses potential energy. That is,

$$W_c = -\Delta PE \tag{4.5.6}$$

Therefore,

$$-\Delta PE = KE \tag{4.5.7}$$

or

$$\Delta PE + \Delta KE = 0. \tag{4.5.8}$$

This equation means that the total kinetic and potential energy is constant for any process involving only conservative forces. That is,

$$KE + PE = constant \tag{4.5.9}$$

or

$$KE_i + PE_i = KE_f + PE_f \tag{4.5.10}$$

$$(conservative\ forces\ only), \tag{4.5.11}$$

where i and f denote initial and final values. This equation is a form of the work-energy theorem for conservative forces; it is known as the **conservation of mechanical energy principle**. Remember that this applies to the extent that all the forces are conservative, so that friction is negligible. The total kinetic plus potential energy of a system is defined to be its **mechanical energy,** $(KE + PE)$. In a system that experiences only conservative forces, there is a potential energy associated with each force, and the energy only changes form between $KE$ and the various types of $PE$, with the total energy remaining constant.

## Example 4.5.1: Using Conservation of Mechanical Energy to Calculate the Speed of a Toy Car

A 0.100-kg toy car is propelled by a compressed spring, as shown in Figure. The car follows a track that rises 0.180 m above the starting point. The spring is compressed 4.00 cm and has a force constant of 250.0 N/m. Assuming work done by friction to be negligible, find (a) how fast the car is going before it starts up the slope and (b) how fast it is going at the top of the slope.

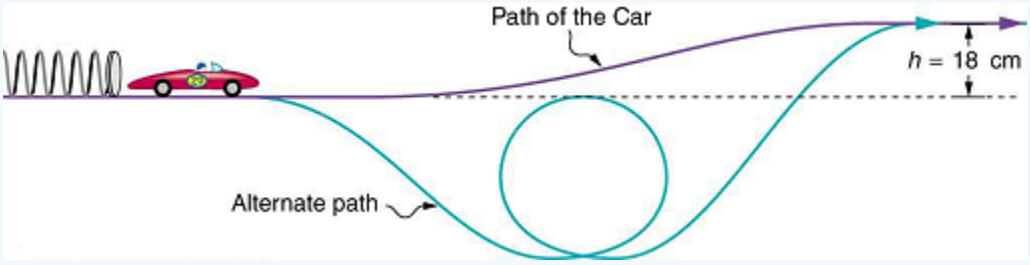



Figure 4.5.3: A toy car is pushed by a compressed spring and coasts up a slope. Assuming negligible friction, the potential energy in the spring is first completely converted to kinetic energy, and then to a combination of kinetic and gravitational potential energy as the car rises. The details of the path are unimportant because all forces are conservative—the car would have the same final speed if it took the alternate path shown.

**Strategy**

The spring force and the gravitational force are conservative forces, so conservation of mechanical energy can be used. Thus,

$$KE_i + PE_i = KE_f + PE_f \tag{4.5.12}$$

or

$$\frac{1}{2}mv_i^2 + mgh_i + \frac{1}{2}kx_i^2 = \frac{1}{2}mv_f^2 + mgh_f + \frac{1}{2}kx_f^2$$

where $h$ is the height (vertical position) and $x$ is the compression of the spring. This general statement looks complex but becomes much simpler when we start considering specific situations. First, we must identify the initial and final conditions in a problem; then, we enter them into the last equation to solve for an unknown.

**Solution for (a)**

This part of the problem is limited to conditions just before the car is released and just after it leaves the spring. Take the initial height to be zero, so that both $h_i$ and $h_f$ are zero. Furthermore, the initial speed $v_i$ is zero and the final compression of the spring $x_f$ is zero, and so several terms in the conservation of mechanical energy equation are zero and it simplifies to

$$\frac{1}{2}kx_i^2 = \frac{1}{2}mv_f^2.$$

In other words, the initial potential energy in the spring is converted completely to kinetic energy in the absence of friction. Solving for the final speed and entering known values yields

$$\begin{aligned} v_f &= \sqrt{\frac{k}{m}}x_i \\ &= \sqrt{\frac{250N/m}{0.100kg}}(0.0400\ m) \\ &= 2.00m/s \end{aligned}$$

**Solution for (b)**

One method of finding the speed at the top of the slope is to consider conditions just before the car is released and just after it reaches the top of the slope, completely ignoring everything in between. Doing the same type of analysis to find which terms are zero, the conservation of mechanical energy becomes

$$\frac{1}{2}kx_i^2 = \frac{1}{2}mv_f^2 + mgh_f.$$

This form of the equation means that the spring's initial potential energy is converted partly to gravitational potential energy and partly to kinetic energy. The final speed at the top of the slope will be less than at the bottom. Solving for $v_f$ and substituting known values gives

$$\begin{aligned} v_f &= \sqrt{\frac{kx^2}{m} - 2gh_f} \\ &= \sqrt{\left(\frac{250N/m}{0.100kg}\right)(0.0400m)^2 - 2(9.80m/s^2)(0.180m)} \\ &= 0.687\,m/s \end{aligned}$$

**Discussion**

Another way to solve this problem is to realize that the car's kinetic energy before it goes up the slope is converted partly to potential energy—that is, to take the final conditions in part (a) to be the initial conditions in part (b).

Note that, for conservative forces, we do not directly calculate the work they do; rather, we consider their effects through their corresponding potential energies, just as we did in Example. Note also that we do not consider details of the path taken—only the starting and ending points are important (as long as the path is not impossible). This assumption is usually a tremendous simplification, because the path may be complicated and forces may vary along the way.

PHET Explorations: Energy Skate Park

Learn about conservation of energy with a skater dude! Build tracks, ramps and jumps for the skater and view the kinetic energy, potential energy and friction as he moves. You can also take the skater to different planets or even space!

## Summary

- A conservative force is one for which work depends only on the starting and ending points of a motion, not on the path taken.
- We can define potential energy ($PE$ for any conservative force, just as we defined $PE_g$ for the gravitational force.

- The potential energy of a spring is $PE_s = \frac{1}{2}kx^2$, where $k$ is the spring's force constant and |(x\) is the displacement from its undeformed position.
- Mechanical energy is defined to be $KE = PE$ for conservative force.
- When only conservative forces act on and within a system, the total mechanical energy is constant. In equation form,

$$KE + PE = constant \tag{4.5.13}$$

or

$$KE_i + PE_i = KE_f + PE_f \tag{4.5.14}$$

where i and f denote initial and final values. This is known as the conservation of mechanical energy.

# Glossary

**conservative force**
: a force that does the same work for any given initial and final configuration, regardless of the path followed

**potential energy**
: energy due to position, shape, or configuration

**potential energy of a spring**
: the stored energy of a spring as a function of its displacement; when Hooke's law applies, it is given by the expression $\frac{1}{2}kx^2$ where $x$ is the distance the spring is compressed or extended and $k$ is the spring constant

**conservation of mechanical energy**
: the rule that the sum of the kinetic energies and potential energies remains constant if only conservative forces act on and within a system

**mechanical energy**
: the sum of kinetic energy and potential energy

This page titled 4.5: Conservative Forces and Potential Energy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.4: Conservative Forces and Potential Energy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.6: Conservation of Energy

**Learning Objectives**

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

- Explain the law of the conservation of energy.
- Describe some of the many forms of energy.
- Define efficiency of an energy conversion process as the fraction left as useful energy or work, rather than being transformed, for example, into thermal energy.

Energy, as we have noted, is conserved, making it one of the most important physical quantities in nature. The law of conservation of energy can be stated as follows:

We have explored some forms of energy and some ways it can be transferred from one system to another. This exploration led to the definition of two major types of energy—mechanical energy $(KE + PE)$ and energy transferred via work done by nonconservative forces $(W_{nc})$ But energy takes *many* other forms, manifesting itself in *many* different ways, and we need to be able to deal with all of these before we can write an equation for the above general statement of the conservation of energy.

## Other Forms of Energy than Mechanical Energy

At this point, we deal with all other forms of energy by lumping them into a single group called other energy $(OE)$. Then we can state the conservation of energy in equation form as

$$KE_i + PE_i + W_{nc} + OE_i = KE_f + PE_f + OE_f. \tag{4.6.1}$$

All types of energy and work can be included in this very general statement of conservation of energy. Kinetic energy is $KE$, work done by a conservative force is represented by $PE$, work done by nonconservative forces is $W_{nc}$ and all other energies are included as $OE$. This equation applies to all previous examples; in those situations $OE$ was constant, and so it subtracted out and was not directly considered.

**Usefulness of the Energy Conservation Principle**

The fact that energy is conserved and has many forms makes it very important. You will find that energy is discussed in many contexts, because it is involved in all processes. It will also become apparent that many situations are best understood in terms of energy and that problems are often most easily conceptualized and solved by considering energy.

When does $OE$ play a role? One example occurs when a person eats. Food is oxidized with the release of carbon dioxide, water, and energy. Some of this chemical energy is converted to kinetic energy when the person moves, to potential energy when the person changes altitude, and to thermal energy (another form of $OE$).

## Some of the Many Forms of Energy

What are some other forms of energy? You can probably name a number of forms of energy not yet discussed. Many of these will be covered in later chapters, but let us detail a few here. **Electrical energy** is a common form that is converted to many other forms and does work in a wide range of practical situations. Fuels, such as gasoline and food, carry **chemical energy** that can be transferred to a system through oxidation. Chemical fuel can also produce electrical energy, such as in batteries. Batteries can in turn produce light, which is a very pure form of energy. Most energy sources on Earth are in fact stored energy from the energy we receive from the Sun. We sometimes refer to this as **radiant energy**, or electromagnetic radiation, which includes visible light, infrared, and ultraviolet radiation. **Nuclear energy** comes from processes that convert measurable amounts of mass into energy. Nuclear energy is transformed into the energy of sunlight, into electrical energy in power plants, and into the energy of the heat transfer and blast in weapons. Atoms and molecules inside all objects are in random motion. This internal mechanical energy

from the random motions is called **thermal energy**, because it is related to the temperature of the object. These and all other forms of energy can be converted into one another and can do work.

Table gives the amount of energy stored, used, or released from various objects and in various phenomena. The range of energies and the variety of types and situations is impressive.

### Problem-Solving Strategies for Energy

You will find the following problem-solving strategies useful whenever you deal with energy. The strategies help in organizing and reinforcing energy concepts. In fact, they are used in the examples presented in this chapter. The familiar general problem-solving strategies presented earlier—involving identifying physical principles, knowns, and unknowns, checking units, and so on —continue to be relevant here.

**Step 1.** Determine the system of interest and identify what information is given and what quantity is to be calculated. A sketch will help.

**Step 2.** Examine all the forces involved and determine whether you know or are given the potential energy from the work done by the forces. Then use step 3 or step 4.

**Step 3.** If you know the potential energies for the forces that enter into the problem, then forces are all conservative, and you can apply conservation of mechanical energy simply in terms of potential and kinetic energy. The equation expressing conservation of energy is

$$KE_i + PE_i = KE_f + PE_f. \tag{4.6.2}$$

**Step 4.** If you know the potential energy for only some of the forces, possibly because some of them are nonconservative and do not have a potential energy, or if there are other energies that are not easily treated in terms of force and work, then the conservation of energy law in its most general form must be used.

$$KE_i + PE_i + W_{nc} + OE_i = KE_f + PE_f + OE_f. \tag{4.6.3}$$

In most problems, one or more of the terms is zero, simplifying its solution. Do not calculate $W_c$, the work done by conservative forces; it is already incorporated in the $PE$ terms.

**Step 5.** You have already identified the types of work and energy involved (in step 2). Before solving for the unknown, *eliminate terms wherever possible* to simplify the algebra. For example, choose $h = 0$ at either the initial or final point, so that $PE_g$ is zero there. Then solve for the unknown in the customary manner.

**Step 6.** *Check the answer to see if it is reasonable*. Once you have solved a problem, reexamine the forms of work and energy to see if you have set up the conservation of energy equation correctly. For example, work done against friction should be negative, potential energy at the bottom of a hill should be less than that at the top, and so on. Also check to see that the numerical value obtained is reasonable. For example, the final speed of a skateboarder who coasts down a 3-m-high ramp could reasonably be 20 km/h, but *not* 80 km/h.

## Transformation of Energy

The transformation of energy from one form into others is happening all the time. The chemical energy in food is converted into thermal energy through metabolism; light energy is converted into chemical energy through photosynthesis. In a larger example, the chemical energy contained in coal is converted into thermal energy as it burns to turn water into steam in a boiler. This thermal energy in the steam in turn is converted to mechanical energy as it spins a turbine, which is connected to a generator to produce electrical energy. (In all of these examples, not all of the initial energy is converted into the forms mentioned. This important point is discussed later in this section.)

Another example of energy conversion occurs in a solar cell. Sunlight impinging on a solar cell (Figure 7.7.1) produces electricity, which in turn can be used to run an electric motor. Energy is converted from the primary source of solar energy into electrical energy and then into mechanical energy.



Figure 4.6.1: Solar energy is converted into electrical energy by solar cells, which is used to run a motor in this solar-power aircraft. (credit: NASA)

| Object/phenomenon | Energy in joules |
|---|---|
| Big Bang | $10^{68}$ |
| Energy released in a supernova | $10^{44}$ |
| Fusion of all the hydrogen in Earth's oceans | $10^{34}$ |
| Annual world energy use | $4 \times 10^{20}$ |
| Large fusion bomb (9 megaton) | $3.8 \times 10^{16}$ |
| 1 kg hydrogen (fusion to helium) | $6.4 \times 10^{14}$ |
| 1 kg uranium (nuclear fission) | $8.0 \times 10^{13}$ |
| Hiroshima-size fission bomb (10 kiloton) | $4.2 \times 10^{13}$ |
| 90,000-ton aircraft carrier at 30 knots | $1.1 \times 10^{10}$ |
| 1 barrel crude oil | $5.9 \times 10^{9}$ |
| 1 ton TNT | $4.2 \times 10^{9}$ |
| 1 gallon of gasoline | $1.2 \times 10^{8}$ |
| Daily home electricity use (developed countries) | $7n \times 10^{7}$ |
| Daily adult food intake (recommended) | $1.2 \times 10^{7}$ |
| 1000-kg car at 90 km/h | $3.1 \times 10^{5}$ |
| 1 g fat (9.3 kcal) | $3.9 \times 10^{4}$ |
| ATP hydrolysis reaction | $3.2 \times 10^{4}$ |
| 1 g carbohydrate (4.1 kcal) | $1.7 \times 10^{4}$ |
| 1 g protein (4.1 kcal) | $1.7 \times 10^{4}$ |
| Tennis ball at 100 km/h | 22 |
| Mosquito g at 0.5 m/s) | $1.3 \times 10^{-6}$ |
| Single electron in a TV tube beam | $4.0 \times 10^{-15}$ |
| Energy to break one DNA strand | $4.0 \times 10^{-19}$ |

# Efficiency

Even though energy is conserved in an energy conversion process, the output of *useful energy* or work will be less than the energy input. The efficiency $E_{ff}$ of an energy conversion process is defined as

$$Efficiency\,(E_{ff}) = \frac{useful\,energy\,or\,work\,output}{total\,energy\,input} = \frac{W_{out}}{E_{in}}. \tag{4.6.4}$$

Table lists some efficiencies of mechanical devices and human activities. In a coal-fired power plant, for example, about 40% of the chemical energy in the coal becomes useful electrical energy. The other 60% transforms into other (perhaps less useful) energy forms, such as thermal energy, which is then released to the environment through combustion gases and cooling towers.

| Activity/device | Efficiency (%) |
|---|---|
| Cycling and climbing | 20 |
| Swimming, surface | 2 |
| Swimming, submerged | 4 |
| Shoveling | 3 |
| Weightlifting | 9 |
| Steam engine | 17 |
| Gasoline engine | 30 |
| Diesel engine | 35 |
| Nuclear power plant | 35 |
| Coal power plant | 42 |
| Electric motor | 98 |
| Compact fluorescent light | 20 |
| Gas heater (residential) | 90 |
| Solar cell | 10 |

Efficiency of the Human Body and Mechanical Devices

**PhET Explorations: Masses and Springs**

A realistic mass and spring laboratory. Hang masses from springs and adjust the spring stiffness and damping. You can even slow time. Transport the lab to different planets. A chart shows the kinetic, potential, and thermal energies for each spring.




Figure 4.6.2: Masses and Springs

# Summary

- The law of conservation of energy states that the total energy is constant in any process. Energy may change in form or be transferred from one system to another, but the total remains the same.
- When all forms of energy are considered, conservation of energy is written in equation form as

$$KE_i + PE_i + W_{nc} + OE_i = KE_f + PE_f + OE_f, \tag{4.6.5}$$

where $OE$ is all **other forms of energy** besides mechanical energy.

- Commonly encountered forms of energy include electric energy, chemical energy, radiant energy, nuclear energy, and thermal energy.
- Energy is often utilized to do work, but it is not possible to convert all the energy of a system to work.

The efficiency $E_{ff}$ of a machine or human is defined to be $E_{ff} = \frac{W_{out}}{E_{in}}$, where $W_{out}$ is useful work output and $E_{in}$ s the energy consumed.

# Glossary

**law of conservation of energy**
: the general law that total energy is constant in any process; energy may change in form or be transferred from one system to another, but the total remains the same

**electrical energy**
: the energy carried by a flow of charge

**chemical energy**
: the energy in a substance stored in the bonds between atoms and molecules that can be released in a chemical reaction

**radiant energy**
: the energy carried by electromagnetic waves

**nuclear energy**
: energy released by changes within atomic nuclei, such as the fusion of two light nuclei or the fission of a heavy nucleus

**thermal energy**
: the energy within an object due to the random motion of its atoms and molecules that accounts for the object's temperature

**efficiency**
: a measure of the effectiveness of the input of energy to do work; useful energy or work divided by the total input of energy

This page titled 4.6: Conservation of Energy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.6: Conservation of Energy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.7: Power

**Learning Objectives**

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

- Calculate power by calculating changes in energy over time.
- Examine power consumption and calculations of the cost of energy consumed.

## What is Power?

*Power*—the word conjures up many images: a professional football player muscling aside his opponent, a dragster roaring away from the starting line, a volcano blowing its lava into the atmosphere, or a rocket blasting off, as in Figure.



Figure 4.7.1: This powerful rocket on the Space Shuttle Endeavor did work and consumed energy at a very high rate. (credit: NASA)

These images of power have in common the rapid performance of work, consistent with the scientific definition of power $P$ as the rate at which work is done.

**Power**

Power is the rate at which work is done.

$$P = \frac{W}{t} \tag{4.7.1}$$

The SI unit for power is the watt $W$, where 1 watt equals 1 joule/second $(1\ W = 1\ J/s)$.

Because work is energy transfer, power is also the rate at which energy is expended. A 60-W light bulb, for example, expends 60 J of energy per second. Great power means a large amount of work or energy developed in a short time. For example, when a powerful car accelerates rapidly, it does a large amount of work and consumes a large amount of fuel in a short time.

# Calculating Power from Energy

## Example 4.7.1: Calculating the Power to Climb Stairs

What is the power output for a 60.0-kg woman who runs up a 3.00 m high flight of stairs in 3.50 s, starting from rest but having a final speed of 2.00 m/s? (See Figure.)

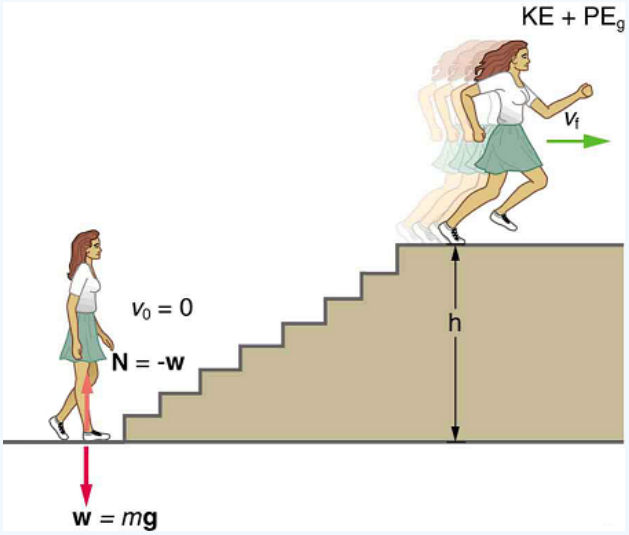



Figure 4.7.2:When this woman runs upstairs starting from rest, she converts the chemical energy originally from food into kinetic energy and gravitational potential energy. Her power output depends on how fast she does this.

**Strategy and Concept**

The work going into mechanical energy is $W = KE + PE$. At the bottom of the stairs, we take both $KE$ and $PE$ as initially zero; thus $W = KE_f + PE_g = \frac{1}{2}mv_f^2 + mgh$, where $h$ is the vertical height of the stairs. Because all terms are given, we can calculate $W$ and then divide it by time to get power.

**Solution**

Substituting the expression for $W$ into the definition of power given in the previous equation, $P = W/t$ yields

$$W = \frac{W}{t} = \frac{\frac{1}{2}mv_f^2 + mgh}{t}. \tag{4.7.2}$$

Entering known values yields

$$P = \frac{0.5(60\,kg)(2.00m/s^2) + (60.0\,kg)(9.80m/s^2)(3.00\,m)}{3.50\,s} \tag{4.7.3}$$

$$= \frac{120\,J + 1764\,J}{3.50\,s} \tag{4.7.4}$$

$$= 538\,W. \tag{4.7.5}$$

**Discussion**

The woman does 1764 J of work to move up the stairs compared with only 120 J to increase her kinetic energy; thus, most of her power output is required for climbing rather than accelerating.

It is impressive that this woman's useful power output is slightly less than 1 horsepower $(1\ hp = 746\ W)$. People can generate more than a horsepower with their leg muscles for short periods of time by rapidly converting available blood sugar and oxygen into work output. (A horse can put out 1 hp for hours on end.) Once oxygen is depleted, power output decreases and the person

begins to breathe rapidly to obtain oxygen to metabolize more food—this is known as the *aerobic* stage of exercise. If the woman climbed the stairs slowly, then her power output would be much less, although the amount of work done would be the same.

### Making Connections: Take-Home Investigation —Measure Your Power Rating

- Determine your own power rating by measuring the time it takes you to climb a flight of stairs. We will ignore the gain in kinetic energy, as the above example showed that it was a small portion of the energy gain. Don't expect that your output will be more than about 0.5 hp.

## Examples of Power

Examples of power are limited only by the imagination, because there are as many types as there are forms of work and energy. (See Table for some examples.) Sunlight reaching Earth's surface carries a maximum power of about 1.3 kilowatts per square meter $kW/m^2$ A tiny fraction of this is retained by Earth over the long term. Our consumption rate of fossil fuels is far greater than the rate at which they are stored, so it is inevitable that they will be depleted. Power implies that energy is transferred, perhaps changing form. It is never possible to change one form completely into another without losing some of it as thermal energy. For example, a 60-W incandescent bulb converts only 5 W of electrical power to light, with 55 W dissipating into thermal energy. Furthermore, the typical electric power plant converts only 35 to 40% of its fuel into electricity. The remainder becomes a huge amount of thermal energy that must be dispersed as heat transfer, as rapidly as it is created. A coal-fired power plant may produce 1000 megawatts; 1 megawatt (MW) is $10^6$ of electric power. But the power plant consumes chemical energy at a rate of about 2500 MW, creating heat transfer to the surroundings at a rate of 1500 MW. (See Figure.)



Figure $4.7.3$: Tremendous amounts of electric power are generated by coal-fired power plants such as this one in China, but an even larger amount of power goes into heat transfer to the surroundings. The large cooling towers here are needed to transfer heat as rapidly as it is produced. The transfer of heat is not unique to coal plants but is an unavoidable consequence of generating electric power from any fuel—nuclear, coal, oil, natural gas, or the like. (credit: Kleinolive, Wikimedia Commons)

| Object or Phenomenon | Power in Watts |
|---|---|
| Supernova (at peak) | $5 \times 10^{37}$ |
| Milky Way galaxy | $10^{37}$ |
| Crab Nebula pulsar | $10^{28}$ |
| The Sun | $4 \times 10^{26}$ |
| Volcanic eruption (maximum) | $4 \times 10^{15}$ |
| Lightning bolt | $2 \times 10^{12}$ |

| Object or Phenomenon | Power in Watts |
|---|---|
| Nuclear power plant (total electric and heat transfer) | $3 \times 10^9$ |
| Aircraft carrier (total useful and heat transfer) | $10^8$ |
| Dragster (total useful and heat transfer) | $2 \times 10^6$ |
| Car (total useful and heat transfer) | $8 \times 10^4$ |
| Football player (total useful and heat transfer) | $5 \times 10^3$ |
| Clothes dryer | $4 \times 10^3$ |
| Person at rest (all heat transfer) | 100 |
| Typical incandescent light bulb (total useful and heat transfer) | 60 |
| Heart, person at rest (total useful and heat transfer) | 8 |
| Electric clock | 3 |
| Pocket calculator | $10^{-3}$ |

## Power and Energy Consumption

We usually have to pay for the energy we use. It is interesting and easy to estimate the cost of energy for an electrical appliance if its power consumption rate and time used are known. The higher the power consumption rate and the longer the appliance is used, the greater the cost of that appliance. The power consumption rate is $P = \frac{W}{t} = \frac{E}{t}$, where $E$ is the energy supplied by the electricity company. So the energy consumed over a time $t$ is

$$E = Pt. \tag{4.7.6}$$

Electricity bills state the energy used in units of kilowatt-hours $(kW \cdot h)$, which is the product of power in kilowatts and time in hours. This unit is convenient because electrical power consumption at the kilowatt level for hours at a time is typical.

### Example4.7.2: Calculating Energy Costs

What is the cost of running a 0.200-kW computer 6.00 h per day for 30.0 d if the cost of electricity is \$0.120 per $kW \cdot h$?

**Strategy**

Cost is based on energy consumed; thus, we must find $E$ from $E = Pt$ and then calculate the cost. Because electrical energy is expressed in $kW \cdot h$ at the start of a problem such as this it is convenient to convert the units into $kW$ and hours.

**Solution**

The energy consumed in $kW \cdot h$ is

$$E = Pt = (0.200\,kW)(6.00\,h/d)(30.0\,d) \tag{4.7.7}$$

$$= 36.0\,kW \cdot h, \tag{4.7.8}$$

and the cost is simply given by

$$cost = (36.0\,kW \cdot h)(\$0.120\,per\,kW \cdot h) = \$4.32\,per\,month. \tag{4.7.9}$$

**Discussion**

The cost of using the computer in this example is neither exorbitant nor negligible. It is clear that the cost is a combination of power and time. When both are high, such as for an air conditioner in the summer, the cost is high.

The motivation to save energy has become more compelling with its ever-increasing price. Armed with the knowledge that energy consumed is the product of power and time, you can estimate costs for yourself and make the necessary value judgments about where to save energy. Either power or time must be reduced. It is most cost-effective to limit the use of high-power devices that normally operate for long periods of time, such as water heaters and air conditioners. This would not include relatively high power devices like toasters, because they are on only a few minutes per day. It would also not include electric clocks, in spite of their 24-hour-per-day usage, because they are very low power devices. It is sometimes possible to use devices that have greater efficiencies—that is, devices that consume less power to accomplish the same task. One example is the compact fluorescent light bulb, which produces over four times more light per watt of power consumed than its incandescent cousin.

Modern civilization depends on energy, but current levels of energy consumption and production are not sustainable. The likelihood of a link between global warming and fossil fuel use (with its concomitant production of carbon dioxide), has made reduction in energy use as well as a shift to non-fossil fuels of the utmost importance. Even though energy in an isolated system is a conserved quantity, the final result of most energy transformations is waste heat transfer to the environment, which is no longer useful for doing work. As we will discuss in more detail in Thermodynamics, the potential for energy to produce useful work has been “degraded” in the energy transformation.

## Summary

- Power is the rate at which work is done, or in equation form, for the average power $P$ for work $W$ done over a time $t$, $P = W/t$ .
- The SI unit for power is the watt (W), where $1\ W = 1\ J/s$.
- The power of many devices such as electric motors is also often expressed in horsepower (hp), where $1\ hp = 746\ W$.

## Glossary

**power**
: the rate at which work is done

**watt**
: (W) SI unit of power, with $1W = 1J/s$

**horsepower**
: an older non-SI unit of power, with $1hp = 746W$

**kilowatt-hour**
: ($kW \cdot h$) unit used primarily for electrical energy provided by electric utility companies

This page titled 4.7: Power is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.7: Power by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.8: Work, Energy, and Power in Humans

**Learning Objectives**

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

- Explain the human body's consumption of energy when at rest vs. when engaged in activities that do useful work.
- Calculate the conversion of chemical energy in food into useful work.

## Energy Conversion in Humans

Our own bodies, like all living organisms, are energy conversion machines. Conservation of energy implies that the chemical energy stored in food is converted into work, thermal energy, and/or stored as chemical energy in fatty tissue. (Figure 7.09.1.) The fraction going into each form depends both on how much we eat and on our level of physical activity. If we eat more than is needed to do work and stay warm, the remainder goes into body fat.

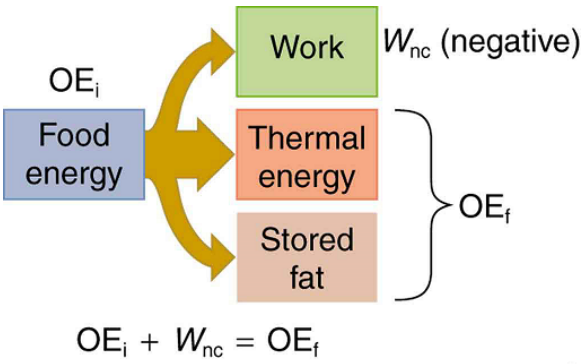



Figure 4.8.1: Energy consumed by humans is converted to work, thermal energy, and stored fat. By far the largest fraction goes to thermal energy, although the fraction varies depending on the type of physical activity.

## Power Consumed at Rest

The *rate* at which the body uses food energy to sustain life and to do different activities is called the metabolic rate. The total energy conversion rate of a person *at rest* is called the basal metabolic rate (BMR) and is divided among various systems in the body, as shown in Table. The largest fraction goes to the liver and spleen, with the brain coming next. Of course, during vigorous exercise, the energy consumption of the skeletal muscles and heart increase markedly. About 75% of the calories burned in a day go into these basic functions. The BMR is a function of age, gender, total body weight, and amount of muscle mass (which burns more calories than body fat). Athletes have a greater BMR due to this last factor.

*Basal Metabolic Rates (BMR):*

| Organ | Power consumed at rest (W) | Oxygen consumption (mL/min) | Percent of BMR |
|---|---|---|---|
| Liver & spleen | 23 | 67 | 27 |
| Brain | 16 | 47 | 19 |
| Skeletal muscle | 15 | 45 | 18 |
| Kidney | 9 | 26 | 10 |
| Heart | 6 | 17 | 7 |
| Other | 16 | 48 | 19 |

| Organ | Power consumed at rest (W) | Oxygen consumption (mL/min) | Percent of BMR |
|---|---|---|---|
| **Totals** | 85 W | 250 mL/min | 100% |

Energy consumption is directly proportional to oxygen consumption because the digestive process is basically one of oxidizing food. We can measure the energy people use during various activities by measuring their oxygen use. (See Figure7.09.1.) Approximately 20 kJ of energy are produced for each liter of oxygen consumed, independent of the type of food. Table shows energy and oxygen consumption rates (power expended) for a variety of activities.

## Power of Doing Useful Work

Work done by a person is sometimes called useful work, which is *work done on the outside world*, such as lifting weights. Useful work requires a force exerted through a distance on the outside world, and so it excludes internal work, such as that done by the heart when pumping blood. Useful work does include that done in climbing stairs or accelerating to a full run, because these are accomplished by exerting forces on the outside world. Forces exerted by the body are non-conservative, so that they can change the mechanical energy $(KE + PE)$ of the system worked upon, and this is often the goal. A baseball player throwing a ball, for example, increases both the ball's kinetic and potential energy.

If a person needs more energy than they consume, such as when doing vigorous work, the body must draw upon the chemical energy stored in fat. So exercise can be helpful in losing fat. However, the amount of exercise needed to produce a loss in fat, or to burn off extra calories consumed that day, can be large, as Example 7.09.1 illustrates.

### Example 4.8.1: Calculating Weight Loss from Exercising

If a person who normally requires an average of 12,000 kJ (3000 kcal) of food energy per day consumes 13,000 kJ per day, he will steadily gain weight. How much bicycling per day is required to work off this extra 1000 kJ?

**Solution**

Table states that 400 W are used when cycling at a moderate speed. The time required to work off 1000 kJ at this rate is then

$$Time = \frac{energy}{\left(\frac{energy}{time}\right)} = \frac{1000\,kJ}{400\,W} = 2500\,s = 42\,min. \tag{4.8.1}$$

**Discussion**

If this person uses more energy than he or she consumes, the person's body will obtain the needed energy by metabolizing body fat. If the person uses 13,000 kJ but consumes only 12,000 kJ, then the amount of fat loss will be

$$Fat\,loss = (1000\,kJ)\left(\frac{1\,g\,fat}{30\,kJ}\right) = 26\,g, \tag{4.8.2}$$

assuming the energy content of fat to be 39 kJ/g.

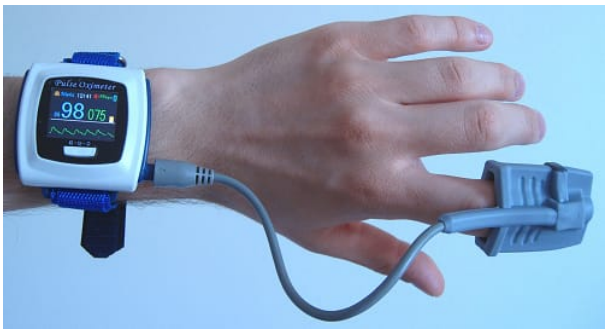



Figure 4.8.2: A pulse oximeter is an apparatus that measures the amount of oxygen in blood. A knowledge of oxygen and carbon dioxide levels indicates a person's metabolic rate, which is the rate at

which food energy is converted to another form. a person's metabolic rate, which is the rate at which food energy is converted to another form. Such measurements can indicate the level of athletic conditioning as well as certain medical problems. (credit: UusiAjaja, Wikimedia Commons)

Energy and Oxygen Consumption Rates:

| Activity | Energy consumption in watts | Oxygen consumption in liters $O_2$/min |
|---|---|---|
| Sleeping | 83 | 0.24 |
| Sitting at rest | 120 | 0.34 |
| Standing relaxed | 125 | 0.36 |
| Sitting in class | 210 | 0.60 |
| Walking (5 km/h) | 280 | 0.80 |
| Cycling (13–18 km/h) | 400 | 1.14 |
| Shivering | 425 | 1.21 |
| Playing tennis | 440 | 1.26 |
| Swimming breaststroke | 475 | 1.36 |
| Ice skating (14.5 km/h) | 545 | 1.56 |
| Climbing stairs (116/min) | 685 | 1.96 |
| Cycling (21 km/h) | 700 | 2.00 |
| Running cross-country | 740 | 2.12 |
| Playing basketball | 800 | 2.28 |
| Cycling, professional racer | 1855 | 5.30 |
| Sprinting | 2415 | 6.90 |

All bodily functions, from thinking to lifting weights, require energy. (See Figure 7.09.3.) The many small muscle actions accompanying all quiet activity, from sleeping to head scratching, ultimately become thermal energy, as do less visible muscle actions by the heart, lungs, and digestive tract. Shivering, in fact, is an involuntary response to low body temperature that pits muscles against one another to produce thermal energy in the body (and do no work). The kidneys and liver consume a surprising amount of energy, but the biggest surprise of all it that a full 25% of all energy consumed by the body is used to maintain electrical potentials in all living cells. (Nerve cells use this electrical potential in nerve impulses.) This bioelectrical energy ultimately becomes mostly thermal energy, but some is utilized to power chemical processes such as in the kidneys and liver, and in fat production.

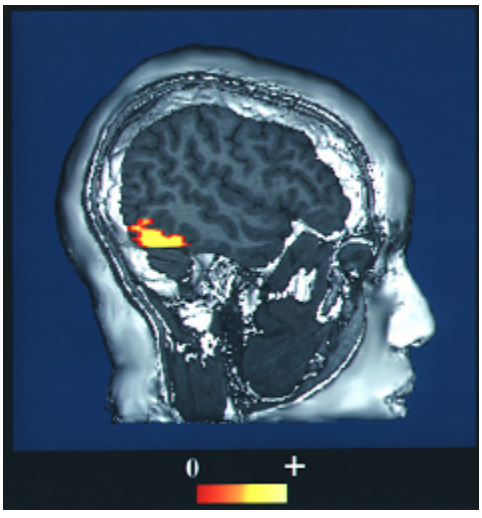



Figure 4.8.3: This MRI scan shows an increased level of energy consumption in the

vision center of the brain. Here, the patient was being asked to recognize faces. (credit: NIH via Wikimedia Commons)

## Summary

- The human body converts energy stored in food into work, thermal energy, and/or chemical energy that is stored in fatty tissue.
- The *rate* at which the body uses food energy to sustain life and to do different activities is called the metabolic rate, and the corresponding rate when at rest is called the basal metabolic rate (BMR)
- The energy included in the basal metabolic rate is divided among various systems in the body, with the largest fraction going to the liver and spleen, and the brain coming next.
- About 75% of food calories are used to sustain basic body functions included in the basal metabolic rate.
- The energy consumption of people during various activities can be determined by measuring their oxygen use, because the digestive process is basically one of oxidizing food.

## Glossary

**metabolic rate**
: the rate at which the body uses food energy to sustain life and to do different activities

**basal metabolic rate**
: the total energy conversion rate of a person at rest

**useful work**
: work done on an external system

This page titled 4.8: Work, Energy, and Power in Humans is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.8: Work, Energy, and Power in Humans by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.9: World Energy Use

**Learning Objectives**

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

- Describe the distinction between renewable and nonrenewable energy sources.
- Explain why the inevitable conversion of energy to less useful forms makes it necessary to conserve energy resources.

Energy is an important ingredient in all phases of society. We live in a very interdependent world, and access to adequate and reliable energy resources is crucial for economic growth and for maintaining the quality of our lives. But current levels of energy consumption and production are not sustainable. About 40% of the world's energy comes from oil, and much of that goes to transportation uses. Oil prices are dependent as much upon new (or foreseen) discoveries as they are upon political events and situations around the world. The U.S., with 4.5% of the world's population, consumes 24% of the world's oil production per year; 66% of that oil is imported!

## Renewable and Nonrenewable Energy Sources

The principal energy resources used in the world are shown in Figure 7.10.1. The fuel mix has changed over the years but now is dominated by oil, although natural gas and solar contributions are increasing. Renewable forms of energy are those sources that cannot be used up, such as water, wind, solar, and biomass. About 85% of our energy comes from nonrenewable fossil fuels—oil, natural gas, coal. The likelihood of a link between global warming and fossil fuel use, with its production of carbon dioxide through combustion, has made, in the eyes of many scientists, a shift to non-fossil fuels of utmost importance—but it will not be easy.

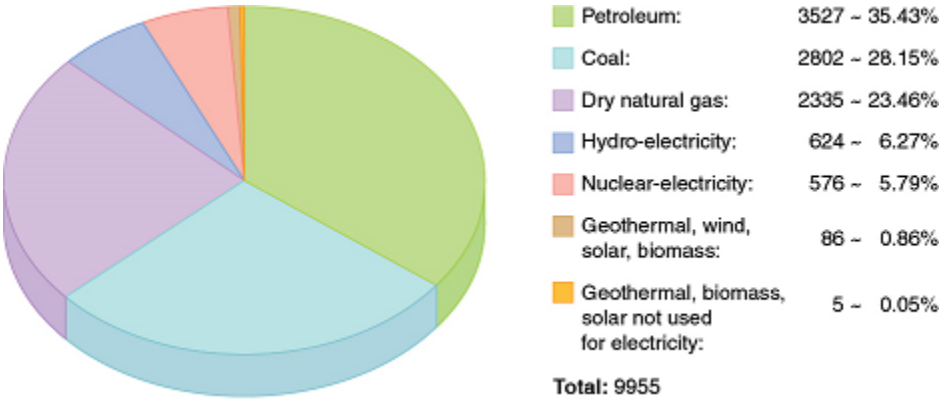



Figure 4.9.1: World energy consumption by source, in billions of kilowatt-hours: 2006. (credit: KVDP)

## The World's Growing Energy Needs

World energy consumption continues to rise, especially in the developing countries. (See Figure 7.10.1.) Global demand for energy has tripled in the past 50 years and might triple again in the next 30 years. While much of this growth will come from the rapidly booming economies of China and India, many of the developed countries, especially those in Europe, are hoping to meet their energy needs by expanding the use of renewable sources. Although presently only a small percentage, renewable energy is growing very fast, especially wind energy. For example, Germany plans to meet 20% of its electricity and 10% of its overall energy needs with renewable resources by the year 2020. (See 7.10.2.) Energy is a key constraint in the rapid economic growth of China and India. In 2003, China surpassed Japan as the world's second largest consumer of oil. However, over 1/3 of this is imported. Unlike most Western countries, coal dominates the commercial energy resources of China, accounting for 2/3 of its energy consumption. In 2009 China surpassed the United States as the largest generator of $CO_2$ In India, the main energy resources are biomass (wood and dung) and coal. Half of India's oil is imported. About 70% of India's electricity is generated by highly polluting coal. Yet there are sizeable strides being made in renewable energy. India has a rapidly growing wind energy

base, and it has the largest solar cooking program in the world.

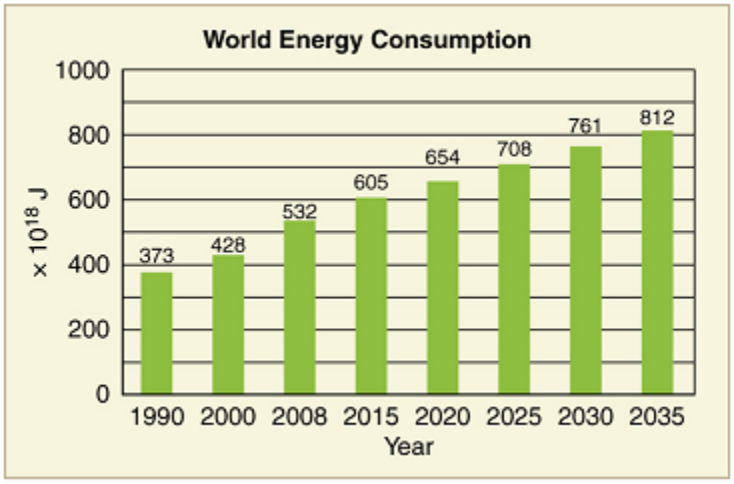



Figure 4.9.2: Past and projected world energy use (source: Based on data from U.S. Energy Information Administration, 2011)

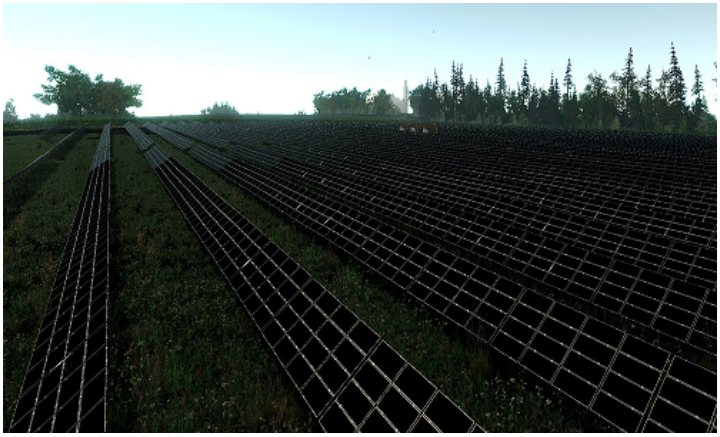


Figure 4.9.3: Solar cell arrays at a power plant in Steindorf, Germany (credit: Michael Betke, Flickr)

Table displays the 2006 commercial energy mix by country for some of the prime energy users in the world. While non-renewable sources dominate, some countries get a sizeable percentage of their electricity from renewable resources. For example, about 67% of New Zealand's electricity demand is met by hydroelectric. Only 10% of the U.S. electricity is generated by renewable resources, primarily hydroelectric. It is difficult to determine total contributihttp://physwiki.ucdavis.ed..._Energy_Useons of renewable energy in some countries with a large rural population, so these percentages in this table are left blank.

*Energy Consumption—Selected Countries (2006):*

| Country | Consumption, in EJ ($10^{18}$ J) | Oil | Natural Gas | Coal | Nuclear | Hydro | Other Renewables | Electricity Use per capita (kWh/yr) | Energy Use per capita (GJ/yr) |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 5.4 | 34% | 17% | 44% | 0% | 3% | 1% | 10000 | 260 |
| Brazil | 9.6 | 48% | 7% | 5% | 1% | 35% | 2% | 2000 | 50 |
| China | 63 | 22% | 3% | 69% | 1% | 6% | | 1500 | 35 |
| Egypt | 2.4 | 50% | 41% | 1% | 0% | 6% | | 990 | 32 |
| Germany | 16 | 37% | 24% | 24% | 11% | 1% | 3% | 6400 | 173 |
| India | 15 | 34% | 7% | 52% | 1% | 5% | | 470 | 13 |
| Indonesia | 4.9 | 51% | 26% | 16% | 0% | 2% | 3% | 420 | 22 |
| Japan | 24 | 48% | 14% | 21% | 12% | 4% | 1% | 7100 | 176 |

| Country | Consumption, in EJ ($10^{18}$ J) | Oil | Natural Gas | Coal | Nuclear | Hydro | Other Renewables | Electricity Use per capita (kWh/yr) | Energy Use per capita (GJ/yr) |
|---|---|---|---|---|---|---|---|---|---|
| New Zealand | 0.44 | 32% | 26% | 6% | 0% | 11% | 19% | 8500 | 102 |
| Russia | 31 | 19% | 53% | 16% | 5% | 6% | | 5700 | 202 |
| U.S. | 105 | 40% | 23% | 22% | 8% | 3% | 1% | 12500 | 340 |
| **World** | **432** | **39%** | **23%** | **24%** | **6%** | **6%** | **2%** | **2600** | **71** |

## Energy and Economic Well-being

The last two columns in this table examine the energy and electricity use per capita. Economic well-being is dependent upon energy use, and in most countries higher standards of living, as measured by GDP (gross domestic product) per capita, are matched by higher levels of energy consumption per capita. This is borne out in Figure 7.10.4. Increased efficiency of energy use will change this dependency. A global problem is balancing energy resource development against the harmful effects upon the environment in its extraction and use.

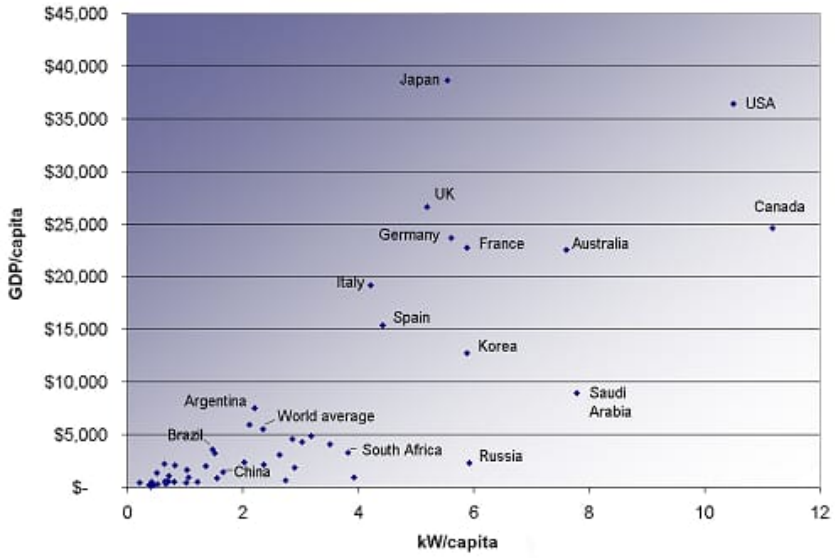



Figure 4.9.4: Power consumption per capita versus GDP per capita for various countries. Note the increase in energy usage with increasing GDP. (2007, credit: Frank van Mierlo, Wikimedia Commons)

## Conserving Energy

As we finish this chapter on energy and work, it is relevant to draw some distinctions between two sometimes misunderstood terms in the area of energy use. As has been mentioned elsewhere, the “law of the conservation of energy” is a very useful principle in analyzing physical processes. It is a statement that cannot be proven from basic principles, but is a very good bookkeeping device, and no exceptions have ever been found. It states that the total amount of energy in an isolated system will always remain constant. Related to this principle, but remarkably different from it, is the important philosophy of energy conservation. This concept has to do with seeking to decrease the amount of energy used by an individual or group through (1) reduced activities (e.g., turning down thermostats, driving fewer kilometers) and/or (2) increasing conversion efficiencies in the performance of a particular task—such as developing and using more efficient room heaters, cars that have greater miles-per-gallon ratings, energy-efficient compact fluorescent lights, etc.

Since energy in an isolated system is not destroyed or created or generated, one might wonder why we need to be concerned about our energy resources, since energy is a conserved quantity. The problem is that the final result of most energy transformations is waste heat transfer to the environment and conversion to energy forms no longer useful for doing work. To state it in another way, the potential for energy to produce useful work has been “degraded” in the energy transformation. (This

will be discussed in more detail in Thermodynamics.)

## Summary

- The relative use of different fuels to provide energy has changed over the years, but fuel use is currently dominated by oil, although natural gas and solar contributions are increasing.
- Although non-renewable sources dominate, some countries meet a sizeable percentage of their electricity needs from renewable resources.
- The United States obtains only about 10% of its energy from renewable sources, mostly hydroelectric power.
- Economic well-being is dependent upon energy use, and in most countries higher standards of living, as measured by GDP (Gross Domestic Product) per capita, are matched by higher levels of energy consumption per capita.
- Even though, in accordance with the law of conservation of energy, energy can never be created or destroyed, energy that can be used to do work is always partly converted to less useful forms, such as waste heat to the environment, in all of our uses of energy for practical purposes.

## Glossary

**renewable forms of energy**
: those sources that cannot be used up, such as water, wind, solar, and biomass

**fossil fuels**
: oil, natural gas, and coal

This page titled 4.9: World Energy Use is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 7.9: World Energy Use by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.10: Prelude to Linear Momentum and Collisions

We use the term momentum in various ways in everyday language, and most of these ways are consistent with its precise scientific definition. We speak of sports teams or politicians gaining and maintaining the momentum to win. We also recognize that momentum has something to do with collisions. For example, looking at the rugby players in the photograph colliding and falling to the ground, we expect their momenta to have great effects in the resulting collisions. Generally, momentum implies a tendency to continue on course—to move in the same direction—and is associated with great mass and speed.



Figure 4.10.1: Each rugby player has great momentum, which will affect the outcome of their collisions with each other and the ground. (credit: ozzzie, Flickr)

Momentum, like energy, is important because it is conserved. Only a few physical quantities are conserved in nature, and studying them yields fundamental insight into how nature works, as we shall see in our study of momentum.

This page titled 4.10: Prelude to Linear Momentum and Collisions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 8.0: Prelude to Linear Momentum and Collisions by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.11: Linear Momentum and Force

## Learning Objectives

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

- Define linear momentum.
- Explain the relationship between momentum and force.
- State Newton's second law of motion in terms of momentum.
- Calculate momentum given mass and velocity.

The scientific definition of linear momentum is consistent with most people's intuitive understanding of momentum: a large, fast-moving object has greater momentum than a smaller, slower object. Linear momentum is defined as the product of a system's mass multiplied by its velocity.

## Linear Momentum

Linear momentum is defined as the product of a system's mass multiplied by its velocity:

$$p = mv \tag{4.11.1}$$

Momentum is directly proportional to the object's mass and also its velocity. Thus the greater an object's mass or the greater its velocity, the greater its momentum. Momentum $p$ is a vector having the same direction as the velocity $v$. The SI unit for momentum is $kg \cdot m/s$.

## Example 4.11.1: Calculating Momentum: A Football Player and a Football

a. Calculate the momentum of a 110-kg football player running at 8.00 m/s.
b. Compare the player's momentum with the momentum of a hard-thrown 0.410-kg football that has a speed of 25.0 m/s.

**Strategy**

No information is given regarding direction, and so we can calculate only the magnitude of the momentum, $p$ (As usual, a symbol that is in italics is a magnitude, whereas one that is italicized, boldfaced, and has an arrow is a vector.) In both parts of this example, the magnitude of momentum can be calculated directly from the definition of momentum given in Equation 4.11.1, which becomes

$$p = mv$$

when only magnitudes are considered.

**Solution for (a)**

To determine the momentum of the player, substitute the known values for the player's mass and speed into the equation.

$$\begin{aligned} p_{player} &= (110\,kg)(8.00\,m/s) \\ &= 880\,kg \cdot m/s \end{aligned}$$

**Solution for (b)**

To determine the momentum of the ball, substitute the known values for the ball's mass and speed into the equation.

$$p_{ball} = (0.410\,kg)(25.0\,m/s)$$
$$= 10.3\,kg \cdot m/s$$

The ratio of the player's momentum to that of the ball is

$$\frac{p_{player}}{p_{ball}} = \frac{880}{10.3} = 85.0$$

**Discussion**

Although the ball has greater velocity, the player has a much greater mass. Thus the momentum of the player is much greater than the momentum of the football, as you might guess. As a result, the player's motion is only slightly affected if he catches the ball. We shall quantify what happens in such collisions in terms of momentum in later sections.

## Momentum and Newton's Second Law

The importance of momentum, unlike the importance of energy, was recognized early in the development of classical physics. Momentum was deemed so important that it was called the "quantity of motion." Newton actually stated his second law of motion in terms of momentum: The net external force equals the change in momentum of a system divided by the time over which it changes.

### Newton's Second Law of Motion in Terms of Momentum

The net external force equals the change in momentum of a system divided by the time over which it changes.

$$F_{net} = \frac{\Delta p}{\Delta t} \tag{4.11.2}$$

where $F_{net}$ is the net external force, $\Delta p$ is the change in momentum, and $\Delta t$ is the change in time.

## Making Connections: Force and Momentum

Force and momentum are intimately related. Force acting over time can change momentum, and Newton's second law of motion, can be stated in its most broadly applicable form in terms of momentum. Momentum continues to be a key concept in the study of atomic and subatomic particles in quantum mechanics.

This statement of Newton's second law of motion includes the more familiar $F_{net} = ma$ as a special case. We can derive this form as follows. First, note that the change in momentum $\Delta p$ is given by

$$\Delta p = \Delta(mv) \tag{4.11.3}$$

If the mass of the system is constant, then

$$\Delta(mv) = m\Delta v. \tag{4.11.4}$$

So that for constant mass, Newton's second law of motion becomes

$$F_{net} = \frac{\Delta p}{\Delta t} = \frac{m\Delta v}{\Delta t}. \tag{4.11.5}$$

Because $\frac{\Delta v}{\Delta t} = a$, we get the familiar equation

$$F_{net} = ma \tag{4.11.6}$$

when the mass of the system is *constant*.

Newton's second law of motion stated in terms of momentum is more generally applicable because it can be applied to systems

where the mass is changing, such as rockets, as well as to systems of constant mass. We will consider systems with varying mass in some detail; however, the relationship between momentum and force remains useful when mass is constant, such as in the following example.

### Example 4.11.2: Calculating Force: Venus Williams' Racquet

During the 2007 French Open, Venus Williams hit the fastest recorded serve in a premier women's match, reaching a speed of 58 m/s (209 km/h). What is the average force exerted on the 0.057-kg tennis ball by Venus Williams' racquet, assuming that the ball's speed just after impact is 58 m/s, that the initial horizontal component of the velocity before impact is negligible, and that the ball remained in contact with the racquet for 5.0 ms (milliseconds)?

**Strategy**

This problem involves only one dimension because the ball starts from having no horizontal velocity component before impact. Newton's second law stated in terms of momentum is then written as

$$F_{net} = \frac{\Delta p}{\Delta t}$$

As noted above, when mass is constant, the change in momentum is given by

$$\Delta p = m\Delta v = m(v_f - v_i).$$

In this example, the velocity just after impact and the change in time are given; thus, once $\Delta p$ s calculated, $F_{net} = \frac{\Delta p}{\Delta t}$ can be used to find the force.

**Solution**

To determine the change in momentum, substitute the values for the initial and final velocities into the equation above.

$$\begin{aligned}\Delta p &= m(v_f - v_i)\\ &= (0.057\,kg)(58\,m/s - 0\,m/s)\\ &= 3.306\,kg \cdot m/s = 3.3\,kg \cdot m/s\end{aligned}$$

Now the magnitude of the net external force can determined by using $F_{net} = \frac{\Delta p}{\Delta t}$

$$\begin{aligned}F_{net} = \frac{\Delta p}{\Delta t} &= \frac{3.306\,kg}{5.0 \times 10^{-3}}\\ &= 661\,N,\end{aligned}$$

where we have retained only two significant figures in the final step.

**Discussion**

This quantity was the average force exerted by Venus Williams' racquet on the tennis ball during its brief impact (note that the ball also experienced the 0.56-N force of gravity, but that force was not due to the racquet). This problem could also be solved by first finding the acceleration and then using $F = ma$ but one additional step would be required compared with the strategy used in this example.

## Summary

- Linear momentum (*momentum* for brevity) is defined as the product of a system's mass multiplied by its velocity.
- In symbols, linear momentum $p$ is defined to be

$$p = mv$$

  where $m$ is the mass of the system and $v$ is its velocity.
- The SI unit for momentum is $kg \cdot m/s$.

- Newton's second law of motion in terms of momentum states that the net external force equals the change in momentum of a system divided by the time over which it changes.
- In symbols, Newton's second law of motion is defined to be

$$F_{net} = \frac{\Delta p}{\Delta t}$$

where $F_{net}$ is the net external force, $\Delta p$ s the change in momentum, and $\Delta t$ is change in time.

# Glossary

**linear momentum**
: the product of mass and velocity

**second law of motion**
: physical law that states that the net external force equals the change in momentum of a system divided by the time over which it changes

This page titled 4.11: Linear Momentum and Force is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 8.1: Linear Momentum and Force by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.12: Conservation of Momentum

## Learning Objectives

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

- Describe the principle of conservation of momentum.
- Derive an expression for the conservation of momentum.
- Explain conservation of momentum with examples.
- Explain the principle of conservation of momentum as it relates to atomic and subatomic particles.

Momentum is an important quantity because it is conserved. Yet it was not conserved in the examples in Impulse and Linear Momentum and Force, where large changes in momentum were produced by forces acting on the system of interest. Under what circumstances is momentum conserved?

The answer to this question entails considering a sufficiently large system. It is always possible to find a larger system in which total momentum is constant, even if momentum changes for components of the system. If a football player runs into the goalpost in the end zone, there will be a force on him that causes him to bounce backward. However, the Earth also recoils —conserving momentum—because of the force applied to it through the goalpost. Because Earth is many orders of magnitude more massive than the player, its recoil is immeasurably small and can be neglected in any practical sense, but it is real nevertheless.

Consider what happens if the masses of two colliding objects are more similar than the masses of a football player and Earth—for example, one car bumping into another, as shown in Figure $4.12.1$. Both cars are coasting in the same direction when the lead car (labeled $m_2$ is bumped by the trailing car (labeled $m_1$. The only unbalanced force on each car is the force of the collision. (Assume that the effects due to friction are negligible.) Car 1 slows down as a result of the collision, losing some momentum, while car 2 speeds up and gains some momentum. We shall now show that the total momentum of the two-car system remains constant.

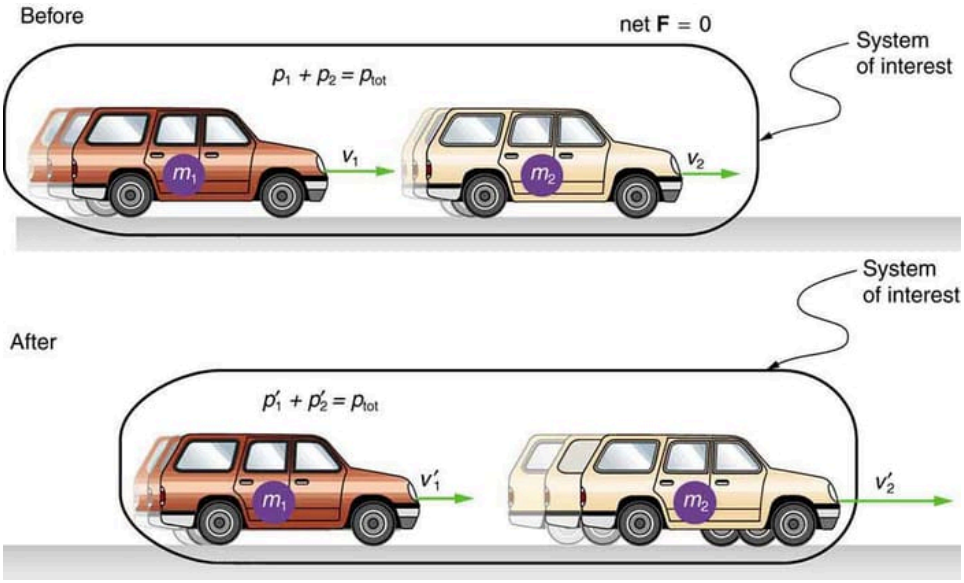



Figure $4.12.1$: A car of mass $m_1$ moving with a velocity of $v_1$ bumps into another car of mass $m_2$ and velocity $v_2$ that it is following. As a result, the first car slows down to a velocity of $v_1'$ and the second speeds up to a velocity of $v_2'$ The momentum of each car is changed, but the total momentum $p_{tot}$ of the two cars is the same before and after the collision (if you assume friction is negligible).

Using the definition of impulse, the change in momentum of car 1 is given by

$$\Delta p_1 = F_1 \Delta t, \tag{4.12.1}$$

is the force on car 1 due to car 2, and $\Delta t$

where $F_1$ is the time the force acts (the duration of the collision). Intuitively, it seems obvious that the collision time is the same for both cars, but it is only true for objects traveling at ordinary speeds. This assumption must be modified for objects travelling near the speed of light, without affecting the result that momentum is conserved.

Similarly, the change in momentum of car 2 is

$$\Delta p_2 = F_2 \Delta t, \tag{4.12.2}$$

where $F_2$ is the force on car 2 due to car 1, and we assume the duration of the collision $\Delta t$ is the same for both cars. We know from Newton's third law that $F_2 = -F_1$, and so

$$\Delta p_2 = -F_1 \Delta t = -\Delta p_1. \tag{4.12.3}$$

Thus, the changes in momentum are equal and opposite, and

$$\Delta p_1 + \Delta p_2 = 0. \tag{4.12.4}$$

Because the changes in momentum add to zero, the total momentum of the two-car system is constant. That is,

$$p_1 + p_2 = constant \tag{4.12.5}$$

$$p_1 + p_2 = p_1' + p_2', \tag{4.12.6}$$

where $p_1'$ and $p_2'$ are the momenta of cars 1 and 2 after the collision. (We often use primes to denote the final state.)

This result—that momentum is conserved—has validity far beyond the preceding one-dimensional case. It can be similarly shown that total momentum is conserved for any isolated system, with any number of objects in it. In equation form, the conservation of momentum principle for an isolated system is written

$$p_{tot} = constant, \tag{4.12.7}$$

or

$$p_{tot} = p_{tot}, \tag{4.12.8}$$

where $p_{tot}$ is the total momentum (the sum of the momenta of the individual objects in the system) and $p_{tot}$, is the total momentum some time later. (The total momentum can be shown to be the momentum of the center of mass of the system.) An isolated system is defined to be one for which the net external force is zero $(F_{net} = 0)$.

## Conservation of Momentum Principle

$$p_{tot} = constant \tag{4.12.9}$$

$$p_{tot} = p'_{tot}\,(isolated\,system) \tag{4.12.10}$$

## Isolated System

An isolated system is defined to be one for which the net external force is zero $(F_{net} = 0)$.

Perhaps an easier way to see that momentum is conserved for an isolated system is to consider Newton's second law in terms of momentum, $F_{net} = \frac{\Delta p_{tot}}{\Delta t}$. For an isolated system, $(F_{net} = 0)$; thus $\Delta p_{tot} = 0$ and $\Delta p$ is constant.

We have noted that the three length dimensions in nature x, y and z are independent, and it is interesting to note that momentum can be conserved in different ways along each dimension. For example, during projectile motion and where air resistance is negligible, momentum is conserved in the horizontal direction because horizontal forces are zero and momentum is unchanged. But along the vertical direction, the net vertical force is not zero and the momentum of the projectile is not conserved (Figure 4.12.2). However, if the momentum of the projectile-Earth system is considered in the vertical direction, we find that the total momentum is conserved.

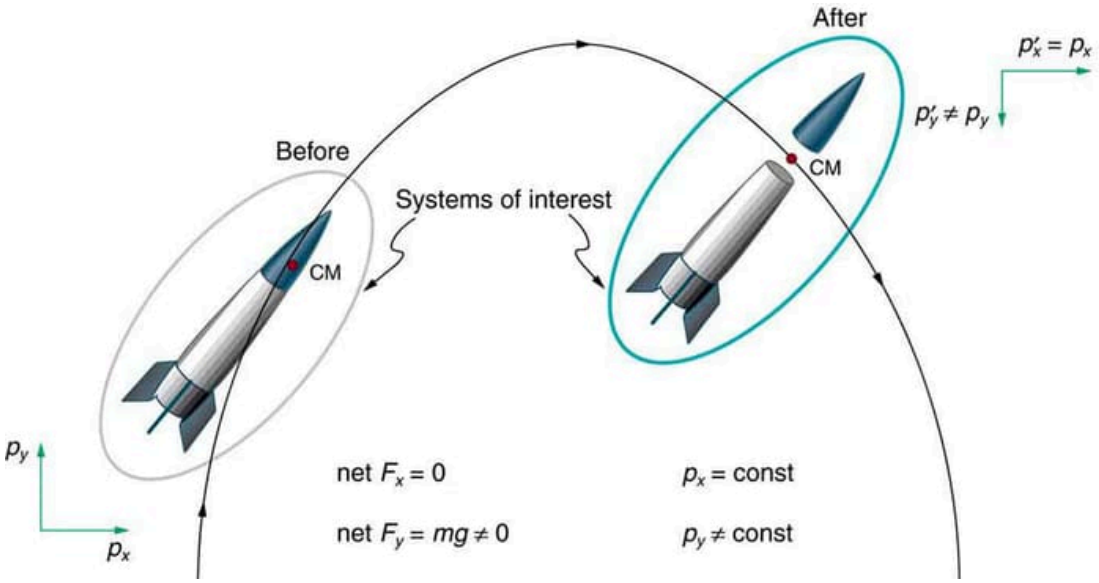



Figure 4.12.2: The horizontal component of a projectile's momentum is conserved if air resistance is negligible, even in this case where a space probe separates. The forces causing the separation are internal to the system, so that the net external horizontal force $F_{x-net}$ is still zero. The vertical component of the momentum is not conserved, because the net vertical force $F_{y-net}$ is not zero. In the vertical direction, the space probe-Earth system needs to be considered and we find that the total momentum is conserved. The center of mass of the space probe takes the same path it would if the separation did not occur.

The conservation of momentum principle can be applied to systems as different as a comet striking Earth and a gas containing huge numbers of atoms and molecules. Conservation of momentum is violated only when the net external force is not zero. But another larger system can always be considered in which momentum is conserved by simply including the source of the external force. For example, in the collision of two cars considered above, the two-car system conserves momentum while each one-car system does not.

## MAKING CONNECTIONS: TAKE-HOME Investigation—Drop of Tennis Ball and a Basketball

Hold a tennis ball side by side and in contact with a basketball. Drop the balls together. (Be careful!) What happens? Explain your observations. Now hold the tennis ball above and in contact with the basketball. What happened? Explain your observations. What do you think will happen if the basketball ball is held above and in contact with the tennis ball?

## MAKING CONNECTIONS: TAKE-HOME Investigation—Two Tennis Balls in a Ballistic Trajectory

Tie two tennis balls together with a string about a foot long. Hold one ball and let the other hang down and throw it in a ballistic trajectory. Explain your observations. Now mark the center of the string with bright ink or attach a brightly colored sticker to it and throw again. What happened? Explain your observations.

Some aquatic animals such as jellyfish move around based on the principles of conservation of momentum. A jellyfish fills its umbrella section with water and then pushes the water out resulting in motion in the opposite direction to that of the jet of water. Squids propel themselves in a similar manner but, in contrast with jellyfish, are able to control the direction in which they move by aiming their nozzle forward or backward. Typical squids can move at speeds of 8 to 12 km/h.

The ballistocardiograph (BCG) was a diagnostic tool used in the second half of the 20th century to study the strength of the heart. About once a second, your heart beats, forcing blood into the aorta. A force in the opposite direction is exerted on the rest of your body (recall Newton's third law). A ballistocardiograph is a device that can measure this reaction force. This measurement is done by using a sensor (resting on the person) or by using a moving table suspended from the ceiling. This technique can gather information on the strength of the heart beat and the volume of blood passing from the heart. However, the electrocardiogram (ECG or EKG) and the echocardiogram (cardiac ECHO or ECHO; a technique that uses ultrasound to see an image of the heart) are more widely used in the practice of cardiology.

**Making Connections: Conservation of Momentum and Collision**

Conservation of momentum is quite useful in describing collisions. Momentum is crucial to our understanding of atomic and subatomic particles because much of what we know about these particles comes from collision experiments.

## Subatomic Collisions and Momentum

The conservation of momentum principle not only applies to the macroscopic objects, it is also essential to our explorations of atomic and subatomic particles. Giant machines hurl subatomic particles at one another, and researchers evaluate the results by assuming conservation of momentum (among other things).

On the small scale, we find that particles and their properties are invisible to the naked eye but can be measured with our instruments, and models of these subatomic particles can be constructed to describe the results. Momentum is found to be a property of all subatomic particles including massless particles such as photons that compose light. Momentum being a property of particles hints that momentum may have an identity beyond the description of an object's mass multiplied by the object's velocity. Indeed, momentum relates to wave properties and plays a fundamental role in what measurements are taken and how we take these measurements. Furthermore, we find that the conservation of momentum principle is valid when considering systems of particles. We use this principle to analyze the masses and other properties of previously undetected particles, such as the nucleus of an atom and the existence of quarks that make up particles of nuclei. Figure 4.12.3 below illustrates how a particle scattering backward from another implies that its target is massive and dense. Experiments seeking evidence that quarks make up protons (one type of particle that makes up nuclei) scattered high-energy electrons off of protons (nuclei of hydrogen atoms). Electrons occasionally scattered straight backward in a manner that implied a very small and very dense particle makes up the proton—this observation is considered nearly direct evidence of quarks. The analysis was based partly on the same conservation of momentum principle that works so well on the large scale.

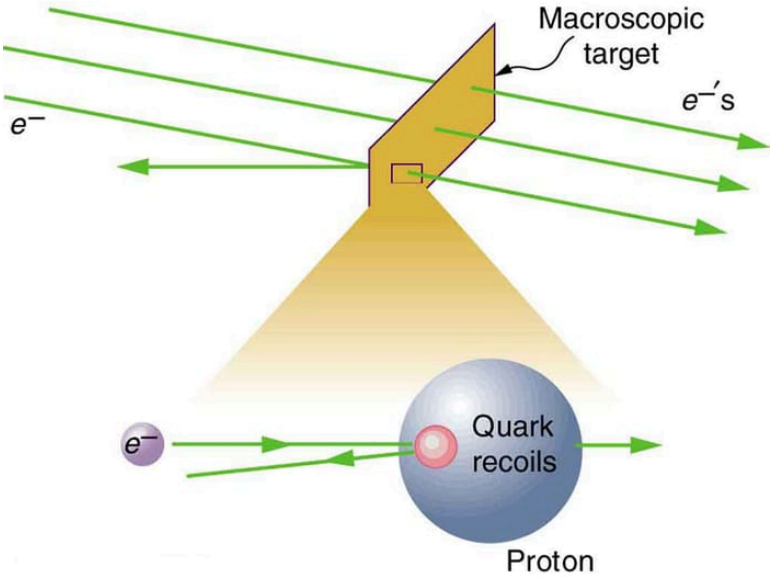



Figure 4.12.3: A subatomic particle scatters straight backward from a target particle. In experiments seeking evidence for quarks, electrons were observed to occasionally scatter straight backward from a proton.

## Summary

- The conservation of momentum principle is written

$$p_{tot} = constant \tag{4.12.11}$$

or

$$p_{tot} = p'_{tot}\,(isolated\,system), \tag{4.12.12}$$

- $p_{tot}$ is the initial total momentum and $p'_{tot}$ is the total momentum some time later. An isolated system is defined to be one for which the net external force is zero $(F_{net} = 0$
- During projectile motion and where air resistance is negligible, momentum is conserved in the horizontal direction because horizontal forces are zero.
- Conservation of momentum applies only when the net external force is zero.
- The conservation of momentum principle is valid when considering systems of particles.

## Glossary

**conservation of momentum principle**
: when the net external force is zero, the total momentum of the system is conserved or constant

**isolated system**
: a system in which the net external force is zero

**quark**
: fundamental constituent of matter and an elementary particle

This page titled 4.12: Conservation of Momentum is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 8.3: Conservation of Momentum by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 4.13: Energy and Momentum

**Learning Objectives**

- Understand the differences between energy and momentum.

Energy and momentum have many similarities as well as many differences. These similarities and differences were confounding for the scientists of the 1700s and 1800s who were trying to formulate ideas about how the movements of masses followed basic repeatable principles.

Momentum and energy can both be formulated around ideas of mass and velocity. However, the formulation of energy is based on mass times velocity squared while the formulation of momentum is based on mass times velocity. Also, energy is not always based on movement while momentum is always based on movement. We have already seen how potential energy can be calculated for objects which are not moving. As you can probably imagine, there is no component of movement involved in most of the other forms of energy we have not considered in great detail yet.

The other similarity between momentum and energy is that they are both conserved. This is an important similarity, as there are very few properties in nature that are conserved. However, these two are conserved for different reasons and in different ways. Momentum is a vector quantity, while energy is a scalar quantity (an important type of scalar quantity called a state function, as we will learn later). As a vector quantity, momentum could cancel out during a collision. For example, if two cars of equal mass were moving at the same speed directly towards each other we might expect them to stop moving when they collide. In this case the momentum of the system defined as both cars is conserved because there overall momentum adds up to zero both before and after the collision. However, the energy does not cancel out. The energy of motion has been converted into potential energy as the bumpers are compressed and into thermal energy as they warm up. It does not cancel out as momentum does.

## Section Summary

- Momentum is based on mass and velocity, and energy can be formulated to also be based on mass and velocity.
- Momentum and energy are both conserved.
- Momentum is a vector property, and energy is a scalar property.
- Momentum and energy are different properties, although they have sometimes been confused with each other.

This page titled 4.13: Energy and Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

# 5: MODULE 4-THERMODYNAMICS AND WAVES


5.1: Prelude to Temperature, Kinetic Theory, and the Gas Laws

5.2: Temperature

5.3: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature

5.4: Phase Changes

5.5: Prelude to Heat and Heat Transfer Methods

5.6: Heat

5.7: Temperature Change and Heat Capacity

5.8: Phase Change and Latent Heat

5.9: Heat Transfer Methods

5.10: Conduction

5.11: Convection

5.12: Radiation

5.13: Prelude to Thermodynamics

5.14: The First Law of Thermodynamics

5.15: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency

5.16: Waves

5.17: Superposition and Interference

5.18: Prelude to the Physics of Hearing

5.19: Sound

5.20: Speed of Sound, Frequency, and Wavelength

5.21: Doppler Effect and Sonic Booms

5.22: Prelude to Electromagnetic Waves

5.23: The Electromagnetic Spectrum



This page titled 5: Module 4-Thermodynamics and Waves is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

# 5.1: Prelude to Temperature, Kinetic Theory, and the Gas Laws

Heat is something familiar to each of us. We feel the warmth of the summer Sun, the chill of a clear summer night, the heat of coffee after a winter stroll, and the cooling effect of our sweat. Heat transfer is maintained by temperature differences. Manifestations of heat transfer—the movement of heat energy from one place or material to another—are apparent throughout the universe. Heat from beneath Earth's surface is brought to the surface in flows of incandescent lava. The Sun warms Earth's surface and is the source of much of the energy we find on it. Rising levels of atmospheric carbon dioxide threaten to trap more of the Sun's energy, perhaps fundamentally altering the ecosphere. In space, supernovas explode, briefly radiating more heat than an entire galaxy does.



Figure 5.1.1:The welder's gloves and helmet protect him from the electric arc that transfers enough thermal energy to melt the rod, spray sparks, and burn the retina of an unprotected eye. The thermal energy can be felt on exposed skin a few meters away, and its light can be seen for kilometers. (credit: Kevin S. O'Brien/U.S. Navy)

What is heat? How do we define it? How is it related to temperature? What are heat's effects? How is it related to other forms of energy and to work? We will find that, in spite of the richness of the phenomena, there is a small set of underlying physical principles that unite the subjects and tie them to other fields.

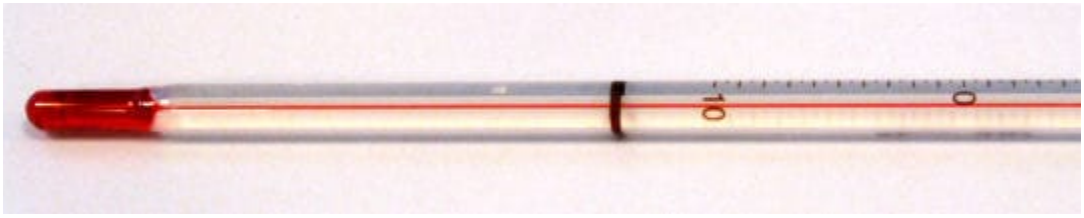



Figure 5.1.2: In a typical thermometer like this one, the alcohol, with a red dye, expands more rapidly than the glass containing it. When the thermometer's temperature increases, the liquid from the bulb is forced into the narrow tube, producing a large change in the length of the column for a small change in temperature. (credit: Chemical Engineer, Wikimedia Commons)

This page titled 5.1: Prelude to Temperature, Kinetic Theory, and the Gas Laws is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 13.0: Prelude to Temperature, Kinetic Theory, and the Gas Laws by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.2: Temperature

## Learning Objectives

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

- Define temperature.
- Convert temperatures between the Celsius, Fahrenheit, and Kelvin scales.
- Define thermal equilibrium.
- State the zeroth law of thermodynamics.

The concept of temperature has evolved from the common concepts of hot and cold. Human perception of what feels hot or cold is a relative one. For example, if you place one hand in hot water and the other in cold water, and then place both hands in tepid water, the tepid water will feel cool to the hand that was in hot water, and warm to the one that was in cold water. The scientific definition of temperature is less ambiguous than your senses of hot and cold. Temperature is operationally defined to be what we measure with a thermometer. (Many physical quantities are defined solely in terms of how they are measured. We shall see later how temperature is related to the kinetic energies of atoms and molecules, a more physical explanation.) Two accurate thermometers, one placed in hot water and the other in cold water, will show the hot water to have a higher temperature. If they are then placed in the tepid water, both will give identical readings (within measurement uncertainties). In this section, we discuss temperature, its measurement by thermometers, and its relationship to thermal equilibrium. Again, temperature is the quantity measured by a thermometer.

## Misconception Alert: Human Perception vs. Reality

On a cold winter morning, the wood on a porch feels warmer than the metal of your bike. The wood and bicycle are in thermal equilibrium with the outside air, and are thus the same temperature. They *feel* different because of the difference in the way that they conduct heat away from your skin. The metal conducts heat away from your body faster than the wood does (see more about conductivity in Conduction). This is just one example demonstrating that the human sense of hot and cold is not determined by temperature alone.

Another factor that affects our perception of temperature is humidity. Most people feel much hotter on hot, humid days than on hot, dry days. This is because on humid days, sweat does not evaporate from the skin as efficiently as it does on dry days. It is the evaporation of sweat (or water from a sprinkler or pool) that cools us off.

Any physical property that depends on temperature, and whose response to temperature is reproducible, can be used as the basis of a thermometer. Because many physical properties depend on temperature, the variety of thermometers is remarkable. For example, volume increases with temperature for most substances. This property is the basis for the common alcohol thermometer, the old mercury thermometer, and the bimetallic strip (Figure). Other properties used to measure temperature include electrical resistance and color, as shown in Figure, and the emission of infrared radiation, as shown in Figure.

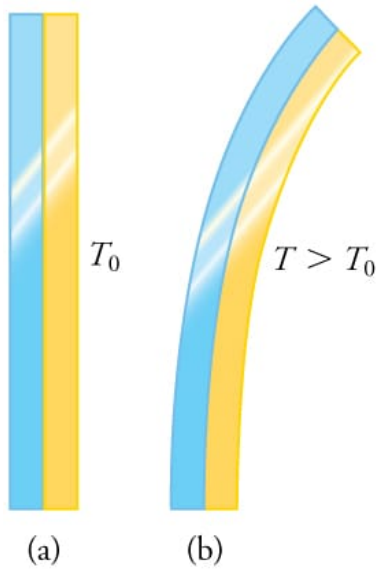



Figure 5.2.1: The curvature of a bimetallic strip depends on temperature. (a) The strip is straight at the starting temperature, where its two components have the same length. (b) At a higher temperature, this strip bends to the right, because the metal on the left has expanded more than the metal on the right.




Figure 5.2.2: Each of the six squares on this plastic (liquid crystal) thermometer contains a film of a different heat-sensitive liquid crystal material. Below $95^o$ $F$ all six squares are black. When the plastic thermometer is exposed to temperature that increases to $95^o$ $F$ the first liquid crystal square changes color. When the temperature increases above $96.8^o$ $F$ the second liquid crystal square also changes color, and so forth. (credit: Arkrishna, Wikimedia Commons)

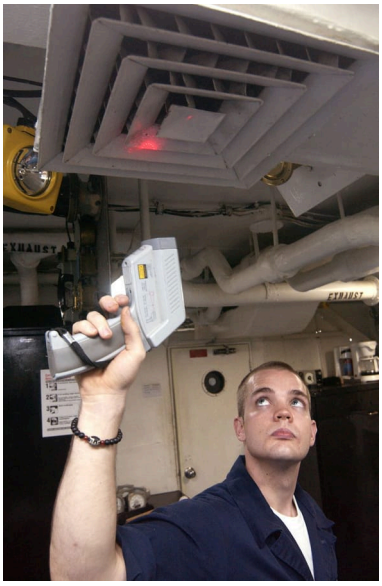


Figure 5.2.3: Fireman Jason Ormand uses a pyrometer to check the temperature of an aircraft carrier's ventilation system. Infrared radiation (whose emission varies with temperature) from the vent is measured and a temperature readout is quickly produced. Infrared measurements are also frequently used as a measure of body temperature. These modern thermometers, placed in the ear canal, are more accurate than alcohol thermometers placed under the tongue or in the armpit. (credit: Lamel J. Hinton/U.S. Navy)

# Temperature Scales

Thermometers are used to measure temperature according to well-defined scales of measurement, which use pre-defined reference points to help compare quantities. The three most common temperature scales are the Fahrenheit, Celsius, and Kelvin scales. A temperature scale can be created by identifying two easily reproducible temperatures. The freezing and boiling temperatures of water at standard atmospheric pressure are commonly used.

The Celsius scale (which replaced the slightly different *centigrade* scale) has the freezing point of water at $0^o\ C$ and the boiling point at $100^o\ C$. Its unit is the degree Celsius $(^oC)$. On the Fahrenheit scale (still the most frequently used in the United States), the freezing point of water is at $37^o\ F$ and the boiling point is at $212^o\ F$. The unit of temperature on this scale is the degree Fahrenheit $(^oF)$. Note that a temperature difference of one degree Celsius is greater than a temperature difference of one degree Fahrenheit. Only 100 Celsius degrees span the same range as 180 Fahrenheit degrees, thus one degree on the Celsius scale is 1.8 times larger than one degree on the Fahrenheit scale $180/100 = 9/5$.

The Kelvin scale is the temperature scale that is commonly used in science. It is an *absolute temperature* scale defined to have 0 K at the lowest possible temperature, called absolute zero. The official temperature unit on this scale is the *kelvin*, which is abbreviated K, and is not accompanied by a degree sign. The freezing and boiling points of water are 273.15 K and 373.15 K, respectively. Thus, the magnitude of temperature differences is the same in units of kelvins and degrees Celsius. Unlike other temperature scales, the Kelvin scale is an absolute scale. It is used extensively in scientific work because a number of physical quantities, such as the volume of an ideal gas, are directly related to absolute temperature. The kelvin is the SI unit used in scientific work.

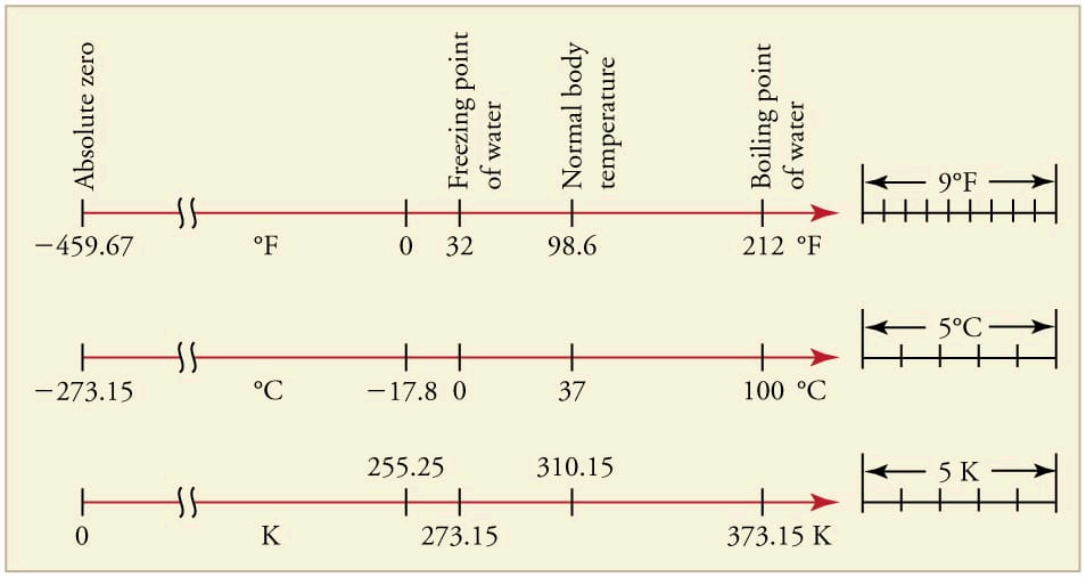



Figure 5.2.4: Relationships between the Fahrenheit, Celsius, and Kelvin temperature scales, rounded to the nearest degree. The relative sizes of the scales are also shown.

The relationships between the three common temperature scales is shown in Figure. Temperatures on these scales can be converted using the equations in Table.

| To convert from . . . | Use this equation . . . | Also written as . . . |
|---|---|---|
| Celsius to Fahrenheit | $T(^oF) = \frac{9}{5}T(C^o) + 32$ | $T_{^oF} = \frac{9}{5}T_{^oC} + 32$ |
| Fahrenheit to Celsius | $T(^oC) = \frac{9}{5}(T(F^o) - 32)$ | $T_{^oC} = \frac{5}{9}(T_{^oF} - 32)$ |

| To convert from . . . | Use this equation . . . | Also written as . . . |
|---|---|---|
| Celsius to Kelvin | $T(K) = T(^oC) + 273.15$ | $T_K = T_{^oC} + 273.15$ |
| Kelvin to Celsius | $T^oC) = T(K) - 273.15$ | $T_{^oC} = T_K - 273.15$ |
| Fahrenheit to Kelvin | $T(K) = \frac{5}{9}(T(^oF) - 32) + 273.15$ | $T_K = \frac{5}{9}(T_{^oF} - 32) + 273.15$ |
| Kelvin to Fahrenheit | $T(^oF) = \frac{9}{5}(T(K) - 273.15) + 32$ | $T_{^oF} = \frac{9}{5}(T_K - 273.15) + 32$ |

Notice that the conversions between Fahrenheit and Kelvin look quite complicated. In fact, they are simple combinations of the conversions between Fahrenheit and Celsius, and the conversions between Celsius and Kelvin.

## Example 5.2.1: Converting between Temperature Scales: Room Temperature

"Room temperature" is generally defined to be $25^oC$ (a) What is room temperature in $^oF$? (b) What is it in K?

**Strategy**

To answer these questions, all we need to do is choose the correct conversion equations and plug in the known values.

**Solution for (a)**

1. Choose the right equation. To convert from $^oC$ to $^oF$, use this equation

$$T_{^oF} = \frac{9}{5}T_{^oC} + 32. \tag{5.2.1}$$

2. Plug the known value into the equation and solve:

$$T_{^oF} = \frac{9}{5}25^oC + 32 = 77^oF. \tag{5.2.2}$$

**Solution for (b)**

1. Choose the right equation. To convert from $^oC$ to $K$, use this equation

$$T_K = T_{^oC} + 273.15. \tag{5.2.3}$$

2. Plug the known value into the equation and solve:

$$T_K = 25^oC + 273.15 = 298\,K. \tag{5.2.4}$$

## Example 5.2.2: Plug the known value into the equation and solve:

The Reaumur scale is a temperature scale that was used widely in Europe in the 18th and 19th centuries. On the Reaumur temperature scale, the freezing point of water is $0^oR$ and the boiling temperature is $80^oR$. If "room temperature" is $25^oC$

on the Celsius scale, what is it on the Reaumur scale?

**Strategy**

To answer this question, we must compare the Reaumur scale to the Celsius scale. The difference between the freezing point and boiling point of water on the Reaumur scale is $80^oR$. On the Celsius scale it is $100^oC$. Therefore $100^oC = 80^oR$. Both scales start at $0^o$ for freezing, so we can derive a simple formula to convert between temperatures on the two scales.

**Solution**

1. Derive a formula to convert from one scale to the other:

$$T_{^oR} = \frac{0.8^oR}{^oC} \times T_{^oC}. \tag{5.2.5}$$

2. Plug the known value into the equation and solve:

$$T_{^oR} = \frac{0.8^oR}{^oC} \times 25^oC = 20^oR. \tag{5.2.6}$$

## Temperature Ranges in the Universe

Figure shows the wide range of temperatures found in the universe. Human beings have been known to survive with body temperatures within a small range, from $24^oC$ to $44^oC$ $(75^oF$ to $111^oF)$. The average normal body temperature is usually given as $37.0^oC$ $(98.6^oF)$, and variations in this temperature can indicate a medical condition: a fever, an infection, a tumor, or circulatory problems (see Figure).

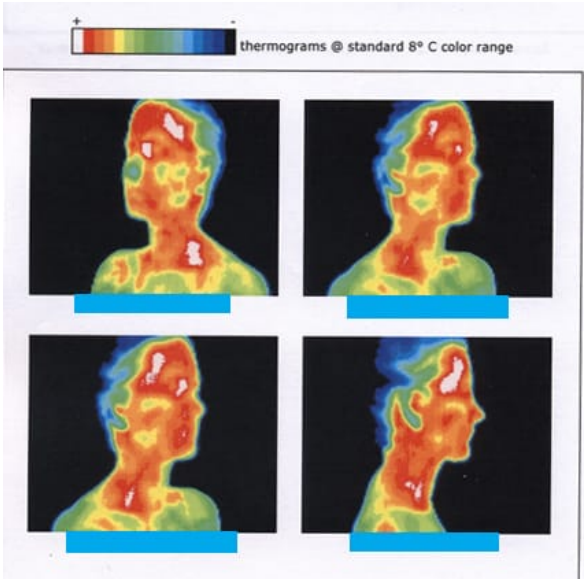



Figure $5.2.5$:This image of radiation from a person's body (an infrared thermograph) shows the location of temperature abnormalities in the upper body. Dark blue corresponds to cold areas and red to white corresponds to hot areas. An elevated temperature might be an indication of malignant tissue (a cancerous tumor in the breast, for example), while a depressed temperature might be due to a decline in blood flow from a clot. In this case, the abnormalities are caused by a condition called hyperhidrosis. (credit: Porcelina81, Wikimedia Commons)

The lowest temperatures ever recorded have been measured during laboratory experiments: $4.5 \times 10^{-10}K$ at the Massachusetts Institute of Technology (USA), and $1.0 \times 10^{-10}K$ at Helsinki University of Technology (Finland). In comparison, the coldest recorded place on Earth's surface is Vostok, Antarctica at 183 K $(-89^oC)$ and the coldest place (outside the lab) known in the universe is the Boomerang Nebula, with a temperature of 1 K.

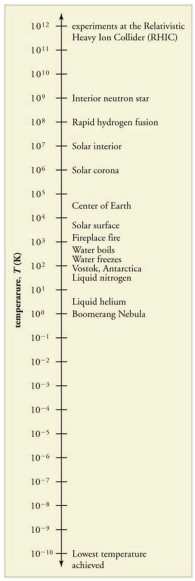



Figure 5.2.6: Each increment on this logarithmic scale indicates an increase by a factor of ten, and thus illustrates the tremendous range of temperatures in nature. Note that zero on a logarithmic scale would occur off the bottom of the page at infinity.

### Making Connections: Absolute Zero

What is absolute zero? Absolute zero is the temperature at which all molecular motion has ceased. The concept of absolute zero arises from the behavior of gases. Figure shows how the pressure of gases at a constant volume decreases as temperature decreases. Various scientists have noted that the pressures of gases extrapolate to zero at the same temperature, $-273.15^oC$. This extrapolation implies that there is a lowest temperature. This temperature is called *absolute zero*. Today we know that most gases first liquefy and then freeze, and it is not actually possible to reach absolute zero. The numerical value of absolute zero temperature is $-273.15^oC$ or 0K.

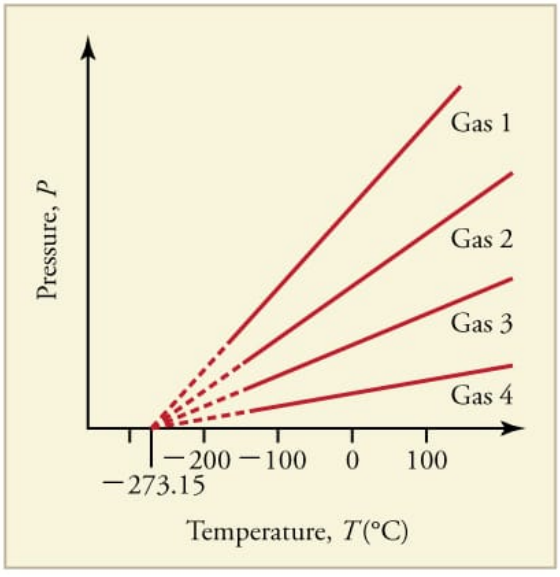



Figure 5.2.7: Graph of pressure versus temperature for various gases kept at a constant volume. Note that all of the graphs extrapolate to zero pressure at the same temperature.

## Thermal Equilibrium and the Zeroth Law of Thermodynamics

Thermometers actually take their *own* temperature, not the temperature of the object they are measuring. This raises the question of how we can be certain that a thermometer measures the temperature of the object with which it is in contact. It is based on the fact that any two systems placed in *thermal contact* (meaning heat transfer can occur between them) will reach the

same temperature. That is, heat will flow from the hotter object to the cooler one until they have exactly the same temperature. The objects are then in thermal equilibrium, and no further changes will occur. The systems interact and change because their temperatures differ, and the changes stop once their temperatures are the same. Thus, if enough time is allowed for this transfer of heat to run its course, the temperature a thermometer registers *does* represent the system with which it is in thermal equilibrium. Thermal equilibrium is established when two bodies are in contact with each other and can freely exchange energy.

Furthermore, experimentation has shown that if two systems, A and B, are in thermal equilibrium with each another, and B is in thermal equilibrium with a third system C, then A is also in thermal equilibrium with C. This conclusion may seem obvious, because all three have the same temperature, but it is basic to thermodynamics. It is called the zeroth **law of thermodynamics**.

### The Zeroth Law of Thermodynamics

If two systems, A and B, are in thermal equilibrium with each other, and B is in thermal equilibrium with a third system, C, then A is also in thermal equilibrium with C.

This law was postulated in the 1930s, after the first and second laws of thermodynamics had been developed and named. It is called the *zeroth law* because it comes logically before the first and second laws (discussed in Thermodynamics). An example of this law in action is seen in babies in incubators: babies in incubators normally have very few clothes on, so to an observer they look as if they may not be warm enough. However, the temperature of the air, the cot, and the baby is the same, because they are in thermal equilibrium, which is accomplished by maintaining air temperature to keep the baby comfortable.

### Check Your Understanding

Does the temperature of a body depend on its size

[Hide Solution]

No, the system can be divided into smaller parts each of which is at the same temperature. We say that the temperature is an *intensive* quantity. Intensive quantities are independent of size.

## Section Summary

- Temperature is the quantity measured by a thermometer.
- Temperature is related to the average kinetic energy of atoms and molecules in a system.
- Absolute zero is the temperature at which there is no molecular motion.
- There are three main temperature scales: Celsius, Fahrenheit, and Kelvin.
- Temperatures on one scale can be converted to temperatures on another scale using the following equations:

$$T_{^\circ F} = \frac{9}{5}T_{^\circ C} + 32 \tag{5.2.7}$$

$$T_{^\circ C} = \frac{5}{9}T_{^\circ F} - 32) \tag{5.2.8}$$

$$T_K = T_{^\circ C} + 273.15 \tag{5.2.9}$$

$$T_{^\circ C} = T_K - 273.15 \tag{5.2.10}$$

- Systems are in thermal equilibrium when they have the same temperature.
- Thermal equilibrium occurs when two bodies are in contact with each other and can freely exchange energy.
- The zeroth law of thermodynamics states that when two systems, A and B, are in thermal equilibrium with each other, and B is in thermal equilibrium with a third system, C, then A is also in thermal equilibrium with C.

## Glossary

**temperature**
: the quantity measured by a thermometer

**Celsius scale**
temperature scale in which the freezing point of water is **0°C** and the boiling point of water is **100°C**

**degree Celsius**
unit on the Celsius temperature scale

**Fahrenheit scale**
temperature scale in which the freezing point of water is **32°F** and the boiling point of water is **212°F**

**degree Fahrenheit**
unit on the Fahrenheit temperature scale

**Kelvin scale**
temperature scale in which 0 K is the lowest possible temperature, representing absolute zero

**absolute zero**
the lowest possible temperature; the temperature at which all molecular motion ceases

**thermal equilibrium**
the condition in which heat no longer flows between two objects that are in contact; the two objects have the same temperature

**zeroth law of thermodynamics**
law that states that if two objects are in thermal equilibrium, and a third object is in thermal equilibrium with one of those objects, it is also in thermal equilibrium with the other object

This page titled 5.2: Temperature is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 13.1: Temperature by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.3: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature

## Learning Objectives

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

- Express the ideal gas law in terms of molecular mass and velocity.
- Define thermal energy.
- Calculate the kinetic energy of a gas molecule, given its temperature.
- Describe the relationship between the temperature of a gas and the kinetic energy of atoms and molecules.
- Describe the distribution of speeds of molecules in a gas.

We have developed macroscopic definitions of pressure and temperature. Pressure is the force divided by the area on which the force is exerted, and temperature is measured with a thermometer. We gain a better understanding of pressure and temperature from the kinetic theory of gases, which assumes that atoms and molecules are in continuous random motion.

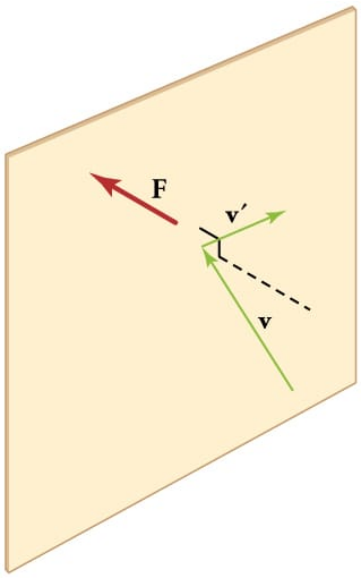



Figure 5.3.1: When a molecule collides with a rigid wall, the component of its momentum perpendicular to the wall is reversed. A force is thus exerted on the wall, creating pressure.

Figure shows an elastic collision of a gas molecule with the wall of a container, so that it exerts a force on the wall (by Newton's third law). Because a huge number of molecules will collide with the wall in a short time, we observe an average force per unit area. These collisions are the source of pressure in a gas. As the number of molecules increases, the number of collisions and thus the pressure increase. Similarly, the gas pressure is higher if the average velocity of molecules is higher. The actual relationship is derived in the Things Great and Small feature below. The following relationship is found:

$$PV = \frac{1}{3}Nm\overline{v^2}, \tag{5.3.1}$$

where $P$ is the pressure (average force per unit area), $V$ is the volume of gas in the container, $N$ is the number of molecules in the container, $m$ is the mass of a molecule, and $\overline{v^2}$ is the average of the molecular speed squared.

What can we learn from this atomic and molecular version of the ideal gas law? We can derive a relationship between temperature and the average translational kinetic energy of molecules in a gas. Recall the previous expression of the ideal gas law:

$$PV = NkT. \tag{5.3.2}$$

Equating the right-hand side of this equation with the right-hand side of $PV = \frac{1}{3}Nmv^2$ gives

$$\frac{1}{3}Nm\overline{v^2} = NkT. \tag{5.3.3}$$

## Making Connections: Things Great and Small—Atomic and Molecular Origin of Pressure in a Gas

Figure shows a box filled with a gas. We know from our previous discussions that putting more gas into the box produces greater pressure, and that increasing the temperature of the gas also produces a greater pressure. But why should increasing the temperature of the gas increase the pressure in the box? A look at the atomic and molecular scale gives us some answers, and an alternative expression for the ideal gas law.

The figure shows an expanded view of an elastic collision of a gas molecule with the wall of a container. Calculating the average force exerted by such molecules will lead us to the ideal gas law, and to the connection between temperature and molecular kinetic energy. We assume that a molecule is small compared with the separation of molecules in the gas, and that its interaction with other molecules can be ignored. We also assume the wall is rigid and that the molecule's direction changes, but that its speed remains constant (and hence its kinetic energy and the magnitude of its momentum remain constant as well). This assumption is not always valid, but the same result is obtained with a more detailed description of the molecule's exchange of energy and momentum with the wall.

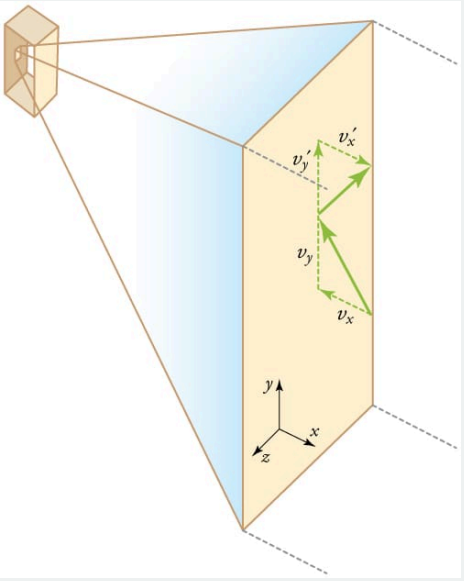



Figure 5.3.2: Gas in a box exerts an outward pressure on its walls. A molecule colliding with a rigid wall has the direction of its velocity and momentum in the x- direction reversed. This direction is perpendicular to the wall. The components of its velocity momentum in the y- and z- directions are not changed, which means there is no force parallel to the wall.

If the molecule's velocity changes in the x-direction, its momentum changes from $-mv_x$ to $+mv_x$. Thus, its change in momentum is $\Delta mv = +mv_x - (-mv_x) = 2mv_x$. The force exerted on the molecule is given by

$$F = \frac{\Delta p}{\Delta t} = \frac{2mv_x}{\Delta t}. \tag{5.3.4}$$

There is no force between the wall and the molecule until the molecule hits the wall. During the short time of the collision, the force between the molecule and wall is relatively large. We are looking for an average force; we take $\Delta t$ to be the average time between collisions of the molecule with this wall. It is the time it would take the molecule to go across the box and back (a distance $2l$) at a speed of $v_x$. Thus $\Delta t = 2l/v_x$, and the expression for the force becomes

$$F = \frac{2mv_x}{2l/v_x} = \frac{mv_x^2}{l}. \tag{5.3.5}$$

This force is due to *one* molecule. We multiply by the number of molecules $N$ and use their average squared velocity to find the force

$$F = N\frac{m\overline{v_x^2}}{l}, \tag{5.3.6}$$

where the bar over a quantity means its average value. We would like to have the force in terms of the speed $v$, rather than the x-component of the velocity. We note that the total velocity squared is the sum of the squares of its x-components, so that

$$\overline{v^2} = \overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2}. \tag{5.3.7}$$

Because the velocities are random, their average components in all directions are the same:

$$\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2} \tag{5.3.8}$$

Thus,

$$\overline{v^2} = 3\overline{v_x^2}, \tag{5.3.9}$$

or

$$\overline{v_x^2} = \frac{1}{3}\overline{v^2}. \tag{5.3.10}$$

Substituting $\frac{1}{3}\overline{v^2}$ into the expression for $F$ gives

$$F = N\frac{m\overline{v^2}}{3l}. \tag{5.3.11}$$

The pressure is $F/A$, so that we obtain

$$P = \frac{F}{A} = N\frac{m\overline{v^2}}{3Al} = \frac{1}{3}\frac{Nm\overline{v^2}}{V}, \tag{5.3.12}$$

where we used $V = Al$ for the volume. This gives the important result.

$$PV = \frac{1}{3}Nm\overline{v^2}. \tag{5.3.13}$$

This equation is another expression of the ideal gas law.

We can get the average kinetic energy of a molecule, $\frac{1}{2}mv^2$, from the right-hand side of the equation by canceling $N$ and multiplying by 3/2. This calculation produces the result that the average kinetic energy of a molecule is directly related to absolute temperature.

$$\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{5.3.14}$$

The average translational kinetic energy of a molecule, $\overline{KE}$, is called thermal energy**.** The equation $\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT$

is a molecular interpretation of temperature, and it has been found to be valid for gases and reasonably accurate in liquids and solids. It is another definition of temperature based on an expression of the molecular energy.

It is sometimes useful to rearrange $\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT$ and solve for the average speed of molecules in a gas in terms of temperature,

$$\sqrt{\overline{v^2}} = v_{rms} = \sqrt{\frac{3kT}{m}}, \tag{5.3.15}$$

where $v_{rms}$ stands for root-mean-square (rms) speed.

## Example 5.3.1: Calculating Kinetic Energy and Speed of a Gas Molecule

(a) What is the average kinetic energy of a gas molecule at $20^oC$ (room temperature)? (b) Find the rms speed of a nitrogen molecule $(N_2)$

at this temperature.

**Strategy for (a)**

The known in the equation for the average kinetic energy is the temperature.

$$\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{5.3.16}$$

Before substituting values into this equation, we must convert the given temperature to kelvins. This conversion gives $T = (20.0 + 273)k = 293\ K$.

**Solution for (a)**

The temperature alone is sufficient to find the average translational kinetic energy. Substituting the temperature into the translational kinetic energy equation gives

$$\overline{KE} = \frac{3}{2}kT = \frac{3}{2}(1.38 \times 10^{-23}\ J/K)(293\ K) = 6.07 \times 10^{-21}\ J. \tag{5.3.17}$$

**Strategy for (b)**

Finding the rms speed of a nitrogen molecule involves a straightforward calculation using the equation

$$\sqrt{\overline{v^2}} = v_{rms} = \sqrt{\frac{3kT}{m}}, \tag{5.3.18}$$

but we must first find the mass of a nitrogen molecule. Using the molecular mass of nitrogen $N_2$ from the periodic table,

$$m = \frac{2(14.0067) \times 10^{-3}\ kg/mol}{6.02 \times 10^{23}\ mol^{-1}} = 4.65 \times 10^{-26}\ kg. \tag{5.3.19}$$

**Solution for (b)**

Substituting this mass and the value for $k$ into the equation for $v_{rms}$ yields

$$v_{rms} = \sqrt{\frac{3(1.38 \times 10^{-23}\ J/K)(293)}{4.65 \times 10^{-26}\ kg}} = 511\ m/s. \tag{5.3.20}$$

**Discussion**

Note that the average kinetic energy of the molecule is independent of the type of molecule. The average translational kinetic energy depends only on absolute temperature. The kinetic energy is very small compared to macroscopic energies, so that we do not feel when an air molecule is hitting our skin. The rms velocity of the nitrogen molecule is surprisingly large. These large molecular velocities do not yield macroscopic movement of air, since the molecules move in all directions with equal likelihood. The *mean free path* (the distance a molecule can move on average between collisions) of molecules in air is very small, and so the molecules move rapidly but do not get very far in a second. The high value for rms speed is reflected in the speed of sound, however, which is about 340 m/s at room temperature. The faster the rms speed of air molecules, the faster that sound vibrations can be transferred through the air. The speed of sound increases with temperature and is greater in gases with small molecular masses, such as helium. (See Figure.)

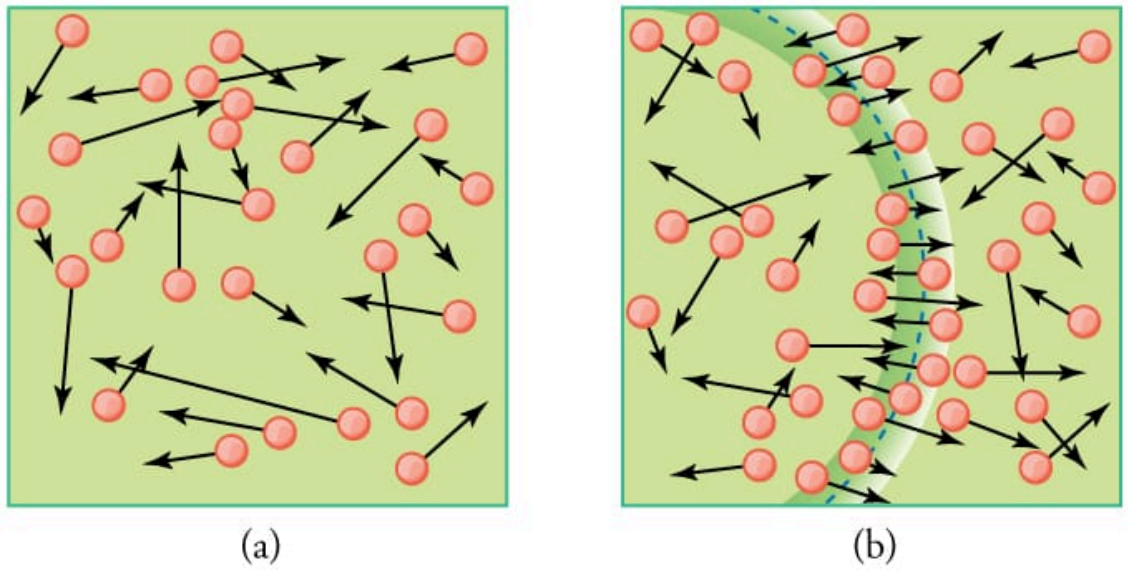



Figure 5.3.3: (a) There are many molecules moving so fast in an ordinary gas that they collide a billion times every second. (b) Individual molecules do not move very far in a small amount of time, but disturbances like sound waves are transmitted at speeds related to the molecular speeds.

**Making Connections: Historical Note—Kinetic Theory of Gases**

- The kinetic theory of gases was developed by Daniel Bernoulli (1700–1782), who is best known in physics for his work on fluid flow (hydrodynamics). Bernoulli's work predates the atomistic view of matter established by Dalton.

## Distribution of Molecular Speeds

The motion of molecules in a gas is random in magnitude and direction for individual molecules, but a gas of many molecules has a predictable distribution of molecular speeds. This distribution is called the *Maxwell-Boltzmann distribution*, after its originators, who calculated it based on kinetic theory, and has since been confirmed experimentally. (See Figure.) The distribution has a long tail, because a few molecules may go several times the rms speed. The most probable speed is less than the rms speed $v_{rms}$. Figure shows that the curve is shifted to higher speeds at higher temperatures, with a broader range of speeds.

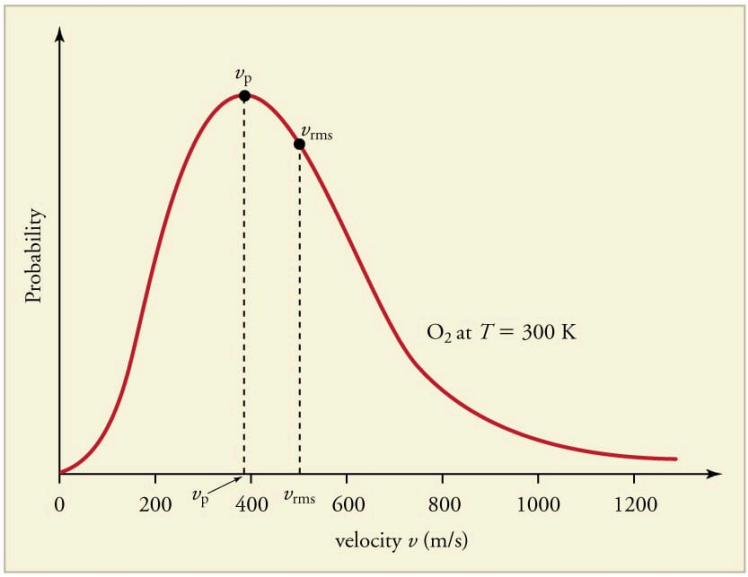



Figure 5.3.4: The Maxwell-Boltzmann distribution of molecular speeds in an ideal gas. The most likely speed $v_p$ is less than the rms speed $v_{rms}$. Although very high speeds are possible, only a tiny fraction of the molecules have speeds that are an order of magnitude greater than $v_{rms}$.

The distribution of thermal speeds depends strongly on temperature. As temperature increases, the speeds are shifted to higher values and the distribution is broadened.

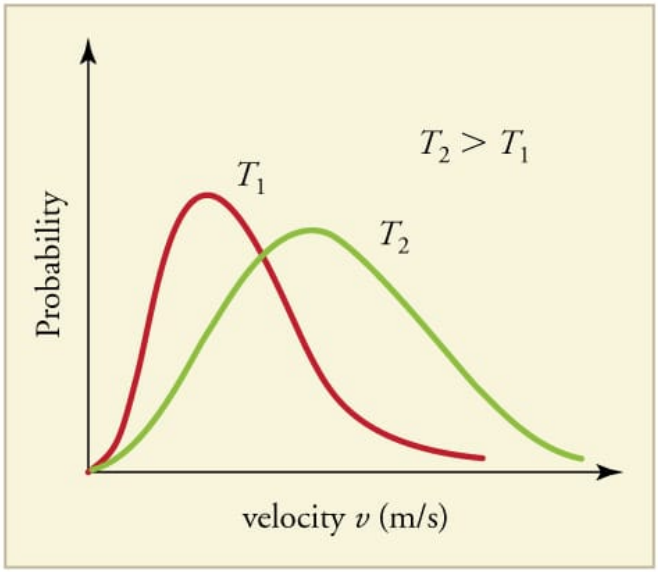



Figure 5.3.5: The Maxwell-Boltzmann distribution is shifted to higher speeds and is broadened at higher temperatures.

What is the implication of the change in distribution with temperature shown in Figure for humans? All other things being equal, if a person has a fever, he or she is likely to lose more water molecules, particularly from linings along moist cavities such as the lungs and mouth, creating a dry sensation in the mouth.

## Example 5.3.2: Calculating Temperature: Escape Velocity of Helium Atoms

In order to escape Earth's gravity, an object near the top of the atmosphere (at an altitude of 100 km) must travel away from Earth at 11.1 km/s. This speed is called the *escape velocity*. At what temperature would helium atoms have an rms speed equal to the escape velocity?

**Strategy**

Identify the knowns and unknowns and determine which equations to use to solve the problem.

**Solution**

1. Identify the knowns: $v$ is the escape velocity, 11.1 km/s.

2. Identify the unknowns: We need to solve for temperature, $T$. We also need to solve for the mass $m$

of the helium atom.

3. Determine which equations are needed.

- To solve for mass $m$ of the helium atom, we can use information from the periodic table:

$$m = \frac{molar\,mass}{number\,of\,atoms\,per\,mole}. \tag{5.3.21}$$

- To solve for temperature $T$, we can rearrange either

$$\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{5.3.22}$$

or

$$\sqrt{\overline{v^2}} = v_{rms} = \sqrt{\frac{3kT}{m}} \tag{5.3.23}$$

- $$T = \frac{\overline{mv^2}}{3k}, \tag{5.3.24}$$

where $k$ is the Boltzmann constant and $m$ is the mass of a helium atom.

4. Plug the known values into the equations and solve for the unknowns.

$$m = \frac{molar\,mass}{number\,of\,atoms\,per\,mole} = \frac{4.0026 \times 10^{-3} kg/mole}{6.02 \times 10^{23} mol} = 6.65 \times 10^{-27}\,kg \tag{5.3.25}$$

$$T = \frac{(6.65 \times 10^{-27}\,kg)(11.1 \times 10^{3}\,m/s)^2}{3(1.38 \times 10^{-23}\,J/K)} = 1.98 \times 10^{4}\,K \tag{5.3.26}$$

**Discussion**

This temperature is much higher than atmospheric temperature, which is approximately 250 K $(-25^oC$ or $-10^oF)$ at high altitude. Very few helium atoms are left in the atmosphere, but there were many when the atmosphere was formed. The reason for the loss of helium atoms is that there are a small number of helium atoms with speeds higher than Earth's escape velocity even at normal temperatures. The speed of a helium atom changes from one instant to the next, so that at any instant, there is a small, but nonzero chance that the speed is greater than the escape speed and the molecule escapes from Earth's gravitational pull. Heavier molecules, such as oxygen, nitrogen, and water (very little of which reach a very high altitude), have smaller rms speeds, and so it is much less likely that any of them will have speeds greater than the escape velocity. In fact, so few have speeds above the escape velocity that billions of years are required to lose significant amounts of the atmosphere. Figure shows the impact of a lack of an atmosphere on the Moon. Because the gravitational pull of the Moon is much weaker, it has lost almost its entire atmosphere. The comparison between Earth and the Moon is discussed in this chapter's Problems and Exercises.



Figure 5.3.6: This photograph of Apollo 17 Commander Eugene Cernan driving the lunar rover on the Moon in 1972 looks as though it was taken at night with a large spotlight. In fact, the light is coming from the Sun. Because the acceleration due to gravity on the Moon is so low (about 1/6 that of Earth), the Moon's escape velocity is much smaller. As a result, gas molecules escape very easily from the Moon, leaving it with virtually no atmosphere. Even during the daytime, the sky is black because there is no gas to scatter sunlight. (credit: Harrison H. Schmitt/NASA)

### Check Your Understanding

If you consider a very small object such as a grain of pollen, in a gas, then the number of atoms and molecules striking its surface would also be relatively small. Would the grain of pollen experience any fluctuations in pressure due to statistical fluctuations in the number of gas atoms and molecules striking it in a given amount of time?

[Hide Solution]

Yes. Such fluctuations actually occur for a body of any size in a gas, but since the numbers of atoms and molecules are immense for macroscopic bodies, the fluctuations are a tiny percentage of the number of collisions, and the averages spoken of in this section vary imperceptibly. Roughly speaking the fluctuations are proportional to the inverse square root of the number of collisions, so for small bodies they can become significant. This was actually observed in the 19th century for pollen grains in water, and is known as the Brownian effect.

### PhET Explorations: Gas Properties

Pump gas molecules into a box and see what happens as you change the volume, add or remove heat, change gravity, and more. Measure the temperature and pressure, and discover how the properties of the gas vary in relation to each other.




Figure 5.3.7: Gas Property

## Summary

- Kinetic theory is the atomistic description of gases as well as liquids and solids.
- Kinetic theory models the properties of matter in terms of continuous random motion of atoms and molecules.
- The ideal gas law can also be expressed as

$$PV = \frac{1}{3}Nm\overline{v^2}, \tag{5.3.27}$$

where $P$ is the pressure (average force per unit area), $V$ is the volume of gas in the container, $N$ is the number of molecules in the container, is the mass of a molecule, and $\overline{v^2}$ is the average of the molecular speed squared.

- The temperature of gases is proportional to the average translational kinetic energy of atoms and molecules.

$$\overline{KE} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{5.3.28}$$

or

$$\sqrt{\overline{v^2}} = v_{rms} = \sqrt{\frac{3kT}{m}}. \tag{5.3.29}$$

- The motion of individual molecules in a gas is random in magnitude and direction. However, a gas of many molecules has a predictable distribution of molecular speeds, known as the *Maxwell-Boltzmann distribution*.

This page titled 5.3: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 13.4: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.4: Phase Changes

**Learning Objectives**

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

- Interpret a phase diagram.
- State Dalton's law.
- Identify and describe the triple point of a gas from its phase diagram.
- Describe the state of equilibrium between a liquid and a gas, a liquid and a solid, and a gas and a solid.

Up to now, we have considered the behavior of ideal gases. Real gases are like ideal gases at high temperatures. At lower temperatures, however, the interactions between the molecules and their volumes cannot be ignored. The molecules are very close (condensation occurs) and there is a dramatic decrease in volume, as seen in Figure $5.4.1$. The substance changes from a gas to a liquid. When a liquid is cooled to even lower temperatures, it becomes a solid. The volume never reaches zero because of the finite volume of the molecules.

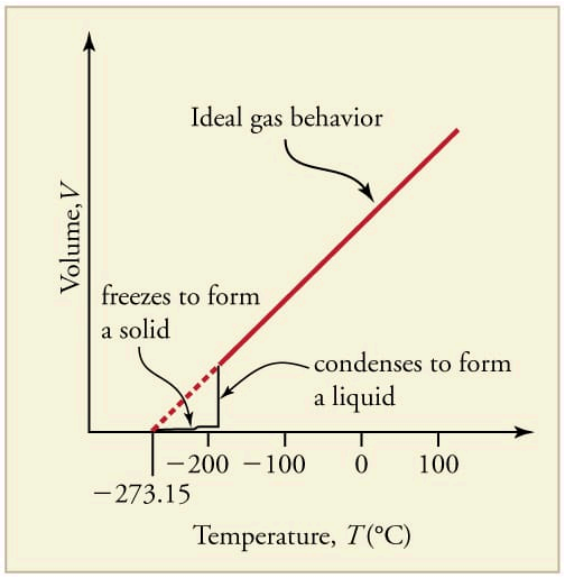



Figure $5.4.1$: A sketch of volume versus temperature for a real gas at constant pressure. The linear (straight line) part of the graph represents ideal gas behavior—volume and temperature are directly and positively related and the line extrapolates to zero volume at $273.15^oC$, or absolute zero. When the gas becomes a liquid, however, the volume actually decreases precipitously at the liquefaction point. The volume decreases slightly once the substance is solid, but it never becomes zero.

High pressure may also cause a gas to change phase to a liquid. Carbon dioxide, for example, is a gas at room temperature and atmospheric pressure, but becomes a liquid under sufficiently high pressure. If the pressure is reduced, the temperature drops and the liquid carbon dioxide solidifies into a snow-like substance at the temperature $-78^oC$. Solid $CO_2$ is called "dry ice." Another example of a gas that can be in a liquid phase is liquid nitrogen $(LN_2)$. $LN_2$ is made by liquefaction of atmospheric air (through compression and cooling). It boils at 77 K $(-196^oC)$ at atmospheric pressure. $LN_2$ is useful as a refrigerant and allows for the preservation of blood, sperm, and other biological materials. It is also used to reduce noise in electronic sensors and equipment, and to help cool down their current-carrying wires. In dermatology, $LN_2$ is used to freeze and painlessly remove warts and other growths from the skin.

## PV Diagrams

We can examine aspects of the behavior of a substance by plotting a graph of pressure versus volume, called a ***PV* diagram**. When the substance behaves like an ideal gas, the ideal gas law describes the relationship between its pressure and volume. That

is,

$$PV = NkT\,(ideal\,gas). \tag{5.4.1}$$

Now, assuming the number of molecules and the temperature are fixed,

$$PV - constant\,(ideal\,gas,\ constant\,temperature). \tag{5.4.2}$$

For example, the volume of the gas will decrease as the pressure increases. If you plot the relationship $PV$ = constant on a $PV$ diagram, you find a hyperbola. Figure 5.4.2 shows a graph of pressure versus volume. The hyperbolas represent ideal-gas behavior at various fixed temperatures, and are called *isotherms*. At lower temperatures, the curves begin to look less like hyperbolas—the gas is not behaving ideally and may even contain liquid. There is a critical point—that is, a critical temperature—above which liquid cannot exist. At sufficiently high pressure above the critical point, the gas will have the density of a liquid but will not condense. Carbon dioxide, for example, cannot be liquefied at a temperature above $31.0^oC$. Critical pressure is the minimum pressure needed for liquid to exist at the critical temperature. Table 5.4.1 lists representative critical temperatures and pressures.

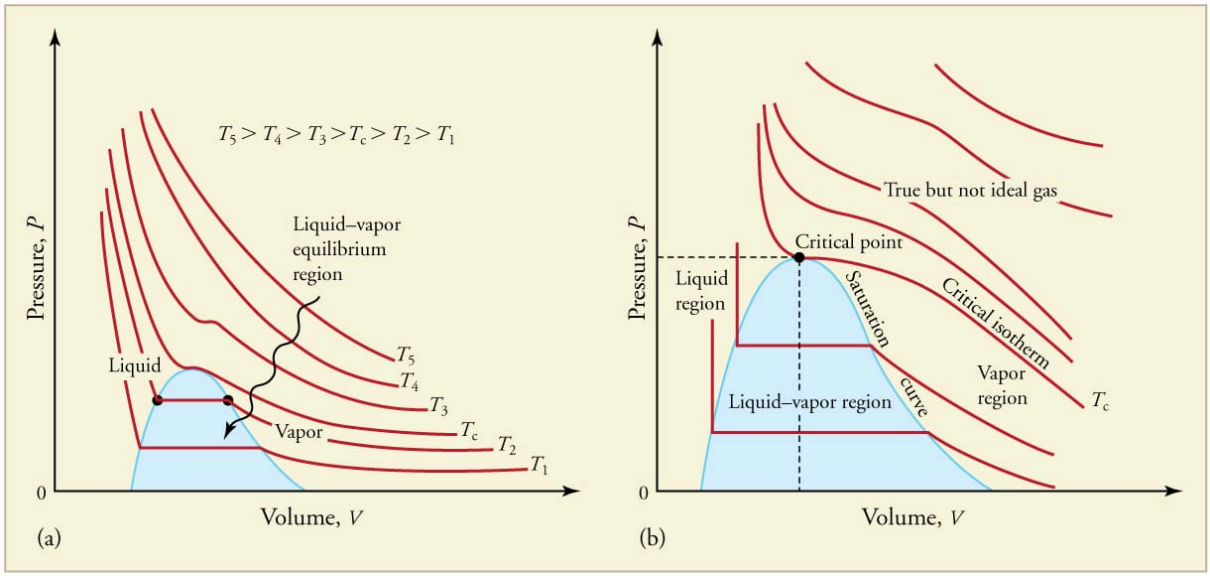



Figure 5.4.2: $PV$ diagrams. (a) Each curve (isotherm) represents the relationship between $P$ and $V$ at a fixed temperature; the upper curves are at higher temperatures. The lower curves are not hyperbolas, because the gas is no longer an ideal gas. (b) An expanded portion of the $PV$ diagram for low temperatures, where the phase can change from a gas to a liquid. The term "vapor" refers to the gas phase when it exists at a temperature below the boiling temperature.

*Table 5.4.1: Critical Temperatures and Pressures*

| Substance | Critical Temperature | | Critical pressure | |
|---|---|---|---|---|
| | $K$ | $^oC$ | $Pa$ | $atm$ |
| Water | 647.4 | 374.3 | $22.12 \times 10^6$ | 219.0 |
| Sulfur dioxide | 430.7 | 157.6 | $7.88 \times 10^6$ | 78.0 |
| Ammonia | 405.5 | 132.4 | $11.28 \times 10^6$ | 111.7 |
| Carbon dioxide | 304.2 | 31.1 | $7.39 \times 10^6$ | 73.2 |
| Oxygen | 154.8 | –118.4 | $5.08 \times 10^6$ | 50.3 |
| Nitrogen | 126.2 | –146.9 | $3.39 \times 10^6$ | 33.6 |
| Hydrogen | 33.3 | –239.9 | $1.30 \times 10^6$ | 12.9 |
| Helium | 5.3 | –267.9 | $0.229 \times 10^6$ | 2.27 |

## Phase Diagrams

The plots of pressure versus temperatures provide considerable insight into thermal properties of substances. There are well-defined regions on these graphs that correspond to various phases of matter, so $PT$ graphs are called **phase diagrams**. Figure $5.4.3$ shows the phase diagram for water. Using the graph, if you know the pressure and temperature you can determine the phase of water. The solid lines—boundaries between phases—indicate temperatures and pressures at which the phases coexist (that is, they exist together in ratios, depending on pressure and temperature). For example, the boiling point of water is $100^oC$ at 1.00 atm. As the pressure increases, the boiling temperature rises steadily to $374^oC$ at a pressure of 218 atm. A pressure cooker (or even a covered pot) will cook food faster because the water can exist as a liquid at temperatures greater than $100^oC$ without all boiling away. The curve ends at a point called the *critical point*, because at higher temperatures the liquid phase does not exist at any pressure. The critical point occurs at the critical temperature, as you can see for water from Table. The critical temperature for oxygen is $-118^oC$, so oxygen cannot be liquefied above this temperature.

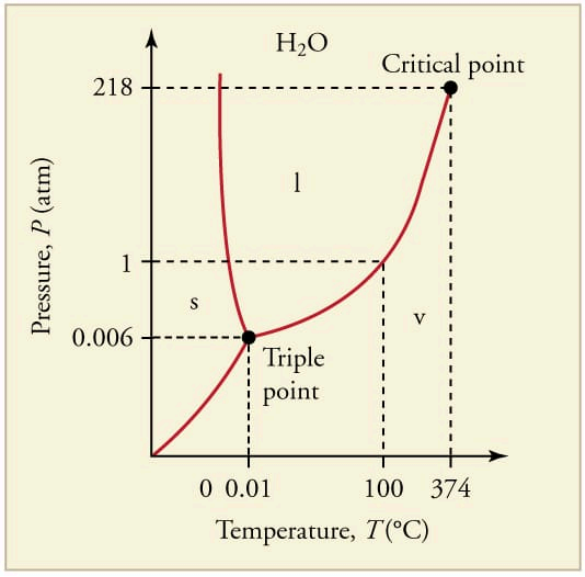



Figure $5.4.3$: The phase diagram ($PT$ graph) for water. Note that the axes are nonlinear and the graph is not to scale. This graph is simplified—there are several other exotic phases of ice at higher pressures.

Similarly, the curve between the solid and liquid regions in Figure gives the melting temperature at various pressures. For example, the melting point is $0^oC$ at 1.00 atm, as expected. Note that, at a fixed temperature, you can change the phase from solid (ice) to liquid (water) by increasing the pressure. Ice melts from pressure in the hands of a snowball maker. From the phase diagram, we can also say that the melting temperature of ice rises with increased pressure. When a car is driven over snow, the increased pressure from the tires melts the snowflakes; afterwards the water refreezes and forms an ice layer.

At sufficiently low pressures there is no liquid phase, but the substance can exist as either gas or solid. For water, there is no liquid phase at pressures below 0.00600 atm. The phase change from solid to gas is called **sublimation**. It accounts for large losses of snow pack that never make it into a river, the routine automatic defrosting of a freezer, and the freeze-drying process applied to many foods. Carbon dioxide, on the other hand, sublimates at standard atmospheric pressure of 1 atm. (The solid form of $CO_2$ is known as dry ice because it does not melt. Instead, it moves directly from the solid to the gas state.)

All three curves on the phase diagram meet at a single point, the **triple point**, where all three phases exist in equilibrium. For water, the triple point occurs at 273.16 K $(0.01^oC)$ and is a more accurate calibration temperature than the melting point of water at 1.00 atm, or 273.15 K $(0.0^oC)$. See Table for the triple point values of other substances.

# Equilibrium

Liquid and gas phases are in equilibrium at the boiling temperature. (See Figure.) If a substance is in a closed container at the boiling point, then the liquid is boiling and the gas is condensing at the same rate without net change in their relative amount. Molecules in the liquid escape as a gas at the same rate at which gas molecules stick to the liquid, or form droplets and become part of the liquid phase. The combination of temperature and pressure has to be “just right”; if the temperature and pressure are

increased, equilibrium is maintained by the same increase of boiling and condensation rates.

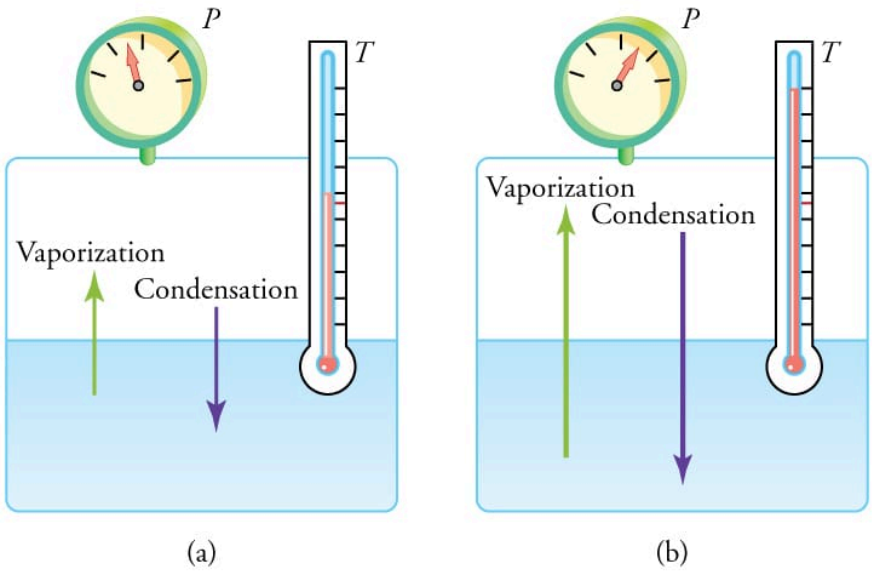



Figure 5.4.4: Equilibrium between liquid and gas at two different boiling points inside a closed container. (a) The rates of boiling and condensation are equal at this combination of temperature and pressure, so the liquid and gas phases are in equilibrium. (b) At a higher temperature, the boiling rate is faster and the rates at which molecules leave the liquid and enter the gas are also faster. Because there are more molecules in the gas, the gas pressure is higher and the rate at which gas molecules condense and enter the liquid is faster. As a result the gas and liquid are in equilibrium at this higher temperature.

*Table 5.4.2: Triple Point Temperatures and Pressures*

| Substance | Temperature | | Pressure | |
|---|---|---|---|---|
| | $K$ | $^oC$ | Pa | atm |
| Water | 273.16 | 0.01 | $6.10 \times 10^2$ | 0.00600 |
| Carbon dioxide | 216.55 | −56.60 | $5.16 \times 10^5$ | 5.11 |
| Sulfur dioxide | 197.68 | −75.47 | $1.67 \times 10^3$ | 0.0167 |
| Ammonia | 195.40 | −77.75 | $6.06 \times 10^3$ | 0.0600 |
| Nitrogen | 63.18 | −210.0 | $1.25 \times 10^4$ | 0.124 |
| Oxygen | 54.36 | −218.8 | $1.52 \times 10^2$ | 0.00151 |
| Hydrogen | 13.84 | −259.3 | $7.04 \times 10^3$ | 0.0697 |

One example of equilibrium between liquid and gas is that of water and steam at $100^oC$ and 1.00 atm. This temperature is the boiling point at that pressure, so they should exist in equilibrium. Why does an open pot of water at $100^oC$ boil completely away? The gas surrounding an open pot is not pure water: it is mixed with air. If pure water and steam are in a closed container at $100^oC$ and 1.00 atm, they would coexist—but with air over the pot, there are fewer water molecules to condense, and water boils. What about water at $20.0^oC$ and 1.00 atm? This temperature and pressure correspond to the liquid region, yet an open glass of water at this temperature will completely evaporate. Again, the gas around it is air and not pure water vapor, so that the reduced evaporation rate is greater than the condensation rate of water from dry air. If the glass is sealed, then the liquid phase remains. We call the gas phase a vapor when it exists, as it does for water at $20.0^oC$, at a temperature below the boiling temperature.

### Exercise 5.4.1: Check your Understanding

Explain why a cup of water (or soda) with ice cubes stays at $0^oC$, even on a hot summer day.

*Answer available at the end of this section.*

# Vapor Pressure, Partial Pressure, and Dalton's Law

**Vapor pressure** is defined as the pressure at which a gas coexists with its solid or liquid phase. Vapor pressure is created by faster molecules that break away from the liquid or solid and enter the gas phase. The vapor pressure of a substance depends on both the substance and its temperature—an increase in temperature increases the vapor pressure.

**Partial pressure** is defined as the pressure a gas would create if it occupied the total volume available. In a mixture of gases, *the total pressure is the sum of partial pressures of the component gases*, assuming ideal gas behavior and no chemical reactions between the components. This law is known as **Dalton's law of partial pressures**, after the English scientist John Dalton (1766–1844), who proposed it. Dalton's law is based on kinetic theory, where each gas creates its pressure by molecular collisions, independent of other gases present. It is consistent with the fact that pressures add according to Pascal's Principle. Thus water evaporates and ice sublimates when their vapor pressures exceed the partial pressure of water vapor in the surrounding mixture of gases. If their vapor pressures are less than the partial pressure of water vapor in the surrounding gas, liquid droplets or ice crystals (frost) form.

**Exercise 5.4.1**

Is energy transfer involved in a phase change? If so, will energy have to be supplied to change phase from solid to liquid and liquid to gas? What about gas to liquid and liquid to solid? Why do they spray the orange trees with water in Florida when the temperatures are near or just below freezing?

*Answer available at the end of this section.*

**PHET EXPLORATIONS: STATES OF MATTER - BASICS**

Heat, cool, and compress atoms and molecules and watch as they change between solid, liquid, and gas phases.

PhET Interactive Simulation

Figure 13.6.5. States of Matter: Basics

## Section Summary

- Most substances have three distinct phases: gas, liquid, and solid.
- Phase changes among the various phases of matter depend on temperature and pressure.
- The existence of the three phases with respect to pressure and temperature can be described in a phase diagram.
- Two phases coexist (i.e., they are in thermal equilibrium) at a set of pressures and temperatures. These are described as a line on a phase diagram.
- The three phases coexist at a single pressure and temperature. This is known as the triple point and is described by a single point on a phase diagram.
- A gas at a temperature below its boiling point is called a vapor.
- Vapor pressure is the pressure at which a gas coexists with its solid or liquid phase.
- Partial pressure is the pressure a gas would create if it existed alone.
- Dalton's law states that the total pressure is the sum of the partial pressures of all of the gases present.

## Glossary

***PV* diagram**
: a graph of pressure vs. volume

**critical point**
: the temperature above which a liquid cannot exist

**critical temperature**
: the temperature above which a liquid cannot exist

**critical pressure**
: the minimum pressure needed for a liquid to exist at the critical temperature

**vapor**
: a gas at a temperature below the boiling temperature

**vapor pressure**
: the pressure at which a gas coexists with its solid or liquid phase

**phase diagram**
: a graph of pressure vs. temperature of a particular substance, showing at which pressures and temperatures the three phases of the substance occur

**triple point**
: the pressure and temperature at which a substance exists in equilibrium as a solid, liquid, and gas

**sublimation**
: the phase change from solid to gas

**partial pressure**
: the pressure a gas would create if it occupied the total volume of space available

**Dalton's law of partial pressures**
: the physical law that states that the total pressure of a gas is the sum of partial pressures of the component gases

## Answers

**Exercise** 5.4.1**: Check your Understanding**

The ice and liquid water are in thermal equilibrium, so that the temperature stays at the freezing temperature as long as ice remains in the liquid. (Once all of the ice melts, the water temperature will start to rise.)

**Exercise** 5.4.1

Yes, energy transfer is involved in a phase change. We know that atoms and molecules in solids and liquids are bound to each other because we know that force is required to separate them. So in a phase change from solid to liquid and liquid to gas, a force must be exerted, perhaps by collision, to separate atoms and molecules. Force exerted through a distance is work, and energy is needed to do work to go from solid to liquid and liquid to gas. This is intuitively consistent with the need for energy to melt ice or boil water. The converse is also true. Going from gas to liquid or liquid to solid involves atoms and molecules pushing together, doing work and releasing energy.

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

- 13.5: Phase Changes by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.5: Prelude to Heat and Heat Transfer Methods

Energy can exist in many forms and heat is one of the most intriguing. Heat is often hidden, as it only exists when in transit, and is transferred by a number of distinctly different methods. Heat transfer touches every aspect of our lives and helps us understand how the universe functions. It explains the chill we feel on a clear breezy night, or why Earth's core has yet to cool. This chapter defines and explores heat transfer, its effects, and the methods by which heat is transferred. These topics are fundamental, as well as practical, and will often be referred to in the chapters ahead.

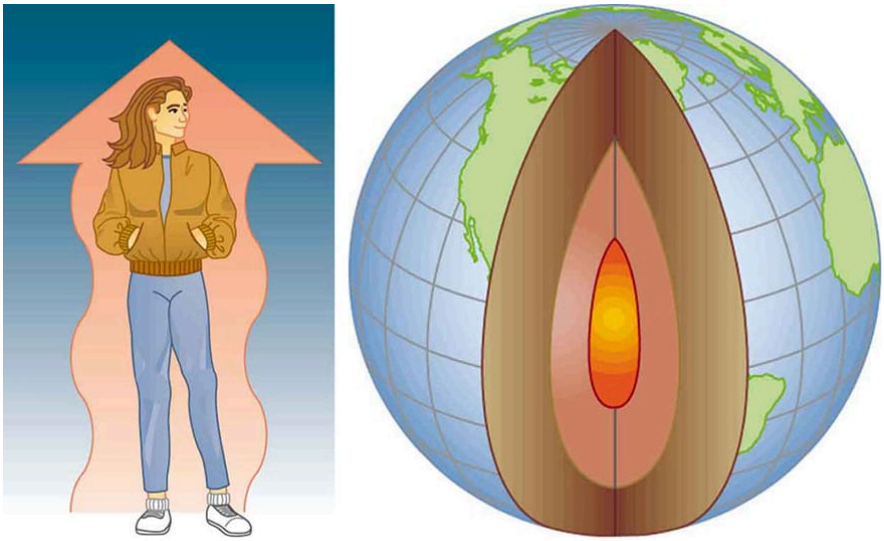



Figure 5.5.1. (a) The chilling effect of a clear breezy night is produced by the wind and by radiative heat transfer to cold outer space. (b) There was once great controversy about the Earth's age, but it is now generally accepted to be about 4.5 billion years old. Much of the debate is centered on the Earth's molten interior. According to our understanding of heat transfer, if the Earth is really that old, its center should have cooled off long ago. The discovery of radioactivity in rocks revealed the source of energy that keeps the Earth's interior molten, despite heat transfer to the surface, and from there to cold outer space.

This page titled 5.5: Prelude to Heat and Heat Transfer Methods is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.0: Prelude to Heat and Heat Transfer Methods by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.6: Heat

## Learning Objectives

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

- Define heat as transfer of energy.

In Work, Energy, and Energy Resources, we defined work as force times distance and learned that work done on an object changes its kinetic energy. We also saw in Temperature, Kinetic Theory, and the Gas Laws that temperature is proportional to the (average) kinetic energy of atoms and molecules. We say that a thermal system has a certain internal energy: its internal energy is higher if the temperature is higher. If two objects at different temperatures are brought in contact with each other, energy is transferred from the hotter to the colder object until equilibrium is reached and the bodies reach thermal equilibrium (i.e., they are at the same temperature). No work is done by either object, because no force acts through a distance. The transfer of energy is caused by the temperature difference, and ceases once the temperatures are equal. These observations lead to the following definition of heat: Heat is the spontaneous transfer of energy due to a temperature difference.

As noted in Temperature, Kinetic Theory, and the Gas Laws, heat is often confused with temperature. For example, we may say the heat was unbearable, when we actually mean that the temperature was high. Heat is a form of energy, whereas temperature is not. The misconception arises because we are sensitive to the flow of heat, rather than the temperature.

Owing to the fact that heat is a form of energy, it has the SI unit of *joule* (J). The *calorie* (cal) is a common unit of energy, defined as the energy needed to change the temperature of 1.00 g of water by $1.00^oC$ —specifically, between $14.5^oC$ and $15.5^oC$, since there is a slight temperature dependence. Perhaps the most common unit of heat is the **kilocalorie** (kcal), which is the energy needed to change the temperature of 1.00 kg of water by $1.00^oC$. Since mass is most often specified in kilograms, kilocalorie is commonly used. Food calories (given the notation Cal, and sometimes called “big calorie”) are actually kilocalories $(1\ kilocalorie = 1000\ calories)$, a fact not easily determined from package labeling.

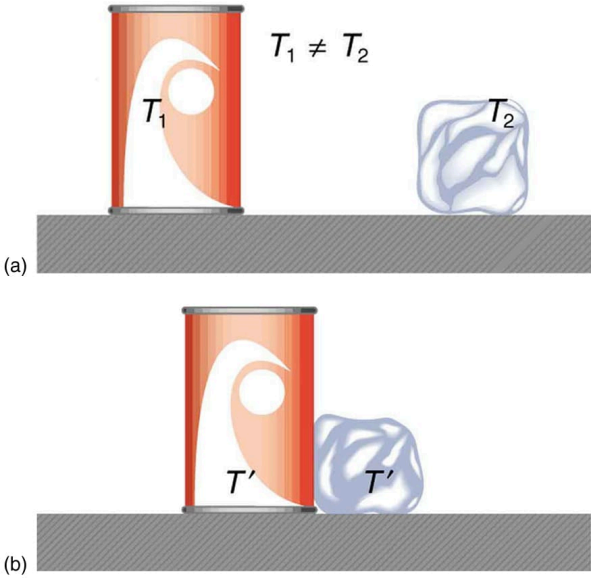



Figure $5.6.1$: In panel (a) the soft drink and the ice have different temperatures, $T_1$ and $T_2$, and are not in thermal equilibrium. In panel (b), when the soft drink and ice are allowed to interact, energy is transferred until they reach the same temperature $T'$, achieving equilibrium. Heat transfer occurs due to the difference in temperatures. In fact, since the soft drink and ice are both in contact with the surrounding air and bench, the equilibrium temperature will be the same for both.

## Mechanical Equivalent of Heat

It is also possible to change the temperature of a substance by doing work. Work can transfer energy into or out of a system. This realization helped establish the fact that heat is a form of energy. James Prescott Joule (1818–1889) performed many experiments to establish the **mechanical equivalent of heat**—*the work needed to produce the same effects as heat transfer*. In terms of the units used for these two terms, the best modern value for this equivalence is

$$1.000\,kcal = 4186\,J. \tag{5.6.1}$$

We consider this equation as the conversion between two different units of energy.

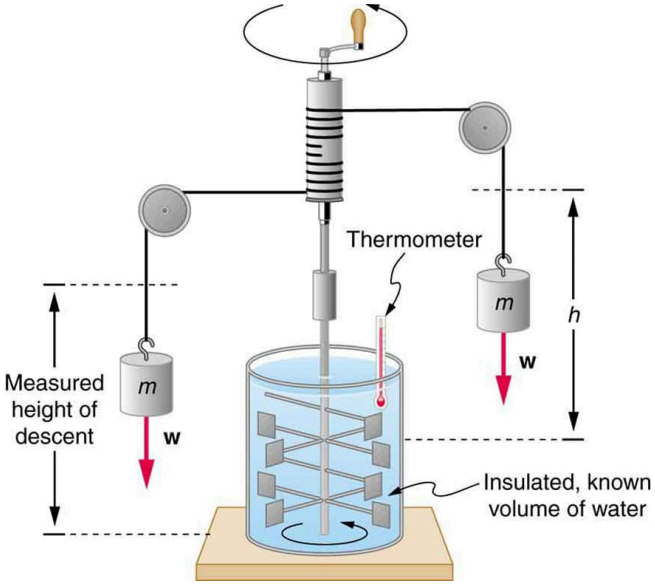



Figure 5.6.2: Schematic depiction of Joule's experiment that established the equivalence of heat and work.

Figure 5.6.2 shows one of Joule's most famous experimental setups for demonstrating the mechanical equivalent of heat. It demonstrated that work and heat can produce the same effects, and helped establish the principle of conservation of energy. Gravitational potential energy (PE) (work done by the gravitational force) is converted into kinetic energy (KE), and then randomized by viscosity and turbulence into increased average kinetic energy of atoms and molecules in the system, producing a temperature increase. His contributions to the field of thermodynamics were so significant that the SI unit of energy was named after him.

Heat added or removed from a system changes its internal energy and thus its temperature. Such a temperature increase is observed while cooking. However, adding heat does not necessarily increase the temperature. An example is melting of ice; that is, when a substance changes from one phase to another. Work done on the system or by the system can also change the internal energy of the system. Joule demonstrated that the temperature of a system can be increased by stirring. If an ice cube is rubbed against a rough surface, work is done by the frictional force. A system has a well-defined internal energy, but we cannot say that it has a certain "heat content" or "work content". We use the phrase "heat transfer" to emphasize its nature.

### Exercise 5.6.1

Two samples (A and B) of the same substance are kept in a lab. Someone adds 10 kilojoules (kJ) of heat to one sample, while 10 kJ of work is done on the other sample. How can you tell to which sample the heat was added?

*Answer available at the end of this section.*

## Summary

- Heat and work are the two distinct methods of energy transfer.

- Heat is energy transferred solely due to a temperature difference.
- Any energy unit can be used for heat transfer, and the most common are kilocalorie (kcal) and joule (J).
- Kilocalorie is defined to be the energy needed to change the temperature of 1.00 kg of water between $14.5^oC$ and $15.5^oC$.
- The mechanical equivalent of this heat transfer is $1.00\ kcal = 4186\ J$.

# Glossary

**heat**
: the spontaneous transfer of energy due to a temperature difference

**kilocalorie**
: 1kilocalorie=1000caloriessize 12{1`"kilocalorie=1000"`"calories"} {}

**mechanical equivalent of heat**
: the work needed to produce the same effects as heat transfer

## Answers

**Exercise** 5.6.1

Heat and work both change the internal energy of the substance. However, the properties of the sample only depend on the internal energy so that it is impossible to tell whether heat was added to sample A or B.

This page titled 5.6: Heat is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.1: Heat by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.7: Temperature Change and Heat Capacity

## Learning Objectives

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

- Observe heat transfer and change in temperature and mass.
- Calculate final temperature after heat transfer between two objects.

One of the major effects of heat transfer is temperature change: heating increases the temperature while cooling decreases it. We assume that there is no phase change and that no work is done on or by the system. Experiments show that the transferred heat depends on three factors—the change in temperature, the mass of the system, and the substance and phase of the substance.

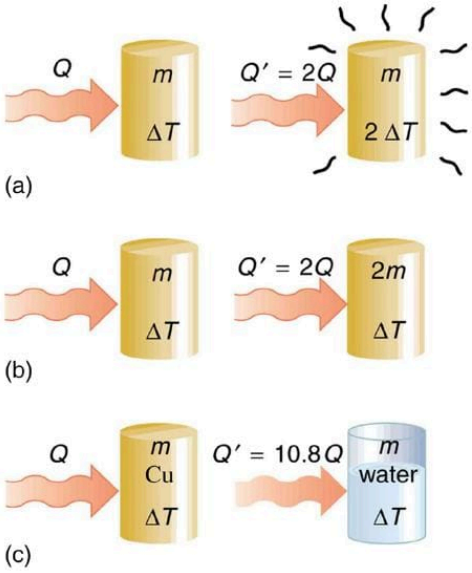



Figure 5.7.1: The heat $Q$ transferred to cause a temperature change depends on the magnitude of the temperature change, the mass of the system, and the substance and phase involved. (a) The amount of heat transferred is directly proportional to the temperature change. To double the temperature change of a mass $m$, you need to add twice the heat. (b) The amount of heat transferred is also directly proportional to the mass. To cause an equivalent temperature change in a doubled mass, you need to add twice the heat. (c) The amount of heat transferred depends on the substance and its phase. If it takes an amount $Q$ of heat to cause a temperature change $\Delta T$ in a given mass of copper, it will take 10.8 times that amount of heat to cause the equivalent temperature change in the same mass of water assuming no phase change in either substance.

The dependence on temperature change and mass are easily understood. Owing to the fact that the (average) kinetic energy of an atom or molecule is proportional to the absolute temperature, the internal energy of a system is proportional to the absolute temperature and the number of atoms or molecules. Owing to the fact that the transferred heat is equal to the change in the internal energy, the heat is proportional to the mass of the substance and the temperature change. The transferred heat also depends on the substance so that, for example, the heat necessary to raise the temperature is less for alcohol than for water. For the same substance, the transferred heat also depends on the phase (gas, liquid, or solid).

## Heat Transfer and Temperature Change

The quantitative relationship between heat transfer and temperature change contains all three factors:

$$Q = mc\Delta T, \tag{5.7.1}$$

where $Q$ is the symbol for heat transfer, $m$ is the mass of the substance, and $\Delta T$ is the change in temperature. The symbol $c$ stands for specific heat and depends on the material and phase. The specific heat is the amount of heat necessary to change the temperature of 1.00 kg of mass by $1.00^oC$. The specific heat $c$ is a property of the substance; its SI unit is $J/(kg \cdot K)$ or $J/(kg \cdot^o C)$. Recall that the temperature change $(\Delta T)$ is the same in units of kelvin and degrees Celsius. If heat transfer is measured in kilocalories, then *the unit of specific heat* is $kcal/(kg \cdot^o C)$.

Values of specific heat must generally be looked up in tables, because there is no simple way to calculate them. In general, the specific heat also depends on the temperature. Table $5.7.1$ lists representative values of specific heat for various substances. Except for gases, the temperature and volume dependence of the specific heat of most substances is weak. We see from this table that the specific heat of water is five times that of glass and ten times that of iron, which means that it takes five times as much heat to raise the temperature of water the same amount as for glass and ten times as much heat to raise the temperature of water as for iron. In fact, water has one of the largest specific heats of any material, which is important for sustaining life on Earth.

## Example $5.7.1$: Calculating the Required Heat: Heating Water in an Aluminum Pan

A 0.500 kg aluminum pan on a stove is used to heat 0.250 liters of water from $20.0^oC$ to $80.0^oC$.

(a) How much heat is required? What percentage of the heat is used to raise the temperature of (b) the pan and (c) the water?

**Strategy**

The pan and the water are always at the same temperature. When you put the pan on the stove, the temperature of the water and the pan is

increased by the same amount. We use the equation for the heat transfer for the given temperature change and mass of water and aluminum. The specific heat values for water and aluminum are given in Table $5.7.1$.

**Solution**

Because water is in thermal contact with the aluminum, the pan and the water are at the same temperature.

1. Calculate the temperature difference:
$$\Delta T = T_f - T_i = 60.0^oC. \tag{5.7.2}$$
2. Calculate the mass of water. Because the density of water is $1000\ kg/m^3$, one liter of water has a mass of 1 kg, and the mass of 0.250 liters of water is $m_w = 0.250\ kg$.
3. Calculate the heat transferred to the water. Use the specific heat of water in Table $5.7.1$
$$Q_w = m_w c_w \Delta T = (0.250\,kg)(4186\,J/kg^oC)(60.0^oC) = 62.8\,kJ. \tag{5.7.3}$$
4. Calculate the heat transferred to the aluminum. Use the specific heat for aluminum in Table \({5{.}7{.}\eatSpaces{1} \endSpaces}\:
$$Q_{Al} = m_{Al} c_{Al} \Delta T = (0.500\,kg)(900\,J/kg^oC)(60.0^oC) = 27.0kJ. \tag{5.7.4}$$
5. Compare the percentage of heat going into the pan versus that going into the water. First, find the total transferred heat:
$$Q_{Total} = Q_W + Q_{Al} = 62.8\,kJ + 27.0\,kJ = 89.8\,kJ. \tag{5.7.5}$$
6. Thus, the amount of heat going into heating the pan is
$$\frac{62.8\,kJ}{89.8\,kJ} \times 100\% = 69.9\% \tag{5.7.6}$$

**Discussion**

In this example, the heat transferred to the container is a significant fraction of the total transferred heat. Although the mass of the pan is twice that of the water, the specific heat of water is over four times greater than that of aluminum. Therefore, it takes a bit more than twice the heat to achieve the given temperature change for the water as compared to the aluminum pan.



Figure 5.7.2: The smoking brakes on this truck are a visible evidence of the mechanical equivalent of heat.

## Example 5.7.2: Calculating the Temperature Increase from the Work Done on a Substance: Truck Brakes Overheat on Downhill Runs

Truck brakes used to control speed on a downhill run do work, converting gravitational potential energy into increased internal energy (higher temperature) of the brake material. This conversion prevents the gravitational potential energy from being converted into kinetic energy of the truck. The problem is that the mass of the truck is large compared with that of the brake material absorbing the energy, and the temperature increase may occur too fast for sufficient heat to transfer from the brakes to the environment.

Calculate the temperature increase of 100 kg of brake material with an average specific heat of $800.0\ J/kg\cdot^o C$ if the material retains 10% of the energy from a 10,000-kg truck descending 75.0 m (in vertical displacement) at a constant speed.

**Strategy**

If the brakes are not applied, gravitational potential energy is converted into kinetic energy. When brakes are applied, gravitational potential energy is converted into internal energy of the brake material. We first calculate the gravitational potential energy $(Mgh)$ that the entire truck loses in its descent and then find the temperature increase produced in the brake material alone.

**Solution**

1. Calculate the change in gravitational potential energy as the truck goes downhill

$$Mgh = (10,000\,kg)(9.80\,m/s^2)(75.0\,m) = 7.35 \times 10^6\,J. \tag{5.7.7}$$

2. Calculate the temperature from the heat transferred using $Q = Mgh$ and

$$\Delta T = \frac{Q}{mc}, \tag{5.7.8}$$

where $m$ is the mass of the brake material. Insert the values $m = 100\ kg$ and $c = 800\ J/kg\cdot^o C$ to find

$$\Delta T = \frac{(7.35v \times 10^6\,J)}{(100\,kg)(800\,J/kg^oC)} = 92^oC. \tag{5.7.9}$$

**Discussion**

This temperature is close to the boiling point of water. If the truck had been traveling for some time, then just before the descent, the brake temperature would likely be higher than the ambient temperature. The temperature increase in the descent would likely raise the temperature of the brake material above the boiling point of water, so this technique is not practical. However, the same idea underlies the recent hybrid technology of cars, where mechanical energy (gravitational potential energy) is converted by the brakes into electrical energy (battery).

*Table 5.7.1: Specific Heats1 of Various Substances*

| Substances | Specific heat (c) | |
|---|---|---|
| Solids | $J/kg \cdot^o C$ | $kcal/kg \cdot^o C$ |
| Aluminum | 900 | 0.215 |
| Asbestos | 800 | 0.19 |
| Concrete, granite (average) | 840 | 0.20 |
| Copper | 387 | 0.0924 |
| Glass | 840 | 0.20 |
| Gold | 129 | 0.0308 |
| Human body (average at 37 °C) | 3500 | 0.83 |
| Ice (average, -50°C to 0°C) | 2090 | 0.50 |
| Iron, steel | 452 | 0.108 |
| Lead | 128 | 0.0305 |
| Silver | 235 | 0.0562 |
| Wood | 1700 | 0.4 |
| *Liquids* | | |
| Benzene | 1740 | 0.415 |
| Ethanol | 2450 | 0.586 |
| Glycerin | 2410 | 0.576 |
| Mercury | 139 | 0.0333 |
| Water (15.0 °C) | 4186 | 1.000 |
| *Gases* [3] | | |
| Air (dry) | 721 (1015) | 0.172 (0.242) |
| Ammonia | 1670 (2190) | 0.399 (0.523) |
| Carbon dioxide | 638 (833) | 0.152 (0.199) |
| Nitrogen | 739 (1040) | 0.177 (0.248) |
| Oxygen | 651 (913) | 0.156 (0.218) |
| Steam (100°C) | 1520 (2020) | 0.363 (0.482) |

Note that Example 5.7.2 is an illustration of the mechanical equivalent of heat. Alternatively, the temperature increase could be produced by a blow torch instead of mechanically.

## Example 5.7.3: Calculating the Final Temperature When Heat Is Transferred Between Two Bodies: Pouring Cold Water in a Hot Pan

Suppose you pour 0.250 kg of $20.0^oC$ water (about a cup) into a 0.500-kg aluminum pan off the stove with a temperature of $150^oC$. Assume that the pan is placed on an insulated pad and that a negligible amount of water boils off. What is the temperature when the water and pan reach thermal equilibrium a short time later?

**Strategy**

The pan is placed on an insulated pad so that little heat transfer occurs with the surroundings. Originally the pan and water are not in thermal equilibrium: the pan is at a higher temperature than the water. Heat transfer then restores thermal equilibrium once the water and pan are in contact. Because heat transfer between the pan and water takes place rapidly, the mass of evaporated water is negligible and the magnitude of the heat lost by the pan is equal to the heat gained by the water. The exchange of heat stops once a thermal equilibrium between the pan and the water is achieved. The heat exchange can be written as $|Q_{hot}| = Q_{cold}$.

**Solution**

1. Use the equation for heat transfer $Q = mc\Delta T$ to express the heat lost by the aluminum pan in terms of the mass of the pan, the specific heat of aluminum, the initial temperature of the pan, and the final temperature:
$$Q_{hot} = m_{Al}c_{Al}(T_f - 150^oC). \tag{5.7.10}$$
2. Express the heat gained by the water in terms of the mass of the water, the specific heat of water, the initial temperature of the water and the final temperature:
$$Q_{cold} = m_w c_w (T_f - 20.0^oC). \tag{5.7.11}$$
3. Note that $Q_{hot} < 0$ and $Q_{cold} > 0$ and that they must sum to zero because the heat lost by the hot pan must be the same as the heat gained by the cold water:
$$Q_{cold} + Q_{hot} = 0, \tag{5.7.12}$$
$$Q_{cold} = -Q_{hot}, \tag{5.7.13}$$
$$m_w c_w (T_f - 20.0^oC) = -m_{Al}c_{Al}(T_f - 150.0^oC). \tag{5.7.14}$$
4. Bring all terms involving $T_f$ on the left hand side and all other terms on the right hand side. Solve for $T_f$,
$$T_f = \frac{m_{Al}c_{Al}(150^oC) + m_W c_W (20.0^oC)}{m_{Al}c_{Al} + m_W c_W},$$
and insert the numerical values:
$$T_f = \frac{(0.500kg)(900J/kg^oC)(150^oC) + (0.250kg)(4186J/kg^oC)(20.0^oC)}{(0.500kg)(900J/kg^oC) + (0.250kg)(4186J/kg^oC)} = \frac{88430J}{1496.5J/^oC} = 59.1\ ^oC.$$

**Discussion**

This is a typical *calorimetry* problem—two bodies at different temperatures are brought in contact with each other and exchange heat until a common temperature is reached. Why is the final temperature so much closer to **20.0°C** than **150°C**? The reason is that water has a greater specific heat than most common substances and thus undergoes a small temperature change for a given heat transfer. A large body of water, such as a lake, requires a large amount of heat to increase its temperature appreciably. This explains why the temperature of a lake stays relatively constant during a day even when the temperature change of the air is large. However, the water temperature does change over longer times (e.g., summer to winter).

## TAKE-HOME EXPERIMENT: TEMPERATURE CHANGE OF LAND AND WATER

What heats faster, land or water?

To study differences in heat capacity:

- Place equal masses of dry sand (or soil) and water at the same temperature into two small jars. (The average density of soil or sand is about 1.6 times that of water, so you can achieve approximately equal masses by using $50$ more water by volume.)
- Heat both (using an oven or a heat lamp) for the same amount of time.
- Record the final temperature of the two masses.
- Now bring both jars to the same temperature by heating for a longer period of time.
- Remove the jars from the heat source and measure their temperature every 5 minutes for about 30 minutes.

Which sample cools off the fastest? This activity replicates the phenomena responsible for land breezes and sea breezes.

**Exercise** $5.7.1$

If 25 kJ is necessary to raise the temperature of a block from **25°C** to **30°C**, how much heat is necessary to heat the block from **45°C** to **50°C**?

*Answer available at the end of this section.*

## Summary

- The transfer of heat $Q$ that leads to a change $\Delta T$ in the temperature of a body with mass $m$ is $Q = mc\Delta T$, where $c$ is the specific heat of the material. This relationship can also be considered as the definition of specific heat.

## Footnotes

1 The values for solids and liquids are at constant volume and at **25°C**, except as noted.
2 These values are identical in units of **cal/g · °C**.
3 cv at constant volume and at **20.0°C**, except as noted, and at 1.00 atm average pressure. Values in parentheses are $c_p$ at a constant pressure of 1.00 atm.

## Glossary

specific heat
the amount of heat necessary to change the temperature of 1.00 kg of a substance by 1.00 °C

### Answers

**Exercise** $5.7.1$

The heat transfer depends only on the temperature difference. Since the temperature differences are the same in both cases, the same 25 kJ is necessary in the second case.

This page titled 5.7: Temperature Change and Heat Capacity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.2: Temperature Change and Heat Capacity by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.8: Phase Change and Latent Heat

**Learning Objectives**

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

- Examine heat transfer.
- Calculate final temperature from heat transfer.

So far we have discussed temperature change due to heat transfer. No temperature change occurs from heat transfer if ice melts and becomes liquid water (i.e., during a phase change). For example, consider water dripping from icicles melting on a roof warmed by the Sun. Conversely, water freezes in an ice tray cooled by lower-temperature surroundings.



Figure $5.8.1$. Heat from the air transfers to the ice causing it to melt. (credit: Mike Brand)

Energy is required to melt a solid because the cohesive bonds between the molecules in the solid must be broken apart such that, in the liquid, the molecules can move around at comparable kinetic energies; thus, there is no rise in temperature. Similarly, energy is needed to vaporize a liquid, because molecules in a liquid interact with each other via attractive forces. There is no temperature change until a phase change is complete. The temperature of a cup of soda initially at $0^oC$ stays at $0^oC$ until all the ice has melted. Conversely, energy is released during freezing and condensation, usually in the form of thermal energy. Work is done by cohesive forces when molecules are brought together. The corresponding energy must be given off (dissipated) to allow them to stay together ( see Figure $5.8.2$)

The energy involved in a phase change depends on two major factors: the number and strength of bonds or force pairs. The number of bonds is proportional to the number of molecules and thus to the mass of the sample. The strength of forces depends on the type of molecules. The heat $Q$ required to change the phase of a sample of mass $m$ is given by

$$Q = mL_f(melting/freezing), \tag{5.8.1}$$

$$Q = mL_v(vaporization/condensation), \tag{5.8.2}$$

where the latent heat of fusion, $L_f$, and latent heat of vaporization, $L_v$, are material constants that are determined experimentally. See (Table $5.8.1$).

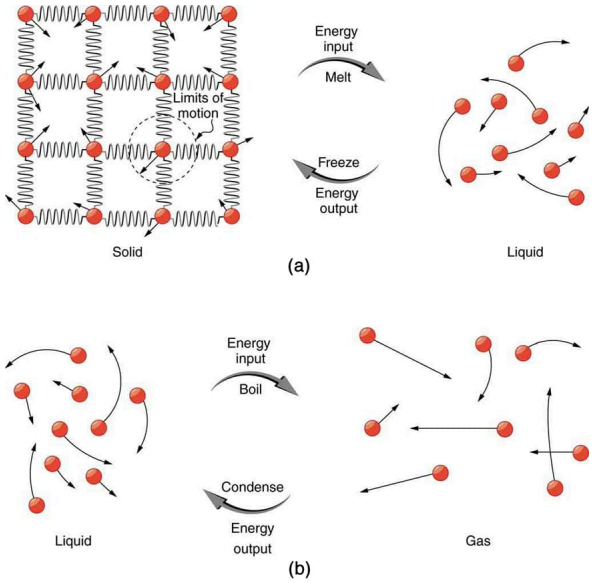



Figure 5.8.2. (a) Energy is required to partially overcome the attractive forces between molecules in a solid to form a liquid. That same energy must be removed for freezing to take place. (b) Molecules are separated by large distances when going from liquid to vapor, requiring significant energy to overcome molecular attraction. The same energy must be removed for condensation to take place. There is no temperature change until a phase change is complete.

Latent heat is measured in units of J/kg. Both $L_f$ and $L_v$ depend on the substance, particularly on the strength of its molecular forces as noted earlier. $L_f$ and $L_v$ are collectively called **latent heat coefficients**. They are *latent*, or hidden, because in phase changes, energy enters or leaves a system without causing a temperature change in the system; so, in effect, the energy is hidden. Table 5.8.1 lists representative values of $L_f$ and $L_v$, together with melting and boiling points.

The table shows that significant amounts of energy are involved in phase changes. Let us look, for example, at how much energy is needed to melt a kilogram of ice at $0^oC$ to produce a kilogram of water at $0^oC$. Using the equation for a change in temperature and the value for water from Table 5.8.1, we find that

$$Q = mL_f = (1.0\,kg)(334\,kJ/kg) = 334\,kJ \tag{5.8.3}$$

is the energy to melt a kilogram of ice. This is a lot of energy as it represents the same amount of energy needed to raise the temperature of 1 kg of liquid water from $0^oC$ to $79.8^oC$. Even more energy is required to vaporize water; it would take 2256 kJ to change 1 kg of liquid water at the normal boiling point $(100^oC$ at atmospheric pressure) to steam (water vapor). This example shows that the energy for a phase change is enormous compared to energy associated with temperature changes without a phase change.

*Table 5.8.1: Heats of Fusion and Vaporization 1*

| | $L_f$ | | | $L_v$ | | |
|---|---|---|---|---|---|---|
| **Substance** | **Melting point (°C)** | **kJ/kg** | **kcal/kg** | **Boiling point (°C)** | **kJ/kg** | **kcal/kg** |
| Helium | –269.7 | 5.23 | 1.25 | –268.9 | 20.9 | 4.99 |
| Hydrogen | –259.3 | 58.6 | 14.0 | –252.9 | 452 | 108 |
| Nitrogen | –210.0 | 25.5 | 6.09 | –195.8 | 201 | 48.0 |
| Oxygen | –218.8 | 13.8 | 3.30 | –183.0 | 213 | 50.9 |
| Ethanol | –114 | 104 | 24.9 | 78.3 | 854 | 204 |
| Ammonia | –75 | | 108 | –33.4 | 1370 | 327 |
| Mercury | –38.9 | 11.8 | 2.82 | 357 | 272 | 65.0 |
| Water | 0.00 | 334 | 79.8 | 100.0 | 2256[2] | 539[3] |
| Sulfur | 119 | 38.1 | 9.10 | 444.6 | 326 | 77.9 |

| | | $L_f$ | | | $L_v$ | |
|---|---|---|---|---|---|---|
| Lead | 327 | 24.5 | 5.85 | 1750 | 871 | 208 |
| Antimony | 631 | 165 | 39.4 | 1440 | 561 | 134 |
| Aluminum | 660 | 380 | 90 | 2450 | 11400 | 2720 |
| Silver | 961 | 88.3 | 21.1 | 2193 | 2336 | 558 |
| Gold | 1063 | 64.5 | 15.4 | 2660 | 1578 | 377 |
| Copper | 1083 | 134 | 32.0 | 2595 | 5069 | 1211 |
| Uranium | 1133 | 84 | 20 | 3900 | 1900 | 454 |
| Tungsten | 3410 | 184 | 44 | 5900 | 4810 | 1150 |

Phase changes can have a tremendous stabilizing effect even on temperatures that are not near the melting and boiling points, because evaporation and condensation (conversion of a gas into a liquid state) occur even at temperatures below the boiling point.Take, for example, the fact that air temperatures in humid climates rarely go above $35.0^oC$,

which is because most heat transfer goes into evaporating water into the air. Similarly, temperatures in humid weather rarely fall below the dew point because enormous heat is released when water vapor condenses.

We examine the effects of phase change more precisely by considering adding heat into a sample of ice at $-20^oC$ (Figure $5.8.3$ ). The temperature of the ice rises linearly, absorbing heat at a constant rate of $0.50\ cal/g\cdot^o C$ until it reaches $0^oC$. Once at this temperature, the ice begins to melt until all the ice has melted, absorbing 79.8 cal/g of heat. The temperature remains constant at $0^oC$ during this phase change. Once all the ice has melted, the temperature of the liquid water rises, absorbing heat at a new constant rate of $1.00\ cal/g\cdot^o C$. At $100^oC$, the water begins to boil and the temperature again remains constant while the water absorbs 539 cal/g of heat during this phase change. When all the liquid has become steam vapor, the temperature rises again, absorbing heat at a rate of $0.482\ cal/g\cdot^o C$.

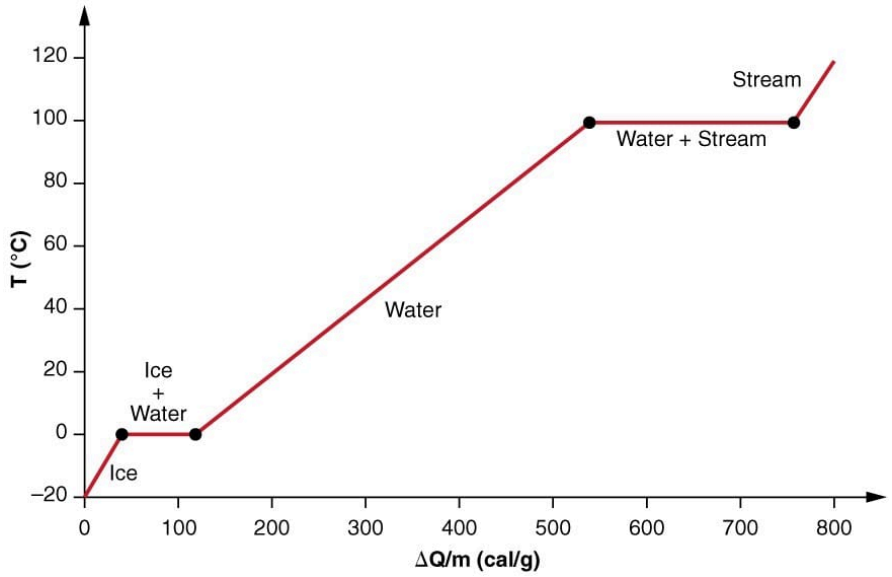



Figure $5.8.3$. A graph of temperature versus energy added. The system is constructed so that no vapor evaporates while ice warms to become liquid water, and so that, when vaporization occurs, the vapor remains in of the system. The long stretches of constant temperature values at $0^oC$ and $100^oC$ reflect the large latent heat of melting and vaporization, respectively.

Water can evaporate at temperatures below the boiling point. More energy is required than at the boiling point, because the kinetic energy of water molecules at temperatures below $100^oC$ is less than that at $100^oC$, hence less energy is available from random thermal motions. Take, for example, the fact that, at body temperature, perspiration from the skin requires a heat input of 2428 kJ/kg, which is about 10 percent higher than the latent heat of vaporization at $100^oC$. This heat comes from the skin, and thus provides an effective cooling mechanism in hot weather. High humidity inhibits evaporation, so that body temperature might rise, leaving unevaporated sweat on your brow.

## Example 5.8.1: Calculate Final Temperature from Phase Change: Cooling Soda with Ice Cubes

Three ice cubes are used to chill a soda at $20^oC$ with mass $m_{soda} = 0.25\ kg$. The ice is at $0^oC$ and each ice cube has a mass of 6.0 g. Assume that the soda is kept in a foam container so that heat loss can be ignored. Assume the soda has the same heat capacity as water. Find the final temperature when all ice has melted.

**Strategy**

The ice cubes are at the melting temperature of $0^oC$. Heat is transferred from the soda to the ice for melting. Melting of ice occurs in two steps: first the phase change occurs and solid (ice) transforms into liquid water at the melting temperature, then the temperature of this water rises. Melting yields water at $0^oC$, so more heat is transferred from the soda to this water until the water plus soda system reaches thermal equilibrium,

$$Q_{ice} = -Q_{soda}. \tag{5.8.4}$$

The heat transferred to the ice is $Q_{ice} = m_{ice}L_f + m_{ice}c_w(T_f - 0^oC)$.. The heat given off by the soda is $Q_{soda} = m_{soda}c_w(T_f - 20^oC)$. Since no heat is lost, $Q_{ice} = -Q_{soda}$, so that

$$m_{ice}Lf + m_{ice}c_w(T_f - 0^oC) = -m_{soda}c_w(T_f - 20^oC). \tag{5.8.5}$$

Bring all terms involving $T_f$ on the left-hand-side and all other terms on the right-hand-side. Solve for the unknown quantity $T_f$:

$$T_f = \frac{m_{soda}c_w(20^oC) - m_{ice}L_f}{(m_{soda} + m_{ice})c_w}. \tag{5.8.6}$$

**Solution**

1. Identify the known quantities. The mass of ice is $m_{ice} = 3 \times 6.0\ g = 0.018\ kg$ and the mass of soda is $m_{soda} = 0.25\ kg$.
2. Calculate the terms in the numerator:

$$m_{soda}c_w(20^oC) = (0.25\,kg)(4186\,J/kg\cdot^o C)(20^oC) = 20,930\,J \tag{5.8.7}$$

and

$$m_{ice}L_f = (0.018\,kg)(334,000\,J/kg) = 6012\,J. \tag{5.8.8}$$

3. Calculate the denominator:

$$(m_{soda} + m_{ice})c_w = (0.25\,kg + 0.018\,kg)(4186\,J/kg\cdot^o C) = 1122\,J/^oC \tag{5.8.9}$$

4. Calculate the final temperature:

$$T_f = \frac{20,930\,J - 6012\,J}{1122\,J/^oC} = 13^oC. \tag{5.8.10}$$

**Discussion**

This example illustrates the enormous energies involved during a phase change. The mass of ice is about 7 percent the mass of water but leads to a noticeable change in the temperature of soda. Although we assumed that the ice was at the freezing temperature, this is incorrect: the typical temperature is $-6^oC$. However, this correction gives a final temperature that is essentially identical to the result we found. Can you explain why?

We have seen that vaporization requires heat transfer to a liquid from the surroundings, so that energy is released by the surroundings. Condensation is the reverse process, increasing the temperature of the surroundings. This increase may seem surprising, since we associate condensation with cold objects—the glass in the figure, for example. However, energy must be removed from the condensing molecules to make a vapor condense. The energy is exactly the same as that required to make the phase change in the other direction, from liquid to vapor, and so it can be calculated from $Q = mL_v$.

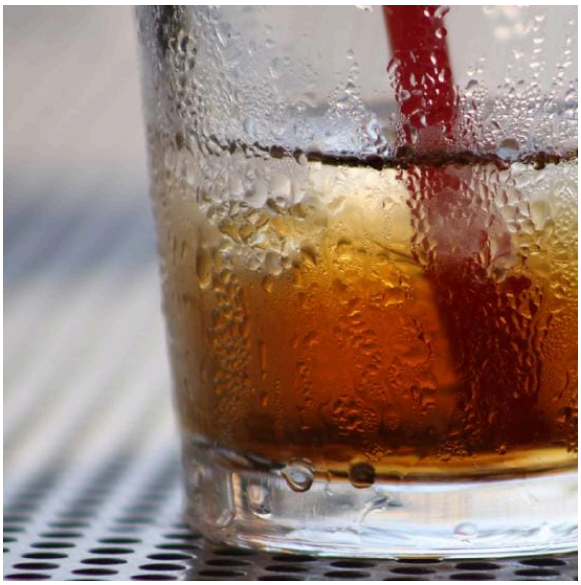


Figure 5.8.4. Condensation forms on this glass of iced tea because the temperature of the nearby air is reduced to below the dew point. The rate at which water molecules join together exceeds the rate at which they separate, and so water condenses. Energy is released when the water condenses, speeding the melting of the ice in the glass. (credit: Jenny Downing)

## Real-World Application

Energy is also released when a liquid freezes. This phenomenon is used by fruit growers in Florida to protect oranges when the temperature is close to the freezing point $(0^oC)$. Growers spray water on the plants in orchards so that the water freezes and heat is released to the growing oranges on the trees. This prevents the temperature inside the orange from dropping below freezing, which would damage the fruit.



Figure 5.8.5. The ice on these trees released large amounts of energy when it froze, helping to prevent the temperature of the trees from dropping below $0^oC$. Water is intentionally sprayed on orchards to help prevent hard frosts. (credit: Hermann Hammer)

**Sublimation** is the transition from solid to vapor phase. You may have noticed that snow can disappear into thin air without a trace of liquid water, or the disappearance of ice cubes in a freezer. The reverse is also true: Frost can form on very cold windows without going through the liquid stage. A popular effect is the making of “smoke” from dry ice, which is solid carbon dioxide. Sublimation occurs because the equilibrium vapor pressure of solids is not zero. Certain air fresheners use the sublimation of a solid to inject a perfume into the room. Moth balls are a slightly toxic example of a phenol (an organic compound) that sublimates,

while some solids, such as osmium tetroxide, are so toxic that they must be kept in sealed containers to prevent human exposure to their sublimation-produced vapors.

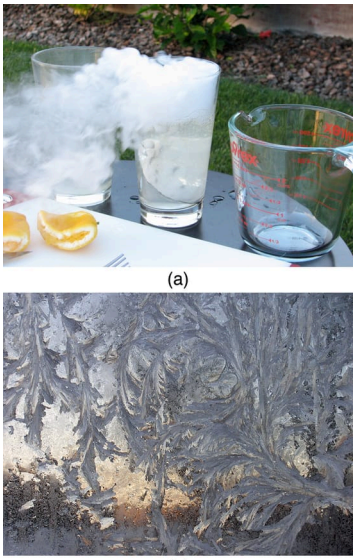



Figure 5.8.6. Direct transitions between solid and vapor are common, sometimes useful, and even beautiful. (a) Dry ice sublimates directly to carbon dioxide gas. The visible vapor is made of water droplets. (credit: Windell Oskay) (b) Frost forms patterns on a very cold window, an example of a solid formed directly from a vapor. (credit: Liz West)

All phase transitions involve heat. In the case of direct solid-vapor transitions, the energy required is given by the equation $Q = mL_s$, where $L_s$ is the heat of sublimation, which is the energy required to change 1.00 kg of a substance from the solid phase to the vapor phase. $L_s$ is analogous to $L_f$ and $L_v$, and its value depends on the substance. Sublimation requires energy input, so that dry ice is an effective coolant, whereas the reverse process (i.e., frosting) releases energy. The amount of energy required for sublimation is of the same order of magnitude as that for other phase transitions.

The material presented in this section and the preceding section allows us to calculate any number of effects related to temperature and phase change. In each case, it is necessary to identify which temperature and phase changes are taking place and then to apply the appropriate equation. Keep in mind that heat transfer and work can cause both temperature and phase changes.

## Problem-Solving Strategies for the Effects of Heat Transfer

1. *Examine the situation to determine that there is a change in the temperature or phase. Is there heat transfer into or out of the system?* When the presence or absence of a phase change is not obvious, you may wish to first solve the problem as if there were no phase changes, and examine the temperature change obtained. If it is sufficient to take you past a boiling or melting point, you should then go back and do the problem in steps—temperature change, phase change, subsequent temperature change, and so on.
2. *Identify and list all objects that change temperature and phase.*
3. *Identify exactly what needs to be determined in the problem (identify the unknowns).* A written list is useful.
4. *Make a list of what is given or what can be inferred from the problem as stated (identify the knowns).*
5. *Solve the appropriate equation for the quantity to be determined (the unknown).* If there is a temperature change, the transferred heat depends on the specific heat (see Temperature Change and Heat Capacity) whereas, for a phase change, the transferred heat depends on the latent heat. See Table 5.8.1.
6. *Substitute the knowns along with their units into the appropriate equation and obtain numerical solutions complete with units.* You will need to do this in steps if there is more than one stage to the process (such as a temperature change followed by a phase change).
7. *Check the answer to see if it is reasonable: Does it make sense?* As an example, be certain that the temperature change does not also cause a phase change that you have not taken into account.

**Exercise 5.8.1**

Why does snow remain on mountain slopes even when daytime temperatures are higher than the freezing temperature?

*Answer available at the end of this section.*

## Summary

- Most substances can exist either in solid, liquid, and gas forms, which are referred to as “phases.”
- Phase changes occur at fixed temperatures for a given substance at a given pressure, and these temperatures are called boiling and freezing (or melting) points.
- During phase changes, heat absorbed or released is given by $Q = mL$, where $L$ is the latent heat coefficient.

## Footnotes

1 Values quoted at the normal melting and boiling temperatures at standard atmospheric pressure (1 atm).
2 At 37.0°C (body temperature), the heat of vaporization $L_v$ for water is 2430 kJ/kg or 580 kcal/kg
3 At 37.0°C (body temperature), the heat of vaporization $L_v$ for water is 2430 kJ/kg or 580 kcal/kg

## Glossary

**heat of sublimation**
the energy required to change a substance from the solid phase to the vapor phase

**latent heat coefficient**
a physical constant equal to the amount of heat transferred for every 1 kg of a substance during the change in phase of the substance

**sublimation**
the transition from the solid phase to the vapor phase

## Contributors and Attributions

Paul Peter Urone (Professor Emeritus at California State University, Sacramento) and Roger Hinrichs (State University of New York, College at Oswego) with Contributing Authors: Kim Dirks (University of Auckland) and Manjula Sharma (University of Sydney). This work is licensed by OpenStax University Physics under a Creative Commons Attribution License (by 4.0).

### Answers

**Exercise** 5.8.1

Snow is formed from ice crystals and thus is the solid phase of water. Because enormous heat is necessary for phase changes, it takes a certain amount of time for this heat to be accumulated from the air, even if the air is above $0^oC$. The warmer the air is, the faster this heat exchange occurs and the faster the snow melts.

This page titled 5.8: Phase Change and Latent Heat is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.3: Phase Change and Latent Heat by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.9: Heat Transfer Methods

## Learning Objectives

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

- Discuss the different methods of heat transfer.

Equally as interesting as the effects of heat transfer on a system are the methods by which this occurs. Whenever there is a temperature difference, heat transfer occurs. Heat transfer may occur rapidly, such as through a cooking pan, or slowly, such as through the walls of a picnic ice chest. We can control rates of heat transfer by choosing materials (such as thick wool clothing for the winter), controlling air movement (such as the use of weather stripping around doors), or by choice of color (such as a white roof to reflect summer sunlight). So many processes involve heat transfer, so that it is hard to imagine a situation where no heat transfer occurs. Yet every process involving heat transfer takes place by only three methods:

1. **Conduction** is heat transfer through stationary matter by physical contact. (The matter is stationary on a macroscopic scale—we know there is thermal motion of the atoms and molecules at any temperature above absolute zero.) Heat transferred between the electric burner of a stove and the bottom of a pan is transferred by conduction.
2. **Convection** is the heat transfer by the macroscopic movement of a fluid. This type of transfer takes place in a forced-air furnace and in weather systems, for example.
3. Heat transfer by **radiation** occurs when microwaves, infrared radiation, visible light, or another form of electromagnetic radiation is emitted or absorbed. An obvious example is the warming of the Earth by the Sun. A less obvious example is thermal radiation from the human body.

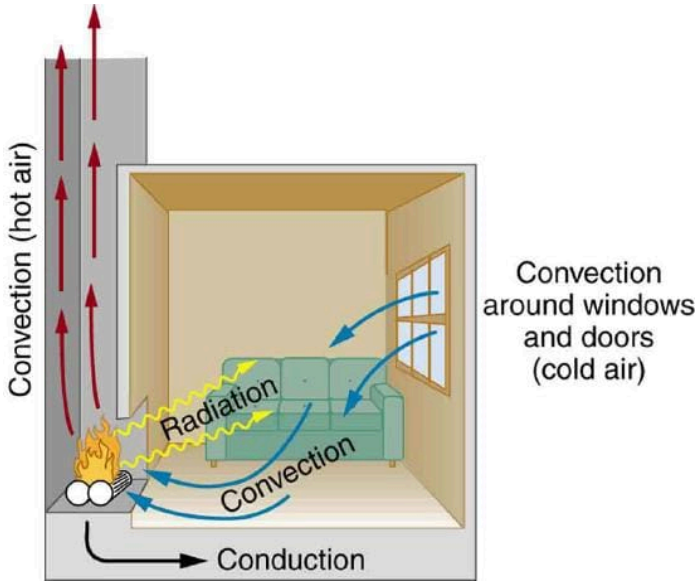



Figure 5.9.1: In a fireplace, heat transfer occurs by all three methods: conduction, convection, and radiation. Radiation is responsible for most of the heat transferred into the room. Heat transfer also occurs through conduction into the room, but at a much slower rate. Heat transfer by convection also occurs through cold air entering the room around windows and hot air leaving the room by rising up the chimney.

We examine these methods in some detail in the three following modules. Each method has unique and interesting characteristics, but all three do have one thing in common: they transfer heat solely because of a temperature difference Figure 5.9.1.

**Exercise 5.9.1**

Name an example from daily life (different from the text) for each mechanism of heat transfer.

*Answer available at the end of this section.*

## Summary

- Heat is transferred by three different methods: conduction, convection, and radiation.

## Glossary

**conduction**
: heat transfer through stationary matter by physical contact

**convection**
: heat transfer by the macroscopic movement of fluid

**radiation**
: heat transfer which occurs when microwaves, infrared radiation, visible light, or other electromagnetic radiation is emitted or absorbed

## Answers

**Exercise** 5.9.1

- Conduction: Heat transfers into your hands as you hold a hot cup of coffee.
- Convection: Heat transfers as the barista “steams” cold milk to make hot cocoa.
- Radiation: Reheating a cold cup of coffee in a microwave oven.

This page titled 5.9: Heat Transfer Methods is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.4: Heat Transfer Methods by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.10: Conduction

## Learning Objectives

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

- Calculate thermal conductivity.
- Observe conduction of heat in collisions.
- Study thermal conductivities of common substances.

Your feet feel cold as you walk barefoot across the living room carpet in your cold house and then step onto the kitchen tile floor. This result is intriguing, since the carpet and tile floor are both at the same temperature. The different sensation you feel is explained by the different rates of heat transfer: the heat loss during the same time interval is greater for skin in contact with the tiles than with the carpet, so the temperature drop is greater on the tiles.



Figure $5.10.1$: Insulation is used to limit the conduction of heat from the inside to the outside (in winters) and from the outside to the inside (in summers). (credit: Giles Douglas)

Some materials conduct thermal energy faster than others. In general, good conductors of electricity (metals like copper, aluminum, gold, and silver) are also good heat conductors, whereas insulators of electricity (wood, plastic, and rubber) are poor heat conductors. Figure $5.10.2$ shows molecules in two bodies at different temperatures. The (average) kinetic energy of a molecule in the hot body is higher than in the colder body. If two molecules collide, an energy transfer from the molecule with greater kinetic energy to the molecule with less kinetic energy occurs. The cumulative effect from all collisions results in a net flux of heat from the hot body to the colder body. The heat flux thus depends on the temperature difference $\Delta T = T_{hot} - T_{cold}$ Therefore, you will get a more severe burn from boiling water than from hot tap water. Conversely, if the temperatures are the same, the net heat transfer rate falls to zero, and equilibrium is achieved. Owing to the fact that the number of collisions increases with increasing area, heat conduction depends on the cross-sectional area. If you touch a cold wall with your palm, your hand cools faster than if you just touch it with your fingertip.

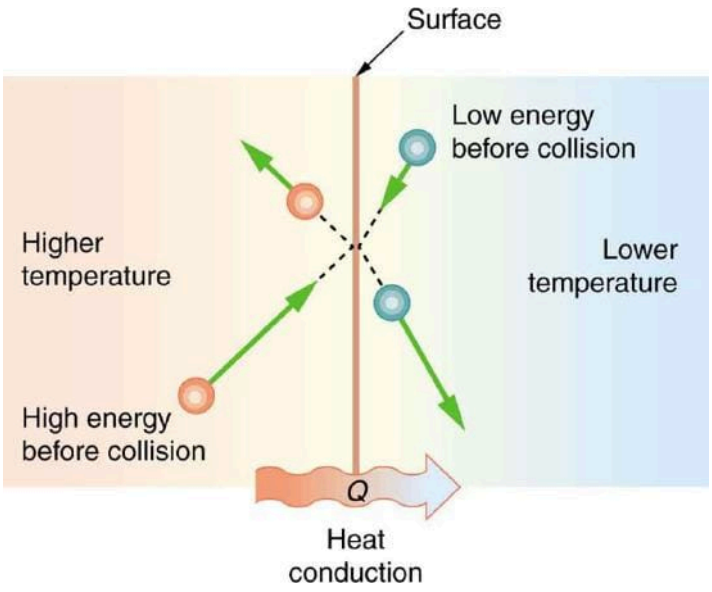



Figure 5.10.2: The molecules in two bodies at different temperatures have different average kinetic energies. Collisions occurring at the contact surface tend to transfer energy from high-temperature regions to low-temperature regions. In this illustration, a molecule in the lower temperature region (right side) has low energy before collision, but its energy increases after colliding with the contact surface. In contrast, a molecule in the higher temperature region (left side) has high energy before collision, but its energy decreases after colliding with the contact surface.

A third factor in the mechanism of conduction is the thickness of the material through which heat transfers. The figure below shows a slab of material with different temperatures on either side. Suppose that $T_2$ is greater than $T_1$ so that heat is transferred from left to right. Heat transfer from the left side to the right side is accomplished by a series of molecular collisions. The thicker the material, the more time it takes to transfer the same amount of heat. This model explains why thick clothing is warmer than thin clothing in winters, and why Arctic mammals protect themselves with thick blubber.

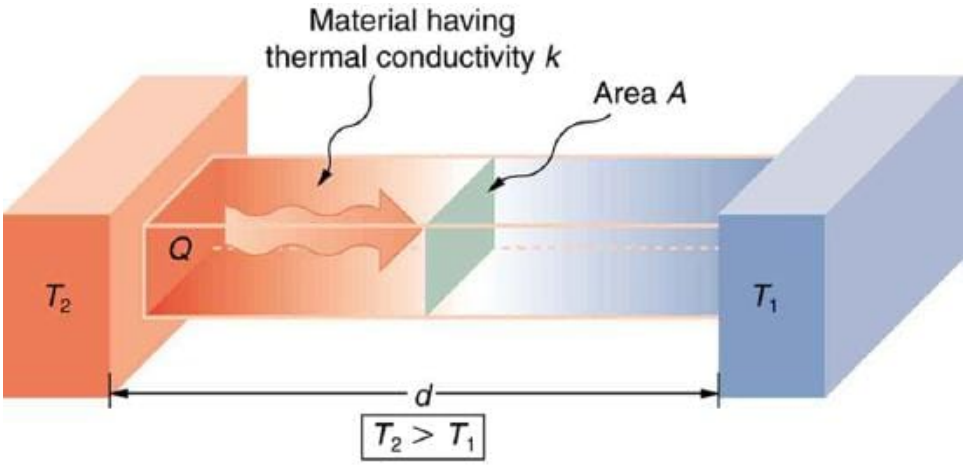



Figure 5.10.3: Heat conduction occurs through any material, represented here by a rectangular bar, whether window glass or walrus blubber. The temperature of the material is $T_2$ on the left and $T_1$ on the right, where $T_2$ is greater than $T_1$ The rate of heat transfer by conduction is directly proportional to the surface area $A$, the temperature difference $T_2 - T_1$, and the substance's conductivity $k$. The rate of heat transfer is inversely proportional to the thickness $d$.

Lastly, the heat transfer rate depends on the material properties described by the coefficient of **thermal conductivity**. All four factors are included in a simple equation that was deduced from and is confirmed by experiments. The rate of conductive heat transfer through a slab of material, such as the one in Figure 5.10.3, is given by

$$\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}, \tag{5.10.1}$$

where $Q/t$ is the rate of heat transfer in watts or kilocalories per second, $k$ s the thermal conductivity of the material, $A$ and $d$ are

its surface area and thickness and $(T_2 - T_1)$ is the temperature difference across the slab. Table 5.10.1gives representative values of thermal conductivity.

## Example 5.10.1: Calculating Heat Transfer Through Conduction: Conduction Rate Through an Ice Box

A Styrofoam ice box has a total area of $0.950\ m^2$ and walls with an average thickness of 2.50 cm. The box contains ice, water, and canned beverages at $0^oC$. The inside of the box is kept cold by melting ice. How much ice melts in one day if the ice box is kept in the trunk of a car at $35.0^oC$?

**Strategy**

This question involves both heat for a phase change (melting of ice) and the transfer of heat by conduction. To find the amount of ice melted, we must find the net heat transferred. This value can be obtained by calculating the rate of heat transfer by conduction and multiplying by time.

**Solution**

1. Identify the knowns.
$$A = 0.950\,m^2;\ d = 2.50\,cm = 0.0250\,m;\ T_1 = 35.0^oC,\ t = 1\,day = 24\,hours = 86,400\,sec. \tag{5.10.2}$$
2. Identify the unknowns. We need to solve for the mass of the ice $m$. We will also need to solve for the net heat transferred to melt the ice, $Q$.
3. Determine which equations to use. The rate of heat transfer by conduction is given by
$$\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}. \tag{5.10.3}$$
4. The heat is used to melt the ice: $Q = mL_f$.
5. Insert the known values:
$$\frac{Q}{t} = \frac{(0.010\,J/s \cdot m \cdot^o C)(0.950\,m^2)(35.0^oC - 0^oC)}{0.0250\,m} = 13.3\,J/s. \tag{5.10.4}$$
6. Multiply the rate of heat transfer by the time $(1\ day = 86,400s)$:
$$Q = (Q/t)t = (13.3\,J/s)(86,400\,s) = 1.15 \times 10^6\,J. \tag{5.10.5}$$
7. Set this equal to the heat transferred to melt the ice: $Q = mL_f$. Solve for the mass $m$:
$$m = \frac{Q}{L_f} = \frac{1.15 \times 10^6\,J}{334 \times 10^3\,J/kg} = 3.44\,kg. \tag{5.10.6}$$

**Discussion**

The result of 3.44 kg, or about 7.6 lbs, seems about right, based on experience. You might expect to use about a 4 kg (7–10 lb) bag of ice per day. A little extra ice is required if you add any warm food or beverages.

Inspecting the conductivities in Table 5.10.1 shows that Styrofoam is a very poor conductor and thus a good insulator. Other good insulators include fiberglass, wool, and goose-down feathers. Like Styrofoam, these all incorporate many small pockets of air, taking advantage of air's poor thermal conductivity.

*Table 5.10.1: Thermal Conductivities of Common Substances*

| Substance | Thermal conductivity |
|---|---|
| Air | 0.023 |
| Aluminum | 220 |
| Asbestos | 0.16 |
| Concrete brick | 0.84 |
| Copper | 390 |

| Substance | Thermal conductivity |
|---|---|
| Cork | 0.042 |
| Down feathers | 0.025 |
| Fatty tissue (without blood) | 0.2 |
| Glass (average) | 0.84 |
| Glass wool | 0.042 |
| Gold | 318 |
| Ice | 2.2 |
| Plasterboard | 0.16 |
| Silver | 420 |
| Snow (dry) | 0.10 |
| Steel (stainless) | 14 |
| Steel iron | 80 |
| Styrofoam | 0.010 |
| Water | 0.6 |
| Wood | 0.08–0.16 |
| Wool | 0.04 |

Thermal Conductivities of Common Substances[1]

A combination of material and thickness is often manipulated to develop good insulators—the smaller the conductivity $k$ and the larger the thickness $d$, the better. The ratio of $d/k$ will thus be large for a good insulator. The ratio $d/k$ is called the $R$ **factor.** The rate of conductive heat transfer is inversely proportional to $R$. The larger the value of $R$, the better the insulation. $R$ factors are most commonly quoted for household insulation, refrigerators, and the like—unfortunately, it is still in non-metric units of $ft^2 \cdot^o F \cdot h/Btu$, although the unit usually goes unstated (1 British thermal unit [Btu] is the amount of energy needed to change the temperature of 1.0 lb of water by 1.0 °F). A couple of representative values are an $R$ factor of 11 for 3.5-in-thick fiberglass batts (pieces) of insulation and an $R$ factor of 19 for 6.5-in-thick fiberglass batts. Walls are usually insulated with 3.5-in batts, while ceilings are usually insulated with 6.5-in batts. In cold climates, thicker batts may be used in ceilings and walls.



Figure $5.10.4$:The fiberglass batt is used for insulation of walls and ceilings to prevent heat transfer between the inside of the building and the outside environment. (CC BY-SA 3.0; Radomil).

Note that in Table $5.10.1$, the best thermal conductors—silver, copper, gold, and aluminum—are also the best electrical conductors, again related to the density of free electrons in them. Cooking utensils are typically made from good conductors.

## Example 5.10.1: Calculating the Temperature Difference Maintained by a Heat Transfer: Conduction Through an Aluminum Pan

Water is boiling in an aluminum pan placed on an electrical element on a stovetop. The sauce pan has a bottom that is 0.800 cm thick and 14.0 cm in diameter. The boiling water is evaporating at the rate of 1.00 g/s. What is the temperature difference across (through) the bottom of the pan?

**Strategy**

Conduction through the aluminum is the primary method of heat transfer here, and so we use the equation for the rate of heat transfer and solve for the temperature difference.

$$T_2 - T_1 = \frac{Q}{t}\left(\frac{d}{kA}\right). \tag{5.10.7}$$

**Solution**

1. Identify the knowns and convert them to the SI units.

   The thickness of the pan, $d = 0.800\ cm = 8.0 \times 10^{-3} m$, the area of the pan, $A = \pi(0.14/2)^2\ m^2 = 1.54 \times 10^{-2}\ m^2$, and the thermal conductivity, $k = 220\ J/s \cdot m \cdot^o C$.

2. Calculate the necessary heat of vaporization of 1 g of water:

$$Q = mL_v = (1.00 \times 10^{-3}\, kg)(2256 \times 10^3\, J/kg) = 2256\, J. \tag{5.10.8}$$

3. Calculate the rate of heat transfer given that 1 g of water melts in one second:

$$Q/t = 2256\, J/s\, or 2.26 kW. \tag{5.10.9}$$

4. Insert the knowns into the equation and solve for the temperature difference:

$$T_2 - T_1 = \frac{Q}{t}\left(\frac{d}{kA}\right) = (2256\, J/s)\frac{8.00 \times 10^{-3} m}{(220\, J/s \cdot m \cdot^o C)(1.54 \times 10^{-2}\, m^2)} = 5.33^o C. \tag{5.10.10}$$

**Discussion**

The value for the heat transfer $Q/t = 2.26\ kW\ or 2256\ J/s$ is typical for an electric stove. This value gives a remarkably small temperature difference between the stove and the pan. Consider that the stove burner is red hot while the inside of the pan is nearly $100^o C$ because of its contact with boiling water. This contact effectively cools the bottom of the pan in spite of its proximity to the very hot stove burner. Aluminum is such a good conductor that it only takes this small temperature difference to produce a heat transfer of 2.26 kW into the pan.

Conduction is caused by the random motion of atoms and molecules. As such, it is an ineffective mechanism for heat transport over macroscopic distances and short time distances. Take, for example, the temperature on the Earth, which would be unbearably cold during the night and extremely hot during the day if heat transport in the atmosphere was to be only through conduction. In another example, car engines would overheat unless there was a more efficient way to remove excess heat from the pistons.

## Exercise 5.10.1: Check your understanding

How does the rate of heat transfer by conduction change when all spatial dimensions are doubled?

*Answer available at the end of this section.*

# Summary

Heat conduction is the transfer of heat between two objects in direct contact with each other. The rate of heat transfer $Q/t$ (energy per unit time) is proportional to the temperature difference $T_2 - T_1$ and the contact area $A$ and inversely proportional to the

distance between the objects:

$$\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}. \tag{5.10.11}$$

## Footnotes

1. At temperatures near 0°C.

# Glossary

***R* factor**
: the ratio of thickness to the conductivity of a material

**rate of conductive heat transfer**
: rate of heat transfer from one material to another

**thermal conductivity**
: the property of a material's ability to conduct heat

## Answers

**Exercise** 5.10.1**: Check your understanding**

Because area is the product of two spatial dimensions, it increases by a factor of four when each dimension is doubled $(A_{final} = (2d)^2 = 4d^2 = 4A_{initial})$. The distance, however, simply doubles. Because the temperature difference and the coefficient of thermal conductivity are independent of the spatial dimensions, the rate of heat transfer by conduction increases by a factor of four divided by two, or two:

$$\left(\frac{Q}{t}\right)_{final} = \frac{kA_{final}(T_2 - T_1)}{d_{final}} = \frac{k(4A_{initial})(T_2 - T_1)}{2d_{initial}} = 2\frac{kA_{initial}(T_2 - T_1)}{d_{initial}} = 2\left(\frac{Q}{t}\right)_{initial} \tag{5.10.12}$$

This page titled 5.10: Conduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.5: Conduction by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.11: Convection

## Learning Objectives

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

- Discuss the method of heat transfer by convection.

Convection is driven by large-scale flow of matter. In the case of Earth, the atmospheric circulation is caused by the flow of hot air from the tropics to the poles, and the flow of cold air from the poles toward the tropics. (Note that Earth's rotation causes the observed easterly flow of air in the northern hemisphere). Car engines are kept cool by the flow of water in the cooling system, with the water pump maintaining a flow of cool water to the pistons. The circulatory system is used the body: when the body overheats, the blood vessels in the skin expand (dilate), which increases the blood flow to the skin where it can be cooled by sweating. These vessels become smaller when it is cold outside and larger when it is hot (so more fluid flows, and more energy is transferred).

The body also loses a significant fraction of its heat through the breathing process.

While convection is usually more complicated than conduction, we can describe convection and do some straightforward, realistic calculations of its effects. Natural convection is driven by buoyant forces: hot air rises because density decreases as temperature increases. The house in Figure 5.11.1 is kept warm in this manner, as is the pot of water on the stove in Figure 5.11.2. Ocean currents and large-scale atmospheric circulation transfer energy from one part of the globe to another. Both are examples of natural convection.

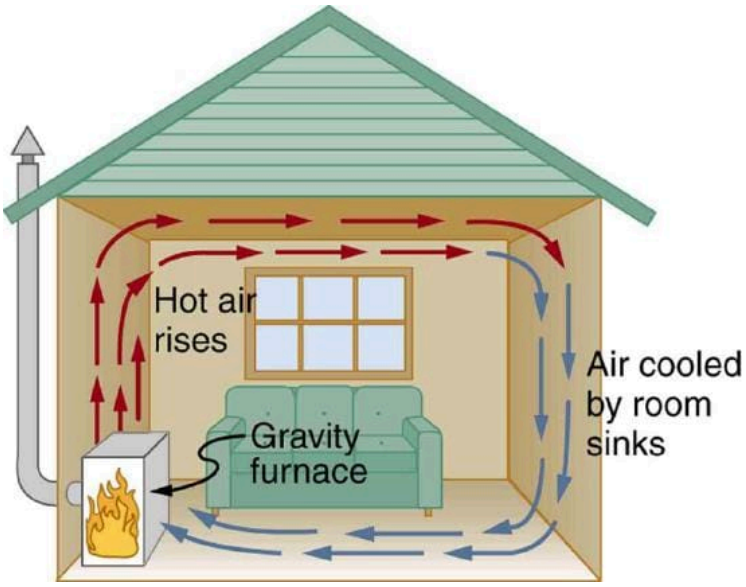



Figure 5.11.1: Air heated by the so-called gravity furnace expands and rises, forming a convective loop that transfers energy to other parts of the room. As the air is cooled at the ceiling and outside walls, it contracts, eventually becoming denser than room air and sinking to the floor. A properly designed heating system using natural convection, like this one, can be quite efficient in uniformly heating a home.

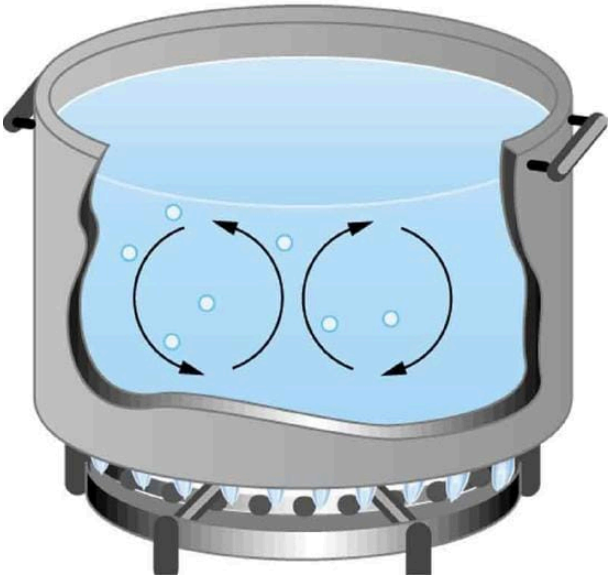


Figure 5.11.2: Convection plays an important role in heat transfer inside this pot of water. Once conducted to the inside, heat transfer to other parts of the pot is mostly by convection. The hotter water expands, decreases in density, and rises to transfer heat to other regions of the water, while colder water sinks to the bottom. This process keeps repeating.

### Take-Home Experiment: Convection Rolls in a Heated Pan

Take two small pots of water and use an eye dropper to place a drop of food coloring near the bottom of each. Leave one on a bench top and heat the other over a stovetop. Watch how the color spreads and how long it takes the color to reach the top. Watch how convective loops form.

### Example 5.11.1: Calculating Heat Transfer by Convection: Convection of Air Through the Walls of a House

Most houses are not airtight: air goes in and out around doors and windows, through cracks and crevices, following wiring to switches and outlets, and so on. The air in a typical house is completely replaced in less than an hour. Suppose that a moderately-sized house has inside dimensions $12.0m \times 18.0m \times 3.00m$ high, and that all air is replaced in 30.0 min. Calculate the heat transfer per unit time in watts needed to warm the incoming cold air by $10.0^oC$ thus replacing the heat transferred by convection alone.

**Strategy**

Heat is used to raise the temperature of air so that $Q = mc\Delta T$. The rate of heat transfer is then $Q/t$, where $t$ is the time for air turnover. We are given that $\Delta T$ is $10.0^oC$, but we must still find values for the mass of air and its specific heat before we can calculate $Q$. The specific heat of air is a weighted average of the specific heats of nitrogen and oxygen, which gives $c = c_p \approx 1000\ J/kg \cdot^o C$ from the table (note that the specific heat at constant pressure must be used for this process).

**Solution**

1. Determine the mass of air from its density and the given volume of the house. The density is given from the density $\rho$ and the volume
$$m = \rho V = (1.29\,kg/m^3)912.0\,m \times 18.0\,m \times 3.00\,m) = 836\,kg \tag{5.11.1}$$
2. Calculate the heat transferred from the change in air temperature $Q = mc\Delta T$ so that
$$Q = (836\,kg)(1000\,J/kg \cdot^o C)(10.0^oC) = 8.36 \times 10^6\,J. \tag{5.11.2}$$
3. Calculate the heat transfer from the heat $Q$ and the turnover time $t$. Since air is turned over in $t = 0.500\ h = 1800\ s$, the heat transferred per unit time is

$$\frac{Q}{t} = \frac{8.36 \times 10^6\, J}{1800\, s} = 4.64\, kW. \tag{5.11.3}$$

**Discussion**

This rate of heat transfer is equal to the power consumed by about forty-six 100-W light bulbs. Newly constructed homes are designed for a turnover time of 2 hours or more, rather than 30 minutes for the house of this example. Weather stripping, caulking, and improved window seals are commonly employed. More extreme measures are sometimes taken in very cold (or hot) climates to achieve a tight standard of more than 6 hours for one air turnover. Still longer turnover times are unhealthy, because a minimum amount of fresh air is necessary to supply oxygen for breathing and to dilute household pollutants. The term used for the process by which outside air leaks into the house from cracks around windows, doors, and the foundation is called “air infiltration.”

A cold wind is much more chilling than still cold air, because convection combines with conduction in the body to increase the rate at which energy is transferred away from the body. The table below gives approximate wind-chill factors, which are the temperatures of still air that produce the same rate of cooling as air of a given temperature and speed. Wind-chill factors are a dramatic reminder of convection’s ability to transfer heat faster than conduction. For example, a 15.0 m/s wind at $0^oC$ has the chilling equivalent of still air at about $-18^oC$.

**Moving air temperature Wind speed (m/s)**

| $^oC$ | 2 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|
| **5** | 3 | -1 | -8 | -10 | -12 |
| **2** | 0 | -7 | -12 | -16 | -18 |
| **0** | -2 | -9 | -15 | -18 | -20 |
| **-5** | -7 | -15 | -22 | -26 | -29 |
| **-10** | -12 | -21 | -29 | -34 | -36 |
| **-20** | -23 | -34 | -44 | -50 | -52 |
| **-40** | -44 | -59 | -73 | -82 | -84 |

Although air can transfer heat rapidly by convection, it is a poor conductor and thus a good insulator. The amount of available space for airflow determines whether air acts as an insulator or conductor. The space between the inside and outside walls of a house, for example, is about 9 cm (3.5 in) —large enough for convection to work effectively. The addition of wall insulation prevents airflow, so heat loss (or gain) is decreased. Similarly, the gap between the two panes of a double-paned window is about 1 cm, which prevents convection and takes advantage of air’s low conductivity to prevent greater loss. Fur, fiber, and fiberglass also take advantage of the low conductivity of air by trapping it in spaces too small to support convection, as shown in the figure. Fur and feathers are lightweight and thus ideal for the protection of animals.

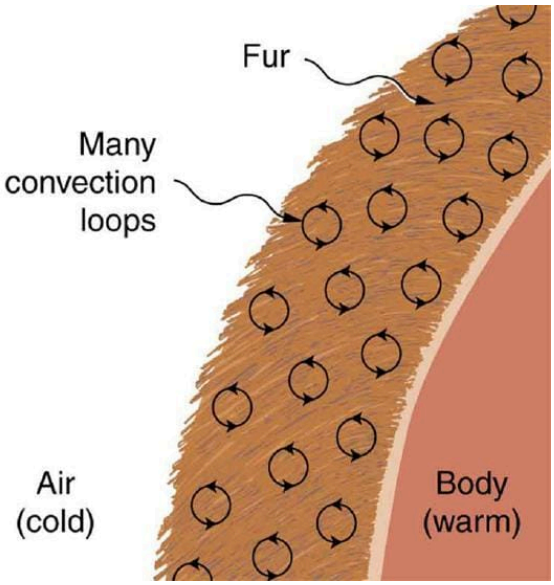



Figure 5.11.3: Fur is filled with air, breaking it up into many small pockets. Convection is very slow here, because the loops are so small. The low conductivity of air makes fur a very good lightweight insulator.

Some interesting phenomena happen *when convection is accompanied by a phase change*. It allows us to cool off by sweating, even if the temperature of the surrounding air exceeds body temperature. Heat from the skin is required for sweat to evaporate from the skin, but without air flow, the air becomes saturated and evaporation stops. Air flow caused by convection replaces the saturated air by dry air and evaporation continues.

## Example 5.11.2: Calculate the Flow of Mass during Convection: Sweat-Heat Transfer away from the Body

The average person produces heat at the rate of about 120 W when at rest. At what rate must water evaporate from the body to get rid of all this energy? (This evaporation might occur when a person is sitting in the shade and surrounding temperatures are the same as skin temperature, eliminating heat transfer by other methods.)

**Strategy**

Energy is needed for a phase change ($Q = mL_v$). Thus, the energy loss per unit time is

$$\frac{Q}{t} = \frac{mL_v}{t} = 120\,W = 120\,J/s. \tag{5.11.4}$$

We divide both sides of the equation by $L_v$ to find that the mass evaporated per unit time is

$$\frac{m}{t} = \frac{120\,J/s}{L_v}. \tag{5.11.5}$$

**Solution**

(1) Insert the value of the latent heat from [link], $L_v = 2430\ kJ/kg = 2430\ J/g$. This yields

$$\frac{m}{t} = \frac{120\,J/s}{2430\,J/g} = 0.0494\,g/s = 2.96\,g/min. \tag{5.11.6}$$

**Discussion**

Evaporating about 3 g/min seems reasonable. This would be about 180 g (about 7 oz) per hour. If the air is very dry, the sweat may evaporate without even being noticed. A significant amount of evaporation also takes place in the lungs and breathing passages.

Another important example of the combination of phase change and convection occurs when water evaporates from the oceans.

Heat is removed from the ocean when water evaporates. If the water vapor condenses in liquid droplets as clouds form, heat is released in the atmosphere. Thus, there is an overall transfer of heat from the ocean to the atmosphere. This process is the driving power behind thunderheads, those great cumulus clouds that rise as much as 20.0 km into the stratosphere. Water vapor carried in by convection condenses, releasing tremendous amounts of energy. This energy causes the air to expand and rise, where it is colder. More condensation occurs in these colder regions, which in turn drives the cloud even higher. Such a mechanism is called positive feedback, since the process reinforces and accelerates itself. These systems sometimes produce violent storms, with lightning and hail, and constitute the mechanism driving hurricanes.



Figure 5.11.4: Cumulus clouds are caused by water vapor that rises because of convection. The rise of clouds is driven by a positive feedback mechanism. (credit: Mike Love)



Figure 5.11.5: Convection accompanied by a phase change releases the energy needed to drive this thunderhead into the stratosphere.



(credit: Gerardo García Moretti ) Figure 5.11.6: The phase change that occurs when this iceberg melts involves tremendous heat transfer. (credit: Dominic Alves)

The movement of icebergs is another example of convection accompanied by a phase change. Suppose an iceberg drifts from Greenland into warmer Atlantic waters. Heat is removed from the warm ocean water when the ice melts and heat is released to the land mass when the iceberg forms on Greenland.

**Exercise 5.11.1**

Explain why using a fan in the summer feels refreshing!

*Answer available at the end of this section.*

## Summary

- Convection is heat transfer by the macroscopic movement of mass. Convection can be natural or forced and generally transfers thermal energy faster than conduction. The table gives wind-chill factors, indicating that moving air has the same chilling effect of much colder stationary air. *Convection that occurs along with a phase change* can transfer energy from cold regions to warm ones.

### Answers

**Exercise** 5.11.1

Using a fan increases the flow of air: warm air near your body is replaced by cooler air from elsewhere. Convection increases the rate of heat transfer so that moving air “feels” cooler than still air.

This page titled 5.11: Convection is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.6: Convection by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.12: Radiation

## Learning Objectives

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

- Discuss heat transfer by radiation.
- Explain the power of different materials.

You can feel the heat transfer from a fire and from the Sun. Similarly, you can sometimes tell that the oven is hot without touching its door or looking inside—it may just warm you as you walk by. The space between the Earth and the Sun is largely empty, without any possibility of heat transfer by convection or conduction. In these examples, heat is transferred by radiation. That is, the hot body emits electromagnetic waves that are absorbed by our skin: no medium is required for electromagnetic waves to propagate. Different names are used for electromagnetic waves of different wavelengths: radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma rays.



Figure 5.12.1: Most of the heat transfer from this fire to the observers is through infrared radiation. The visible light, although dramatic, transfers relatively little thermal energy. Convection transfers energy away from the observers as hot air rises, while conduction is negligibly slow here. Skin is very sensitive to infrared radiation, so that you can sense the presence of a fire without looking at it directly. (credit: Daniel X. O'Neil)

The energy of electromagnetic radiation depends on the wavelength (color) and varies over a wide range: a smaller wavelength (or higher frequency) corresponds to a higher energy. Because more heat is radiated at higher temperatures, a temperature change is accompanied by a color change. Take, for example, an electrical element on a stove, which glows from red to orange, while the higher-temperature steel in a blast furnace glows from yellow to white. The radiation you feel is mostly infrared, which corresponds to a lower temperature than that of the electrical element and the steel. The radiated energy depends on its intensity, which is represented in the figure below by the height of the distribution.

Electromagnetic Waves explains more about the electromagnetic spectrum and Introduction to Quantum Physics discusses how the decrease in wavelength corresponds to an increase in energy.

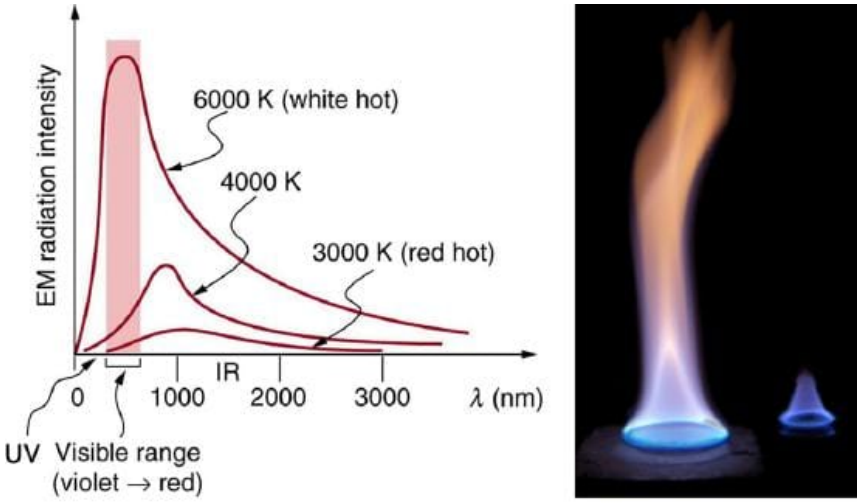



Figure 5.12.2: (a) A graph of the spectra of electromagnetic waves emitted from an ideal radiator at three different temperatures. The intensity or rate of radiation emission increases dramatically with temperature, and the spectrum shifts toward the visible and ultraviolet parts of the spectrum. The shaded portion denotes the visible part of the spectrum. It is apparent that the shift toward the ultraviolet with temperature makes the visible appearance shift from red to white to blue as temperature increases. (b) Note the variations in color corresponding to variations in flame temperature. (credit: Tuohirulla)

All objects absorb and emit electromagnetic radiation. The rate of heat transfer by radiation is largely determined by the color of the object. Black is the most effective, and white is the least effective. People living in hot climates generally avoid wearing black clothing, for instance. Similarly, black asphalt in a parking lot will be hotter than adjacent gray sidewalk on a summer day, because black absorbs better than gray. The reverse is also true—black radiates better than gray. Thus, on a clear summer night, the asphalt will be colder than the gray sidewalk, because black radiates the energy more rapidly than gray. An *ideal radiator* is the same color as an *ideal absorber*, and captures all the radiation that falls on it. In contrast, white is a poor absorber and is also a poor radiator. A white object reflects all radiation, like a mirror. (A perfect, polished white surface is mirror-like in appearance, and a crushed mirror looks white.)

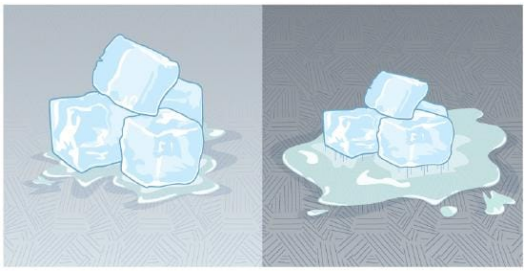


Figure 5.12.3: This illustration shows that the darker pavement is hotter than the lighter pavement (much more of the ice on the right has melted), although both have been in the sunlight for the same time. The thermal conductivities of the pavements are the same.

Gray objects have a uniform ability to absorb all parts of the electromagnetic spectrum. Colored objects behave in similar but more complex ways, which gives them a particular color in the visible range and may make them special in other ranges of the nonvisible spectrum. Take, for example, the strong absorption of infrared radiation by the skin, which allows us to be very sensitive to it.

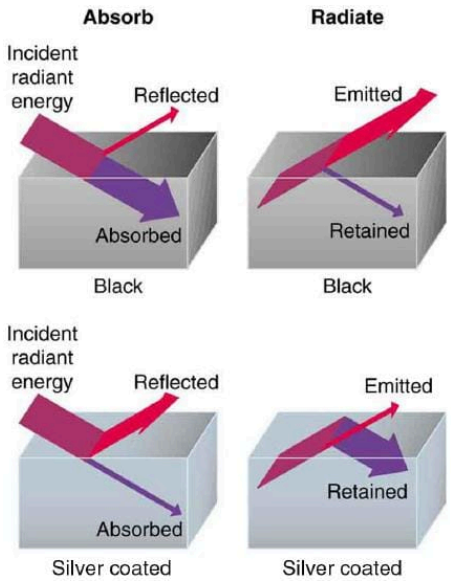



Figure 5.12.4: A black object is a good absorber and a good radiator, while a white (or silver) object is a poor absorber and a poor radiator. It is as if radiation from the inside is reflected back into the silver object, whereas radiation from the inside of the black object is "absorbed" when it hits the surface and finds itself on the outside and is strongly emitted.

The rate of heat transfer by emitted radiation is determined by the **Stefan-Boltzmann law of radiation:**

$$\frac{Q}{t} = \sigma e A T^4, \tag{5.12.1}$$

where $\sigma = 5.67 \times 10^{-8}\ J/s \cdot m^2 \cdot k^4$ is the Stefan-Boltzmann constant, $A$ is the surface area of the object, and $T$ is its absolute temperature in kelvin. The symbol $e$ stands for the emissivity of the object, which is a measure of how well it radiates. An ideal jet-black (or black body) radiator has $e = 1$, whereas a perfect reflector has $e = 0$. Real objects fall between these two values. Take, for example, tungsten light bulb filaments which have an $e$ of about 0.5, and carbon black (a material used in printer toner), which has the (greatest known) emissivity of about 0.99.

The radiation rate is directly proportional to the *fourth power* of the absolute temperature—a remarkably strong temperature dependence. Furthermore, the radiated heat is proportional to the surface area of the object. If you knock apart the coals of a fire, there is a noticeable increase in radiation due to an increase in radiating surface area.



Figure 5.12.5: A thermograph of part of a building shows temperature variations, indicating where heat transfer to the outside is most severe. Windows are a major region of heat transfer to the outside of homes. (credit: U.S. Army)

Skin is a remarkably good absorber and emitter of infrared radiation, having an emissivity of 0.97 in the infrared spectrum. Thus, we are all nearly (jet) black in the infrared, in spite of the obvious variations in skin color. This high infrared emissivity is why we can so easily feel radiation on our skin. It is also the basis for the use of night scopes used by law enforcement and the military to detect human beings. Even small temperature variations can be detected because of the $T^4$ dependence. Images, called *thermographs*, can be used medically to detect regions of abnormally high temperature in the body, perhaps indicative of disease. Similar techniques can be used to detect heat leaks in homes Figure 5.12.5, optimize performance of blast furnaces, improve comfort levels in work environments, and even remotely map the Earth's temperature profile.

All objects emit and absorb radiation. The *net* rate of heat transfer by radiation (absorption minus emission) is related to both the temperature of the object and the temperature of its surroundings. Assuming that an object with a temperature $T_1$ is surrounded by an environment with uniform temperature $T_2$ the net rate of heat transfer by radiation is

$$\frac{Q_{net}}{t} = \sigma e A(T_2^4 - T_1^4), \tag{5.12.2}$$

where $e$ is the emissivity of the object alone. In other words, it does not matter whether the surroundings are white, gray, or black; the balance of radiation into and out of the object depends on how well it emits and absorbs radiation. When $T_2 > T_1$, the quantity $Q_{net}/t$ is positive; that is, the net heat transfer is from hot to cold.

### Take-Home Experiment: Temperature in the Sun

Place a thermometer out in the sunshine and shield it from direct sunlight using an aluminum foil. What is the reading? Now remove the shield, and note what the thermometer reads. Take a handkerchief soaked in nail polish remover, wrap it around the thermometer and place it in the sunshine. What does the thermometer read?

### Example 5.12.1: Calculate the Net Heat Transfer of a Person: Heat Transfer by Radiation

What is the rate of heat transfer by radiation, with an unclothed person standing in a dark room whose ambient temperature is $22.0^oC$. The person has a normal skin temperature of $33.0^oC$ and a surface area of $1.50\ m^2$. The emissivity of skin is 0.97 in the infrared, where the radiation takes place.

**Strategy**

We can solve this by using the equation for the rate of radiative heat transfer.

**Solution**

Insert the temperatures values $T_2 = 295\ K$ and $T_1 = 306\ K$, so that

$$\frac{Q}{t} = \sigma e A(T_2^4 - T_1^4) \tag{5.12.3}$$

$$= (5.67 \times 10^{-8}\, J/s \cdot m^2 \cdot k^4)(0.97)(1.50\, m^2)[(295\, K)^4 - (306\, K)^4] \tag{5.12.4}$$

$$= -99\, J/s = -99\, W. \tag{5.12.5}$$

**Discussion**

This value is a significant rate of heat transfer to the environment (note the minus sign), considering that a person at rest may produce energy at the rate of 125 W and that conduction and convection will also be transferring energy to the environment. Indeed, we would probably expect this person to feel cold. Clothing significantly reduces heat transfer to the environment by many methods, because clothing slows down both conduction and convection, and has a lower emissivity (especially if it is white) than skin.

The Earth receives almost all its energy from radiation of the Sun and reflects some of it back into outer space. Because the Sun is hotter than the Earth, the net energy flux is from the Sun to the Earth. However, the rate of energy transfer is less than the equation for the radiative heat transfer would predict because the Sun does not fill the sky. The average emissivity (e) of the Earth is about 0.65, but the calculation of this value is complicated by the fact that the highly reflective cloud coverage varies greatly from day to day. There is a negative feedback (one in which a change produces an effect that opposes that change) between clouds and heat transfer; greater temperatures evaporate more water to form more clouds, which reflect more radiation

back into space, reducing the temperature. The often mentioned **greenhouse effect** is directly related to the variation of the Earth's emissivity with radiation type (see the figure given below). The greenhouse effect is a natural phenomenon responsible for providing temperatures suitable for life on Earth. The Earth's relatively constant temperature is a result of the energy balance between the incoming solar radiation and the energy radiated from the Earth. Most of the infrared radiation emitted from the Earth is absorbed by carbon dioxide $(CO_2)$ and water $(H_2O)$ in the atmosphere and then re-radiated back to the Earth or into outer space. Re-radiation back to the Earth maintains its surface temperature about $40^oC$ higher than it would be if there was no atmosphere, similar to the way glass increases temperatures in a greenhouse.

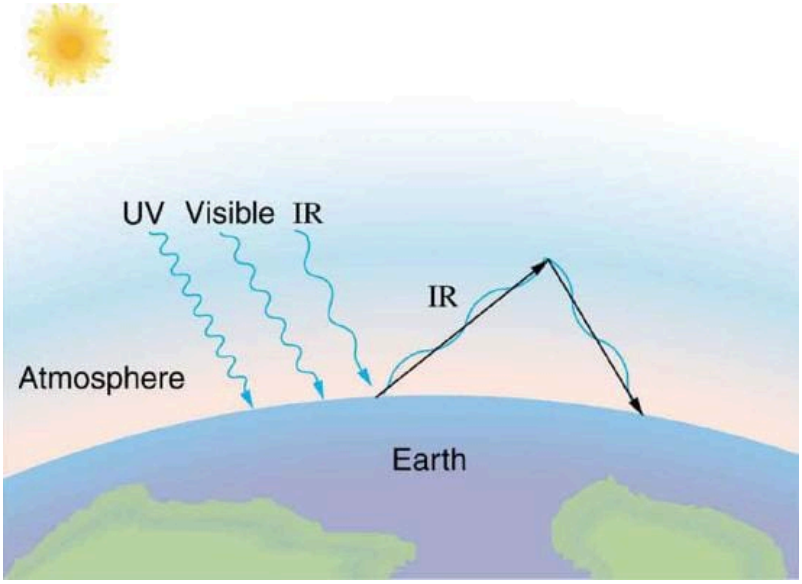



Figure $5.12.6$: The greenhouse effect is a name given to the trapping of energy in the Earth's atmosphere by a process similar to that used in greenhouses. The atmosphere, like window glass, is transparent to incoming visible radiation and most of the Sun's infrared. These wavelengths are absorbed by the Earth and re-emitted as infrared. Since Earth's temperature is much lower than that of the Sun, the infrared radiated by the Earth has a much longer wavelength. The atmosphere, like glass, traps these longer infrared rays, keeping the Earth warmer than it would otherwise be. The amount of trapping depends on concentrations of trace gases like carbon dioxide, and a change in the concentration of these gases is believed to affect the Earth's surface temperature.

The greenhouse effect is also central to the discussion of global warming due to emission of carbon dioxide and methane (and other so-called greenhouse gases) into the Earth's atmosphere from industrial production and farming. Changes in global climate could lead to more intense storms, precipitation changes (affecting agriculture), reduction in rain forest biodiversity, and rising sea levels.

Heating and cooling are often significant contributors to energy use in individual homes. Current research efforts into developing environmentally friendly homes quite often focus on reducing conventional heating and cooling through better building materials, strategically positioning windows to optimize radiation gain from the Sun, and opening spaces to allow convection. It is possible to build a zero-energy house that allows for comfortable living in most parts of the United States with hot and humid summers and cold winters.

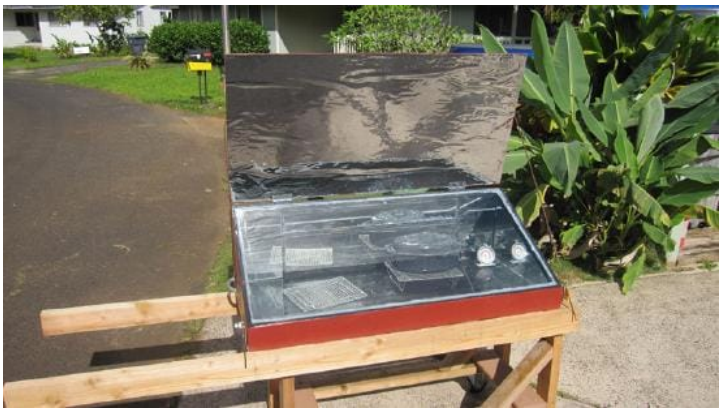


Figure $5.12.7$: This simple but effective solar cooker uses the

greenhouse effect and reflective material to trap and retain solar energy. Made of inexpensive, durable materials, it saves money and labor, and is of particular economic value in energy-poor developing countries. (credit: E.B. Kauai)

Conversely, dark space is very cold, about $3K(-454^oF)$, so that the Earth radiates energy into the dark sky. Owing to the fact that clouds have lower emissivity than either oceans or land masses, they reflect some of the radiation back to the surface, greatly reducing heat transfer into dark space, just as they greatly reduce heat transfer into the atmosphere during the day. The rate of heat transfer from soil and grasses can be so rapid that frost may occur on clear summer evenings, even in warm latitudes.

### Exercise 5.12.1

What is the change in the rate of the radiated heat by a body at the temperature $T_1 = 20^oC$ compared to when the body is at the temperature $T_2 = 40^oC$?

*Answer available at the end of this section.*

### Career Connection: Energy Conservation Consultation

The cost of energy is generally believed to remain very high for the foreseeable future. Thus, passive control of heat loss in both commercial and domestic housing will become increasingly important. Energy consultants measure and analyze the flow of energy into and out of houses and ensure that a healthy exchange of air is maintained inside the house. The job prospects for an energy consultant are strong.

### Problem Solving Strategies for the Methods of Heat Transfer

1. Examine the situation to determine what type of heat transfer is involved.
2. Identify the type(s) of heat transfer—conduction, convection, or radiation.
3. Identify exactly what needs to be determined in the problem (identify the unknowns). A written list is very useful.
4. Make a list of what is given or can be inferred from the problem as stated (identify the knowns).
5. Solve the appropriate equation for the quantity to be determined (the unknown).
6. For conduction, equation $\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}$ is appropriate. [link] lists thermal conductivities. For convection, determine the amount of matter moved and use equation $Q = mc\Delta T$, to calculate the heat transfer involved in the temperature change of the fluid. If a phase change accompanies convection, equation $Q = mL_f$ or $Q = mL_v$ is appropriate to find the heat transfer involved in the phase change. [link] lists information relevant to phase change. For radiation, equation $\frac{Q_{net}}{t} = \sigma eA(T_2^4 - T_1^4)$ gives the net heat transfer rate.
7. Insert the knowns along with their units into the appropriate equation and obtain numerical solutions complete with units.
8. Check the answer to see if it is reasonable. Does it make sense?

## Summary

- Radiation is the rate of heat transfer through the emission or absorption of electromagnetic waves.
- The rate of heat transfer depends on the surface area and the fourth power of the absolute temperature:

$$\frac{Q}{t} = \sigma eAT^4, \tag{5.12.6}$$

where $\sigma = 5.67 \times 10^{-8}\ J/s \cdot m^2 \cdot K^4$ is the Stefan-Boltzmann constant and $e$ is the emissivity of the body. For a black body, $e = 1$ whereas a shiny white or perfect reflector has $e = 0$, with real objects having values of e between 1 and 0. The net rate of heat transfer by radiation is

$$\frac{Q_{net}}{t} = \sigma eA(T_2^4 - T_1^4) \tag{5.12.7}$$

where $T_1$ is the temperature of an object surrounded by an environment with uniform temperature $T_2$ and $e$ is the emissivity of the *object*.

# Glossary

**emissivity**
: measure of how well an object radiates

**greenhouse effect**
: warming of the Earth that is due to gases such as carbon dioxide and methane that absorb infrared radiation from the Earth's surface and reradiate it in all directions, thus sending a fraction of it back toward the surface of the Earth

**net rate of heat transfer by radiation**
: is $\frac{Q_{net}}{t} = \sigma e A(T_2^4 - T_1^4)$

**radiation**
: energy transferred by electromagnetic waves directly as a result of a temperature difference

**Stefan-Boltzmann law of radiation**
: $\frac{Q}{t} = \sigma e A T^4$, where σ is the Stefan-Boltzmann constant, $A$ is the surface area of the object, $T$ is the absolute temperature, and $e$ is the emissivity

## Answers

**Exercise** 5.12.1

The radiated heat is proportional to the fourth power of the absolute temperature. Because $T_1 = 293\ K$ and $T_2 = 313\ K$, the rate of heat transfer increases by about 30 percent of the original rate.

This page titled 5.12: Radiation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 14.7: Radiation by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.13: Prelude to Thermodynamics

Heat transfer is energy in transit, and it can be used to do work. It can also be converted to any other form of energy. A car engine, for example, burns fuel for heat transfer into a gas. Work is done by the gas as it exerts a force through a distance, converting its energy into a variety of other forms—into the car's kinetic or gravitational potential energy; into electrical energy to run the spark plugs, radio, and lights; and back into stored energy in the car's battery. But most of the heat transfer produced from burning fuel in the engine does not do work on the gas. Rather, the energy is released into the environment, implying that the engine is quite inefficient.



Figure 5.13.1. A steam engine uses heat transfer to do work. Tourists regularly ride this narrow-gauge steam engine train near the San Juan Skyway in Durango, Colorado, part of the National Scenic Byways Program. (credit: Dennis Adams)

It is often said that modern gasoline engines cannot be made to be significantly more efficient. We hear the same about heat transfer to electrical energy in large power stations, whether they are coal, oil, natural gas, or nuclear powered. Why is that the case? Is the inefficiency caused by design problems that could be solved with better engineering and superior materials? Is it part of some money-making conspiracy by those who sell energy? Actually, the truth is more interesting, and reveals much about the nature of heat transfer.

Basic physical laws govern how heat transfer for doing work takes place and place insurmountable limits onto its efficiency. This chapter will explore these laws as well as many applications and concepts associated with them. These topics are part of *thermodynamics*—the study of heat transfer and its relationship to doing work.

This page titled 5.13: Prelude to Thermodynamics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 15.0: Prelude to Thermodynamics by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.14: The First Law of Thermodynamics

## Learning Objectives

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

- Define the first law of thermodynamics.
- Describe how conservation of energy relates to the first law of thermodynamics.
- Identify instances of the first law of thermodynamics working in everyday situations, including biological metabolism.
- Calculate changes in the internal energy of a system, after accounting for heat transfer and work done.

If we are interested in how heat transfer is converted into doing work, then the conservation of energy principle is important. The **first law of thermodynamics** applies the conservation of energy principle to systems where heat transfer and doing work are the methods of transferring energy into and out of the system.



Figure $5.14.1$: This boiling tea kettle represents energy in motion. The water in the kettle is turning to water vapor because heat is being transferred from the stove to the kettle. As the entire system gets hotter, work is done—from the evaporation of the water to the whistling of the kettle. (credit: Gina Hamilton)

The first law of thermodynamics states that the change in internal energy of a system equals the net heat transfer *into* the system minus the net work done *by* the system. In equation form, the first law of thermodynamics is

$$\Delta U = Q - W. \tag{5.14.1}$$

Here $\Delta U$ is the *change in internal energy* $U$ of the system. $Q$ is the *net heat transferred into the system*—that is, $Q$ is the sum of all heat transfer into and out of the system. $W$ is the *net work done by the system*—that is, $W$ is the sum of all work done on or by the system. We use the following sign conventions: if $Q$ is positive, then there is a net heat transfer into the system; if $W$ is positive, then there is net work done by the system. So positive $Q$ adds energy to the system and positive $W$ takes energy from the system. Thus $\Delta U = Q - W$. Note also that if more heat transfer into the system occurs than work done, the difference is stored as internal energy. Heat engines are a good example of this—heat transfer into them takes place so that they can do work (Figure $5.14.2$). We will now examine $Q,\ W$ and $\Delta U$ further.

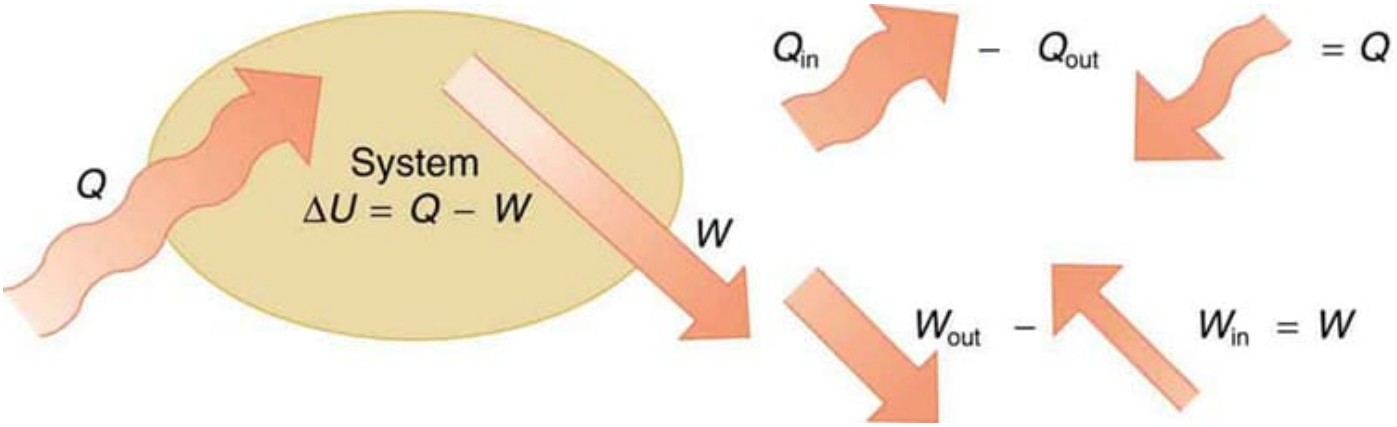



Figure 5.14.2: The first law of thermodynamics is the conservation-of-energy principle stated for a system where heat and work are the methods of transferring energy for a system in thermal equilibrium. $Q$ represents the net heat transfer—it is the sum of all heat transfers into and out of the system. $Q$ is positive for net heat transfer into the system. $W$ is the total work done on and by the system. $W$ is positive when more work is done by the system than on it. The change in the internal energy of the system, $\Delta U$, is related to heat and work by the first law of thermodynamics (Equation 5.14.1).

**LAW OF THERMODYNAMICS AND LAW OF CONSERVATION OF ENERGY**

The first law of thermodynamics is actually the law of conservation of energy stated in a form most useful in thermodynamics. The first law gives the relationship between heat transfer, work done, and the change in internal energy of a system.

## Heat Q and Work *W*

Heat transfer $Q$ and doing work $W$ are the two everyday means of bringing energy into or taking energy out of a system. The processes are quite different. Heat transfer, a less organized process, is driven by temperature differences. Work, a quite organized process, involves a macroscopic force exerted through a distance. Nevertheless, heat and work can produce identical results.For example, both can cause a temperature increase. Heat transfer into a system, such as when the Sun warms the air in a bicycle tire, can increase its temperature, and so can work done on the system, as when the bicyclist pumps air into the tire. Once the temperature increase has occurred, it is impossible to tell whether it was caused by heat transfer or by doing work. This uncertainty is an important point. Heat transfer and work are both energy in transit—neither is stored as such in a system. However, both can change the internal energy $U$ of a system. Internal energy is a form of energy completely different from either heat or work.

## Internal Energy *U*

We can think about the internal energy of a system in two different but consistent ways. The first is the atomic and molecular view, which examines the system on the atomic and molecular scale. The **internal energy** $U$ of a system is the sum of the kinetic and potential energies of its atoms and molecules. Recall that kinetic plus potential energy is called mechanical energy. Thus internal energy is the sum of atomic and molecular mechanical energy. Because it is impossible to keep track of all individual atoms and molecules, we must deal with averages and distributions. A second way to view the internal energy of a system is in terms of its macroscopic characteristics, which are very similar to atomic and molecular average values.

Macroscopically, we define the change in internal energy $\Delta U$ to be that given by the first law of thermodynamics (Equation 5.14.1 ):

$$\Delta U = Q - W$$

Many detailed experiments have verified that $\Delta U = Q - W$, where $\Delta U$ is the change in total kinetic and potential energy of all atoms and molecules in a system. It has also been determined experimentally that the internal energy $U$ of a system depends only on the state of the system and *not how it reached that state*. More specifically, $U$ is found to be a function of a few macroscopic quantities (pressure, volume, and temperature, for example), independent of past history such as whether there has been heat transfer or work done. This independence means that if we know the state of a system, we can calculate changes in its internal

energy $U$ from a few macroscopic variables.

## MACROSCOPIC vs. MICROSCOPIC

In thermodynamics, we often use the macroscopic picture when making calculations of how a system behaves, while the atomic and molecular picture gives underlying explanations in terms of averages and distributions. We shall see this again in later sections of this chapter. For example, in the topic of entropy, calculations will be made using the atomic and molecular view.

To get a better idea of how to think about the internal energy of a system, let us examine a system going from State 1 to State 2. The system has internal energy $U_1$ in State 1, and it has internal energy $U_2$ in State 2, no matter how it got to either state. So the change in internal energy

$$\Delta U = U_2 - U_1 \tag{5.14.2}$$

is independent of what caused the change. In other words, $\delta U$ *is independent of path*. By path, we mean the method of getting from the starting point to the ending point. Why is this independence important? Both $Q$ and $W$ *depend on path*, but $\Delta U$ does not (Equation 5.14.1). This path independence means that internal energy $U$ is easier to consider than either heat transfer or work done.

## Example 5.14.1: Calculating Change in Internal Energy - The Same Change in $U$ is Produced by Two Different Processes

a. Suppose there is heat transfer of 40.00 J to a system, while the system does 10.00 J of work. Later, there is heat transfer of 25.00 J out of the system while 4.00 J of work is done on the system. What is the net change in internal energy of the system?
b. What is the change in internal energy of a system when a total of 150.00 J of heat transfer occurs out of (from) the system and 159.00 J of work is done on the system (Figure 5.14.3)?

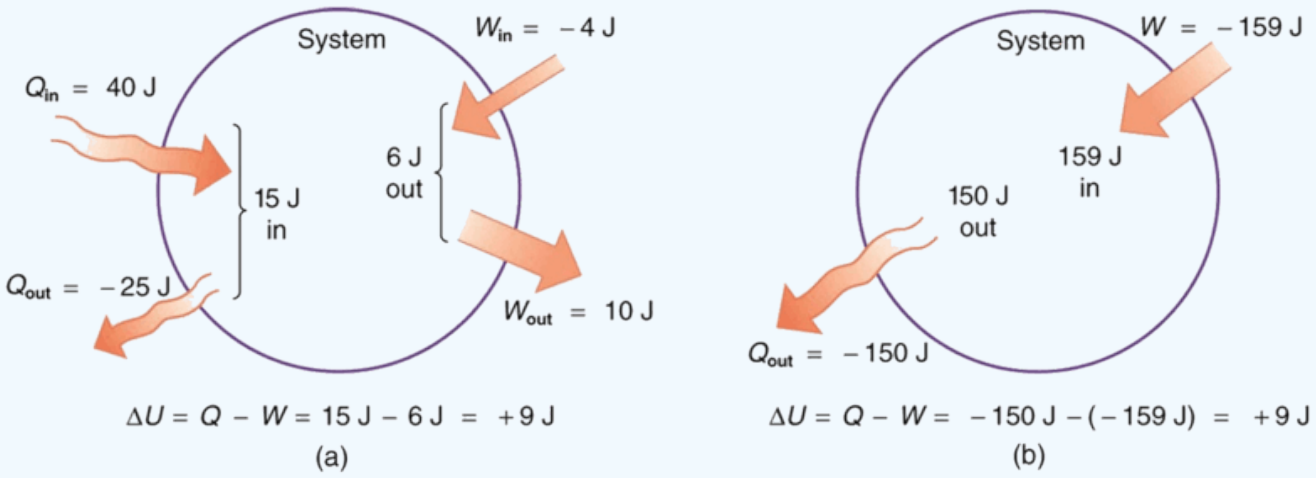



Figure 5.14.3: Two different processes produce the same change in a system. (a) A total of 15.00 J of heat transfer occurs into the system, while work takes out a total of 6.00 J. The change in internal energy is $\delta U = Q - W = 9.00\ J$. (b) Heat transfer removes 150.00 J from the system while work puts 159.00 J into it, producing an increase of 9.00 J in internal energy. If the system starts out in the same state in (a) and (b), it will end up in the same final state in either case—its final state is related to internal energy, not how that energy was acquired.

**Strategy**

In part (a), we must first find the net heat transfer and net work done from the given information. Then the first law of thermodynamics (Equation 5.14.1).

can be used to find the change in internal energy. In part (b), the net heat transfer and work done are given, so the equation can be used directly.

**Solution for (a)**

The net heat transfer is the heat transfer into the system minus the heat transfer out of the system, or

$$\begin{aligned} Q &= 40.00\,J - 25.00\,J \\ &= 15.00\,J \end{aligned}$$

Similarly, the total work is the work done by the system minus the work done on the system, or

$$\begin{aligned} W &= 10.00\,J - 4.00\,J \\ &= 6.00\,J. \end{aligned}$$

**Discussion on (a)**

No matter whether you look at the overall process or break it into steps, the change in internal energy is the same.

**Solution for (b)**

Here the net heat transfer and total work are given directly to be $Q = -150.00\ J$ and $W = -159.00\ J$, so that

$$\begin{aligned} \Delta U = Q - W &= -150.00 - (-159.00) \\ &= 9.00\,J. \end{aligned}$$

**Discussion on (b)**

A very different process in part (b) produces the same 9.00-J change in internal energy as in part (a). Note that the change in the system in both parts is related to $\Delta U$ and not to the individual $Q$s or $W$s involved. The system ends up in the *same* state in both (a) and (b). Parts (a) and (b) present two different paths for the system to follow between the same starting and ending points, and the change in internal energy for each is the same—it is independent of path.

## Human Metabolism and the First Law of Thermodynamics

Human metabolism is the conversion of food into heat transfer, work, and stored fat. Metabolism is an interesting example of the first law of thermodynamics in action. We now take another look at these topics via the first law of thermodynamics. Considering the body as the system of interest, we can use the first law to examine heat transfer, doing work, and internal energy in activities ranging from sleep to heavy exercise. What are some of the major characteristics of heat transfer, doing work, and energy in the body? For one, body temperature is normally kept constant by heat transfer to the surroundings. This means $Q$ is negative. Another fact is that the body usually does work on the outside world. This means $W$ is positive. In such situations, then, the body loses internal energy, since $\Delta U = Q - W$ is negative.

Now consider the effects of eating. Eating increases the internal energy of the body by adding chemical potential energy (this is an unromantic view of a good steak). The body *metabolizes* all the food we consume. Basically, metabolism is an oxidation process in which the chemical potential energy of food is released. This implies that food input is in the form of work. Food energy is reported in a special unit, known as the Calorie. This energy is measured by burning food in a calorimeter, which is how the units are determined.

In chemistry and biochemistry, one calorie (spelled with a *lowercase* c) is defined as the energy (or heat transfer) required to raise the temperature of one gram of pure water by one degree Celsius. Nutritionists and weight-watchers tend to use the *dietary* calorie, which is frequently called a Calorie (spelled with a *capital* C). One food Calorie is the energy needed to raise the temperature of one *kilogram* of water by one degree Celsius. This means that one dietary Calorie is equal to one kilocalorie for the chemist, and one must be careful to avoid confusion between the two.

Again, consider the internal energy the body has lost. There are three places this internal energy can go—to heat transfer, to doing work, and to stored fat (a tiny fraction also goes to cell repair and growth). Heat transfer and doing work take internal energy out of the body, and food puts it back. If you eat just the right amount of food, then your average internal energy remains constant.

Whatever you lose to heat transfer and doing work is replaced by food, so that, in the long run, $\Delta U = 0$. If you overeat repeatedly, then $\Delta U$ is always positive, and your body stores this extra internal energy as fat. The reverse is true if you eat too little. If $\Delta U$ is negative for a few days, then the body metabolizes its own fat to maintain body temperature and do work that takes energy from the body. This process is how dieting produces weight loss.

Life is not always this simple, as any dieter knows. The body stores fat or metabolizes it only if energy intake changes for a period of several days. Once you have been on a major diet, the next one is less successful because your body alters the way it responds to low energy intake. Your basal metabolic rate (BMR) is the rate at which food is converted into heat transfer and work done while the body is at complete rest. The body adjusts its basal metabolic rate to partially compensate for over-eating or under-eating. The body will decrease the metabolic rate rather than eliminate its own fat to replace lost food intake. You will chill more easily and feel less energetic as a result of the lower metabolic rate, and you will not lose weight as fast as before. Exercise helps to lose weight, because it produces both heat transfer from your body and work, and raises your metabolic rate even when you are at rest. Weight loss is also aided by the quite low efficiency of the body in converting internal energy to work, so that the loss of internal energy resulting from doing work is much greater than the work done.It should be noted, however, that living systems are not in thermal equilibrium.

The body provides us with an excellent indication that many thermodynamic processes are *irreversible*. An irreversible process can go in one direction but not the reverse, under a given set of conditions. For example, although body fat can be converted to do work and produce heat transfer, work done on the body and heat transfer into it cannot be converted to body fat. Otherwise, we could skip lunch by sunning ourselves or by walking down stairs. Another example of an irreversible thermodynamic process is photosynthesis. This process is the intake of one form of energy—light—by plants and its conversion to chemical potential energy. Both applications of the first law of thermodynamics are illustrated in Figure $5.14.4$. One great advantage of conservation laws such as the first law of thermodynamics is that they accurately describe the beginning and ending points of complex processes, such as metabolism and photosynthesis, without regard to the complications in between.

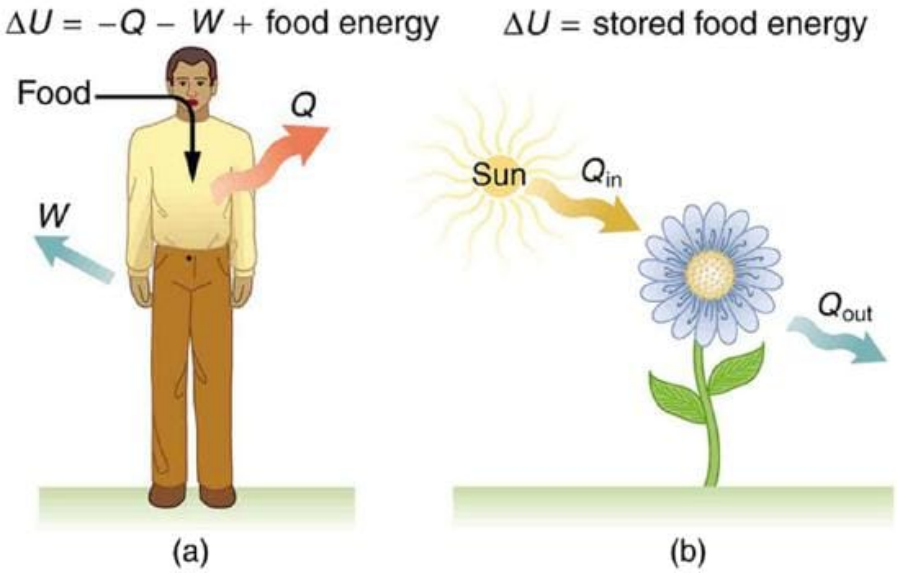



Figure $5.14.4$: (a) The first law of thermodynamics applied to metabolism. Heat transferred out of the body $Q$ and work done by the body $W$ remove internal energy, while food intake replaces it. (Food intake may be considered as work done on the body.) (b) Plants convert part of the radiant heat transfer in sunlight to stored chemical energy, a process called photosynthesis.

## Summary

The table presents a summary of terms relevant to the first law of thermodynamics.

| Term | Definition |
|---|---|
| $U$ | Internal energy—the sum of the kinetic and potential energies of a system's atoms and molecules. Can be divided into many subcategories, such as thermal and chemical energy. Depends only on the state of a system (such as its $P$, $V$ and $T$, not on how the energy entered the system. Change in internal energy is path independent. |

| Term | Definition |
|---|---|
| $Q$ | Heat—energy transferred because of a temperature difference. Characterized by random molecular motion. Highly dependent on path. $Q$ entering a system is positive. |
| $W$ | Work—energy transferred by a force moving through a distance. An organized, orderly process. Path dependent. $W$ done by a system (either against an external force or to increase the volume of the system) is positive. |

- The first law of thermodynamics is given as $\Delta U = Q - W$, where $\Delta U$ is the change in internal energy of a system, $Q$ is the net heat transfer (the sum of all heat transfer into and out of the system), and $W$ is the net work done (the sum of all work done on or by the system).
- Both $Q$ and $W$ are energy in transit; only $\Delta U$ represents an independent quantity capable of being stored.
- The internal energy $U$ of a system depends only on the state of the system and not how it reached that state.
- Metabolism of living organisms, and photosynthesis of plants, are specialized types of heat transfer, doing work, and internal energy of systems.

## Glossary

**first law of thermodynamics**
: states that the change in internal energy of a system equals the net heat transfer *into* the system minus the net work done *by* the system

**internal energy**
: the sum of the kinetic and potential energies of a system's atoms and molecules

**human metabolism**
: conversion of food into heat transfer, work, and stored fat

This page titled 5.14: The First Law of Thermodynamics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 15.1: The First Law of Thermodynamics by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.15: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency

## Learning Objectives

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

- State the expressions of the second law of thermodynamics.
- Calculate the efficiency and carbon dioxide emission of a coal-fired electricity plant, using second law characteristics.
- Describe and define the Otto cycle.



Figure 5.15.1: These ice floes melt during the Arctic summer. Some of them refreeze in the winter, but the second law of thermodynamics predicts that it would be extremely unlikely for the water molecules contained in these particular floes to reform the distinctive alligator-like shape they formed when the picture was taken in the summer of 2009. (credit: Patrick Kelley, U.S. Coast Guard, U.S. Geological Survey)

The second law of thermodynamics deals with the direction taken by spontaneous processes. Many processes occur spontaneously in one direction only—that is, they are irreversible, under a given set of conditions. Although irreversibility is seen in day-to-day life—a broken glass does not resume its original state, for instance—complete irreversibility is a statistical statement that cannot be seen during the lifetime of the universe. More precisely, an **irreversible process** is one that depends on path. If the process can go in only one direction, then the reverse path differs fundamentally and the process cannot be reversible. For example, as noted in the previous section, heat involves the transfer of energy from higher to lower temperature. A cold object in contact with a hot one never gets colder, transferring heat to the hot object and making it hotter. Furthermore, mechanical energy, such as kinetic energy, can be completely converted to thermal energy by friction, but the reverse is impossible. A hot stationary object never spontaneously cools off and starts moving. Yet another example is the expansion of a puff of gas introduced into one corner of a vacuum chamber. The gas expands to fill the chamber, but it never regroups in the corner. The random motion of the gas molecules could take them all back to the corner, but this is never observed to happen. (See Figure 5.15.2)

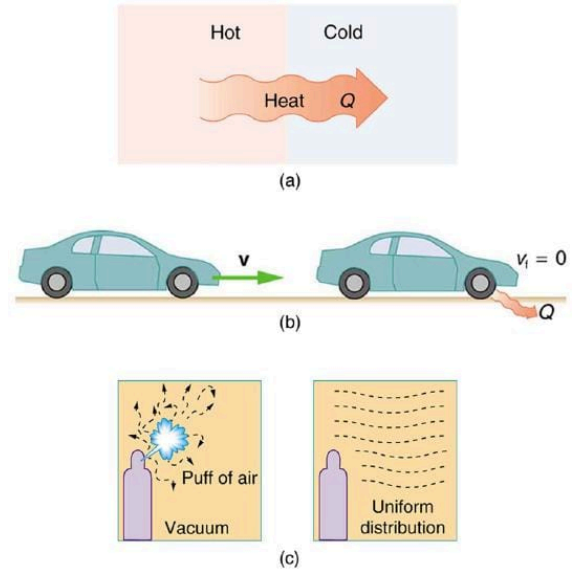



Figure 5.15.2: Examples of one-way processes in nature. (a) Heat transfer occurs spontaneously from hot to cold and not from cold to hot. (b) The brakes of this car convert its kinetic energy to heat transfer to the environment. The reverse process is impossible. (c) The burst of gas let into this vacuum chamber quickly expands to uniformly fill every part of the chamber. The random motions of the gas molecules will never return them to the corner.

The fact that certain processes never occur suggests that there is a law forbidding them to occur. The first law of thermodynamics would allow them to occur—none of those processes violate conservation of energy. The law that forbids these processes is called the second law of thermodynamics. We shall see that the second law can be stated in many ways that may seem different, but which in fact are equivalent. Like all natural laws, the second law of thermodynamics gives insights into nature, and its several statements imply that it is broadly applicable, fundamentally affecting many apparently disparate processes.

The already familiar direction of heat transfer from hot to cold is the basis of our first version of the **second law of thermodynamics**.

**The Second Law of Thermodynamics (first expression)**

Heat transfer occurs spontaneously from higher- to lower-temperature bodies but never spontaneously in the reverse direction.

Another way of stating this: It is impossible for any process to have as its sole result heat transfer from a cooler to a hotter object.

## Heat Engines

Now let us consider a device that uses heat transfer to do work. As noted in the previous section, such a device is called a heat engine, and one is shown schematically in 5.15.3(b). Gasoline and diesel engines, jet engines, and steam turbines are all heat engines that do work by using part of the heat transfer from some source. Heat transfer from the hot object (or hot reservoir) is denoted as $Q_f$, while heat transfer into the cold object (or cold reservoir) is $Q_c$, and the work done by the engine is $W$. The temperatures of the hot and cold reservoirs are $T_h$ and $T_c$, respectively.

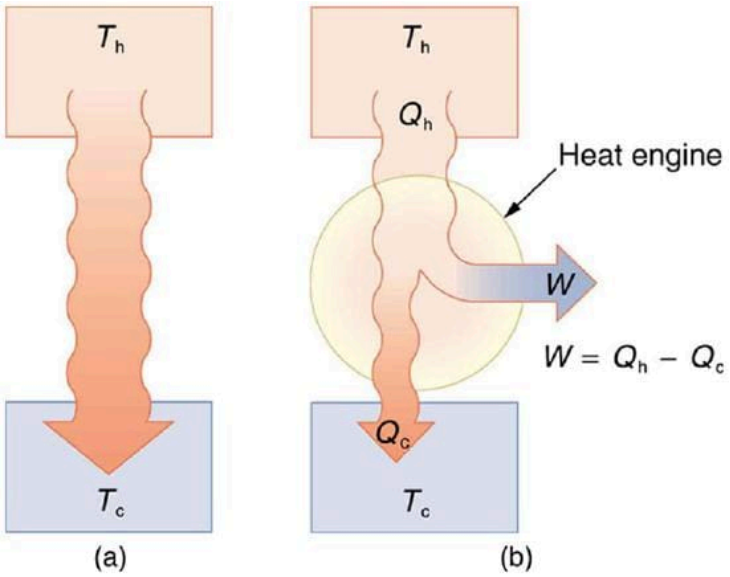



Figure $5.15.3$: (a) Heat transfer occurs spontaneously from a hot object to a cold one, consistent with the second law of thermodynamics. (b) A heat engine, represented here by a circle, uses part of the heat transfer to do work. The hot and cold objects are called the hot and cold reservoirs $Q_h$ is the heat transfer out of the hot reservoir, $W$ is the work output, and $Q_c$ is the heat transfer into the cold reservoir.

Because the hot reservoir is heated externally, which is energy intensive, it is important that the work is done as efficiently as possible. In fact, we would like $W$ to equal $Q_h$ and for there to be no heat transfer to the environment $(Q_c = 0)$. Unfortunately, this is impossible. The second law of thermodynamics also states, with regard to using heat transfer to do work (the second expression of the second law):

**The Second Law of Thermodynamics (second expression)**

It is impossible in any system for heat transfer from a reservoir to completely convert to work in a cyclical process in which the system returns to its initial state.

A cyclical process brings a system, such as the gas in a cylinder, back to its original state at the end of every cycle. Most heat engines, such as reciprocating piston engines and rotating turbines, use cyclical processes. The second law, just stated in its second form, clearly states that such engines cannot have perfect conversion of heat transfer into work done. Before going into the underlying reasons for the limits on converting heat transfer into work, we need to explore the relationships among $W$, $Q_h$, and $Q_c$ and to define the efficiency of a cyclical heat engine. As noted, a cyclical process brings the system back to its original condition at the end of every cycle. Such a system's internal energy $U$ is the same at the beginning and end of every cycle—that is, $\Delta U = 0$. The first law of thermodynamics states that

$$\Delta U = Q - W, \tag{5.15.1}$$

where $Q$ is the *net* heat transfer during the cycle $(Q = Q_h - Q_c)$ and $W$ is the net work done by the system. Since $\Delta U = 0$ for a complete cycle, we have

$$0 = Q - W, \tag{5.15.2}$$

so that

$$W = Q. \tag{5.15.3}$$

Thus the net work done by the system equals the net heat transfer into the system, or

$$W = Q_h - Q_c \, (cyclical \, process), \tag{5.15.4}$$

just as shown schematically in Figure $5.15.3$(b). The problem is that in all processes, there is some heat transfer $Q_c$ to the environment—and usually a very significant amount at that.

In the conversion of energy to work, we are always faced with the problem of getting less out than we put in. We define *conversion efficiency* $Eff$ to be the ratio of useful work output to the energy input (or, in other words, the ratio of what we get to what we spend). In that spirit, we define the efficiency of a heat engine to be its net work output $W$ divided by heat transfer to the engine $Q_h$, that is,

$$Eff = \frac{W}{Q_h}. \tag{5.15.5}$$

Since $W = Q_h - Q_c$ in a cyclical process, we can also express this as

$$Eff = \frac{Q_h - Q_c}{Q_h} = 1 - \frac{Q_c}{Q_h}\,(cyclical\,process), \tag{5.15.6}$$

making it clear that an efficiency of 1, or 100%, is possible only if there is no heat transfer to the environment $(Q_c = 0)$. Note that all $Q$s are positive. The direction of heat transfer is indicated by a plus or minus sign. For example, $Q - c$ is out of the system and so is preceded by a minus sign.

## Example 5.15.1: Daily Work Done by a Coal-Fired Power Station, Its Efficiency and Carbon Dioxide Emissions

A coal-fired power station is a huge heat engine. It uses heat transfer from burning coal to do work to turn turbines, which are used to generate electricity. In a single day, a large coal power station has $2.50 \times 10^{14}\ J$ of heat transfer from coal and $1.48 \times 10^{1}4\ J$ of heat transfer into the environment. (a) What is the work done by the power station? (b) What is the efficiency of the power station? (c) In the combustion process, the following chemical reaction occurs: $C + O_2 \rightarrow CO_2$ This implies that every 12 kg of coal puts 12 kg + 16 kg + 16 kg = 44 kg of carbon dioxide into the atmosphere. Assuming that 1 kg of coal can provide $2.5 \times 10^{6}\ J$ of heat transfer upon combustion, how much $CO_2$ is emitted per day by this power plant?

**Strategy for (a)**

We can use $W = Q_h - Q_c$ to find the work output $W$,

assuming a cyclical process is used in the power station. In this process, water is boiled under pressure to form high-temperature steam, which is used to run steam turbine-generators, and then condensed back to water to start the cycle again.

**Solution for (a)**

Work output is given by:

$$W = Q_h - Q_c.$$

Substituting the given values:

$$\begin{aligned} W &= 2.50 \times 10^{1}4\,J - 1.48 \times 10^{14}\,J \\ &= 1.02 \times 10^{14}\,J. \end{aligned}$$

**Strategy for (b)**

The efficiency can be calculated with $Eff = \frac{W}{Q_h}$ since $Q_H$ is given and work $W$ was found in the first part of this example.

**Solution for (b)**

Efficiency is given by $Eff = \frac{W}{Q_h}$. The work $W$ was just found to be $1.01 \times 10^{14}$, and $Q_h$ is given, so the efficiency is

$$Eff = \frac{1.01 \times 10^{14}\,J}{2.50 \times 10^{14}\,J} = 0.408,\ or\ 40.8\% \tag{5.15.7}$$

**Strategy for (c)**

The daily consumption of coal is calculated using the information that each day there is

$$\frac{2.50 \times 10^{14}\ J}{2.50 \times 10^{6}\ J/kg} = 1.0 \times 10^{8}\ kg.$$

Assuming that the coal is pure and that all the coal goes toward producing carbon dioxide, the carbon dioxide produced per day is

$$1.0 \times 10^{8}\ kg\ coal \times \frac{44\ kg\ CO_2}{12\ kg\ coal} = 3.7 \times 10^{8}\ kg\ CO_2.$$

This is 370,000 metric tons of $CO2$ produced every day.

**Discussion**

If all the work output is converted to electricity in a period of one day, the average power output is 1180 MW (this is left to you as an end-of-chapter problem). This value is about the size of a large-scale conventional power plant. The efficiency found is acceptably close to the value of 42% given for coal power stations. It means that fully 59.2% of the energy is heat transfer to the environment, which usually results in warming lakes, rivers, or the ocean near the power station, and is implicated in a warming planet generally. While the laws of thermodynamics limit the efficiency of such plants—including plants fired by nuclear fuel, oil, and natural gas—the heat transfer to the environment could be, and sometimes is, used for heating homes or for industrial processes. The generally low cost of energy has not made it economical to make better use of the waste heat transfer from most heat engines. Coal-fired power plants produce the greatest amount of $\mathrm{CO_2}$ per unit energy output (compared to natural gas or oil), making coal the least efficient fossil fuel.

With the information given in Example 5.15.1, we can find characteristics such as the efficiency of a heat engine without any knowledge of how the heat engine operates, but looking further into the mechanism of the engine will give us greater insight. Figure 5.15.4 illustrates the operation of the common four-stroke gasoline engine. The four steps shown complete this heat engine's cycle, bringing the gasoline-air mixture back to its original condition.

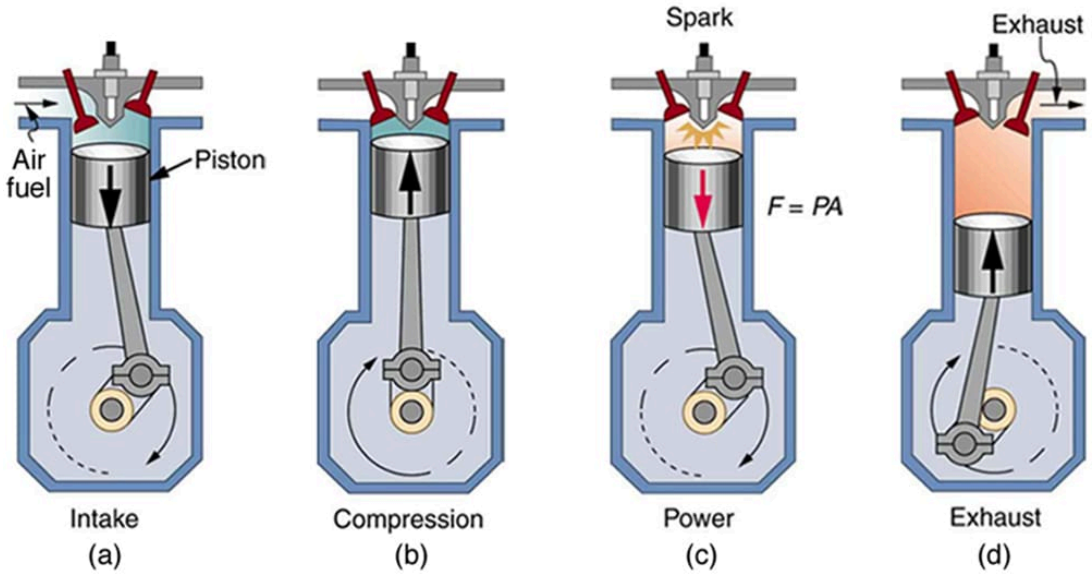



Figure 5.15.4: In the four-stroke internal combustion gasoline engine, heat transfer into work takes place in the cyclical process shown here. The piston is connected to a rotating crankshaft, which both takes work out of and does work on the gas in the cylinder. (a) Air is mixed with fuel during the intake stroke. (b) During the compression stroke, the air-fuel mixture is rapidly compressed in a nearly adiabatic process, as the piston rises with the valves closed. Work is done on the gas. (c) The power stroke has two distinct parts. First, the air-fuel mixture is ignited, converting chemical potential energy into thermal energy almost instantaneously, which leads to a great increase in pressure. Then the piston descends, and the gas does work by exerting a force through a distance in a nearly adiabatic process. (d) The exhaust stroke expels the hot gas to prepare the engine for another cycle, starting again with the intake

stroke.

The Otto cycle shown in Figure $5.15.5a$ is used in four-stroke internal combustion engines, although in fact the true Otto cycle paths do not correspond exactly to the strokes of the engine.

The adiabatic process AB corresponds to the nearly adiabatic compression stroke of the gasoline engine. In both cases, work is done on the system (the gas mixture in the cylinder), increasing its temperature and pressure. Along path BC of the Otto cycle, heat transfer $Q_h$ into the gas occurs at constant volume, causing a further increase in pressure and temperature. This process corresponds to burning fuel in an internal combustion engine, and takes place so rapidly that the volume is nearly constant. Path CD in the Otto cycle is an adiabatic expansion that does work on the outside world, just as the power stroke of an internal combustion engine does in its nearly adiabatic expansion. The work done by the system along path CD is greater than the work done on the system along path AB, because the pressure is greater, and so there is a net work output. Along path DA in the Otto cycle, heat transfer $Q_c$

from the gas at constant volume reduces its temperature and pressure, returning it to its original state. In an internal combustion engine, this process corresponds to the exhaust of hot gases and the intake of an air-gasoline mixture at a considerably lower temperature. In both cases, heat transfer into the environment occurs along this final path.

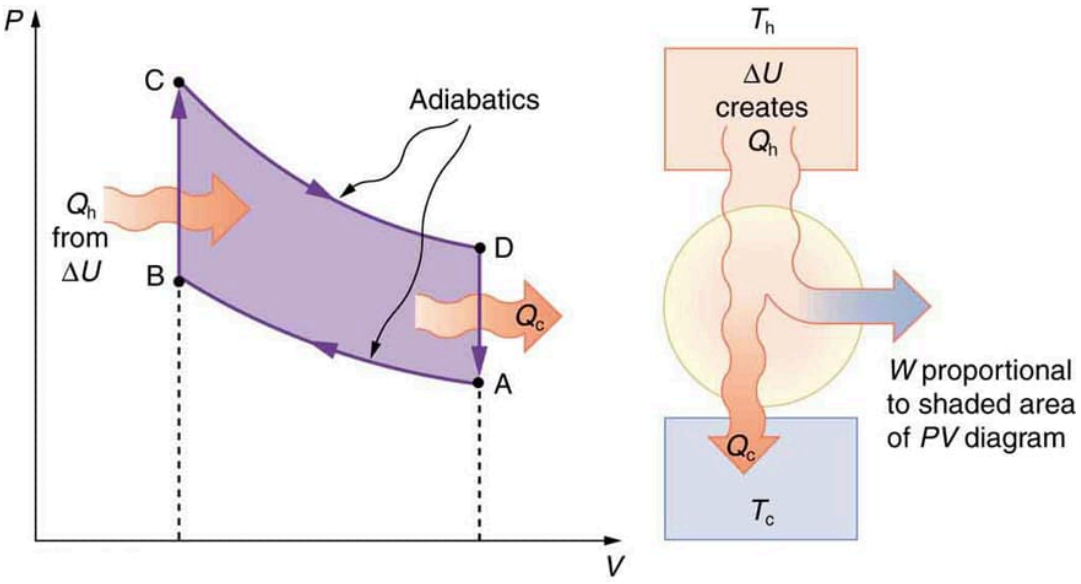



Figure $5.15.5$: $PV$ diagram for a simplified Otto cycle, analogous to that employed in an internal combustion engine. Point A corresponds to the start of the compression stroke of an internal combustion engine. Paths AB and CD are adiabatic and correspond to the compression and power strokes of an internal combustion engine, respectively. Paths BC and DA are isochoric and accomplish similar results to the ignition and exhaust-intake portions, respectively, of the internal combustion engine's cycle. Work is done on the gas along path AB, but more work is done by the gas along path CD, so that there is a net work output.

The net work done by a cyclical process is the area inside the closed path on a $PV$ diagram, such as that inside path ABCDA in Figure $5.15.5$. Note that in every imaginable cyclical process, it is absolutely necessary for heat transfer from the system to occur in order to get a net work output. In the Otto cycle, heat transfer occurs along path DA. If no heat transfer occurs, then the return path is the same, and the net work output is zero. The lower the temperature on the path AB, the less work has to be done to compress the gas. The area inside the closed path is then greater, and so the engine does more work and is thus more efficient. Similarly, the higher the temperature along path CD, the more work output there is (Figure $5.15.6$. So efficiency is related to the temperatures of the hot and cold reservoirs. In the next section, we shall see what the absolute limit to the efficiency of a heat engine is, and how it is related to temperature.

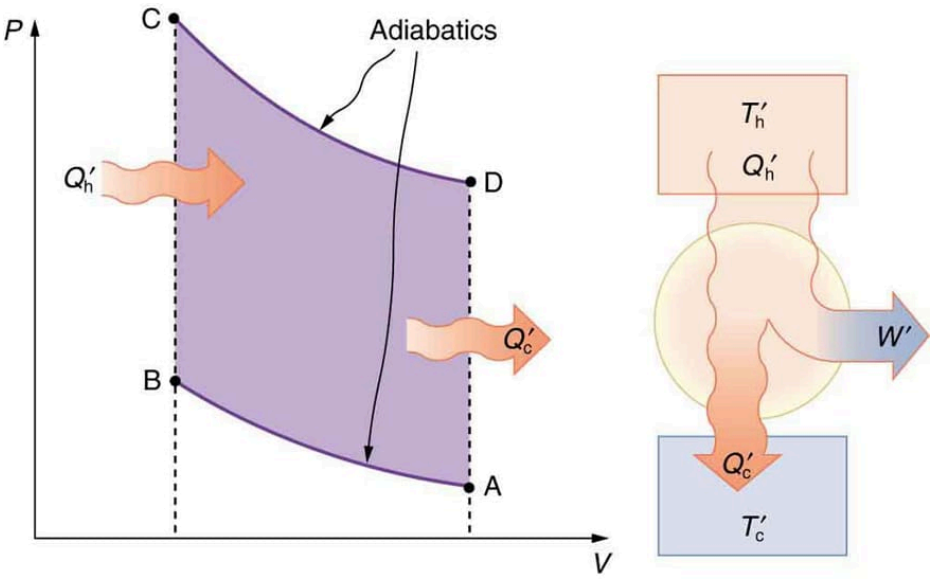



Figure 5.15.5: This Otto cycle produces a greater work output than the one in Figure 5, because the starting temperature of path CD is higher because the starting temperature of path CD is higher and the starting temperature of path AB is lower. The area inside the loop is greater, corresponding to greater net work output.

## Summary

- The two expressions of the second law of thermodynamics are: (i) Heat transfer occurs spontaneously from higher- to lower-temperature bodies but never spontaneously in the reverse direction; and (ii) It is impossible in any system for heat transfer from a reservoir to completely convert to work in a cyclical process in which the system returns to its initial state.
- Irreversible processes depend on path and do not return to their original state. Cyclical processes are processes that return to their original state at the end of every cycle.
- In a cyclical process, such as a heat engine, the net work done by the system equals the net heat transfer into the system, or $W = Q_h - Q_c$, where $Q_h$ is the heat transfer from the hot object (hot reservoir), and $Q_c$ is the heat transfer into the cold object (cold reservoir).
- Efficiency can be expressed as $Eff = \frac{W}{Q_h}$, the ratio of work output divided by the amount of energy input.
- The four-stroke gasoline engine is often explained in terms of the Otto cycle, which is a repeating sequence of processes that convert heat into work.

## Glossary

**irreversible process**
: any process that depends on path direction

**second law of thermodynamics**
: heat transfer flows from a hotter to a cooler object, never the reverse, and some heat energy in any process is lost to available work in a cyclical process

**cyclical process**
: a process in which the path returns to its original state at the end of every cycle

**Otto cycle**
: a thermodynamic cycle, consisting of a pair of adiabatic processes and a pair of isochoric processes, that converts heat into work, e.g., the four-stroke engine cycle of intake, compression, ignition, and exhaust

This page titled 5.15: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 15.3: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.16: Waves

## Learning Objectives

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

- State the characteristics of a wave.
- Calculate the velocity of wave propagation.

What do we mean when we say something is a wave? The most intuitive and easiest wave to imagine is the familiar water wave. More precisely, a **wave** is a disturbance that propagates, or moves from the place it was created. For water waves, the disturbance is in the surface of the water, perhaps created by a rock thrown into a pond or by a swimmer splashing the surface repeatedly. For sound waves, the disturbance is a change in air pressure, perhaps created by the oscillating cone inside a speaker. For earthquakes, there are several types of disturbances, including disturbance of Earth's surface and pressure disturbances under the surface. Even radio waves are most easily understood using an analogy with water waves. Visualizing water waves is useful because there is more to it than just a mental image. Water waves exhibit characteristics common to all waves, such as amplitude, period, frequency and energy. All wave characteristics can be described by a small set of underlying principles.



Figure $5.16.1$: Waves in the ocean behave similarly to all other types of waves. (credit: Steve Jurveston, Flickr)

A wave is a disturbance that propagates, or moves from the place it was created. The simplest waves repeat themselves for several cycles and are associated with simple harmonic motion. Let us start by considering the simplified water wave in Figure $5.16.2$. The wave is an up and down disturbance of the water surface. It causes a sea gull to move up and down in simple harmonic motion as the wave crests and troughs (peaks and valleys) pass under the bird. The time for one complete up and down motion is the wave's period $T$. The wave's frequency is $f = 1/T$, as usual. The wave itself moves to the right in the figure. This movement of the wave is actually the disturbance moving to the right, not the water itself (or the bird would move to the right). We define **wave velocity** $v_w$ to be the speed at which the disturbance moves. Wave velocity is sometimes also called the *propagation velocity or propagation speed,* because the disturbance propagates from one location to another.

## MISCONCEPTION ALERT

Many people think that water waves push water from one direction to another. In fact, the particles of water tend to stay in one location, save for moving up and down due to the energy in the wave. The energy moves forward through the water, but the water stays in one place. If you feel yourself pushed in an ocean, what you feel is the energy of the wave, not a rush of water.

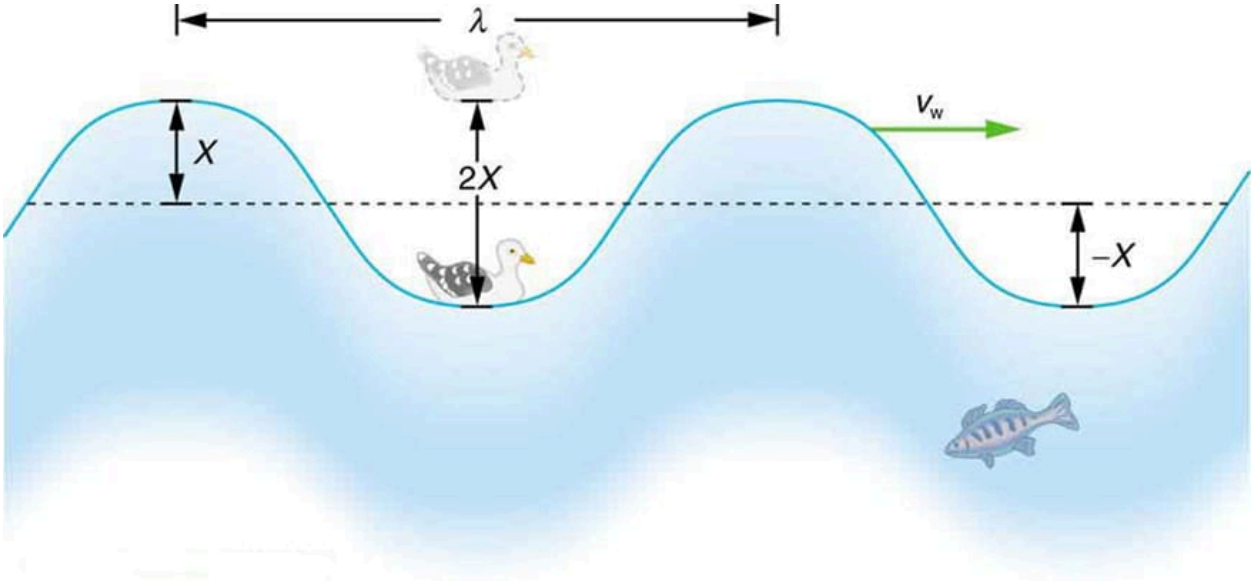



Figure 5.16.2 : An idealized ocean wave passes under a sea gull that bobs up and down in simple harmonic motion. The wave has a wavelength $\lambda$ which is the distance between adjacent identical parts of the wave. The up and down disturbance of the surface propagates parallel to the surface at a speed $v_w$.

The water wave in Figure 5.16.2 also has a length associated with it, called its wavelength $\lambda$ the distance between adjacent identical parts of a wave. ($\lambda$ is the distance parallel to the direction of propagation.) The speed of propagation $v_w$ is the distance the wave travels in a given time, which is one wavelength in the time of one period. In equation form, that is

$$v_w = \frac{\lambda}{T} \tag{5.16.1}$$

or

$$v_w = f\lambda. \tag{5.16.2}$$

This fundamental relationship holds for all types of waves. For water waves, $v_w$ is the speed of a surface wave; for sound, $v_w$ is the speed of sound; and for visible light, $v_w$ is the speed of light, for example.

## TAKE-HOME EXPERIMENT: WAVES IN A BOWL

Fill a large bowl or basin with water and wait for the water to settle so there are no ripples. Gently drop a cork into the middle of the bowl. Estimate the wavelength and period of oscillation of the water wave that propagates away from the cork. Remove the cork from the bowl and wait for the water to settle again. Gently drop the cork at a height that is different from the first drop. Does the wavelength depend upon how high above the water the cork is dropped?

## Example 5.16.1: Calculate the Velocity of Wave Propagation: Gull in the Ocean

Calculate the wave velocity of the ocean wave in Figure 5.16.2 if the distance between wave crests is 10.0 m and the time for a sea gull to bob up and down is 5.00 s.

**Strategy**

We are asked to find $v_w$. The given information tells us that $\lambda = 10.0\ m$ and $T = 5.00\ s$. Therefore, we can use Equation 5.16.1 to find the wave velocity.

**Solution**

1. Enter the known values into Equation 5.16.1:

$$v_w = \frac{\lambda}{T}$$
$$= \frac{10.0\,m}{5.00\,s}.$$

2. Solve for $v_w$ to find $v_w = 2.00\ m/s$.

**Discussion**

This slow speed seems reasonable for an ocean wave. Note that the wave moves to the right in the figure at this speed, not the varying speed at which the sea gull moves up and down.

## Transverse and Longitudinal Waves

A simple wave consists of a periodic disturbance that propagates from one place to another. The wave in Figure $5.16.3$ propagates in the horizontal direction while the surface is disturbed in the vertical direction. Such a wave is called a **transverse wave** or shear wave; in such a wave, the disturbance is perpendicular to the direction of propagation. In contrast, in a **longitudinal wave** or compressional wave, the disturbance is parallel to the direction of propagation. Figure $5.16.4$ shows an example of a longitudinal wave. The size of the disturbance is its amplitude *X* and is completely independent of the speed of propagation $v_w$.

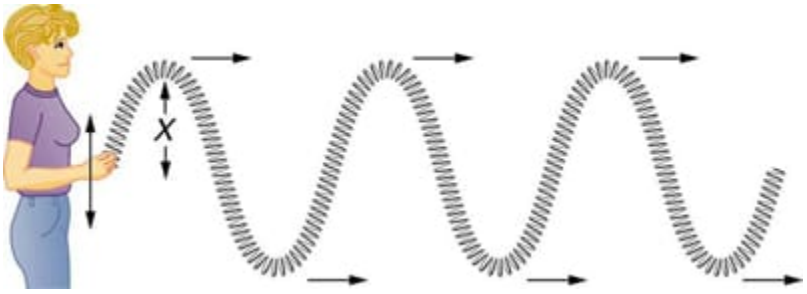



Figure $5.16.3$: In this example of a transverse wave, the wave propagates horizontally, and the disturbance in the cord is in the vertical direction.

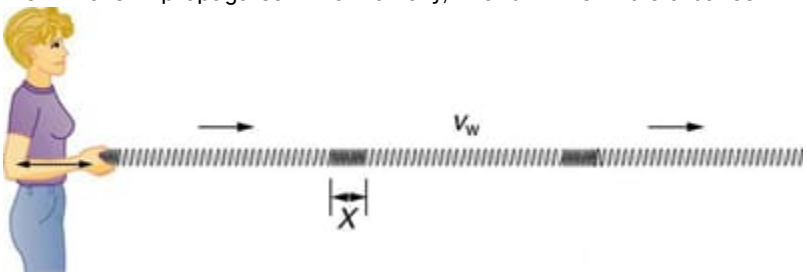



Figure $5.16.4$: In this example of a longitudinal wave, the wave propagates horizontally, and the disturbance in the cord is also in the horizontal direction.

Waves may be transverse, longitudinal, or *a combination of the two*. (Water waves are actually a combination of transverse and longitudinal. The simplified water wave illustrated in Figure $5.16.2$shows no longitudinal motion of the bird.) The waves on the strings of musical instruments are transverse—so are electromagnetic waves, such as visible light.

Sound waves in air and water are longitudinal. Their disturbances are periodic variations in pressure that are transmitted in fluids. Fluids do not have appreciable shear strength, and thus the sound waves in them must be longitudinal or compressional. Sound in solids can be both longitudinal and transverse.

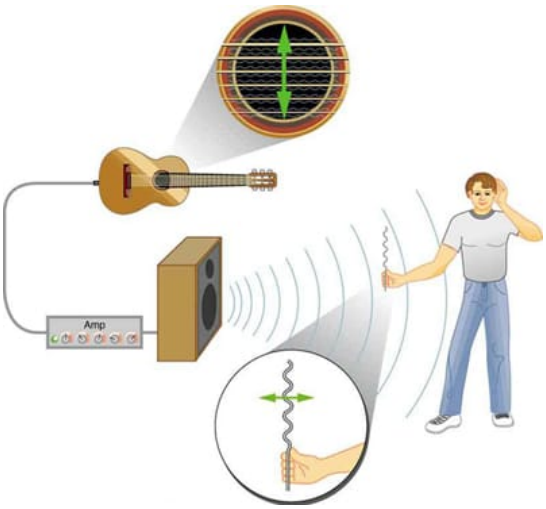



Figure 5.16.5: The wave on a guitar string is transverse. The sound wave rattles a sheet of paper in a direction that shows the sound wave is longitudinal.

Earthquake waves under Earth's surface also have both longitudinal and transverse components (called compressional or P-waves and shear or S-waves, respectively). These components have important individual characteristics—they propagate at different speeds, for example. Earthquakes also have surface waves that are similar to surface waves on water.

### Exercise 5.16.1:Check Your Understanding

Why is it important to differentiate between longitudinal and transverse waves?

*Answer available at the end of this section.*

### PHET EXPLORATIONS: WAVE ON A STRING

Watch a string vibrate in slow motion with this PhET simulation. Wiggle the end of the string and make waves, or adjust the frequency and amplitude of an oscillator. Adjust the damping and tension. The end can be fixed, loose, or open.




Figure 5.16.6: Wave on a string

## Summary

- A wave is a disturbance that moves from the point of creation with a wave velocity $v_w$.
- A wave has a wavelength $\lambda$ which is the distance between adjacent identical parts of the wave.
- Wave velocity and wavelength are related to the wave's frequency and period by $v_w = \frac{\lambda}{T}$ or $v_w = f\lambda$.
- A transverse wave has a disturbance perpendicular to its direction of propagation, whereas a longitudinal wave has a disturbance parallel to its direction of propagation.

# Glossary

**longitudinal wave**
: a wave in which the disturbance is parallel to the direction of propagation

**transverse wave**
: a wave in which the disturbance is perpendicular to the direction of propagation

**wave velocity**
: the speed at which the disturbance moves. Also called the propagation velocity or propagation speed

**wavelength**
: the distance between adjacent identical parts of a wave

## Answers

**Exercise** 5.16.1**:Check Your Understanding**

In the different types of waves, energy can propagate in a different direction relative to the motion of the wave. This is important to understand how different types of waves affect the materials around them.

This page titled 5.16: Waves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 16.9: Waves by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.17: Superposition and Interference

## Learning Objectives

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

- Explain standing waves.
- Describe the mathematical representation of overtones and beat frequency.

Most waves do not look very simple. They look more like the waves in Figure 5.17.1 than like the simple water wave considered in Waves. (Simple waves may be created by a simple harmonic oscillation, and thus have a sinusoidal shape). Complex waves are more interesting, even beautiful, but they look formidable. Most waves appear complex because they result from several simple waves adding together. Luckily, the rules for adding waves are quite simple.



Figure 5.17.1: These waves result from the superposition of several waves from different sources, producing a complex pattern. (credit: waterborough, Wikimedia Commons)

When two or more waves arrive at the same point, they superimpose themselves on one another. More specifically, the disturbances of waves are superimposed when they come together—a phenomenon called **superposition**. Each disturbance corresponds to a force, and forces add. If the disturbances are along the same line, then the resulting wave is a simple addition of the disturbances of the individual waves—that is, their amplitudes add. Figure 5.17.2 and Figure 5.17.3 illustrate superposition in two special cases, both of which produce simple results.

Figure 5.17.2 shows two identical waves that arrive at the same point exactly in phase. The crests of the two waves are precisely aligned, as are the troughs. This superposition produces pure **constructive interference**. Because the disturbances add, pure constructive interference produces a wave that has twice the amplitude of the individual waves, but has the same wavelength.

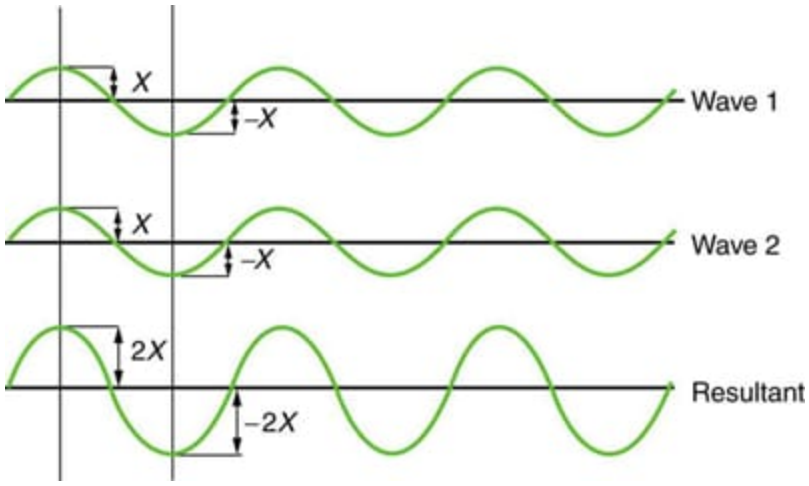



Figure 5.17.2: Pure constructive interference of two identical waves produces one with twice the amplitude, but the same wavelength.

Figure 5.17.3 shows two identical waves that arrive exactly out of phase—that is, precisely aligned crest to trough—producing pure **destructive interference**. Because the disturbances are in the opposite direction for this superposition, the resulting amplitude is zero for pure destructive interference—the waves completely cancel.

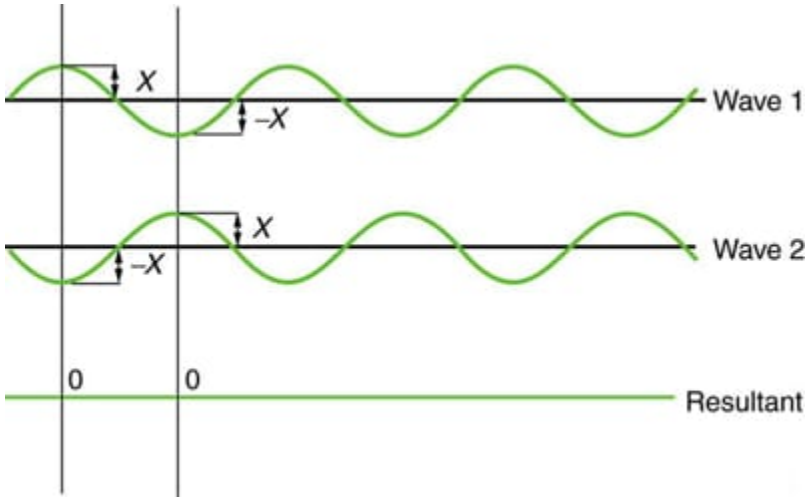



Figure 5.17.3: Pure destructive interference of two identical waves produces zero amplitude, or complete cancellation.

While pure constructive and pure destructive interference do occur, they require precisely aligned identical waves. The superposition of most waves produces a combination of constructive and destructive interference and can vary from place to place and time to time. Sound from a stereo, for example, can be loud in one spot and quiet in another. Varying loudness means the sound waves add partially constructively and partially destructively at different locations. A stereo has at least two speakers creating sound waves, and waves can reflect from walls. All these waves superimpose. An example of sounds that vary over time from constructive to destructive is found in the combined whine of airplane jets heard by a stationary passenger. The combined sound can fluctuate up and down in volume as the sound from the two engines varies in time from constructive to destructive. These examples are of waves that are similar.

An example of the superposition of two dissimilar waves is shown in Figure 5.17.4. Here again, the disturbances add and subtract, producing a more complicated looking wave.

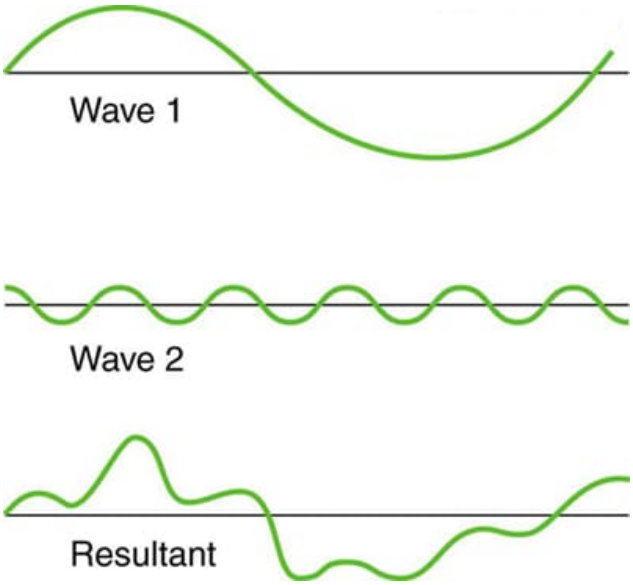



Figure 5.17.4: Superposition of non-identical waves exhibits both constructive and destructive interference.

## Standing Waves

Sometimes waves do not seem to move; rather, they just vibrate in place. Unmoving waves can be seen on the surface of a glass of milk in a refrigerator, for example. Vibrations from the refrigerator motor create waves on the milk that oscillate up and down but do not seem to move across the surface. These waves are formed by the superposition of two or more moving waves, such as illustrated in Figure 5.17.5 for two identical waves moving in opposite directions. The waves move through each other with their disturbances adding as they go by. If the two waves have the same amplitude and wavelength, then they alternate between constructive and destructive interference. The resultant looks like a wave standing in place and, thus, is called a **standing wave**. Waves on the glass of milk are one example of standing waves. There are other standing waves, such as on guitar strings and in organ pipes. With the glass of milk, the two waves that produce standing waves may come from reflections from the side of the glass.

A closer look at earthquakes provides evidence for conditions appropriate for resonance, standing waves, and constructive and destructive interference. A building may be vibrated for several seconds with a driving frequency matching that of the natural frequency of vibration of the building—producing a resonance resulting in one building collapsing while neighboring buildings do not. Often buildings of a certain height are devastated while other taller buildings remain intact. The building height matches the condition for setting up a standing wave for that particular height. As the earthquake waves travel along the surface of Earth and reflect off denser rocks, constructive interference occurs at certain points. Often areas closer to the epicenter are not damaged while areas farther away are damaged.

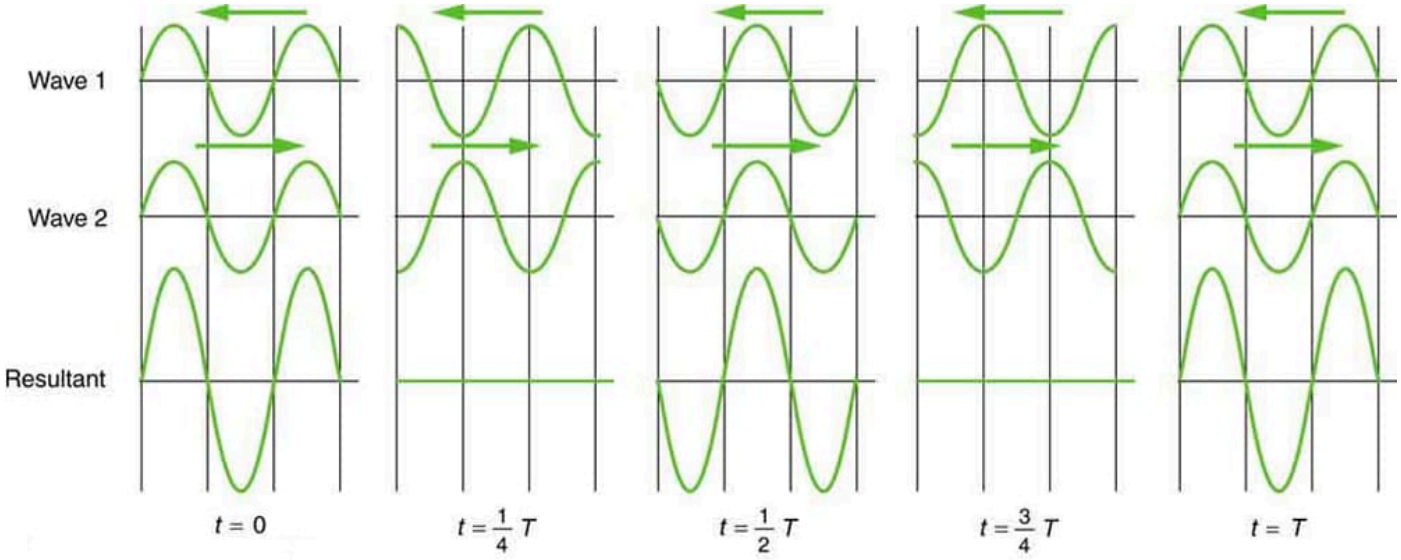



Figure 5.17.5: Standing wave created by the superposition of two identical waves moving in opposite directions. The oscillations are at fixed locations in space and result from alternately constructive and destructive interference.

Standing waves are also found on the strings of musical instruments and are due to reflections of waves from the ends of the string. Figure 5.17.6 and Figure 5.17.7 show three standing waves that can be created on a string that is fixed at both ends. **Nodes** are the points where the string does not move; more generally, nodes are where the wave disturbance is zero in a standing wave. The fixed ends of strings must be nodes, too, because the string cannot move there. The word **antinode** is used to denote the location of maximum amplitude in standing waves. Standing waves on strings have a frequency that is related to the propagation speed $v_w$ of the disturbance on the string. The wavelength $\lambda$ is determined by the distance between the points where the string is fixed in place.

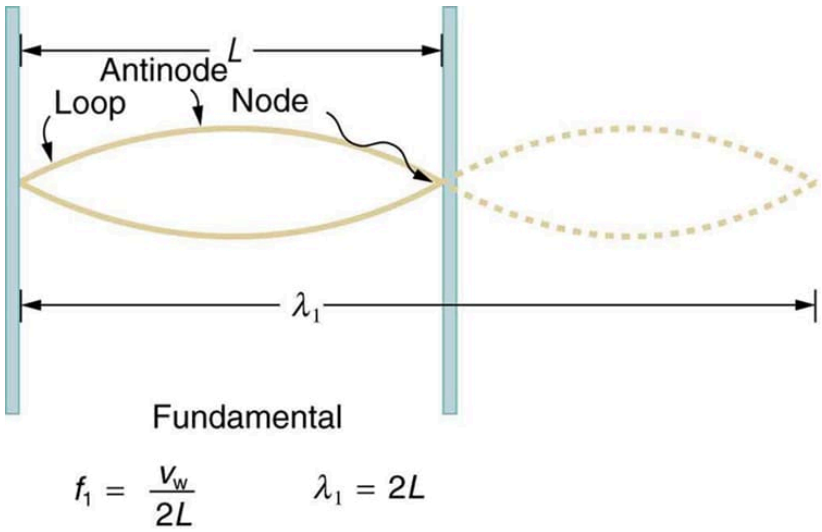



Figure 5.17.6: The figure shows a string oscillating at its fundamental frequency.

The lowest frequency, called the fundamental frequency, is thus for the longest wavelength, which is seen to be $\lambda_1 = 2L$. Therefore, the fundamental frequency is $f_1 = v_w/\lambda_1 = v_w/2L$. In this case, the overtones or harmonics are multiples of the fundamental frequency. As seen in Figure 5.17.7, the first harmonic can easily be calculated since $\lambda_2 = L$. Thus, $f_2 = v_w/\lambda_2 = v_w/2L = 2f_1$. Similarly, $f_3 = 3f_1$, and so on. All of these frequencies can be changed by adjusting the tension in the string. The greater the tension, the greater $v_w$ is and the higher the frequencies. This observation is familiar to anyone who has ever observed a string instrument being tuned. We will see in later chapters that standing waves are crucial to many resonance phenomena, such as in sounding boxes on string instruments.

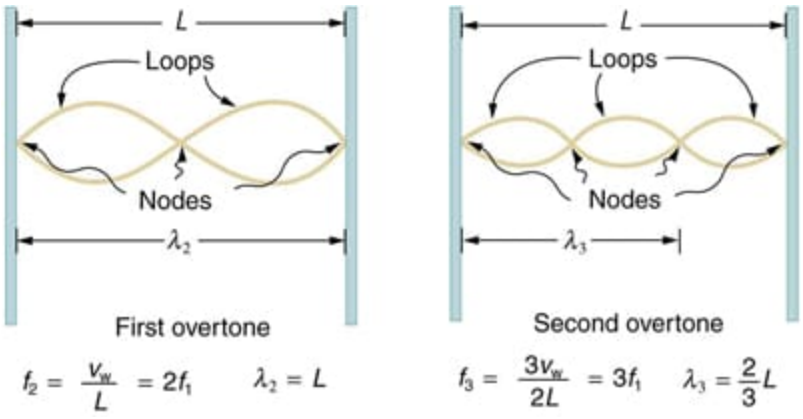



Figure 5.17.7: First and second harmonic frequencies are shown.

## Beats

Striking two adjacent keys on a piano produces a warbling combination usually considered to be unpleasant. The superposition of two waves of similar but not identical frequencies is the culprit. Another example is often noticeable in jet aircraft, particularly the two-engine variety, while taxiing. The combined sound of the engines goes up and down in loudness. This varying loudness happens because the sound waves have similar but not identical frequencies. The discordant warbling of the piano and the fluctuating loudness of the jet engine noise are both due to alternately constructive and destructive interference as the two waves go in and out of phase. Figure 5.17.8 illustrates this graphically.

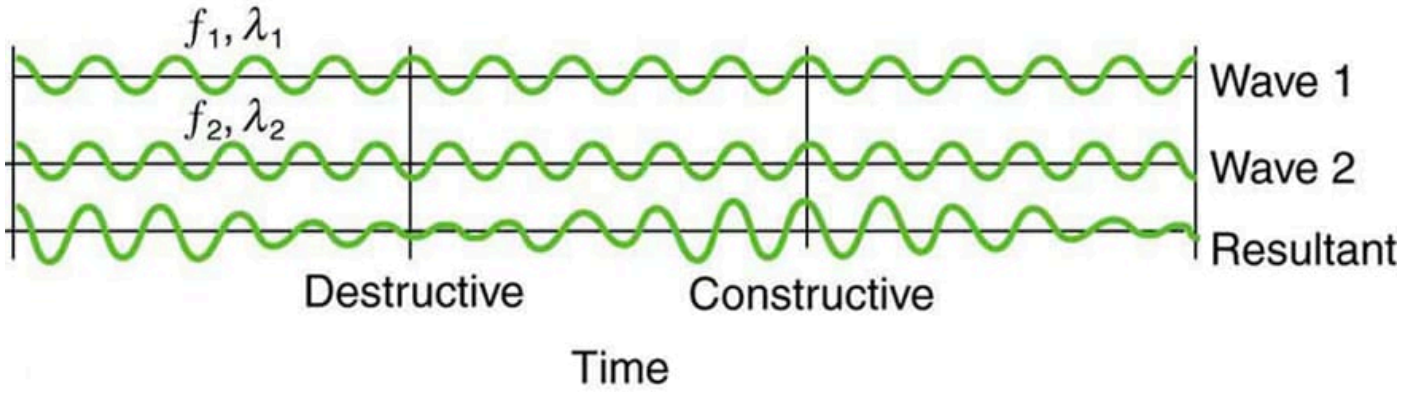



Figure 5.17.8: Beats are produced by the superposition of two waves of slightly different frequencies but identical amplitudes. The waves alternate in time between constructive interference and destructive interference, giving the resulting wave a time-varying amplitude.

The wave resulting from the superposition of two similar-frequency waves has a frequency that is the average of the two. This wave fluctuates in amplitude, or *beats*, with a frequency called the beat frequency. We can determine the beat frequency by adding two waves together mathematically. Note that a wave can be represented at one point in space as

$$x = X\,cos\left(\frac{2\pi t}{T}\right) = X\,cos\,(2\pi\,ft), \tag{5.17.1}$$

where $f = 1/T$ is the frequency of the wave. Adding two waves that have different frequencies but identical amplitudes produces a resultant

$$x = x_1 + x_2. \tag{5.17.2}$$

More specifically,

$$x = X\,cos(2\pi\,f_1 t) + X\,cos(2\pi\,f_2 t). \tag{5.17.3}$$

Using a trigonometric identity, it can be shown that

$$x = 2X\,cos(\pi\, f_B t)cos(2\pi\, f_{ave} t), \tag{5.17.4}$$

where

$$f_B = |f_1 - f_2| \tag{5.17.5}$$

is the beat frequency, and $f_{ave}$ is the average of $f_1$ and $f_2$. These results mean that the resultant wave has twice the amplitude and the average frequency of the two superimposed waves, but it also fluctuates in overall amplitude at the beat frequency $f_B$. The first cosine term in the expression effectively causes the amplitude to go up and down. The second cosine term is the wave with frequency $f_{ave}$. This result is valid for all types of waves. However, if it is a sound wave, providing the two frequencies are similar, then what we hear is an average frequency that gets louder and softer (or warbles) at the beat frequency.

## MAKING CAREER CONNECTIONS

Piano tuners use beats routinely in their work. When comparing a note with a tuning fork, they listen for beats and adjust the string until the beats go away (to zero frequency). For example, if the tuning fork has a $256hZ$ frequency and two beats per second are heard, then the other frequency is either $254$ or $258\ Hz$. Most keys hit multiple strings, and these strings are actually adjusted until they have nearly the same frequency and give a slow beat for richness. Twelve-string guitars and mandolins are also tuned using beats.

While beats may sometimes be annoying in audible sounds, we will find that beats have many applications. Observing beats is a very useful way to compare similar frequencies. There are applications of beats as apparently disparate as in ultrasonic imaging and radar speed traps.

### Exercise 5.17.1

Imagine you are holding one end of a jump rope, and your friend holds the other. If your friend holds her end still, you can move your end up and down, creating a transverse wave. If your friend then begins to move her end up and down, generating a wave in the opposite direction, what resultant wave forms would you expect to see in the jump rope?

*Answer available at the end of this section.*

### Exercise 5.17.2

Add text here. For the automatic number to work, you need to add the "AutoNum" template (preferably at the end) to the page.

*Answer available at the end of this section.*

### Exercise 5.17.3

You hook up a stereo system. When you test the system, you notice that in one corner of the room, the sounds seem dull. In another area, the sounds seem excessively loud. Describe how the sound moving about the room could result in these effects.

*Answer available at the end of this section.*

## PHET EXPLORATIONS: WAVE INTERFERENCE

Make waves with a dripping faucet, audio speaker, or laser! Add a second source or a pair of slits to create an interference pattern.




Figure 5.17.9: Wave Interference

## Summary

- Superposition is the combination of two waves at the same location.
- Constructive interference occurs when two identical waves are superimposed in phase.
- Destructive interference occurs when two identical waves are superimposed exactly out of phase.
- A standing wave is one in which two waves superimpose to produce a wave that varies in amplitude but does not propagate.
- Nodes are points of no motion in standing waves.
- An antinode is the location of maximum amplitude of a standing wave.
- Waves on a string are resonant standing waves with a fundamental frequency and can occur at higher multiples of the fundamental, called overtones or harmonics.
- Beats occur when waves of similar frequencies $f_1$ and $f_2$ are superimposed. The resulting amplitude oscillates with a beat frequency given by $f_B = |f_1 - f_2|$.

## Glossary

**antinode**
: the location of maximum amplitude in standing waves

**beat frequency**
: the frequency of the amplitude fluctuations of a wave

**constructive interference**
: when two waves arrive at the same point exactly in phase; that is, the crests of the two waves are precisely aligned, as are the troughs

**destructive interference**
: when two identical waves arrive at the same point exactly out of phase; that is, precisely aligned crest to trough

**fundamental frequency**
: the lowest frequency of a periodic waveform

**nodes**
: the points where the string does not move; more generally, nodes are where the wave disturbance is zero in a standing wave

**overtones**
: multiples of the fundamental frequency of a sound

**superposition**
: the phenomenon that occurs when two or more waves arrive at the same point

## Answers

**Exercise** 5.17.1

The rope would alternate between having waves with amplitudes two times the original amplitude and reaching equilibrium with no amplitude at all. The wavelengths will result in both constructive and destructive interference.

**Exercise** 5.17.2

Nodes are areas of wave interference where there is no motion. Antinodes are areas of wave interference where the motion is at its maximum point.

**Exercise** 5.17.3

With multiple speakers putting out sounds into the room, and these sounds bouncing off walls, there is bound to be some wave interference. In the dull areas, the interference is probably mostly destructive. In the louder areas, the interference is probably mostly constructive.

This page titled 5.17: Superposition and Interference is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 16.10: Superposition and Interference by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.18: Prelude to the Physics of Hearing

If a tree falls in the forest and no one is there to hear it, does it make a sound? The answer to this old philosophical question depends on how you define sound. If sound only exists when someone is around to perceive it, then there was no sound. However, if we define sound in terms of physics; that is, a disturbance of the atoms in matter transmitted from its origin outward (in other words, a wave), then there *was* a sound, even if nobody was around to hear it.



Figure 5.18.1: This tree fell some time ago. When it fell, atoms in the air were disturbed. Physicists would call this disturbance sound whether someone was around to hear it or not. (credit: B.A. Bowen Photography)

Such a wave is the physical phenomenon we call *sound*. Its perception is hearing. Both the physical phenomenon and its perception are interesting and will be considered in this text. We shall explore both sound and hearing; they are related, but are not the same thing. We will also explore the many practical uses of sound waves, such as in medical imaging.

This page titled 5.18: Prelude to the Physics of Hearing is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 17.0: Prelude to the Physics of Hearing by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.19: Sound

## Learning Objectives

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

- Define sound and hearing.
- Describe sound as a longitudinal wave.

Sound can be used as a familiar illustration of waves. Because hearing is one of our most important senses, it is interesting to see how the physical properties of sound correspond to our perceptions of it. **Hearing** is the perception of sound, just as vision is the perception of visible light. But sound has important applications beyond hearing. Ultrasound, for example, is not heard but can be employed to form medical images and is also used in treatment.



Figure 5.19.1: This glass has been shattered by a high-intensity sound wave of the same frequency as the resonant frequency of the glass. While the sound is not visible, the effects of the sound prove its existence. (credit: ||read||, Flickr)

The physical phenomenon of **sound** is defined to be a disturbance of matter that is transmitted from its source outward. Sound is a wave. On the atomic scale, it is a disturbance of atoms that is far more ordered than their thermal motions. In many instances, sound is a periodic wave, and the atoms undergo simple harmonic motion. In this text, we shall explore such periodic sound waves.

A vibrating string produces a sound wave as illustrated in Figures 5.19.2-5.19.4. As the string oscillates back and forth, it transfers energy to the air, mostly as thermal energy created by turbulence. But a small part of the string's energy goes into compressing and expanding the surrounding air, creating slightly higher and lower local pressures. These compressions (high pressure regions) and rarefactions (low pressure regions) move out as longitudinal pressure waves having the same frequency as the string—they are the disturbance that is a sound wave. (Sound waves in air and most fluids are longitudinal, because fluids have almost no shear strength. In solids, sound waves can be both transverse and longitudinal.) Figure 5.19.4 shows a graph of gauge pressure versus distance from the vibrating string.

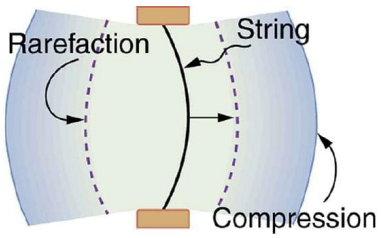



Figure 5.19.2: A vibrating string moving to the right compresses the air in front of it and

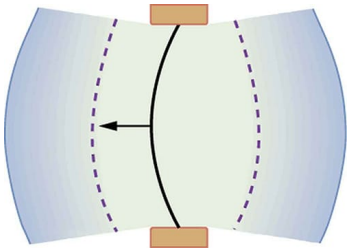


Figure 5.19.3: As the string moves to the left, it creates another compression and rarefaction as the ones on the right move away from the string.

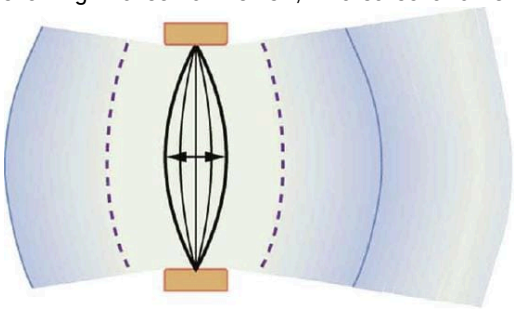



Figure 5.19.4: After many vibrations, there are a series of compressions and rarefactions moving out from the string as a sound wave. The graph shows gauge pressure versus distance from the source. Pressures vary only slightly from atmospheric for ordinary sounds.

The amplitude of a sound wave decreases with distance from its source, because the energy of the wave is spread over a larger and larger area. But it is also absorbed by objects, such as the eardrum in Figure 5.19.5, and converted to thermal energy by the viscosity of air. In addition, during each compression a little heat transfers to the air and during each rarefaction even less heat transfers from the air, so that the heat transfer reduces the organized disturbance into random thermal motions. (These processes can be viewed as a manifestation of the second law of thermodynamics presented in Introduction to the Second Law of Thermodynamics: Heat Engines and Their Efficiency.) Whether the heat transfer from compression to rarefaction is significant depends on how far apart they are—that is, it depends on wavelength. Wavelength, frequency, amplitude, and speed of propagation are important for sound, as they are for all waves.

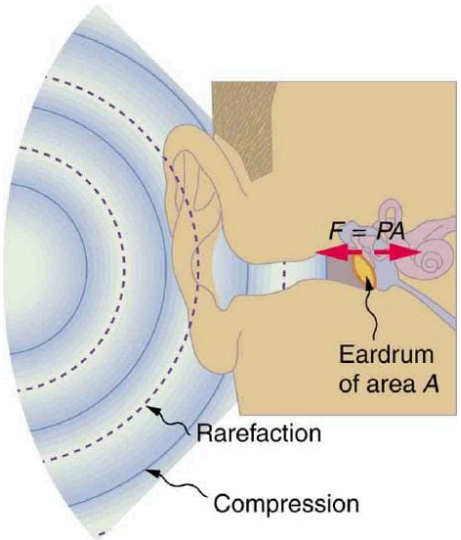



Figure 5.19.5: Sound wave compressions and rarefactions travel up the ear canal and force the eardrum to vibrate. There is a net force on the eardrum, since the sound wave pressures differ from the atmospheric pressure found behind the eardrum. A complicated mechanism converts the vibrations to nerve impulses, which are perceived by the person.

PHET EXPLORATIONS: WAVE INTERFERENCE

Make waves with a dripping faucet, audio speaker, or laser! Add a second source or a pair of slits to create an interference pattern.




Figure 5.19.6: Wave Interference

## Summary

- Sound is a disturbance of matter that is transmitted from its source outward.
- Sound is one type of wave.
- Hearing is the perception of sound.

## Glossary

**sound**
: a disturbance of matter that is transmitted from its source outward

**hearing**
: the perception of sound

This page titled 5.19: Sound is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 17.1: Sound by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.20: Speed of Sound, Frequency, and Wavelength

## Learning Objectives

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

- Define pitch.
- Describe the relationship between the speed of sound, its frequency, and its wavelength.
- Describe the effects on the speed of sound as it travels through various media.
- Describe the effects of temperature on the speed of sound.

Sound, like all waves, travels at a certain speed and has the properties of frequency and wavelength. You can observe direct evidence of the speed of sound while watching a fireworks display. The flash of an explosion is seen well before its sound is heard, implying both that sound travels at a finite speed and that it is much slower than light. You can also directly sense the frequency of a sound. Perception of frequency is called pitch. The wavelength of sound is not directly sensed, but indirect evidence is found in the correlation of the size of musical instruments with their pitch. Small instruments, such as a piccolo, typically make high-pitch sounds, while large instruments, such as a tuba, typically make low-pitch sounds. High pitch means small wavelength, and the size of a musical instrument is directly related to the wavelengths of sound it produces. So a small instrument creates short-wavelength sounds. Similar arguments hold that a large instrument creates long-wavelength sounds.



Figure $5.20.1$: When a firework explodes, the light energy is perceived before the sound energy. Sound travels more slowly than light does. (credit: Dominic Alves, Flickr)

The relationship of the speed of sound, its frequency, and wavelength is the same as for all waves:

$$v_w = f\lambda, \tag{5.20.1}$$

where $v_w$ is the speed of sound, $f$ is its frequency, and $\lambda$ is its wavelength. The wavelength of a sound is the distance between adjacent identical parts of a wave—for example, between adjacent compressions as illustrated in Figure $5.20.2$. The frequency is the same as that of the source and is the number of waves that pass a point per unit time.

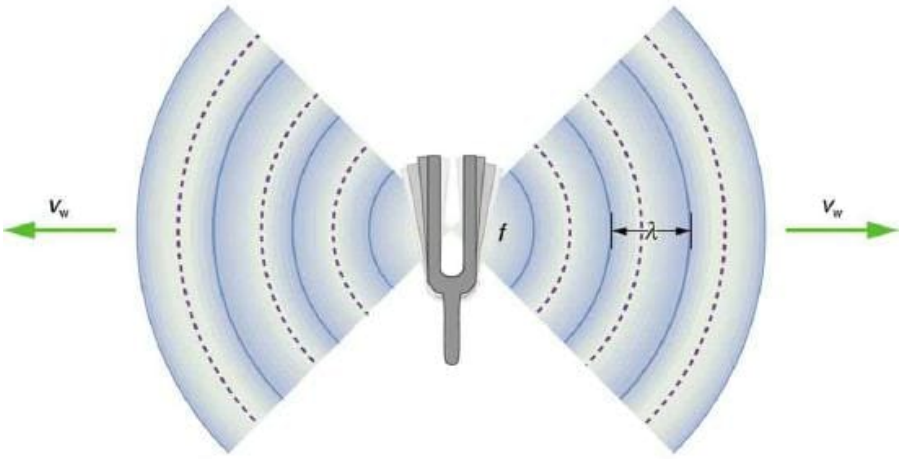



Figure 5.20.2: A sound wave emanates from a source vibrating at a frequency $f$, propagates at $v_w$, and has a wavelength $\lambda$.

Table 5.20.1 makes it apparent that the speed of sound varies greatly in different media. The speed of sound in a medium is determined by a combination of the medium's rigidity (or compressibility in gases) and its density. The more rigid (or less compressible) the medium, the faster the speed of sound. For materials that have similar rigidities, sound will travel faster through the one with the lower density because the sound energy is more easily transferred from particle to particle. The speed of sound in air is low, because air is compressible. Because liquids and solids are relatively rigid and very difficult to compress, the speed of sound in such media is generally greater than in gases.

*Table 5.20.1: Speed of Sound in Various Media*

| Medium | $v_w(m/s)$ |
|---|---|
| **Gases at $0^oC$** | |
| Air | 331 |
| Carbon dioxide | 259 |
| Oxygen | 316 |
| Helium | 965 |
| Hydrogen | 1290 |
| **Liquids at $20^oC$** | |
| Ethanol | 1160 |
| Mercury | 1450 |
| Water, fresh | 1480 |
| Sea water | 1540 |
| Human tissue | 1540 |
| **Solids (longitudinal or bulk)** | |
| Vulcanized rubber | 54 |
| Polyethylene | 920 |
| Marble | 3810 |
| Glass, Pyrex | 5640 |
| Lead | 1960 |
| Aluminum | 5120 |
| Steel | 5960 |

Earthquakes, essentially sound waves in Earth's crust, are an interesting example of how the speed of sound depends on the rigidity of the medium. Earthquakes have both longitudinal and transverse components, and these travel at different speeds. The bulk modulus of granite is greater than its shear modulus. For that reason, the speed of longitudinal or pressure waves (P-waves) in earthquakes in granite is significantly higher than the speed of transverse or shear waves (S-waves). Both components of earthquakes travel slower in less rigid material, such as sediments. P-waves have speeds of 4 to 7 km/s, and S-waves correspondingly range in speed from 2 to 5 km/s, both being faster in more rigid material. The P-wave gets progressively farther ahead of the S-wave as they travel through Earth's crust. The time between the P- and S-waves is routinely used to determine the distance to their source, the epicenter of the earthquake.

The speed of sound is affected by temperature in a given medium. For air at sea level, the speed of sound is given by

$$v_w = (331\, m/s)\sqrt{\frac{T}{273\, K}}, \tag{5.20.2}$$

where the temperature (denoted as $T$) is in units of kelvin. The speed of sound in gases is related to the average speed of particles in the gas, $v_{rms}$, and that

$$v_{rms} = \sqrt{\frac{3\, kT}{m}}, \tag{5.20.3}$$

where $k$ is the Boltzmann constant $(1.38 \times 10^{-23}\ J/K)$ and $m$ is the mass of each (identical) particle in the gas. So, it is reasonable that the speed of sound in air and other gases should depend on the square root of temperature. While not negligible, this is not a strong dependence. At $0^oC$, the speed of sound is 331 m/s, whereas at $20^oC$ it is 343 m/s, less than a 4% increase. Figure $5.20.3$ shows a use of the speed of sound by a bat to sense distances. Echoes are also used in medical imaging.

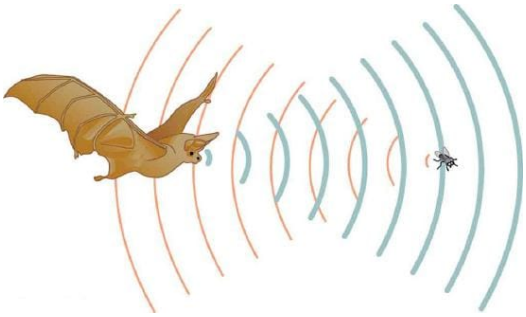


Figure $5.20.3$: A bat uses sound echoes to find its way about and to catch prey. The time for the echo to return is directly proportional to the distance.

One of the more important properties of sound is that its speed is nearly independent of frequency. This independence is certainly true in open air for sounds in the audible range of 20 to 20,000 Hz. If this independence were not true, you would certainly notice it for music played by a marching band in a football stadium, for example. Suppose that high-frequency sounds traveled faster—then the farther you were from the band, the more the sound from the low-pitch instruments would lag that from the high-pitch ones. But the music from all instruments arrives in cadence independent of distance, and so all frequencies must travel at nearly the same speed. Recall that

$$v_w = f\lambda. \tag{5.20.4}$$

In a given medium under fixed conditions, $v_w$ is constant, so that there is a relationship between $f$ and $\lambda$; the higher the frequency, the smaller the wavelength. See Figure $5.20.4$ and consider the following example.

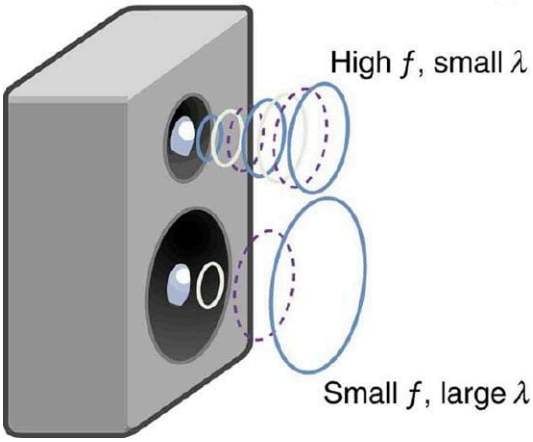



Figure 5.20.4: Because they travel at the same speed in a given medium, low-frequency sounds must have a greater wavelength than high-frequency sounds. Here, the lower-frequency sounds are emitted by the large speaker, called a woofer, while the higher-frequency sounds are emitted by the small speaker, called a tweeter.

## Example 5.20.1: Calculating Wavelengths: What Are the Wavelengths of Audible Sounds?

Calculate the wavelengths of sounds at the extremes of the audible range, 20 and 20,000 Hz, in $30.0^oC$ air. (Assume that the frequency values are accurate to two significant figures.)

**Strategy**

To find wavelength from frequency, we can use $v_w = f\lambda$.

**Solution**

1. Identify knowns. The value for $v_w$, is given by
$$v_w = (331\,m/s)\sqrt{\frac{T}{273\,K}}.$$
2. Convert the temperature into kelvin and then enter the temperature into the equation
$$v_w = (331\,m/s)\sqrt{\frac{303\,K}{273\,K}} = 348.7\,m/s.$$
3. Solve the relationship between speed and wavelength for $\lambda$:
$$\lambda = \frac{v_w}{f}.$$
4. Enter the speed and the minimum frequency to give the maximum wavelength:
$$\lambda_{max} = \frac{348.7\,m/s}{20\,Hz} = 17\,m.$$
5. Enter the speed and the maximum frequency to give the minimum wavelength:
$$\lambda_{min} = \frac{348.7\,m/s}{20,000\,Hz} = 0.017\,m = 1.7\,cm.$$

**Discussion**

Because the product of $f$ multiplied by $\lambda$ equals a constant, the smaller $f$ is, the larger $\lambda$ must be, and vice versa.

The speed of sound can change when sound travels from one medium to another. However, the frequency usually remains the same because it is like a driven oscillation and has the frequency of the original source. If $v_w$ changes and $f$ remains the same, then the wavelength $\lambda$ must change. That is, because $v_w = f\lambda$, the higher the speed of a sound, the greater its wavelength for a given frequency.

### MAKING CONNECTIONS: TAKE-HOME INVESTIGATION - VOICE AS A SOUND WAVE

Suspend a sheet of paper so that the top edge of the paper is fixed and the bottom edge is free to move. You could tape the top edge of the paper to the edge of a table. Gently blow near the edge of the bottom of the sheet and note how the sheet moves. Speak softly and then louder such that the sounds hit the edge of the bottom of the paper, and note how the sheet moves. Explain the effects.

### Exercise $5.20.1A$

Imagine you observe two fireworks explode. You hear the explosion of one as soon as you see it. However, you see the other firework for several milliseconds before you hear the explosion. Explain why this is so.

*Answer available at the end of this section.*

### Exercise $5.20.1B$

You observe two musical instruments that you cannot identify. One plays high-pitch sounds and the other plays low-pitch sounds. How could you determine which is which without hearing either of them play?

*Answer available at the end of this section.*

## Summary

- The relationship of the speed of sound $v_w$, its frequency $f$, and its wavelength $\lambda$ is given by $v_w = f\lambda$, which is the same relationship given for all waves.
- In air, the speed of sound is related to air temperature $T$ by $v_w = (331\ m/s)\sqrt{\frac{T}{273\,K}}$. $v_w$ is the same for all frequencies and wavelengths.

## Glossary

**pitch**
: the perception of the frequency of a sound

## Answers

**Exercise** $5.20.1A$

Sound and light both travel at definite speeds. The speed of sound is slower than the speed of light. The first firework is probably very close by, so the speed difference is not noticeable. The second firework is farther away, so the light arrives at your eyes noticeably sooner than the sound wave arrives at your ears.

**Exercise** $5.20.1B$

Compare their sizes. High-pitch instruments are generally smaller than low-pitch instruments because they generate a smaller wavelength.

This page titled 5.20: Speed of Sound, Frequency, and Wavelength is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 17.2: Speed of Sound, Frequency, and Wavelength by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.21: Doppler Effect and Sonic Booms

## Learning Objectives

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

- Define Doppler effect, Doppler shift, and sonic boom.
- Calculate the frequency of a sound heard by someone observing Doppler shift.
- Describe the sounds produced by objects moving faster than the speed of sound.

The characteristic sound of a motorcycle buzzing by is an example of the **Doppler effect**. The high-pitch scream shifts dramatically to a lower-pitch roar as the motorcycle passes by a stationary observer. The closer the motorcycle brushes by, the more abrupt the shift. The faster the motorcycle moves, the greater the shift. We also hear this characteristic shift in frequency for passing race cars, airplanes, and trains. It is so familiar that it is used to imply motion and children often mimic it in play.

The Doppler effect is an alteration in the observed frequency of a sound due to motion of either the source or the observer. Although less familiar, this effect is easily noticed for a stationary source and moving observer. For example, if you ride a train past a stationary warning bell, you will hear the bell's frequency shift from high to low as you pass by. The actual change in frequency due to relative motion of source and observer is called a **Doppler shift**. The Doppler effect and Doppler shift are named for the Austrian physicist and mathematician Christian Johann Doppler (1803–1853), who did experiments with both moving sources and moving observers. Doppler, for example, had musicians play on a moving open train car and also play standing next to the train tracks as a train passed by. Their music was observed both on and off the train, and changes in frequency were measured.

What causes the Doppler shift? Figure 5.21.1, Figure 5.21.2, and Figure 5.21.3compare sound waves emitted by stationary and moving sources in a stationary air mass. Each disturbance spreads out spherically from the point where the sound was emitted. If the source is stationary, then all of the spheres representing the air compressions in the sound wave centered on the same point, and the stationary observers on either side see the same wavelength and frequency as emitted by the source, as in Figure 5.21.1. If the source is moving, as in Figure 5.21.2, then the situation is different. Each compression of the air moves out in a sphere from the point where it was emitted, but the point of emission moves. This moving emission point causes the air compressions to be closer together on one side and farther apart on the other. Thus, the wavelength is shorter in the direction the source is moving (on the right in Figure 5.21.2), and longer in the opposite direction (on the left in Figure 5.21.2). Finally, if the observers move, as in Figure 5.21.3, the frequency at which they receive the compressions changes. The observer moving toward the source receives them at a higher frequency, and the person moving away from the source receives them at a lower frequency.




Figure 5.21.1: Sounds emitted by a source spread out in spherical waves. Because the source, observers, and air are stationary, the wavelength and frequency are the same in all directions and to all observers.

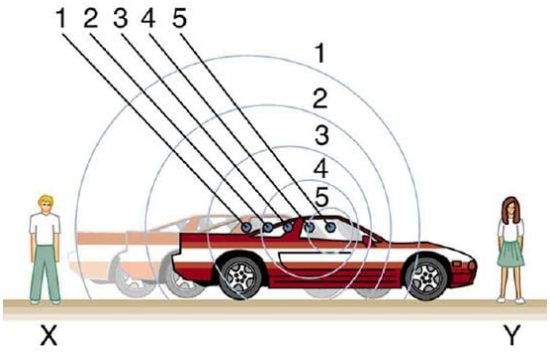



Figure 5.21.2: Sounds emitted by a source moving to the right spread out from the points at which they were emitted. The wavelength is reduced and, consequently, the frequency is increased in the direction of motion, so that the observer on the right hears a higher-pitch sound. The opposite is true for the observer on the left, where the wavelength is increased and the frequency is reduced.

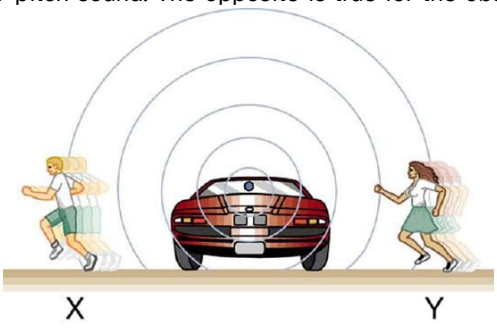



Figure 5.21.3: The same effect is produced when the observers move relative to the source. Motion toward the source increases frequency as the observer on the right passes through more wave crests than she would if stationary. Motion away from the source decreases frequency as the observer on the left passes through fewer wave crests than he would if stationary.

We know that wavelength and frequency are related by $v_w = f\lambda$, where $v_w$ is the fixed speed of sound. The sound moves in a medium and has the same speed $v_w$ in that medium whether the source is moving or not. Thus $f$ multiplied by $\lambda$ is a constant. Because the observer on the right in Figure 5.21.2 receives a shorter wavelength, the frequency she receives must be higher. Similarly, the observer on the left receives a longer wavelength, and hence he hears a lower frequency. The same thing happens in Figure 5.21.3. A higher frequency is received by the observer moving toward the source, and a lower frequency is received by an observer moving away from the source. In general, then, relative motion of source and observer toward one another increases the received frequency. Relative motion apart decreases frequency. The greater the relative speed is, the greater the effect.

## Doppler Effect

The Doppler effect occurs not only for sound but for any wave when there is relative motion between the observer and the source. There are Doppler shifts in the frequency of sound, light, and water waves, for example. Doppler shifts can be used to determine velocity, such as when ultrasound is reflected from blood in a medical diagnostic. The recession of galaxies is determined by the shift in the frequencies of light received from them and has implied much about the origins of the universe. Modern physics has been profoundly affected by observations of Doppler shifts.

For a stationary observer and a moving source, the frequency $f_{obs}$ received by the observer can be shown to be

$$f_{obs} = f_s \left( \frac{v_w}{v_w \pm v_s} \right), \tag{5.21.1}$$

where $f_s$ is the frequency of the source, v_s\) is the speed of the source along a line joining the source and observer, and $v_s$ is the speed of sound. The minus sign is used for motion toward the observer and the plus sign for motion away from the observer, producing the appropriate shifts up and down in frequency. Note that the greater the speed of the source, the greater the effect. Similarly, for a stationary source and moving observer, the frequency received by the observer $f_{obs}$ is given by

$$f_{obs} = f_s \left( \frac{v_w \pm v_{obs}}{v_w} \right), \tag{5.21.2}$$

where $v_{obs}$ is the speed of the observer along a line joining the source and observer. Here the plus sign is for motion toward the

source, and the minus is for motion away from the source.

## Example 5.21.1: Calculate Doppler Shift: A Train Horn

Suppose a train that has a 150-Hz horn is moving at 35.0 m/s in still air on a day when the speed of sound is 340 m/s.

(a) What frequencies are observed by a stationary person at the side of the tracks as the train approaches and after it passes?

(b) What frequency is observed by the train's engineer traveling on the train?

**Strategy**

To find the observed frequency in (a), $f_{obs} = f_s \left(\frac{v_s}{v_w \pm v_s}\right)$, must be used because the source is moving. The minus sign is used for the approaching train, and the plus sign for the receding train. In (b), there are two Doppler shifts—one for a moving source and the other for a moving observer.

**Solution for (a)**

(1) Enter known values into $f_{obs} = f_s \left(\frac{v_w}{v_w - v_s}\right)$.

$$f_{obs} = f_s \left(\frac{v_w}{v_w - v_s}\right) = (150\,Hz)\left(\frac{340\,m/s}{340\,m/s - 35.0\,m/s}\right) \tag{5.21.3}$$

(2) Calculate the frequency observed by a stationary person as the train approaches.

$$f_{obs} = (150\,Hz)(1.11) = 167\,Hz \tag{5.21.4}$$

(3) Use the same equation with the plus sign to find the frequency heard by a stationary person as the train recedes.

$$f_{obs} = f_s \left(\frac{v_w}{v_w - v_s}\right) = (150\,Hz)\left(\frac{340\,m/s}{340n\,m/s + 35.0\,m/s}\right) \tag{5.21.5}$$

(4) Calculate the second frequency.

$$f_{obs} = (150\,Hz)(0.907) = 136\,Hz \tag{5.21.6}$$

**Discussion on (a)**

The numbers calculated are valid when the train is far enough away that the motion is nearly along the line joining train and observer. In both cases, the shift is significant and easily noticed. Note that the shift is 17.0 Hz for motion toward and 14.0 Hz for motion away. The shifts are not symmetric.

**Solution for (b)**

(1) Identify knowns:

- It seems reasonable that the engineer would receive the same frequency as emitted by the horn, because the relative velocity between them is zero.
- Relative to the medium (air), the speeds are $v_s = v_{obs} = 35.0\ m/s$.
- The first Doppler shift is for the moving observer; the second is for the moving source.

(2) Use the following equation

$$f_{obs} = \left[f_s \left(\frac{v_w \pm v_{obs}}{v_w}\right)\right]\left(\frac{v_w}{v_w \pm v_s}\right). \tag{5.21.7}$$

The quantity in the square brackets is the Doppler-shifted frequency due to a moving observer. The factor on the right is the effect of the moving source.

(3) Because the train engineer is moving in the direction toward the horn, we must use the plus sign for $v_{obs}$; however, because the horn is also moving in the direction away from the engineer, we also use the plus sign for $v_s$. But the train is carrying both the engineer and the horn at the same velocity, so $v_s = v_{obs}$. As a result, everything but $f_s$ cancels, yielding

$$f_{obs} = f_s. \tag{5.21.8}$$

**Discussion for (b)**

We may expect that there is no change in frequency when source and observer move together because it fits your experience. For example, there is no Doppler shift in the frequency of conversations between driver and passenger on a motorcycle. People talking when a wind moves the air between them also observe no Doppler shift in their conversation. The crucial point is that source and observer are not moving relative to each other.

## Sonic Booms to Bow Wakes

What happens to the sound produced by a moving source, such as a jet airplane, that approaches or even exceeds the speed of sound? The answer to this question applies not only to sound but to all other waves as well.

Suppose a jet airplane is coming nearly straight at you, emitting a sound of frequency $f_s$. The greater the plane's speed $v_s$, the greater the Doppler shift and the greater the value observed for $f_{obs}$. Now, as $v_s$ approaches the speed of sound, $f_{obs}$ approaches infinity, because the denominator in $f_{obs} = f_s\left(\frac{v_w}{v_w \pm v_s}\right)$ approaches zero. At the speed of sound, this result means that in front of the source, each successive wave is superimposed on the previous one because the source moves forward at the speed of sound. The observer gets them all at the same instant, and so the frequency is infinite. (Before airplanes exceeded the speed of sound, some people argued it would be impossible because such constructive superposition would produce pressures great enough to destroy the airplane.) If the source exceeds the speed of sound, no sound is received by the observer until the source has passed, so that the sounds from the approaching source are mixed with those from it when receding. This mixing appears messy, but something interesting happens—a sonic boom is created. (See Figure 5.21.4.)

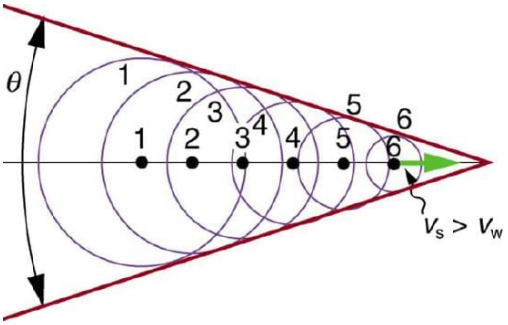



Figure 5.21.4: Sound waves from a source that moves faster than the speed of sound spread spherically from the point where they are emitted, but the source moves ahead of each. Constructive interference along the lines shown (actually a cone in three dimensions) creates a shock wave called a sonic boom. The faster the speed of the source, the smaller the angle $\theta$.

There is constructive interference along the lines shown (a cone in three dimensions) from similar sound waves arriving there simultaneously. This superposition forms a disturbance called a **sonic boom**, a constructive interference of sound created by an object moving faster than sound. Inside the cone, the interference is mostly destructive, and so the sound intensity there is much less than on the shock wave. An aircraft creates two sonic booms, one from its nose and one from its tail. (See Figure 5.21.5.) During television coverage of space shuttle landings, two distinct booms could often be heard. These were separated by exactly the time it would take the shuttle to pass by a point. Observers on the ground often do not see the aircraft creating the sonic boom, because it has passed by before the shock wave reaches them, as seen in Figure 5.21.5. If the aircraft flies close by at low altitude, pressures in the sonic boom can be destructive and break windows as well as rattle nerves. Because of how destructive sonic booms can be, supersonic flights are banned over populated areas of the United States.

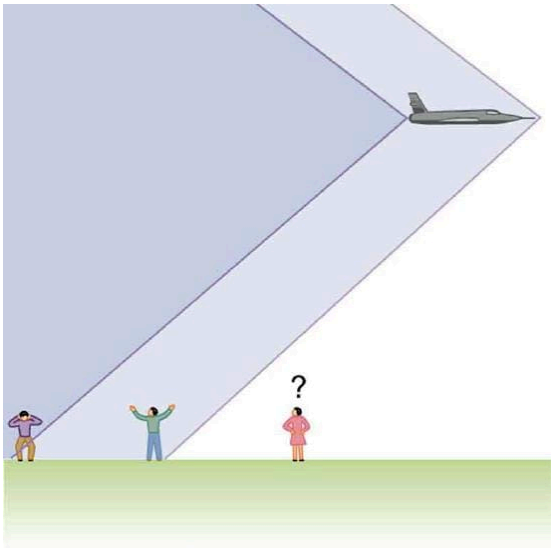


Figure 5.21.5: Two sonic booms, created by the nose and tail of an aircraft, are observed on the ground after the plane has passed by.

Sonic booms are one example of a broader phenomenon called bow wakes. A bow wake, such as the one in Figure 5.21.6, is created when the wave source moves faster than the wave propagation speed. Water waves spread out in circles from the point where created, and the bow wake is the familiar V-shaped wake trailing the source. A more exotic bow wake is created when a subatomic particle travels through a medium faster than the speed of light travels in that medium. (In a vacuum, the maximum speed of light will be $3.99 \times 10^8\ m/s$; in the medium of water, the speed of light is closer to $0.75c$. If the particle creates light in its passage, that light spreads on a cone with an angle indicative of the speed of the particle, as illustrated in Figure 5.21.7. Such a bow wake is called Cerenkov radiation and is commonly observed in particle physics.


Figure 5.21.6: Bow wake created by a duck. Constructive interference produces the rather structured wake, while there is relatively little wave action inside the wake, where interference is mostly destructive. (credit: Horia Varlan, Flickr)


Figure 5.21.7: The blue glow in this research reactor pool is Cerenkov radiation caused by subatomic particles traveling faster than the speed of light in water. (credit: U.S. Nuclear Regulatory Commission)

Doppler shifts and sonic booms are interesting sound phenomena that occur in all types of waves. They can be of considerable use. For example, the Doppler shift in ultrasound can be used to measure blood velocity, while police use the Doppler shift in radar (a microwave) to measure car velocities. In meteorology, the Doppler shift is used to track the motion of storm clouds; such “Doppler Radar” can give velocity and direction and rain or snow potential of imposing weather fronts. In astronomy, we can examine the light emitted from distant galaxies and determine their speed relative to ours. As galaxies move away from us, their light is shifted to a lower frequency, and so to a longer wavelength—the so-called red shift. Such information from galaxies far, far

away has allowed us to estimate the age of the universe (from the Big Bang) as about 14 billion years.

**Exercise 5.21.1**

Why did scientist Christian Doppler observe musicians both on a moving train and also from a stationary point not on the train?

*Answer available at the end of this section.*

**Exercise 5.21.2**

Describe a situation in your life when you might rely on the Doppler shift to help you either while driving a car or walking near traffic.

*Answer available at the end of this section.*

## Summary

- The Doppler effect is an alteration in the observed frequency of a sound due to motion of either the source or the observer.
- The actual change in frequency is called the Doppler shift.
- A sonic boom is constructive interference of sound created by an object moving faster than sound.
- A sonic boom is a type of bow wake created when any wave source moves faster than the wave propagation speed.
- For a stationary observer and a moving source, the observed frequency $f_{obs}$ is: $f_{obs} = f_s \left( \frac{v_w}{v_w \pm v_s} \right)$, where $f_{obs}$ is the frequency of the source, $v_s$ is the speed of the source, and $v_w$ is the speed of sound. The minus sign is used for motion toward the observer and the plus sign for motion away.
- For a stationary source and moving observer, the observed frequency is: $f_{obs} = f_s \left( \frac{v_w \pm v_{obs}}{v_w} \right)$, where $v_{obs}$ is the speed of the observer.

## Glossary

**Doppler effect**
: an alteration in the observed frequency of a sound due to motion of either the source or the observer

**Doppler shift**
: the actual change in frequency due to relative motion of source and observer

**sonic boom**
: a constructive interference of sound created by an object moving faster than sound

**bow wake**
: V-shaped disturbance created when the wave source moves faster than the wave propagation speed

## Contributors and Attributions

{{template.ContribOpenStaxCollege()}}

### Answers

**Exercise** 5.21.1

Doppler needed to compare the perception of sound when the observer is stationary and the sound source moves, as well as when the sound source and the observer are both in motion.

**Exercise** 5.21.2

If I am driving and I hear Doppler shift in an ambulance siren, I would be able to tell when it was getting closer and also if it has passed by. This would help me to

know whether I needed to pull over and let the ambulance through.

This page titled 5.21: Doppler Effect and Sonic Booms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 17.4: Doppler Effect and Sonic Booms by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.22: Prelude to Electromagnetic Waves

The beauty of a coral reef, the warm radiance of sunshine, the sting of sunburn, the X-ray revealing a broken bone, even microwave popcorn—all are brought to us by **electromagnetic waves**. The list of the various types of electromagnetic waves, ranging from radio transmission waves to nuclear gamma-ray ($\gamma$-ray) emissions, is interesting in itself.

Even more intriguing is that all of these widely varied phenomena are different manifestations of the same thing—electromagnetic waves. (See Figure 2.) What are electromagnetic waves? How are they created, and how do they travel? How can we understand and organize their widely varying properties? What is their relationship to electric and magnetic effects? These and other questions will be explored.



Figure 5.22.1: Human eyes detect these orange "sea goldie" fish swimming over a coral reef in the blue waters of the Gulf of Eilat (Red Sea) using visible light. (credit: Daviddarom, Wikimedia Commons)

MISCONCEPTION ALERT: SOUND WAVES VS. RADIO WAVES:

Many people confuse sound waves with **radio waves**, one type of electromagnetic (EM) wave. However, sound and radio waves are completely different phenomena. Sound creates pressure variations (waves) in matter, such as air or water, or your eardrum. Conversely, radio waves are *electromagnetic waves*, like visible light, infrared, ultraviolet, X-rays, and gamma rays. EM waves don't need a medium in which to propagate; they can travel through a vacuum, such as outer space.

A radio works because sound waves played by the D.J. at the radio station are converted into electromagnetic waves, then encoded and transmitted in the radio-frequency range. The radio in your car receives the radio waves, decodes the information, and uses a speaker to change it back into a sound wave, bringing sweet music to your ears.

## Discovering a New Phenomenon

It is worth noting at the outset that the general phenomenon of electromagnetic waves was predicted by theory before it was realized that light is a form of electromagnetic wave. The prediction was made by James Clerk Maxwell in the mid-19th century when he formulated a single theory combining all the electric and magnetic effects known by scientists at that time. "Electromagnetic waves" was the name he gave to the phenomena his theory predicted.

 Figure 5.22.2: The electromagnetic waves sent and received by this 50-foot radar dish antenna at Kennedy Space Center in Florida are not visible, but help track expendable launch vehicles with high-definition imagery. The first use of this C-band radar dish was for the launch of the Atlas V rocket sending the New Horizons probe toward Pluto. (credit: NASA)

Such a theoretical prediction followed by experimental verification is an indication of the power of science in general, and physics in particular. The underlying connections and unity of physics allow certain great minds to solve puzzles without having all the pieces. The prediction of electromagnetic waves is one of the most spectacular examples of this power. Certain others, such as the prediction of antimatter, will be discussed in later modules.

This page titled 5.22: Prelude to Electromagnetic Waves is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 24.0: Prelude to Electromagnetic Waves by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 5.23: The Electromagnetic Spectrum

## Learning Objectives

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

- List three “rules of thumb” that apply to the different frequencies along the electromagnetic spectrum.
- Explain why the higher the frequency, the shorter the wavelength of an electromagnetic wave.
- Draw a simplified electromagnetic spectrum, indicating the relative positions, frequencies, and spacing of the different types of radiation bands.
- List and explain the different methods by which electromagnetic waves are produced across the spectrum.

In this module we examine how electromagnetic waves are classified into categories such as radio, infrared, ultraviolet, and so on, so that we can understand some of their similarities as well as some of their differences. We will also find that there are many connections with previously discussed topics, such as wavelength and resonance. A brief overview of the production and utilization of electromagnetic waves is found in Table $5.23.1$.

*Table* $5.23.1$*: Electromagnetic Wave*

| Type of EM wave | Production | Applications | Life sciences aspect | Issues |
|---|---|---|---|---|
| Radio & TV | Acceleration charges | Communications Remote controls | MRI | Requires controls for band use |
| Microwaves | Accelerating charges & thermal agitation | Communication Ovens Radar | Deep heating | Cell phone use |
| Infrared | Thermal agitations & electronic transitions | Thermal imaging Heating | Absorbed by atmosphere | Greenhouse effect |
| Visible light | Thermal agitations & electronic transitions | All pervasive | Photosynthesis Human vision | |
| Ultraviolet | Thermal agitations & electronic transitions | Sterilization Cancer control | Vitamin D production | Ozone depletion Cancer causing |
| X-rays | Inner electronic transitions and fast collisions | Medical Security | Medical diagnosis Cancer therapy | Cancer causing |
| Gamma rays | Nuclear decay | Nuclear medicineSecurity | Medical diagnosis Cancer therapy | Cancer causing Radiation damage |

## Wave

There are many types of waves, such as water waves and even earthquakes. Among the many shared attributes of waves are propagation speed, frequency, and wavelength. These are always related by the expression $vw = f\lambda$. This module concentrates on EM waves, but other modules contain examples of all of these characteristics for sound waves and submicroscopic particles.

As noted before, an electromagnetic wave has a frequency and a wavelength associated with it and travels at the speed of light, or $c$. The relationship among these wave characteristics can be described by $vw = f\lambda$, where $vw$ is the propagation speed of the wave, $f$ is the frequency, and $\lambda$ is the wavelength. Here $vw = c$, so that for all electromagnetic waves,

$$c = f\lambda. \tag{5.23.1}$$

Thus, for all electromagnetic waves, the greater the frequency, the smaller the wavelength.

Figure 5.23.1 shows how the various types of electromagnetic waves are categorized according to their wavelengths and frequencies -- that is, it shows the electromagnetic spectrum. Many of the characteristics of the various types of electromagnetic waves are related to their frequencies and wavelengths, as we shall see.

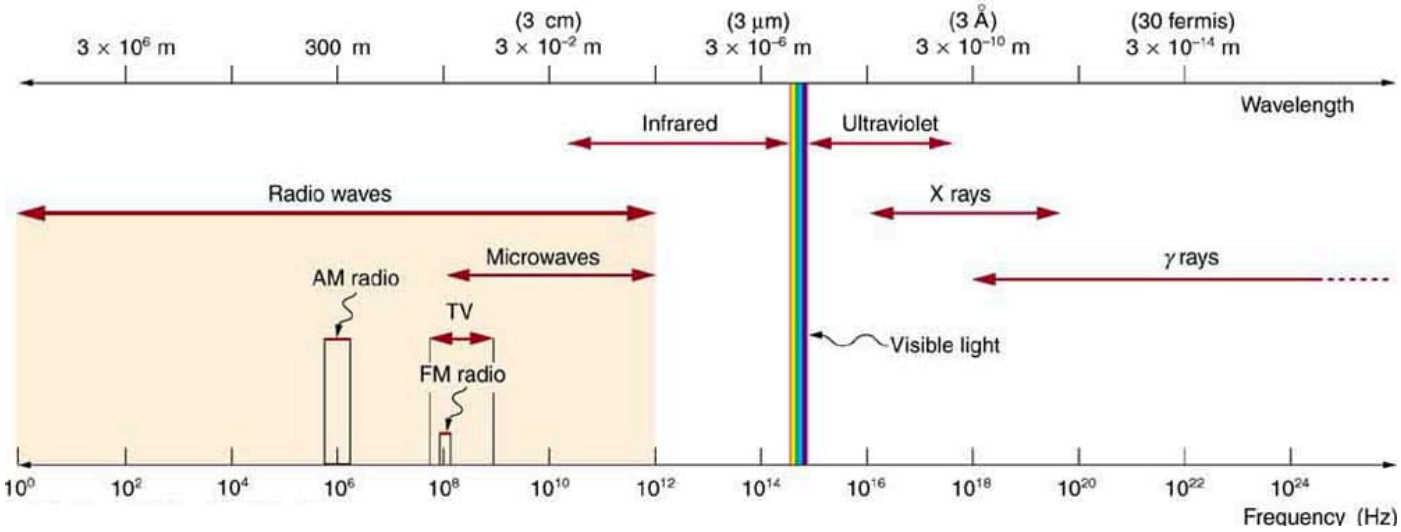



Figure 5.23.1: The electromagnetic spectrum, showing the major categories of electromagnetic waves. The range of frequencies and wavelengths is remarkable. The dividing line between some categories is distinct, whereas other categories overlap.

### ELECTROMAGNETIC SPECTRUM: RULES OF THUMB

Three rules that apply to electromagnetic waves in general are as follows:

- High-frequency electromagnetic waves are more energetic and are more able to penetrate than low-frequency waves.
- High-frequency electromagnetic waves can carry more information per unit time than low-frequency waves.
- The shorter the wavelength of any electromagnetic wave probing a material, the smaller the detail it is possible to resolve.

Note that there are exceptions to these rules of thumb.

## Transmission, Reflection, and Absorption

What happens when an electromagnetic wave impinges on a material? If the material is transparent to the particular frequency, then the wave can largely be transmitted. If the material is opaque to the frequency, then the wave can be totally reflected. The wave can also be absorbed by the material, indicating that there is some interaction between the wave and the material, such as the thermal agitation of molecules.

Of course it is possible to have partial transmission, reflection, and absorption. We normally associate these properties with visible light, but they do apply to all electromagnetic waves. What is not obvious is that something that is transparent to light may be opaque at other frequencies. For example, ordinary glass is transparent to visible light but largely opaque to ultraviolet radiation. Human skin is opaque to visible light -- we cannot see through people -- but transparent to X-rays.

## Radio and TV Waves

The broad category of **radio waves** is defined to contain any electromagnetic wave produced by currents in wires and circuits. Its name derives from their most common use as a carrier of audio information (i.e., radio). The name is applied to electromagnetic waves of similar frequencies regardless of source. Radio waves from outer space, for example, do not come from alien radio stations. They are created by many astronomical phenomena, and their study has revealed much about nature on the largest

scales.

There are many uses for radio waves, and so the category is divided into many subcategories, including microwaves and those electromagnetic waves used for AM and FM radio, cellular telephones, and TV.

The lowest commonly encountered radio frequencies are produced by high-voltage AC power transmission lines at frequencies of 50 or 60 Hz. (Figure 5.23.2). These extremely long wavelength electromagnetic waves (about 6000 km!) are one means of energy loss in long-distance power transmission.



Figure 5.23.2: This high-voltage traction power line running to Eutingen Railway Substation in Germany radiates electromagnetic waves with very long wavelengths. (credit: Zonk43, Wikimedia Commons)

There is an ongoing controversy regarding potential health hazards associated with exposure to these electromagnetic fields ( $E$-fields). Some people suspect that living near such transmission lines may cause a variety of illnesses, including cancer. But demographic data are either inconclusive or simply do not support the hazard theory. Recent reports that have looked at many European and American epidemiological studies have found no increase in risk for cancer due to exposure to $E$-fields.

**Extremely low frequency (ELF)** radio waves of about 1 kHz are used to communicate with submerged submarines. The ability of radio waves to penetrate salt water is related to their wavelength (much like ultrasound penetrating tissue) -- the longer the wavelength, the farther they penetrate. Since salt water is a good conductor, radio waves are strongly absorbed by it, and very long wavelengths are needed to reach a submarine under the surface. (Figure 5.23.3).

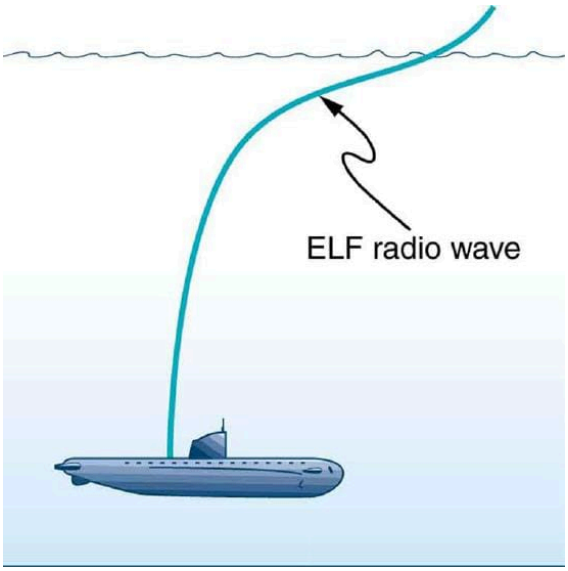



Figure 5.23.3: Very long wavelength radio waves are needed to reach this submarine, requiring extremely low frequency signals (ELF). Shorter wavelengths do not penetrate to any significant depth.

AM radio waves are used to carry commercial radio signals in the frequency range from 540 to 1600 kHz. The abbreviation AM stands for **amplitude modulation**, which is the method for placing information on these waves (Figure 5.23.4). A **carrier wave** having the basic frequency of the radio station, say 1530 kHz, is varied or modulated in amplitude by an audio signal. The resulting wave has a constant frequency, but a varying amplitude.

A radio receiver tuned to have the same resonant frequency as the carrier wave can pick up the signal, while rejecting the many other frequencies impinging on its antenna. The receiver's circuitry is designed to respond to variations in amplitude of the carrier wave to replicate the original audio signal. That audio signal is amplified to drive a speaker or perhaps to be recorded.

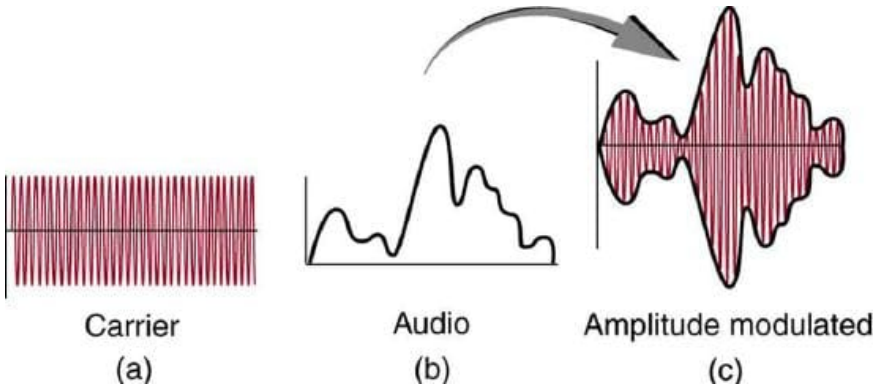



Figure 5.23.4: Amplitude modulation for AM radio. (a) A carrier wave at the station's basic frequency. (b) An audio signal at much lower audible frequencies. (c) The amplitude of the carrier is modulated by the audio signal without changing its basic frequency.

## FM Radio Waves

FM radio waves are also used for commercial radio transmission, but in the frequency range of 88 to 108 MHz. FM stands for **frequency modulation**, another method of carrying information (Figure 5.23.5). Here a carrier wave having the basic frequency of the radio station, perhaps 105.1 MHz, is modulated in frequency by the audio signal, producing a wave of constant amplitude but varying frequency.

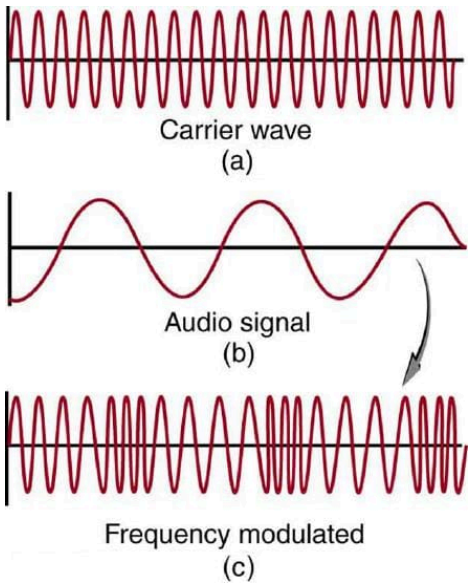



Figure 5.23.5: Frequency modulation for FM radio. (a) A carrier wave at the station's basic frequency. (b) An audio signal at much lower audible frequencies. (c) The frequency of the carrier is modulated by the audio signal without changing its amplitude.

Since audible frequencies range up to 20 kHz (or 0.020 MHz) at most, the frequency of the FM radio wave can vary from the carrier by as much as 0.020 MHz. Thus the carrier frequencies of two different radio stations cannot be closer than 0.020 MHz. An FM receiver is tuned to resonate at the carrier frequency and has circuitry that responds to variations in frequency, reproducing the audio information.

FM radio is inherently less subject to noise from stray radio sources than AM radio. The reason is that amplitudes of waves add. So an AM receiver would interpret noise added onto the amplitude of its carrier wave as part of the information. An FM receiver can be made to reject amplitudes other than that of the basic carrier wave and only look for variations in frequency. It is thus easier to reject noise from FM, since noise produces a variation in amplitude.

**Television** is also broadcast on electromagnetic waves. Since the waves must carry a great deal of visual as well as audio information, each channel requires a larger range of frequencies than simple radio transmission. TV channels utilize frequencies in the range of 54 to 88 MHz and 174 to 222 MHz. (The entire FM radio band lies between channels 88 MHz and 174 MHz.) These TV channels are called VHF (for **very high frequency**). Other channels called UHF (for **ultra high frequency**) utilize an even higher frequency range of 470 to 1000 MHz.

The TV video signal is AM, while the TV audio is FM. Note that these frequencies are those of free transmission with the user utilizing an old-fashioned roof antenna. Satellite dishes and cable transmission of TV occurs at significantly higher frequencies and is rapidly evolving with the use of the high-definition or HD format.

## Example 5.23.1: Calculating Wavelengths of Radio Waves:

Calculate the wavelengths of a 1530-kHz AM radio signal, a 105.1-MHz FM radio signal, and a 1.90-GHz cell phone signal.

**Strategy**

The relationship between wavelength and frequency is $c = f\lambda$, where $c = 3.00 \times 10^8 m/s$ is the speed of light (the speed of light is only very slightly smaller in air than it is in a vacuum). We can rearrange this equation to find the wavelength for all three frequencies.

**Solution**

Rearranging gives

$$\lambda = \frac{c}{f}. \tag{5.23.2}$$

a. For the $f = 1530kHz$ AM radio signal, then,

$$\lambda = \frac{3.00 \times 10^8 m/s}{1530 \times 10^3 cycles/s} \tag{5.23.3}$$

$$= 196m. \tag{5.23.4}$$

b. For the $f = 105.1MHz$ AM radio signal, then,

$$\lambda = \frac{3.00 \times 10^8 m/s}{105.1 \times 10^6 cycles/s} \tag{5.23.5}$$

$$= 2.85m. \tag{5.23.6}$$

c. For the $f = 1.90GHz$ AM radio signal, then,

$$\lambda = \frac{3.00 \times 10^8 m/s}{1.90 \times 10^9 cycles/s} \tag{5.23.7}$$

$$= 0.158m. \tag{5.23.8}$$

**Discussion**

These wavelengths are consistent with the spectrum in Figure 5.23.1. The wavelengths are also related to other properties of these electromagnetic waves, as we shall see.

The wavelengths found in the preceding example are representative of AM, FM, and cell phones, and account for some of the differences in how they are broadcast and how well they travel. The most efficient length for a linear antenna, such as discussed in 24.3, is $\lambda/2$, half the wavelength of the electromagnetic wave. Thus a very large antenna is needed to efficiently broadcast typical AM radio with its carrier wavelengths on the order of hundreds of meters.

One benefit to these long AM wavelengths is that they can go over and around rather large obstacles (like buildings and hills), just as ocean waves can go around large rocks. FM and TV are best received when there is a line of sight between the broadcast antenna and receiver, and they are often sent from very tall structures. FM, TV, and mobile phone antennas themselves are much smaller than those used for AM, but they are elevated to achieve an unobstructed line of sight (Figure $5.23.6$).




Figure $5.23.6$: (a) A large tower is used to broadcast TV signals. The actual antennas are small structures on top of the tower -- they are placed at great heights to have a clear line of sight over a large broadcast area. (credit: Ozizo, Wikimedia Commons) (b) The NTT Dokomo mobile phone tower at Tokorozawa City, Japan. (credit: tokoroten, Wikimedia Commons)

## Radio Wave Interference

Astronomers and astrophysicists collect signals from outer space using electromagnetic waves. A common problem for astrophysicists is the “pollution” from electromagnetic radiation pervading our surroundings from communication systems in general. Even everyday gadgets like our car keys having the facility to lock car doors remotely and being able to turn TVs on and off using remotes involve radio-wave frequencies. In order to prevent interference between all these electromagnetic signals, strict regulations are drawn up for different organizations to utilize different radio frequency bands.

One reason why we are sometimes asked to switch off our mobile phones (operating in the range of 1.9 GHz) on airplanes and in hospitals is that important communications or medical equipment often uses similar radio frequencies and their operation can be affected by frequencies used in the communication devices.

For example, radio waves used in magnetic resonance imaging (MRI) have frequencies on the order of 100 MHz, although this varies significantly depending on the strength of the magnetic field used and the nuclear type being scanned. MRI is an important medical imaging and research tool, producing highly detailed two- and three-dimensional images. Radio waves are broadcast, absorbed, and reemitted in a resonance process that is sensitive to the density of nuclei (usually protons or hydrogen nuclei).

The wavelength of 100-MHz radio waves is 3 m, yet using the sensitivity of the resonant frequency to the magnetic field strength, details smaller than a millimeter can be imaged. This is a good example of an exception to a rule of thumb (in this case, the rubric that details much smaller than the probe’s wavelength cannot be detected). The intensity of the radio waves used in MRI presents little or no hazard to human health.

## Microwaves

**Microwaves** are the highest-frequency electromagnetic waves that can be produced by currents in macroscopic circuits and devices. Microwave frequencies range from about $10^9 Hz$ to the highest practical $LC$ resonance at nearly $10^{12} Hz$. Since they have high frequencies, their wavelengths are short compared with those of other radio waves -- hence the name "microwave."

Microwaves can also be produced by atoms and molecules. They are, for example, a component of electromagnetic radiation generated by **thermal agitation.** The thermal motion of atoms and molecules in any object at a temperature above absolute zero causes them to emit and absorb radiation.

Since it is possible to carry more information per unit time on high frequencies, microwaves are quite suitable for communications. Most satellite-transmitted information is carried on microwaves, as are land-based long-distance transmissions. A clear line of sight between transmitter and receiver is needed because of the short wavelengths involved.

**Radar** is a common application of microwaves that was first developed in World War II. By detecting and timing microwave echoes, radar systems can determine the distance to objects as diverse as clouds and aircraft. A Doppler shift in the radar echo can be used to determine the speed of a car or the intensity of a rainstorm. Sophisticated radar systems are used to map the Earth and other planets, with a resolution limited by wavelength. (Figure 5.23.7). The shorter the wavelength of any probe, the smaller the detail it is possible to observe.

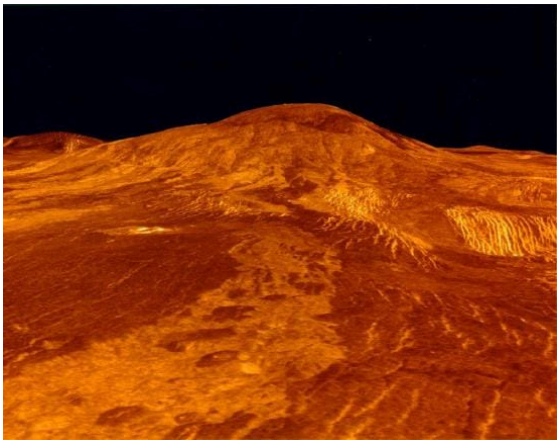


Figure 5.23.7: An image of Sif Mons with lava flows on Venus, based on Magellan synthetic aperture radar data combined with radar altimetry to produce a three-dimensional map of the surface. The Venusian atmosphere is opaque to visible light, but not to the microwaves that were used to create this image. (credit: NSSDC, NASA/JPL)

## Heating with Microwaves

How does the ubiquitous microwave oven produce microwaves electronically, and why does food absorb them preferentially? Microwaves at a frequency of 2.45 GHz are produced by accelerating electrons. The microwaves are then used to induce an alternating electric field in the oven.

Water and some other constituents of food have a slightly negative charge at one end and a slightly positive charge at one end (called polar molecules). The range of microwave frequencies is specially selected so that the polar molecules, in trying to keep orienting themselves with the electric field, absorb these energies and increase their temperatures -- called dielectric heating.

The energy thereby absorbed results in thermal agitation heating food and not the plate, which does not contain water. Hot spots in the food are related to constructive and destructive interference patterns. Rotating antennas and food turntables help spread out the hot spots.

Another use of microwaves for heating is within the human body. Microwaves will penetrate more than shorter wavelengths into tissue and so can accomplish “deep heating” (called microwave diathermy). This is used for treating muscular pains, spasms, tendonitis, and rheumatoid arthritis.

### MAKING CONNECTIONS: TAKE-HOME EXPERIMENT - MICROWAVE OVENS

1. Look at the door of a microwave oven. Describe the structure of the door. Why is there a metal grid on the door? How does the size of the holes in the grid compare with the wavelengths of microwaves used in microwave ovens? What is this wavelength?
2. Place a glass of water (about 250 ml) in the microwave and heat it for 30 seconds. Measure the temperature gain (the $\Delta T$). Assuming that the power output of the oven is 1000 W, calculate the efficiency of the heat-transfer process.
3. Remove the rotating turntable or moving plate and place a cup of water in several places along a line parallel with the opening. Heat for 30 seconds and measure the $\Delta T$ for each position. Do you see cases of destructive interference?

Microwaves generated by atoms and molecules far away in time and space can be received and detected by electronic circuits. Deep space acts like a blackbody with a 2.7 K temperature, radiating most of its energy in the microwave frequency range. In 1964, Penzias and Wilson detected this radiation and eventually recognized that it was the radiation of the Big Bang's cooled remnants.

## Infrared Radiation

The microwave and infrared regions of the electromagnetic spectrum overlap (Figure $5.23.1$). **Infrared radiation** is generally produced by thermal motion and the vibration and rotation of atoms and molecules. Electronic transitions in atoms and molecules can also produce infrared radiation.

The range of infrared frequencies extends up to the lower limit of visible light, just below red. In fact, infrared means "below red." Frequencies at its upper limit are too high to be produced by accelerating electrons in circuits, but small systems, such as atoms and molecules, can vibrate fast enough to produce these waves.

Water molecules rotate and vibrate particularly well at infrared frequencies, emitting and absorbing them so efficiently that the emissivity for skin is $e = 0.97$ in the infrared. Night-vision scopes can detect the infrared emitted by various warm objects, including humans, and convert it to visible light.

We can examine radiant heat transfer from a house by using a camera capable of detecting infrared radiation. Reconnaissance satellites can detect buildings, vehicles, and even individual humans by their infrared emissions, whose power radiation is proportional to the fourth power of the absolute temperature. More mundanely, we use infrared lamps, some of which are called quartz heaters, to preferentially warm us because we absorb infrared better than our surroundings.

The Sun radiates like a nearly perfect blackbody (that is, it has $e = 1$), with a 6000 K surface temperature. About half of the solar energy arriving at the Earth is in the infrared region, with most of the rest in the visible part of the spectrum, and a relatively small amount in the ultraviolet. On average, 50 percent of the incident solar energy is absorbed by the Earth.

The relatively constant temperature of the Earth is a result of the energy balance between the incoming solar radiation and the energy radiated from the Earth. Most of the infrared radiation emitted from the Earth is absorbed by $CO_2$ and $H_2O$ in the atmosphere and then radiated back to Earth or into outer space. This radiation back to Earth is known as the greenhouse effect, and it maintains the surface temperature of the Earth about $40^{\circ}C$ higher than it would be if there is no absorption. Some scientists think that the increased concentration of $CO_2$ and other greenhouse gases in the atmosphere, resulting from increases in fossil fuel burning, has increased global average temperatures.

## Visible Light

**Visible light** is the narrow segment of the electromagnetic spectrum to which the normal human eye responds. Visible light is produced by vibrations and rotations of atoms and molecules, as well as by electronic transitions within atoms and molecules. The receivers or detectors of light largely utilize electronic transitions. We say the atoms and molecules are excited when they absorb and relax when they emit through electronic transitions.

Figure $5.23.8$ shows this part of the spectrum, together with the colors associated with particular pure wavelengths. We usually refer to visible light as having wavelengths of between 400 nm and 750 nm. (The retina of the eye actually responds to the lowest ultraviolet frequencies, but these do not normally reach the retina because they are absorbed by the cornea and lens of the eye.)

Red light has the lowest frequencies and longest wavelengths, while violet has the highest frequencies and shortest wavelengths. Blackbody radiation from the Sun peaks in the visible part of the spectrum but is more intense in the red than in the violet, making the Sun yellowish in appearance.

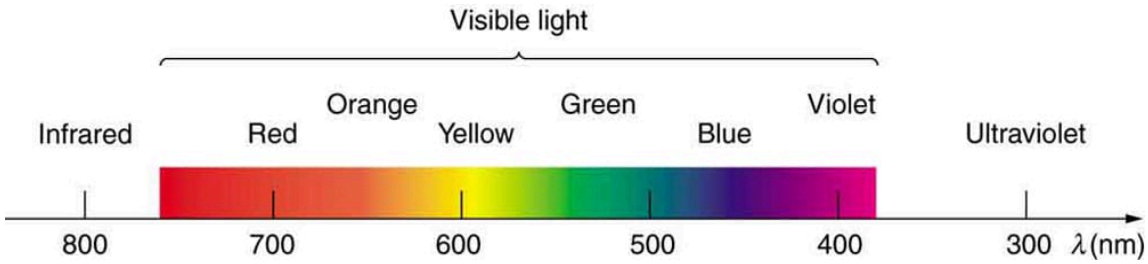



Figure 5.23.8: A small part of the electromagnetic spectrum that includes its visible components. The divisions between infrared, visible, and ultraviolet are not perfectly distinct, nor are those between the seven rainbow colors.

Living things--plants and animals-- have evolved to utilize and respond to parts of the electromagnetic spectrum they are embedded in. Visible light is the most predominant and we enjoy the beauty of nature through visible light. Plants are more selective. Photosynthesis makes use of parts of the visible spectrum to make sugars.

## Example 5.23.2: Integrated Concept Problem: Correcting Vision with Lasers

During laser vision correction, a brief burst of 193-nm ultraviolet light is projected onto the cornea of a patient. It makes a spot 0.80 mm in diameter and evaporates a layer of cornea $0.30\mu m$ thick. Calculate the energy absorbed, assuming the corneal tissue has the same properties as water; it is initially at $34^{\circ}C$. Assume the evaporated tissue leaves at a temperature of $100^{\circ}C$.

**Strategy**

The energy from the laser light goes toward raising the temperature of the tissue and also toward evaporating it. Thus we have two amounts of heat to add together. Also, we need to find the mass of corneal tissue involved.

**Solution**

To figure out the heat required to raise the temperature of the tissue to $100^{\circ}C$, we can apply concepts of thermal energy. We know that

$$Q = mc\Delta T, \tag{5.23.9}$$

where Q is the heat required to raise the temperature, $\Delta T$ is the desired change in temperature, $m$ is the mass of tissue to be heated, and $c$ is the specific heat of water equal to 4186 J/kg/K.

Without knowing the mass $m$ at this point, we have

$$\begin{aligned} Q &= m\left(4186 J/kg/K\right)\left(100^{\circ}C - 34^{\circ}C\right) \\ &= m\left(276,276 J/kg\right) \\ &= m\left(276 J/kg\right). \end{aligned}$$

The latent heat of vaporization of water is 2256 kJ/kg, so that the energy needed to evaporate mass $m$ is

$$Q_v = mL_v = m\left(2256 kJ/kg\right).$$

To find the mass $m$, we use the equation $\rho = m/V$, where $\rho$ is the density of the tissue and $V$ is its volume. For this case,

$$\begin{aligned} m &= \rho V \\ &= \left(1000 kg/m^3\right)\left(area \times thickness\left(m^3\right)\right) \\ &= \left(1000 kg/m^3\right)\left(\pi\left(0.80 \times 10^{-3} m\right)^2/4\right)\left(0.30 \times 10^{-6} m\right) \\ &= 0.151 \times 10^{-9} kg. \end{aligned}$$

Therefore, the total energy absorbed by the tissue in the eye is the sum of $Q$ and $Q_v$:

$$\begin{aligned} Q_{tot} &= m\left(c\Delta T + L_v\right) \\ &= \left(0.151 \times 10^{-9} kg\right)\left(276 kJ/kg + 2256 kJ/kg\right) \\ &= 382 \times 10^{-9} kJ. \end{aligned}$$

**Discussion**

The lasers used for this eye surgery are excimer lasers, whose light is well absorbed by biological tissue. They evaporate rather than burn the tissue, and can be used for precision work. Most lasers used for this type of eye surgery have an average power rating of about one watt. For our example, if we assume that each laser burst from this pulsed laser lasts for 10 ns, and there are 400 bursts per second, then the average power is

$$Q_{tot} \times 400 = 150\, mW$$

Optics is the study of the behavior of visible light and other forms of electromagnetic waves. Optics falls into two distinct categories. When electromagnetic radiation, such as visible light, interacts with objects that are large compared with its wavelength, its motion can be represented by straight lines like rays. Ray optics is the study of such situations and includes lenses and mirrors.

When electromagnetic radiation interacts with objects about the same size as the wavelength or smaller, its wave nature becomes apparent. For example, observable detail is limited by the wavelength, and so visible light can never detect individual atoms, because they are so much smaller than its wavelength. Physical or wave optics is the study of such situations and includes all wave characteristics.

TAKE-HOME EXPERIMENT: COLORS THAT MATCH

When you light a match you see largely orange light; when you light a gas stove you see blue light. Why are the colors different? What other colors are present in these?

## Ultraviolet Radiation

Ultraviolet means “above violet.” The electromagnetic frequencies of **ultraviolet radiation (UV)** extend upward from violet, the highest-frequency visible light. Ultraviolet is also produced by atomic and molecular motions and electronic transitions. The wavelengths of ultraviolet extend from 400 nm down to about 10 nm at its highest frequencies, which overlap with the lowest X-ray frequencies. It was recognized as early as 1801 by Johann Ritter that the solar spectrum had an invisible component beyond the violet range.

Solar UV radiation is broadly subdivided into three regions: UV-A (320–400 nm), UV-B (290–320 nm), and UV-C (220–290 nm), ranked from long to shorter wavelengths (from smaller to larger energies). Most UV-B and all UV-C is absorbed by ozone ($O_3$) molecules in the upper atmosphere. Consequently, 99% of the solar UV radiation reaching the Earth’s surface is UV-A.

## Human Exposure to UV Radiation

It is largely exposure to UV-B that causes skin cancer. It is estimated that as many as 20% of adults will develop skin cancer over the course of their lifetime. Again, treatment is often successful if caught early. Despite very little UV-B reaching the Earth’s surface, there are substantial increases in skin-cancer rates in countries such as Australia, indicating how important it is that UV-B and UV-C continue to be absorbed by the upper atmosphere.

All UV radiation can damage collagen fibers, resulting in an acceleration of the aging process of skin and the formation of wrinkles. Because there is so little UV-B and UV-C reaching the Earth’s surface, sunburn is caused by large exposures, and skin cancer from repeated exposure. Some studies indicate a link between overexposure to the Sun when young and melanoma later in life.

The tanning response is a defense mechanism in which the body produces pigments to absorb future exposures in inert skin layers above living cells. Basically UV-B radiation excites DNA molecules, distorting the DNA helix, leading to mutations and the

possible formation of cancerous cells.

Repeated exposure to UV-B may also lead to the formation of cataracts in the eyes -- a cause of blindness among people living in the equatorial belt where medical treatment is limited. Cataracts, clouding in the eye's lens and a loss of vision, are age related; 60% of those between the ages of 65 and 74 will develop cataracts. However, treatment is easy and successful, as one replaces the lens of the eye with a plastic lens. Prevention is important. Eye protection from UV is more effective with plastic sunglasses than those made of glass.

A major acute effect of extreme UV exposure is the suppression of the immune system, both locally and throughout the body.

Low-intensity ultraviolet is used to sterilize haircutting implements, implying that the energy associated with ultraviolet is deposited in a manner different from lower-frequency electromagnetic waves. (Actually this is true for all electromagnetic waves with frequencies greater than visible light.)

Flash photography is generally not allowed of precious artworks and colored prints because the UV radiation from the flash can cause photo-degradation in the artworks. Often artworks will have an extra-thick layer of glass in front of them, which is especially designed to absorb UV radiation.

## UV Light and the Ozone Layer

If all of the Sun's ultraviolet radiation reached the Earth's surface, there would be extremely grave effects on the biosphere from the severe cell damage it causes. However, the layer of ozone ($O_3$) in our upper atmosphere (10 to 50 km above the Earth) protects life by absorbing most of the dangerous UV radiation.

Unfortunately, today we are observing a depletion in ozone concentrations in the upper atmosphere. This depletion has led to the formation of an "ozone hole" in the upper atmosphere. The hole is more centered over the southern hemisphere, and changes with the seasons, being largest in the spring. This depletion is attributed to the breakdown of ozone molecules by refrigerant gases called chlorofluorocarbons (CFCs).

The UV radiation helps dissociate the CFC's, releasing highly reactive chlorine (Cl) atoms, which catalyze the destruction of the ozone layer. For example, the reaction of $CFCl_3$ with a photon of light ($hv$) can be written as:

$$\mathrm{CFCl_3 + hv \ \rightarrow CFCl_2 + Cl.} \tag{5.23.10}$$

The Cl atom then catalyzes the breakdown of ozone as follows:

$$\mathrm{Cl + O_3 \ \rightarrow ClO + O_2} \tag{5.23.11}$$

and

$$\mathrm{ClO + O_3 \ \rightarrow Cl + 2\,O_2.} \tag{5.23.12}$$

A single chlorine atom could destroy ozone molecules for up to two years before being transported down to the surface. The CFCs are relatively stable and will contribute to ozone depletion for years to come. CFCs are found in refrigerants, air conditioning systems, foams, and aerosols.

International concern over this problem led to the establishment of the "Montreal Protocol" agreement (1987) to phase out CFC production in most countries. However, developing-country participation is needed if worldwide production and elimination of CFCs is to be achieved. Probably the largest contributor to CFC emissions today is India. But the protocol seems to be working, as there are signs of an ozone recovery (Figure 5.23.9).

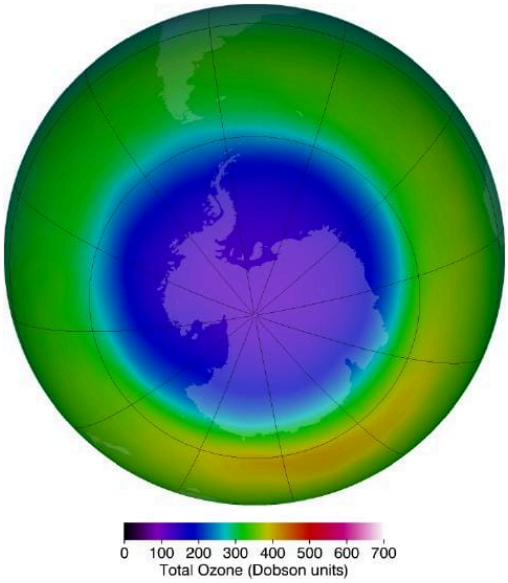



Figure 5.23.9: This map of ozone concentration over Antarctica in October 2011 shows severe depletion suspected to be caused by CFCs. Less dramatic but more general depletion has been observed over northern latitudes, suggesting the effect is global. With less ozone, more ultraviolet radiation from the Sun reaches the surface, causing more damage. (credit: NASA Ozone Watch)

## Benefits of UV Light

Besides the adverse effects of ultraviolet radiation, there are also benefits of exposure in nature and uses in technology. Vitamin D production in the skin (epidermis) results from exposure to UVB radiation, generally from sunlight. A number of studies indicate lack of vitamin D can result in the development of a range of cancers (prostate, breast, colon), so a certain amount of UV exposure is helpful. Lack of vitamin D is also linked to osteoporosis. Exposures (with no sunscreen) of 10 minutes a day to arms, face, and legs might be sufficient to provide the accepted dietary level. However, in the winter time north of about $37^{\circ}$ latitude, most UVB gets blocked by the atmosphere.

UV radiation is used in the treatment of infantile jaundice and in some skin conditions. It is also used in sterilizing workspaces and tools, and killing germs in a wide range of applications. It is also used as an analytical tool to identify substances.

When exposed to ultraviolet, some substances, such as minerals, glow in characteristic visible wavelengths, a process called fluorescence. So-called black lights emit ultraviolet to cause posters and clothing to fluoresce in the visible. Ultraviolet is also used in special microscopes to detect details smaller than those observable with longer-wavelength visible-light microscopes.

### THINGS GREAT AND SMALL: A SUBMICROSCOPIC VIEW OF X-RAY PRODUCTION

X-rays can be created in a high-voltage discharge. They are emitted in the material struck by electrons in the discharge current. There are two mechanisms by which the electrons create X-rays.

The first method is illustrated in the Figure 5.23.10. An electron is accelerated in an evacuated tube by a high positive voltage. The electron strikes a metal plate (e.g., copper) and produces X-rays. Since this is a high-voltage discharge, the electron gains sufficient energy to ionize the atom.

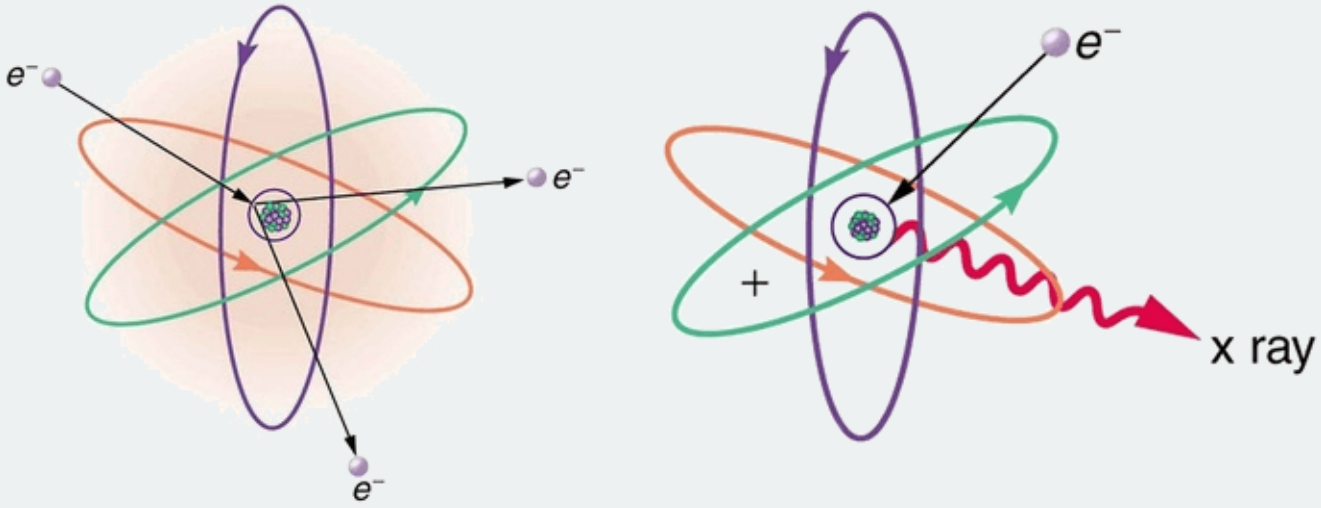



Figure 5.23.10: Artist's conception of an electron ionizing an atom followed by the recapture of an electron and emission of an X-ray. An energetic electron strikes an atom and knocks an electron out of one of the orbits closest to the nucleus. Later, the atom captures another electron, and the energy released by its fall into a low orbit generates a high-energy EM wave called an X-ray.

In the case shown, an inner-shell electron (one in an orbit relatively close to and tightly bound to the nucleus) is ejected. A short time later, another electron is captured and falls into the orbit in a single great plunge. The energy released by this fall is given to an EM wave known as an X-ray. Since the orbits of the atom are unique to the type of atom, the energy of the X-ray is characteristic of the atom, hence the name characteristic X-ray.

The second method by which an energetic electron creates an X-ray when it strikes a material is illustrated in the figure below. The electron interacts with charges in the material as it penetrates. These collisions transfer kinetic energy from the electron to the electrons and atoms in the material.

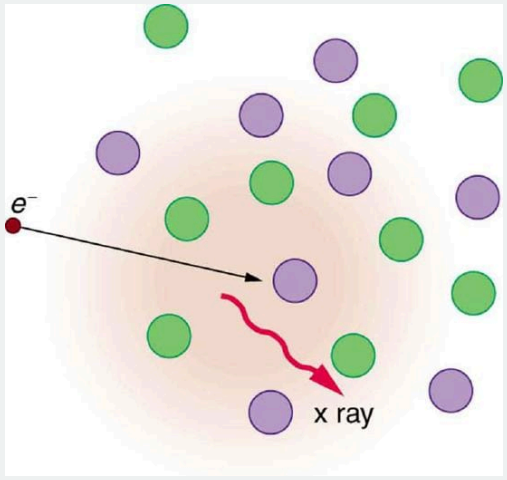



Figure 5.23.11: Artist's conception of an electron being slowed by collisions in a material and emitting X-ray radiation. This energetic electron makes numerous collisions with electrons and atoms in a material it penetrates. An accelerated charge radiates EM waves, a second method by which X-rays are created.

A loss of kinetic energy implies an acceleration, in this case decreasing the electron's velocity. Whenever a charge is accelerated, it radiates EM waves. Given the high energy of the electron, these EM waves can have high energy. We call them X-rays. Since the process is random, a broad spectrum of X-ray energy is emitted that is more characteristic of the electron energy than the type of material the electron encounters. Such EM radiation is called "bremsstrahlung" (German for "braking radiation").

# X-Rays

In the 1850s, scientists (such as Faraday) began experimenting with high-voltage electrical discharges in tubes filled with rarefied gases. It was later found that these discharges created an invisible, penetrating form of very high frequency electromagnetic radiation. This radiation was called an **X-ray**, because its identity and nature were unknown.

As described in "Things Great and Small," there are two methods by which X-rays are created --both are submicroscopic processes and can be caused by high-voltage discharges. While the low-frequency end of the X-ray range overlaps with the ultraviolet, X-rays extend to much higher frequencies (and energies).

X-rays have adverse effects on living cells similar to those of ultraviolet radiation, and they have the additional liability of being more penetrating, affecting more than the surface layers of cells. Cancer and genetic defects can be induced by exposure to X-rays. Because of their effect on rapidly dividing cells, X-rays can also be used to treat and even cure cancer.

The widest use of X-rays is for imaging objects that are opaque to visible light, such as the human body or aircraft parts. In humans, the risk of cell damage is weighed carefully against the benefit of the diagnostic information obtained. However, questions have risen in recent years as to accidental overexposure of some people during CT scans --a mistake at least in part due to poor monitoring of radiation dose.

The ability of X-rays to penetrate matter depends on density, and so an X-ray image can reveal very detailed density information. Figure 5.23.12 shows an example of the simplest type of X-ray image, an X-ray shadow on film. The amount of information in a simple X-ray image is impressive, but more sophisticated techniques, such as CT scans, can reveal three-dimensional information with details smaller than a millimeter.

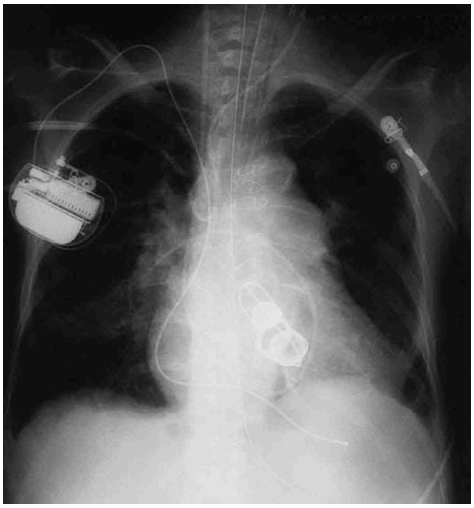


Figure 5.23.12: This shadow X-ray image shows many interesting features, such as artificial heart valves, a pacemaker, and the wires used to close the sternum. (credit: P. P. Urone)

The use of X-ray technology in medicine is called radiology -- an established and relatively cheap tool in comparison to more sophisticated technologies. Consequently, X-rays are widely available and used extensively in medical diagnostics. During World War I, mobile X-ray units, advocated by Madame Marie Curie, were used to diagnose soldiers.

Because they can have wavelengths less than 0.01 nm, X-rays can be scattered (a process called X-ray diffraction) to detect the shape of molecules and the structure of crystals. X-ray diffraction was crucial to Crick, Watson, and Wilkins in the determination of the shape of the double-helix DNA molecule.

X-rays are also used as a precise tool for trace-metal analysis in X-ray induced fluorescence, in which the energy of the X-ray emissions are related to the specific types of elements and amounts of materials present.

# Gamma Rays

Soon after nuclear radioactivity was first detected in 1896, it was found that at least three distinct types of radiation were being emitted. The most penetrating nuclear radiation was called **a gamma ray** ( $\gamma$ ray) (again a name given because its identity and

character were unknown), and it was later found to be an extremely high frequency electromagnetic wave.

In fact, $\gamma$ rays are any electromagnetic radiation emitted by a nucleus. This can be from natural nuclear decay or induced nuclear processes in nuclear reactors and weapons. The lower end of the $\gamma$**-ray** frequency range overlaps the upper end of the X-ray range, but $\gamma$ rays can have the highest frequency of any electromagnetic radiation.

Gamma rays have characteristics identical to X-rays of the same frequency -- they differ only in source. At higher frequencies, $\gamma$ rays are more penetrating and more damaging to living tissue. They have many of the same uses as X-rays, including cancer therapy. Gamma radiation from radioactive materials is used in nuclear medicine.

Figure 13 shows a medical image based on $\gamma$ rays. Food spoilage can be greatly inhibited by exposing it to large doses of $\gamma$ radiation, thereby obliterating responsible microorganisms. Damage to food cells through irradiation occurs as well, and the long-term hazards of consuming radiation-preserved food are unknown and controversial for some groups. Both X-ray and $\gamma$**-ray** technologies are also used in scanning luggage at airports.

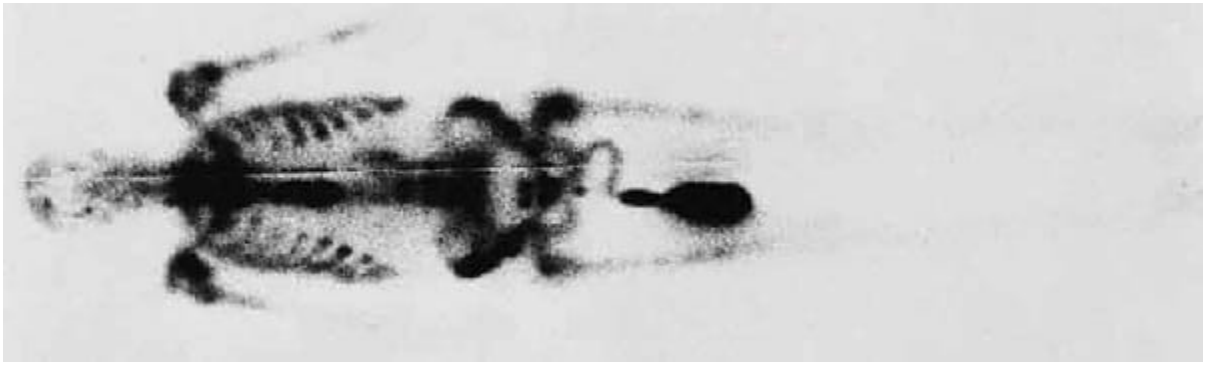


Figure 5.23.13: This is an image of the $\gamma$ rays emitted by nuclei in a compound that is concentrated in the bones and eliminated through the kidneys. Bone cancer is evidenced by nonuniform concentration in similar structures. For example, some ribs are darker than others. (credit: P. P. Urone)

## Detecting Electromagnetic Waves from Space

A final note on star gazing. The entire electromagnetic spectrum is used by researchers for investigating stars, space, and time. As noted earlier, Penzias and Wilson detected microwaves to identify the background radiation originating from the Big Bang. Radio telescopes such as the Arecibo Radio Telescope in Puerto Rico and Parkes Observatory in Australia were designed to detect radio waves.

Infrared telescopes need to have their detectors cooled by liquid nitrogen to be able to gather useful signals. Since infrared radiation is predominantly from thermal agitation, if the detectors were not cooled, the vibrations of the molecules in the antenna would be stronger than the signal being collected.

The most famous of these infrared sensitive telescopes is the James Clerk Maxwell Telescope in Hawaii. The earliest telescopes, developed in the seventeenth century, were optical telescopes, collecting visible light. Telescopes in the ultraviolet, X-ray, and $\gamma$ -ray regions are placed outside the atmosphere on satellites orbiting the Earth.

The Hubble Space Telescope (launched in 1990) gathers ultraviolet radiation as well as visible light. In the X-ray region, there is the Chandra X-ray Observatory (launched in 1999), and in the $\gamma$-ray region, there is the new Fermi Gamma-ray Space Telescope (launched in 2008—taking the place of the Compton Gamma Ray Observatory, 1991–2000.).

### PHET EXPLORATIONS: COLOR VISION

Make a whole rainbow by mixing red, green, and blue light. Change the wavelength of a monochromatic beam or filter white light. View the light as a solid beam, or see the individual photons.

# Summary

- The relationship among the speed of propagation, wavelength, and frequency for any wave is given by $v_w = f\lambda$, so that for electromagnetic waves,

$$c = f\lambda,$$

  where $f$ is the frequency, $\lambda$ is the wavelength, and $c$ is the speed of light.
- The electromagnetic spectrum is separated into many categories and subcategories, based on the frequency and wavelength, source, and uses of the electromagnetic waves.
- Any electromagnetic wave produced by currents in wires is classified as a radio wave, the lowest frequency electromagnetic waves. Radio waves are divided into many types, depending on their applications, ranging up to microwaves at their highest frequencies.
- Infrared radiation lies below visible light in frequency and is produced by thermal motion and the vibration and rotation of atoms and molecules. Infrared's lower frequencies overlap with the highest-frequency microwaves.
- Visible light is largely produced by electronic transitions in atoms and molecules, and is defined as being detectable by the human eye. Its colors vary with frequency, from red at the lowest to violet at the highest.
- Ultraviolet radiation starts with frequencies just above violet in the visible range and is produced primarily by electronic transitions in atoms and molecules.
- X-rays are created in high-voltage discharges and by electron bombardment of metal targets. Their lowest frequencies overlap the ultraviolet range but extend to much higher values, overlapping at the high end with gamma rays.
- Gamma rays are nuclear in origin and are defined to include the highest-frequency electromagnetic radiation of any type.

# Glossary

**electromagnetic spectrum**
: the full range of wavelengths or frequencies of electromagnetic radiation

**radio waves**
: electromagnetic waves with wavelengths in the range from 1 mm to 100 km; they are produced by currents in wires and circuits and by astronomical phenomena

**microwaves**
: electromagnetic waves with wavelengths in the range from 1 mm to 1 m; they can be produced by currents in macroscopic circuits and devices

**thermal agitation**
: the thermal motion of atoms and molecules in any object at a temperature above absolute zero, which causes them to emit and absorb radiation

**radar**
: a common application of microwaves. Radar can determine the distance to objects as diverse as clouds and aircraft, as well as determine the speed of a car or the intensity of a rainstorm

**infrared radiation (IR)**
: a region of the electromagnetic spectrum with a frequency range that extends from just below the red region of the visible light spectrum up to the microwave region, or from $0.74\mu m$ to $300\mu m$

**ultraviolet radiation (UV)**
: electromagnetic radiation in the range extending upward in frequency from violet light and overlapping with the lowest X-ray frequencies, with wavelengths from 400 nm down to about 10 nm

**visible light**
: the narrow segment of the electromagnetic spectrum to which the normal human eye responds

**amplitude modulation (AM)**
: a method for placing information on electromagnetic waves by modulating the amplitude of a carrier wave with an audio signal, resulting in a wave with constant frequency but varying amplitude

**extremely low frequency (ELF)**
: electromagnetic radiation with wavelengths usually in the range of 0 to 300 Hz, but also about 1kHz

**carrier wave**
: an electromagnetic wave that carries a signal by modulation of its amplitude or frequency

**frequency modulation (FM)**
: a method of placing information on electromagnetic waves by modulating the frequency of a carrier wave with an audio signal, producing a wave of constant amplitude but varying frequency

**TV**
: video and audio signals broadcast on electromagnetic waves

**very high frequency (VHF)**
: TV channels utilizing frequencies in the two ranges of 54 to 88 MHz and 174 to 222 MHz

**ultra-high frequency (UHF)**
: TV channels in an even higher frequency range than VHF, of 470 to 1000 MHz

**X-ray**
: invisible, penetrating form of very high frequency electromagnetic radiation, overlapping both the ultraviolet range and the $\gamma$-ray range

**gamma ray**
: ($\gamma$ ray); extremely high frequency electromagnetic radiation emitted by the nucleus of an atom, either from natural nuclear decay or induced nuclear processes in nuclear reactors and weapons. The lower end of the $\gamma$-ray frequency range overlaps the upper end of the X-ray range, but $\gamma$ rays can have the highest frequency of any electromagnetic radiation

This page titled 5.23: The Electromagnetic Spectrum is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 24.3: The Electromagnetic Spectrum by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6: MODULE 5- OPTICS


6.1: Prelude to Geometric Optics

6.2: The Ray Aspect of Light

6.3: The Law of Reflection

6.4: The Law of Refraction

6.5: Dispersion - Rainbows and Prisms

6.6: Image Formation by Lenses

6.7: Image Formation by Mirrors


This page titled 6: Module 5- Optics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

# 6.1: Prelude to Geometric Optics

## Geometric Optics

Light from this page or screen is formed into an image by the lens of your eye, much as the lens of the camera that made this photograph. Mirrors, like lenses, can also form images that in turn are captured by your eye.



Figure 6.1.1: Image seen as a result of reflection of light on a plane smooth surface. (credit: NASA Goddard Photo and Video, via Flickr)

Our lives are filled with light. Through vision, the most valued of our senses, light can evoke spiritual emotions, such as when we view a magnificent sunset or glimpse a rainbow breaking through the clouds. Light can also simply amuse us in a theater, or warn us to stop at an intersection. It has innumerable uses beyond vision. Light can carry telephone signals through glass fibers or cook a meal in a solar oven. Life itself could not exist without light's energy. From photosynthesis in plants to the sun warming a cold-blooded animal, its supply of energy is vital.



Figure 6.1.2: Double Rainbow over the bay of Pocitos in Montevideo, Uruguay. (credit: Madrax, Wikimedia Commons)

We already know that visible light is the type of electromagnetic waves to which our eyes respond. That knowledge still leaves many questions regarding the nature of light and vision. What is color, and how do our eyes detect it? Why do diamonds sparkle? How does light travel? How do lenses and mirrors form images? These are but a few of the questions that are answered by the study of optics. Optics is the branch of physics that deals with the behavior of visible light and other electromagnetic waves. In particular, optics is concerned with the generation and propagation of light and its interaction with matter. What we have already learned about the generation of light in our study of heat transfer by radiation will be expanded upon in later topics, especially

those on atomic physics. Now, we will concentrate on the propagation of light and its interaction with matter.

It is convenient to divide optics into two major parts based on the size of objects that light encounters. When light interacts with an object that is several times as large as the light's wavelength, its observable behavior is like that of a ray; it does not prominently display its wave characteristics. We call this part of optics "geometric optics." This chapter will concentrate on such situations. When light interacts with smaller objects, it has very prominent wave characteristics, such as constructive and destructive interference. "Wave Optics" will concentrate on such situations.

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

- 25.0: Prelude to Geometric Optics by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.2: The Ray Aspect of Light

## Learning Objectives

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

- List the ways by which light travels from a source to another location.

There are three ways in which light can travel from a source to another location (Figure 6.2.1). It can come directly from the source through empty space, such as from the Sun to Earth. Or light can travel through various media, such as air and glass, to the person. Light can also arrive after being reflected, such as by a mirror. In all of these cases, light is modeled as traveling in straight lines called rays. Light may change direction when it encounters objects (such as a mirror) or in passing from one material to another (such as in passing from air to glass), but it then continues in a straight line or as a ray. The word **ray** comes from mathematics and here means a straight line that originates at some point. It is acceptable to visualize light rays as laser rays (or even science fiction depictions of ray guns).

## Definition: RAY

The word “ray” comes from mathematics and here means a straight line that originates at some point.

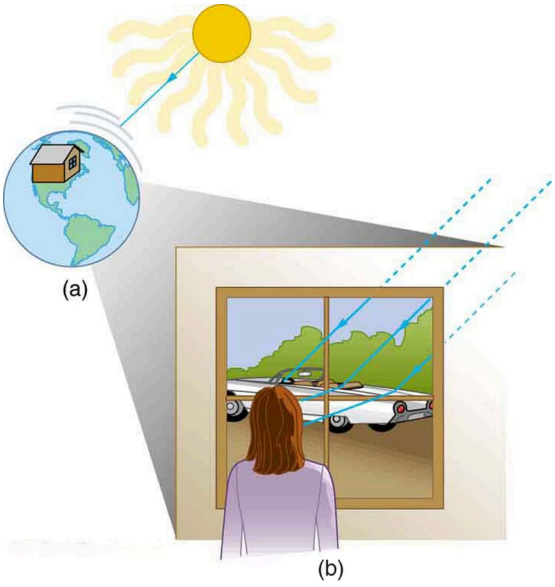



Figure 6.2.1: Three methods for light to travel from a source to another location. (a) Light reaches the upper atmosphere of Earth traveling through empty space directly from the source. (b) Light can reach a person in one of two ways. It can travel through media like air and glass. It can also reflect from an object like a mirror. In the situations shown here, light interacts with objects large enough that it travels in straight lines, like a ray.

Experiments, as well as our own experiences, show that when light interacts with objects several times as large as its wavelength, it travels in straight lines and acts like a ray. Its wave characteristics are not pronounced in such situations. Since the wavelength of light is less than a micron (a thousandth of a millimeter), it acts like a ray in the many common situations in which it encounters objects larger than a micron. For example, when light encounters anything we can observe with unaided eyes, such as a mirror, it acts like a ray, with only subtle wave characteristics. We will concentrate on the ray characteristics in this chapter.

Since light moves in straight lines, changing directions when it interacts with materials, it is described by geometry and simple trigonometry. This part of optics, where the ray aspect of light dominates, is therefore called **geometric optics**. There are two laws that govern how light changes direction when it interacts with matter. These are the law of reflection, for situations in which light bounces off matter, and the law of refraction, for situations in which light passes through matter.

Definition: GEOMETRIC OPTICS

The part of optics dealing with the ray aspect of light is called geometric optics.

## Summary

- A straight line that originates at some point is called a ray.
- The part of optics dealing with the ray aspect of light is called geometric optics.
- Light can travel in three ways from a source to another location: (1) directly from the source through empty space; (2) through various media; (3) after being reflected from a mirror.

## Glossary

**ray**
straight line that originates at some point

**geometric optics**
part of optics dealing with the ray aspect of light

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

- 25.1: The Ray Aspect of Light by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.3: The Law of Reflection

## Learning Objectives

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

- Explain reflection of light from polished and rough surfaces.

Whenever we look into a mirror, or squint at sunlight glinting from a lake, we are seeing a reflection. When you look at this page, too, you are seeing light reflected from it. Large telescopes use reflection to form an image of stars and other astronomical objects.

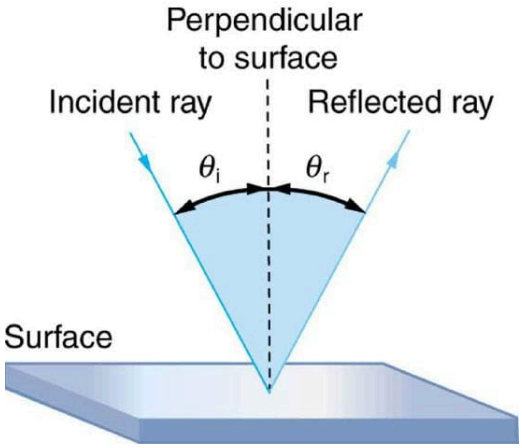



Figure $6.3.1$: The law of reflection states that the angle of reflection equals the angle of incidence -- $\theta_r = \theta_i$. The angles are measured relative to the perpendicular to the surface at the point where the ray strikes the surface.

The law of reflection is illustrated in Figure $6.3.1$, which also shows how the angles are measured relative to the perpendicular to the surface at the point where the light ray strikes. We expect to see reflections from smooth surfaces, but Figure $6.3.2$ illustrates how a rough surface reflects light. Since the light strikes different parts of the surface at different angles, it is reflected in many different directions, or diffused.

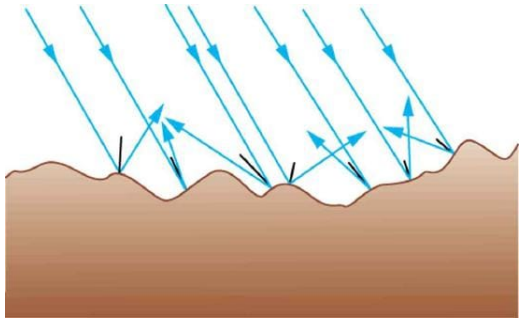


Figure $6.3.2$: Light is diffused when it reflects from a rough surface. Here many parallel rays are incident, but they are reflected at many different angles since the surface is rough.

Diffused light is what allows us to see a sheet of paper from any angle, as illustrated in Figure $6.3.3a$. Many objects, such as people, clothing, leaves, and walls, have rough surfaces and can be seen from all sides. A mirror, on the other hand, has a smooth surface (compared with the wavelength of light) and reflects light at specific angles, as illustrated in Figure $6.3.3b$. When the moon reflects from a lake, as shown in Figure $6.3.4$, a combination of these effects takes place.

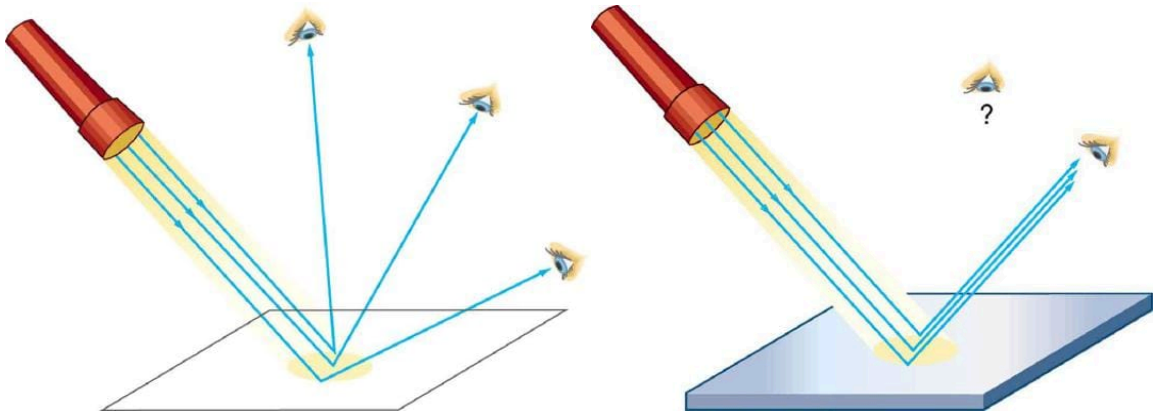



Figure 6.3.3: (left) When a sheet of paper is illuminated with many parallel incident rays, it can be seen at many different angles, because its surface is rough and diffuses the light. (right) A mirror illuminated by many parallel rays reflects them in only one direction, since its surface is very smooth. Only the observer at a particular angle will see the reflected light.

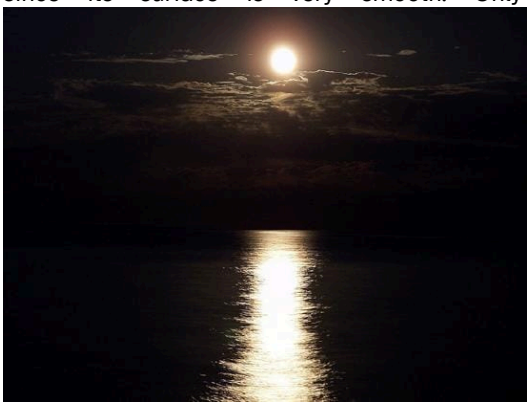


Figure 6.3.4: Moonlight is spread out when it is reflected by the lake, since the surface is shiny but uneven. (credit: Diego Torres Silvestre, Flickr)

The law of reflection is very simple: The angle of reflection equals the angle of incidence.

**Definition: THE LAW OF REFLECTION**

The angle of reflection equals the angle of incidence.

When we see ourselves in a mirror, it appears that our image is actually behind the mirror. This is illustrated in Figure 6.3.5. We see the light coming from a direction determined by the law of reflection. The angles are such that our image is exactly the same distance behind the mirror as we stand away from the mirror. If the mirror is on the wall of a room, the images in it are all behind the mirror, which can make the room seem bigger. Although these mirror images make objects appear to be where they cannot be (like behind a solid wall), the images are not figments of our imagination. Mirror images can be photographed and videotaped by instruments and look just as they do with our eyes (optical instruments themselves). The precise manner in which images are formed by mirrors and lenses will be treated in later sections of this chapter

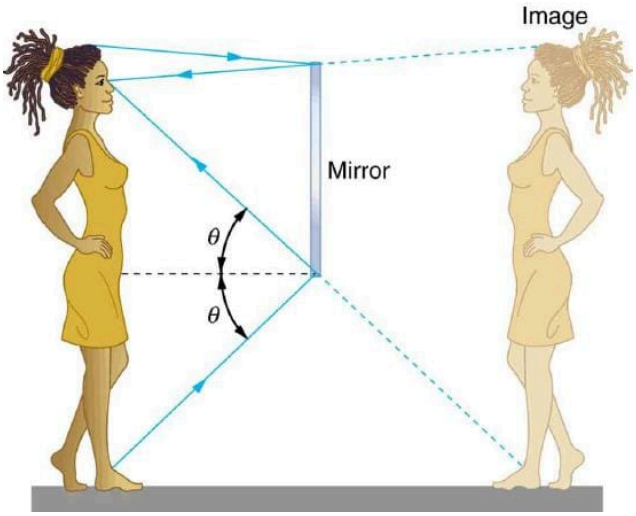



Figure 6.3.5: Our image in a mirror is behind the mirror. The two rays shown are those that strike the mirror at just the correct angles to be reflected into the eyes of the person. The image appears to be in the direction the rays are coming from when they enter the eyes.

### TAKE-HOME EXPERIMENT: LAW OF REFLECTION

Take a piece of paper and shine a flashlight at an angle at the paper, as shown in Figure 6.3.5*a*. Now shine the flashlight at a mirror at an angle. Do your observations confirm the predictions in Figure 6.3.3? Shine the flashlight on various surfaces and determine whether the reflected light is diffuse or not. You can choose a shiny metallic lid of a pot or your skin. Using the mirror and flashlight, can you confirm the law of reflection? You will need to draw lines on a piece of paper showing the incident and reflected rays. (This part works even better if you use a laser pencil.)

## Summary

- The angle of reflection equals the angle of incidence.
- A mirror has a smooth surface and reflects light at specific angles.
- Light is diffused when it reflects from a rough surface.
- Mirror images can be photographed and videotaped by instruments.

## Glossary

**mirror**
: smooth surface that reflects light at specific angles, forming an image of the person or object in front of it

**law of reflection**
: angle of reflection equals the angle of incidence

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

- 25.2: The Law of Reflection by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.4: The Law of Refraction

## Learning Objectives

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

- Determine the index of refraction, given the speed of light in a medium.

It is easy to notice some odd things when looking into a fish tank. For example, you may see the same fish appearing to be in two different places (Figure 6.4.1). This is because light coming from the fish to us changes direction when it leaves the tank, and in this case, it can travel two different paths to get to our eyes. The changing of a light ray's direction (loosely called bending) when it passes through variations in matter is called **refraction**. Refraction is responsible for a tremendous range of optical phenomena, from the action of lenses to voice transmission through optical fibers.

## Definition: REFRACTION

The changing of a light ray's direction (loosely called bending) when it passes through variations in matter is called refraction.

## SPEED OF LIGHT

The speed of light $c$ not only affects refraction, it is one of the central concepts of Einstein's theory of relativity. As the accuracy of the measurements of the speed of light were improved, $c$ was found not to depend on the velocity of the source or the observer. However, the speed of light does vary in a precise manner with the material it traverses. These facts have far-reaching implications, as we will see in "Special Relativity." It makes connections between space and time and alters our expectations that all observers measure the same time for the same event, for example. The speed of light is so important that its value in a vacuum is one of the most fundamental constants in nature as well as being one of the four fundamental SI units.

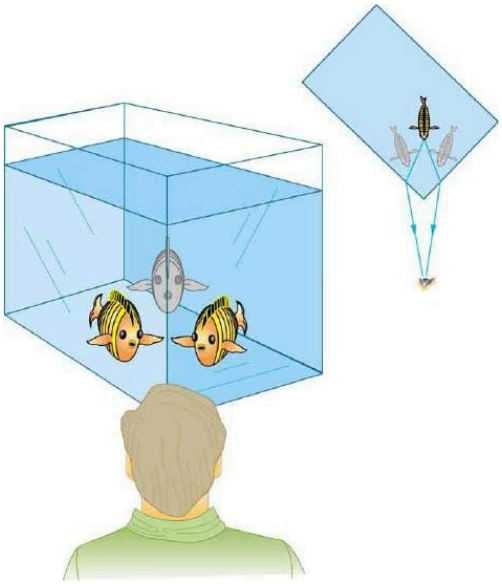


Figure 6.4.1: Looking at the fish tank as shown, we can see the same fish in two different locations, because light changes directions when it passes from water to air. In this case, the light can reach the observer by two different paths, and so the fish seems to be in two different places. This bending of light is called refraction and is responsible for many optical phenomena.

Why does light change direction when passing from one material (medium) to another? It is because light changes speed when

going from one material to another. So before we study the law of refraction, it is useful to discuss the speed of light and how it varies in different media.

## The Speed of Light

Early attempts to measure the speed of light, such as those made by Galileo, determined that light moved extremely fast, perhaps instantaneously. The first real evidence that light traveled at a finite speed came from the Danish astronomer Ole Roemer in the late 17th century. Roemer had noted that the average orbital period of one of Jupiter's moons, as measured from Earth, varied depending on whether Earth was moving toward or away from Jupiter. He correctly concluded that the apparent change in period was due to the change in distance between Earth and Jupiter and the time it took light to travel this distance. From his 1676 data, a value of the speed of light was calculated to be $2.26 \times 10^8 m/s$ (only 25% different from today's accepted value). In more recent times, physicists have measured the speed of light in numerous ways and with increasing accuracy. One particularly direct method, used in 1887 by the American physicist Albert Michelson (1852–1931), is illustrated in Figure $6.4.2$. Light reflected from a rotating set of mirrors was reflected from a stationary mirror 35 km away and returned to the rotating mirrors. The time for the light to travel can be determined by how fast the mirrors must rotate for the light to be returned to the observer's eye.

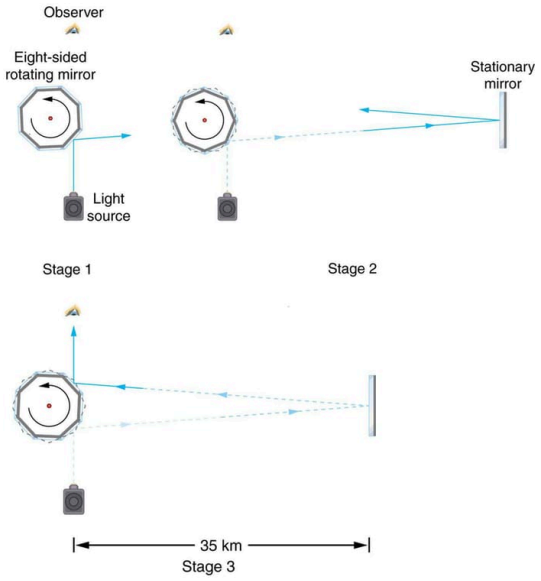



Figure $6.4.2$: A schematic of early apparatus used by Michelson and others to determine the speed of light. As the mirrors rotate, the reflected ray is only briefly directed at the stationary mirror. The returning ray will be reflected into the observer's eye only if the next mirror has rotated into the correct position just as the ray returns. By measuring the correct rotation rate, the time for the round trip can be measured and the speed of light calculated. Michelson's calculated value of the speed of light was only 0.04% different from the value used today.

The speed of light is now known to great precision. In fact, the speed of light in a vacuum $c$ is so important that it is accepted as one of the basic physical quantities and has the fixed value.

**VALUE OF THE SPEED OF LIGHT**

$$c \equiv 2.99792458 \times 10^8 \tag{6.4.1}$$

$$\sim 3.00 \times 10^8 m/s \tag{6.4.2}$$

The approximate value of $3.00 \times 10^8 m/s$ is used whenever three-digit accuracy is sufficient. The speed of light through matter is less than it is in a vacuum, because light interacts with atoms in a material. The speed of light depends strongly on the type of material, since its interaction with different atoms, crystal lattices, and other substructures varies.

## Definition: INDEX OF REFRACTION

We define the *index of refraction* $n$ of a material to be

$$n = \frac{c}{v}, \tag{6.4.3}$$

where $v$ is the observed speed of light in the material. Since the speed of light is always less than $c$ in matter and equals $c$ only in a vacuum, the index of refraction is always greater than or equal to one. That is, $n > 1$.

Table $6.4.1$ gives the indices of refraction for some representative substances. The values are listed for a particular wavelength of light, because they vary slightly with wavelength. (This can have important effects, such as colors produced by a prism.) Note that for gases, $n$ is close to 1.0. This seems reasonable, since atoms in gases are widely separated and light travels at $c$ in the vacuum between atoms. It is common to take $n = 1$ for gases unless great precision is needed. Although the speed of light $v$ in a medium varies considerably from its value $c$ in a vacuum, it is still a large speed.

*Table* $6.4.1$*: Index of Refraction in Various Media*

| Medium | n |
|---|---|
| ***Gases at*** $0\ ^{o}C, 1atm$ | |
| Air | 1.000293 |
| Carbon dioxide | 1.00045 |
| Hydrogen | 1.000139 |
| Oxygen | 1.000271 |
| ***Liquids at 20°C*** | |
| Benzene | 1.501 |
| Carbon disulfide | 1.628 |
| Carbon tetrachloride | 1.461 |
| Ethanol | 1.361 |
| Glycerine | 1.473 |
| Water, fresh | 1.333 |
| ***Solids at 20°C*** | |
| Diamond | 2.419 |
| Fluorite | 1.434 |
| Glass, crown | 1.52 |
| Glass, flint | 1.66 |
| ***Ice at 20°C*** | 1.309 |
| Polystyrene | 1.49 |
| Plexiglas | 1.51 |
| Quartz, crystalline | 1.544 |
| Quartz, fused | 1.458 |
| Sodium chloride | 1.544 |
| Zircon | 1.923 |

### Example 6.4.1: Speed of Light in Matter

Calculate the speed of light in zircon, a material used in jewelry to imitate diamond.

**Strategy:**

The speed of light in a material, $v$, can be calculated from the index of refraction $n$ of the material using the equation $n = c/v$ .

**Solution**

The equation for index of refraction (Equation 6.4.3) can be rearranged to determine $v$

$$v = \frac{c}{n}.$$

The index of refraction for zircon is given as 1.923 in Table 6.4.1, and $c$ is given in the equation for speed of light. Entering these values in the last expression gives

$$\begin{aligned} v &= \frac{3.00 \times 10^8 m/s}{1.923} \\ &= 1.56 \times 10^8 m/s. \end{aligned}$$

**Discussion:**

This speed is slightly larger than half the speed of light in a vacuum and is still high compared with speeds we normally experience. The only substance listed in Table 6.4.1 that has a greater index of refraction than zircon is diamond. We shall see later that the large index of refraction for zircon makes it sparkle more than glass, but less than diamond.

## Law of Refraction

Figure 6.4.3 shows how a ray of light changes direction when it passes from one medium to another. As before, the angles are measured relative to a perpendicular to the surface at the point where the light ray crosses it. (Some of the incident light will be reflected from the surface, but for now we will concentrate on the light that is transmitted.) The change in direction of the light ray depends on how the speed of light changes. The change in the speed of light is related to the indices of refraction of the media involved. In the situations shown in Figure 6.4.3, medium 2 has a greater index of refraction than medium 1. This means that the speed of light is less in medium 2 than in medium 1. Note that as shown in Figure 6.4.3$a$, the direction of the ray moves closer to the perpendicular when it slows down. Conversely, as shown in Figure 6.4.3$b$, the direction of the ray moves away from the perpendicular when it speeds up. The path is exactly reversible. In both cases, you can imagine what happens by thinking about pushing a lawn mower from a footpath onto grass, and vice versa. Going from the footpath to grass, the front wheels are slowed and pulled to the side as shown. This is the same change in direction as for light when it goes from a fast medium to a slow one. When going from the grass to the footpath, the front wheels can move faster and the mower changes direction as shown. This, too, is the same change in direction as for light going from slow to fast.

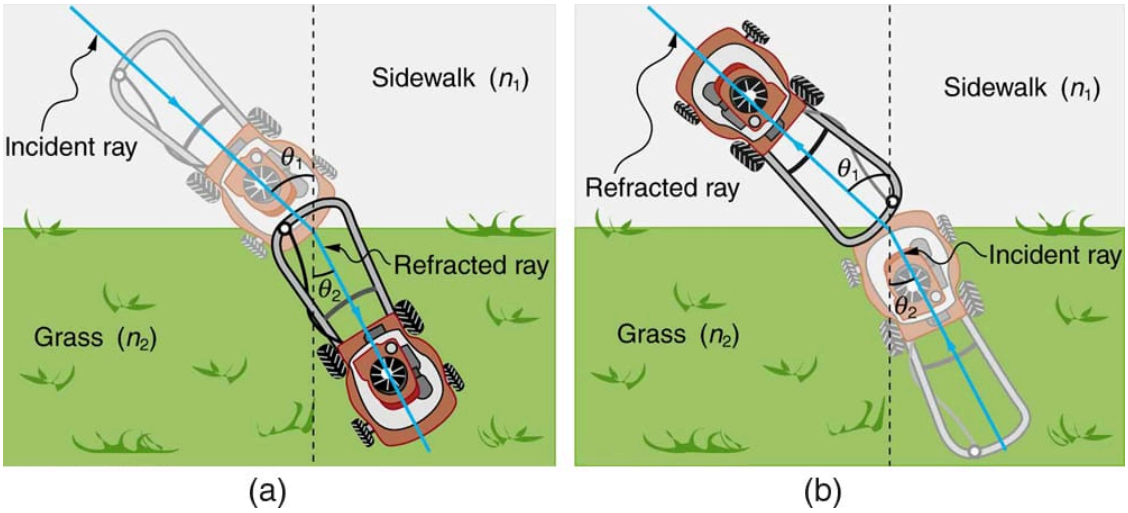



Figure 6.4.3: The change in direction of a light ray depends on how the speed of light changes when it crosses from one medium to another. The speed of light is greater in medium 1 than in medium 2 in the situations shown here. (a) A ray of light moves closer to the perpendicular when it slows down. This is analogous to what happens when a lawn mower goes from a footpath to grass. (b) A ray of light moves away from the perpendicular when it speeds up. This is analogous to what happens when a lawn mower goes from grass to footpath. The paths are exactly reversible.

The amount that a light ray changes its direction depends both on the incident angle and the amount that the speed changes. For a ray at a given incident angle, a large change in speed causes a large change in direction, and thus a large change in angle. The exact mathematical relationship is the **law of refraction**, or "Snell's Law," which is stated in equation form as

### THE LAW OF REFRACTION (Snell's Law)

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{6.4.4}$$

Here, $n_1$ and $n_2$ are the indices of refraction for medium 1 and 2, and $\theta_1$ and $\theta_2$ are the angles between the rays and the perpendicular in medium 1 and 2, as shown in Figure 6.4.3. The incoming ray is called the incident ray and the outgoing ray the refracted ray, and the associated angles the incident angle and the refracted angle. The law of refraction is also called Snell's law after the Dutch mathematician Willebrord Snell (1591–1626), who discovered it in 1621. Snell's experiments showed that the law of refraction was obeyed and that a characteristic index of refraction $n$ could be assigned to a given medium. Snell was not aware that the speed of light varied in different media, but through experiments he was able to determine indices of refraction from the way light rays changed direction.

### TAKE-HOME EXPERIMENT: A BROKEN PENCIL

A classic observation of refraction occurs when a pencil is placed in a glass half filled with water. Do this and observe the shape of the pencil when you look at the pencil sideways, that is, through air, glass, water. Explain your observations. Draw ray diagrams for the situation.

### Example 6.4.2: Determine the Index of Refraction from Refraction Data

Find the index of refraction for medium 2 in Figure 6.4.3*a*, assuming medium 1 is air and given the incident angle is $30.0^\circ$ and the angle of refraction is $22.0^\circ$.

**Strategy**

The index of refraction for air is taken to be 1 in most cases (and up to four significant figures, it is 1.000). Thus $n_1 = 1.00$ here. From the given information, $\theta_1 = 30.0^\circ$ and $\theta_2 = 22.0^\circ$ With this information, the only unknown in Snell's law is $n_2$, so that it can be used to find this unknown.

**Solution**

Snell's law (Equation 6.4.4) can be rearranged to isolate $n_2$ gives

$$n_2 = n_1 \frac{\sin \theta_1}{\sin \theta_2}. \tag{6.4.5}$$

Entering known values,

$$\begin{aligned} n_2 &= n_1 \frac{\sin 30.0^\circ}{\sin 22.0^\circ} \\ &= \frac{0.500}{0.375} \\ &= 1.33. \end{aligned}$$

**Discussion**

This is the index of refraction for water, and Snell could have determined it by measuring the angles and performing this calculation. He would then have found 1.33 to be the appropriate index of refraction for water in all other situations, such as when a ray passes from water to glass. Today we can verify that the index of refraction is related to the speed of light in a medium by measuring that speed directly.

## Example 6.4.3: A Larger Change in Direction

Suppose that in a situation like that in the previous example, light goes from air to diamond and that the incident angle is $30.0^\circ$. Calculate the angle of refraction $\theta_2$ in the diamond.

**Strategy**

Again the index of refraction for air is taken to be $n_1 = 1.00$, and we are given $\theta_1 = 30.0^\circ$. We can look up the index of refraction for diamond in Table 6.4.1, finding $n_2 = 2.419$. The only unknown in Snell's law is $\theta_2$, which we wish to determine.

**Solution**

Solving Snell's law (Equation 6.4.4) for $\sin \theta_2$ yields

$$\sin \theta_2 = \frac{n_1}{n_2} \sin \theta_1. \tag{6.4.6}$$

Entering known values,

$$\begin{aligned} \sin \theta_2 &= \frac{1.00}{2.419} \sin 30.0^\circ \\ &= (0.413)\,(0.500) \\ &= 0.207. \end{aligned}$$

The angle is thus

$$\theta_2 = \sin 0.207^{-1} = 11.9^\circ. \tag{6.4.7}$$

**Discussion**

For the same $30^\circ$ angle of incidence, the angle of refraction in diamond is significantly smaller than in water ($11.9^\circ$ rather than $22^\circ$ -- see the preceding example).

## Summary


- The changing of a light ray's direction when it passes through variations in matter is called refraction.
- The speed of light in vacuuum $c = 2.99792458 \times 10^8 \sim 3.00 \times 10^8 m/s$
- Index of refraction $n = \frac{c}{v}$, where $v$ is the speed of light in the material, $c$ is the speed of light in vacuum, and $n$ is the index of refraction.
- Snell's law, the law of refraction, is stated in equation form as $n_1 \sin_{\theta_1} = n_2 \sin_{\theta_2}$.


## Glossary

**refraction**
: changing of a light ray's direction when it passes through variations in matter

**index of refraction**
: for a material, the ratio of the speed of light in vacuum to that in the material

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

- 25.3: The Law of Refraction by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.5: Dispersion - Rainbows and Prisms

## Learning Objectives

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

- Explain the phenomenon of dispersion and discuss its advantages and disadvantages.

Everyone enjoys the spectacle of a rainbow glimmering against a dark stormy sky. How does sunlight falling on clear drops of rain get broken into the rainbow of colors we see? The same process causes white light to be broken into colors by a clear glass prism or a diamond. (See Figure 1.)

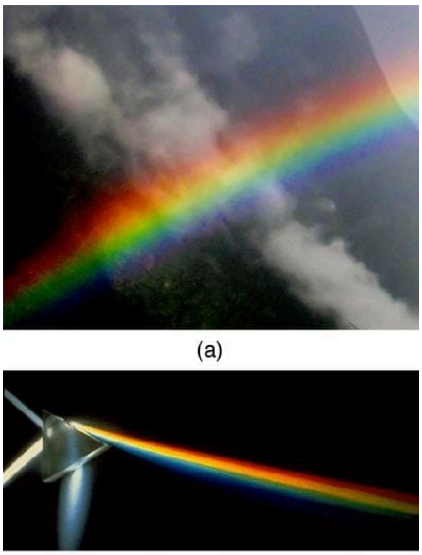



Figure 6.5.1: The colors of the rainbow (a) and those produced by a prism (b) are identical. (credit: Alfredo55, Wikimedia Commons; NASA)

We see about six colors in a rainbow -- red, orange, yellow, green, blue, and violet; sometimes indigo is listed, too. Those colors are associated with different wavelengths of light, as shown in Figure 1. When our eye receives pure-wavelength light, we tend to see only one of the six colors, depending on wavelength. The thousands of other hues we can sense in other situations are our eye's response to various mixtures of wavelengths. White light, in particular, is a fairly uniform mixture of all visible wavelengths. Sunlight, considered to be white, actually appears to be a bit yellow because of its mixture of wavelengths, but it does contain all visible wavelengths. The sequence of colors in rainbows is the same sequence as the colors plotted versus wavelength in Figure 2. What this implies is that white light is spread out according to wavelength in a rainbow. **Dispersion** is defined as the spreading of white light into its full spectrum of wavelengths. More technically, dispersion occurs whenever there is a process that changes the direction of light in a manner that depends on wavelength. Dispersion, as a general phenomenon, can occur for any type of wave and always involves wavelength-dependent processes.

## DISPERSION

Dispersion is defined to be the spreading of white light into its full spectrum of wavelengths.

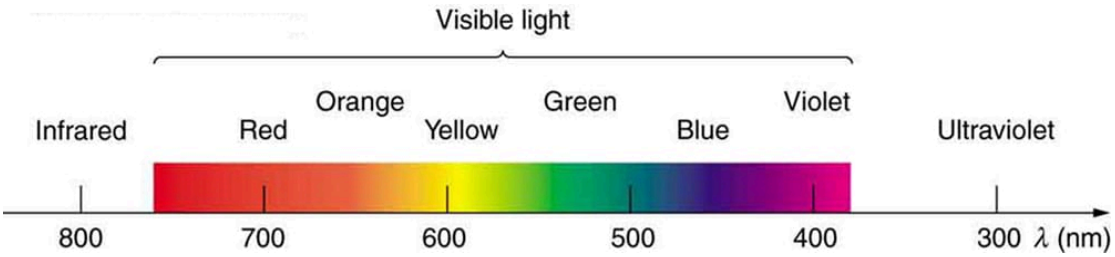



Figure 6.5.2: Even though rainbows are associated with seven colors, the rainbow is a continuous distribution of colors according to wavelengths.

Refraction is responsible for dispersion in rainbows and many other situations. The angle of refraction depends on the index of refraction, as we saw in "The Law of Refraction." We know that the index of refraction $n$ depends on the medium. But for a given medium, $n$ also depends on wavelength (Table 6.5.1). Note that, for a given medium, $n$ increases as wavelength decreases and is greatest for violet light. Thus violet light is bent more than red light, as shown for a prism in Figure 3 and the light is dispersed into the same sequence of wavelengths as seen in 1 and 2.

## MAKING CONNECTIONS: DISPERSION

Any type of wave can exhibit dispersion. Sound waves, all types of electromagnetic waves, and water waves can be dispersed according to wavelength. Dispersion occurs whenever the speed of propagation depends on wavelength, thus separating and spreading out various wavelengths. Dispersion may require special circumstances and can result in spectacular displays such as in the production of a rainbow. This is also true for sound, since all frequencies ordinarily travel at the same speed. If you listen to sound through a long tube, such as a vacuum cleaner hose, you can easily hear it is dispersed by interaction with the tube. Dispersion, in fact, can reveal a great deal about what the wave has encountered that disperses its wavelengths. The dispersion of electromagnetic radiation from outer space, for example, has revealed much about what exists between the stars -- the so-called empty space.

*Table 6.5.1: Index of Refraction n in Selected Media at Various Wavelengths*

| Medium | Red (660 nm) | Orange (610 nm) | Yellow (580 nm) | Green (550 nm) | Blue (470 nm) | Violet (410 nm) |
|---|---|---|---|---|---|---|
| Water | 1.331 | 1.332 | 1.333 | 1.335 | 1.338 | 1.342 |
| Diamond | 2.410 | 2.415 | 2.417 | 2.426 | 2.444 | 2.458 |
| Glass, crown | 1.512 | 1.514 | 1.518 | 1.519 | 1.524 | 1.530 |
| Glass, flint | 1.662 | 1.665 | 1.667 | 1.674 | 1.684 | 1.698 |
| Polustyrene | 1.488 | 1.490 | 1.492 | 1.493 | 1.499 | 1.506 |
| Quartz, fused | 1.455 | 1.456 | 1.458 | 1.459 | 1.462 | 1.468 |

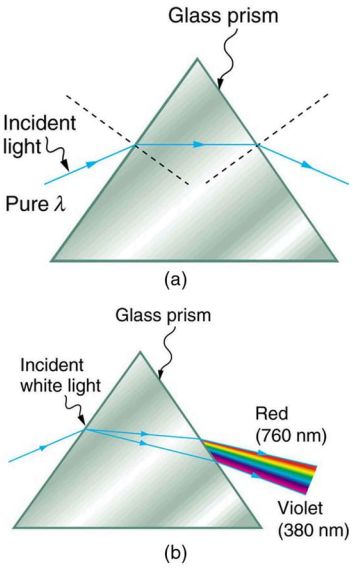



Figure 6.5.3: (a) A pure wavelength of light falls onto a prism and is refracted at both surfaces. (b) White light is dispersed by the prism (shown exaggerated). Since the index of refraction varies with wavelength, the angles of refraction vary with wavelength. A sequence of red to violet is produced, because the index of refraction increases steadily with decreasing wavelength.

Rainbows are produced by a combination of refraction and reflection. You may have noticed that you see a rainbow only when you look away from the sun. Light enters a drop of water and is reflected from the back of the drop, as shown in Figure 4. The light is refracted both as it enters and as it leaves the drop. Since the index of refraction of water varies with wavelength, the light is dispersed, and a rainbow is observed, as shown in Figure 5a. (There is no dispersion caused by reflection at the back surface, since the law of reflection does not depend on wavelength.) The actual rainbow of colors seen by an observer depends on the myriad of rays being refracted and reflected toward the observer's eyes from numerous drops of water. The effect is most spectacular when the background is dark, as in stormy weather, but can also be observed in waterfalls and lawn sprinklers. The arc of a rainbow comes from the need to be looking at a specific angle relative to the direction of the sun, as illustrated in Figure 5b. (If there are two reflections of light within the water drop, another "secondary" rainbow is produced. This rare event produces an arc that lies above the primary rainbow arc -- see Figure 5c.)

## RAINBOWS

Rainbows are produced by a combination of refraction and reflection.

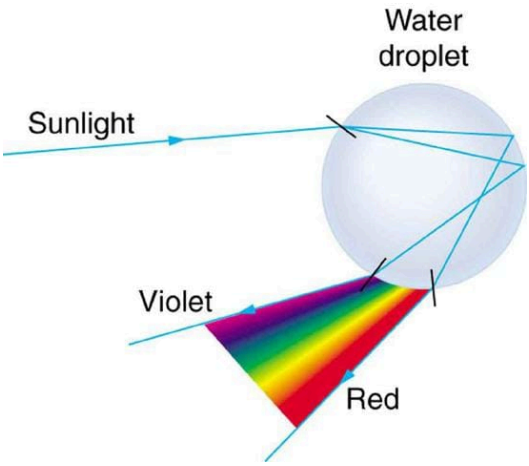
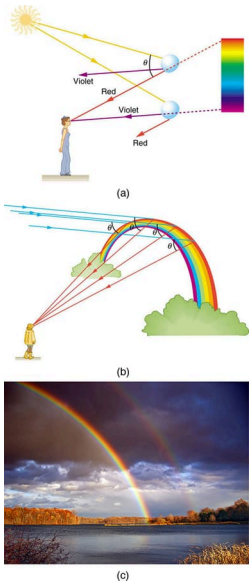



Figure 6.5.4: Part of the light falling on this water drop enters and is reflected




from the back of the drop. This light is refracted and dispersed both as it enters and as it leaves the drop.
Figure 6.5.5: (a) Different colors emerge in different directions, and so you must look at different locations to see the various colors of a rainbow. (b) The arc of a rainbow results from the fact that a line between the observer and any point on the arc must make the correct angle with the parallel rays of sunlight to receive the refracted rays. (c) Double rainbow. (credit: Nicholas, Wikimedia Commons)

Dispersion may produce beautiful rainbows, but it can cause problems in optical systems. White light used to transmit messages in a fiber is dispersed, spreading out in time and eventually overlapping with other messages. Since a laser produces a nearly pure wavelength, its light experiences little dispersion, an advantage over white light for transmission of information. In contrast, dispersion of electromagnetic waves coming to us from outer space can be used to determine the amount of matter they pass through. As with many phenomena, dispersion can be useful or a nuisance, depending on the situation and our human goals.

### PHET EXPLORATIONS: GEOMETRIC OPTICS

How does a lens form an image? See how light rays are refracted by a lens. Watch how the image changes when you adjust the focal length of the lens, move the object, move the lens, or move the screen.

## Summary

- The spreading of white light into its full spectrum of wavelengths is called dispersion.
- Rainbows are produced by a combination of refraction and reflection and involve the dispersion of sunlight into a continuous distribution of colors.
- Dispersion produces beautiful rainbows but also causes problems in certain optical systems.

## Glossary

**dispersion**
: spreading of white light into its full spectrum of wavelengths

**rainbow**
: dispersion of sunlight into a continuous distribution of colors according to wavelength, produced by the refraction and reflection of sunlight by water droplets in the sky

This page titled 6.5: Dispersion - Rainbows and Prisms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 25.5: Dispersion - Rainbows and Prisms by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.6: Image Formation by Lenses

## Learning Objectives

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

- List the rules for ray tracking for thin lenses.
- Illustrate the formation of images using the technique of ray tracking.
- Determine power of a lens given the focal length.

Lenses are found in a huge array of optical instruments, ranging from a simple magnifying glass to the eye to a camera's zoom lens. In this section, we will use the law of refraction to explore the properties of lenses and how they form images.

The word *lens* derives from the Latin word for a lentil bean, the shape of which is similar to the convex lens in Figure 1. The convex lens shown has been shaped so that all light rays that enter it parallel to its axis cross one another at a single point on the opposite side of the lens. (The axis is defined to be a line normal to the lens at its center, as shown in Figure 1.) Such a lens is called a **converging (or convex) lens** for the converging effect it has on light rays. An expanded view of the path of one ray through the lens is shown, to illustrate how the ray changes direction both as it enters and as it leaves the lens. Since the index of refraction of the lens is greater than that of air, the ray moves towards the perpendicular as it enters and away from the perpendicular as it leaves. (This is in accordance with the law of refraction.) Due to the lens's shape, light is thus bent toward the axis at both surfaces. The point at which the rays cross is defined to be the **focal point** F of the lens. The distance from the center of the lens to its focal point is defined to be the **focal length** $f$ of the lens. Figure 2 shows how a converging lens, such as that in a magnifying glass, can converge the nearly parallel light rays from the sun to a small spot.

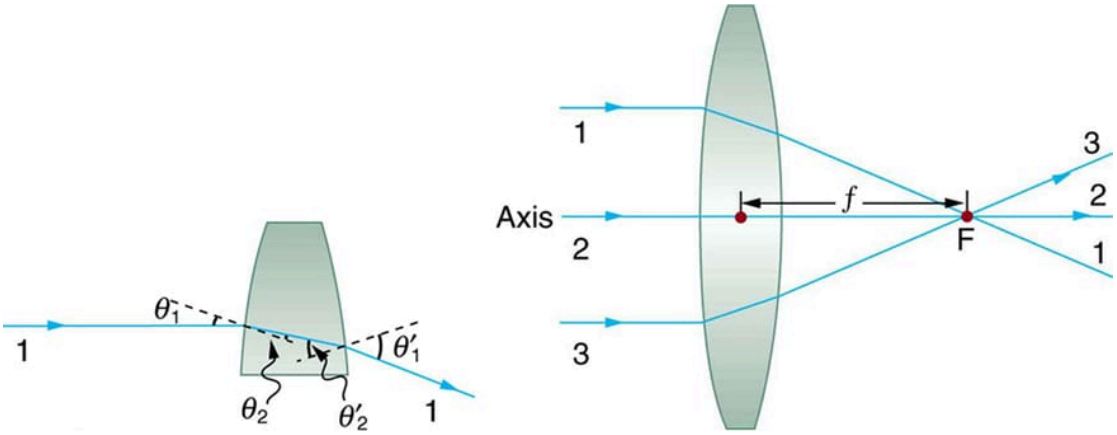



Figure 6.6.1: Rays of light entering a converging lens parallel to its axis converge at its focal point F. (Ray 2 lies on the axis of the lens.) The distance from the center of the lens to the focal point is the lens's focal length $f$. An expanded view of the path taken by ray 1 shows the perpendiculars and the angles of incidence and refraction at both surfaces.

## Definition: CONVERGING OR CONVEX LENS

The lens in which light rays that enter it parallel to its axis cross one another at a single point on the opposite side with a converging effect is called converging lens.

## Definition: FOCAL POINT F

The point at which the light rays cross is called the focal point F of the lens.

## Definition: FOCAL LENGTH $f$

The distance from the center of the lens to its focal point is called focal length $f$.



Figure 6.6.2: Sunlight focused by a converging magnifying glass can burn paper. Light rays from the sun are nearly parallel and cross at the focal point of the lens. The more powerful the lens, the closer to the lens the rays will cross.

The greater effect a lens has on light rays, the more powerful it is said to be. For example, a powerful converging lens will focus parallel light rays closer to itself and will have a smaller focal length than a weak lens. The light will also focus into a smaller and more intense spot for a more powerful lens. The **power** $P$ of a lens is defined to be the inverse of its focal length:

## Definition: POWER $P$

The power $P$ of a lens is defined to be the inverse of its focal length. In equation form, this is

$$P = \frac{1}{f}. \tag{6.6.1}$$

where $f$ is the focal length of the lens, which must be given in meters (and not cm or mm). The power of a lens $P$ has the unit diopters (D), provided that the focal length is given in meters. That is $1D = 1/m$ or $1m^{-1}$. (Note that this power (optical power, actually) is not the same as power in watts defined in "Work, Energy, and Energy Resources." It is a concept related to the effect of optical devices on light.) Optometrists prescribe common spectacles and contact lenses in units of diopters.

## Example 6.6.1: What is the Power of a Common Magnifying Glass?

Suppose you take a magnifying glass out on a sunny day and you find that it concentrates sunlight to a small spot 8.00 cm away from the lens. What are the focal length and power of the lens?

**Strategy:**

The situation here is the same as those shown in Figure 1 and Figure 2. The Sun is so far away that the Sun's rays are nearly parallel when they reach Earth. The magnifying glass is a convex (or converging) lens, focusing the nearly parallel rays of sunlight. Thus the focal length of the lens is the distance from the lens to the spot, and its power is the inverse of this distance (in m).

**Solution**

The focal length of the lens is the distance from the center of the lens to the spot, given to be 8.00 cm. Thus,

$$f = 8.00cm. \tag{6.6.2}$$

To find the power of the lens, we must first convert the focal length to meters; then, we substitute this value into the equation for power. This gives

$$P = \frac{1}{f} = \frac{1}{0.0800m} = 12.5D. \tag{6.6.3}$$

**Discussion:**

This is a relatively powerful lens. The power of a lens in diopters should not be confused with the familiar concept of power in watts. It is an unfortunate fact that the word “power” is used for two completely different concepts. If you examine a prescription for eyeglasses, you will note lens powers given in diopters. If you examine the label on a motor, you will note energy consumption rate given as a power in watts.

Figure $6.6.3$ shows a concave lens and the effect it has on rays of light that enter it parallel to its axis (the path taken by ray 2 in the figure is the axis of the lens). The concave lens is a **diverging lens**, because it causes the light rays to bend away (diverge) from its axis. In this case, the lens has been shaped so that all light rays entering it parallel to its axis appear to originate from the same point, $F$, defined to be the focal point of a diverging lens. The distance from the center of the lens to the focal point is again called the focal length $f$ of the lens. Note that the focal length and power of a diverging lens are defined to be negative. For example, if the distance to $F$ in Figure 3 is 5.00 cm, then the focal length is $f = -5.00cm$ and the power of the lens is $P = -20D$ . An expanded view of the path of one ray through the lens is shown in the figure to illustrate how the shape of the lens, together with the law of refraction, causes the ray to follow its particular path and be diverged.

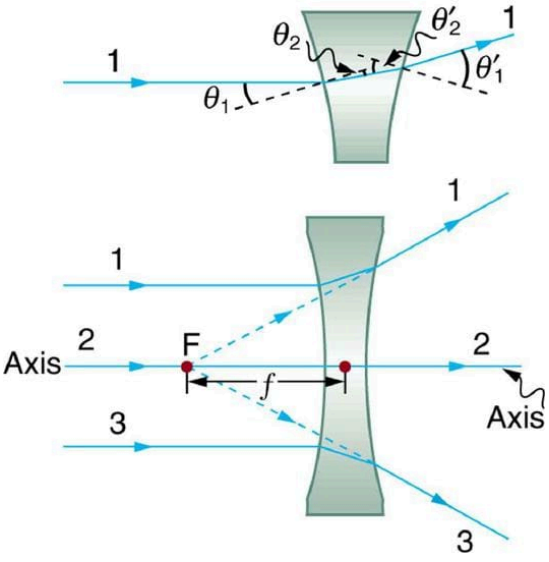



Figure $6.6.3$: Rays of light entering a diverging lens parallel to its axis are diverged, and all appear to originate at its focal point $F$. The dashed lines are not rays -- they indicate the directions from which the rays appear to come. The focal length $f$ of a diverging lens is negative. An expanded view of the path taken by ray 1 shows the perpendiculars and the angles of incidence and refraction at both surfaces.

**Definition: DIVERGING LENS**

A lens that causes the light rays to bend away from its axis is called a diverging lens.

As noted in the initial discussion of the law of refraction in "The Law of Refraction," the paths of light rays are exactly reversible. This means that the direction of the arrows could be reversed for all of the rays in Figures $6.6.1$ and Figure $6.6.3$. For example, if a point light source is placed at the focal point of a convex lens, as shown in Figure 4, parallel light rays emerge from the other side.

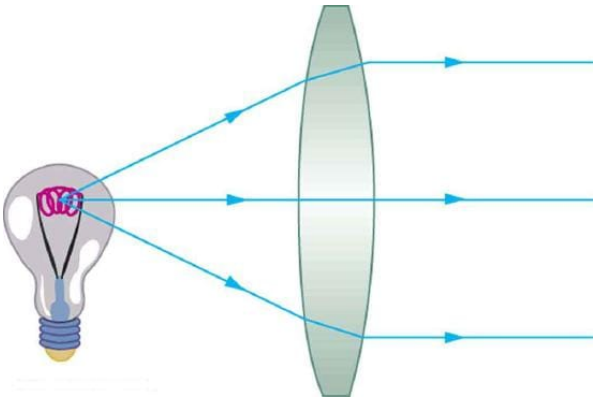


Figure 1. This technique is used in lighthouses and sometimes in traffic lights to produce a directional beam of light from a source that emits light in all directions.

## Ray Tracing and Thin Lenses

**Ray tracing** is the technique of determining or following (tracing) the paths that light rays take. For rays passing through matter, the law of refraction is used to trace the paths. Here we use ray tracing to help us understand the action of lenses in situations ranging from forming images on film to magnifying small print to correcting nearsightedness. While ray tracing for complicated lenses, such as those found in sophisticated cameras, may require computer techniques, there is a set of simple rules for tracing rays through thin lenses. A **thin lens** is defined to be one whose thickness allows rays to refract, as illustrated in Figure 6.6.1, but does not allow properties such as dispersion and aberrations. An ideal thin lens has two refracting surfaces but the lens is thin enough to assume that light rays bend only once. A thin symmetrical lens has two focal points, one on either side and both at the same distance from the lens (Figure 6.6.5). Another important characteristic of a thin lens is that light rays through its center are deflected by a negligible amount, as seen in Figure 6.6.6.

**Definition: THIN LENS**

A thin lens is defined to be one whose thickness allows rays to refract but does not allow properties such as dispersion and aberrations.

**TAKE-HOME EXPERIMENT: A VISIT TO THE OPTICIAN**

Look through your eyeglasses (or those of a friend) backward and forward and comment on whether they act like thin lenses.

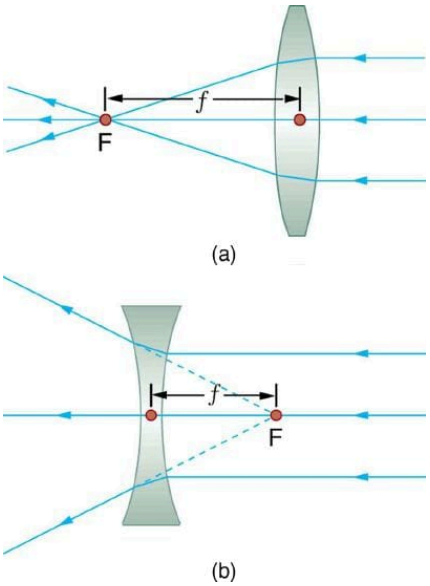



Figure 6.6.5: Thin lenses have the same focal length on either side. (a) Parallel light rays

entering a converging lens from the right cross at its focal point on the left. (b) Parallel light rays entering a diverging lens from the right seem to come from the focal point on the right.

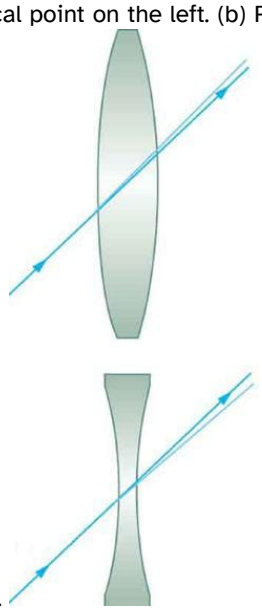


Figure $6.6.6$: The light ray through the center of a thin lens is deflected by a negligible amount and is assumed to emerge parallel to its original path (shown as a shaded line).

Using paper, pencil, and a straight edge, ray tracing can accurately describe the operation of a lens. The Ruls for Ray Tracing for thin lenses are based on the illustrations already discussed:

**RULES FOR RAY TRACING (for thin lenses)**

1. A ray entering a converging lens parallel to its axis passes through the focal point F of the lens on the other side. (See rays 1 and 3 in Figure $6.6.1$).
2. A ray entering a diverging lens parallel to its axis seems to come from the focal point F. (See rays 1 and 3 in Figure $6.6.3$ ).
3. A ray passing through the center of either a converging or a diverging lens does not change direction. (See Figure $6.6.6$, and see ray 2 in Figure $6.6.1$ and $6.6.3$).
4. A ray entering a converging lens through its focal point exits parallel to its axis. (The reverse of rays 1 and 3 in Figure $6.6.1$).
5. A ray that enters a diverging lens by heading toward the focal point on the opposite side exits parallel to the axis. (The reverse of rays 1 and 3 in Figure $6.6.3$).

## Image Formation by Thin Lenses

In some circumstances, a lens forms an obvious image, such as when a movie projector casts an image onto a screen. In other cases, the image is less obvious. Where, for example, is the image formed by eyeglasses? We will use ray tracing for thin lenses to illustrate how they form images, and we will develop equations to describe the image formation quantitatively.

Consider an object some distance away from a converging lens, as shown in Figure $6.6.7$. To find the location and size of the image formed, we trace the paths of selected light rays originating from one point on the object, in this case the top of the person's head. The figure shows three rays from the top of the object that can be traced using the ray tracing rules given above. (Rays leave this point going in many directions, but we concentrate on only a few with paths that are easy to trace.) The first ray is one that enters the lens parallel to its axis and passes through the focal point on the other side (rule 1). The second ray passes through the center of the lens without changing direction (rule 3). The third ray passes through the nearer focal point on its way into the lens and leaves the lens parallel to its axis (rule 4). The three rays cross at the same point on the other side of the lens. The image of the top of the person's head is located at this point. All rays that come from the same point on the top of the person's head are refracted in such a way as to cross at the point shown. Rays from another point on the object, such as her belt buckle, will also cross at another common point, forming a complete image, as shown. Although three rays are traced in Figure $6.6.7$, only two are necessary to locate the image. It is best to trace rays for which there are simple ray tracing rules. Before applying ray tracing to other situations, let us consider the example shown in Figure 7 in more detail.

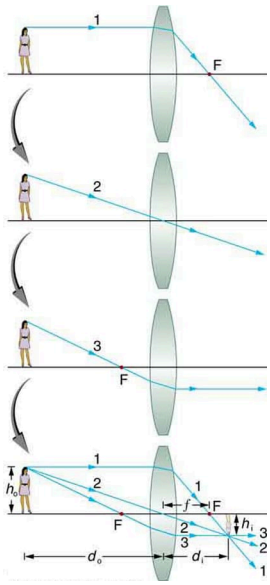



Figure $6.6.7$: Ray tracing is used to locate the image formed by a lens. Rays originating from the same point on the object are traced -- the three chosen rays each follow one of the rules for ray tracing, so that their paths are easy to determine. The image is located at the point where the rays cross. In this case, a real image -- one that can be projected on a screen -- is formed.

The image formed in Figure $6.6.7$ is a **real image**, meaning that it can be projected. That is, light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye, for example. Figure $6.6.8$ shows how such an image would be projected onto film by a camera lens. This figure also shows how a real image is projected onto the retina by the lens of an eye. Note that the image is there whether it is projected onto a screen or not.

**Definition: REAL IMAGE**

The image in which light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye is called a real image.

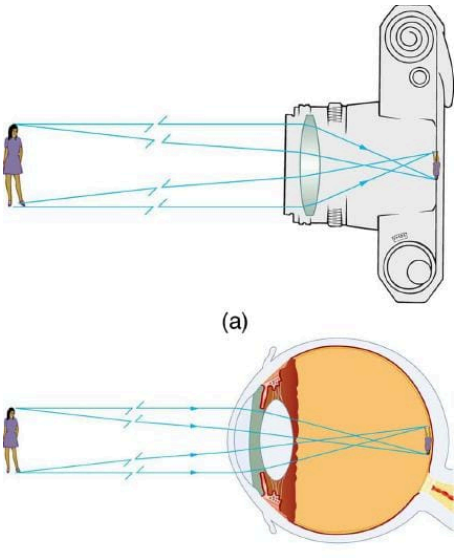



Figure $6.6.8$: Real images can be projected. (a) A real image of the person is projected onto film. (b) The converging nature of the multiple surfaces that make up the eye result in the projection of a real image on the retina.

Several important distances appear in Figure 7. We define $d_0$ to be the object of distance, the distance of an object from the center of a lens. **Image distance** $d_i$ is defined to be the distance of the image from the center of a lens. The height of the object

and height of the image are given the symbols $h_0$ and $h_i$, respectively. Images that appear upright relative to the object have heights that are positive and those that are inverted have negative heights. Using the rules of ray tracing and making a scale drawing with paper and pencil, like that in Figure 7, we can accurately describe the location and size of an image. But the real benefit of ray tracing is in visualizing how images are formed in a variety of situations. To obtain numerical information, we use a pair of equations that can be derived from a geometric analysis of ray tracing for thin lenses. The **thin lens equations** are

$$\frac{1}{d_0} + \frac{1}{d_i} = \frac{1}{f} \tag{6.6.4}$$

and

$$\frac{h_i}{h_0} = -\frac{d_i}{d_0} = m. \tag{6.6.5}$$

We define the ratio of image height to object height ($h_i/h_0$) to be the **magnification** $m$. (The minus sign in the equation above will be discussed shortly.) The thin lens equations are broadly applicable to all situations involving thin lenses (and "thin" mirrors, as we will see later). We will explore many features of image formation in the following worked examples.

## Definition: IMAGE DISTANCE

The distance of the image from the center of the lens is called image distance.

## Definition: THIN LENS EQUATIONS AND MAGNIFICATION

Label '25.7.1' multiply defined

Label '25.7.2' multiply defined

## Example 6.6.2: Finding the Image of a Light Bulb Filament by Ray Tracing and by the Thin Lens Equations

A clear glass light bulb is placed 0.750 m from a convex lens having a 0.500 m focal length, as shown in the figure. Use ray tracing to get an approximate location for the image. Then use the thin lens equations to calculate

a. the location of the image and
b. its magnification.

Verify that ray tracing and the thin lens equations produce consistent results.

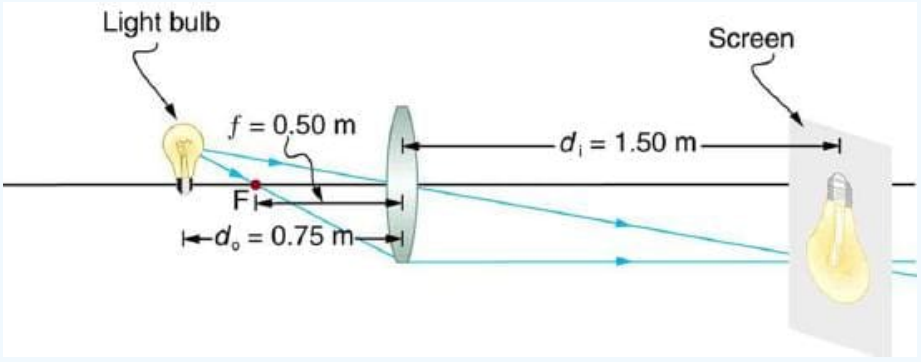



Figure 6.6.9: A light bulb placed 0.750 m from a lens having a 0.500 m focal length produces a real image on a poster board as discussed in the example above. Ray tracing predicts the image location and size.

**Strategy and Concept**

Since the object is placed farther away from a converging lens than the focal length of the lens, this situation is analogous to those illustrated in Figure 7 and Figure 8. Ray tracing to scale should produce similar results for $d_i$. Numerical solutions for $d_i$ and $m$ can be obtained using the thin lens equations, noting that $d_o = 0.750m$ and $f = 0.500m$.

**Solutions (Ray Tracing)**

The ray tracing to scale in Figure 9 shows two rays from a point on the bulb's filament crossing about 1.50 m on the far side of the lens. Thus the image distance $d_i$ is about 1.50 m. Similarly, the image height based on ray tracing is greater than the object height by about a factor of 2, and the image is inverted. Thus $m$ is about –2. The minus sign indicates that the image is inverted.

The thin lens equations can be used to find $d_i$ from the given information:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}. \tag{6.6.6}$$

Rearranging to isolate $d_i$ gives

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}. \tag{6.6.7}$$

Entering known quantities gives a value for $1/d_i$:

$$\frac{1}{d_i} = \frac{1}{0.500m} - \frac{1}{0.750m} = \frac{0.667}{m}. \tag{6.6.8}$$

This must be inverted to find $d_i$:

$$d_i = \frac{m}{0.667} = 1.50m. \tag{6.6.9}$$

Note that another way to find $d_i$ is to rearrange equation:

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}. \tag{6.6.10}$$

This yields the equation for the image distance as:

$$d_i = \frac{fd_o}{d_o - f}. \tag{6.6.11}$$

Note that there is no inverting here.

The thin lens equations can be used to find the magnification $m$, since both $d_i$ and $d_o$ are known. Entering their values gives

$$m = -\frac{d_i}{d_o} = -\frac{1.50m}{0.750m} = -2.00. \tag{6.6.12}$$

**Discussion**

Note that the minus sign causes the magnification to be negative when the image is inverted. Ray tracing and the use of the thin lens equations produce consistent results. The thin lens equations give the most precise results, being limited only by the accuracy of the given information. Ray tracing is limited by the accuracy with which you can draw, but it is highly useful both conceptually and visually.

Real images, such as the one considered in the previous example, are formed by converging lenses whenever an object is farther from the lens than its focal length. This is true for movie projectors, cameras, and the eye. We shall refer to these as *case 1* images. A case 1 image is formed when $d_o > f$ and $f$ is positive, as in Figure $6.6.10a$. (A summary of the three cases or types of image formation appears at the end of this section.)

A different type of image is formed when an object, such as a person's face, is held close to a convex lens. The image is upright and larger than the object, as seen in Figure $6.6.10b$, and so the lens is called a magnifier. If you slowly pull the magnifier away from the face, you will see that the magnification steadily increases until the image begins to blur. Pulling the magnifier even farther away produces an inverted image as seen in Figure $6.6.10a$. The distance at which the image blurs, and beyond which it inverts, is the focal length of the lens. To use a convex lens as a magnifier, the object must be closer to the converging lens than its focal length. This is called a *case 2* image. A case 2 image is formed when $d_o < f$ and $f$ is positive.



(a)



(b)

Figure 6.6.10: (a) When a converging lens is held farther away from the face than the lens's focal length, an inverted image is formed. This is a case 1 image. Note that the image is in focus but the face is not, because the image is much closer to the camera taking this photograph than the face. (credit: DaMongMan, Flickr) (b) A magnified image of a face is produced by placing it closer to the converging lens than its focal length. This is a case 2 image. (credit: Casey Fleser, Flickr)

Figure 6.6.11 uses ray tracing to show how an image is formed when an object is held closer to a converging lens than its focal length. Rays coming from a common point on the object continue to diverge after passing through the lens, but all appear to originate from a point at the location of the image. The image is on the same side of the lens as the object and is farther away from the lens than the object. This image, like all case 2 images, cannot be projected and, hence, is called a **virtual image**. Light rays only appear to originate at a virtual image; they do not actually pass through that location in space. A screen placed at the location of a virtual image will receive only diffuse light from the object, not focused rays from the lens. Additionally, a screen placed on the opposite side of the lens will receive rays that are still diverging, and so no image will be projected on it. We can see the magnified image with our eyes, because the lens of the eye converges the rays into a real image projected on our retina. Finally, we note that a virtual image is upright and larger than the object, meaning that the magnification is positive and greater than 1.

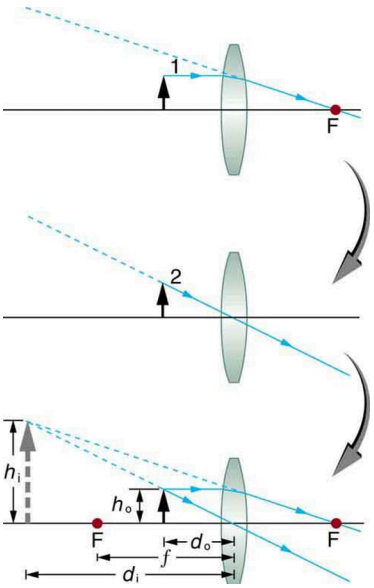



Figure 6.6.11: Ray tracing predicts the image location and size for an object held closer to a converging lens than its focal length. Ray 1 enters parallel to the axis and exits through the focal point on the opposite side, while ray 2 passes through the center of the lens without changing path. The two rays continue to diverge on the other side of the lens, but both appear to come from a common point, locating the upright, magnified, virtual image. This is a case 2 image.

**Definition: VIRTUAL IMAGE**

An image that is on the same side of the lens as the object and cannot be projected on a screen is called a virtual image.

**Example 6.6.3: Image Produced by a Magnifying Glass**

Suppose the book in Figure 6.6.11*a* is held 7.50 cm from a convex lens of focal length 10.0 cm, such as a typical magnifying glass might have. What magnification is produced?

**Strategy and Concept**

We are given that $d_o = 7.50cm$ and $f = 10.0cm$, so we have a situation where the object is placed closer to the lens than its focal length. We therefore expect to get a case 2 virtual image with a positive magnification that is greater than 1. Ray tracing produces an image like that shown in Figure 11 but we will use the thin lens equations to get numerical solutions in this example.

**Solution**

To find the magnification $m$, we try to use magnification equation, $m = -d_i/d_o$. We do not have a value for $d_i$, so that we must first find the location of the image using lens equation. (The procedure is the same as followed in the preceding example, where $d_o$ and $f$ were known.) Rearranging the magnification equation to isolate $d_i$ gives

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}. \tag{6.6.13}$$

Entering known values, we obtain a value for $1/d_i$:

$$\frac{1}{d_i} = \frac{1}{10.0cm} - \frac{1}{7.50cm} = \frac{-0.0333}{cm}. \tag{6.6.14}$$

This must be inverted to find $d_i$:

$$d_i = -\frac{cm}{0.0333} = -30.0cm. \tag{6.6.15}$$

Now the thin lens equation can be used to find the magnification $m$, since both $d_i$ and $d_o$ are known. Entering their values gives

$$m = -\frac{d_i}{d_o} = -\frac{-30.0cm}{7.50cm} = 4.00. \tag{6.6.16}$$

**Discussion**

A number of results in this example are true of all case 2 images, as well as being consistent with Figure 11. Magnification is indeed positive (as predicted), meaning the image is upright. The magnification is also greater than 1, meaning that the image is larger than the object—in this case, by a factor of 4. Note that the image distance is negative. This means the image is on the same side of the lens as the object. Thus the image cannot be projected and is virtual. (Negative values of $d_i$ occur for virtual images.) The image is farther from the lens than the object, since the image distance is greater in magnitude than the object distance. The location of the image is not obvious when you look through a magnifier. In fact, since the image is bigger than the object, you may think the image is closer than the object. But the image is farther away, a fact that is useful in correcting farsightedness, as we shall see in a later section.

A third type of image is formed by a diverging or concave lens. Try looking through eyeglasses meant to correct nearsightedness. (Figure 6.6.12). You will see an image that is upright but smaller than the object. This means that the magnification is positive but less than 1. The ray diagram in Figure 13 shows that the image is on the same side of the lens as the object and, hence, cannot be projected—it is a virtual image. Note that the image is closer to the lens than the object. This is a *case 3* image, formed for any object by a negative focal length or diverging lens.



Figure 6.6.12: A car viewed through a concave or diverging lens looks upright. This is a case 3 image. (credit: Daniel Oines, Flickr)

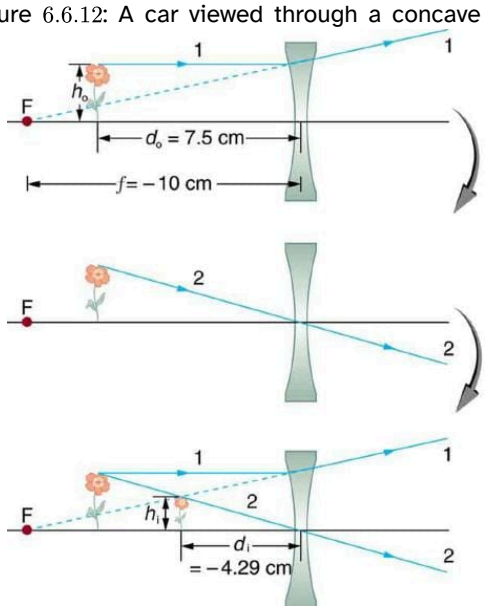



Figure 6.6.13: Ray tracing predicts the image location and size for a concave or diverging lens. Ray 1 enters parallel to the axis and is bent so that it appears to originate from the focal point. Ray 2 passes through the center of the lens without changing path. The two rays appear to come from a common point, locating the upright image. This is a case 3 image, which is closer to the lens than the object and smaller in height.

## Example 6.6.4: Image Produced by a Concave Lens

Suppose an object such as a book page is held 7.50 cm from a concave lens of focal length –10.0 cm. Such a lens could be used in eyeglasses to correct pronounced nearsightedness. What magnification is produced?

**Strategy and Concept**

This example is identical to the preceding one, except that the focal length is negative for a concave or diverging lens. The method of solution is thus the same, but the results are different in important ways.

**Solution**

To find the magnification $m$, we must first find the image distance $d_i$ using thin lens equation

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}, \tag{6.6.17}$$

or its alternative rearrangement

$$d_i = \frac{f d_o}{d_o - f}. \tag{6.6.18}$$

We are given that $f = -10.00cm$ and $d_o = 7.50cm$. Entering these yields a value for $1/d_i$:

$$\frac{1}{d_i} = \frac{1}{-10.0cm} - \frac{1}{7.50cm} = \frac{-0.2333}{cm}. \tag{6.6.19}$$

This must be inverted to find $d_i$:

$$d_i = -\frac{cm}{0.2333} = -4.29cm. \tag{6.6.20}$$

Or

$$d_i = \frac{(7.5)(-10)}{(7.5 - (-10))} = -75/17.5 = -4.29cm. \tag{6.6.21}$$

Now the magnification equation can be used to find the magnification $m$, since both $d_i$ and $d_o$ are known. Entering their values gives

$$m = -\frac{d_i}{d_o} = -\frac{-4.29cm}{7.50cm} = 0.571. \tag{6.6.22}$$

**Discussion:**

A number of results in this example are true of all case 3 images, as well as being consistent with Figure 13. Magnification is positive (as predicted), meaning the image is upright. The magnification is also less than 1, meaning the image is smaller than the object -- in this case, a little over half its size. The image distance is negative, meaning the image is on the same side of the lens as the object. (The image is virtual.) The image is closer to the lens than the object, since the image distance is smaller in magnitude than the object distance. The location of the image is not obvious when you look through a concave lens. In fact, since the image is smaller than the object, you may think it is farther away. But the image is closer than the object, a fact that is useful in correcting nearsightedness, as we shall see in a later section.

The table summarizes the three types of images formed by single thin lenses. These are referred to as case 1, 2, and 3 images. Convex (converging) lenses can form either real or virtual images (cases 1 and 2, respectively), whereas concave (diverging) lenses can form only virtual images (always case 3). Real images are always inverted, but they can be either larger or smaller than the object. For example, a slide projector forms an image larger than the slide, whereas a camera makes an image smaller than the object being photographed. Virtual images are always upright and cannot be projected. Virtual images are larger than the object only in case 2, where a convex lens is used. The virtual image produced by a concave lens is always smaller than the object -- a case 3 image. We can see and photograph virtual images only by using an additional lens to form a real image.

*Table 6.6.1: Three Types of Images Formed By Thin Lenses*

| Type | Formed When | Image type | $d_i$ | m |
|---|---|---|---|---|
| Case 1 | positive **f** $d_0 > f$ | real | positive | negative |
| Case 2 | positive **f** $d_0 < f$ | virtual | negative | positive **m > 1** |
| Case 3 | negative **f** | virtual | negative | positive |

In "Image Formation by Mirrors," we shall see that mirrors can form exactly the same types of images as lenses.

## TAKE-HOME EXPERIMENT: CONCENTRATING SUNLIGHT

Find several lenses and determine whether they are converging or diverging. In general those that are thicker near the edges are diverging and those that are thicker near the center are converging. On a bright sunny day take the converging lenses outside and try focusing the sunlight onto a piece of paper. Determine the focal lengths of the lenses. Be careful because the paper may start to burn, depending on the type of lens you have selected.

### Problem-Solving Strategies for Lenses

- Step 1. Examine the situation to determine that image formation by a lens is involved.
- Step 2. Determine whether ray tracing, the thin lens equations, or both are to be employed. A sketch is very useful even if ray tracing is not specifically required by the problem. Write symbols and values on the sketch.
- Step 3. Identify exactly what needs to be determined in the problem (identify the unknowns).
- Step 4. Make alist of what is given or can be inferred from the problem as stated (identify the knowns). It is helpful to determine whether the situation involves a case 1, 2, or 3 image. While these are just names for types of images, they have certain characteristics (given in the table) that can be of great use in solving problems.
- Step 5. If ray tracing is required, use the ray tracing rules listed near the beginning of this section.
- Step 6. Most quantitative problems require the use of the thin lens equations. These are solved in the usual manner by substituting knowns and solving for unknowns. Several worked examples serve as guides.
- Step 7. Check to see if the answer is reasonable: Does it make sense? If you have identified the type of image (case 1, 2, or 3), you should assess whether your answer is consistent with the type of image, magnification, and so on.

### MISCONCEPTION ALERT:

We do not realize that light rays are coming from every part of the object, passing through every part of the lens, and all can be used to form the final image. We generally feel the entire lens, or mirror, is needed to form an image. Actually, half a lens will form the same, though a fainter, image.

## Summary

- Light rays entering a converging lens parallel to its axis cross one another at a single point on the opposite side.
- For a converging lens, the focal point is the point at which converging light rays cross; for a diverging lens, the focal point is the point from which diverging light rays appear to originate.
- The distance from the center of the lens to its focal point is called the focal length $f$.
- Power $P$ of a lens is defined to be the inverse of its focal length, $P = \frac{1}{f}$.
- A lens that causes the light rays to bend away from its axis is called a diverging lens.
- Ray tracing is the technique of graphically determining the paths that light rays take.
- The image in which light rays from one point on the object actually cross at the location of the image and can be projected onto a screen, a piece of film, or the retina of an eye is called a real image.
- Thin lens equations are $\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$ and $\frac{h_i}{h_o} = m$ (magnification).
- The distance of the image from the center of the lens is called image distance.
- An image that is on the same side of the lens as the object and cannot be projected on a screen is called a virtual image.

## Glossary

**converging lens**
: a convex lens in which light rays that enter it parallel to its axis converge at a single point on the opposite side

**diverging lens**
: a concave lens in which light rays that enter it parallel to its axis bend away (diverge) from its axis

**focal point**
: for a converging lens or mirror, the point at which converging light rays cross; for a diverging lens or mirror, the point from which diverging light rays appear to originate

**focal length**
: distance from the center of a lens or curved mirror to its focal point

**magnification**
: ratio of image height to object height

**power**
: inverse of focal length

**real image**
: image that can be projected

**virtual image**
: image that cannot be projected

This page titled 6.6: Image Formation by Lenses is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 25.6: Image Formation by Lenses by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 6.7: Image Formation by Mirrors

## Learning Objectives

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

- Illustrate image formation in a flat mirror.
- Explain with ray diagrams the formation of an image using spherical mirrors.
- Determine focal length and magnification given radius of curvature, distance of object and image.

We only have to look as far as the nearest bathroom to find an example of an image formed by a mirror. Images in flat mirrors are the same size as the object and are located behind the mirror. Like lenses, mirrors can form a variety of images. For example, dental mirrors may produce a magnified image, just as makeup mirrors do. Security mirrors in shops, on the other hand, form images that are smaller than the object. We will use the law of reflection to understand how mirrors form images, and we will find that mirror images are analogous to those formed by lenses.

Figure $6.7.1$ helps illustrate how a flat mirror forms an image. Two rays are shown emerging from the same point, striking the mirror, and being reflected into the observer's eye. The rays can diverge slightly, and both still get into the eye. If the rays are extrapolated backward, they seem to originate from a common point behind the mirror, locating the image. (The paths of the reflected rays into the eye are the same as if they had come directly from that point behind the mirror.) Using the law of reflection -- the angle of reflection equals the angle of incidence -- we can see that the image and object are the same distance from the mirror. This is a virtual image, since it cannot be projected -- the rays only appear to originate from a common point behind the mirror. Obviously, if you walk behind the mirror, you cannot see the image, since the rays do not go there. But in front of the mirror, the rays behave exactly as if they had come from behind the mirror, so that is where the image is situated.

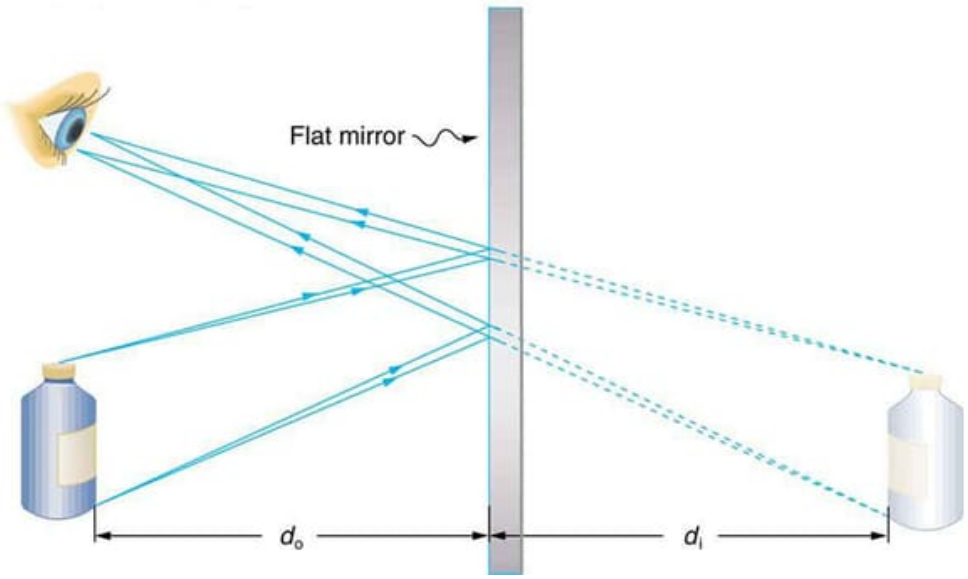



Figure $6.7.1$: Two sets of rays from common points on an object are reflected by a flat mirror into the eye of an observer. The reflected rays seem to originate from behind the mirror, locating the virtual image.

Now let us consider the focal length of a mirror -- for example, the concave spherical mirrors in Figure $6.7.2$. Rays of light that strike the surface follow the law of reflection. For a mirror that is large compared with its radius of curvature, as in Figure $6.7.2a$, we see that the reflected rays do not cross at the same point, and the mirror does not have a well-defined focal point. If the mirror had the shape of a parabola, the rays would all cross at a single point, and the mirror would have a well-defined focal point. But parabolic mirrors are much more expensive to make than spherical mirrors. The solution is to use a mirror that is small compared with its radius of curvature, as shown in Figure $6.7.2b$. (This is the mirror equivalent of the thin lens approximation.) To a very good approximation, this mirror has a well-defined focal point at F that is the focal distance $f$ from the center of the mirror. The focal

length $f$ of a concave mirror is positive, since it is a converging mirror.

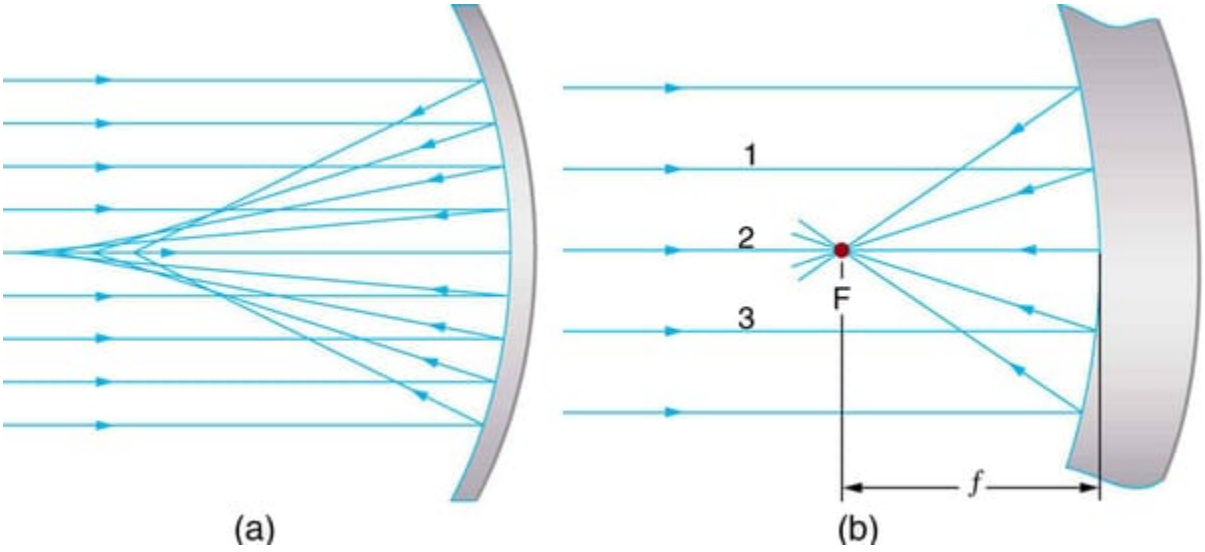



Figure $6.7.2$: (a) Parallel rays reflected from a large spherical mirror do not all cross at a common point. (b) If a spherical mirror is small compared with its radius of curvature, parallel rays are focused to a common point. The distance of the focal point from the center of the mirror is its focal length $f$. Since this mirror is converging, it has a positive focal length.

Just as for lenses, the shorter the focal length, the more powerful the mirror; thus, $P = 1/f$ for a mirror, too. A more strongly curved mirror has a shorter focal length and a greater power. Using the law of reflection and some simple trigonometry, it can be shown that the focal length is half the radius of curvature, or

$$f = \frac{R}{2}, \tag{6.7.1}$$

where $R$ is the radius of curvature of a spherical mirror. The smaller the radius of curvature, the smaller the focal length and, thus, the more powerful the mirror.

The convex mirror shown in Figure $6.7.3$ also has a focal point. Parallel rays of light reflected from the mirror seem to originate from the point F at the focal distance $f$ behind the mirror. The focal length and power of a convex mirror are negative, since it is a diverging mirror.

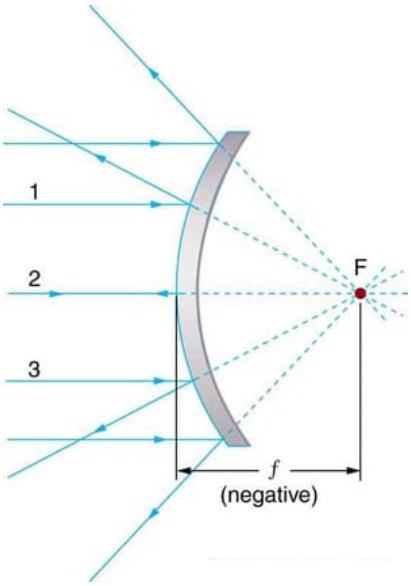



Figure $6.7.3$: Parallel rays of light reflected from a convex spherical mirror (small in size compared with its radius of curvature) seem to originate from a well-defined focal point at the focal distance $f$ behind the mirror. Convex mirrors diverge light rays and, thus, have a negative focal length.

Ray tracing is as useful for mirrors as for lenses. The rules for ray tracing for mirrors are based on the illustrations just discussed:

### Ray Tracing Rules

1. A ray approaching a concave converging mirror parallel to its axis is reflected through the focal point F of the mirror on the same side. (See rays 1 and 3 in Figure 6.7.2)
2. A ray approaching a convex diverging mirror parallel to its axis is reflected so that it seems to come from the focal point F behind the mirror. (See rays 1 and 3 in Figure 6.7.3).
3. Any ray striking the center of a mirror is followed by applying the law of reflection; it makes the same angle with the axis when leaving as when approaching. (See ray 2 in Figures 6.7.2 and 6.7.3).
4. A ray approaching a concave converging mirror through its focal point is reflected parallel to its axis. (The reverse of rays 1 and 3 in Figure 6.7.3).
5. A ray approaching a convex diverging mirror by heading toward its focal point on the opposite side is reflected parallel to the axis. (The reverse of rays 1 and 3 in Figure 6.7.3).

We will use ray tracing to illustrate how images are formed by mirrors, and we can use ray tracing quantitatively to obtain numerical information. But since we assume each mirror is small compared with its radius of curvature, we can use the thin lens equations for mirrors just as we did for lenses.

Consider the situation shown in Figure 6.7.4, concave spherical mirror reflection, in which an object is placed farther from a concave (converging) mirror than its focal length. That is, $f$ is positive and $d_o > f$, so that we may expect an image similar to the case 1 real image formed by a converging lens. Ray tracing in Figure 4 shows that the rays from a common point on the object all cross at a point on the same side of the mirror as the object. Thus a real image can be projected onto a screen placed at this location. The image distance is positive, and the image is inverted, so its magnification is negative. This is a *case 1 image for mirrors*. It differs from the case 1 image for lenses only in that the image is on the same side of the mirror as the object. It is otherwise identical.

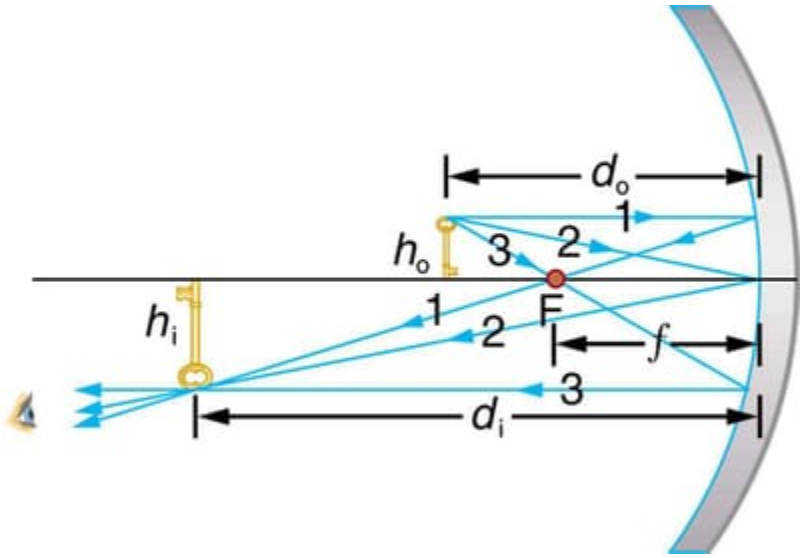



Figure 6.7.4: A case 1 image for a mirror. An object is farther from the converging mirror than its focal length. Rays from a common point on the object are traced using the rules in the text. Ray 1 approaches parallel to the axis, ray 2 strikes the center of the mirror, and ray 3 goes through the focal point on the way toward the mirror. All three rays cross at the same point after being reflected, locating the inverted real image. Although three rays are shown, only two of the three are needed to locate the image and determine its height.

### Example 6.7.1: A Concave Reflector

Electric room heaters use a concave mirror to reflect infrared (IR) radiation from hot coils. Note that IR follows the same law of reflection as visible light. Given that the mirror has a radius of curvature of 50.0 cm and produces an image of the coils 3.00 m away from the mirror, where are the coils?

**Strategy and Concept**

We are given that the concave mirror projects a real image of the coils at an image distance $d_i = 3.00m$. The coils are the object, and we are asked to find their location -- that is, to find the object distance $d_o$. We are also given the radius of curvature of the mirror, so that its focal length is $f = R/2 = 25.0cm$ (positive since the mirror is concave or converging). Assuming the mirror is small compared with its radius of curvature, we can use the thin lens equations, to solve this problem.

**Solution**

Since $d_i$ and $f$ are known, thin lens equation can be used to find $d_o$:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}.$$

Rearranging to isolate $d_o$ gives

$$\frac{1}{d_o} = \frac{1}{f} - \frac{1}{d_i}.$$

Entering known quantities gives a value for $1/d_o$:

$$\frac{1}{d_o} = \frac{1}{0.250m} - \frac{1}{3.00m} = \frac{3.667}{m}.$$

This must be inverted to find $d_o$:

$$d_o = \frac{1m}{3.667} = 27.3cm.$$

**Discussion**

Note that the object (the filament) is farther from the mirror than the mirror's focal length. This is a case 1 image ($d_o > f$ and $f$ positive), consistent with the fact that a real image is formed. You will get the most concentrated thermal energy directly in front of the mirror and 3.00 m away from it. Generally, this is not desirable, since it could cause burns. Usually, you want the rays to emerge parallel, and this is accomplished by having the filament at the focal point of the mirror.

Note that the filament here is not much farther from the mirror than its focal length and that the image produced is considerably farther away. This is exactly analogous to a slide projector. Placing a slide only slightly farther away from the projector lens than its focal length produces an image significantly farther away. As the object gets closer to the focal distance, the image gets farther away. In fact, as the object distance approaches the focal length, the image distance approaches infinity and the rays are sent out parallel to one another.

## Example 6.7.2: Solar Electric Generating System

One of the solar technologies used today for generating electricity is a device (called a parabolic trough or concentrating collector) that concentrates the sunlight onto a blackened pipe that contains a fluid. This heated fluid is pumped to a heat exchanger, where its heat energy is transferred to another system that is used to generate steam -- and so generate electricity through a conventional steam cycle. Figure 6.7.5 shows such a working system in southern California. Concave mirrors are used to concentrate the sunlight onto the pipe. The mirror has the approximate shape of a section of a cylinder. For the problem, assume that the mirror is exactly one-quarter of a full cylinder.

a. If we wish to place the fluid-carrying pipe 40.0 cm from the concave mirror at the mirror's focal point, what will be the radius of curvature of the mirror?
b. Per meter of pipe, what will be the amount of sunlight concentrated onto the pipe, assuming the insolation (incident solar radiation) is $0.900kW/m^2$?
c. If the fluid-carrying pipe has a 2.00-cm diameter, what will be the temperature increase of the fluid per meter of pipe over a period of one minute? Assume all the solar radiation incident on the reflector is absorbed by the pipe, and that the fluid is mineral oil.

**Strategy**

To solve an *Integrated Concept Problem* we must first identify the physical principles involved. Part (a) is related to the current topic. Part (b) involves a little math, primarily geometry. Part (c) requires an understanding of heat and density.

**Solution**

**(a)** To a good approximation for a concave or semi-spherical surface, the point where the parallel rays from the sun converge will be at the focal point, so $R = 2f = 80.0cm$.

**(b)** The insolation is $900W/m^2$. We must find the cross-sectional area $A$ of the concave mirror, since the power delivered is $900W/m^2 \times A$. The mirror in this case is a quarter-section of a cylinder, so the area for a length $L$ of the mirror is $A = \frac{1}{4}(2\pi R)L$. The area for a length of 1.00 m is then

$$\begin{aligned} A &= \frac{\pi}{2} R\,(1.00m) \\ &= \frac{(3.14)}{2}\,(0.800m)\,(1.00m) \\ &= 1.26m^2. \end{aligned}$$

The insolation on the 1.00-m length of pipe is then

$$\left(9.00 \times 10^2 \frac{W}{m^2}\right)\left(1.26m^2\right) = 1130W.$$

**(c)** The increase in temperature is given by $Q = mc\Delta T$. the mass $m$ of the mineral oil in the one-meter section of pipe is

$$\begin{aligned} m &= \rho V \\ &= \rho\pi\left(\frac{d}{2}\right)^2 (1.00m) \\ &= \left(8.00 \times 10^2 kg/m^3\right)(3.14)\,(0.0100m)^2\,(1.00m) \\ &= 0.251kg. \end{aligned}$$

Therefore, the increase in temperature in one minute is

$$\begin{aligned} \Delta T &= Q/mc \\ &= \frac{(1130W)(60.0s)}{(0.251kg)(1670J \cdot kg/^{\circ}C)} \\ &= 162^{\circ}. \end{aligned}$$

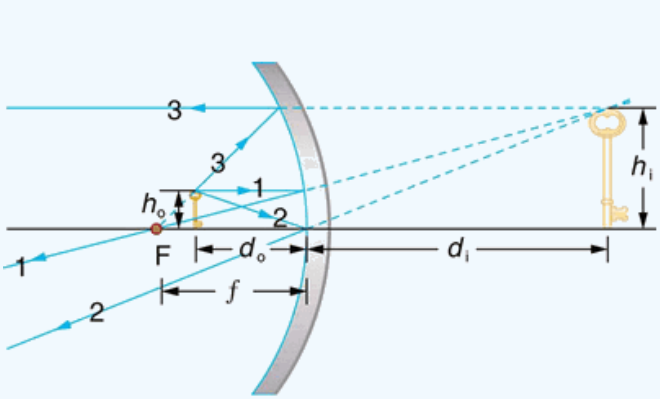
**Discussion (c)**

An array of such pipes in the California desert can provide a thermal output of 250 MW on a sunny day, with fluids reaching temperatures as high as $400^{\circ}$ We are considering only one meter of pipe here, and ignoring heat losses along the pipe.



Figure 6.7.5: Parabolic trough collectors are used to generate electricity in southern California. (credit: kjkolb, Wikimedia Commons)

What happens if an object is closer to a concave mirror than its focal length? This is analogous to a case 2 image for lenses ($d_o < f$ and $f$ positive), which is a magnifier. In fact, this is how makeup mirrors act as magnifiers. Figure 6.7.6$a$ uses ray tracing to locate the image of an object placed close to a concave mirror. Rays from a common point on the object are reflected in such a manner that they appear to be coming from behind the mirror, meaning that the image is virtual and cannot be projected. As with a magnifying glass, the image is upright and larger than the object. This is a *case 2 image for mirrors* and is exactly analogous to that for lenses.




(a)

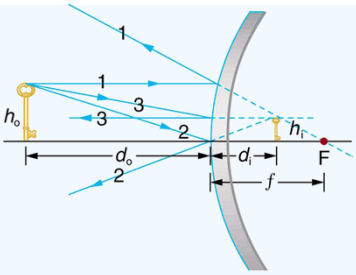



(b)

Figure 6.7.6: (a) Case 2 images for mirrors are formed when a converging mirror has an object closer to it than its focal length. Ray 1 approaches parallel to the axis, ray 2 strikes the center of the mirror, and ray 3 approaches the mirror as if it came from the focal point. (b) A magnifying mirror showing the reflection. (credit: Mike Melrose, Flickr)

All three rays appear to originate from the same point after being reflected, locating the upright virtual image behind the mirror and showing it to be larger than the object. (b) Makeup mirrors are perhaps the most common use of a concave mirror to produce a larger, upright image.

A convex mirror is a diverging mirror ($f$ is negative) and forms only one type of image. It is a *case 3 image* -- one that is upright and smaller than the object, just as for diverging lenses. Figure 6.7.7$a$ uses ray tracing to illustrate the location and size of the case 3 image for mirrors. Since the image is behind the mirror, it cannot be projected and is thus a virtual image. It is also seen to be smaller than the object.




(a)



(b)

Figure 6.7.7: uses ray tracing to illustrate the location and size of the case 3 image for mirrors. Since the image is behind the mirror, it cannot be projected and is thus a virtual image. It is also seen to be smaller than the object.

## Example 6.7.2: Image in a Convex Mirror

A keratometer is a device used to measure the curvature of the cornea, particularly for fitting contact lenses. Light is reflected from the cornea, which acts like a convex mirror, and the keratometer measures the magnification of the image. The smaller the magnification, the smaller the radius of curvature of the cornea. If the light source is 12.0 cm from the cornea and the image's magnification is 0.0320, what is the cornea's radius of curvature?

**Strategy**

If we can find the focal length of the convex mirror formed by the cornea, we can find its radius of curvature (the radius of curvature is twice the focal length of a spherical mirror). We are given that the object distance is $d_o = 12.0cm$ and that $m = 0.0320$. We first solve for the image distance $d_i$, and then for $f$.

**Solution**

$m = -d_i/d_o$. Solving this expression for $d_i$ gives

$$d_i = -md_o.$$

Entering known values yields

$$d_i = -(0.0320)(12.0cm) = -0.384cm.$$

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \tag{6.7.2}$$

Substituting known values,

$$\frac{1}{f} = \frac{1}{12.0cm} + \frac{1}{-0.384cm} = \frac{-2.52}{cm}.$$

This must be inverted to find $f$.

$$f = \frac{cm}{-2.52} = -0.400cm.$$

The radius of curvature is twice the focal length, so that

$$R = 2|f| = 0.800cm.$$

**Discussion:**

Although the focal length $f$ of a convex mirror is defined to be negative, we take the absolute value to give us a positive value for $R$.

The radius of curvature found here is reasonable for a cornea. The distance from cornea to retina in an adult eye is about 2.0 cm. In practice, many corneas are not spherical, complicating the job of fitting contact lenses. Note that the image distance here is negative, consistent with the fact that the image is behind the mirror, where it cannot be projected. In this section's Problems and Exercises, you will show that for a fixed object distance, the smaller the radius of curvature, the smaller the magnification.

The three types of images formed by mirrors (cases 1, 2, and 3) are exactly analogous to those formed by lenses, as summarized in the table at the end of "Image Formation by Lenses." It is easiest to concentrate on only three types of images -- then remember that concave mirrors act like convex lenses, whereas convex mirrors act like concave lenses.

### TAKE-HOME EXPERIMENT: CONCAVE MIRRORS CLOSE TO HOME

Find a flashlight and identify the curved mirror used in it. Find another flashlight and shine the first flashlight onto the second one, which is turned off. Estimate the focal length of the mirror. You might try shining a flashlight on the curved mirror behind the headlight of a car, keeping the headlight switched off, and determine its focal length.

### Problem-Solving Strategy for Mirrors

- Step 1. Examine the situation to determine that image formation by a mirror is involved.
- Step 2. Refer to the "Problem-Solving Strategies for Lenses." The same strategies are valid for mirrors as for lenses with one qualification-- use the ray tracing rules for mirrors listed earlier in this section.

## Summary

- The characteristics of an image formed by a flat mirror are: (a) The image and object are the same distance from the mirror, (b) The image is a virtual image, and (c) The image is situated behind the mirror.
- Image length is half the radius of curvature.

$$f = \frac{R}{2}$$

- A convex mirror is a diverging mirror and forms only one type of image, namely a virtual image.

## Glossary

**converging mirror**
: a concave mirror in which light rays that strike it parallel to its axis converge at one or more points along the axis

**diverging mirror**
: a convex mirror in which light rays that strike it parallel to its axis bend away (diverge) from its axis

**law of reflection**
: angle of reflection equals the angle of incidence

---

This page titled 6.7: Image Formation by Mirrors is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 25.7: Image Formation by Mirrors by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7: MODULE 6A- ELECTRICITY, MAGNETISM


7.1: Prelude to Electric Charge and Electric Field

7.2: Static Electricity and Charge - Conservation of Charge

7.3: Conductors and Insulators

7.4: Coulomb's Law

7.5: Electric Field- Concept of a Field Revisited

7.7: Electric Potential Energy- Potential Difference

7.8: Prelude to Electric Current, Resistance, and Ohm's Law

7.9: Current

7.10: Ohm’s Law - Resistance and Simple Circuits

7.11: Resistance and Resistivity

7.12: 20.4 Electric Power and Energy

7.13: Alternating Current versus Direct Current

7.14: Electric Hazards and the Human Body

7.15: Prelude to Circuits and DC Instruments

7.16: Resistors in Series and Parallel

7.17: Prelude to Magnetism

7.18: Magnets

7.19: Ferromagnets and Electromagnets

7.20: Magnetic Fields and Magnetic Field Lines

7.21: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies

7.06: Introduction to Electric Potential and Electric Energy



This page titled 7: Module 6a- Electricity, Magnetism is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Arunava Roy.

# 7.1: Prelude to Electric Charge and Electric Field

The image of American politician and scientist Benjamin Franklin (1706–1790) flying a kite in a thunderstorm is familiar to every schoolchild. (Figure 7.1.1) In this experiment, Franklin demonstrated a connection between lightning and **static electricity**. Sparks were drawn from a key hung on a kite string during an electrical storm. These sparks were like those produced by static electricity, such as the spark that jumps from your finger to a metal doorknob after you walk across a wool carpet. What Franklin demonstrated in his dangerous experiment was a connection between phenomena on two different scales: one the grand power of an electrical storm, the other an effect of more human proportions. Connections like this one reveal the underlying unity of the laws of nature, an aspect we humans find particularly appealing.



Figure 7.1.1: Static electricity from this plastic slide causes the child's hair to stand on end. The sliding motion stripped electrons away from the child's body, leaving an excess of positive charges, which repel each other along each strand of hair. (credit: Ken Bosma/ Wikimedia Commons)

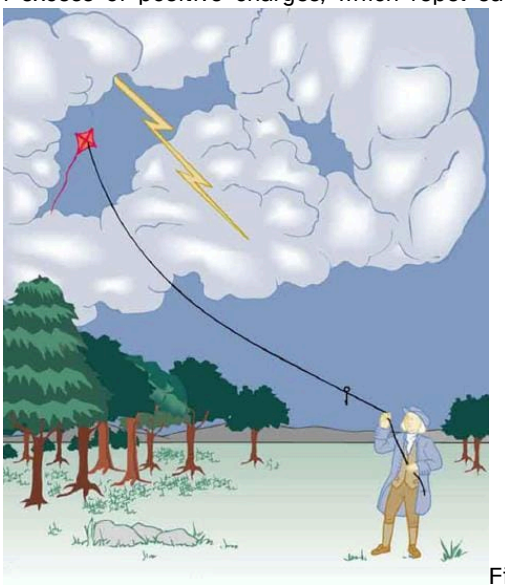


Figure 7.1.2: When Benjamin Franklin demonstrated that lightning was related to static electricity, he made a connection that is now part of the evidence that all directly experienced forces except the gravitational force are manifestations of the electromagnetic force.

Much has been written about Franklin. His experiments were only part of the life of a man who was a scientist, inventor, revolutionary, statesman, and writer. Franklin's experiments were not performed in isolation, nor were they the only ones to reveal connections.

For example, the Italian scientist Luigi Galvani (1737–1798) performed a series of experiments in which static electricity was used to stimulate contractions of leg muscles of dead frogs, an effect already known in humans subjected to static discharges. But Galvani also found that if he joined two metal wires (say copper and zinc) end to end and touched the other ends to muscles, he produced the same effect in frogs as static discharge. Alessandro Volta (1745–1827), partly inspired by Galvani's work, experimented with various combinations of metals and developed the battery.

During the same era, other scientists made progress in discovering fundamental connections. The periodic table was developed as the systematic properties of the elements were discovered. This influenced the development and refinement of the concept of atoms as the basis of matter. Such submicroscopic descriptions of matter also help explain a great deal more.

Atomic and molecular interactions, such as the forces of friction, cohesion, and adhesion, are now known to be manifestations of the **electromagnetic force**. Static electricity is just one aspect of the electromagnetic force, which also includes moving electricity and magnetism.

All the macroscopic forces that we experience directly, such as the sensations of touch and the tension in a rope, are due to the electromagnetic force, one of the four fundamental forces in nature. The gravitational force, another fundamental force, is actually sensed through the electromagnetic interaction of molecules, such as between those in our feet and those on the top of a bathroom scale. (The other two fundamental forces, the strong nuclear force and the weak nuclear force, cannot be sensed on the human scale.)

This chapter begins the study of electromagnetic phenomena at a fundamental level. The next several chapters will cover static electricity, moving electricity, and magnetism—collectively known as electromagnetism. In this chapter, we begin with the study of electric phenomena due to charges that are at least temporarily stationary, called electrostatics, or static electricity.

## Glossary

**static electricity**
: a buildup of electric charge on the surface of an object

**electromagnetic force**
: one of the four fundamental forces of nature; the electromagnetic force consists of static electricity, moving electricity and magnetism

This page titled 7.1: Prelude to Electric Charge and Electric Field is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 18.0: Prelude to Electric Charge and Electric Field by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.2: Static Electricity and Charge - Conservation of Charge

## Learning Objectives

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

- Define electric charge, and describe how the two types of charge interact.
- Describe three common situations that generate static electricity.
- State the law of conservation of charge.

What makes plastic wrap cling? Static electricity. Not only are applications of static electricity common these days, its existence has been known since ancient times. The first record of its effects dates to ancient Greeks who noted more than 500 years B.C. that polishing amber temporarily enabled it to attract bits of straw (Figure 7.2.1). The very word *electric* derives from the Greek word for amber (*electron*).



Figure 7.2.1: Borneo amber was mined in Sabah, Malaysia, from shale-sandstone-mudstone veins. When a piece of amber is rubbed with a piece of silk, the amber gains more electrons, giving it a net negative charge. At the same time, the silk, having lost electrons, becomes positively charged. (credit: Sebakoamber, Wikimedia Commons).

Many of the characteristics of static electricity can be explored by rubbing things together. Rubbing creates the spark you get from walking across a wool carpet, for example. Static cling generated in a clothes dryer and the attraction of straw to recently polished amber also result from rubbing. Similarly, lightning results from air movements under certain weather conditions. You can also rub a balloon on your hair, and the static electricity created can then make the balloon cling to a wall. We also have to be cautious of static electricity, especially in dry climates. When we pump gasoline, we are warned to discharge ourselves (after sliding across the seat) on a metal surface before grabbing the gas nozzle. Attendants in hospital operating rooms must wear booties with a conductive strip of aluminum foil on the bottoms to avoid creating sparks which may ignite flammable anesthesia gases combined with the oxygen being used.

Some of the most basic characteristics of static electricity include:

- The effects of static electricity are explained by a physical quantity not previously introduced, called electric charge.
- There are only two types of charge, one called positive and the other called negative.
- Like charges repel, whereas unlike charges attract.
- The force between charges decreases with distance.

How do we know there are two types of **electric charge**? When various materials are rubbed together in controlled ways, certain combinations of materials always produce one type of charge on one material and the opposite type on the other. By convention, we call one type of charge “positive”, and the other type “negative.” For example, when glass is rubbed with silk, the glass becomes positively charged and the silk negatively charged. Since the glass and silk have opposite charges, they attract one another like clothes that have rubbed together in a dryer. Two glass rods rubbed with silk in this manner will repel one another, since each rod has positive charge on it. Similarly, two silk cloths so rubbed will repel, since both cloths have negative charge. Figure 7.2.2 shows how these simple materials can be used to explore the nature of the force between charges.

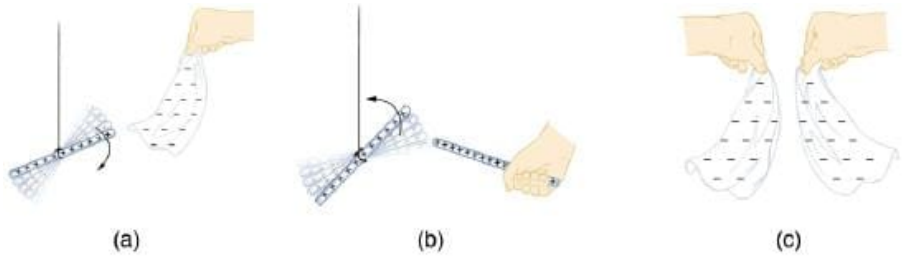



Figure 7.2.2: A glass rod becomes positively charged when rubbed with silk, while the silk becomes negatively charged. (a) The glass rod is attracted to the silk because their charges are opposite. (b) Two similarly charged glass rods repel. (c) Two similarly charged silk cloths repel.

More sophisticated questions arise. Where do these charges come from? Can you create or destroy charge? Is there a smallest unit of charge? Exactly how does the force depend on the amount of charge and the distance between charges? Such questions obviously occurred to Benjamin Franklin and other early researchers, and they interest us even today.

## Charge Carried by Electrons and Protons

Franklin wrote in his letters and books that he could see the effects of electric charge but did not understand what caused the phenomenon. Today we have the advantage of knowing that normal matter is made of atoms, and that atoms contain positive and negative charges, usually in equal amounts.

Figure 7.2.3 shows a simple model of an atom with negative **electrons** orbiting its positive nucleus. The nucleus is positive due to the presence of positively charged **protons**. Nearly all charge in nature is due to electrons and protons, which are two of the three building blocks of most matter. (The third is the neutron, which is neutral, carrying no charge.) Other charge-carrying particles are observed in cosmic rays and nuclear decay, and are created in particle accelerators. All but the electron and proton survive only a short time and are quite rare by comparison.

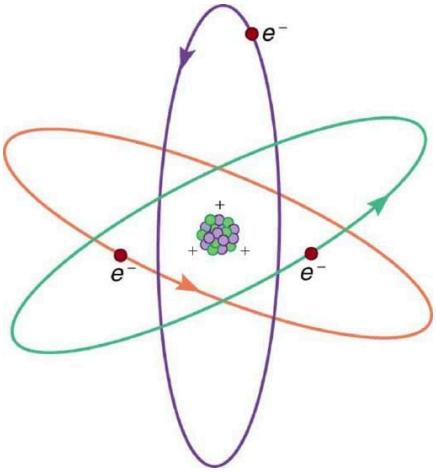



Figure 7.2.3: This simplified (and not to scale) view of an atom is called the planetary model of the atom. Negative electrons orbit a much heavier positive nucleus, as the planets orbit the much heavier sun. There the similarity ends, because forces in the atom are electromagnetic, whereas those in the planetary system are gravitational. Normal macroscopic amounts of matter contain immense numbers of atoms and molecules and, hence, even greater numbers of individual negative and positive charges.

The charges of electrons and protons are identical in magnitude but opposite in sign. Furthermore, all charged objects in nature are integral multiples of this basic quantity of charge, meaning that all charges are made of combinations of a basic unit of charge. Usually, charges are formed by combinations of electrons and protons. The magnitude of this basic charge is

$$|q_e| = 1.60 \times 10^{-19} C$$

The symbol $q$ is commonly used for charge and the subscript $e$ indicates the charge of a single electron (or proton).

The SI unit of charge is the coulomb (C). The number of protons needed to make a charge of 1.00 C is

$$1.00C \times \frac{1proton}{1.60 \times 10^{-19}C} = 6.25 \times 10^{18} protons$$

Similarly, $6.25 \times 10^{18}$ electrons have a combined charge of –1.00 coulomb. Just as there is a smallest bit of an element (an atom), there is a smallest bit of charge. There is no directly observed charge smaller than $| q_e |$ (see Things Great and Small: The Submicroscopic Origin of Charge), and all observed charges are integral multiples of \(|q_e|\).

## THINGS GREAT AND SMALL: THE SUBMICROSCOPIC ORIGIN OF CHARGE

With the exception of exotic, short-lived particles, all charge in nature is carried by electrons and protons. Electrons carry the charge we have named negative. Protons carry an equal-magnitude charge that we call positive. (Figure 7.2.4) Electron and proton charges are considered fundamental building blocks, since all other charges are integral multiples of those carried by electrons and protons. Electrons and protons are also two of the three fundamental building blocks of ordinary matter. The neutron is the third and has zero total charge.

Figure 7.2.4 shows a person touching a Van de Graaff generator and receiving excess positive charge. The expanded view of a hair shows the existence of both types of charges but an excess of positive. The repulsion of these positive like charges causes the strands of hair to repel other strands of hair and to stand up. The further blowup shows an artist's conception of an electron and a proton perhaps found in an atom in a strand of hair.

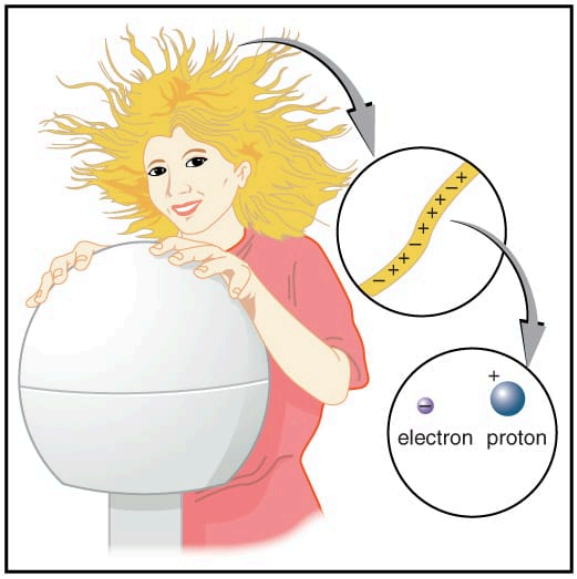



Figure 7.2.4: When this person touches a Van de Graaff generator, some electrons are attracted to the generator, resulting in an excess of positive charge, causing her hair to stand on end. The charges in one hair are shown. An artist's conception of an electron and a proton illustrate the particles carrying the negative and positive charges. We cannot really see these particles with visible light because they are so small (the electron seems to be an infinitesimal point), but we know a great deal about their measurable properties, such as the charges they carry.

The electron seems to have no substructure; in contrast, when the substructure of protons is explored by scattering extremely energetic electrons from them, it appears that there are point-like particles inside the proton. These sub-particles, named quarks, have never been directly observed, but they are believed to carry fractional charges as seen in Figure 7.2.5. Charges on electrons and protons and all other directly observable particles are unitary, but these quark substructures carry charges of either $-\frac{1}{3}$ or $+\frac{2}{3}$. There are continuing attempts to observe fractional charge directly and to learn of the properties of quarks, which are perhaps the ultimate substructure of matter.

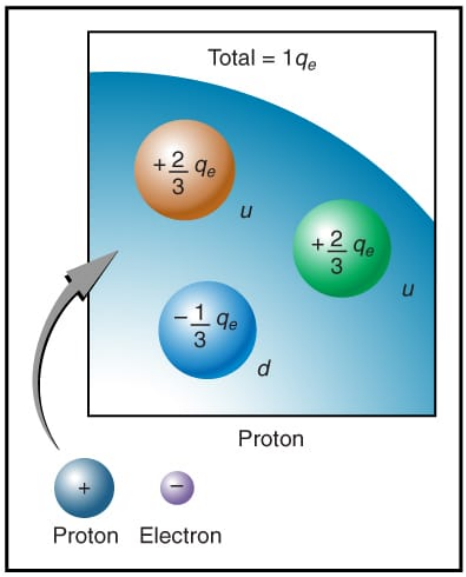



Figure 7.2.5: Artist's conception of fractional quark charges inside a proton. A group of three quark charges add up to the single positive charge on the proton: $-\frac{1}{3}q_e + \frac{2}{3}q_e + \frac{2}{3}q_e = +1q_e$.

## Separation of Charge in Atoms

Charges in atoms and molecules can be separated—for example, by rubbing materials together. Some atoms and molecules have a greater affinity for electrons than others and will become negatively charged by close contact in rubbing, leaving the other material positively charged. (Figure 7.2.6) Positive charge can similarly be induced by rubbing. Methods other than rubbing can also separate charges. Batteries, for example, use combinations of substances that interact in such a way as to separate charges. Chemical interactions may transfer negative charge from one substance to the other, making one battery terminal negative and leaving the first one positive.

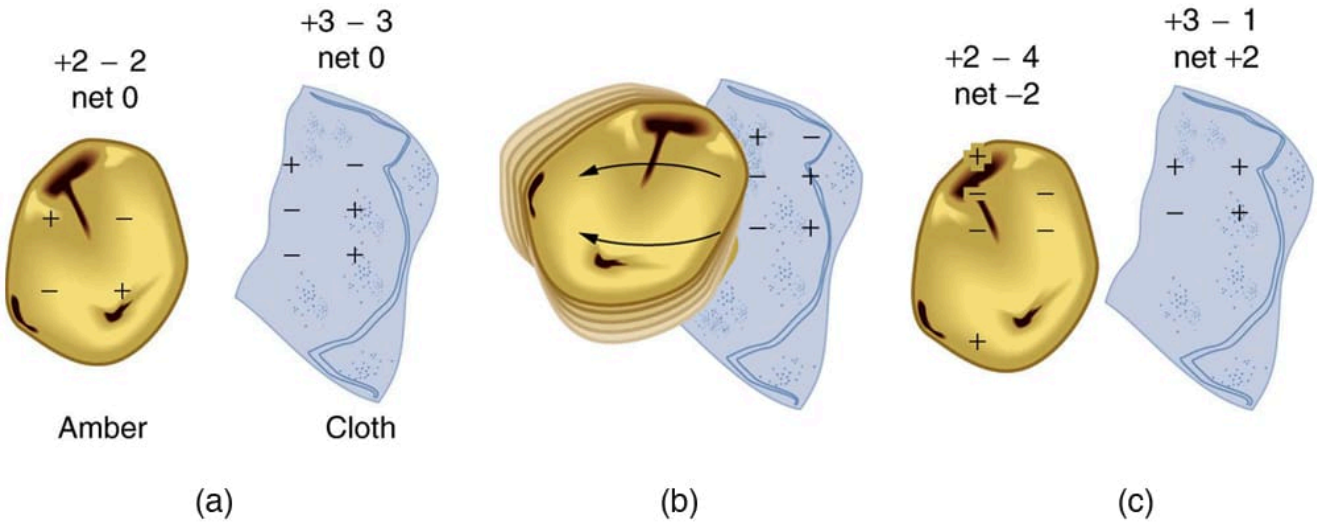



Figure 7.2.6: When materials are rubbed together, charges can be separated, particularly if one material has a greater affinity for electrons than another. (a) Both the amber and cloth are originally neutral, with equal positive and negative charges. Only a tiny fraction of the charges are involved, and only a few of them are shown here. (b) When rubbed together, some negative charge is transferred to the amber, leaving the cloth with a net positive charge. (c) When separated, the amber and cloth now have net charges, but the absolute value of the net positive and negative charges will be equal.

No charge is actually created or destroyed when charges are separated as we have been discussing. Rather, existing charges are moved about. In fact, in all situations the total amount of charge is always constant. This universally obeyed law of nature is called the **law of conservation of charge**.

## LAW OF CONSERVATION OF CHARGE

Total charge is constant in any process.

In more exotic situations, such as in particle accelerators, mass, $\Delta m$, can be created from energy in the amount $\Delta m = \frac{E}{c^2}$. Sometimes, the created mass is charged, such as when an electron is created. Whenever a charged particle is created, another having an opposite charge is always created along with it, so that the total charge created is zero. Usually, the two particles are "matter-antimatter" counterparts. For example, an antielectron would usually be created at the same time as an electron. The antielectron has a positive charge (it is called a positron), and so the total charge created is zero. (Figure 7.2.7) All particles have antimatter counterparts with opposite signs. When matter and antimatter counterparts are brought together, they completely annihilate one another. By annihilate, we mean that the mass of the two particles is converted to energy *E*, again obeying the relationship $\Delta m = \frac{E}{c^2}$. Since the two particles have equal and opposite charge, the total charge is zero before and after the annihilation; thus, total charge is conserved.

## MAKING CONNECTIONS: CONSERVATION LAWS

Only a limited number of physical quantities are universally conserved. Charge is one—energy, momentum, and angular momentum are others. Because they are conserved, these physical quantities are used to explain more phenomena and form more connections than other, less basic quantities. We find that conserved quantities give us great insight into the rules followed by nature and hints to the organization of nature. Discoveries of conservation laws have led to further discoveries, such as the weak nuclear force and the quark substructure of protons and other particles.

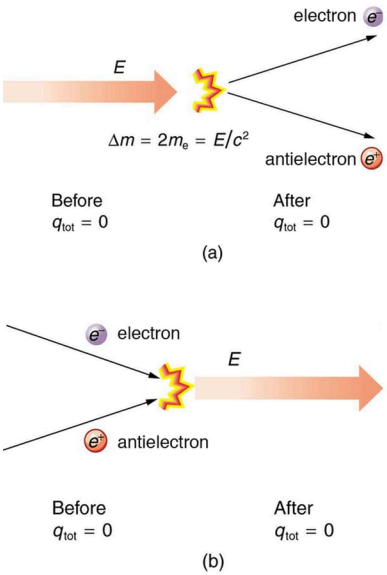



Figure 7.2.7: (a) When enough energy is present, it can be converted into matter. Here the matter created is an electron–antielectron pair. ($m_e$ is the electron's mass.) The total charge before and after this event is zero. (b) When matter and antimatter collide, they annihilate each other; the total charge is conserved at zero before and after the annihilation.

The law of conservation of charge is absolute—it has never been observed to be violated. Charge, then, is a special physical quantity, joining a very short list of other quantities in nature that are always conserved. Other conserved quantities include energy, momentum, and angular momentum.

## PHET EXPLORATIONS: BALLOONS AND STATIC ELECTRICITY

Why does a balloon stick to your sweater? Rub a balloon on a sweater, then let go of the balloon and it flies over and sticks to the sweater. View the charges in the sweater, balloons, and the wall.

## Summary

- There are only two types of charge, which we call positive and negative.
- Like charges repel, unlike charges attract, and the force between charges decreases with the square of the distance.
- The vast majority of positive charge in nature is carried by protons, while the vast majority of negative charge is carried by electrons.
- The electric charge of one electron is equal in magnitude and opposite in sign to the charge of one proton.
- An ion is an atom or molecule that has nonzero total charge due to having unequal numbers of electrons and protons.
- The SI unit for charge is the coulomb (C), with protons and electrons having charges of opposite sign but equal magnitude; the magnitude of this basic charge $|q_e|$ is
- $|q_e| = 1.60 \times 10^{-19} C$
- Whenever charge is created or destroyed, equal amounts of positive and negative are involved.
- Most often, existing charges are separated from neutral objects to obtain some net charge.
- Both positive and negative charges exist in neutral objects and can be separated by rubbing one object with another. For macroscopic objects, negatively charged means an excess of electrons and positively charged means a depletion of electrons.
- The law of conservation of charge ensures that whenever a charge is created, an equal charge of the opposite sign is created at the same time.

# Glossary

**electric charge**
: a physical property of an object that causes it to be attracted toward or repelled from another charged object; each charged object generates and is influenced by a force called an electromagnetic force

**law of conservation of charge**
: states that whenever a charge is created, an equal amount of charge with the opposite sign is created simultaneously

**electron**
: a particle orbiting the nucleus of an atom and carrying the smallest unit of negative charge

**proton**
: a particle in the nucleus of an atom and carrying a positive charge equal in magnitude and opposite in sign to the amount of negative charge carried by an electron

This page titled 7.2: Static Electricity and Charge - Conservation of Charge is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 18.1: Static Electricity and Charge - Conservation of Charge by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.3: Conductors and Insulators

**Learning Objectives**

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

- Define conductor and insulator, explain the difference, and give examples of each.
- Describe three methods for charging an object.
- Explain what happens to an electric force as you move farther from the source.
- Define polarization.

Some substances, such as metals and salty water, allow charges to move through them with relative ease. Some of the electrons in metals and similar conductors are not bound to individual atoms or sites in the material. These **free electrons** can move through the material much as air moves through loose sand. Any substance that has free electrons and allows charge to move relatively freely through it is called a **conductor**. The moving electrons may collide with fixed atoms and molecules, losing some energy, but they can move in a conductor. Superconductors allow the movement of charge without any loss of energy. Salty water and other similar conducting materials contain free ions that can move through them. An ion is an atom or molecule having a positive or negative (nonzero) total charge. In other words, the total number of electrons is not equal to the total number of protons.



Figure 7.3.1: This power adapter uses metal wires and connectors to conduct electricity from the wall socket to a laptop computer. The conducting wires allow electrons to move freely through the cables, which are shielded by rubber and plastic. These materials act as insulators that don't allow electric charge to escape outward. (credit: Evan-Amos, Wikimedia Commons)

Other substances, such as glass, do not allow charges to move through them. These are called **insulators**. Electrons and ions in insulators are bound in the structure and cannot move easily—as much as $10^{23}$ times more slowly than in conductors. Pure water and dry table salt are insulators, for example, whereas molten salt and salty water are conductors.

## Charging by Contact

Figure 7.3.2 shows an electroscope being charged by touching it with a positively charged glass rod. Because the glass rod is an insulator, it must actually touch the electroscope to transfer charge to or from it. (Note that the extra positive charges reside on the surface of the glass rod as a result of rubbing it with silk before starting the experiment.) Since only electrons move in metals, we see that they are attracted to the top of the electroscope. There, some are transferred to the positive rod by touch, leaving the electroscope with a net positive charge.

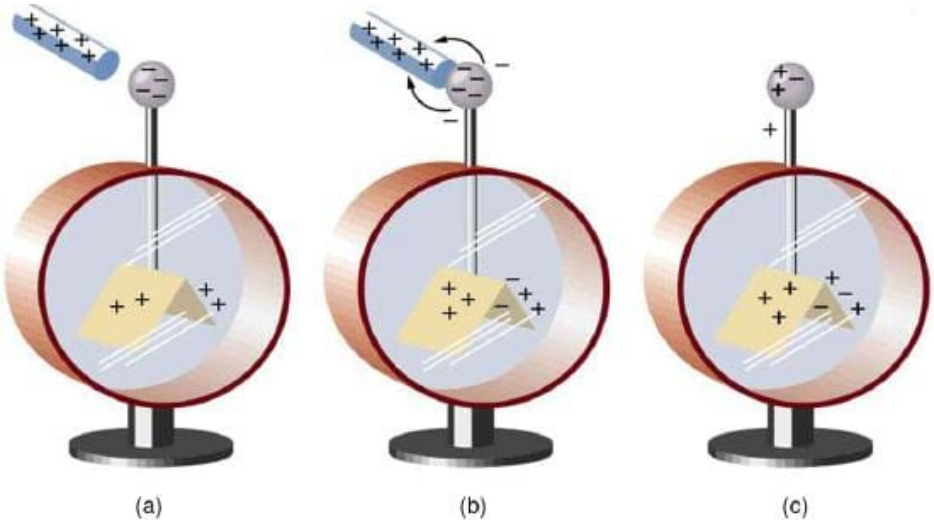



Figure 7.3.2: An electroscope is a favorite instrument in physics demonstrations and student laboratories. It is typically made with gold foil leaves hung from a (conducting) metal stem and is insulated from the room air in a glass-walled container. (a) A positively charged glass rod is brought near the tip of the electroscope, attracting electrons to the top and leaving a net positive charge on the leaves. Like charges in the light flexible gold leaves repel, separating them. (b) When the rod is touched against the ball, electrons are attracted and transferred, reducing the net charge on the glass rod but leaving the electroscope positively charged. (c) The excess charges are evenly distributed in the stem and leaves of the electroscope once the glass rod is removed.

Electrostatic repulsion in the leaves of the charged electroscope separates them. The electrostatic force has a horizontal component that results in the leaves moving apart as well as a vertical component that is balanced by the gravitational force. Similarly, the electroscope can be negatively charged by contact with a negatively charged object.

## Charging by Induction

It is not necessary to transfer excess charge directly to an object in order to charge it. Figure 7.3.3 shows a method of induction wherein a charge is created in a nearby object, without direct contact. Here we see two neutral metal spheres in contact with one another but insulated from the rest of the world. A positively charged rod is brought near one of them, attracting negative charge to that side, leaving the other sphere positively charged.

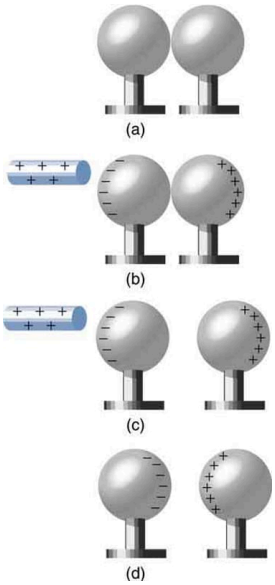



Figure 7.3.3: Charging by induction. (a) Two uncharged or neutral metal spheres are in contact with each other but insulated from the rest of the world. (b) A positively charged glass rod is brought near the sphere on the left, attracting negative charge and leaving the other sphere positively charged. (c) The spheres are separated before the rod is removed, thus separating negative and positive charge. (d) The spheres retain net charges after the inducing rod is removed—without ever

having been touched by a charged object.

This is an example of induced polarization of neutral objects. Polarization is the separation of charges in an object that remains neutral. If the spheres are now separated (before the rod is pulled away), each sphere will have a net charge. Note that the object closest to the charged rod receives an opposite charge when charged by induction. Note also that no charge is removed from the charged rod, so that this process can be repeated without depleting the supply of excess charge.

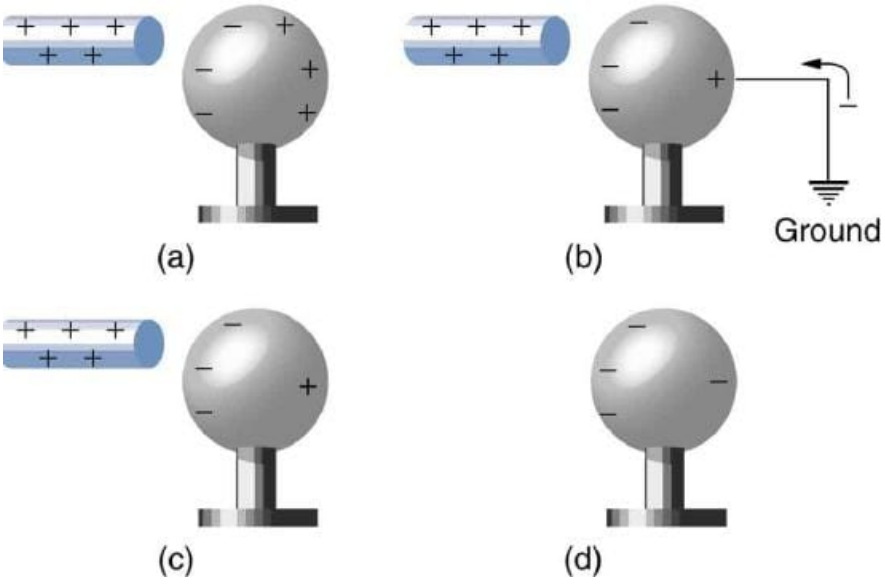



Figure 7.3.4: Charging by induction, using a ground connection. (a) A positively charged rod is brought near a neutral metal sphere, polarizing it. (b) The sphere is grounded, allowing electrons to be attracted from the earth's ample supply. (c) The ground connection is broken. (d) The positive rod is removed, leaving the sphere with an induced negative charge.

Another method of charging by induction is shown in Figure 7.3.4. The neutral metal sphere is polarized when a charged rod is brought near it. The sphere is then grounded, meaning that a conducting wire is run from the sphere to the ground. Since the earth is large and most ground is a good conductor, it can supply or accept excess charge easily. In this case, electrons are attracted to the sphere through a wire called the ground wire, because it supplies a conducting path to the ground. The ground connection is broken before the charged rod is removed, leaving the sphere with an excess charge opposite to that of the rod. Again, an opposite charge is achieved when charging by induction and the charged rod loses none of its excess charge.

Neutral objects can be attracted to any charged object. The pieces of straw attracted to polished amber are neutral, for example. If you run a plastic comb through your hair, the charged comb can pick up neutral pieces of paper. Figure 7.3.5 shows how the polarization of atoms and molecules in neutral objects results in their attraction to a charged object.

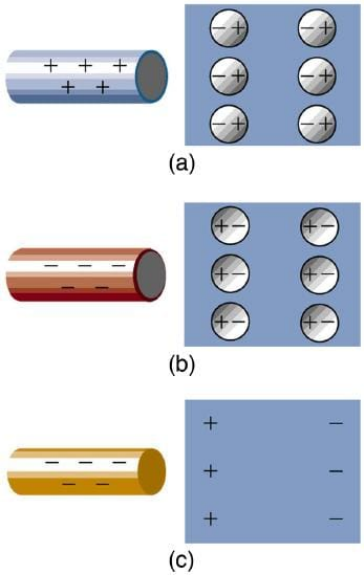



Figure 7.3.5: Both positive and negative objects attract a neutral object by polarizing its molecules. (a) A positive object brought near a neutral insulator polarizes its molecules. There is a slight shift in the distribution of the electrons orbiting the molecule, with unlike charges being brought nearer and like charges moved away. Since the electrostatic force decreases with distance, there is a net attraction. (b) A negative object produces the opposite polarization, but again attracts the neutral object. (c) The same effect occurs for a conductor; since the unlike charges are closer, there is a net attraction.

When a charged rod is brought near a neutral substance, an insulator in this case, the distribution of charge in atoms and molecules is shifted slightly. Opposite charge is attracted nearer the external charged rod, while like charge is repelled. Since the electrostatic force decreases with distance, the repulsion of like charges is weaker than the attraction of unlike charges, and so there is a net attraction. Thus a positively charged glass rod attracts neutral pieces of paper, as will a negatively charged rubber rod. Some molecules, like water, are polar molecules. Polar molecules have a natural or inherent separation of charge, although they are neutral overall. Polar molecules are particularly affected by other charged objects and show greater polarization effects than molecules with naturally uniform charge distributions.

## Check Your Understanding

Can you explain the attraction of water to the charged rod in the figure below?

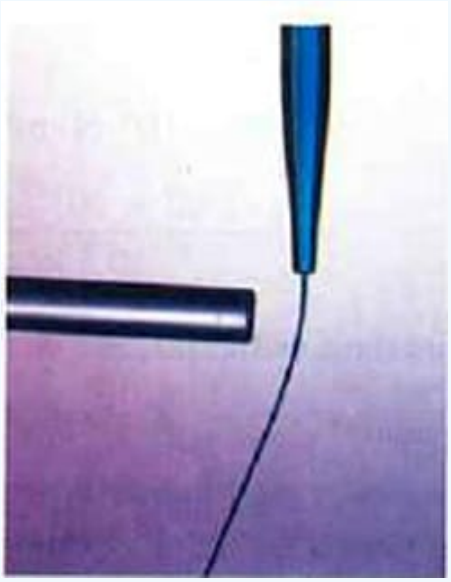


Figure 7.3.6.

**Solution**

**Water molecules are polarized, giving them slightly positive and slightly negative sides. This makes water even more susceptible to a charged rod's attraction. As the water flows downward, due to the force of gravity, the charged conductor exerts a net attraction to the opposite charges in the stream of water, pulling it closer.**

PHET EXPLORATIONS: JOHN TRAVOLTAGE

Make sparks fly with John Travoltage. Wiggle Johnnie's foot and he picks up charges from the carpet. Bring his hand close to the door knob and get rid of the excess charge.

## Summary

- Polarization is the separation of positive and negative charges in a neutral object.
- A conductor is a substance that allows charge to flow freely through its atomic structure.
- An insulator holds charge within its atomic structure.
- Objects with like charges repel each other, while those with unlike charges attract each other.
- A conducting object is said to be grounded if it is connected to the Earth through a conductor. Grounding allows transfer of charge to and from the earth's large reservoir.
- Objects can be charged by contact with another charged object and obtain the same sign charge.
- If an object is temporarily grounded, it can be charged by induction, and obtains the opposite sign charge.
- Polarized objects have their positive and negative charges concentrated in different areas, giving them a non-symmetrical charge.
- Polar molecules have an inherent separation of charge.

## Glossary

**free electron**
an electron that is free to move away from its atomic orbit

**conductor**
a material that allows electrons to move separately from their atomic orbits

**insulator**
a material that holds electrons securely within their atomic orbits

**grounded**
when a conductor is connected to the Earth, allowing charge to freely flow to and from Earth's unlimited reservoir

**induction**
the process by which an electrically charged object brought near a neutral object creates a charge in that object

**polarization**
slight shifting of positive and negative charges to opposite sides of an atom or molecule

**electrostatic repulsion**
the phenomenon of two objects with like charges repelling each other

This page titled 7.3: Conductors and Insulators is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 18.2: Conductors and Insulators by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.4: Coulomb's Law

## Learning Objectives

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

- State Coulomb's law in terms of how the electrostatic force changes with the distance between two objects.
- Calculate the electrostatic force between two charged point forces, such as electrons or protons.
- Compare the electrostatic force to the gravitational attraction for a proton and an electron; for a human and the Earth.

Through the work of scientists in the late 18th century, the main features of the **electrostatic force**—the existence of two types of charge, the observation that like charges repel, unlike charges attract, and the decrease of force with distance—were eventually refined, and expressed as a mathematical formula. The mathematical formula for the electrostatic force is called **Coulomb's law** after the French physicist Charles Coulomb (1736–1806), who performed experiments and first proposed a formula to calculate it.



Figure 7.4.1: This NASA image of Arp 87 shows the result of a strong gravitational attraction between two galaxies. In contrast, at the subatomic level, the electrostatic attraction between two objects, such as an electron and a proton, is far greater than their mutual attraction due to gravity. (credit: NASA/HST)

## Definition: Coulomb's Law

Coulomb's law calculates the magnitude of the force $F$ between two point charges, $q_1$ and $q_2$, separated by a distance $r$.

$$F = k\frac{|q_1 q_2|}{r^2}. \tag{7.4.1}$$

In SI units, the constant$k$ is equal to

$$k = 8.988 \times 10^9 \frac{N \cdot m^2}{C^2} \approx 8.99 \times 10^9 \frac{N \cdot m^2}{C^2}. \tag{7.4.2}$$

The electrostatic force is a vector quantity and is expressed in units of newtons. The force is understood to be along the line joining the two charges. (Figure 7.4.2)

Although the formula for Coulomb's law is simple, it was no mean task to prove it. The experiments Coulomb did, with the primitive equipment then available, were difficult. Modern experiments have verified Coulomb's law to great precision. For example, it has been shown that the force is inversely proportional to distance between two objects squared $(F \propto 1/r^2)$ to an accuracy of 1 part in $10^{16}$. No exceptions have ever been found, even at the small distances within the atom.

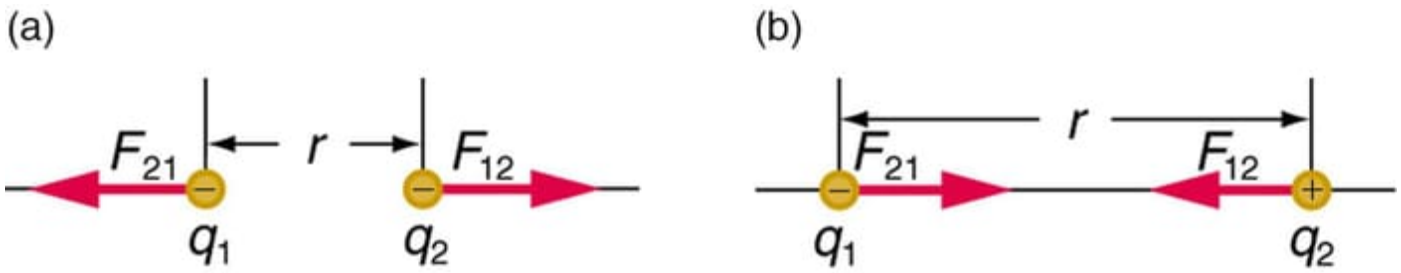



Figure 7.4.2: The magnitude of the electrostatic force$F$ between point charges $q_1$ and $q_2$ separated by a distance $r$ is given by Coulomb's law. Note that Newton's third law (every force exerted creates an equal and opposite force) applies as usual—the force on $q_1$ is equal in magnitude and opposite in direction to the force it exerts on $q_2$. (a) Like charges. (b) Unlike charges.

## Example 7.4.1: How Strong is the Coulomb Force Relative to the Gravitational Force?

Compare the electrostatic force between an electron and proton separated by $0.530 \times 10^{-10} m$ with the gravitational force between them. This distance is their average separation in a hydrogen atom.

**Strategy**

To compare the two forces, we first compute the electrostatic force using Coulomb's law, $F = k\frac{|q_1 q_2}{r^2}$. We then calculate the gravitational force using Newton's universal law of gravitation. Finally, we take a ratio to see how the forces compare in magnitude.

**Solution**

Entering the given and known information about the charges and separation of the electron and proton into the expression of Coulomb's law yields

$$F = k\frac{|q_1 q_2|}{r^2}$$

$$= (8.99 \times 10^9 N \cdot m^2/C^2) \times \frac{(1.60 \times 10^{-19} C)(1.60 \times 10^{-19} C)}{(0.530 \times 10^{-10} m)^2}$$

Thus the Coulomb force is

$$F = 8.19 \times 10^{-8} N.$$

The charges are opposite in sign, so this is an attractive force. This is a very large force for an electron—it would cause an acceleration of $8.99 \times 10^{22} m/s^2$(verification is left as an end-of-section problem).The gravitational force is given by Newton's law of gravitation as:

$$F_G = G\frac{mM}{r^2},$$

where $G = 6.67 \times 10^{-11} N \cdot m^2/kg^2$. Here $m$ and $M$ represent the electron and proton masses, which can be found in the appendices. Entering values for the knowns yields

$$F_G = (6.67 \times 10^{-11} N \cdot m^2/kg^2) \times \frac{(9.11 \times 10^{-31} kg)(1.67 \times 10^{-27} kg)}{(0.530 \times 10^{-10} m)^2} = 3.61 \times 10^{-47} N$$

This is also an attractive force, although it is traditionally shown as positive since gravitational force is always attractive. The ratio of the magnitude of the electrostatic force to gravitational force in this case is, thus,

$$\frac{F}{F_G} = 2.27 \times 10^{39}.$$

**Discussion**

> This is a remarkably large ratio! Note that this will be the ratio of electrostatic force to gravitational force for an electron and a proton at any distance (taking the ratio before entering numerical values shows that the distance cancels). This ratio gives some indication of just how much larger the Coulomb force is than the gravitational force between two of the most common particles in nature.

As the example implies, gravitational force is completely negligible on a small scale, where the interactions of individual charged particles are important. On a large scale, such as between the Earth and a person, the reverse is true. Most objects are nearly electrically neutral, and so attractive and repulsive **Coulomb forces** nearly cancel. Gravitational force on a large scale dominates interactions between large objects because it is always attractive, while Coulomb forces tend to cancel.

## Summary

- Frenchman Charles Coulomb was the first to publish the mathematical equation that describes the electrostatic force between two objects.
- Coulomb's law gives the magnitude of the force between point charges. It is $F = k\frac{|q_1 q_2|}{r^2}$, where $q_1$ and $q_2$ are two point charges separated by a distance $r$, and $k \approx 8.99 \times 10^9 N \cdot m^2/C^2$
- This Coulomb force is extremely basic, since most charges are due to point-like particles. It is responsible for all electrostatic effects and underlies most macroscopic forces.
- The Coulomb force is extraordinarily strong compared with the gravitational force, another basic force—but unlike gravitational force it can cancel, since it can be either attractive or repulsive.
- The electrostatic force between two subatomic particles is far greater than the gravitational force between the same two particles.

## Glossary

**Coulomb's law**
: the mathematical equation calculating the electrostatic force vector between two charged particles

**Coulomb force**
: another term for the electrostatic force

**electrostatic force**
: the amount and direction of attraction or repulsion between two charged bodies

This page titled 7.4: Coulomb's Law is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 18.3: Coulomb's Law by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.5: Electric Field- Concept of a Field Revisited

**Learning Objectives**

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

- Describe a force field and calculate the strength of an electric field due to a point charge.
- Calculate the force exerted on a test charge by an electric field.
- Explain the relationship between electrical force (**F**) on a test charge and electrical field strength (**E**).

Contact forces, such as between a baseball and a bat, are explained on the small scale by the interaction of the charges in atoms and molecules in close proximity. They interact through forces that include the **Coulomb force**. Action at a distance is a force between objects that are not close enough for their atoms to "touch." That is, they are separated by more than a few atomic diameters.

For example, a charged rubber comb attracts neutral bits of paper from a distance via the Coulomb force. It is very useful to think of an object being surrounded in space by a **force field**. The force field carries the force to another object (called a test object) some distance away.

## Concept of a Field

A field is a way of conceptualizing and mapping the force that surrounds any object and acts on another object at a distance without apparent physical connection. For example, the gravitational field surrounding the earth (and all other masses) represents the gravitational force that would be experienced if another mass were placed at a given point within the field.

In the same way, the Coulomb force field surrounding any charge extends throughout space. Using Coulomb's law, $F = k|q_1 q_2|/r^2$ , its magnitude is given by the equation $F = k|qQ|/r^2$, for a **point charge** (a particle having a charge $Q$) acting on a **test charge** $q$ at a distance $r$ (Figure $7.5.1$). Both the magnitude and direction of the Coulomb force field depend on $Q$ and the test charge $q$.

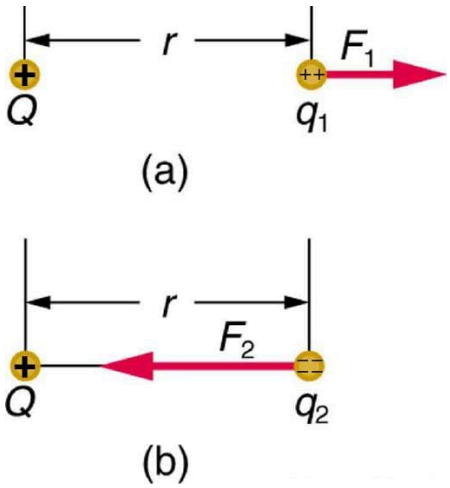



Figure $7.5.1$: The Coulomb force field due to a positive charge $Q$ is shown acting on two different charges. Both charges are the same distance from $Q$. (a) Since $q_1$ is positive, the force $F_1$ acting on it is repulsive. (b) The charge $q_2$ is negative and greater in magnitude than $q_1$, and so the force $F_2$ acting on it is attractive and stronger than $F_1$. The Coulomb force field is thus not unique at any point in space, because it depends on the test charges $q_1$ and $q_2$ as well as the charge $Q$.

To simplify things, we would prefer to have a field that depends only on $Q$ and not on the test charge $q$. The electric field is defined in such a manner that it represents only the charge creating it and is unique at every point in space. Specifically, the electric field $E$ is defined to be the ratio of the Coulomb force to the test charge:

$$\mathbf{E} = \frac{\mathbf{F}}{q}, \tag{7.5.1}$$

where $\mathbf{F}$ is the electrostatic force (or Coulomb force) exerted on a positive test charge $q$. It is understood that $\mathbf{E}$ is in the same direction as $\mathbf{F}$. It is also assumed that $q$ is so small that it does not alter the charge distribution creating the electric field. The units of electric field are newtons per coulomb (N/C). If the electric field is known, then the electrostatic force on any charge $q$ is simply obtained by multiplying charge times electric field, or $\mathbf{F} = q\mathbf{E}$. Consider the electric field due to a point charge $Q$. According to Coulomb's law, the force it exerts on a test charge $q$ is $F = k|qQ|/r^2$. Thus the magnitude of the electric field, $E$, for a point charge is

$$E = |\frac{F}{q}| = k|\frac{qQ}{qr^2}| = k\frac{|Q|}{r^2}. \tag{7.5.2}$$

Since the test charge cancels, we see that

$$E = k\frac{|Q|}{r^2}. \tag{7.5.3}$$

The electric field is thus seen to depend only on the charge $Q$ and the distance $r$; it is completely independent of the test charge $q$.

## Example 7.5.1: Calculating the Electric Field of a Point Charge

Calculate the strength and direction of the electric field $E$ due to a point charge of 2.00 nC (nano-Coulombs) at a distance of 5.00 mm from the charge.

**Strategy**

We can find the electric field created by a point charge by using the equation $E = kQ/r^2$.

**Solution**

Here $Q = 2.00 \times 10^{-9}C$ and $r = 5.00 \times 10^{-3}m$. Entering those values into the above equation gives

$$\begin{aligned} E &= k\frac{Q}{r^2} \\ &= (8.99 \times 10^9 N \cdot m^2/C^2) \times \frac{(2.00 \times 10^{-9}C)}{(5.00 \times 10^{-3}m)^2} \\ &= 7.19 \times 10^5 N/C. \end{aligned}$$

**Discussion**

This **electric field strength** is the same at any point 5.00 mm away from the charge $Q$ that creates the field. It is positive, meaning that it has a direction pointing away from the charge $Q$.

## Example 7.5.2: Calculating the Force Exerted on a Point Charge by an Electric Field

What force does the electric field found in the previous example exert on a point charge of$-0.250\mu C$?

**Strategy**

Since we know the electric field strength and the charge in the field, the force on that charge can be calculated using the definition of electric field $\mathbf{E} = \mathbf{F}/q$ rearranged to $\mathbf{F} = q\mathbf{E}$.

**Solution**

The magnitude of the force on a charge $q = -.250\mu C$ exerted by a field of strength $E = 7.20 \times 10^5 N/C$ is thus,

$$\begin{aligned} F &= -qE \\ &= (0.250 \times 10^{-6} C)(7.20 \times 10^{5} N/C) \\ &= 0.180 N. \end{aligned}$$

Because $q$ is negative, the force is directed opposite to the direction of the field.

**Discussion**

The force is attractive, as expected for unlike charges. (The field was created by a positive charge and here acts on a negative charge.) The charges in this example are typical of common static electricity, and the modest attractive force obtained is similar to forces experienced in static cling and similar situations.

**PHET EXPLORATIONS: ELECTRIC FIELD OF DREAMS**

Play ball! Add charges to the Field of Dreams and see how they react to the electric field. Turn on a background electric field and adjust the direction and magnitude.




Figure 7.5.2: Electric Field of Dreams

## Summary

- The electrostatic force field surrounding a charged object extends out into space in all directions.
- The electrostatic force exerted by a point charge on a test charge at a distance$r$ depends on the charge of both charges, as well as the distance between the two.
- The electric field $\mathbf{E}$ is defined to be $\mathbf{E} = \frac{\mathbf{F}}{q}$, where $\mathbf{F}$ is the Coulomb or electrostatic force exerted on a small positive test charge $q$. $\mathbf{E}$ has units of N/C.
- The magnitude of the electric field $\mathbf{E}$ created by a point charge $Q$ is $\mathbf{E} = k\frac{|Q|}{r^2}$. where $r$ is the distance from $Q$. The electric field $\mathbf{E}$ is a vector and fields due to multiple charges add like vectors.

# Glossary

**field**
: a map of the amount and direction of a force acting on other objects, extending out into space

**point charge**
: A charged particle, designated Q, generating an electric field

**test charge**
: A particle (designated q) with either a positive or negative charge set down within an electric field generated by a point charge

This page titled 7.5: Electric Field- Concept of a Field Revisited is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 18.4: Electric Field- Concept of a Field Revisited by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.7: Electric Potential Energy- Potential Difference

## Learning Objectives

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

- Define electric potential and electric potential energy.
- Describe the relationship between potential difference and electrical potential energy.
- Explain electron volt and its usage in submicroscopic process.
- Determine electric potential energy given potential difference and amount of charge.

When a free positive charge $q$ is accelerated by an electric field, such as shown in Figure 7.7.1, it is given kinetic energy. The process is analogous to an object being accelerated by a gravitational field. It is as if the charge is going down an electrical hill where its electric potential energy is converted to kinetic energy. Let us explore the work done on a charge $q$ by the electric field in this process, so that we may develop a definition of electric potential energy.

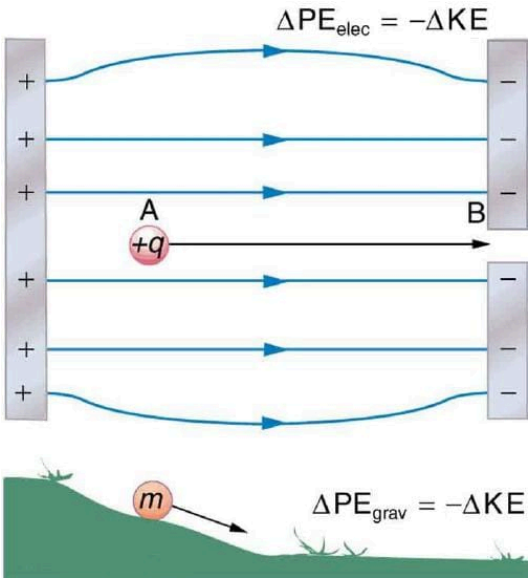



Figure 7.7.1: A charge accelerated by an electric field is analogous to a mass going down a hill. In both cases potential energy is converted to another form. Work is done by a force, but since this force is conservative, we can write $W = -\Delta \mathrm{PE}$.

The electrostatic or Coulomb force is conservative, which means that the work done on $q$ is independent of the path taken. This is exactly analogous to the gravitational force in the absence of dissipative forces such as friction. When a force is conservative, it is possible to define a potential energy associated with the force, and it is usually easier to deal with the potential energy (because it depends only on position) than to calculate the work directly.

We use the letters PE to denote electric potential energy, which has units of joules (J). The change in potential energy, $\Delta \mathrm{PE}$, is crucial, since the work done by a conservative force is the negative of the change in potential energy; that is, $W = -\Delta \mathrm{PE}$. For example, work $W$ done to accelerate a positive charge from rest is positive and results from a loss in PE, or a negative $\Delta \mathrm{PE}$. There must be a minus sign in front of $\Delta \mathrm{PE}$ to make $W$ positive. PE can be found at any point by taking one point as a reference and calculating the work needed to move a charge to the other point.

## POTENTIAL ENERGY

$W = -\Delta\text{PE}$. For example, work $W$ done to accelerate a positive charge from rest is positive and results from a loss in PE, or a negative $\Delta\text{PE}$ There must be a minus sign in front of $\Delta\text{PE}$ to make $W$ positive. PE can be found at any point by taking one point as a reference and calculating the work needed to move a charge to the other point.

Gravitational potential energy and electric potential energy are quite analogous. Potential energy accounts for work done by a conservative force and gives added insight regarding energy and energy transformation without the necessity of dealing with the force directly. It is much more common, for example, to use the concept of voltage (related to electric potential energy) than to deal with the Coulomb force directly.

Calculating the work directly is generally difficult, since $W = Fd\cos\theta$ and the direction and magnitude of $F$ can be complex for multiple charges, for odd-shaped objects, and along arbitrary paths. But we do know that, since $F = qE$, the work, and hence $\Delta\text{PE}$ , is proportional to the test charge $q$ To have a physical quantity that is independent of test charge, we define **electric potential** $V$ (or simply potential, since electric is understood) to be the potential energy per unit charge:

$$V = \frac{\text{PE}}{q}. \tag{7.7.1}$$

## ELECTRIC POTENTIAL

This is the electric potential energy per unit charge.

$$V = \frac{\text{PE}}{q} \tag{7.7.2}$$

Since PE is proportional to $q$, the dependence on $q$ cancels. Thus $V$ does not depend on $q$. The change in potential energy $\Delta\text{PE}$ is crucial, and so we are concerned with the difference in potential or potential difference $\Delta V$ between two points, where

$$\Delta V = V_B - V_A = \frac{\Delta\text{PE}}{q}. \tag{7.7.3}$$

The **potential difference** between points A and B, $V_B - V_A$, is thus defined to be the change in potential energy of a charge $q$ moved from A to B, divided by the charge. Units of potential difference are joules per coulomb, given the name volt (V) after Alessandro Volta.

$$1\text{V} = 1\frac{\text{J}}{\text{C}} \tag{7.7.4}$$

## POTENTIAL DIFFERENCE

The potential difference between points A and B, $V_B - V_A$, is defined to be the change in potential energy of a charge $q$ moved from A to B, divided by the charge. Units of potential difference are joules per coulomb, given the name volt (V) after Alessandro Volta.

$$1\text{V} = 1\frac{\text{J}}{\text{C}} \tag{7.7.5}$$

The familiar term **voltage** is the common name for potential difference. Keep in mind that whenever a voltage is quoted, it is understood to be the potential difference between two points. For example, every battery has two terminals, and its voltage is the potential difference between them. More fundamentally, the point you choose to be zero volts is arbitrary. This is analogous to the fact that gravitational potential energy has an arbitrary zero, such as sea level or perhaps a lecture hall floor.

In summary, the relationship between potential difference (or voltage) and electrical potential energy is given by

$$\Delta V = \frac{\Delta \text{PE}}{q} \text{ and } \Delta \text{PE} = q\Delta V. \tag{7.7.6}$$

## POTENTIAL DIFFERENCE AND ELECTRICAL POTENTIAL ENERGY

The relationship between potential difference (or voltage) and electrical potential energy is given by

$$\Delta = \frac{\Delta \text{PE}}{q} \text{ and } \Delta \text{PE} = q\Delta V. \tag{7.7.7}$$

The second equation is equivalent to the first.

Voltage is not the same as energy. Voltage is the energy per unit charge. Thus a motorcycle battery and a car battery can both have the same voltage (more precisely, the same potential difference between battery terminals), yet one stores much more energy than the other since $\Delta PE = q\Delta V$. The car battery can move more charge than the motorcycle battery, although both are 12 V batteries.

## Example 7.7.1:Calculating Energy

Suppose you have a 12.0 V motorcycle battery that can move 5000 C of charge, and a 12.0 V car battery that can move 60,000 C of charge. How much energy does each deliver? (Assume that the numerical value of each charge is accurate to three significant figures.)

**Strategy**

To say we have a 12.0 V battery means that its terminals have a 12.0 V potential difference. When such a battery moves charge, it puts the charge through a potential difference of 12.0 V, and the charge is given a change in potential energy equal to $\Delta PE = q\Delta V$.

So to find the energy output, we multiply the charge moved by the potential difference.

**Solution**

For the motorcycle battery, $q = 5000\text{C}$ and $\Delta = 12.0\text{V}$. The total energy delivered by the motorcycle battery is

$$\Delta \text{PE}_{cycle} = (5000\text{C})(12.0\text{V}) \tag{7.7.8}$$

$$= (5000\text{C})(12.0\text{J/C}) \tag{7.7.9}$$

$$= 6.00 \times 10^4\text{J}. \tag{7.7.10}$$

Similarly, for the car battery, $q = 60,000\text{C}$ and

$$\Delta \text{PE}_{car} = (60,000\text{C})(12.0\text{V}) \tag{7.7.11}$$

$$= 7.20 \times 10^5\text{J}. \tag{7.7.12}$$

**Discussion**

While voltage and energy are related, they are not the same thing. The voltages of the batteries are identical, but the energy supplied by each is quite different. Note also that as a battery is discharged, some of its energy is used internally and its terminal voltage drops, such as when headlights dim because of a low car battery. The energy supplied by the battery is still calculated as in this example, but not all of the energy is available for external use.

Note that the energies calculated in the previous example are absolute values. The change in potential energy for the battery is negative, since it loses energy. These batteries, like many electrical systems, actually move negative charge—electrons in particular. The batteries repel electrons from their negative terminals (A) through whatever circuitry is involved and attract them to their positive terminals (B) as shown in Figure 7.7.2. The change in potential is $\Delta V = V_B - V_A = +12\text{V}$ and the charge $q$ is

negative, so that $\Delta \mathrm{PE} = q\Delta V$ is negative, meaning the potential energy of the battery has decreased when $q$ has moved from A to B.

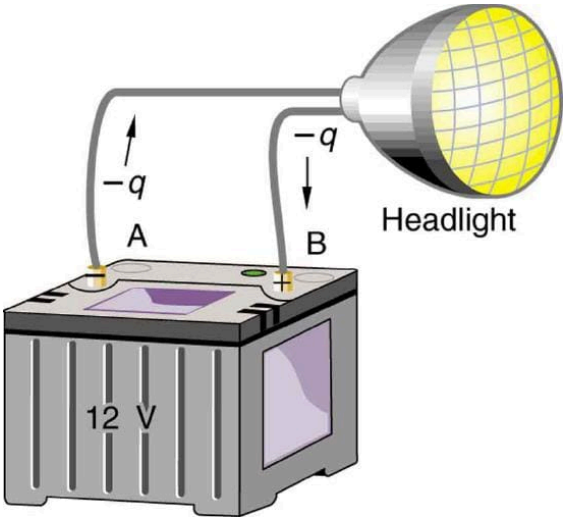



Figure 7.7.2: A battery moves negative charge from its negative terminal through a headlight to its positive terminal. Appropriate combinations of chemicals in the battery separate charges so that the negative terminal has an excess of negative charge, which is repelled by it and attracted to the excess positive charge on the other terminal. In terms of potential, the positive terminal is at a higher voltage than the negative. Inside the battery, both positive and negative charges move.

## Example 7.7.2: How Many Electrons Move through a Headlight Each Second?

When a 12.0 V car battery runs a single 30.0 W headlight, how many electrons pass through it each second?

**Strategy**

To find the number of electrons, we must first find the charge that moved in 1.00 s. The charge moved is related to voltage and energy through the equation $\Delta \mathrm{PE} = q\Delta V$. A 30.0 W lamp uses 30.0 joules per second. Since the battery loses energy, we have $\Delta \mathrm{PE} = -30.0J$ and, since the electrons are going from the negative terminal to the positive, we see that $\Delta V = +12.0V$.

**Solution**

To find the charge $q$ moved, we solve the equation $\Delta \mathrm{PE} = q\Delta V$:

$$q = \frac{\Delta \mathrm{PE}}{\Delta V}. \tag{7.7.13}$$

Entering the values for $\Delta PE$ and $\Delta V$, we get

$$q = \frac{-30.0\mathrm{J}}{+12.0\mathrm{V}} = \frac{-30.0\mathrm{J}}{+12.0\mathrm{J/C}} = -2.50\mathrm{C}. \tag{7.7.14}$$

The number of electrons $n_e$ is the total charge divided by the charge per electron. That is,

$$n_e = \frac{-2.50\mathrm{C}}{-1.60 \times 10^{-19}\mathrm{C/e^-}} = 1.56 \times 10^{19}\mathrm{electrons}. \tag{7.7.15}$$

**Discussion**

This is a very large number. It is no wonder that we do not ordinarily observe individual electrons with so many being present in ordinary systems. In fact, electricity had been in use for many decades before it was determined that the moving charges in many circumstances were negative. Positive charge moving in the opposite direction of negative charge often produces identical effects; this makes it difficult to determine which is moving or whether both are moving.

# The Electron Volt

The energy per electron is very small in macroscopic situations like that in the previous example—a tiny fraction of a joule. But on a submicroscopic scale, such energy per particle (electron, proton, or ion) can be of great importance. For example, even a tiny fraction of a joule can be great enough for these particles to destroy organic molecules and harm living tissue. The particle may do its damage by direct collision, or it may create harmful x rays, which can also inflict damage. It is useful to have an energy unit related to submicroscopic effects. Figure $7.7.3$ shows a situation related to the definition of such an energy unit. An electron is accelerated between two charged metal plates as it might be in an old-model television tube or oscilloscope. The electron is given kinetic energy that is later converted to another form—light in the television tube, for example. (Note that downhill for the electron is uphill for a positive charge.) Since energy is related to voltage by $\Delta PE = q\Delta V$ we can think of the joule as a coulomb-volt.

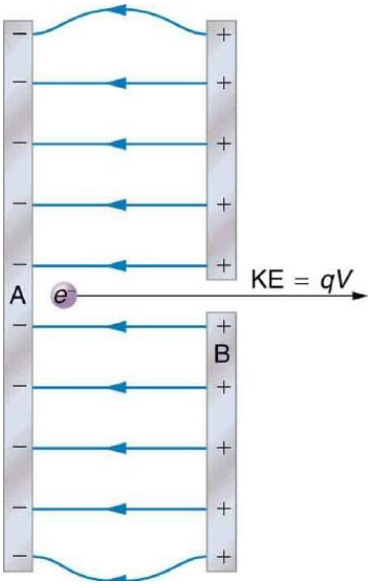



Figure $7.7.3$: A typical electron gun accelerates electrons using a potential difference between two metal plates. The energy of the electron in electron volts is numerically the same as the voltage between the plates. For example, a 5000 V potential difference produces 5000 eV electrons.

On the submicroscopic scale, it is more convenient to define an energy unit called the **electron volt** (eV), which is the energy given to a fundamental charge accelerated through a potential difference of 1 V. In equation form,

$$1\text{ev} = (1.60 \times 10^{-19}\text{C})(1\text{V}) = (1.60 \times 10^{-19}\text{C})(1\text{J/C}) \tag{7.7.16}$$

$$= 1.60 \times 10^{-19} J. \tag{7.7.17}$$

## ELECTRON VOLT

On the submicroscopic scale, it is more convenient to define an energy unit called the electron volt (eV), which is the energy given to a fundamental charge accelerated through a potential difference of 1 V. In equation form,

$$1\text{eV} = (1.60 \times 10^{-19}\text{C})(1\text{V}) = (1.60 \times 10^{-19}\text{C})(1\text{J/C}) \tag{7.7.18}$$

$$= 1.60 \times 10^{-19}\text{C} \tag{7.7.19}$$

An electron accelerated through a potential difference of 1 V is given an energy of 1 eV. It follows that an electron accelerated through 50 V is given 50 eV. A potential difference of 100,000 V (100 kV) will give an electron an energy of 100,000 eV (100 keV), and so on. Similarly, an ion with a double positive charge accelerated through 100 V will be given 200 eV of energy. These simple relationships between accelerating voltage and particle charges make the electron volt a simple and convenient energy unit in such circumstances.

CONNECTIONS: ENERGY UNITS

The electron volt (eV) is the most common energy unit for submicroscopic processes. This will be particularly noticeable in the chapters on modern physics. Energy is so important to so many subjects that there is a tendency to define a special energy unit for each major topic. There are, for example, calories for food energy, kilowatt-hours for electrical energy, and therms for natural gas energy.

The electron volt is commonly employed in submicroscopic processes—chemical valence energies and molecular and nuclear binding energies are among the quantities often expressed in electron volts. For example, about 5 eV of energy is required to break up certain organic molecules. If a proton is accelerated from rest through a potential difference of 30 kV, it is given an energy of 30 keV (30,000 eV) and it can break up as many as 6000 of these molecules ( $30,000\text{eV} \div 5\text{eV}$ per molecule $= 6000$ molecules). Nuclear decay energies are on the order of 1 MeV (1,000,000 eV) per event and can, thus, produce significant biological damage.

## Conservation of Energy

The total energy of a system is conserved if there is no net addition (or subtraction) of work or heat transfer. For conservative forces, such as the electrostatic force, conservation of energy states that mechanical energy is a constant.

**Mechanical energy** is the sum of the kinetic energy and potential energy of a system; that is, $KE + PE = \text{ constant}$. A loss of PE of a charged particle becomes an increase in its KE. Here PE is the electric potential energy. Conservation of energy is stated in equation form as

$$\text{KE} + \text{PE} = \text{constant} \tag{7.7.20}$$

or

$$\text{KE}_i + \text{PE}_i = \text{KE}_f + \text{PE}_f, \tag{7.7.21}$$

where i and f stand for initial and final conditions. As we have found many times before, considering energy can give us insights and facilitate problem solving.

### Example 7.7.3: Electrical Potential Energy Converted to Kinetic Energy

Calculate the final speed of a free electron accelerated from rest through a potential difference of 100 V. (Assume that this numerical value is accurate to three significant figures.)

**Strategy**

We have a system with only conservative forces. Assuming the electron is accelerated in a vacuum, and neglecting the gravitational force (we will check on this assumption later), all of the electrical potential energy is converted into kinetic energy. We can identify the initial and final forms of energy to be $\text{KE}_i = 0, \text{KE}_f = \frac{1}{2}mv^2, \text{PE}_i = qV, \text{ and } \text{PE}_f = 0$.

**Solution**

Conservation of energy states that

$$\text{KE}_i + \text{PE}_i = \text{KE}_f + \text{PE}_f \tag{7.7.22}$$

Entering the forms identified above, we obtain

$$qV = \frac{mv^2}{2}. \tag{7.7.23}$$

We solve this for $v$:

$$v = \sqrt{\frac{2qV}{m}}. \tag{7.7.24}$$

Entering values for $q,\ V,\ \text{and}\ m$ gives

$$v = \sqrt{\frac{2(-1.60 \times 10^{-19}\text{C})(-100\text{J/C})}{9.11 \times 10^{-31}\text{kg}}} \tag{7.7.25}$$

$$= 5.93 \times 10^{6}\text{m/s}. \tag{7.7.26}$$

**Discussion**

Note that both the charge and the initial voltage are negative, as in Figure. From the discussions in Electric Charge and Electric Field, we know that electrostatic forces on small particles are generally very large compared with the gravitational force. The large final speed confirms that the gravitational force is indeed negligible here. The large speed also indicates how easy it is to accelerate electrons with small voltages because of their very small mass. Voltages much higher than the 100 V in this problem are typically used in electron guns. Those higher voltages produce electron speeds so great that relativistic effects must be taken into account. That is why a low voltage is considered (accurately) in this example.

## Summary

- Electric potential is potential energy per unit charge.
- The potential difference between points A and B, $V_\text{B} - V_\text{A}$, defined to be the change in potential energy of a charge $q$ moved from A to B, is equal to the change in potential energy divided by the charge, Potential difference is commonly called voltage, represented by the symbol $\Delta V$.

$$\Delta V = \frac{\Delta\text{PE}}{q} \ \text{and}\ \ \Delta\text{PE} = q\Delta V.$$

- An electron volt is the energy given to a fundamental charge accelerated through a potential difference of 1 V. In equation form,

$$1\text{eV} = (1.60 \times 10^{-19}\text{C})(1\text{V}) = (1.60 \times 10^{-19}\text{C})(1\text{J/C})$$

$$= 1.60 \times 10^{-19}\text{J}.$$

- Mechanical energy is the sum of the kinetic energy and potential energy of a system, that is, $\text{KE} + \text{PE}$ This sum is a constant.

## Glossary

**electric potential**
: potential energy per unit charge

**potential difference (or voltage)**
: change in potential energy of a charge moved from one point to another, divided by the charge; units of potential difference are joules per coulomb, known as volt

**electron volt**
: the energy given to a fundamental charge accelerated through a potential difference of one volt

**mechanical energy**
: sum of the kinetic energy and potential energy of a system; this sum is a constant

This page titled 7.7: Electric Potential Energy- Potential Difference is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 19.1: Electric Potential Energy- Potential Difference by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.8: Prelude to Electric Current, Resistance, and Ohm's Law

The flicker of numbers on a handheld calculator, nerve impulses carrying signals of vision to the brain, an ultrasound device sending a signal to a computer screen, the brain sending a message for a baby to twitch its toes, an electric train pulling its load over a mountain pass, a hydroelectric plant sending energy to metropolitan and rural users—these and many other examples of electricity involve *electric current, the movement of charge*. Humankind has indeed harnessed electricity, the basis of technology, to improve our quality of life. Whereas the previous two chapters concentrated on static electricity and the fundamental force underlying its behavior, the next few chapters will be devoted to electric and magnetic phenomena involving current. In addition to exploring applications of electricity, we shall gain new insights into nature—in particular, the fact that all magnetism results from electric current.



Figure 20.1.1 Electric energy in massive quantities is transmitted from this hydroelectric facility, the Srisailam power station located along the Krishna River in India, by the movement of charge—that is, by electric current. (credit: Chintohere, Wikimedia Commons)

This page titled 7.8: Prelude to Electric Current, Resistance, and Ohm's Law is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.0: Prelude to Electric Current, Resistance, and Ohm's Law by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.9: Current

**Learning Objectives**

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

- Define electric current, ampere, and drift velocity
- Describe the direction of charge flow in conventional current.
- Use drift velocity to calculate current and vice versa.

## Electric Current

Electric current is defined to be the rate at which charge flows. A large current, such as that used to start a truck engine, moves a large amount of charge in a small time, whereas a small current, such as that used to operate a hand-held calculator, moves a small amount of charge over a long period of time. In equation form, **electric current** $I$ is defined to be

$$I = \frac{\Delta Q}{\Delta t}, \tag{7.9.1}$$

where $\Delta Q$ is the amount of charge passing through a given area in time $\Delta t$. As in previous chapters, initial time is often taken to be zero, in which case $\Delta t = t$. (7.9.1). The SI unit for current is the **ampere** (A), named for the French physicist André-Marie Ampère (1775–1836). From Equation 7.9.1, we see that an ampere is one coulomb per second:

$$1A = 1C/s \tag{7.9.2}$$

Not only are fuses and circuit breakers rated in amperes (or amps), so are many electrical appliances.

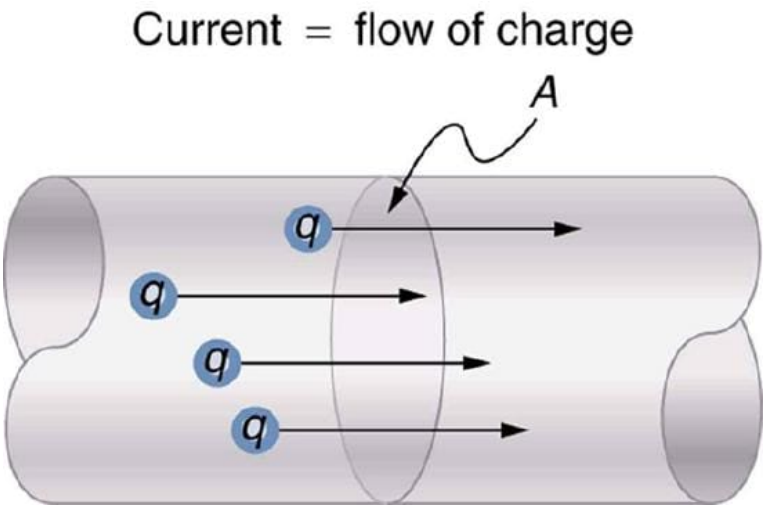



Figure 7.9.1: The rate of flow of charge is current. An ampere is the flow of one coulomb through an area in one second.

**Example 7.9.1: Calculating Currents: Current in a Truck Battery and a Handheld Calculator**

a. What is the current involved when a truck battery sets in motion 720 C of charge in 4.00 s while starting an engine?
b. How long does it take 1.00 C of charge to flow through a handheld calculator if a 0.300-mA current is flowing?

**Strategy**

We can use the definition of current in the equation $I = \Delta Q/\Delta t$ to find the current in part (a), since charge and time are given. In part (b), we rearrange the definition of current and use the given values of charge and current to find the time required.

**Solution (a)**

Entering the given values for charge and time into the definition of current gives

$$\begin{aligned} I &= \frac{\Delta Q}{\Delta t} \\ &= \frac{720C}{4.00s} \\ &= 180C/s \\ &= 180A. \end{aligned}$$

**Discussion (a)**

This large value for current illustrates the fact that a large charge is moved in a small amount of time. The currents in these "starter motors" are fairly large because large frictional forces need to be overcome when setting something in motion.

**Solution (b)**

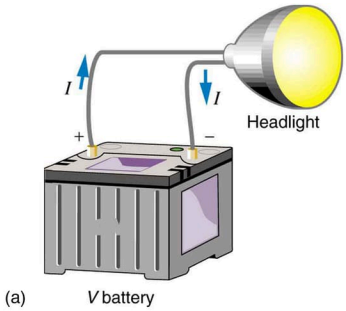
Solving the relationship $I = \Delta Q/\Delta t$ for time $\Delta t$, and entering the known values for charge and current gives

$$\begin{aligned} \Delta t &= \frac{\Delta Q}{I} \\ &= \frac{1.00C}{0.300 \times 10^{-3}C/s} \\ &= 3.33 \times 10^{3}s. \end{aligned}$$

**Discussion (b)**

This time is slightly less than an hour. The small current used by the hand-held calculator takes a much longer time to move a smaller charge than the large current of the truck starter. So why can we operate our calculators only seconds after turning them on? It's because calculators require very little energy. Such small current and energy demands allow handheld calculators to operate from solar cells or to get many hours of use out of small batteries. Remember, calculators do not have moving parts in the same way that a truck engine has with cylinders and pistons, so the technology requires smaller currents.

Figure $7.9.2$ shows a simple circuit and the standard schematic representation of a battery, conducting path, and load (a resistor). Schematics are very useful in visualizing the main features of a circuit. A single schematic can represent a wide variety of situations. The schematic in Figure $7.9.2b$, for example, can represent anything from a truck battery connected to a headlight lighting the street in front of the truck to a small battery connected to a penlight lighting a keyhole in a door. Such schematics are useful because the analysis is the same for a wide variety of situations. We need to understand a few schematics to apply the concepts and analysis to many more situations.



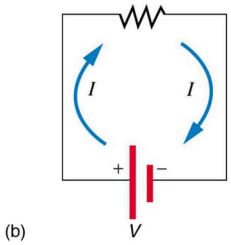




Figure 7.9.2: (a) A simple electric circuit. A closed path for current to flow through is supplied by conducting wires connecting a load to the terminals of a battery. (b) In this schematic, the battery is represented by the two parallel red lines, conducting wires are shown as straight lines, and the zigzag represents the load. The schematic represents a wide variety of similar circuits.

Note that the direction of current flow in Figure 7.9.2 is from positive to negative. *The direction of conventional current is the direction that positive charge would flow.* Depending on the situation, positive charges, negative charges, or both may move. In metal wires, for example, current is carried by electrons—that is, negative charges move. In ionic solutions, such as salt water, both positive and negative charges move. This is also true in nerve cells. A Van de Graaff generator used for nuclear research can produce a current of pure positive charges, such as protons. Figure 7.9.3 illustrates the movement of charged particles that compose a current. The fact that conventional current is taken to be in the direction that positive charge would flow can be traced back to American politician and scientist Benjamin Franklin in the 1700s. He named the type of charge associated with electrons negative, long before they were known to carry current in so many situations. Franklin, in fact, was totally unaware of the small-scale structure of electricity.

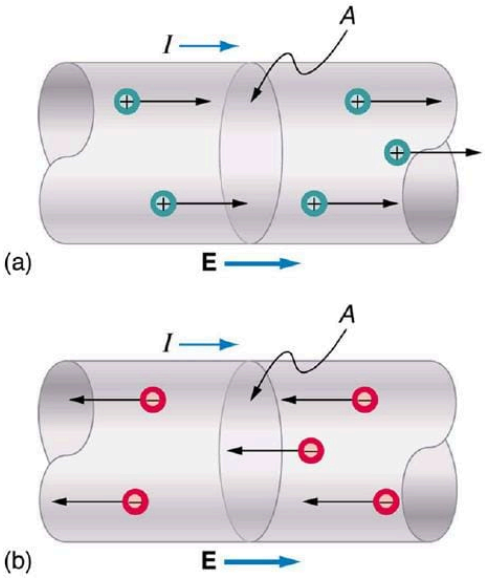



Figure 7.9.3: Current $I$ is the rate at which charge moves through an area $A$, such as the cross-section of a wire. Conventional current is defined to move in the direction of the electric field. (a) Positive charges move in the direction of the electric field and the same direction as conventional current. (b) Negative charges move in the direction opposite to the electric field. Conventional current is in the direction opposite to the movement of negative charge. The flow of electrons is sometimes referred to as electronic flow.

It is important to realize that there is an electric field in conductors responsible for producing the current, as illustrated in Figure

7.9.3 Unlike static electricity, where a conductor in equilibrium cannot have an electric field in it, conductors carrying a current have an electric field and are not in static equilibrium. An electric field is needed to supply energy to move the charges.

### MAKING CONNECTIONS: TAKE-HOME INVESTIGATION-ELECTRIC CURRENT ILLUSTRATION

Find a straw and little peas that can move freely in the straw. Place the straw flat on a table and fill the straw with peas. When you pop one pea in at one end, a different pea should pop out the other end. This demonstration is an analogy for an electric current. Identify what compares to the electrons and what compares to the supply of energy. What other analogies can you find for an electric current?

Note that the flow of peas is based on the peas physically bumping into each other; electrons flow due to mutually repulsive electrostatic forces.

### Example 7.9.2: Calculating the Number of Electrons that Move through a Calculator

If the 0.300-mA current through the calculator mentioned in the example is carried by electrons, how many electrons per second pass through it?

**Strategy**

The current calculated in the previous example was defined for the flow of positive charge. For electrons, the magnitude is the same, but the sign is opposite, $I_{electrons} = -0.300 \times 10^{-3} C/s$. Since each electron $(e^-)$ has a charge of $-1.60 \times 10^{-19} C$, we can convert the current in coulombs per second to electrons per second.

**Solution**:

Starting with the definition of current, we have

$$I_{electrons} = \frac{\Delta Q_{electrons}}{\Delta t} = \frac{-0.300 \times 10^{-3} C}{s}.$$

We divide this by the charge per electron, so that

$$\frac{e}{s} = \frac{-0.300 \times 10{-}3C}{s} \times \frac{1e}{-1.60 \times 10^{-19} C} = 1.88 \times 10^{15} \frac{e}{s}.$$

**Discussion:**

There are so many charged particles moving, even in small currents, that individual charges are not noticed, just as individual water molecules are not noticed in water flow. Even more amazing is that they do not always keep moving forward like soldiers in a parade. Rather they are like a crowd of people with movement in different directions but a general trend to move forward. There are lots of collisions with atoms in the metal wire and, of course, with other electrons.

## Drift Velocity

Electrical signals are known to move very rapidly. Telephone conversations carried by currents in wires cover large distances without noticeable delays. Lights come on as soon as a switch is flicked. Most electrical signals carried by currents travel at speeds on the order of $10^8 m/s$, a significant fraction of the speed of light. Interestingly, the individual charges that make up the current move *much* more slowly on average, typically drifting at speeds on the order of $10^{-4} m/s$. How do we reconcile these two speeds, and what does it tell us about standard conductors?

The high speed of electrical signals results from the fact that the force between charges acts rapidly at a distance. Thus, when a

free charge is forced into a wire, as in 7.9.4, the incoming charge pushes other charges ahead of it, which in turn push on charges farther down the line. The density of charge in a system cannot easily be increased, and so the signal is passed on rapidly. The resulting electrical shock wave moves through the system at nearly the speed of light. To be precise, this rapidly moving signal or shock wave is a rapidly propagating change in electric field.

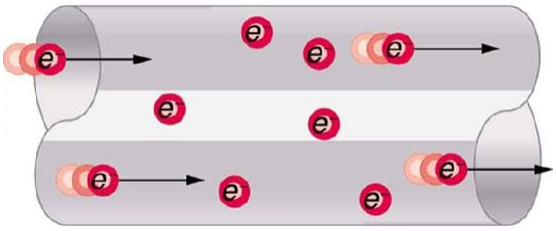


Figure 7.9.4: When charged particles are forced into this volume of a conductor, an equal number are quickly forced to leave. The repulsion between like charges makes it difficult to increase the number of charges in a volume. Thus, as one charge enters, another leaves almost immediately, carrying the signal rapidly forward.

Good conductors have large numbers of free charges in them. In metals, the free charges are free electrons. 7.9.5 shows how free electrons move through an ordinary conductor. The distance that an individual electron can move between collisions with atoms or other electrons is quite small. The electron paths thus appear nearly random, like the motion of atoms in a gas. But there is an electric field in the conductor that causes the electrons to drift in the direction shown (opposite to the field, since they are negative). The **drift velocity** $v_d$ i s the average velocity of the free charges. Drift velocity is quite small, since there are so many free charges. If we have an estimate of the density of free electrons in a conductor, we can calculate the drift velocity for a given current. The larger the density, the lower the velocity required for a given current.

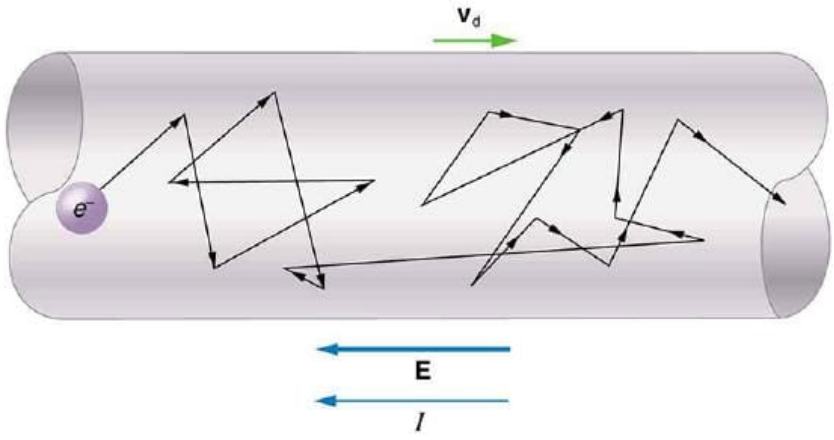



Figure 7.9.5: Free electrons moving in a conductor make many collisions with other electrons and atoms. The path of one electron is shown. The average velocity of the free charges is called the drift velocity, $v_d$, and it is in the direction opposite to the electric field for electrons. The collisions normally transfer energy to the conductor, requiring a constant supply of energy to maintain a steady current.

## CONDUCTION OF ELECTRICITY AND HEAT

Good electrical conductors are often good heat conductors, too. This is because large numbers of free electrons can carry electrical current and can transport thermal energy.

The free-electron collisions transfer energy to the atoms of the conductor. The electric field does work in moving the electrons through a distance, but that work does not increase the kinetic energy (nor speed, therefore) of the electrons. The work is transferred to the conductor's atoms, possibly increasing temperature. Thus a continuous power input is required to keep a current flowing. An exception, of course, is found in superconductors, for reasons we shall explore in a later chapter. Superconductors can have a steady current without a continual supply of energy—a great energy savings. In contrast, the supply of energy can be useful, such as in a lightbulb filament. The supply of energy is necessary to increase the temperature of the tungsten filament, so that the filament glows.

## MAKING CONNECTIONS: TAKE-HOME INVESTIGATION--FILAMENT OBSERVATIONS

Find a lightbulb with a filament. Look carefully at the filament and describe its structure. To what points is the filament connected?

We can obtain an expression for the relationship between current and drift velocity by considering the number of free charges in a segment of wire, as illustrated in Figure 7.9.6. *The number of free charges per unit volume* is given the symbol $n$ and depends on the material. The shaded segment has a volume $Ax$, so that the number of free charges in it is $nAx$. The charge $\Delta Q$ in this segment is thus $qnAx$, where $q$ is the amount of charge on each carrier. (Recall that for electrons, $q$ is $-1.60 \times 10^{-19} C$.) Current is charge moved per unit time; thus, if all the original charges move out of this segment in time $\Delta t$, the current is

$$I = \frac{\Delta Q}{\Delta t} = \frac{qnAx}{\Delta t}. \tag{7.9.3}$$

Note that $x/\Delta t$ is the magnitude of the drift velocity, $v_d$, since the charges move an average distance $x$ in a time $\Delta t$. Rearranging terms gives

$$I = nqAv_d, \tag{7.9.4}$$

where $I$ is the current through a wire of cross-sectional area $A$ made of a material with a free charge density $n$. The carriers of the current each have charge $q$ and move with a drift velocity of magnitude $v_d$.

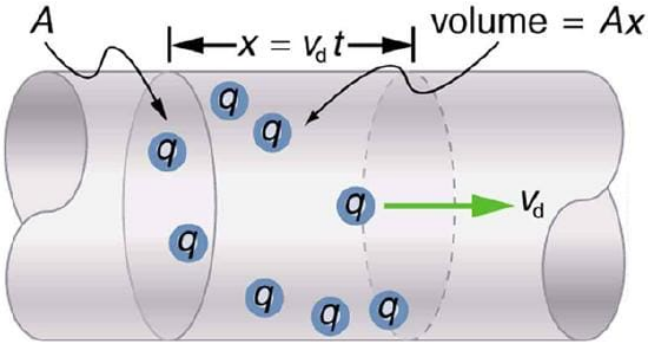



Figure 7.9.6: All the charges in the shaded volume of this wire move out in a time $t$, having a drift velocity of magnitude $v_d = x/t$. See text for further discussion.

Note that simple drift velocity is not the entire story. The speed of an electron is much greater than its drift velocity. In addition, not all of the electrons in a conductor can move freely, and those that do might move somewhat faster or slower than the drift velocity. So what do we mean by free electrons? Atoms in a metallic conductor are packed in the form of a lattice structure. Some electrons are far enough away from the atomic nuclei that they do not experience the attraction of the nuclei as much as the inner electrons do. These are the free electrons. They are not bound to a single atom but can instead move freely among the atoms in a "sea" of electrons. These free electrons respond by accelerating when an electric field is applied. Of course as they move they collide with the atoms in the lattice and other electrons, generating thermal energy, and the conductor gets warmer. In an insulator, the organization of the atoms and the structure do not allow for such free electrons.

## Example 7.9.3: Calculating Drift Velocity in a Common Wire

Calculate the drift velocity of electrons in a 12-gauge copper wire (which has a diameter of 2.053 mm) carrying a 20.0-A current, given that there is one free electron per copper atom. (Household wiring often contains 12-gauge copper wire, and the maximum current allowed in such wire is usually 20 A.) The density of copper is $8.80 \times 10^3 kg/m^3$.

**Strategy**

We can calculate the drift velocity using the equation $I = nqAv_d$. The current $I = 20.0A$ is given, and $q = -1.60 \times 10^{-19}C$ is the charge of an electron. We can calculate the area of a cross-section of the wire using the formula $A = \pi r^2$, where $r$ is one-half the given diameter, 2.053 mm. We are given the density of copper, $8.80 \times 10^3 kg/m^3$, and the periodic table shows that the atomic mass of copper is 63.54 g/mol. We can use these two quantities along with Avogadro's number, $6.02 \times 10^{23} atoms/mol$, to determine $n$, the number of free electrons per cubic meter.

**Solution**

First, calculate the density of free electrons in copper. There is one free electron per copper atom. Therefore, is the same as the number of copper atoms per $m^3$. We can now find $n$ as follows:

$$\begin{aligned} n &= \frac{1e^-}{atom} \times \frac{6.02 \times 10^{23} atoms}{mol} \times \frac{1mol}{63.54g} \times \frac{1000g}{kg} \times \frac{8.80 \times 10^3 kg}{1m^3} \\ &= 8.342 \times 10^{28} e^-/m^3. \end{aligned}$$

The cross-sectional area of the wire is

Missing \end{align*}

Misplaced &

Rearranging $I = nqAv_d$ to isolate drift velocity gives

$$\begin{aligned} v_d &= \frac{I}{nqA} \\ &= \frac{20.0A}{\left(8.342 \times 10^{28}/m^3\right)\left(-1.60 \times 10^{-19}C\right)\left(3.310 \times 10^{-6}m^2\right)} \\ &= -4.53 \times 10^{-4} m/s. \end{aligned}$$

**Discussion**

The minus sign indicates that the negative charges are moving in the direction opposite to conventional current. The small value for drift velocity (on the order of $10^{-4} m/s$) confirms that the signal moves on the order of $10^{12}$ times faster (about $10^8 m/s$) than the charges that carry it.

# Summary

- Electric current $I$ is the rate at which charge flows, given by

$$I = \frac{\Delta Q}{\Delta t,} \tag{7.9.5}$$

  where $\Delta Q$ is the amount of charge passing through an area in time $\Delta t$.
- The direction of conventional current is taken as the direction in which positive charge moves.
- The SI unit for current is the ampere (A), where $1A = 1C/s$.
- Current is the flow of free charges, such as electrons and ions.
- Drift velocitiy $v_d$ is the average speed at which these charges move.
- Current $I$ is proportional to drift velocity $v_d$, as expressed in the relationship $I = nqAv_d$. Here, $I$ is the current through a wire of cross-sectional area $A$. The wire's material as a free-charge density $n$, and each carrier has charge $q$ and a drift velocity $v_d$.
- Electrical signals travel at speeds about $10^{12}$ times greater than the drift velocity of free electrons.

# Glossary

**electric current**
: the rate at which charge flows, $I$ = ΔQ/Δ$t$

**ampere**
: (amp) the SI unit for current; 1 A = 1 C/s

**drift velocity**
: the average velocity at which free charges flow in response to an electric field

This page titled 7.9: Current is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.1: Current by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.10: Ohm's Law - Resistance and Simple Circuits

**Learning Objectives**

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

- Explain the origin of Ohm's law.
- Calculate voltages, currents, or resistances with Ohm's law.
- Explain what an ohmic material is.
- Describe a simple circuit.

What drives current? We can think of various devices—such as batteries, generators, wall outlets, and so on—which are necessary to maintain a current. All such devices create a potential difference and are loosely referred to as voltage sources. When a voltage source is connected to a conductor, it applies a potential difference $V$ that creates an electric field. The electric field in turn exerts force on charges, causing current.

## Ohm's Law

The current that flows through most substances is directly proportional to the voltage $V$ applied to it. The German physicist Georg Simon Ohm (1787–1854) was the first to demonstrate experimentally that the current in a metal wire is *directly proportional to the voltage applied*:

$$I \propto V. \tag{7.10.1}$$

This important relationship is known as **Ohm's law**. It can be viewed as a cause-and-effect relationship, with voltage the cause and current the effect. This is an empirical law like that for friction—an experimentally observed phenomenon. Such a linear relationship doesn't always occur.

## Resistance and Simple Circuits

If voltage drives current, what impedes it? The electric property that impedes current (crudely similar to friction and air resistance) is called resistance $R$. Collisions of moving charges with atoms and molecules in a substance transfer energy to the substance and limit current. Resistance is defined as inversely proportional to current, or

$$I \propto \frac{1}{R}. \tag{7.10.2}$$

Thus, for example, current is cut in half if resistance doubles. Combining the relationships of current to voltage and current to resistance gives

$$I = \frac{V}{R}. \tag{7.10.3}$$

This relationship is also called Ohm's law. Ohm's law in this form really defines resistance for certain materials. Ohm's law (like Hooke's law) is not universally valid. The many substances for which Ohm's law holds are called **ohmic**. These include good conductors like copper and aluminum, and some poor conductors under certain circumstances. Ohmic materials have a resistance $R$ that is independent of voltage $V$ and current $I$. An object that has simple resistance is called a *resistor*, even if its resistance is small. The unit for resistance is an **ohm** and is given the symbol $\Omega$ (upper case Greek omega). Rearranging $I = V/R$ gives $R = V/I$, and so the units of resistance are 1 ohm = 1 volt per ampere:

$$1\Omega = 1\frac{V}{A}. \tag{7.10.4}$$

Figure $7.10.1$ shows the schematic for a simple circuit. A **simple circuit** has a single voltage source and a single resistor. The wires connecting the voltage source to the resistor can be assumed to have negligible resistance, or their resistance can be

included in $R$.

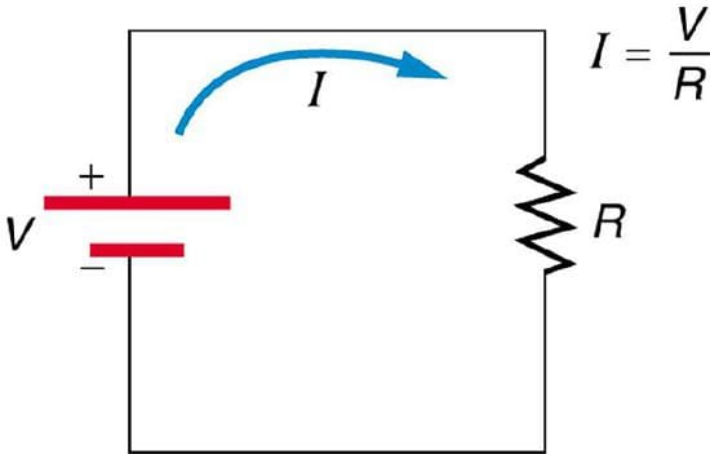



Figure 7.10.1: A simple electric circuit in which a closed path for current to flow is supplied by conductors (usually metal wires) connecting a load to the terminals of a battery, represented by the red parallel lines. The zigzag symbol represents the single resistor and includes any resistance in the connections to the voltage source.

### Example 7.10.1: Calculating Resistance: An Automobile Headlight:

What is the resistance of an automobile headlight through which 2.50 A flows when 12.0 V is applied to it?

**Strategy**

We can rearrange Ohm's law as stated by $I = V/R$ and use it to find the resistance.

**Solution:**

Rearranging Equation 7.10.3 and substituting known values gives

$$\begin{aligned} R &= \frac{V}{I} \\ &= \frac{12.0V}{2.50A} \\ &= 4.80\Omega. \end{aligned}$$

**Discussion:**

This is a relatively small resistance, but it is larger than the cold resistance of the headlight. As we shall see, resistance of metals usually *increases* with*increasing* temperature, and so the bulb has a lower resistance when it is first switched on and will draw considerably more current during its brief warm-up period.

Resistances range over many orders of magnitude. Some ceramic insulators, such as those used to support power lines, have resistances of $10^{12}\Omega$ or more. A dry person may have a hand-to-foot resistance of $10^{5}\Omega$, whereas the resistance of the human heart is about $10^{3}\Omega$. A meter-long piece of large-diameter copper wire may have a resistance of $10^{-5}\Omega$, and superconductors have no resistance at all (they are non-ohmic). Resistance is related to the shape of an object and the material of which it is composed, as will be seen in Resistance and Resistivity.

Additional insight is gained by solving $I = V/R$ for $V$, yielding

$$V = IR. \tag{7.10.5}$$

The expression for $V$ can be interpreted as the *voltage drop across a resistor produced by the low of current* $I$. The phrase $IR$ *drop* is often used for this voltage. For instance, the headlight in the example has an $IR$ drop of 12.0 V. If voltage is measured at various points in a circuit, it will be seen to increase at the voltage source and decrease at the resistor. Voltage is similar to fluid

pressure. The voltage source is like a pump, creating a pressure difference, causing current—the flow of charge. The resistor is like a pipe that reduces pressure and limits flow because of its resistance. Conservation of energy has important consequences here. The voltage source supplies energy (causing an electric field and a current), and the resistor converts it to another form (such as thermal energy). In a simple circuit (one with a single simple resistor), the voltage supplied by the source equals the voltage drop across the resistor, since $PE = q\Delta V$, and the same $q$ flows through each. Thus the energy supplied by the voltage source and the energy converted by the resistor are equal ( Figure 7.10.2).

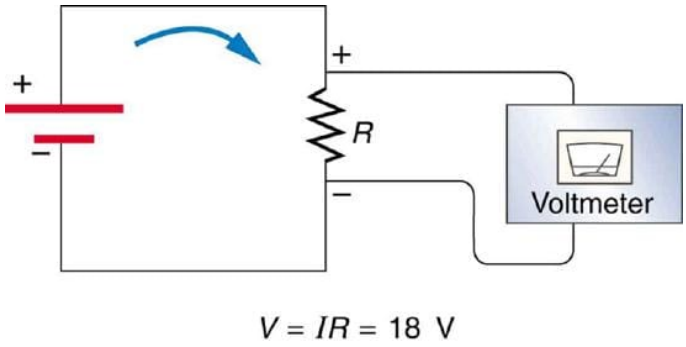



Figure 7.10.2: The voltage drop across a resistor in a simple circuit equals the voltage output of the battery.

**MAKING CONNECTIONS: CONSERVATION OF ENERGY**

In a simple electrical circuit, the sole resistor converts energy supplied by the source into another form. Conservation of energy is evidenced here by the fact that all of the energy supplied by the source is converted to another form by the resistor alone. We will find that conservation of energy has other important applications in circuits and is a powerful tool in circuit analysis.

## Summary

- A simple circuit is one in which there is a single voltage source and a single resistance.
- One statement of Ohm's law gives the relationship between current $I$, voltage $V$, and resistance $R$ in an simple circuit to be $I = \frac{V}{R}$.
- Resistance has units of ohms ($\Omega$), related to volts and amperes by $1\Omega = 1V/A$.
- There is a voltage or $IR$ drop across a resistor, caused by the current flowing through it, given by $V = IR$.

## Glossary

**Ohm's law**
: an empirical relation stating that the current *I* is proportional to the potential difference *V*, $\propto$ *V*; it is often written as *I* = *V*/*R*, where *R* is the resistance

**resistance**
: the electric property that impedes current; for ohmic materials, it is the ratio of voltage to current, *R* = *V*/*I*

**ohm**
: the unit of resistance, given by 1Ω = 1 V/A

**ohmic**
: a type of a material for which Ohm's law is valid

**simple circuit**
: a circuit with a single voltage source and a single resistor

This page titled 7.10: Ohm's Law - Resistance and Simple Circuits is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.2: Ohm's Law - Resistance and Simple Circuits by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.11: Resistance and Resistivity

**Learning Objectives**

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

- Explain the concept of resistivity.
- Use resistivity to calculate the resistance of specified configurations of material.
- Use the thermal coefficient of resistivity to calculate the change of resistance with temperature.

## Material and Shape Dependence of Resistance

The resistance of an object depends on its shape and the material of which it is composed. The cylindrical resistor in Figure 1 is easy to analyze, and, by so doing, we can gain insight into the resistance of more complicated shapes. As you might expect, the cylinder's electric resistance $R$ is directly proportional to its length $L$, similar to the resistance of a pipe to fluid flow. The longer the cylinder, the more collisions charges will make with its atoms. The greater the diameter of the cylinder, the more current it can carry (again similar to the flow of fluid through a pipe). In fact, $R$ is inversely proportional to the cylinder's cross-sectional area $A$.

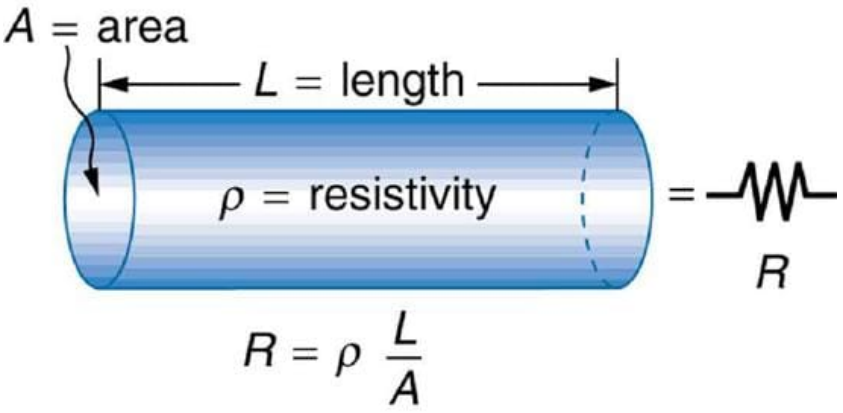



Figure 7.11.1: A uniform cylinder of length $L$ and cross-sectional area $A$. Its resistance to the flow of current is similar to the resistance posed by a pipe to fluid flow. The longer the cylinder, the greater its resistance. The larger its cross-sectional area $A$, the smaller its resistance.

For a given shape, the resistance depends on the material of which the object is composed. Different materials offer different resistance to the flow of charge. We define the **resistivity** $\rho$ of a substance so that the **resistance** $R$ of an object is directly proportional to $\rho$. Resistivity $\rho$ is an *intrinsic* property of a material, independent of its shape or size. The resistance $R$ of a uniform cylinder of length $L$, of cross-sectional area $A$, and made of a material with resistivity $\rho$, is

$$R = \frac{\rho L}{A}. \tag{7.11.1}$$

The table below gives representative values of $\rho$. The materials listed in the table are separated into categories of conductors, semiconductors, and insulators, based on broad groupings of resistivities. Conductors have the smallest resistivities, and insulators have the largest; semiconductors have intermediate resistivities. Conductors have varying but large free charge densities, whereas most charges in insulators are bound to atoms and are not free to move. Semiconductors are intermediate, having far fewer free charges than conductors, but having properties that make the number of free charges depend strongly on the type and amount of impurities in the semiconductor. These unique properties of semiconductors are put to use in modern electronics, as will be explored in later chapters.

Table 7.11.1 gives representative values of ρ. The materials listed in the table are separated into categories of conductors, semiconductors, and insulators, based on broad groupings of resistivities. Conductors have the smallest resistivities, and insulators have the largest; semiconductors have intermediate resistivities. Conductors have varying but large free charge densities, whereas most charges in insulators are bound to atoms and are not free to move. Semiconductors are intermediate,

having far fewer free charges than conductors, but having properties that make the number of free charges depend strongly on the type and amount of impurities in the semiconductor. These unique properties of semiconductors are put to use in modern electronics, as will be explored in later chapters.

*Table 7.11.1: Resistivities ρ of Various materials at 20°C*

| Material | Resistivity ρ(Ω · m) |
|---|---|
| *Conductors* | |
| Silver | $1.59x10^{-8}$ |
| Copper | $1.72x10^{-8}$ |
| Gold | $2.44x10^{-8}$ |
| Aluminum | $2.65x10^{-8}$ |
| Tungsten | $5.6x10^{-8}$ |
| Iron | $9.71x10^{-8}$ |
| Platinum | $10.6x10^{-8}$ |
| Steel | $20x10^{-8}$ |
| Lead | $22x10^{-8}$ |
| Manganin (Cu, Mn, Ni alloy) | $44x10^{-8}$ |
| Constantan (Cu, Ni alloy) | $49x10^{-8}$ |
| Mercury | $96x10^{-8}$ |
| Nichrome (Ni, Fe, Cr alloy) | $100x10^{-8}$ |
| *Semiconductors* | |
| Carbon (pure) | $3.5x10^{-5}$ |
| Carbon | $(3.5 - 6.0)x10^{-5}$ |
| Germanium (pure) | $600x10^{-3}$ |
| Germanium | $(1 - 600)x10^{-3}$ |
| Silicon (pure) | 2300 |
| Silion | 0.1–2300 |
| *Insulators* | |
| Amber | $5x10^{14}$ |
| Glass | $10^{9} - 10^{14}$ |
| Lucite | $> 10^{13}$ |
| Mica | $10^{11} - 10^{15}$ |
| Quartz (fused) | $75x10^{16}$ |
| Rubber (hard) | $10^{13} - 10^{16}$ |
| Sulfur | $10^{15}$ |
| Teflon | $> 10^{13}$ |
| Wood | $10^{8} - 10^{11}$ |

### Example 7.11.1:Calculating Resistor Diameter: A Headlight Filament

A car headlight filament is made of tungsten and has a cold resistance of **0.350Ω**. If the filament is a cylinder 4.00 cm long (it may be coiled to save space), what is its diameter?

**Strategy**

We can rearrange the equation $R = \frac{\rho L}{A}$ to find the cross-sectional area $A$ of the filament from the given information. Then its diameter can be found by assuming it has a circular cross-section.

**Solution**

The cross-sectional area, found by rearranging the expression for the resistance of a cylinder given in $R = \frac{\rho L}{A}$, is

$$A = \frac{\rho L}{R}.$$

Substituting the given values, and taking ρ from Table 7.11.1, yields

$$A = \frac{(5.6\times10^{-8}\Omega\cdot m)(4.00\times10^{-2}m)}{0.350\Omega} = 6.40 \times 10^{-9}m^2.$$

The area of a circle is related to its diameter $D$ by

$$A = \frac{\pi D^2}{4}.$$

Solving for the diameter $D$, and substituting the value found for $A$, gives

$$D = 2(\frac{A}{p})^{\frac{1}{2}} = 2(\frac{6.40\times10^{-9}m^2}{3.14})^{\frac{1}{2}} = 9.0 \times 10^{-5}m.$$

**Discussion**

The diameter is just under a tenth of a millimeter. It is quoted to only two digits, because ρ is known to only two digits.

## Temperature Variation of Resistance

The resistivity of all materials depends on temperature. Some even become superconductors (zero resistivity) at very low temperatures. (See Figure 2.) Conversely, the resistivity of conductors increases with increasing temperature. Since the atoms vibrate more rapidly and over larger distances at higher temperatures, the electrons moving through a metal make more collisions, effectively making the resistivity higher. Over relatively small temperature changes (about $100^\circ C$ or less), resistivity $\rho$ varies with temperature change $\Delta T$ as expressed in the following equation

$$\rho = \rho_0 \left(1 + \alpha\Delta T\right), \tag{7.11.2}$$

where $\rho_0$ is the original resistivity and $\alpha$ is the **temperature coefficient of resistivity.** (See the values of $\alpha$ in the table below.) For larger temperature changes, $\alpha$ may vary or a nonlinear equation may be needed to find $\rho$. Note that $\alpha$ is positive for metals, meaning their resistivity increases with temperature. Some alloys have been developed specifically to have a small temperature dependence. Manganin (which is made of copper, manganese and nickel), for example, has $\alpha$ close to zero (to three digits on the scale in the table), and so its resistivity varies only slightly with temperature. This is useful for making a temperature-independent resistance standard, for example.

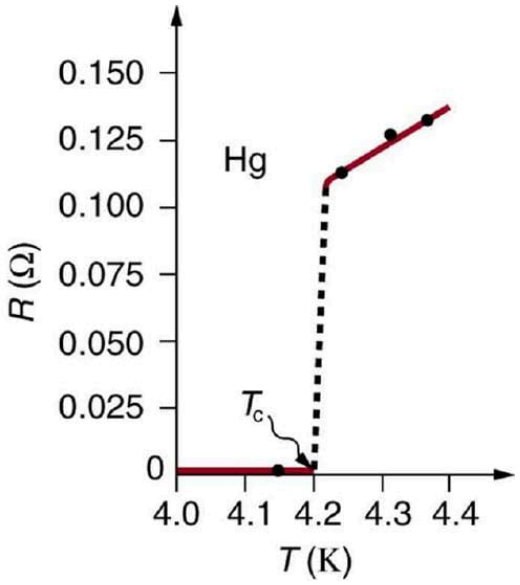



Figure 7.11.2: The resistance of a sample of mercury is zero at very low temperatures—it is a superconductor up to about 4.2 K. Above that critical temperature, its resistance makes a sudden jump and then increases nearly linearly with temperature.

*Table 7.11.2: Tempature Coefficients of Resistivity α*

| Material | Coefficient α(1/°C) |
|---|---|
| *Conductors* | |
| Silver | $3.8x10^{-3}$ |
| Copper | $3.9x10^{-3}$ |
| Gold | $3.4x10^{-3}$ |
| Aluminum | $3.9x10^{-3}$ |
| Tungsten | $4.5x10^{-3}$ |
| Iron | $5.0x10^{-3}$ |
| Platinum | $3.93x10^{-3}$ |
| Lead | $3.9x10^{-3}$ |
| Manganin (Cu, Mn, Ni alloy) | $0.000x10^{-3}$ |
| Constantan (Cu, Ni, alloy) | $0.002x10^{-3}$ |
| Mercury | $0.89x10^{-3}$ |
| Nichrome (Ni, Fe, Cr alloy) | $0.4x10^{-3}$ |
| *Semiconductors* | |
| Carbon (pure) | $-0.5x10^{-3}$ |
| Germanium (pure) | $-50x10^{-3}$ |
| Silicon (pure) | $-70x10^{-3}$ |

Note also that $\alpha$ is negative for the semiconductors listed in the table, meaning that their resistivity decreases with increasing temperature. They become better conductors at higher temperature, because increased thermal agitation increases the number of free charges available to carry current. This property of decreasing $\rho$ with temperature is also related to the type and amount of impurities present in the semiconductors.

The resistance of an object also depends on temperature, since $R_0$ is directly proportional to $\rho$. For a cylinder we know $R = \rho L/A$ , and so, if $L$ and $A$ do not change greatly with temperature, $R$ will have the same temperature dependence as $\rho$. (Examination of

the coefficients of linear expansion shows them to be about two orders of magnitude less than typical temperature coefficients of resistivity, and so the effect of temperature on $L$ and $A$ is about two orders of magnitude less than on $\rho$.) Thus,

$$R = R_0 (1 + \alpha \Delta T) \tag{7.11.3}$$

is the temperature dependence of the resistance of an object, where $R_0$ is the original resistance and $R$ is the resistance after a temperature change $\Delta T$. Numerous thermometers are based on the effect of temperature on resistance. (See Figure 3.) One of the most common is the thermistor, a semiconductor crystal with a strong temperature dependence, the resistance of which is measured to obtain its temperature. The device is small, so that it quickly comes into thermal equilibrium with the part of a person it touches.

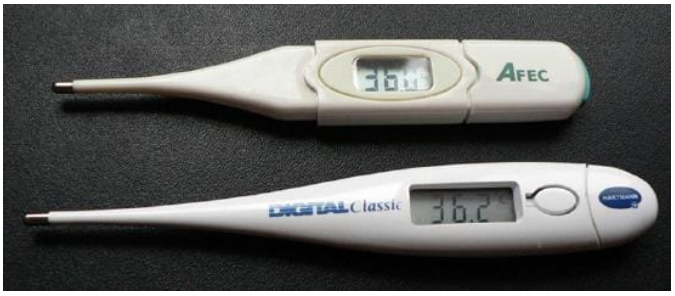



Figure 7.11.3:These familiar thermometers are based on the automated measurement of a thermistor's temperature-dependent resistance. (credit: Biol, Wikimedia Commons)

## Example 7.11.2:Calculating Resistance: Hot-Filament Resistance:

Although caution must be used in applying $\rho = \rho_0 (1 + \alpha \Delta T)$ and $R = R_0 (1 + \alpha \Delta T)$ for temperature changes greater than $100^\circ C$, for tungsten the equations work reasonably well for very large temperature changes. What, then, is the resistance of the tungsten filament in the previous example if its temperature is increased from room temperature ($20^\circ C$) to a typical operating temperature of $2850^\circ C$?

**Strategy**

This is a straightforward application of $R = R_0 (1 + \alpha \Delta T)$, since the original resistance of the filament was givent o eb $R_0 = 0.350\Omega$, and the temperature change is $\Delta T = 2830^\circ C$.

**Solution**

The hot resistance $R$ is obtained by entering known values into the above equation:

$$\begin{aligned} R &= R_0 (1 + \alpha \Delta T) \\ &= (0.350\Omega) \left[1 + \left(4.5 \times 10^{-3}/^\circ C\right)\right] \\ &= 4.8\Omega \end{aligned}$$

**Discussion**

This value is consistent with the headlight resistance example in 20.3.

## PHET EXPLORATIONS: RESISTANCE IN A WIRE

Learn about the physics of resistance in a wire. Change its resistivity, length, and area to see how they affect the wire's resistance. The sizes of the symbols in the equation change along with the diagram of a wire.




Figure 7.11.4: Resistance in a Wire

## Summary

- The resistance $R$ of a cylinder of length $L$ and cross-sectional area $A$ is $R = \frac{\rho L}{A}$, where $\rho$ is the resistivity of the material.
- Values of $\rho$ in the Table show that materials fall into three groups—*conductors, semiconductors, and insulators*.
- Temperature affects resistivity; for relatively small temperature changes $\Delta T$, resistivity is $\rho = \rho_0 \left(1 + \alpha \Delta T\right), where \backslash (\rho_0$ is the original resistivity and $\alpha$ is the temperature coefficient of resistivity.
- The Table gives values for $\alpha$, the temperature coefficient of resistivity.
- The resistance $R$ of an object also varies with temperature: $R = R_0 \left(1 + \alpha \Delta T\right)$, where $R_0$ is the original resistance, and $R$ is the resistance after the temperature change.

## Footnotes

1 Values depend strongly on amounts and types of impurities

2 Values at 20°C.

## Glossary

**resistivity**
: an intrinsic property of a material, independent of its shape or size, directly proportional to the resistance, denoted by *ρ*

**temperature coefficient of resistivity**
: an empirical quantity, denoted by *α*, which describes the change in resistance or resistivity of a material with temperature

This page titled 7.11: Resistance and Resistivity is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.3: Resistance and Resistivity by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.12: 20.4 Electric Power and Energy

**Learning Objectives**

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

- Calculate the power dissipated by a resistor and power supplied by a power supply.
- Calculate the cost of electricity under various circumstances.

## Power in Electric Circuits

Power is associated by many people with electricity. Knowing that power is the rate of energy use or energy conversion, what is the expression for **electric power**? Power transmission lines might come to mind. We also think of lightbulbs in terms of their power ratings in watts. Let us compare a 25-W bulb with a 60-W bulb (Figure $7.12.1a$.) Since both operate on the same voltage, the 60-W bulb must draw more current to have a greater power rating. Thus the 60-W bulb's resistance must be lower than that of a 25-W bulb. If we increase voltage, we also increase power. For example, when a 25-W bulb that is designed to operate on 120 V is connected to 240 V, it briefly glows very brightly and then burns out. Precisely how are voltage, current, and resistance related to electric power?

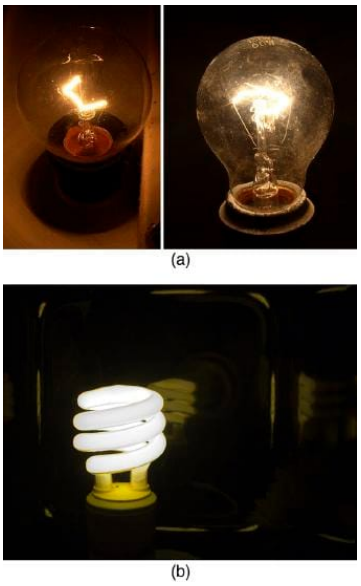



***Figure*** $7.12.1$**:** *(a) Which of these lightbulbs, the 25-W bulb (upper left) or the 60-W bulb (upper right), has the higher resistance? Which draws more current? Which uses the most energy? Can you tell from the color that the 25-W filament is cooler? Is the brighter bulb a different color and if so why? (credits: Dickbauch, Wikimedia Commons; Greg Westfall, Flickr) (b) This compact fluorescent light (CFL) puts out the same intensity of light as the 60-W bulb, but at 1/4 to 1/10 the input power. (credit: dbgg1979, Flickr)*

Electric energy depends on both the voltage involved and the charge moved. This is expressed most simply as $PE = qV$, where $q$ is the charge moved and $V$ is the voltage (or more precisely, the potential difference the charge moves through). Power is the rate at which energy is moved, and so electric power is

$$P = \frac{PE}{t} = \frac{qV}{t}. \tag{7.12.1}$$

Recognizing that current is $I = q/t$ (note that $\Delta t = t$ here), the expression for power becomes

$$P = IV. \tag{7.12.2}$$

Electric power ($P$) is simply the product of current times voltage. Power has familiar units of watts. Since the SI unit for potential energy (PE) is the joule, power has units of joules per second, or watts. Thus, $1A \cdot V = 1W$. For example, cars often have one or more auxiliary power outlets with which you can charge a cell phone or other electronic devices. These outlets may be rated at 20 A, so that the circuit can deliver a maximum power $P = IV = (20A)(12V) = 240W$. In some applications, electric power may be expressed as volt-amperes or even kilovolt-amperes ( $1kA \cdot V = 1kW$ ).

To see the relationship of power to resistance, we combine Ohm's law with $P = IV$. Substituting Ohm Law ($I = V/R$) into Equation 7.12.2 gives

$$P = (V/R)\,V = V^2/R. \tag{7.12.3}$$

Similarly, substituting $V = IR$ gives

$$P = (IR) = I^2R. \tag{7.12.4}$$

Three expressions for electric power are listed together here for convenience:

$$P = IV \tag{7.12.5}$$

$$P = \frac{V^2}{R} \tag{7.12.6}$$

$$P = I^2R. \tag{7.12.7}$$

Note that the first equation is always valid, whereas the other two can be used only for resistors. In a simple circuit, with one voltage source and a single resistor, the power supplied by the voltage source and that dissipated by the resistor are identical. (In more complicated circuits, $P$ can be the power dissipated by a single device and not the total power in the circuit.)

Different insights can be gained from the three different expressions for electric power. For example, $P = V^2/R$ implies that the lower the resistance connected to a given voltage source, the greater the power delivered. Furthermore, since voltage is squared in $P = V^2/R$, the effect of applying a higher voltage is perhaps greater than expected. Thus, when the voltage is doubled to a 25-W bulb, its power nearly quadruples to about 100 W, burning it out. If the bulb's resistance remained constant, its power would be exactly 100 W, but at the higher temperature its resistance is higher, too.

## Example 7.12.1: Calculating Power Dissipation and Current: Hot and Cold Power

(a) Consider the examples given in 20.3 and 20.4. Then find the power dissipated by the car headlight in these examples, both when it is hot and when it is cold.

**Strategy**

For the hot headlight, we know voltage and current, so we can use $P = IV$ to find the power. For the cold headlight, we know the voltage and resistance, so we can use $P = V^2/R$ to find the power.

**Solution**

Entering the known values of current and voltage for the hot headlight, we obtain

$$P = IV = (2.50A)(12.0V) = 30.0W. \tag{7.12.8}$$

The cold resistance was $0.350\Omega$, and so the power it uses when first switched on is

$$P = \frac{V^2}{R} = \frac{(12.0V)^2}{0.350\Omega} = 411W. \tag{7.12.9}$$

**Discussion**

The 30 W dissipated by the hot headlight is typical. But the 411 W when cold is surprisingly higher. The initial power quickly decreases as the bulb's temperature increases and its resistance increases.

(b) What current does it draw when cold?

**Solution**

The current when the bulb is cold can be found several different ways. We rearrange one of the power equations, $P = I^2R$, and enter known values, obtaining

$$I = \sqrt{\frac{P}{R}} = \sqrt{\frac{411W}{0.350\Omega}} = 34.3A. \tag{7.12.10}$$

**Discussion**

The cold current is remarkably higher than the steady-state value of 2.50 A, but the current will quickly decline to that value as the bulb's temperature increases. Most fuses and circuit breakers (used to limit the current in a circuit) are designed to tolerate very high currents briefly as a device comes on. In some cases, such as with electric motors, the current remains high for several seconds, necessitating special "slow blow" fuses.

## The Cost of Electricity

The more electric appliances you use and the longer they are left on, the higher your electric bill. This familiar fact is based on the relationship between energy and power. You pay for the energy used. Since $P = E/t$, we see that

$$E = Pt \tag{7.12.11}$$

is the energy used by a device using power $P$ for a time interval $t$. For example, the more lightbulbs burning, the greater $P$ used; the longer they are on, the greater $t$ is. The energy unit on electric bills is the killowatt-hour ( $kW \cdot h$ ), consistent with the relationship $E = Pt$. It is easy to estimate the cost of operating electric appliances if you have some idea of their power consumption rate in watts or kilowatts, the time they are on in hours, and the cost per kilowatt-hour for your electric utility. Kilowatt-hours, like all other specialized energy units such as food calories, can be converted to joules. You can prove to yourself that $1kW \cdot h = 3.6 \times 10^6 J$.

The electric energy ( $E$ ) used can be reduced either by reducing the time of use or by reducing the power consumption of that appliance or fixture. This will not only reduce the cost, but it will also result in a reduced impact on the environment. Improvements to lighting are some of the fastest ways to reduce the electrical energy used in a home or business. About 20% of a home's use of energy goes to lighting, while the number for commercial establishments is closer to 40%. Fluorescent lights are about four times more efficient than incandescent lights—this is true for both the long tubes and the compact fluorescent lights (CFL). (See Figure 1b.) Thus, a 60-W incandescent bulb can be replaced by a 15-W CFL, which has the same brightness and color. CFLs have a bent tube inside a globe or a spiral-shaped tube, all connected to a standard screw-in base that fits standard incandescent light sockets. (Original problems with color, flicker, shape, and high initial investment for CFLs have been addressed in recent years.) The heat transfer from these CFLs is less, and they last up to 10 times longer. The significance of an investment in such bulbs is addressed in the next example. New white LED lights (which are clusters of small LED bulbs) are even more efficient (twice that of CFLs) and last 5 times longer than CFLs. However, their cost is still high.

### Making Connections: Energy, Power, and Time

The relationship $E = Pt$ is one that you will find useful in many different contexts. The energy your body uses in exercise is related to the power level and duration of your activity, for example. The amount of heating by a power source is related to the power level and time it is applied. Even the radiation dose of an X-ray image is related to the power and time of exposure.

### Example 7.12.2: Calculating the Cost Effectiveness of Compact Fluorescent Lights (CFL)

(a) If the cost of electricity in your area is 12 cents per kWh, what is the total cost (capital plus operation) of using a 60-W incandescent bulb for 1000 hours (the lifetime of that bulb) if the bulb cost 25 cents?

**Strategy**

To find the operating cost, we first find the energy used in kilowatt-hours and then multiply by the cost per kilowatt-hour.

**Solution**

The energy used in kilowatt-hours is found by entering the power and time into the expression for energy:

$$E = Pt = (60W)\,(1000h) = 60,000W \cdot h. \tag{7.12.12}$$

In kilowatt-hours, this is \]E = 60.0 kW \cdot h.\] Now the electricity cost is

$$cost = (60.0kW \cdot h)\,(\$0.12/kW \cdot h) = \$7.20. \tag{7.12.13}$$

The total cost will be $7.20 for 1000 hours (about one-half year at 5 hours per day).

(b) If we replace this bulb with a compact fluorescent light that provides the same light output, but at one-quarter the wattage, and which costs $1.50 but lasts 10 times longer (10,000 hours), what will that total cost be?

**Solution**

Since the CFL uses only 15 W and not 60 W, the electricity cost will be $7.20/4 = $1.80. The CFL will last 10 times longer than the incandescent, so that the investment cost will be 1/10 of the bulb cost for that time period of use, or 0.1($1.50) = $0.15. Therefore, the total cost will be $1.95 for 1000 hours.

**Discussion**

Therefore, it is much cheaper to use the CFLs, even though the initial investment is higher. The increased cost of labor that a business must include for replacing the incandescent bulbs more often has not been figured in here.

### Making Connections: Take-Home Experiment—Electrical Energy Use Inventory

1) Make a list of the power ratings on a range of appliances in your home or room. Explain why something like a toaster has a higher rating than a digital clock. Estimate the energy consumed by these appliances in an average day (by estimating their time of use). Some appliances might only state the operating current. If the household voltage is 120 V, then use $P = IV$.

2) Check out the total wattage used in the rest rooms of your school's floor or building. (You might need to assume the long fluorescent lights in use are rated at 32 W.) Suppose that the building was closed all weekend and that these lights were left on from 6 p.m. Friday until 8 a.m. Monday. What would this oversight cost? How about for an entire year of weekends?

## Summary

- Electric power $P$ is the rate (in watts) that energy is supplied by a source or dissipated by a device.
- Three expressions for electrical power are

$$P = IV, \tag{7.12.14}$$

$$P = \frac{V^2}{R}, \tag{7.12.15}$$

and

$$P = I^2R. \tag{7.12.16}$$

- The energy used by a device with a power $P$ over a time $t$ is $E = Pt$.

## Glossary

**electric power**
: the rate at which electrical energy is supplied by a source or dissipated by a device; it is the product of current times voltage

This page titled 7.12: 20.4 Electric Power and Energy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.4: 20.4 Electric Power and Energy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.13: Alternating Current versus Direct Current

**Learning Objectives**

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

- Explain the differences and similarities between AC and DC current.
- Calculate rms voltage, current, and average power.
- Explain why AC current is used for power transmission.

## Alternating Current

Most of the examples dealt with so far, and particularly those utilizing batteries, have constant voltage sources. Once the current is established, it is thus also a constant. **Direct current** (DC) is the flow of electric charge in only one direction. It is the steady state of a constant-voltage circuit. Most well-known applications, however, use a time-varying voltage source. **Alternating current** (AC) is the flow of electric charge that periodically reverses direction. If the source varies periodically, particularly sinusoidally, the circuit is known as an alternating current circuit. Examples include the commercial and residential power that serves so many of our needs. Figure 7.13.1 shows graphs of voltage and current versus time for typical DC and AC power. The AC voltages and frequencies commonly used in homes and businesses vary around the world.

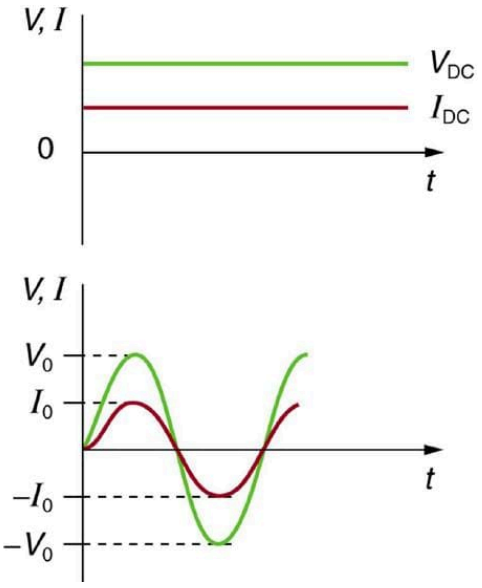



Figure 7.13.1: (a) DC voltage and current are constant in time, once the current is established. (b) A graph of voltage and current versus time for 60-Hz AC power. The voltage and current are sinusoidal and are in phase for a simple resistance circuit. The frequencies and peak voltages of AC sources differ greatly.

Figure 7.13.2 shows a schematic of a simple circuit with an AC voltage source. The voltage between the terminals fluctuates as shown, with the **AC voltage** given by

$$V = V_0 sin 2\pi ft, \tag{7.13.1}$$

where $V$ is the voltage at time $t$, $V_0$, $V_0$ is the peak voltage, and $f$ is the frequency in hertz. For this simple resistance circuit, $I = V/R$, and so the **AC current** is

$$I = I_0 sin 2\pi ft, \tag{7.13.2}$$

where $I$ is the current at time $t$, and $I_0 = V_0/R$ is the peak current. For this example, the voltage and current are said to be in phase, as seen in Figure 7.13.1$b$.

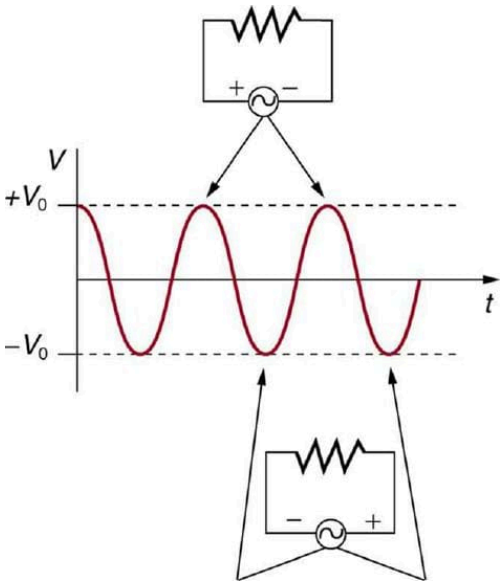



Figure 7.13.2: The potential difference $V$ between the terminals of an AC voltage source fluctuates as shown. The mathematical expression for $V$ is given by $V = V_0 sin 2\pi ft$.

Current in the resistor alternates back and forth just like the driving voltage, since $I = V/R$. If the resistor is a fluorescent light bulb, for example, it brightens and dims 120 times per second as the current repeatedly goes through zero. A 120-Hz flicker is too rapid for your eyes to detect, but if you wave your hand back and forth between your face and a fluorescent light, you will see a stroboscopic effect evidencing AC. The fact that the light output fluctuates means that the power is fluctuating. The power supplied is $P = IV$. Using the expressions for $I$ and $V$ above, we see that the time dependence of power is $P = I_0 V_0 sin^2 2\pi ft$, as shown in Figure 7.13.3.

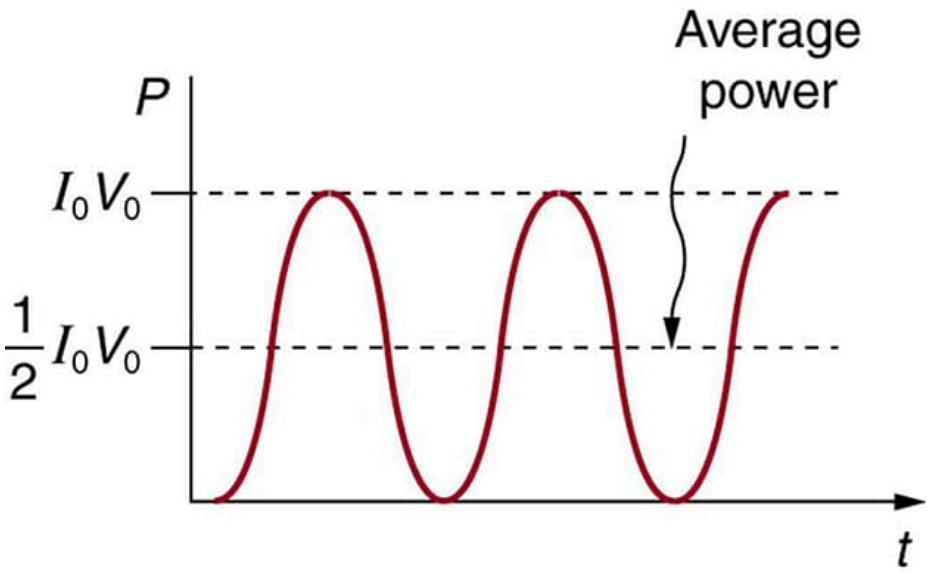



Figure 7.13.3: AC power as a function of time. Since the voltage and current are in phase here, their product is non-negative and fluctuates between zero and $I_0 V_0$. Average power is $(1/2)\, I_0 V_0$.

### Making Connections: Take-Home Experiment—AC/DC Lights

Wave your hand back and forth between your face and a fluorescent light bulb. Do you observe the same thing with the headlights on your car? Explain what you observe. *Warning: Do not look directly at very bright light.*

We are most often concerned with average power rather than its fluctuations—that 60-W light bulb in your desk lamp has an average power consumption of 60 W, for example. As illustrated in Figure 3, the average power $P_{ave}$ is

$$P_{ave} = \frac{1}{2} I_0 V_0. \tag{7.13.3}$$

This is evident from the graph, since the areas above and below the $(1/2)\, I_0 V_0$ line are equal, but it can also be proven using trigonometric identities. Similarly, we define an average or **rms current** $I_{rms}$ and average or **rms voltage** $V_{rms}$ to be, respectively,

$$I_{rms} = \frac{I_0}{\sqrt{2}} \tag{7.13.4}$$

and

$$V_{rms} = \frac{V_0}{\sqrt{2}}. \tag{7.13.5}$$

where rms stands for root mean square, a particular kind of average. In general, to obtain a root mean square, the particular quantity is squared, its mean (or average) is found, and the square root is taken. This is useful for AC, since the average value is zero. Now,

$$P_{ave} = I_{rms} V_{rms}, \tag{7.13.6}$$

which gives

$$P_{ave} = \frac{I_0}{\sqrt{2}} \cdot \frac{V_0}{\sqrt{2}} = \frac{1}{2} I_0 V_0, \tag{7.13.7}$$

as stated above. It is standard practice to quote $I_{rms}$, $V_{rms}$, and $P_{ave}$ rather than the peak values. For example, most household electricity is 120 V AC, which means that $V_{rms}$ is 120 V. The common 10-A circuit breaker will interrupt a sustained $I_{rms}$ greater than 10 A. Your 1.0-kW microwave oven consumes $P_{ave} = 1.0kW$, and so on. You can think of these rms and average values as the equivalent DC values for a simple resistive circuit.

To summarize, when dealing with AC, Ohm's law and the equations for power are completely analogous to those for DC, but rms and average values are used for AC. Thus, for AC, Ohm's law is written

$$I_{rms} = \frac{V_{rms}}{R}. \tag{7.13.8}$$

The various expressions for AC power $P_{ave}$ are

$$P_{ave} = I_{rms} V_{rms}, \tag{7.13.9}$$

$$P_{ave} = \frac{V_{rms}^2}{R}, \tag{7.13.10}$$

and

$$P_{ave} = I_{rms}^2 R. \tag{7.13.11}$$

## Example 7.13.1: Peak Voltage and Power for AC

(a) What is the value of the peak voltage for 120-V AC power?

**Strategy**

We are told that $V_{rms}$ is 120V and $P_{ave}$ is 60.0 W. We can use $V_{rms} = \frac{V_0}{\sqrt{2}}$ to find the peak voltage, and we can manipulate the definition of power to find the peak power from the given average power.

**Solution**

Solving the equation $V_{rms} = \frac{V_0}{\sqrt{2}}$ for the peak voltage $V_0$ and substituting the known value for $V_{rms}$ gives

$$V_0 = \sqrt{2} V_{rms} = 1.414\,(120V) = 170V. \tag{7.13.12}$$

**Discussion**

This means that the AC voltage swings from 170 V to $-170V$ and back 60 times every second. An equivalent DC voltage is a constant 120 V.

(b) What is the peak power consumption rate of a 60.0-W AC light bulb?

**Solution**

Peak power is peak current times peak voltage. Thus,

$$P_0 = I_0 V_0 = 2\left(\frac{1}{2} I_0 V_0\right) = 2P_{ave}. \tag{7.13.13}$$

We know the average power is 60.0 W, and so

$$P_0 = 2\,(60.0W) = 120W. \tag{7.13.14}$$

**Discussion**

So the power swings from zero to 120 W one hundred twenty times per second (twice each cycle), and the power averages 60 W.

## Why Use AC for Power Distribution?

Most large power-distribution systems are AC. Moreover, the power is transmitted at much higher voltages than the 120-V AC (240 V in most parts of the world) we use in homes and on the job. Economies of scale make it cheaper to build a few very large electric power-generation plants than to build numerous small ones. This necessitates sending power long distances, and it is obviously important that energy losses en route be minimized. High voltages can be transmitted with much smaller power losses than low voltages, as we shall see. (See Figure 4.) For safety reasons, the voltage at the user is reduced to familiar values. The crucial factor is that it is much easier to increase and decrease AC voltages than DC, so AC is used in most large power distribution systems.




Figure 7.13.4: Power is distributed over large distances at high voltage to reduce power loss in the transmission lines. The voltages generated at the power plant are stepped up by passive devices called transformers (see Transformers) to 330,000 volts (or more in some places worldwide). At the point of use, the transformers reduce the voltage

### Example 7.13.2: Power Losses are Less for High-Voltage Transmission

(a) What current is needed to transmit 100 MW of power at 200 kV?

**Strategy**

We are given $P_{ave} = 100MW$, $V_{rms} = 200kV$, and the resistance of the lines is $R = 1.00\Omega$. Using these givens, we can find the current flowing (from $P = IV$ ) and then the power dissipated in the lines ($P = I^2R$), and we take the ratio to the total power transmitted.

**Solution**

To find the current, we rearrange the relationship $P_{ave} = I_{rms}V_{rms}$ and substitute known values. This gives

$$I_{rms} = \frac{P_{ave}}{V_{rms}} = \frac{100 \times 10^6 W}{200 \times 10^3 V} = 500A. \tag{7.13.15}$$

(b) What is the power dissipated by the transmission lines if they have a resistance of $1.00\Omega$?

**Solution**

Knowing the current and given the resistance of the lines, the power dissipated in them is found from $P_{ave} = I_{rms}^2 R$.. Substituting the known value gives

$$P_{ave} = I_{rms}^2 R = (500A)^2 (1.00\Omega) = 250kW. \tag{7.13.16}$$

(c) What percentage of the power is lost in the transmission lines?

**Solution**

The percent loss is the ratio of this lost power to the total or input power, multiplied by 100:

$$\%\text{loss} = \frac{250\,\text{kW}}{100\,\text{MW}} \times 100 = 0.250\%. \tag{7.13.17}$$

**Discussion**

One-fourth of a percent is an acceptable loss. Note that if 100 MW of power had been transmitted at 25 kV, then a current of 4000 A would have been needed. This would result in a power loss in the lines of 16.0 MW, or 16.0% rather than 0.250%. The lower the voltage, the more current is needed, and the greater the power loss in the fixed-resistance transmission lines. Of course, lower-resistance lines can be built, but this requires larger and more expensive wires. If superconducting lines could be economically produced, there would be no loss in the transmission lines at all. But, as we shall see in a later chapter, there is a limit to current in superconductors, too. In short, high voltages are more economical for transmitting power, and AC voltage is much easier to raise and lower, so that AC is used in most large-scale power distribution systems.

It is widely recognized that high voltages pose greater hazards than low voltages. But, in fact, some high voltages, such as those associated with common static electricity, can be harmless. So it is not voltage alone that determines a hazard. It is not so widely recognized that AC shocks are often more harmful than similar DC shocks. Thomas Edison thought that AC shocks were more harmful and set up a DC power-distribution system in New York City in the late 1800s. There were bitter fights, in particular between Edison and George Westinghouse and Nikola Tesla, who were advocating the use of AC in early power-distribution systems. AC has prevailed largely due to transformers and lower power losses with high-voltage transmission.

PHET EXPLORATIONS: GENERATOR

Generate electricity with a bar magnet! Discover the physics behind the phenomena by exploring magnets and how you can use them to make a bulb light.

***Figure*** 7.13.5***:*** Generator

## Summary

- Direct current (DC) is the flow of electric current in only one direction. It refers to systems where the source voltage is

constant.

- The voltage source of an alternating current (AC) system puts out $V = V_0 \sin 2\pi ft$, where $V$ is the voltage at time $t$, $V_0$ is the peak voltage, and $f$ is the frequency in hertz.
- In a simple circuit, $I = V/R$ and AC current is $I = I_0 \sin 2\pi ft$, where $I$ is the current at time $t$, and $I_0 = V_0/R$ is the peak current.
- The average AC power is $P_{ave} = \frac{1}{2} I_0 V_0$.
- Average (rms) current $I_{rms}$ and average (rms) voltage $V_{rms}$ and $I_{rms} = \frac{I_0}{\sqrt{2}}$ and $V_{rms} = \frac{V_0}{\sqrt{2}}$, where rms stands for root mean square.
- Thus, $P_{ave} = I_{rms} V_{rms}$.
- Ohm's law for AC is $I_{rms} = \frac{V_{rms}}{R}$.
- Expressions for the average power of an AC circuit are $P_{ave} = I_{rms} V_{rms}$, $P_{ave} = \frac{V_{rms}^2}{R}$, and $P_{ave} = I_{rms}^2 R$, analogous to the expressions for DC circuits.

# Glossary

**direct current**
: (DC) the flow of electric charge in only one direction

**alternating current**
: (AC) the flow of electric charge that periodically reverses direction

**AC voltage**
: voltage that fluctuates sinusoidally with time, expressed as $V = V_0 \sin 2\pi ft$, where $V$ is the voltage at time $t$, $V_0$ is the peak voltage, and $f$ is the frequency in hertz

**AC current**
: current that fluctuates sinusoidally with time, expressed as $I = I_0 \sin 2\pi ft$, where $I$ is the current at time $t$, $I_0$ is the peak current, and $f$ is the frequency in hertz

**rms current**
: the root mean square of the current, $I_{rms} = I_0/\sqrt{2}$, where $I_0$ is the peak current, in an AC system

**rms voltage**
: the root mean square of the voltage, $V_{rms} = V_0/\sqrt{2}$, where $V_0$ is the peak voltage, in an AC system

This page titled 7.13: Alternating Current versus Direct Current is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.5: Alternating Current versus Direct Current by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.14: Electric Hazards and the Human Body

**Learning Objectives**

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

- Define thermal hazard, shock hazard, and short circuit.
- Explain what effects various levels of current have on the human body.

There are two known hazards of electricity—thermal and shock. A **thermal hazard** is one where excessive electric power causes undesired thermal effects, such as starting a fire in the wall of a house. A **shock hazard** occurs when electric current passes through a person. Shocks range in severity from painful, but otherwise harmless, to heart-stopping lethality. This section considers these hazards and the various factors affecting them in a quantitative manner. Electrical Safety: Systems and Devices will consider systems and devices for preventing electrical hazards.

## Thermal Hazards

Electric power causes undesired heating effects whenever electric energy is converted to thermal energy at a rate faster than it can be safely dissipated. A classic example of this is the **short circuit**, a low-resistance path between terminals of a voltage source. An example of a short circuit is shown in Figure 1. Insulation on wires leading to an appliance has worn through, allowing the two wires to come into contact. Such an undesired contact with a high voltage is called a *short*. Since the resistance of the short, $r$, is very small, the power dissipated in the short, $P = V^2/r$, is very large. for example, if $V$ is 120 V and $r$ is $0.100\Omega$, then the power is 144kW, *much* greater than that used by a typical household appliance. Thermal energy delivered at this rate will very quickly raise the temperature of surrounding materials, melting or perhaps igniting them.

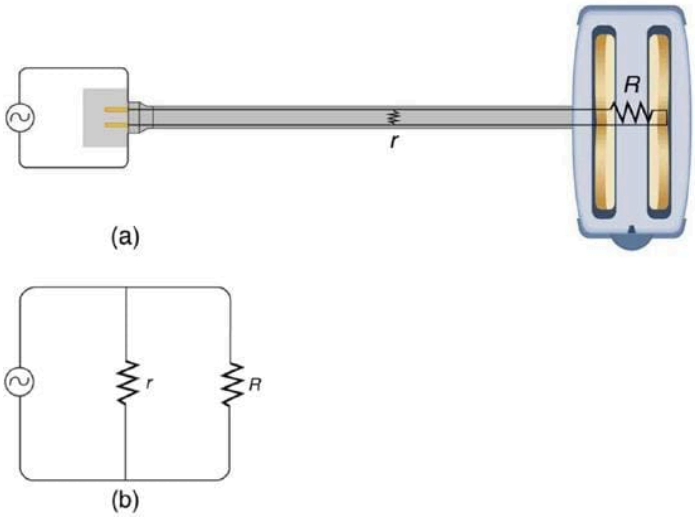



Figure 7.14.1: A short circuit is an undesired low-resistance path across a voltage source. (a) Worn insulation on the wires of a toaster allow them to come into contact with a low resistance $r$. Since $P = V^2/r$, thermal power is created so rapidly that the cord melts or burns. (b) A schematic of the short circuit.

One particularly insidious aspect of a short circuit is that its resistance may actually be decreased due to the increase in temperature. This can happen if the short creates ionization. These charged atoms and molecules are free to move and, thus, lower the resistance $r$. Since $P = V^2/r$, the power dissipated in the short rises, possibly causing more ionization, more power, and so on. High voltages, such as the 480-V AC used in some industrial applications, lend themselves to this hazard, because higher voltages create higher initial power production in a short.

Another serious, but less dramatic, thermal hazard occurs when wires supplying power to a user are overloaded with too great a current. As discussed in the previous section, the power dissipated in the supply wires is $P = I^2R_W$, where $R_W$ is the resistance of

the wires and $I$ the current flowing through them. If either $I$ or $R_W$ is too large, the wires overheat. For example, a worn appliance cord (with some of its braided wires broken) may have $R_W = 2.00\Omega$ rather than the $0.100\Omega$ it should be. If 10.0 A of current passes through the cord, then $P = I^2 R_W = 200W$ is dissipated in the cord—much more than is safe. Similarly, if a wire with a $0.100 - \Omega$ resistance is meant to carry a few amps, but is instead carrying 100 A, it will severely overheat. The power dissipated in the wire will in that case be $P = 1000W$. Fuses and circuit breakers are used to limit excessive currents. (See Figure 2 and Figure 3.) Each device opens the circuit automatically when a sustained current exceeds safe limits.

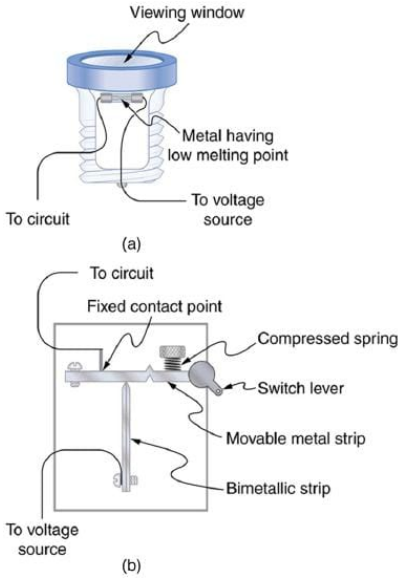



Figure 7.14.2: (a) A fuse has a metal strip with a low melting point that, when overheated by an excessive current, permanently breaks the connection of a circuit to a voltage source. (b) A circuit breaker is an automatic but restorable electric switch. The one shown here has a bimetallic strip that bends to the right and into the notch if overheated. The spring then forces the metal strip downward, breaking the electrical connection at the points.

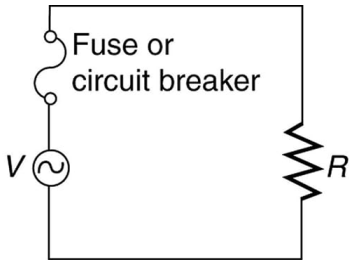



Figure 7.14.3: Schematic of a circuit with a fuse or circuit breaker in it. Fuses and circuit breakers act like automatic switches that open when sustained current exceeds desired limits.

Fuses and circuit breakers for typical household voltages and currents are relatively simple to produce, but those for large voltages and currents experience special problems. For example, when a circuit breaker tries to interrupt the flow of high-voltage electricity, a spark can jump across its points that ionizes the air in the gap and allows the current to continue flowing. Large circuit breakers found in power-distribution systems employ insulating gas and even use jets of gas to blow out such sparks. Here AC is safer than DC, since AC current goes through zero 120 times per second, giving a quick opportunity to extinguish these arcs.

## Shock Hazards

Electrical currents through people produce tremendously varied effects. An electrical current can be used to block back pain. The possibility of using electrical current to stimulate muscle action in paralyzed limbs, perhaps allowing paraplegics to walk, is under study. TV dramatizations in which electrical shocks are used to bring a heart attack victim out of ventricular fibrillation (a massively irregular, often fatal, beating of the heart) are more than common. Yet most electrical shock fatalities occur because a current put the heart into fibrillation. A pacemaker uses electrical shocks to stimulate the heart to beat properly. Some fatal shocks do not produce burns, but warts can be safely burned off with electric current (though freezing using liquid nitrogen is now more common). Of course, there are consistent explanations for these disparate effects. The major factors upon which the effects of electrical shock depend are

1. The amount current $I$
2. The path taken by the current
3. The duration of the shock
4. The frequency $f$ of the current ($f = 0$ for DC)

The table below gives the effects of electrical shocks as a function of current for a typical accidental shock. The effects are for a shock that passes through the trunk of the body, has a duration of 1 s, and is caused by 60-Hz power.

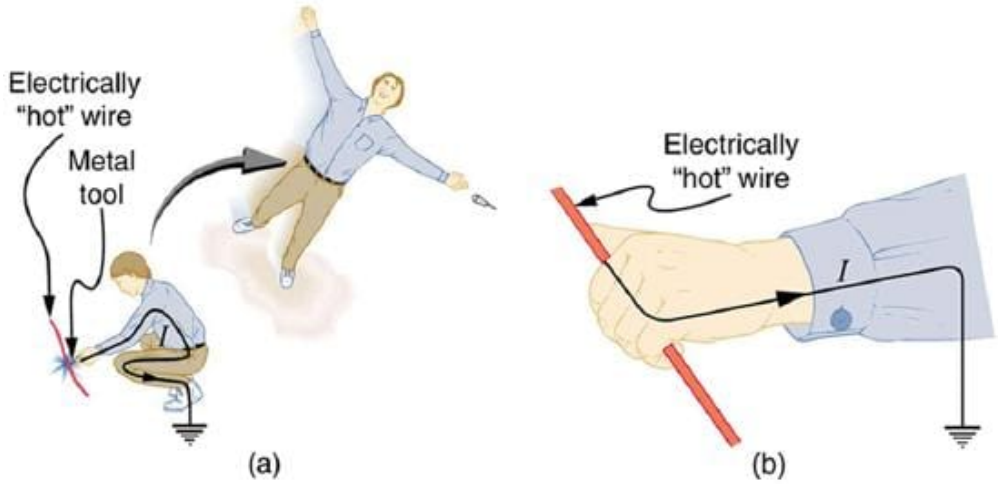



Figure 7.14.4: An electric current can cause muscular contractions with varying effects. (a) The victim is "thrown" backward by involuntary muscle contractions that extend the legs and torso. (b) The victim can't let go of the wire that is stimulating all the muscles in the hand. Those that close the fingers are stronger than those that open them.

| **Current (mA)** | **Effect** |
|---|---|
| 1 | Threshold of sensation |
| 5 | Maximum harmless current |
| 10-20 | Onset of sustained muscular contraction; cannot let go for duration of shock; contraction of chest muscles may stop breathing during shock |
| 50 | Onset of pain |
| 100-300+ | Ventricular fibrillation possible; often fatal |
| 300 | Onset of burns depending on concentration of current |
| 6000 (6A) | Onset of sustained ventricular contraction and respiratory paralysis; both cease when shock ends; heartbeat may return to normal; used to defibrillate the heart |

Effects of Electrical Shock as a Function of Current

Our bodies are relatively good conductors due to the water in our bodies. Given that larger currents will flow through sections with lower resistance (to be further discussed in the next chapter), electric currents preferentially flow through paths in the human body that have a minimum resistance in a direct path to earth. The earth is a natural electron sink. Wearing insulating shoes, a requirement in many professions, prohibits a pathway for electrons by providing a large resistance in that path. Whenever working with high-power tools (drills), or in risky situations, ensure that you do not provide a pathway for current flow (especially through the heart).

Very small currents pass harmlessly and unfelt through the body. This happens to you regularly without your knowledge. The threshold of sensation is only 1 mA and, although unpleasant, shocks are apparently harmless for currents less than 5 mA. A great number of safety rules take the 5-mA value for the maximum allowed shock. At 10 to 20 mA and above, the current can stimulate sustained muscular contractions much as regular nerve impulses do. People sometimes say they were knocked across the room by a shock, but what really happened was that certain muscles contracted, propelling them in a manner not of their own choosing. (See Figure 4a.) More frightening, and potentially more dangerous, is the "can't let go" effect illustrated in Figure 4b.

The muscles that close the fingers are stronger than those that open them, so the hand closes involuntarily on the wire shocking it. This can prolong the shock indefinitely. It can also be a danger to a person trying to rescue the victim, because the rescuer's hand may close about the victim's wrist. Usually the best way to help the victim is to give the fist a hard knock/blow/jar with an insulator or to throw an insulator at the fist. Modern electric fences, used in animal enclosures, are now pulsed on and off to allow people who touch them to get free, rendering them less lethal than in the past.

Greater currents may affect the heart. Its electrical patterns can be disrupted, so that it beats irregularly and ineffectively in a condition called "ventricular fibrillation." This condition often lingers after the shock and is fatal due to a lack of blood circulation. The threshold for ventricular fibrillation is between 100 and 300 mA. At about 300 mA and above, the shock can cause burns, depending on the concentration of current—the more concentrated, the greater the likelihood of burns.

Very large currents cause the heart and diaphragm to contract for the duration of the shock. Both the heart and breathing stop. Interestingly, both often return to normal following the shock. The electrical patterns on the heart are completely erased in a manner that the heart can start afresh with normal beating, as opposed to the permanent disruption caused by smaller currents that can put the heart into ventricular fibrillation. The latter is something like scribbling on a blackboard, whereas the former completely erases it. TV dramatizations of electric shock used to bring a heart attack victim out of ventricular fibrillation also show large paddles. These are used to spread out current passed through the victim to reduce the likelihood of burns.

Current is the major factor determining shock severity (given that other conditions such as path, duration, and frequency are fixed, such as in the table and preceding discussion). A larger voltage is more hazardous, but since $I = V/R$, the severity of the shock depends on the combination of voltage and resistance. For example, a person with dry skin has a resistance of about $200k\Omega$. If he comes into contact with 120-V AC, a current $I = (120V)/(200k\Omega) = .6mA$ passes harmlessly through him. The same person soaking wet may have a resistance of $10.0k\Omega$ and the same 120 V will produce a current of 12 mA—above the "can't let go" threshold and potentially dangerous.

Most of the body's resistance is in its dry skin. When wet, salts go into ion form, lowering the resistance significantly. The interior of the body has a much lower resistance than dry skin because of all the ionic solutions and fluids it contains. If skin resistance is bypassed, such as by an intravenous infusion, a catheter, or exposed pacemaker leads, a person is rendered **microshock sensitive**. In this condition, currents about 1/1000 those listed in the table above produce similar effects. During open-heart surgery, currents as small as $20\mu A$ can be used to still the heart. Stringent electrical safety requirements in hospitals, particularly in surgery and intensive care, are related to the doubly disadvantaged microshock-sensitive patient. The break in the skin has reduced his resistance, and so the same voltage causes a greater current, and a much smaller current has a greater effect.

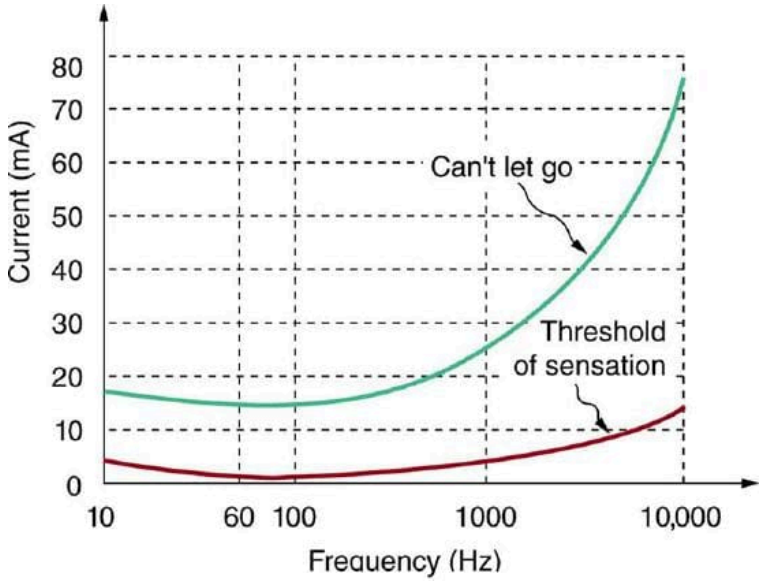



Figure 7.14.5: Graph of average values for the threshold of sensation and the "can't let go" current as a function of frequency. The lower the value, the more sensitive the body is at that frequency.

Factors other than current that affect the severity of a shock are its path, duration, and AC frequency. Path has obvious consequences. For example, the heart is unaffected by an electric shock through the brain, such as may be used to treat manic depression. And it is a general truth that the longer the duration of a shock, the greater its effects. Figure 5 presents a graph that illustrates the effects of frequency on a shock. The curves show the minimum current for two different effects, as a function of

frequency. The lower the current needed, the more sensitive the body is at that frequency. Ironically, the body is most sensitive to frequencies near the 50- or 60-Hz frequencies in common use. The body is slightly less sensitive for DC ($f = 0$), mildly confirming Edison's claims that AC presents a greater hazard. At higher and higher frequencies, the body becomes progressively less sensitive to any effects that involve nerves. This is related to the maximum rates at which nerves can fire or be stimulated. At very high frequencies, electrical current travels only on the surface of a person. Thus a wart can be burned off with very high frequency current without causing the heart to stop. (Do not try this at home with 60-Hz AC!) Some of the spectacular demonstrations of electricity, in which high-voltage arcs are passed through the air and over people's bodies, employ high frequencies and low currents. (See Figure 6.) Electrical safety safety devices and techniques are discussed in detail in Electrical Safety: Systems and Devices.

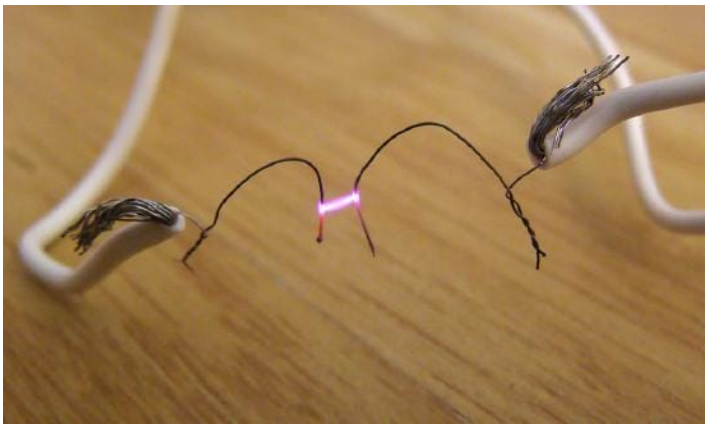


Figure 7.14.6: Is this electric arc dangerous? The answer depends on the AC frequency and the power involved. (credit: Khimich Alex, Wikimedia Commons)

## Summary

- The two types of electric hazards are thermal (excessive power) and shock (current through a person).
- Shock severity is determined by current, path, duration, and AC frequency.
- The table lists shock hazards as a function of current.
- Figure 5 graphs the threshold current for two hazards as a function of frequency.

## Footnotes

1 For an average male shocked through trunk of body for 1 s by 60-Hz AC. Values for females are 60–80% of those listed.

## Glossary

**thermal hazard**
: a hazard in which electric current causes undesired thermal effects

**shock hazard**
: when electric current passes through a person

**short circuit**
: also known as a "short," a low-resistance path between terminals of a voltage source

**microshock sensitive**
: a condition in which a person's skin resistance is bypassed, possibly by a medical procedure, rendering the person vulnerable to electrical shock at currents about 1/1000 the normally required level

---

This page titled 7.14: Electric Hazards and the Human Body is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 20.6: Electric Hazards and the Human Body by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.15: Prelude to Circuits and DC Instruments

Electric circuits are commonplace. Some are simple, such as those in flashlights. Others, such as those used in supercomputers, are extremely complex.



Figure 7.15.1: Electric circuits in a computer allow large amounts of data to be quickly and accurately analyzed.. (credit: Airman 1st Class Mike Meares, United States Air Force)

This collection of modules takes the topic of electric circuits a step beyond simple circuits. When the circuit is purely resistive, everything in this module applies to both DC and AC. Matters become more complex when capacitance is involved. We do consider what happens when capacitors are connected to DC voltage sources, but the interaction of capacitors and other nonresistive devices with AC is left for a later chapter. Finally, a number of important DC instruments, such as meters that measure voltage and current, are covered in this chapter.

This page titled 7.15: Prelude to Circuits and DC Instruments is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 21.0: Prelude to Circuits and DC Instruments by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.16: Resistors in Series and Parallel

## Learning Objectives

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

- Draw a circuit with resistors in parallel and in series.
- Calculate the voltage drop of a current across a resistor using Ohm's law.
- Contrast the way total resistance is calculated for resistors in series and in parallel.
- Explain why total resistance of a parallel circuit is less than the smallest resistance of any of the resistors in that circuit.
- Calculate total resistance of a circuit that contains a mixture of resistors connected in series and in parallel.

sMost circuits have more than one component, called a **resistor** that limits the flow of charge in the circuit. A measure of this limit on charge flow is called **resistance**. The simplest combinations of resistors are the series and parallel connections illustrated in Figure $7.16.1$. The total resistance of a combination of resistors depends on both their individual values and how they are connected.

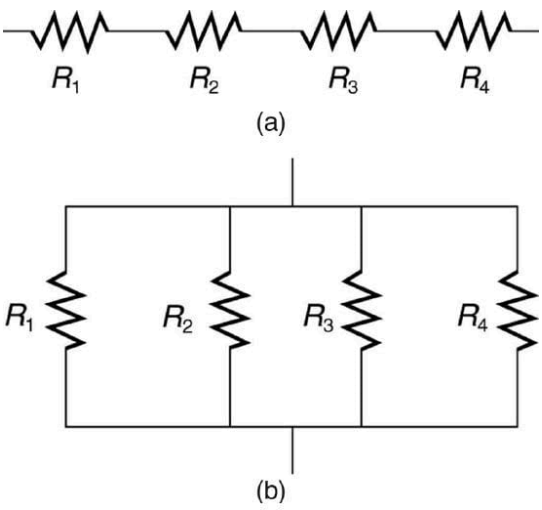



Figure $7.16.1$: (a) A series connection of resistors. (b) A parallel connection of resistors.

## Resistors in Series

When are resistors in **series**? Resistors are in series whenever the flow of charge, called the **current**, must flow through devices sequentially. For example, if current flows through a person holding a screwdriver and into the Earth, then $R_1$ in Figure $7.16.1$(a) could be the resistance of the screwdriver's shaft, $R_2$ the resistance of its handle, $R_3$ the person's body resistance, and $R_4$ the resistance of her shoes.

Figure $2$ shows resistors in series connected to a **voltage** source. It seems reasonable that the total resistance is the sum of the individual resistances, considering that the current has to pass through each resistor in sequence. (This fact would be an advantage to a person wishing to avoid an electrical shock, who could reduce the current by wearing high-resistance rubber-soled shoes. It could be a disadvantage if one of the resistances were a faulty high-resistance cord to an appliance that would reduce the operating current.)

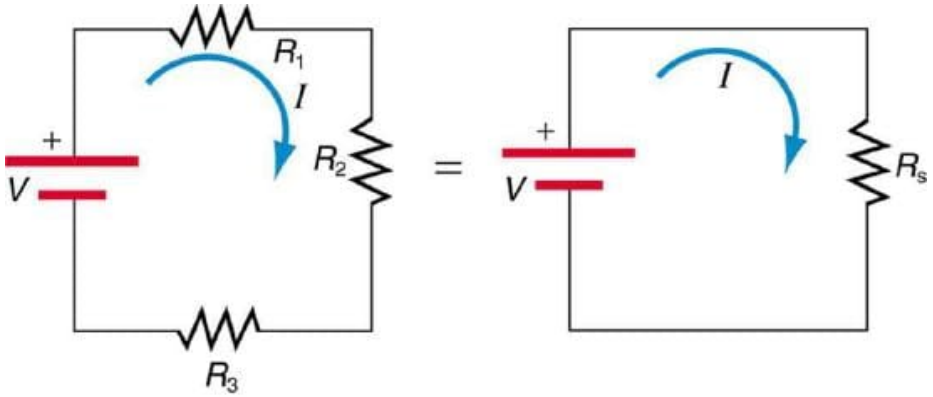



Figure 7.16.2: Three resistors connected in series to a battery (left) and the equivalent single or series resistance (right).

To verify that resistances in series do indeed add, let us consider the loss of electrical power, called a **voltage drop**, in each resistor in Figure 2.

According to **Ohm's law**, the voltage drop, $V$, across a resistor when a current flows through it is calculated using the equation $V = IR$, where $I$ equals the current in amps (A) and $R$ is the resistance in ohms($\Omega$). Another way to think of this is that $V$ is the voltage necessary to make a current $I$ flow through a resistance $R$.

So the voltage drop across $R_1$ is $V_1 = IR_1$, that across $V_2 = IR_2$, and that across $R_3$ is $V_3 = IR_3$. The sum of these voltages equals the voltage output of the source; that is,

$$V = V_1 + V_2 + V_3. \tag{7.16.1}$$

This equation is based on the conservation of energy and conservation of charge. Electrical potential energy can be described by the equation $\mathrm{PE} = qV$, where $q$ is the electric charge and $V$ is the voltage. Thus the energy supplied by the source is $qV$, while that dissipated by the resistors is

$$qV_1 + qV_2 + qV_3. \tag{7.16.2}$$

CONNECTIONS: CONSERVATION LAWS

The derivations of the expressions for series and parallel resistance are based on the laws of conservation of energy and conservation of charge, which state that total charge and total energy are constant in any process. These two laws are directly involved in all electrical phenomena and will be invoked repeatedly to explain both specific effects and the general behavior of electricity.

These energies must be equal, because there is no other source and no other destination for energy in the circuit. Thus, $qV = qV_1 + qV_2 + qV_3$. The charge $q$ cancels, yielding $V = V_1 + V_2 + V_3$, as stated. (Note that the same amount of charge passes through the battery and each resistor in a given amount of time, since there is no capacitance to store charge, there is no place for charge to leak, and charge is conserved.)

Now substituting the values for the individual voltages gives

$$V = IR_1 + IR_2 + IR_3 = I(R_1 + R_2 + R_3). \tag{7.16.3}$$

Note that for the equivalent single series resistance $R_\mathrm{S}$, we have

$$V = IR_\mathrm{S}. \tag{7.16.4}$$

This implies that the total or equivalent series resistance $R_\mathrm{S}$ of three resistors is $R_\mathrm{S} = R_1 + R_2 + R_3$.

This logic is valid in general for any number of resistors in series; thus, the total resistance $R_\mathrm{S}$ of a series connection is

$$R_\mathrm{S} = R_1 + R_2 + R_3 + \ldots, \tag{7.16.5}$$

as proposed. Since all of the current must pass through each resistor, it experiences the resistance of each, and resistances in series simply add up.

## Example 7.16.1: Calculating Resistance, Current, Voltage Drop, and Power Dissipation: Analysis of a Series Circuit

Suppose the voltage output of the battery in Figure 7.16.2 is 12.0V, and the resistances are $R_1 = 1.00\Omega$, $R_2 = 6.00\Omega$, and $R_3 = 13.0\Omega$. (a) What is the total resistance? (b) Find the current. (c) Calculate the voltage drop in each resistor, and show these add to equal the voltage output of the source. (d) Calculate the power dissipated by each resistor. (e) Find the power output of the source, and show that it equals the total power dissipated by the resistors.

**Strategy and Solution for (a)**

The total resistance is simply the sum of the individual resistances, as given by this equation:

$$R_S = R_1 + R_2 + R_3 \tag{7.16.6}$$

$$= 1.00\Omega + 6.00\Omega + 13.0\Omega \tag{7.16.7}$$

$$= 20.0\Omega. \tag{7.16.8}$$

**Strategy and Solution for (b)**

The current is found using Ohm's law, $V = IR$. Entering the value of the applied voltage and the total resistance yields the current for the circuit:

$$I = \frac{V}{R_S} = \frac{12.0\Omega}{20.0\Omega} = 0.600\text{A}. \tag{7.16.9}$$

**Strategy and Solution for (c)**

The voltage—or $IR$ drop—in a resistor is given by Ohm's law. Entering the current and the value of the first resistance yields

$$V_1 = IR_1 = (0.600\text{A})(1.0\Omega) = 0.600\text{V}. \tag{7.16.10}$$

Similarly,

$$V_2 = IR_2 = (0.600\text{A})(6.0\Omega) = 3.60\text{V} \tag{7.16.11}$$

and

$$V_3 = IR_3 = (0.600\text{A})(13.0\Omega) = 7.80\text{V}. \tag{7.16.12}$$

**Discussion for (c)**

The three $IR$ drops add to 12.0V, as predicted:

$$V_1 + V_2 + V_3 = (0.600 + 3.60 + 7.80)\text{V} = 12.0\text{V}. \tag{7.16.13}$$

**Strategy and Solution for (d)**

The easiest way to calculate power in watts (W) dissipated by a resistor in a DC circuit is to use **Joule's law**, $P = IV$, where $P$ is electric power. In this case, each resistor has the same full current flowing through it. By substituting Ohm's law $V = IR$ into Joule's law, we get the power dissipated by the first resistor as

$$P_1 = I^2R_1 = (0.600\text{A})^2(1.00\Omega) = 0.360\text{W}. \tag{7.16.14}$$

Similarly,

$$P_2 = I^2R_2 = (0.600\text{A})^2(6.00\Omega) = 2.16\text{W} \tag{7.16.15}$$

and

$$P_3 = I^2 R_3 = (0.600\text{A})^2(13.0\Omega) = 4.68\text{W}. \tag{7.16.16}$$

**Discussion for (d)**

Power can also be calculated using either $P = IV$ or $P = \frac{V^2}{R}$, where $V$ is the voltage drop across the resistor (not the full voltage of the source). The same values will be obtained.

**Strategy and Solution for (e)**

The easiest way to calculate power output of the source is to use $P = IV$, where $V$ is the source voltage. This gives

$$P = (0.600\text{A})(12.0\text{V}) = 7.20\text{W}. \tag{7.16.17}$$

**Discussion for (e)**

Note, coincidentally, that the total power dissipated by the resistors is also 7.20 W, the same as the power put out by the source. That is,

$$P_1 + P_2 + P_3 = (0.360 + 2.16 + 4.68)\text{W} = 7.20\text{W}. \tag{7.16.18}$$

Power is energy per unit time (watts), and so conservation of energy requires the power output of the source to be equal to the total power dissipated by the resistors.

MAJOR FEATURES OF RESISTORS IN SERIES

1. Series resistances add: $R_\text{S} = R_1 + R_2 + R_3 + \ldots$
2. The same current flows through each resistor in series.
3. Individual resistors in series do not get the total source voltage, but divide it.

## Resistors in Parallel

Figure $7.16.3$ shows resistors in **parallel**, wired to a voltage source. Resistors are in parallel when each resistor is connected directly to the voltage source by connecting wires having negligible resistance. Each resistor thus has the full voltage of the source applied to it.

Each resistor draws the same current it would if it alone were connected to the voltage source (provided the voltage source is not overloaded). For example, an automobile's headlights, radio, and so on, are wired in parallel, so that they utilize the full voltage of the source and can operate completely independently. The same is true in your house, or any building. (See Figure $7.16.3$(b).)

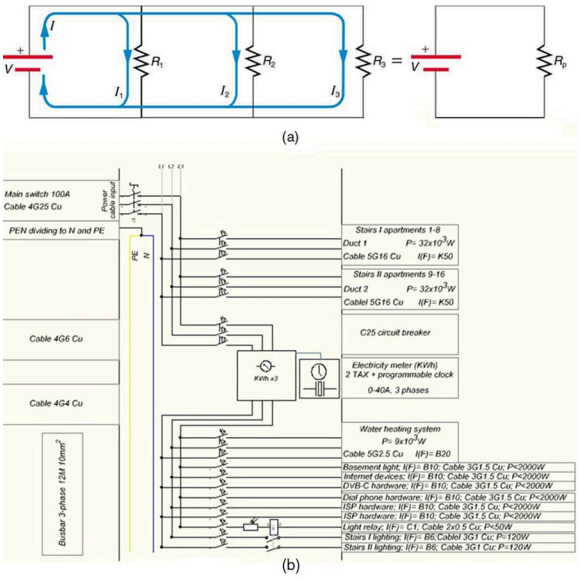



Figure 7.16.3: (a) Three resistors connected in parallel to a battery and the equivalent single or parallel resistance. (b) Electrical power setup in a house. (credit: Dmitry G, Wikimedia Commons)

To find an expression for the equivalent parallel resistance $R_{\text{p}}$, let us consider the currents that flow and how they are related to resistance. Since each resistor in the circuit has the full voltage, the currents flowing through the individual resistors are $I_1 = \frac{V}{R_1}$, $I_2 = \frac{V}{R_2}$, and $I_3 = \frac{V}{R_3}$. Conservation of charge implies that the total current $I$ produced by the source is the sum of these currents:

$$I = I_1 + I_2 + I_3. \tag{7.16.19}$$

Substituting the expressions for the individual currents gives

$$I = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3} = V(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}). \tag{7.16.20}$$

Note that Ohm's law for the equivalent single resistance gives

$$I = \frac{V}{R_{\text{p}}} = V(\frac{1}{R_{\text{p}}}). \tag{7.16.21}$$

The terms inside the parentheses in the last two equations must be equal. Generalizing to any number of resistors, the total resistance $R_{\text{p}}$ of a parallel connection is related to the individual resistances by

$$\frac{1}{R_{\text{p}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots \tag{7.16.22}$$

This relationship results in a total resistance $R_{\text{p}}$ that is less than the smallest of the individual resistances. (This is seen in the next example.) When resistors are connected in parallel, more current flows from the source than would flow for any of them individually, and so the total resistance is lower.

## Example 7.16.2: Calculating Resistance, Current, Power Dissipation, and Power Output: Analysis of a Parallel Circuit

Let the voltage output of the battery and resistances in the parallel connection in Figure 7.16.3 be the same as the previously considered series connection: $V = 12.0\text{V}$, $R_1 = 1.00\Omega$, $R_2 = 6.00\Omega$, and $R_3 = 13.0\Omega$. (a) What is the total resistance? (b) Find the total current. (c) Calculate the currents in each resistor, and show these add to equal the total current output of the source. (d) Calculate the power dissipated by each resistor. (e) Find the power output of the source, and show that it equals the total power dissipated by the resistors.

**Strategy and Solution for (a)**

The total resistance for a parallel combination of resistors is found using the equation below. Entering known values gives

$$\frac{1}{R_{\text{p}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{1.00\Omega} + \frac{1}{6.00\Omega} + \frac{1}{13.0\Omega}. \tag{7.16.23}$$

Thus,

$$\frac{1}{R_{\text{p}}} = \frac{1.00}{\Omega} + \frac{0.1667}{\Omega} + \frac{0.07692}{\Omega} = \frac{1.2436}{\Omega}. \tag{7.16.24}$$

(Note that in these calculations, each intermediate answer is shown with an extra digit.)

We must invert this to find the total resistance $R_{\text{p}}$. This yields

$$R_{\text{p}} = \frac{1}{1.2436}\Omega = 0.8041\Omega. \tag{7.16.25}$$

The total resistance with the correct number of significant digits is $R_{\text{p}} = 0.804\Omega$.

**Discussion for (a)**

$R_{\text{p}}$ is, as predicted, less than the smallest individual resistance.

**Strategy and Solution for (b)**

The total current can be found from Ohm's law, substituting $R_{\text{p}}$ for the total resistance. This gives

$$I = \frac{V}{R_{\text{p}}} = \frac{12.0\text{V}}{0.8041\Omega} = 14.92\text{A}. \tag{7.16.26}$$

**Discussion for (b)**

Current $i$ for each device is much larger than for the same devices connected in series (see the previous example). A circuit with parallel connections has a smaller total resistance than the resistors connected in series.

**Strategy and Solution for (c)**

The individual currents are easily calculated from Ohm's law, since each resistor gets the full voltage. Thus,

$$I_1 = \frac{V}{R_1} = \frac{12.0\text{V}}{1.00\Omega} = 12.0\text{A}. \tag{7.16.27}$$

Similarly,

$$I_2 = \frac{V}{R_2} = \frac{12.0\text{V}}{6.00\Omega} = 2.00\text{A} \tag{7.16.28}$$

and

$$I_3 = \frac{V}{R_3} = \frac{12.0\text{V}}{13.0\Omega} = 0.92\text{A}. \tag{7.16.29}$$

**Discussion for (c)**

The total current is the sum of the individual currents:

$$I_1 + I_2 + I_3 = 14.92\text{A}. \tag{7.16.30}$$

This is consistent with conservation of charge.

**Strategy and Solution for (d)**

The power dissipated by each resistor can be found using any of the equations relating power to current, voltage, and resistance, since all three are known. Let us use $P = \frac{V^2}{R}$, since each resistor gets full voltage. Thus,

$$P_1 = \frac{V^2}{R_1} = \frac{(12.0\text{V})^2}{1.00\Omega} = 144\text{W}. \tag{7.16.31}$$

Similarly,

$$P_2 = \frac{V^2}{R_2} = \frac{(12.0\text{V})^2}{6.00\Omega} = 24.0\text{W} \tag{7.16.32}$$

and

$$P_3 = \frac{V^2}{R_3} = \frac{(12.0\text{V})^2}{13.0\Omega} = 11.1\text{W}. \tag{7.16.33}$$

**Discussion for (d)**

The power dissipated by each resistor is considerably higher in parallel than when connected in series to the same voltage source.

**Strategy and Solution for (e)**

The total power can also be calculated in several ways. Choosing $P = IV$, and entering the total current, yields

$$P = IV = (14.92\text{A})(12.0\text{V}) = 179\text{W}. \tag{7.16.34}$$

**Discussion for (e)**

Total power dissipated by the resistors is also 179 W:

$$P_1 + P_2 + P_3 = 144\text{W} + 24.0\text{W} + 11.1\text{W} = 179\text{W}. \tag{7.16.35}$$

This is consistent with the law of conservation of energy.

**Overall Discussion**

Note that both the currents and powers in parallel connections are greater than for the same devices in series.

MAJOR FEATURES OF RESISTORS IN PARALLEL

1. Parallel resistance is found from $\frac{1}{R_\text{p}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots$, and it is smaller than any individual resistance in the combination.
2. Each resistor in parallel has the same full voltage of the source applied to it. (Power distribution systems most often use parallel connections to supply the myriad devices served with the same voltage and to allow them to operate independently.)
3. Parallel resistors do not each get the total current; they divide it.

## Combinations of Series and Parallel

More complex connections of resistors are sometimes just combinations of series and parallel. These are commonly encountered, especially when wire resistance is considered. In that case, wire resistance is in series with other resistances that are in parallel.

Combinations of series and parallel can be reduced to a single equivalent resistance using the technique illustrated in Figure

7.16.4. Various parts are identified as either series or parallel, reduced to their equivalents, and further reduced until a single resistance is left. The process is more time consuming than difficult.

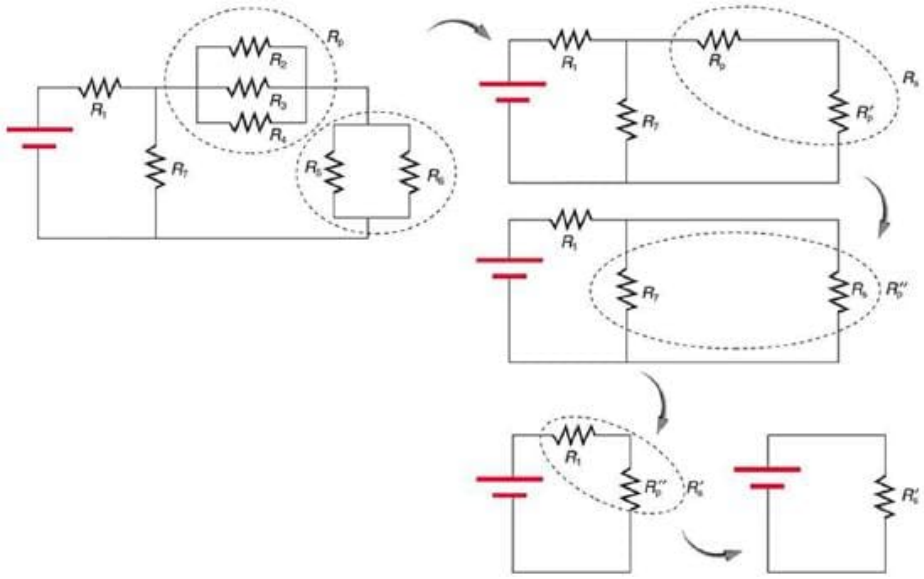


Figure 7.16.4: This combination of seven resistors has both series and parallel parts. Each is identified and reduced to an equivalent resistance, and these are further reduced until a single equivalent resistance is reached.

The simplest combination of series and parallel resistance, shown in Figure 7.16.5, is also the most instructive, since it is found in many applications. For example, $R_1$ could be the resistance of wires from a car battery to its electrical devices, which are in parallel. $R_2$ and $R_3$ could be the starter motor and a passenger compartment light. We have previously assumed that wire resistance is negligible, but, when it is not, it has important effects, as the next example indicates.

### Example 7.16.3: Calculating Resistance, $IR$ Drop, Current, and Power Dissipation: Combining Series and Parallel Circuits

Figure 7.16.5 shows the resistors from the previous two examples wired in a different way—a combination of series and parallel. We can consider $R_1$ to be the resistance of wires leading to $R_2$ and $R_3$. (a) Find the total resistance. (b) What is the $IR$ drop in $R_1$? (c) Find the current $I_2$ through $R_2$. (d) What power is dissipated by $R_2$?

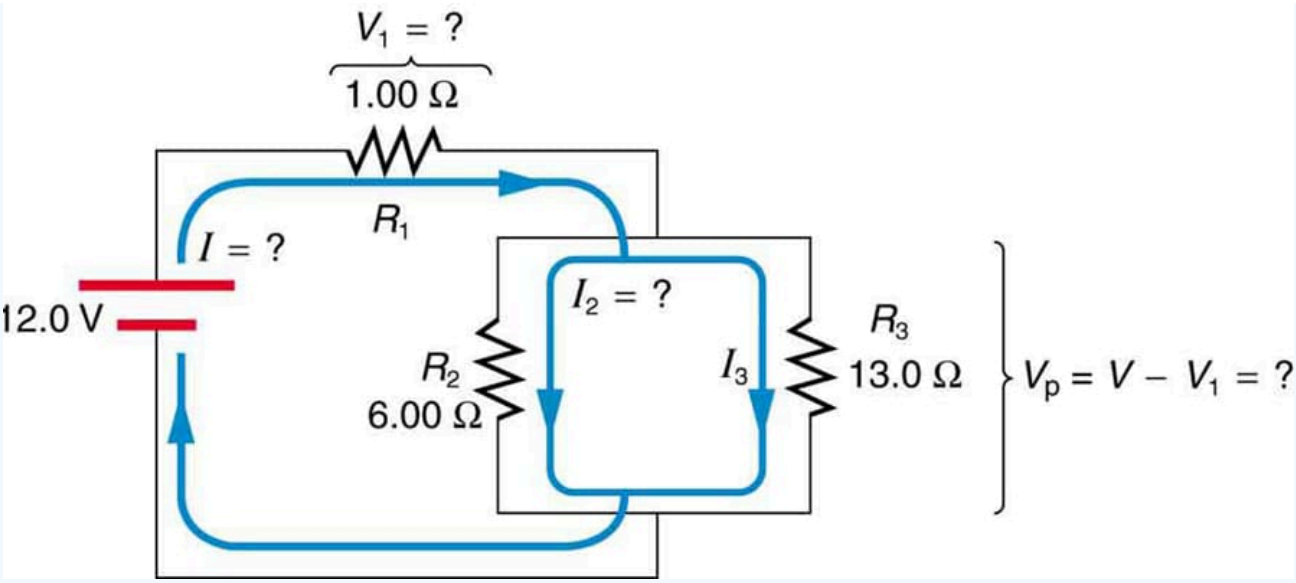



Figure 7.16.5. These three resistors are connected to a voltage source so that $R_2$ and $R_3$ are in parallel with one another and that combination is in series with $R_1$.

**Strategy and Solution for (a)**

To find the total resistance, we note that $R_2$ and $R_3$ are in parallel and their combination $R_p$ is in series with $R_1$. Thus the total (equivalent) resistance of this combination is

$$R_{tot} = R_1 + R_2. \tag{7.16.36}$$

First, we find $R_p$ using the equation for resistors in parallel and entering known values:

$$\frac{1}{R_p} = \frac{1}{R_2} + \frac{1}{6.00\,\Omega} + \frac{1}{13.0\,\Omega} = \frac{0.2436}{\Omega}. \tag{7.16.37}$$

Inverting gives

$$R_p = \frac{1}{0.2436}\Omega = 4.11\,\Omega. \tag{7.16.38}$$

So the total resistance is

$$R_{tot} = R_1 + R_p = 1.00\Omega + 4.11\Omega = 5.11\,\Omega. \tag{7.16.39}$$

**Discussion for (a)**

The total resistance of this combination is intermediate between the pure series and pure parallel values ($20.0\ \Omega$ and $0.804\ \Omega$, respectively) found for the same resistors in the two previous examples.

**Strategy and Solution for (b)**

To find the $IR$ drop in $R_1$, we note that the full current $I$ flows through $R_1$. Thus, its $IR$ drop is

$$V_1 = IR_1. \tag{7.16.40}$$

We must find $I$ before we can calculate $V_1$. The total current $I$ is found using Ohm's law for the circuit. That is,

$$I = \frac{V}{R_{tot}} = \frac{12.0\,V}{5.11\,\Omega} = 2.35\,A. \tag{7.16.41}$$

Entering this into the expression above, we get

$$V_1 = IR_1 = (2.35\,A)(1.00\,\Omega) = 2.35\,V. \tag{7.16.42}$$

**Discussion for (b)**

The voltage applied to $R_2$ and $R_3$ is less than the total voltage by an amount $V_1$. When wire resistance is large, it can significantly affect the operation of the devices represented by $R_2$ and $R_3$.

**Strategy and Solution for (c)**

To find the current through $R_2$, we must first find the voltage applied to it. We call this voltage $V_p$, because it is applied to a parallel combination of resistors. The voltage applied to both $R_2$ and $R_3$ is reduced by the amount $V_1$, and so it is

$$V_p = V - V_1 = 12.0\,V - 2.35\,V = 9.65\,V. \tag{7.16.43}$$

Now the current $I_2$ through resistance $R_2$ is found using Ohm's law:

$$I_2 = \frac{V_p}{R_2} = \frac{9.65\,V}{6.00\,\Omega} = 1.61\,A. \tag{7.16.44}$$

**Discussion for (c)**

The current is less than the 2.00 A that flowed through $R_2$ when it was connected in parallel to the battery in the previous parallel circuit example.

**Strategy and Solution for (d)**

The power dissipated by $R_2$ is given by

$$P_2 = (I_2)^2 R_2 = (161\,A)^2(6.00\,\Omega) = 15.5\,W. \tag{7.16.45}$$

**Discussion for (d)**

The power is less than the 24.0 W this resistor dissipated when connected in parallel to the 12.0-V source.

## Practical Implications

One implication of this last example is that resistance in wires reduces the current and power delivered to a resistor. If wire resistance is relatively large, as in a worn (or a very long) extension cord, then this loss can be significant. If a large current is drawn, the $IR$ drop in the wires can also be significant.

For example, when you are rummaging in the refrigerator and the motor comes on, the refrigerator light dims momentarily. Similarly, you can see the passenger compartment light dim when you start the engine of your car (although this may be due to resistance inside the battery itself).

What is happening in these high-current situations is illustrated in Figure 7.16.6. The device represented by $R_3$ has a very low resistance, and so when it is switched on, a large current flows. This increased current causes a larger $IR$ drop in the wires represented by $R_1$ reducing the voltage across the light bulb (which is $R_2$), which then dims noticeably.

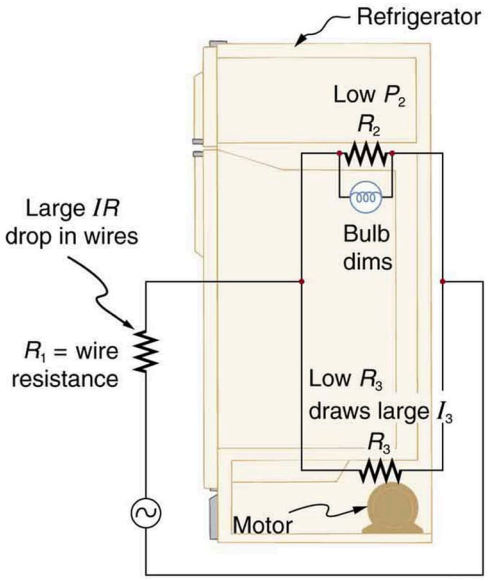



Figure 7.16.6: Why do lights dim when a large appliance is switched on? The answer is that the large current the appliance motor draws causes a significant $IR$ drop in the wires and reduces the voltage across the light.

### Exercise 7.16.1

Can any arbitrary combination of resistors be broken down into series and parallel combinations? See if you can draw a circuit diagram of resistors that cannot be broken down into combinations of series and parallel.

*Answer available at the end of this section.*

### Problem Solving Strategies For Series and Parallel Resistors

1. Draw a clear circuit diagram, labeling all resistors and voltage sources. This step includes a list of the knowns for the problem, since they are labeled in your circuit diagram.
2. Identify exactly what needs to be determined in the problem (identify the unknowns). A written list is useful.
3. Determine whether resistors are in series, parallel, or a combination of both series and parallel. Examine the circuit diagram to make this assessment. Resistors are in series if the same current must pass sequentially through them.
4. Use the appropriate list of major features for series or parallel connections to solve for the unknowns. There is one list for series and another for parallel. If your problem has a combination of series and parallel, reduce it in steps by considering individual groups of series or parallel connections, as done in this module and the examples. Special note: When finding $R_p$, the reciprocal must be taken with care.
5. Check to see whether the answers are reasonable and consistent. Units and numerical results must be reasonable. Total series resistance should be greater, whereas total parallel resistance should be smaller, for example. Power should be greater for the same devices in parallel compared with series, and so on.

## Summary

- The total resistance of an electrical circuit with resistors wired in a series is the sum of the individual resistances: $R_s = R_1 + R_2 + R_3 + ....$
- Each resistor in a series circuit has the same amount of current flowing through it.
- The voltage drop, or power dissipation, across each individual resistor in a series is different, and their combined total adds up to the power source input.
- The total resistance of an electrical circuit with resistors wired in parallel is less than the lowest resistance of any of the components and can be determined using the formula: $\frac{1}{R_p} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + ....$
- Each resistor in a parallel circuit has the same full voltage of the source applied to it.

- The current flowing through each resistor in a parallel circuit is different, depending on the resistance.
- If a more complex connection of resistors is a combination of series and parallel, it can be reduced to a single equivalent resistance by identifying its various parts as series or parallel, reducing each to its equivalent, and continuing until a single resistance is eventually reached.

# Glossary

**series**
: a sequence of resistors or other components wired into a circuit one after the other

**resistor**
: a component that provides resistance to the current flowing through an electrical circuit

**resistance**
: causing a loss of electrical power in a circuit

**Ohm's law**
: the relationship between current, voltage, and resistance within an electrical circuit: V=IR

**voltage**
: the electrical potential energy per unit charge; electric pressure created by a power source, such as a battery

**voltage drop**
: the loss of electrical power as a current travels through a resistor, wire or other component

**current**
: the flow of charge through an electric circuit past a given point of measurement

**Joule's law**
: tthe relationship between potential electrical power, voltage, and resistance in an electrical circuit, given by: $P_e = IV$

**parallel**
: the wiring of resistors or other components in an electrical circuit such that each component receives an equal voltage from the power source; often pictured in a ladder-shaped diagram, with each component on a rung of the ladder

## Answers

**Exercise** 7.16.1

No, there are many ways to connect resistors that are not combinations of series and parallel, including loops and junctions. In such cases Kirchhoff's rules, to be introduced in Kirchhoff's Rules, will allow you to analyze the circuit.

This page titled 7.16: Resistors in Series and Parallel is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 21.1: Resistors in Series and Parallel by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.17: Prelude to Magnetism

One evening, an Alaskan sticks a note to his refrigerator with a small magnet. Through the kitchen window, the Aurora Borealis glows in the night sky. This grand spectacle is shaped by the same force that holds the note to the refrigerator.



Figure 7.17.1: The magnificent spectacle of the Aurora Borealis, or northern lights, glows in the northern sky above Bear Lake near Eielson Air Force Base, Alaska. Shaped by the Earth's magnetic field, this light is produced by radiation spewed from solar storms. (credit: Senior Airman Joshua Strang, via Flickr)

People have been aware of magnets and magnetism for thousands of years. The earliest records date to well before the time of Christ, particularly in a region of Asia Minor called Magnesia (the name of this region is the source of words like *magnetic*). Magnetic rocks found in Magnesia, which is now part of western Turkey, stimulated interest during ancient times. A practical application for magnets was found later, when they were employed as navigational compasses. The use of magnets in compasses resulted not only in improved long-distance sailing, but also in the names of "north" and "south" being given to the two types of magnetic poles.

Today magnetism plays many important roles in our lives. Physicists' understanding of magnetism has enabled the development of technologies that affect our everyday lives. The iPod in your purse or backpack, for example, wouldn't have been possible without the applications of magnetism and electricity on a small scale.

The discovery that weak changes in a magnetic field in a thin film of iron and chromium could bring about much larger changes in electrical resistance was one of the first large successes of nanotechnology. The 2007 Nobel Prize in Physics went to Albert Fert from France and Peter Grunberg from Germany for this discovery of *giant magnetoresistance* and its applications to computer memory.

All electric motors, with uses as diverse as powering refrigerators, starting cars, and moving elevators, contain magnets. Generators, whether producing hydroelectric power or running bicycle lights, use magnetic fields. Recycling facilities employ magnets to separate iron from other refuse. Hundreds of millions of dollars are spent annually on magnetic containment of fusion as a future energy source. Magnetic resonance imaging (MRI) has become an important diagnostic tool in the field of medicine, and the use of magnetism to explore brain activity is a subject of contemporary research and development. The list of applications also includes computer hard drives, tape recording, detection of inhaled asbestos, and levitation of high-speed trains. Magnetism is used to explain atomic energy levels, cosmic rays, and charged particles trapped in the Van Allen belts. Once again, we will find all these disparate phenomena are linked by a small number of underlying physical principles.

 Figure 7.17.2: Engineering of technology like iPods would not be possible without a deep understanding magnetism. (credit: Jesse! S?, Flickr)

## Contributors and Attributions

*Paul Peter Urone (Professor Emeritus at California State University, Sacramento) and Roger Hinrichs (State University of New York, College at Oswego) with Contributing Authors: Kim Dirks (University of Auckland) and Manjula Sharma (University of Sydney). This work is licensed by OpenStax University Physics under a Creative Commons Attribution License (by 4.0).*

-

This page titled 7.17: Prelude to Magnetism is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 22.0: Prelude to Magnetism by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.18: Magnets

## Learning Objectives

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

- Describe the difference between the north and south poles of a magnet.
- Describe how magnetic poles interact with each other.

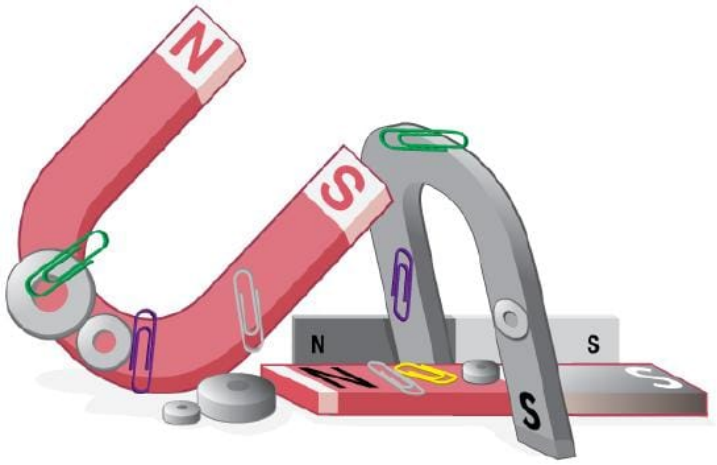



Figure 7.18.1: Magnets come in various shapes, sizes, and strengths. All have both a north pole and a south pole. There is never an isolated pole (a monopole).

All magnets attract iron, such as that in a refrigerator door. However, magnets may attract or repel other magnets. Experimentation shows that all magnets have two poles. If freely suspended, one pole will point toward the north. The two poles are thus named the north magnetic pole and the south magnetic pole (or more properly, north-seeking and south-seeking poles, for the attractions in those directions).

UNIVERSAL CHARACTERISTICS OF MAGNETS AND MAGNET POLES

It is a universal characteristic of all magnets that *like poles repel and unlike poles attract*. (Note the similarity with electrostatics: unlike charges attract and like charges repel.)

Further experimentation shows that it is *impossible to separate north and south poles* in the manner that + and – charges can be separated.

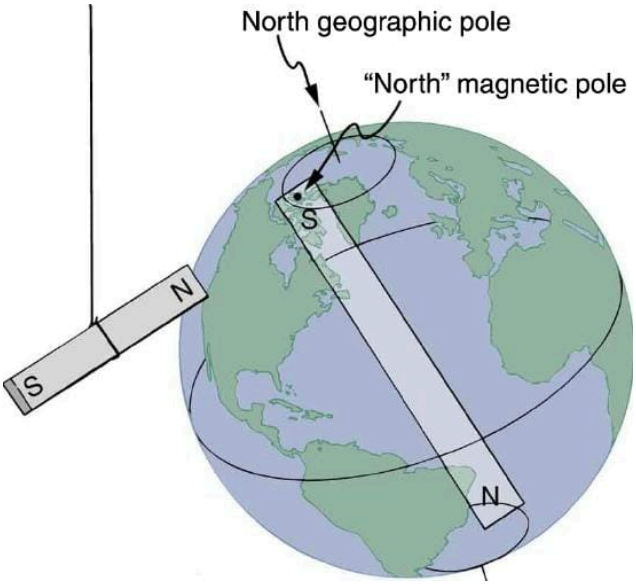



Figure 7.18.2: One end of a bar magnet is suspended from a thread that points toward north. The magnet's two poles are labeled N and S for north-seeking and south-seeking poles, respectively.

MISCONCEPTION ALERT: EARTH'S GEOGRAPHIC NORTH POLE HIDES AN S

The Earth acts like a very large bar magnet with its south-seeking pole near the geographic North Pole. That is why the north pole of your compass is attracted toward the geographic north pole of the Earth—because the magnetic pole that is near the geographic North Pole is actually a south magnetic pole! Confusion arises because the geographic term "North Pole" has come to be used (incorrectly) for the magnetic pole that is near the North Pole. Thus, "North magnetic pole" is actually a misnomer—it should be called the South magnetic pole.

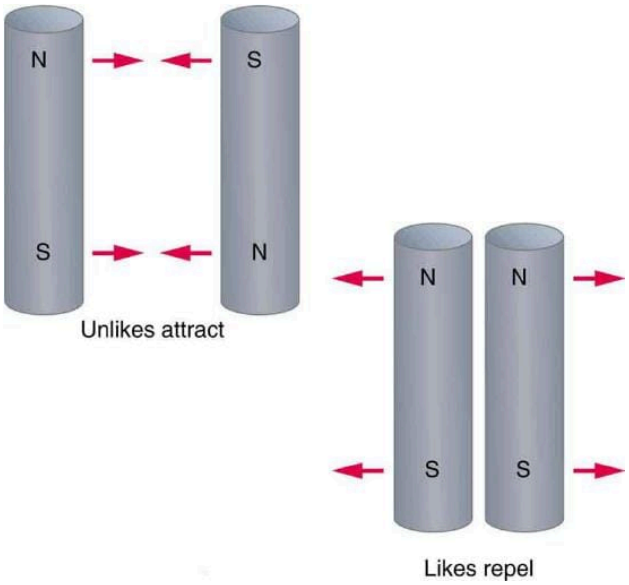



Figure 7.18.3: Unlike poles attract, whereas like poles repel.

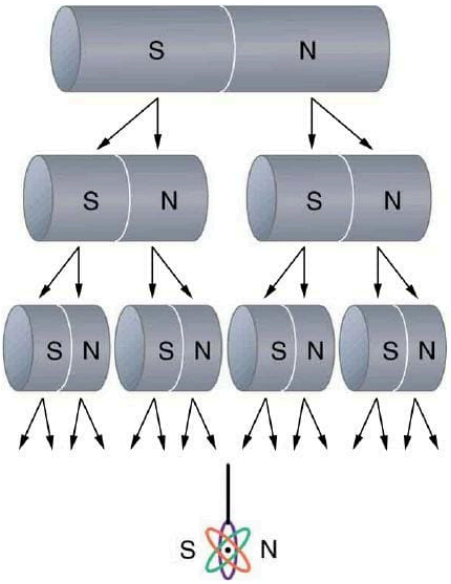



Figure 7.18.4: North and south poles always occur in pairs. Attempts to separate them result in more pairs of poles. If we continue to split the magnet, we will eventually get down to an iron atom with a north pole and a south pole—these, too, cannot be separated.

The fact that magnetic poles always occur in pairs of north and south is true from the very large scale—for example, sunspots always occur in pairs that are north and south magnetic poles—all the way down to the very small scale. Magnetic atoms have both a north pole and a south pole, as do many types of subatomic particles, such as electrons, protons, and neutrons.

### MAKING CONNECTIONS: TAKE-HOME EXPERIMENT -- REFRIGERATOR MAGNETS

We know that like magnetic poles repel and unlike poles attract. See if you can show this for two refrigerator magnets. Will the magnets stick if you turn them over? Why do they stick to the door anyway? What can you say about the magnetic properties of the door next to the magnet? Do refrigerator magnets stick to metal or plastic spoons? Do they stick to all types of metal?

## Summary

- Magnetism is a subject that includes the properties of magnets, the effect of the magnetic force on moving charges and currents, and the creation of magnetic fields by currents.
- There are two types of magnetic poles, called the north magnetic pole and south magnetic pole.
- North magnetic poles are those that are attracted toward the Earth's geographic north pole.
- Like poles repel and unlike poles attract.
- Magnetic poles always occur in pairs of north and south—it is not possible to isolate north and south poles.

## Glossary

**north magnetic pole**
: the end or the side of a magnet that is attracted toward Earth's geographic north pole

**south magnetic pole**
: the end or the side of a magnet that is attracted toward Earth's geographic south pole

This page titled 7.18: Magnets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 22.1: Magnets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.19: Ferromagnets and Electromagnets

**Learning Objectives**

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

- Define ferromagnet.
- Describe the role of magnetic domains in magnetization.
- Explain the significance of the Curie temperature.
- Describe the relationship between electricity and magnetism.

## Ferromagnets

Only certain materials, such as iron, cobalt, nickel, and gadolinium, exhibit strong magnetic effects. Such materials are called ferromagnetic, after the Latin word for iron, *ferrum*. A group of materials made from the alloys of the rare earth elements are also used as strong and permanent magnets; a popular one is neodymium. Other materials exhibit weak magnetic effects, which are detectable only with sensitive instruments. Not only do ferromagnetic materials respond strongly to magnets (the way iron is attracted to magnets), they can also be magnetized themselves—that is, they can be induced to be magnetic or made into permanent magnets.

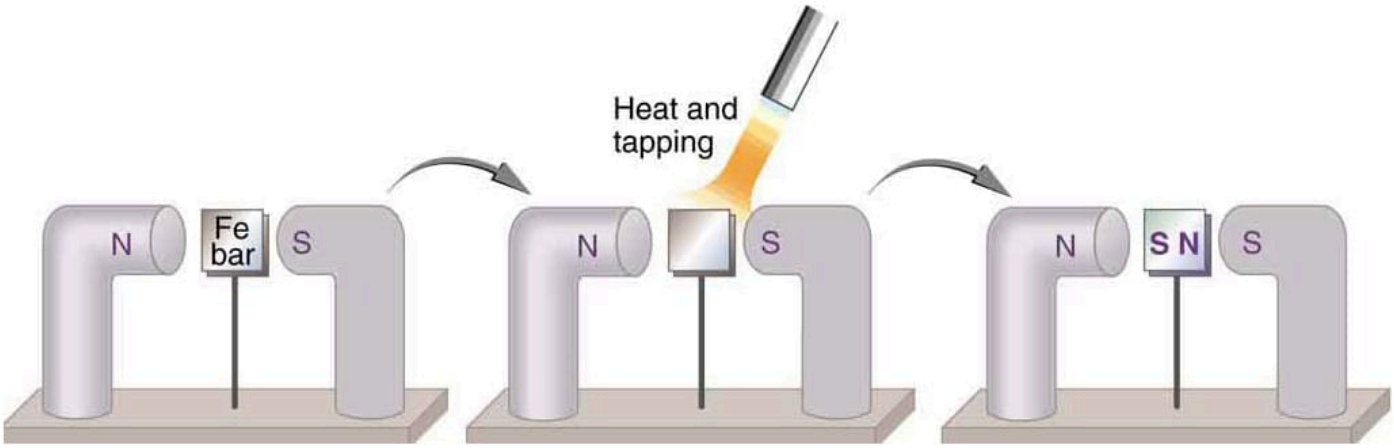



Figure 7.19.1: An unmagnetized piece of iron is placed between two magnets, heated, and then cooled, or simply tapped when cold. The iron becomes a permanent magnet with the poles aligned as shown: its south pole is adjacent to the north pole of the original magnet, and its north pole is adjacent to the south pole of the original magnet. Note that there are attractive forces between the magnets.

When a magnet is brought near a previously unmagnetized ferromagnetic material, it causes local magnetization of the material with unlike poles closest, as in Figure 7.19.1. (This results in the attraction of the previously unmagnetized material to the magnet.) What happens on a microscopic scale is illustrated in Figure 7.19.2. The regions within the material called domains act like small bar magnets. Within domains, the poles of individual atoms are aligned. Each atom acts like a tiny bar magnet. Domains are small and randomly oriented in an unmagnetized ferromagnetic object. In response to an external magnetic field, the domains may grow to millimeter size, aligning themselves as shown in Figure 2b. This induced magnetization can be made permanent if the material is heated and then cooled, or simply tapped in the presence of other magnets.

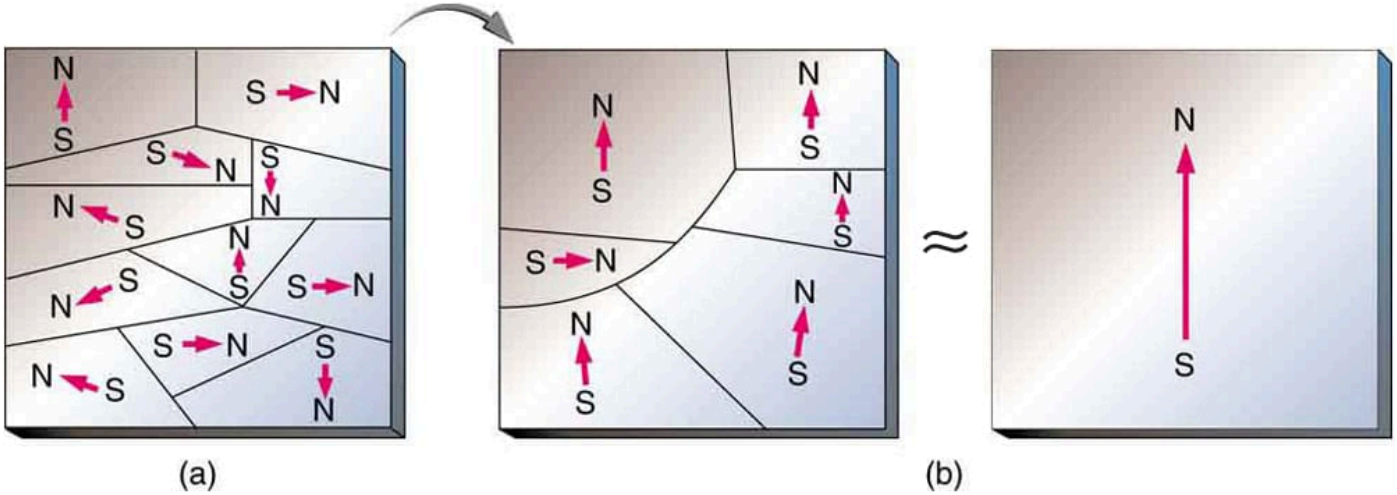



Figure 7.19.2: (a) An unmagnetized piece of iron (or other ferromagnetic material) has randomly oriented domains. (b) When magnetized by an external field, the domains show greater alignment, and some grow at the expense of others. Individual atoms are aligned within domains; each atom acts like a tiny bar magnet.

Conversely, a permanent magnet can be demagnetized by hard blows or by heating it in the absence of another magnet. Increased thermal motion at higher temperature can disrupt and randomize the orientation and the size of the domains. There is a well-defined temperature for ferromagnetic materials, which is called the Curie temperature, above which they cannot be magnetized. The Curie temperature for iron is 1043 K $(770^{\circ}C)$, which is well above room temperature. There are several elements and alloys that have Curie temperatures much lower than room temperature and are ferromagnetic only below those temperatures.

## Electromagnets

Early in the 19th century, it was discovered that electrical currents cause magnetic effects. The first significant observation was by the Danish scientist Hans Christian Oersted (1777–1851), who found that a compass needle was deflected by a current-carrying wire. This was the first significant evidence that the movement of charges had any connection with magnets. Electromagnetism is the use of electric current to make magnets. These temporarily induced magnets are called electromagnets. Electromagnets are employed for everything from a wrecking yard crane that lifts scrapped cars to controlling the beam of a 90-km-circumference particle accelerator to the magnets in medical imaging machines (Figure 7.19.3).



Figure 7.19.3: Instrument for magnetic resonance imaging (MRI). The device uses a superconducting cylindrical coil for the main magnetic field. The patient goes into this “tunnel” on the gurney. (credit: Bill McChesney, Flickr)

Figure 7.19.4 shows that the response of iron filings to a current-carrying coil and to a permanent bar magnet. The patterns are similar. In fact, electromagnets and ferromagnets have the same basic characteristics—for example, they have north and south poles that cannot be separated and for which like poles repel and unlike poles attract.

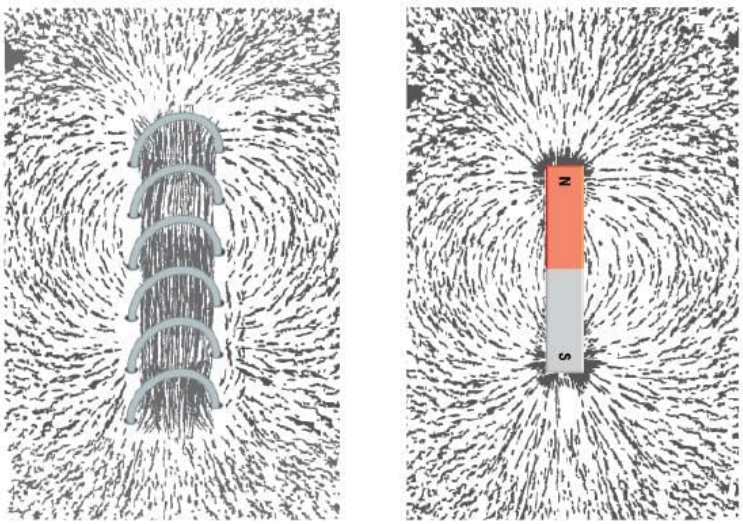



Figure 7.19.4: Iron filings near (a) a current-carrying coil and (b) a magnet act like tiny compass needles, showing the shape of their fields. Their response to a current-carrying coil and a permanent magnet is seen to be very similar, especially near the ends of the coil and the magnet.

Combining a ferromagnet with an electromagnet can produce particularly strong magnetic effects (Figure 7.19.5). Whenever strong magnetic effects are needed, such as lifting scrap metal, or in particle accelerators, electromagnets are enhanced by ferromagnetic materials. Limits to how strong the magnets can be made are imposed by coil resistance (it will overheat and melt at sufficiently high current), and so superconducting magnets may be employed. These are still limited, because superconducting properties are destroyed by too great a magnetic field.

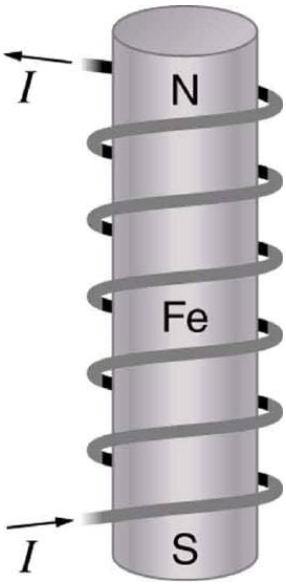



Figure 7.19.5: An electromagnet with a ferromagnetic core can produce very strong magnetic effects. Alignment of domains in the core produces a magnet, the poles of which are aligned with the electromagnet.

Figure 7.19.6 shows a few uses of combinations of electromagnets and ferromagnets. Ferromagnetic materials can act as memory devices, because the orientation of the magnetic fields of small domains can be reversed or erased. Magnetic information storage on videotapes and computer hard drives are among the most common applications. This property is vital in our digital world.

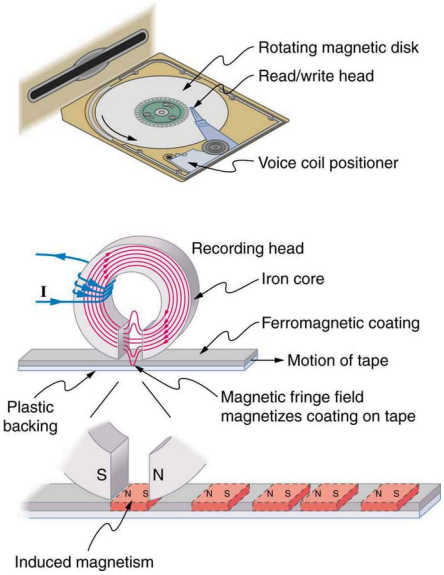



Figure 7.19.6: An electromagnet induces regions of permanent magnetism on a floppy disk coated with a ferromagnetic material. The information stored here is digital (a region is either magnetic or not); in other applications, it can be analog (with a varying strength), such as on audiotapes.

## Current: The Source of All Magnetism

An electromagnet creates magnetism with an electric current. In later sections we explore this more quantitatively, finding the strength and direction of magnetic fields created by various currents. But what about ferromagnets? Figure 7.19.7 shows models of how electric currents create magnetism at the submicroscopic level. (Note that we cannot directly observe the paths of individual electrons about atoms, and so a model or visual image, consistent with all direct observations, is made. We can directly observe the electron's orbital angular momentum, its spin momentum, and subsequent magnetic moments, all of which are explained with electric-current-creating subatomic magnetism.) Currents, including those associated with other submicroscopic particles like protons, allow us to explain ferromagnetism and all other magnetic effects. Ferromagnetism, for example, results from an internal cooperative alignment of electron spins, possible in some materials but not in others.

Crucial to the statement that electric current is the source of all magnetism is the fact that it is impossible to separate north and south magnetic poles. (This is far different from the case of positive and negative charges, which are easily separated.) A current loop always produces a magnetic dipole—that is, a magnetic field that acts like a north pole and south pole pair. Since isolated north and south magnetic poles, called magnetic monopoles, are not observed, currents are used to explain all magnetic effects. If magnetic monopoles did exist, then we would have to modify this underlying connection that all magnetism is due to electrical current. There is no known reason that magnetic monopoles should not exist—they are simply never observed—and so searches at the subnuclear level continue. If they do *not* exist, we would like to find out why not. If they *do* exist, we would like to see evidence of them.

**ELECTRIC CURRENTS AND MAGNETISM**

Electric current is the source of all magnetism.

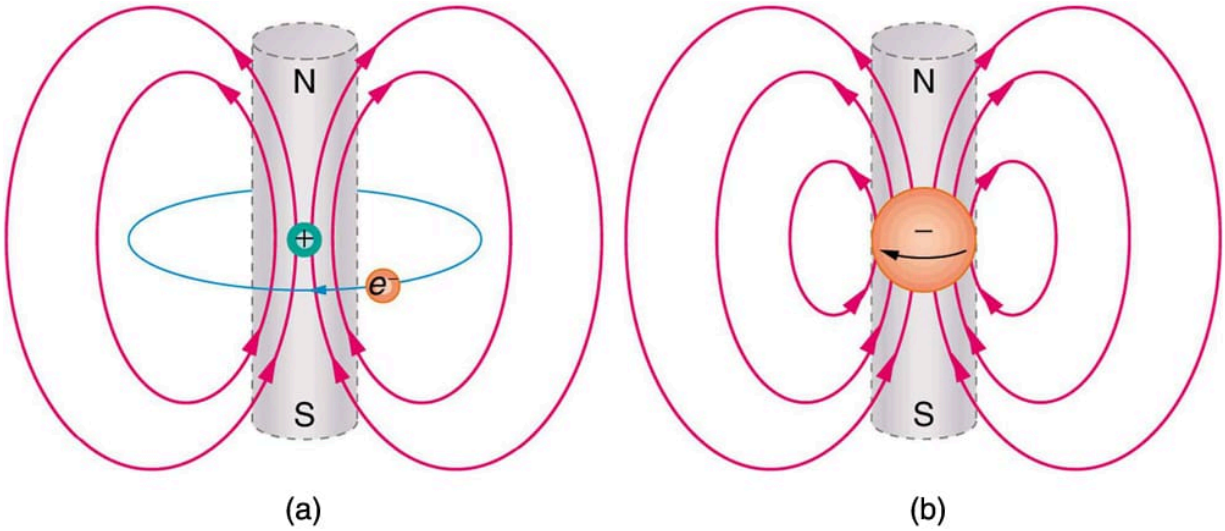



Figure 7.19.7: (a) In the planetary model of the atom, an electron orbits a nucleus, forming a closed-current loop and producing a magnetic field with a north pole and a south pole. (b) Electrons have spin and can be crudely pictured as rotating charge, forming a current that produces a magnetic field with a north pole and a south pole. Neither the planetary model nor the image of a spinning electron is completely consistent with modern physics. However, they do provide a useful way of understanding phenomena.

### PHET EXPLORATIONS: MAGNETS AND ELECTROMAGNETS

Explore the interactions between a compass and bar magnet. Discover how you can use a battery and wire to make a magnet! Can you make it a stronger magnet? Can you make the magnetic field reverse?

## Summary

- Magnetic poles always occur in pairs of north and south—it is not possible to isolate north and south poles.
- All magnetism is created by electric current.
- Ferromagnetic materials, such as iron, are those that exhibit strong magnetic effects.
- The atoms in ferromagnetic materials act like small magnets (due to currents within the atoms) and can be aligned, usually in millimeter-sized regions called domains.
- Domains can grow and align on a larger scale, producing permanent magnets. Such a material is magnetized, or induced to be magnetic.
- Above a material's Curie temperature, thermal agitation destroys the alignment of atoms, and ferromagnetism disappears.
- Electromagnets employ electric currents to make magnetic fields, often aided by induced fields in ferromagnetic materials.

## Glossary

**ferromagnetic**
: materials, such as iron, cobalt, nickel, and gadolinium, that exhibit strong magnetic effects

**magnetized**
: to be turned into a magnet; to be induced to be magnetic

**domains**
: regions within a material that behave like small bar magnets

**Curie temperature**
: the temperature above which a ferromagnetic material cannot be magnetized

**electromagnetism**
: the use of electrical currents to induce magnetism

**electromagnet**
: an object that is temporarily magnetic when an electrical current is passed through it

**magnetic monopoles**
: an isolated magnetic pole; a south pole without a north pole, or vice versa (no magnetic monopole has ever been observed)

This page titled 7.19: Ferromagnets and Electromagnets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 22.2: Ferromagnets and Electromagnets by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.20: Magnetic Fields and Magnetic Field Lines

## Learning Objectives

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

- Define magnetic field and describe the magnetic field lines of various magnetic fields.

Einstein is said to have been fascinated by a compass as a child, perhaps musing on how the needle felt a force without direct physical contact. His ability to think deeply and clearly about action at a distance, particularly for gravitational, electric, and magnetic forces, later enabled him to create his revolutionary theory of relativity. Since magnetic forces act at a distance, we define a magnetic field to represent magnetic forces. The pictorial representation of magnetic field lines is very useful in visualizing the strength and direction of the magnetic field. As shown in Figure 7.20.1, the direction of magnetic field lines is defined to be the direction in which the north end of a compass needle points. The magnetic field is traditionally called the *B*-field.

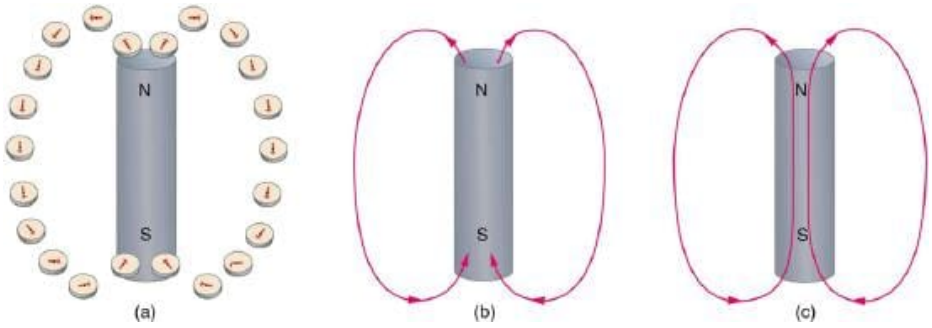



Figure 7.20.1: Magnetic field lines are defined to have the direction that a small compass points when placed at a location. (a) If small compasses are used to map the magnetic field around a bar magnet, they will point in the directions shown: away from the north pole of the magnet, toward the south pole of the magnet. (Recall that the Earth's north magnetic pole is really a south pole in terms of definitions of poles on a bar magnet.) (b) Connecting the arrows gives continuous magnetic field lines. The strength of the field is proportional to the closeness (or density) of the lines. (c) If the interior of the magnet could be probed, the field lines would be found to form continuous closed loops.

Small compasses used to test a magnetic field will not disturb it. (This is analogous to the way we tested electric fields with a small test charge. In both cases, the fields represent only the object creating them and not the probe testing them.) Figure 7.20.2 shows how the magnetic field appears for a current loop and a long straight wire, as could be explored with small compasses. A small compass placed in these fields will align itself parallel to the field line at its location, with its north pole pointing in the direction of *B*. Note the symbols used for field into and out of the paper.

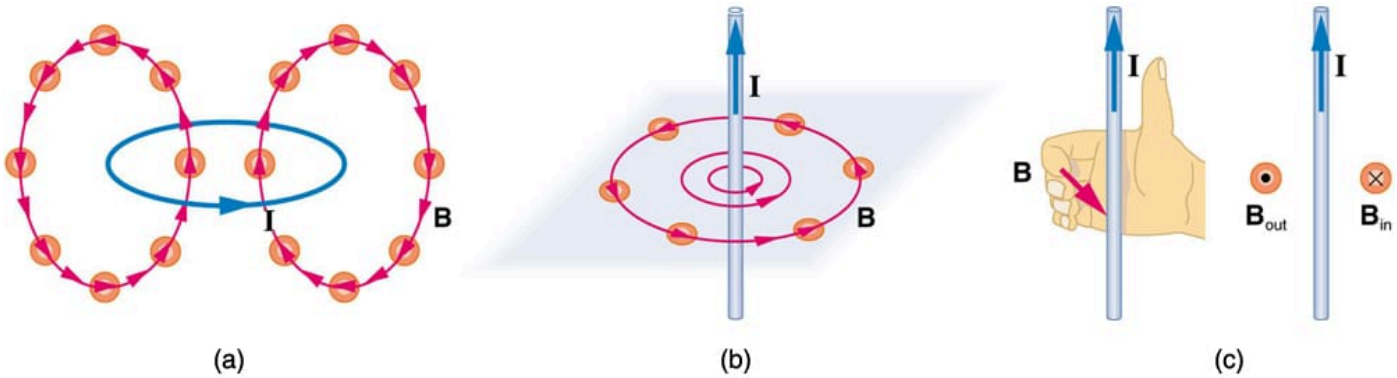



Figure 7.20.2: Small compasses could be used to map the fields shown here. (a) The magnetic field of a circular current loop is similar to that of a bar magnet. (b) A long and straight wire creates a field with magnetic field lines forming circular loops. (c)

When the wire is in the plane of the paper, the field is perpendicular to the paper. Note that the symbols used for the field pointing inward (like the tail of an arrow) and the field pointing outward (like the tip of an arrow).

### MAKING CONNECTIONS: CONCEPT OF A FIELD

A field is a way of mapping forces surrounding any object that can act on another object at a distance without apparent physical connection. The field represents the object generating it. Gravitational fields map gravitational forces, electric fields map electrical forces, and magnetic fields map magnetic forces.

Extensive exploration of magnetic fields has revealed a number of hard-and-fast rules. We use magnetic field lines to represent the field (the lines are a pictorial tool, not a physical entity in and of themselves). The properties of magnetic field lines can be summarized by these rules:

1. The direction of the magnetic field is tangent to the field line at any point in space. A small compass will point in the direction of the field line.
2. The strength of the field is proportional to the closeness of the lines. It is exactly proportional to the number of lines per unit area perpendicular to the lines (called the areal density).
3. Magnetic field lines can never cross, meaning that the field is unique at any point in space.
4. Magnetic field lines are continuous, forming closed loops without beginning or end. They go from the north pole to the south pole.

The last property is related to the fact that the north and south poles cannot be separated. It is a distinct difference from electric field lines, which begin and end on the positive and negative charges. If magnetic monopoles existed, then magnetic field lines would begin and end on them.

## Summary

- Magnetic fields can be pictorially represented by magnetic field lines, the properties of which are as follows:
  - The field is tangent to the magnetic field line.
  - Field strength is proportional to the line density.
  - Field lines cannot cross.
  - Field lines are continuous loops.

## Glossary

**magnetic field**
: the representation of magnetic forces

***B*-field**
: another term for magnetic field

**magnetic field lines**
: the pictorial representation of the strength and the direction of a magnetic field

**direction of magnetic field lines**
: the direction that the north end of a compass needle points

This page titled 7.20: Magnetic Fields and Magnetic Field Lines is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 22.3: Magnetic Fields and Magnetic Field Lines by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.21: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies

Nature's displays of symmetry are beautiful and alluring. A butterfly's wings exhibit an appealing symmetry in a complex system. (See Figure 2.) The laws of physics display symmetries at the most basic level—these symmetries are a source of wonder and imply deeper meaning. Since we place a high value on symmetry, we look for it when we explore nature. The remarkable thing is that we find it.



Figure 7.21.1: These wind turbines in the Thames Estuary in the UK are an example of induction at work. Wind pushes the blades of the turbine, spinning a shaft attached to magnets. The magnets spin around a conductive coil, inducing an electric current in the coil, and eventually feeding the electrical grid. (credit: modification of work by Petr Kratochvil)

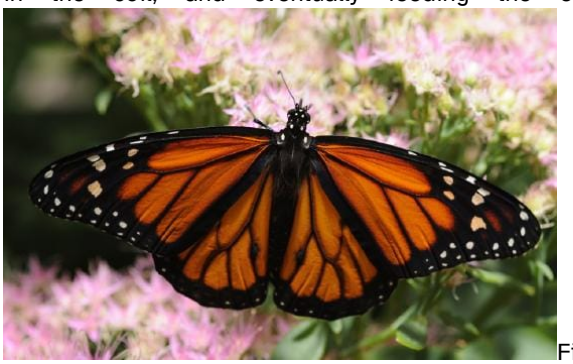


Figure 7.21.2: Physics, like this butterfly, has inherent symmetries. (credit: Thomas Bresson)

The hint of symmetry between electricity and magnetism found in the preceding chapter will be elaborated upon in this chapter. Specifically, we know that a current creates a magnetic field. If nature is symmetric here, then perhaps a magnetic field can create a current. The Hall effect is a voltage caused by a magnetic force. That voltage could drive a current. Historically, it was very shortly after Oersted discovered currents cause magnetic fields that other scientists asked the following question: Can magnetic fields cause currents? The answer was soon found by experiment to be yes. In 1831, some 12 years after Oersted's discovery, the English scientist Michael Faraday (1791–1862) and the American scientist Joseph Henry (1797–1878) independently demonstrated that magnetic fields can produce currents. The basic process of generating emfs (electromotive force) and, hence, currents with magnetic fields is known as **induction**; this process is also called magnetic induction to distinguish it from charging by induction, which utilizes the Coulomb force.

Today, currents induced by magnetic fields are essential to our technological society. The ubiquitous generator—found in

automobiles, on bicycles, in nuclear power plants, and so on—uses magnetism to generate current. Other devices that use magnetism to induce currents include pickup coils in electric guitars, transformers of every size, certain microphones, airport security gates, and damping mechanisms on sensitive chemical balances. Not so familiar perhaps, but important nevertheless, is that the behavior of AC circuits depends strongly on the effect of magnetic fields on currents.

## Glossary

**induction**

(magnetic induction) the creation of emfs and hence currents by magnetic fields

This page titled 7.21: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 23.0: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 7.06: Introduction to Electric Potential and Electric Energy

In Electric Charge and Electric Field, we just scratched the surface (or at least rubbed it) of electrical phenomena. Two of the most familiar aspects of electricity are its energy and *voltage*. We know, for example, that great amounts of electrical energy can be stored in batteries, are transmitted cross-country through power lines, and may jump from clouds to explode the sap of trees. In a similar manner, at molecular levels, *ions* cross cell membranes and transfer information. We also know about voltages associated with electricity. Batteries are typically a few volts, the outlets in your home produce 120 volts, and power lines can be as high as hundreds of thousands of volts. But energy and voltage are not the same thing. A motorcycle battery, for example, is small and would not be very successful in replacing the much larger car battery, yet each has the same voltage. In this chapter, we shall examine the relationship between voltage and electrical energy and begin to explore some of the many applications of electricity.



Figure 7.6.1: Automated external defibrillator unit (AED) (credit: U.S. Defense Department photo/Tech. Sgt. Suzanne M. Day)

This page titled 7.06: Introduction to Electric Potential and Electric Energy is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 19.0: Introduction to Electric Potential and Electric Energy by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/college-physics.

# 8: MODULE 6B-INTRODUCTION TO CHEMISTRY


8.1: Chemistry in Context

8.4: Prelude to Atoms

8.5: Early Ideas in Atomic Theory

8.6: Evolution of Atomic Theory

8.7: Atomic Structure and Symbolism

8.8: Chemical Formulas

8.9: The Periodic Table

**8.10: Organizing the Elements**

8.10.1: Mendeleev's Periodic Table

8.10.2: Periodic Law

8.10.3: Modern Periodic Table- Periods and Groups

8.11: Chemical Nomenclature

8.02: Phases and Classification of Matter

8.03: Physical and Chemical Properties



This page titled 8: Module 6b-Introduction to Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Arunava Roy.

# 8.1: Chemistry in Context

## Learning Objectives

- Outline the historical development of chemistry
- Provide examples of the importance of chemistry in everyday life
- Describe the scientific method and informs the scientific process
- Differentiate among hypotheses, theories, and laws
- Provide examples illustrating macroscopic, microscopic, and symbolic domains

Throughout human history, people have tried to convert matter into more useful forms. Our Stone Age ancestors chipped pieces of flint into useful tools and carved wood into statues and toys. These endeavors involved changing the shape of a substance without changing the substance itself. But as our knowledge increased, humans began to change the composition of the substances as well—clay was converted into pottery, hides were cured to make garments, copper ores were transformed into copper tools and weapons, and grain was made into bread.

Humans began to practice chemistry when they learned to control fire and use it to cook, make pottery, and smelt metals. Subsequently, they began to separate and use specific components of matter. A variety of drugs such as aloe, myrrh, and opium were isolated from plants. Dyes, such as indigo and Tyrian purple, were extracted from plant and animal matter. Metals were combined to form alloys—for example, copper and tin were mixed together to make bronze—and more elaborate smelting techniques produced iron. Alkalis were extracted from ashes, and soaps were prepared by combining these alkalis with fats. Alcohol was produced by fermentation and purified by distillation.

Attempts to understand the behavior of matter extend back for more than 2500 years. As early as the sixth century BC, Greek philosophers discussed a system in which water was the basis of all things. You may have heard of the Greek postulate that matter consists of four elements: earth, air, fire, and water. Subsequently, an amalgamation of chemical technologies and philosophical speculations were spread from Egypt, China, and the eastern Mediterranean by alchemists, who endeavored to transform “base metals” such as lead into “noble metals” like gold, and to create elixirs to cure disease and extend life (Figure 8.1.1).



Figure 8.1.1: This portrayal shows an alchemist’s workshop circa 1580. Although alchemy made some useful contributions to how to manipulate matter, it was not scientific by modern standards. (credit: Chemical Heritage Foundation). A sketch depicts 4 people stirring and handling chemicals. The chemicals are held in a variety of barrels and large cylinders. Several of the containers are being heated over burning embers. A large stove in the laboratory is filled with burning embers. There is also a large chest in the corner that is producing steam.

From alchemy came the historical progressions that led to modern chemistry: the isolation of drugs from natural sources,

metallurgy, and the dye industry. Today, chemistry continues to deepen our understanding and improve our ability to harness and control the behavior of matter. This effort has been so successful that many people do not realize either the central position of chemistry among the sciences or the importance and universality of chemistry in daily life.

## Chemistry: The Central Science

Chemistry is sometimes referred to as “the central science” due to its interconnectedness with a vast array of other STEM disciplines (STEM stands for areas of study in the science, technology, engineering, and math fields). Chemistry and the language of chemists play vital roles in biology, medicine, materials science, forensics, environmental science, and many other fields (Figure 8.1.2). The basic principles of physics are essential for understanding many aspects of chemistry, and there is extensive overlap between many subdisciplines within the two fields, such as chemical physics and nuclear chemistry. Mathematics, computer science, and information theory provide important tools that help us calculate, interpret, describe, and generally make sense of the chemical world. Biology and chemistry converge in biochemistry, which is crucial to understanding the many complex factors and processes that keep living organisms (such as us) alive. Chemical engineering, materials science, and nanotechnology combine chemical principles and empirical findings to produce useful substances, ranging from gasoline to fabrics to electronics. Agriculture, food science, veterinary science, and brewing and wine making help provide sustenance in the form of food and drink to the world’s population. Medicine, pharmacology, biotechnology, and botany identify and produce substances that help keep us healthy. Environmental science, geology, oceanography, and atmospheric science incorporate many chemical ideas to help us better understand and protect our physical world. Chemical ideas are used to help understand the universe in astronomy and cosmology.

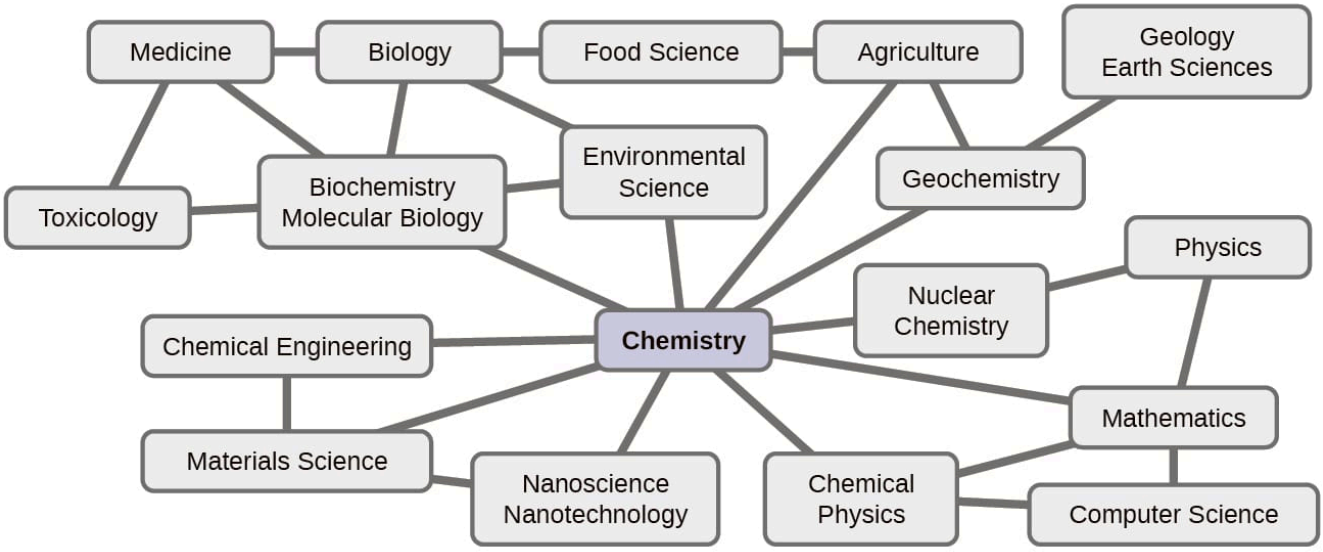


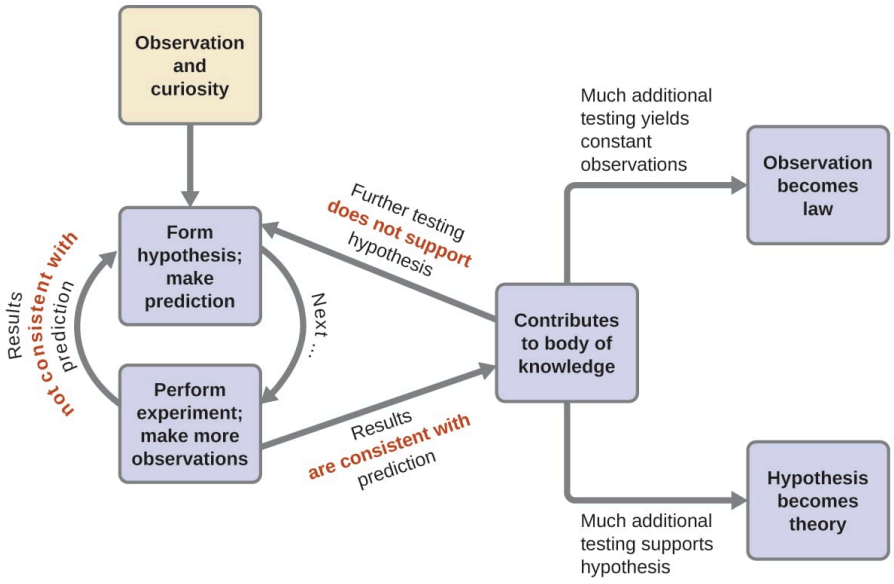

Figure 8.1.2: Knowledge of chemistry is central to understanding a wide range of scientific disciplines. This diagram shows just some of the interrelationships between chemistry and other fields. A flowchart shows a box containing chemistry at its center. Chemistry is connected to geochemistry, nuclear chemistry, chemical physics, nanoscience and nanotechnology, materials science, chemical engineering, biochemistry and molecular biology, environmental science, agriculture, and mathematics. Each of these disciplines is further connected to other related fields including medicine, biology, food science, geology earth sciences, toxicology, physics, and computer science.

What are some changes in matter that are essential to daily life? Digesting and assimilating food, synthesizing polymers that are used to make clothing, containers, cookware, and credit cards, and refining crude oil into gasoline and other products are just a few examples. As you proceed through this course, you will discover many different examples of changes in the composition and structure of matter, how to classify these changes and how they occurred, their causes, the changes in energy that accompany them, and the principles and laws involved. As you learn about these things, you will be learning chemistry, the study of the composition, properties, and interactions of matter. The practice of chemistry is not limited to chemistry books or laboratories: It happens whenever someone is involved in changes in matter or in conditions that may lead to such changes.

# The Scientific Method

Chemistry is a science based on observation and experimentation. Doing chemistry involves attempting to answer questions and explain observations in terms of the laws and theories of chemistry, using procedures that are accepted by the scientific community. There is no single route to answering a question or explaining an observation, but there is an aspect common to every approach: Each uses knowledge based on experiments that can be reproduced to verify the results. Some routes involve a hypothesis, a tentative explanation of observations that acts as a guide for gathering and checking information. We test a hypothesis by experimentation, calculation, and/or comparison with the experiments of others and then refine it as needed.

Some hypotheses are attempts to explain the behavior that is summarized in laws. The laws of science summarize a vast number of experimental observations, and describe or predict some facet of the natural world. If such a hypothesis turns out to be capable of explaining a large body of experimental data, it can reach the status of a theory. Scientific theories are well-substantiated, comprehensive, testable explanations of particular aspects of nature. Theories are accepted because they provide satisfactory explanations, but they can be modified if new data become available. The path of discovery that leads from question and observation to law or hypothesis to theory, combined with experimental verification of the hypothesis and any necessary modification of the theory, is called the scientific method (Figure 8.1.3).




*In this flowchart, the observation and curiosity box has an arrow pointing to a box labeled form hypothesis; make prediction. A curved arrow labeled next connects this box to a box labeled perform experiment; make more observations. Another arrow points back to the box that says form hypothesis; make prediction. This arrow is labeled results not consistent with prediction. Another arrow, labeled results are consistent with prediction points from the perform experiment box to a box labeled contributes to body of knowledge. However, an arrow also points from contributes to body of knowledge back to the form hypothesis; make prediction box. This arrow is labeled further testing does not support hypothesis. There are also two other arrows leading out from contributes to body of knowledge. One arrow is labeled much additional testing yields constant observations. This leads to the observation becomes law box. The other arrow is labeled much additional testing supports hypothesis. This arrow leads to the hypothesis becomes theory box.* Figure 8.1.3: The scientific method follows a process similar to the one shown in this diagram. All the key components are shown, in roughly the right order. Scientific progress is seldom neat and clean: It requires open inquiry and the reworking of questions and ideas in response to findings.

# The Domains of Chemistry

Chemists study and describe the behavior of matter and energy in three different domains: macroscopic, microscopic, and symbolic. These domains provide different ways of considering and describing chemical behavior.

*Macro* is a Greek word that means “large.” The macroscopic domain is familiar to us: It is the realm of everyday things that are large enough to be sensed directly by human sight or touch. In daily life, this includes the food you eat and the breeze you feel on your face. The macroscopic domain includes everyday and laboratory chemistry, where we observe and measure physical and chemical properties, or changes such as density, solubility, and flammability.

The microscopic domain of chemistry is almost always visited in the imagination. *Micro* also comes from Greek and means "small." Some aspects of the microscopic domains are visible through a microscope, such as a magnified image of graphite or bacteria. Viruses, for instance, are too small to be seen with the naked eye, but when we're suffering from a cold, we're reminded of how real they are.

However, most of the subjects in the microscopic domain of chemistry—such as atoms and molecules—are too small to be seen even with standard microscopes and often must be pictured in the mind. Other components of the microscopic domain include ions and electrons, protons and neutrons, and chemical bonds, each of which is far too small to see. This domain includes the individual metal atoms in a wire, the ions that compose a salt crystal, the changes in individual molecules that result in a color change, the conversion of nutrient molecules into tissue and energy, and the evolution of heat as bonds that hold atoms together are created.

The symbolic domain contains the specialized language used to represent components of the macroscopic and microscopic domains. Chemical symbols (such as those used in the periodic table), chemical formulas, and chemical equations are part of the symbolic domain, as are graphs and drawings. We can also consider calculations as part of the symbolic domain. These symbols play an important role in chemistry because they help interpret the behavior of the macroscopic domain in terms of the components of the microscopic domain. One of the challenges for students learning chemistry is recognizing that the same symbols can represent different things in the macroscopic and microscopic domains, and one of the features that makes chemistry fascinating is the use of a domain that must be imagined to explain behavior in a domain that can be observed.

A helpful way to understand the three domains is via the essential and ubiquitous substance of water. That water is a liquid at moderate temperatures, will freeze to form a solid at lower temperatures, and boil to form a gas at higher temperatures (Figure 1.1.4) are macroscopic observations. But some properties of water fall into the microscopic domain—what we cannot observe with the naked eye. The description of water as comprised of two hydrogen atoms and one oxygen atom, and the explanation of freezing and boiling in terms of attractions between these molecules, is within the microscopic arena. The formula $H_2O$, which can describe water at either the macroscopic or microscopic levels, is an example of the symbolic domain. The abbreviations (*g*) for gas, (s) for solid, and (*l*) for liquid are also symbolic.

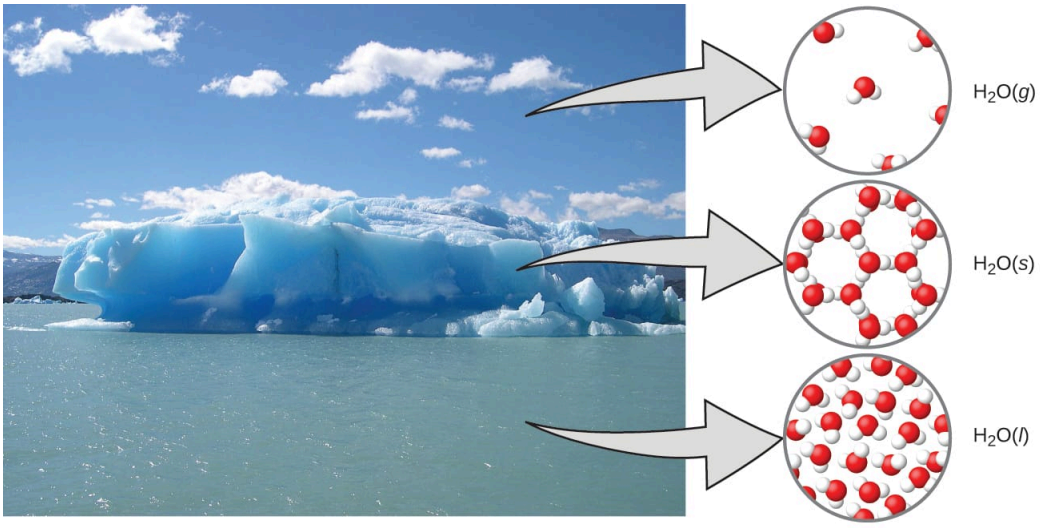



Figure 8.1.4: (a) Moisture in the air, icebergs, and the ocean represent water in the macroscopic domain. (b) At the molecular level (microscopic domain), gas molecules are far apart and disorganized, solid water molecules are close together and organized, and liquid molecules are close together and disorganized. (c) The formula $H_2O$ symbolizes water, and (g), (s), and (l) symbolize its phases. Note that clouds are actually comprised of either very small liquid water droplets or solid water crystals; gaseous water in our atmosphere is not visible to the naked eye, although it may be sensed as humidity. (credit a: modification of work by "Gorkaazk"/Wikimedia Commons). Figure A shows a photo of an iceberg floating in a sea has three arrows. Each arrow points to figure B, which contains three diagrams showing how the water molecules are organized in the air, ice, and sea. In the air, which contains the gaseous form of water, H subscript 2 O gas, the water molecules are disconnected and widely spaced. In the ice, which is the solid form of water, H subscript 2 O solid, the water molecules are bonded together into rings, with each ring containing six water molecules. Three of these rings are connected to each other. In the sea, which is the liquid form of water, H subscript 2 O liquid, the water molecules are very densely packed. The molecules are not bonded together.

## Key Concepts and Summary

Chemistry deals with the composition, structure, and properties of matter, and the ways by which various forms of matter may be interconverted. Thus, it occupies a central place in the study and practice of science and technology. Chemists use the scientific method to perform experiments, pose hypotheses, and formulate laws and develop theories, so that they can better understand the behavior of the natural world. To do so, they operate in the macroscopic, microscopic, and symbolic domains. Chemists measure, analyze, purify, and synthesize a wide variety of substances that are important to our lives.

## Glossary

**chemistry**
: study of the composition, properties, and interactions of matter

**hypothesis**
: tentative explanation of observations that acts as a guide for gathering and checking information

**law**
: statement that summarizes a vast number of experimental observations, and describes or predicts some aspect of the natural world

**macroscopic domain**
: realm of everyday things that are large enough to sense directly by human sight and touch

**microscopic domain**
: realm of things that are much too small to be sensed directly

**scientific method**
: path of discovery that leads from question and observation to law or hypothesis to theory, combined with experimental verification of the hypothesis and any necessary modification of the theory

**symbolic domain**
: specialized language used to represent components of the macroscopic and microscopic domains, such as chemical symbols, chemical formulas, chemical equations, graphs, drawings, and calculations

**theory**
: well-substantiated, comprehensive, testable explanation of a particular aspect of nature

This page titled 8.1: Chemistry in Context is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 11.34: Particle Accelerators and Detectors (Problems Q11.4.3) by OpenStax is licensed CC BY 4.0.

# 8.4: Prelude to Atoms

Your overall health and susceptibility to disease depends upon the complex interaction between your genetic makeup and environmental exposure, with the outcome difficult to predict. Early detection of biomarkers, substances that indicate an organism's disease or physiological state, could allow diagnosis and treatment before a condition becomes serious or irreversible. Recent studies have shown that your exhaled breath can contain molecules that may be biomarkers for recent exposure to environmental contaminants or for pathological conditions ranging from asthma to lung cancer.

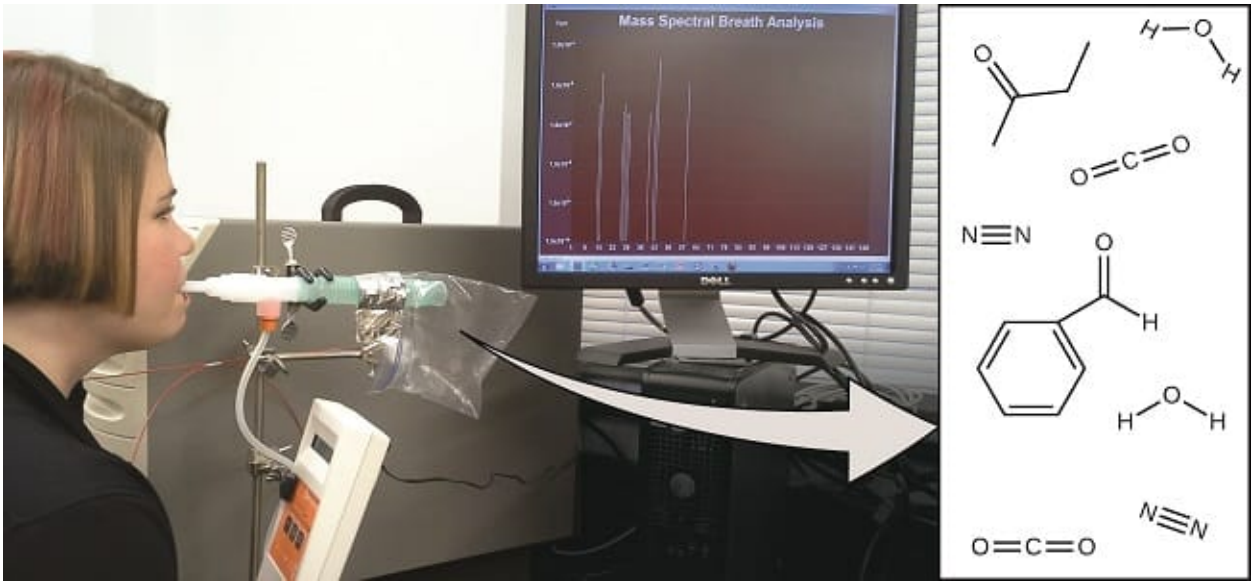



Figure 8.4.1: Analysis of molecules in an exhaled breath can provide valuable information, leading to early diagnosis of diseases or detection of environmental exposure to harmful substances. (credit: modification of work by Paul Flowers) A person is shown blowing into a tube connected to a plastic bag. There is a computer screen displaying data that reads "Mass Spectral Breath Analysis." An arrow from the plastic bag points to an illustration of different molecular compounds contained in the person's exhalation

Scientists are working to develop biomarker "fingerprints" that could be used to diagnose a specific disease based on the amounts and identities of certain molecules in a patient's exhaled breath. An essential concept underlying this goal is that of a molecule's identity, which is determined by the numbers and types of atoms it contains, and how they are bonded together. This chapter will describe some of the fundamental chemical principles related to the composition of matter, including those central to the concept of molecular identity.

This page titled 8.4: Prelude to Atoms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.5: Early Ideas in Atomic Theory

**Learning Objectives**

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

- State the postulates of Dalton's atomic theory
- Use postulates of Dalton's atomic theory to explain the laws of definite and multiple proportions

The language used in chemistry is seen and heard in many disciplines, ranging from medicine to engineering to forensics to art. The language of chemistry includes its own vocabulary as well as its own form of shorthand. Chemical symbols are used to represent atoms and elements. Chemical formulas depict molecules as well as the composition of compounds. Chemical equations provide information about the quality and quantity of the changes associated with chemical reactions.

This chapter will lay the foundation for our study of the language of chemistry. The concepts of this foundation include the atomic theory, the composition and mass of an atom, the variability of the composition of isotopes, ion formation, chemical bonds in ionic and covalent compounds, the types of chemical reactions, and the naming of compounds. We will also introduce one of the most powerful tools for organizing chemical knowledge: the periodic table.

## Atomic Theory through the Nineteenth Century

The earliest recorded discussion of the basic structure of matter comes from ancient Greek philosophers, the scientists of their day. In the fifth century BC, Leucippus and Democritus argued that all matter was composed of small, finite particles that they called *atomos*, a term derived from the Greek word for "indivisible." They thought of atoms as moving particles that differed in shape and size, and which could join together. Later, Aristotle and others came to the conclusion that matter consisted of various combinations of the four "elements"—fire, earth, air, and water—and could be infinitely divided. Interestingly, these philosophers thought about atoms and "elements" as philosophical concepts, but apparently never considered performing experiments to test their ideas.

The Aristotelian view of the composition of matter held sway for over two thousand years, until English schoolteacher John Dalton helped to revolutionize chemistry with his hypothesis that the behavior of matter could be explained using an atomic theory. First published in 1807, many of Dalton's hypotheses about the microscopic features of matter are still valid in modern atomic theory. Here are the postulates of Dalton's atomic theory.

1. Matter is composed of exceedingly small particles called atoms. An atom is the smallest unit of an element that can participate in a chemical change.
2. An element consists of only one type of atom, which has a mass that is characteristic of the element and is the same for all atoms of that element (Figure 8.5.1). A macroscopic sample of an element contains an incredibly large number of atoms, all of which have identical chemical properties.

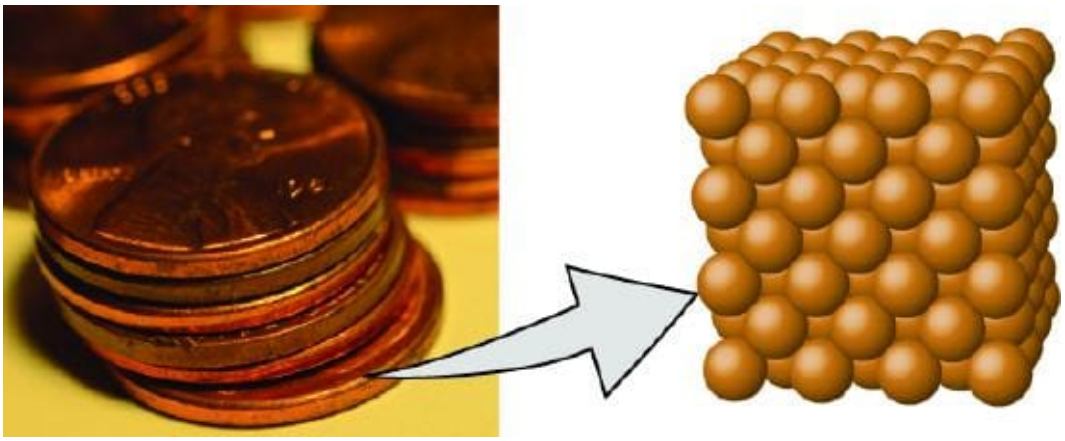


Figure 8.5.1: A pre-1982

copper penny (left) contains approximately $3 \times 10^{22}$ copper atoms (several dozen are represented as brown spheres at the right), each of which has the same chemical properties. (credit: modification of work by “slgckgc”/Flickr) The left image shows a photograph of a stack of pennies. The right image calls out an area of one of the pennies, which is made up of many sphere-shaped copper atoms. The atoms are densely organized.

3. Atoms of one element differ in properties from atoms of all other elements.
4. A compound consists of atoms of two or more elements combined in a small, whole-number ratio. In a given compound, the numbers of atoms of each of its elements are always present in the same ratio (Figure 8.5.2).

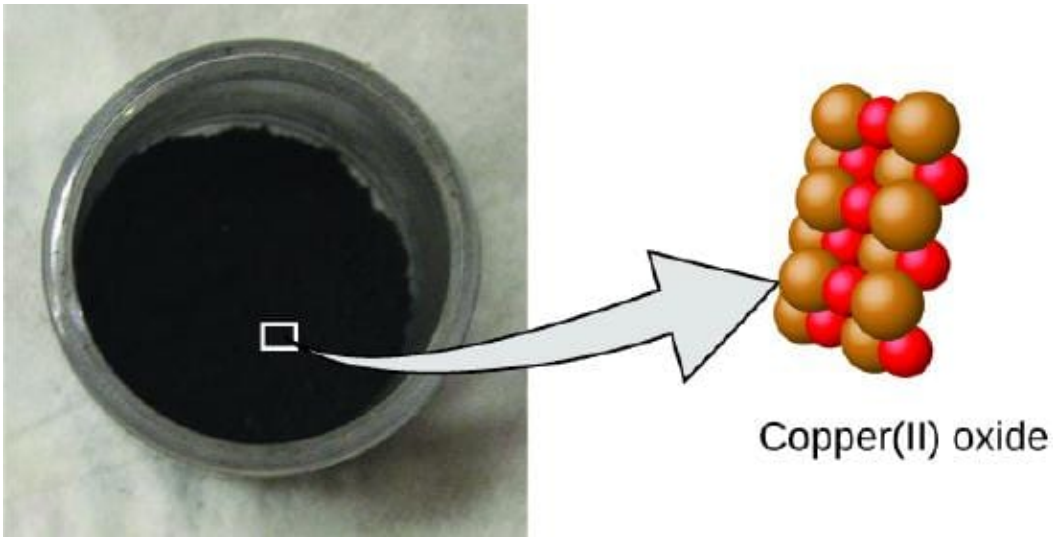



Figure 8.5.2: Copper(II) oxide, a powdery, black compound, results from the combination of two types of atoms—copper (brown spheres) and oxygen (red spheres)—in a 1:1 ratio. (credit: modification of work by “Chemicalinterest”/Wikimedia Commons) The left image shows a container with a black, powdery compound. The right image calls out the molecular structure of the powder which contains copper atoms that are clustered together with an equal number of oxygen atoms.

5. *Atoms are neither created nor destroyed during a chemical change, but are instead rearranged to yield substances that are different from those present before the change (Figure 8.5.3).*

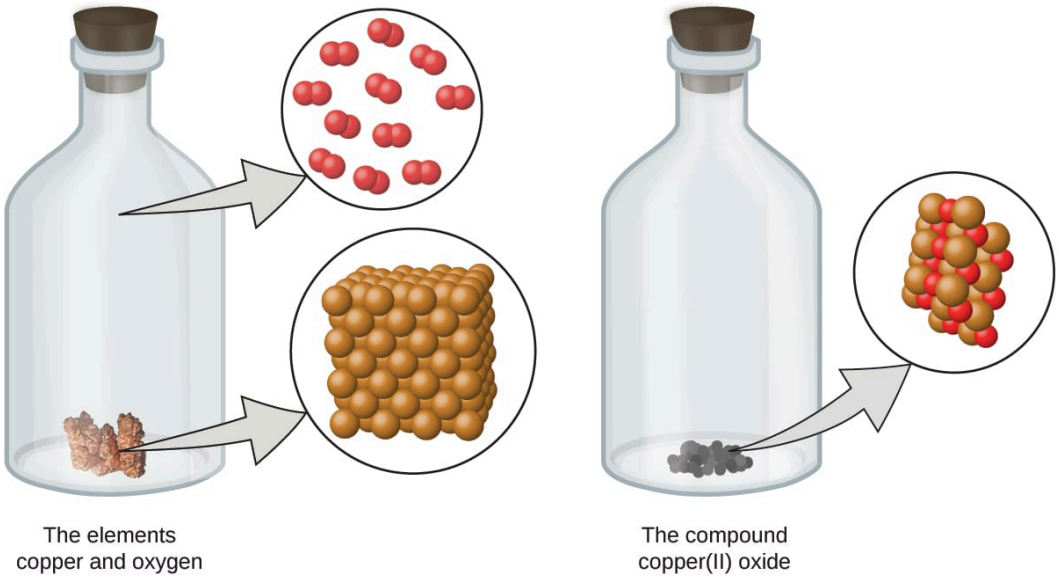



Figure 8.5.3:When the elements copper (a shiny, red-brown solid, shown here as brown spheres) and oxygen (a clear and colorless gas, shown here as red spheres) react, their atoms rearrange to form a compound containing copper and oxygen (a powdery, black solid). (credit copper: modification of work by Copper [images-of-elements.com]). The left stoppered bottle contains copper and oxygen. There is a callout which shows that copper is made up of many sphere-shaped atoms. The atoms are densely organized. The open space of the bottle contains oxygen gas, which is made up of bonded pairs of oxygen atoms that are evenly spaced. The right stoppered

bottle shows the compound copper two oxide. A callout from the powder shows a molecule of copper two oxide, which contains copper atoms that are clustered together with an equal number of oxygen atoms.

Dalton's atomic theory provides a microscopic explanation of the many macroscopic properties of matter that you've learned about. For example, if an element such as copper consists of only one kind of atom, then it cannot be broken down into simpler substances, that is, into substances composed of fewer types of atoms. And if atoms are neither created nor destroyed during a chemical change, then the total mass of matter present when matter changes from one type to another will remain constant (the law of conservation of matter).

## Example 8.5.1: Testing Dalton's Atomic Theory

In the following drawing, the green spheres represent atoms of a certain element. The purple spheres represent atoms of another element. If the spheres touch, they are part of a single unit of a compound. Does the following chemical change represented by these symbols violate any of the ideas of Dalton's atomic theory? If so, which one?



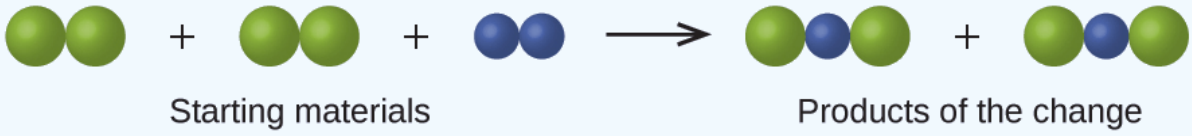

This equation shows that the starting materials of the reaction are two bonded, green spheres, which are being combined with two smaller, bonded purple spheres. The product of the change is one purple sphere that is bonded to one green sphere.

**Solution**

The starting materials consist of two green spheres and two purple spheres. The products consist of only one green sphere and one purple sphere. This violates Dalton's postulate that atoms are neither created nor destroyed during a chemical change, but are merely redistributed. (In this case, atoms appear to have been destroyed.)

## Exercise 8.5.1

In the following drawing, the green spheres represent atoms of a certain element. The purple spheres represent atoms of another element. If the spheres touch, they are part of a single unit of a compound. Does the following chemical change represented by these symbols violate any of the ideas of Dalton's atomic theory? If so, which one




This equation shows that the starting materials of the reaction are two sets of bonded, green spheres which are each being combined with two smaller, bonded purple spheres. The products of the change are two molecules that each contain one purple sphere bonded between two green spheres.

*Answer available at the end of this section.*

Dalton knew of the experiments of French chemist Joseph Proust, who demonstrated that *all samples of a pure compound contain the same elements in the same proportion by mass*. This statement is known as the law of definite proportions or the law of constant composition. The suggestion that the numbers of atoms of the elements in a given compound always exist in the same ratio is consistent with these observations. For example, when different samples of isooctane (a component of gasoline and one of the standards used in the octane rating system) are analyzed, they are found to have a carbon-to-hydrogen mass ratio of 5.33:1, as shown in Table 8.5.1.

*Table 8.5.1: Constant Composition of Isooctane*

| Sample | Carbon | Hydrogen | Mass Ratio |
|---|---|---|---|
| A | 14.82 g | 2.78 g | $\frac{14.82\text{ g carbon}}{2.78\text{ g hydrogen}} = \frac{5.33\text{ g carbon}}{1.00\text{ g hydrogen}}$ |
| B | 22.33 g | 4.19 g | $\frac{22.33\text{ g carbon}}{4.19\text{ g hydrogen}} = \frac{5.33\text{ g carbon}}{1.00\text{ g hydrogen}}$ |
| C | 19.40 g | 3.64 g | $\frac{19.40\text{ g carbon}}{3.63\text{ g hydrogen}} = \frac{5.33\text{ g carbon}}{1.00\text{ g hydrogen}}$ |

It is worth noting that although all samples of a particular compound have the same mass ratio, the converse is not true in general. That is, samples that have the same mass ratio are not necessarily the same substance. For example, there are many compounds other than isooctane that also have a carbon-to-hydrogen mass ratio of 5.33:1.00.

Dalton also used data from Proust, as well as results from his own experiments, to formulate another interesting law. The law of multiple proportions states that *when two elements react to form more than one compound, a fixed mass of one element will react with masses of the other element in a ratio of small, whole numbers*. For example, copper and chlorine can form a green, crystalline solid with a mass ratio of 0.558 g chlorine to 1 g copper, as well as a brown crystalline solid with a mass ratio of 1.116 g chlorine to 1 g copper. These ratios by themselves may not seem particularly interesting or informative; however, if we take a ratio of these ratios, we obtain a useful and possibly surprising result: a small, whole-number ratio.

$$\frac{\frac{1.116\text{ g Cl}}{1\text{ g Cu}}}{\frac{0.558\text{ g Cl}}{1\text{ g Cu}}} = \frac{2}{1}$$

This 2-to-1 ratio means that the brown compound has twice the amount of chlorine per amount of copper as the green compound.

This can be explained by atomic theory if the copper-to-chlorine ratio in the brown compound is 1 copper atom to 2 chlorine atoms, and the ratio in the green compound is 1 copper atom to 1 chlorine atom. The ratio of chlorine atoms (and thus the ratio of their masses) is therefore 2 to 1 (Figure 8.5.4).

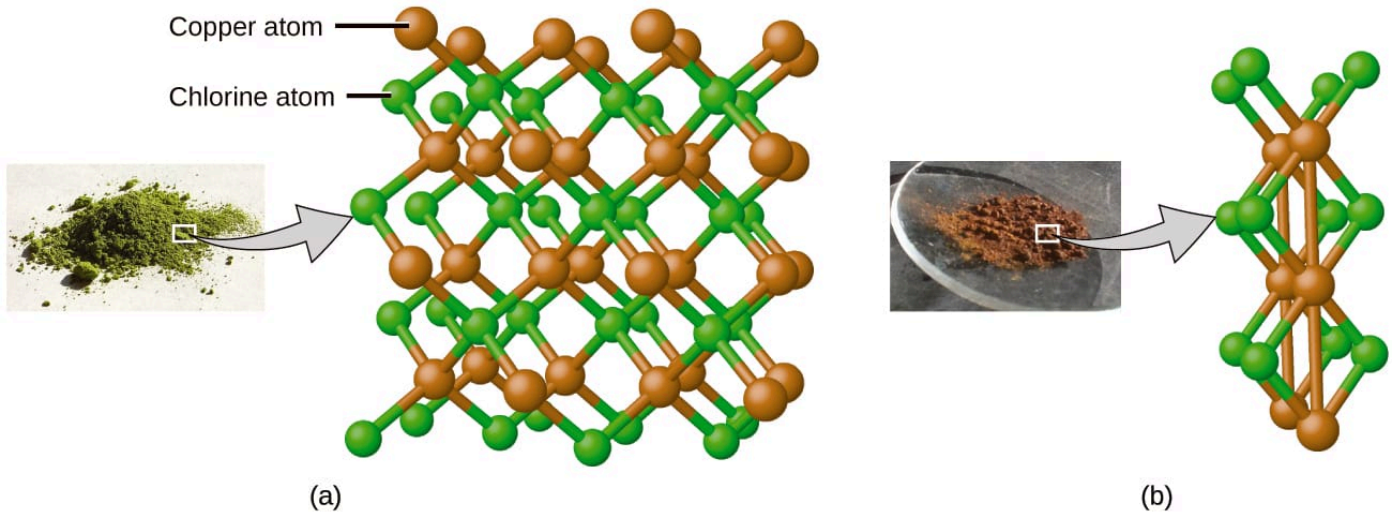



Figure 8.5.4: Compared to the copper chlorine compound in (a), where copper is represented by brown spheres and chlorine by green spheres, the copper chlorine compound in (b) has twice as many chlorine atoms per copper atom. (credit a: modification of work by “Benjah-bmm27”/Wikimedia Commons; credit b: modification of work by “Walkerma”/Wikimedia Commons) Figure A shows a pile of green powder. A callout shows that the green powder is made up of a lattice of copper atoms interspersed with chlorine atoms. The atoms are color coded brown for copper and green for chlorine. The number of copper atoms is equal to the number of chlorine atoms in the molecule. Figure B shows a pile of brown powder. A callout shows that the brown powder is also made up of copper and chlorine atoms similar to the molecule shown in figure A.

### Example 8.5.2: Laws of Definite and Multiple Proportions

A sample of compound A (a clear, colorless gas) is analyzed and found to contain 4.27 g carbon and 5.69 g oxygen. A sample of compound B (also a clear, colorless gas) is analyzed and found to contain 5.19 g carbon and 13.84 g oxygen. Are these data an example of the law of definite proportions, the law of multiple proportions, or neither? What do these data tell you about substances A and B?

**Solution**

In compound A, the mass ratio of carbon to oxygen is:

$$\frac{1.33\text{ g O}}{1\text{ g C}}$$

In compound B, the mass ratio of carbon to oxygen is:

$$\frac{2.67\text{ g O}}{1\text{ g C}}$$

The ratio of these ratios is:

$$\frac{\frac{1.33\text{ g O}}{1\text{ g C}}}{\frac{2.67\text{ g O}}{1\text{ g C}}} = \frac{1}{2}$$

This supports the law of multiple proportions. This means that A and B are different compounds, with A having one-half as much carbon per amount of oxygen (or twice as much oxygen per amount of carbon) as B. A possible pair of compounds that would fit this relationship would be A = $CO_2$ and B = CO.

### Exercise 8.5.1

A sample of compound X (a clear, colorless, combustible liquid with a noticeable odor) is analyzed and found to contain 14.13 g carbon and 2.96 g hydrogen. A sample of compound Y (a clear, colorless, combustible liquid with a noticeable odor that is slightly different from X's odor) is analyzed and found to contain 19.91 g carbon and 3.34 g hydrogen. Are these data an example of the law of definite proportions, the law of multiple proportions, or neither? What do these data tell you about substances X and Y?

*Answer available at the end of this section.*

## Summary

The ancient Greeks proposed that matter consists of extremely small particles called atoms. Dalton postulated that each element has a characteristic type of atom that differs in properties from atoms of all other elements, and that atoms of different elements can combine in fixed, small, whole-number ratios to form compounds. Samples of a particular compound all have the same elemental proportions by mass. When two elements form different compounds, a given mass of one element will combine with masses of the other element in a small, whole-number ratio. During any chemical change, atoms are neither created nor destroyed.

## Glossary

**Dalton's atomic theory**
: set of postulates that established the fundamental properties of atoms

**law of constant composition**
: (also, law of definite proportions) all samples of a pure compound contain the same elements in the same

proportions by mass

**law of multiple proportions**
: when two elements react to form more than one compound, a fixed mass of one element will react with masses of the other element in a ratio of small whole numbers

**law of definite proportions**
: (also, law of constant composition) all samples of a pure compound contain the same elements in the same proportions by mass


- Paul Flowers (University of North Carolina - Pembroke), Klaus Theopold (University of Delaware) and Richard Langley (Stephen F. Austin State University) with contributing authors. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/85abf193-2bd...a7ac8df6@9.110).


## Answers

**Exercise** 8.5.1

The starting materials consist of four green spheres and two purple spheres. The products consist of four green spheres and two purple spheres. This does not violate any of Dalton's postulates: Atoms are neither created nor destroyed, but are redistributed in small, whole-number ratios.

**Exercise** 8.5.1

In compound X, the mass ratio of carbon to hydrogen is $\frac{14.13\ \text{g C}}{2.96\ \text{g H}}$.

In compound Y, the mass ratio of carbon to oxygen is $\frac{19.91\ \text{g C}}{3.34\ \text{g H}}$.

The ratio of these ratios is

$$\frac{\frac{14.13\ \text{g C}}{2.96\ \text{g H}}}{\frac{19.91\ \text{g C}}{3.34\ \text{g H}}} = \frac{4.77\ \text{g C/g H}}{5.96\ \text{g C/g H}} = 0.800 = \frac{4}{5}.$$

This small, whole-number ratio supports the law of multiple proportions. This means that X and Y are different compounds.


This page titled 8.5: Early Ideas in Atomic Theory is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.6: Evolution of Atomic Theory

**Learning Objectives**

- Outline milestones in the development of modern atomic theory
- Summarize and interpret the results of the experiments of Thomson, Millikan, and Rutherford
- Describe the three subatomic particles that compose atoms
- Introduce the term isotopes

In the two centuries since Dalton developed his ideas, scientists have made significant progress in furthering our understanding of atomic theory. Much of this came from the results of several seminal experiments that revealed the details of the internal structure of atoms. Here, we will discuss some of those key developments, with an emphasis on application of the scientific method, as well as understanding how the experimental evidence was analyzed. While the historical persons and dates behind these experiments can be quite interesting, it is most important to understand the concepts resulting from their work.

## Atomic Theory after the Nineteenth Century

If matter were composed of atoms, what were atoms composed of? Were they the smallest particles, or was there something smaller? In the late 1800s, a number of scientists interested in questions like these investigated the electrical discharges that could be produced in low-pressure gases, with the most significant discovery made by English physicist J. J. Thomson using a cathode ray tube. This apparatus consisted of a sealed glass tube from which almost all the air had been removed; the tube contained two metal electrodes. When high voltage was applied across the electrodes, a visible beam called a cathode ray appeared between them. This beam was deflected toward the positive charge and away from the negative charge, and was produced in the same way with identical properties when different metals were used for the electrodes. In similar experiments, the ray was simultaneously deflected by an applied magnetic field, and measurements of the extent of deflection and the magnetic field strength allowed Thomson to calculate the charge-to-mass ratio of the cathode ray particles. The results of these measurements indicated that these particles were much lighter than atoms (Figure 8.6.1).

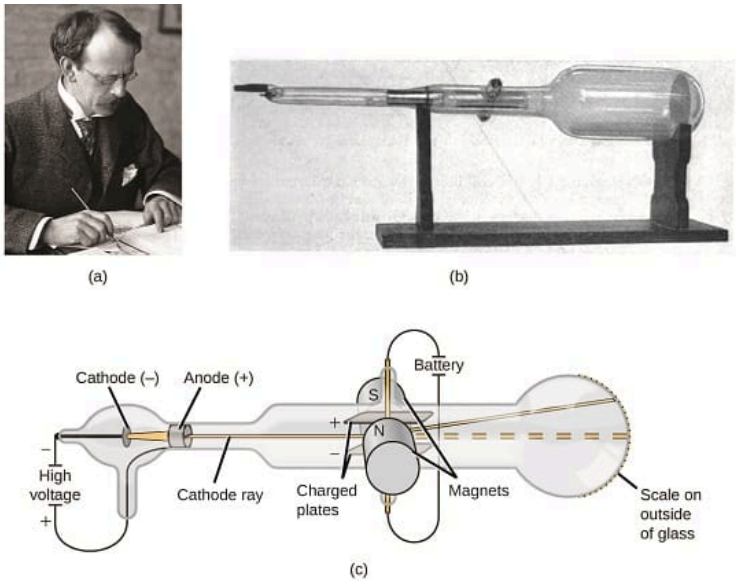
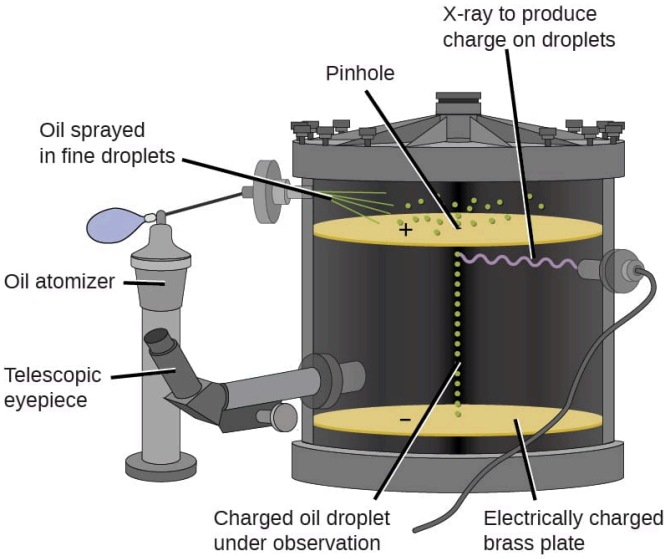



Figure 8.6.1: (a) J. J. Thomson produced a visible beam in a cathode ray tube. (b) This is an early cathode ray tube, invented in 1897 by Ferdinand Braun. (c) In the cathode ray, the beam (shown in yellow) comes from the cathode and is accelerated past the anode toward a fluorescent scale at the end of the tube. Simultaneous deflections by applied electric and magnetic fields permitted Thomson to calculate the mass-to-charge ratio of the particles composing the cathode ray. (credit a: modification of work by Nobel Foundation; credit b: modification of work by Eugen Nesper; credit c: modification of work by “Kurzon”/Wikimedia Commons).

Based on his observations, here is what Thomson proposed and why: The particles are attracted by positive (+) charges and repelled by negative (–) charges, so they must be negatively charged (like charges repel and unlike charges attract); they are less massive than atoms and indistinguishable, regardless of the source material, so they must be fundamental, subatomic constituents of all atoms. Although controversial at the time, Thomson's idea was gradually accepted, and his cathode ray particle is what we now call an electron, a negatively charged, subatomic particle with a mass more than one thousand-times less that of an atom. The term "electron" was coined in 1891 by Irish physicist George Stoney, from "*electr*ic i*on*."

In 1909, more information about the electron was uncovered by American physicist Robert A. Millikan via his "oil drop" experiments. Millikan created microscopic oil droplets, which could be electrically charged by friction as they formed or by using X-rays. These droplets initially fell due to gravity, but their downward progress could be slowed or even reversed by an electric field lower in the apparatus. By adjusting the electric field strength and making careful measurements and appropriate calculations, Millikan was able to determine the charge on individual drops (Figure 8.6.2).




| Oil drop | Charge in coulombs (C) |
|---|---|
| A | $4.8 \times 10^{-19}$ C |
| B | $3.2 \times 10^{-19}$ C |
| C | $6.4 \times 10^{-19}$ C |
| D | $1.6 \times 10^{-19}$ C |
| E | $4.8 \times 10^{-19}$ C |

Figure 8.6.2: Millikan's experiment measured the charge of individual oil drops. The tabulated data are examples of a few possible values.

Looking at the charge data that Millikan gathered, you may have recognized that the charge of an oil droplet is always a multiple of a specific charge, $1.6 \times 10^{-19}$ C. Millikan concluded that this value must therefore be a fundamental charge—the charge of a single electron—with his measured charges due to an excess of one electron (1 times $1.6 \times 10^{-19}$ C), two electrons (2 times $1.6 \times 10^{-19}$ C), three electrons (3 times $1.6 \times 10^{-19}$ C), and so on, on a given oil droplet. Since the charge of an electron was now known due to Millikan's research, and the charge-to-mass ratio was already known due to Thomson's research ($1.759 \times 10^{11}$ C/kg), it only required a simple calculation to determine the mass of the electron as well.

$$\text{Mass of electron} = 1.602 \times 10^{-19} \cancel{\text{C}} \times \frac{1 \text{ kg}}{1.759 \times 10^{11} \cancel{\text{C}}} = 9.107 \times 10^{-31} \text{ kg} \tag{2.3.1}$$

Scientists had now established that the atom was not indivisible as Dalton had believed, and due to the work of Thomson, Millikan, and others, the charge and mass of the negative, subatomic particles—the electrons—were known. However, the positively charged part of an atom was not yet well understood. In 1904, Thomson proposed the "plum pudding" model of atoms, which described a positively charged mass with an equal amount of negative charge in the form of electrons embedded in it, since all atoms are electrically neutral. A competing model had been proposed in 1903 by Hantaro Nagaoka, who postulated a Saturn-like atom, consisting of a positively charged sphere surrounded by a halo of electrons (Figure 8.6.3).

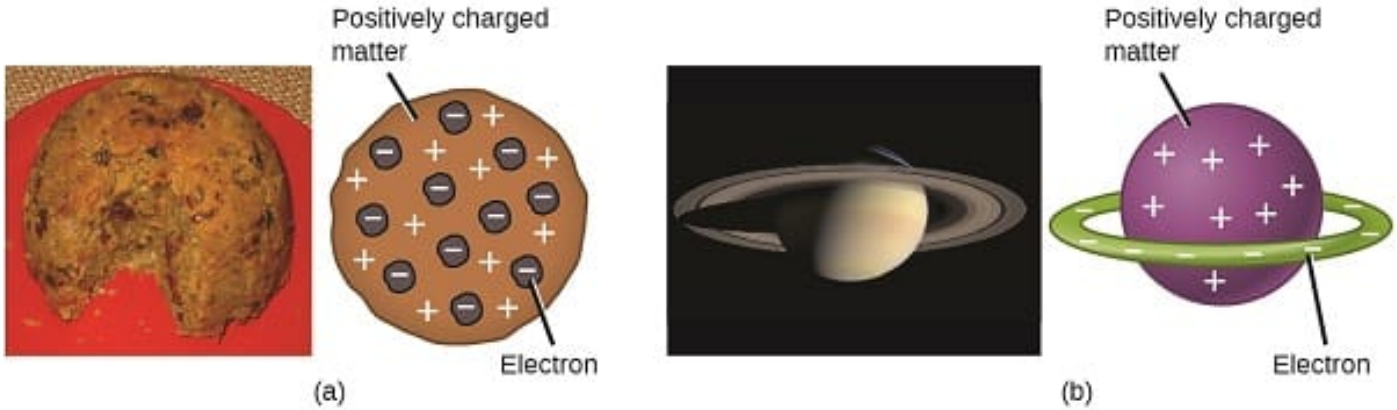



Figure 8.6.3: (a) Thomson suggested that atoms resembled plum pudding, an English dessert consisting of moist cake with embedded raisins (“plums”). (b) Nagaoka proposed that atoms resembled the planet Saturn, with a ring of electrons surrounding a positive “planet.” (credit a: modification of work by “Man vyi”/Wikimedia Commons; credit b: modification of work by “NASA”/Wikimedia Commons).

The next major development in understanding the atom came from Ernest Rutherford, a physicist from New Zealand who largely spent his scientific career in Canada and England. He performed a series of experiments using a beam of high-speed, positively charged alpha particles (α particles) that were produced by the radioactive decay of radium; α particles consist of two protons and two neutrons (you will learn more about radioactive decay in the chapter on nuclear chemistry). Rutherford and his colleagues Hans Geiger (later famous for the Geiger counter) and Ernest Marsden aimed a beam of α particles, the source of which was embedded in a lead block to absorb most of the radiation, at a very thin piece of gold foil and examined the resultant scattering of the α particles using a luminescent screen that glowed briefly where hit by an α particle.

What did they discover? Most particles passed right through the foil without being deflected at all. However, some were diverted slightly, and a very small number were deflected almost straight back toward the source (Figure 8.6.4). Rutherford described finding these results: “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”[1] (p. 68).

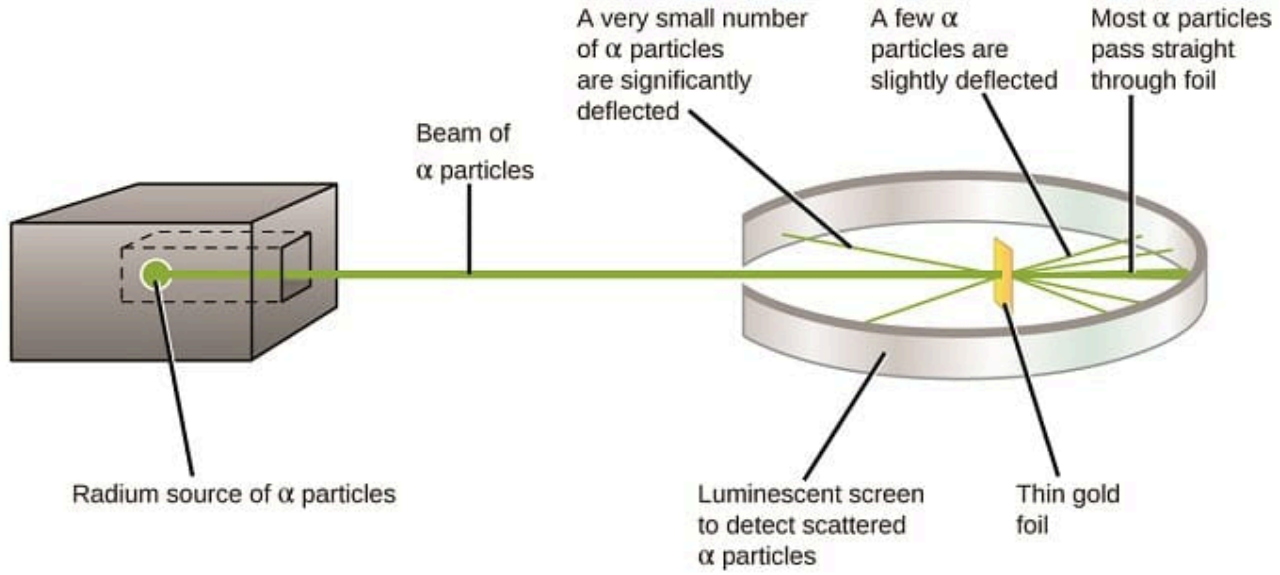



Figure 8.6.4: Geiger and Rutherford fired α particles at a piece of gold foil and detected where those particles went, as shown in this schematic diagram of their experiment. Most of the particles passed straight through the foil, but a few were deflected slightly and a very small number were significantly deflected.

Here is what Rutherford deduced: Because most of the fast-moving α particles passed through the gold atoms undeflected, they must have traveled through essentially empty space inside the atom. Alpha particles are positively charged, so deflections arose when they encountered another positive charge (like charges repel each other). Since like charges repel one another, the few

positively charged α particles that changed paths abruptly must have hit, or closely approached, another body that also had a highly concentrated, positive charge. Since the deflections occurred a small fraction of the time, this charge only occupied a small amount of the space in the gold foil. Analyzing a series of such experiments in detail, Rutherford drew two conclusions:

1. The volume occupied by an atom must consist of a large amount of empty space.
2. A small, relatively heavy, positively charged body, the nucleus, must be at the center of each atom.

This analysis led Rutherford to propose a model in which an atom consists of a very small, positively charged nucleus, in which most of the mass of the atom is concentrated, surrounded by the negatively charged electrons, so that the atom is electrically neutral (Figure 8.6.5).

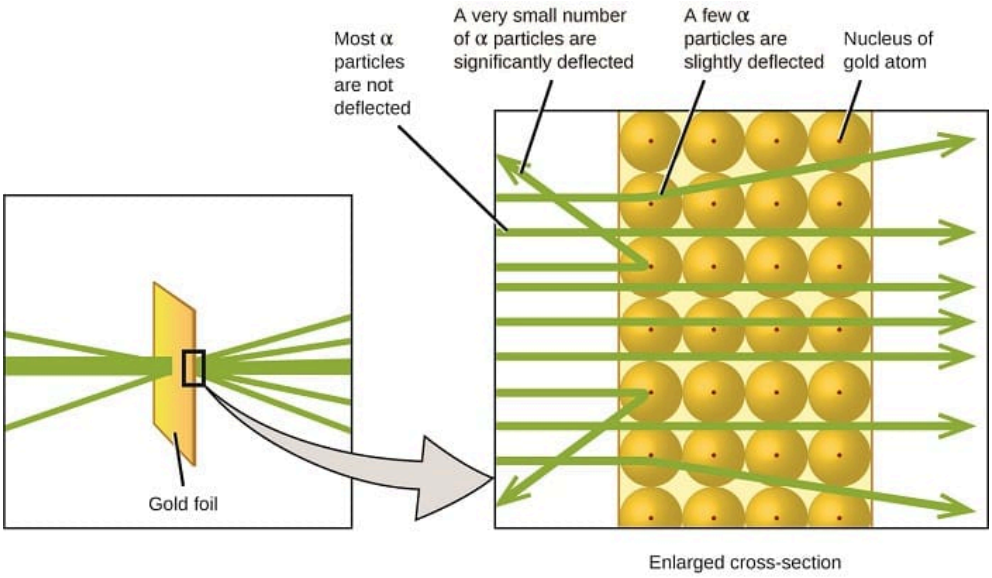



Figure 8.6.5: The α particles are deflected only when they collide with or pass close to the much heavier, positively charged gold nucleus. Because the nucleus is very small compared to the size of an atom, very few α particles are deflected. Most pass through the relatively large region occupied by electrons, which are too light to deflect the rapidly moving particles.

After many more experiments, Rutherford also discovered that the nuclei of other elements contain the hydrogen nucleus as a "building block," and he named this more fundamental particle the proton, the positively charged, subatomic particle found in the nucleus. With one addition, which you will learn next, this nuclear model of the atom, proposed over a century ago, is still used today.

Another important finding was the discovery of isotopes. During the early 1900s, scientists identified several substances that appeared to be new elements, isolating them from radioactive ores. For example, a "new element" produced by the radioactive decay of thorium was initially given the name mesothorium. However, a more detailed analysis showed that mesothorium was chemically identical to radium (another decay product), despite having a different atomic mass. This result, along with similar findings for other elements, led the English chemist Frederick Soddy to realize that an element could have types of atoms with different masses that were chemically indistinguishable. These different types are called isotopes—atoms of the same element that differ in mass. Soddy was awarded the Nobel Prize in Chemistry in 1921 for this discovery.

One puzzle remained: The nucleus was known to contain almost all of the mass of an atom, with the number of protons only providing half, or less, of that mass. Different proposals were made to explain what constituted the remaining mass, including the existence of neutral particles in the nucleus. As you might expect, detecting uncharged particles is very challenging, and it was not until 1932 that James Chadwick found evidence of neutrons, uncharged, subatomic particles with a mass approximately the same as that of protons. The existence of the neutron also explained isotopes: They differ in mass because they have different numbers of neutrons, but they are chemically identical because they have the same number of protons. This will be explained in more detail later in this chapter.

## Summary

Although no one has actually seen the inside of an atom, experiments have demonstrated much about atomic structure.

Thomson's cathode ray tube showed that atoms contain small, negatively charged particles called electrons. Millikan discovered that there is a fundamental electric charge—the charge of an electron. Rutherford's gold foil experiment showed that atoms have a small, dense, positively charged nucleus; the positively charged particles within the nucleus are called protons. Chadwick discovered that the nucleus also contains neutral particles called neutrons. Soddy demonstrated that atoms of the same element can differ in mass; these are called isotopes.

## Footnotes

1. Ernest Rutherford, "The Development of the Theory of Atomic Structure," ed. J. A. Ratcliffe, in *Background to Modern Science*, eds. Joseph Needham and Walter Pagel, (Cambridge, UK: Cambridge University Press, 1938), 61–74. Accessed September 22, 2014, https://ia600508.us.archive.org/3/it...e032734mbp.pdf.

## Glossary

**alpha particle (α particle)**
: positively charged particle consisting of two protons and two neutrons

**electron**
: negatively charged, subatomic particle of relatively low mass located outside the nucleus

**isotopes**
: atoms that contain the same number of protons but different numbers of neutrons

**neutron**
: uncharged, subatomic particle located in the nucleus

**proton**
: positively charged, subatomic particle located in the nucleus

**nucleus**
: massive, positively charged center of an atom made up of protons and neutrons

This page titled 8.6: Evolution of Atomic Theory is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.7: Atomic Structure and Symbolism

## Learning Objectives

- Write and interpret symbols that depict the atomic number, mass number, and charge of an atom or ion
- Define the atomic mass unit and average atomic mass
- Calculate average atomic mass and isotopic abundance

The development of modern atomic theory revealed much about the inner structure of atoms. It was learned that an atom contains a very small nucleus composed of positively charged protons and uncharged neutrons, surrounded by a much larger volume of space containing negatively charged electrons. The nucleus contains the majority of an atom's mass because protons and neutrons are much heavier than electrons, whereas electrons occupy almost all of an atom's volume. The diameter of an atom is on the order of $10^{-10}$ m, whereas the diameter of the nucleus is roughly $10^{-15}$ m—about 100,000 times smaller. For a perspective about their relative sizes, consider this: If the nucleus were the size of a blueberry, the atom would be about the size of a football stadium (Figure 8.7.1).

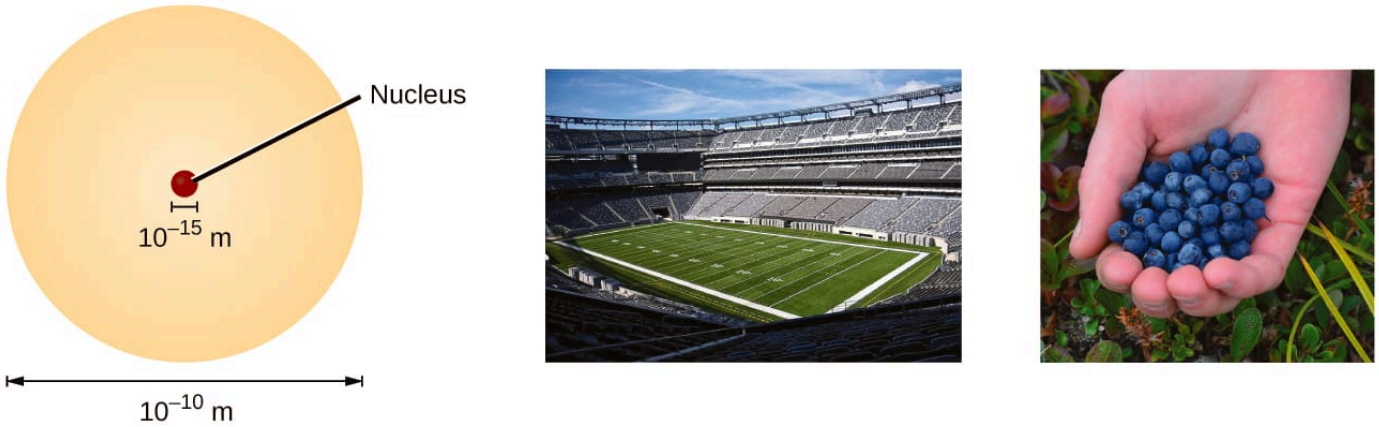



Figure 8.7.1: If an atom could be expanded to the size of a football stadium, the nucleus would be the size of a single blueberry. (credit middle: modification of work by "babyknight"/Wikimedia Commons; credit right: modification of work by Paxson Woelber). The diagram on the left shows a picture of an atom with a nucleus. The central figure shows a photograph of an American football stadium. The figure on the right shows a photograph of a handful of blueberries.

Atoms—and the protons, neutrons, and electrons that compose them—are extremely small. For example, a carbon atom weighs less than $2 \times 10^{-23}$ g, and an electron has a charge of less than $2 \times 10^{-19}$ C (coulomb). When describing the properties of tiny objects such as atoms, we use appropriately small units of measure, such as the atomic mass unit (amu) and the fundamental unit of charge (e). The amu was originally defined based on hydrogen, the lightest element, then later in terms of oxygen. Since 1961, it has been defined with regard to the most abundant isotope of carbon, atoms of which are assigned masses of exactly 12 amu. (This isotope is known as "carbon-12" as will be discussed later in this module.) Thus, one amu is exactly $1/12$ of the mass of one carbon-12 atom: 1 amu = $1.6605 \times 10^{-24}$ g. (The Dalton (Da) and the unified atomic mass unit (u) are alternative units that are equivalent to the amu.) The fundamental unit of charge (also called the elementary charge) equals the magnitude of the charge of an electron (e) with e = $1.602 \times 10^{-19}$ C.

A proton has a mass of 1.0073 amu and a charge of 1+. A neutron is a slightly heavier particle with a mass 1.0087 amu and a charge of zero; as its name suggests, it is neutral. The electron has a charge of 1– and is a much lighter particle with a mass of about 0.00055 amu (it would take about 1800 electrons to equal the mass of one proton. The properties of these fundamental particles are summarized in Table 8.7.1. (An observant student might notice that the sum of an atom's subatomic particles does not equal the atom's actual mass: The total mass of six protons, six neutrons, and six electrons is 12.0993 amu, slightly larger than the 12.00 amu of an actual carbon-12 atom. This "missing" mass is known as the mass defect, and you will learn about it in the chapter on nuclear chemistry.)

*Table 8.7.1: Properties of Subatomic Particles*

| Name | Location | Charge (C) | Unit Charge | Mass (amu) | Mass (g) |
|---|---|---|---|---|---|
| electron | outside nucleus | $-1.602 \times 10^{-19}$ | 1– | 0.00055 | $0.00091 \times 10^{-24}$ |
| proton | nucleus | $1.602 \times 10^{-19}$ | 1+ | 1.00727 | $1.67262 \times 10^{-24}$ |
| neutron | nucleus | 0 | 0 | 1.00866 | $1.67493 \times 10^{-24}$ |

The number of protons in the nucleus of an atom is its**atomic number** ($Z$). This is the defining trait of an element: Its value determines the identity of the atom. For example, any atom that contains six protons is the element carbon and has the atomic number 6, regardless of how many neutrons or electrons it may have. A neutral atom must contain the same number of positive and negative charges, so the number of protons equals the number of electrons. Therefore, the atomic number also indicates the number of electrons in an atom. The total number of protons and neutrons in an atom is called its mass number (A). The number of neutrons is therefore the difference between the mass number and the atomic number: A – Z = number of neutrons.

$$\begin{aligned} \text{atomic number (Z)} &= \text{number of protons} \\ \text{mass number (A)} &= \text{number of protons} + \text{number of neutrons} \\ \text{A} - \text{Z} &= \text{number of neutrons} \end{aligned}$$

Atoms are electrically neutral if they contain the same number of positively charged protons and negatively charged electrons. When the numbers of these subatomic particles are *not* equal, the atom is electrically charged and is called an ion. The charge of an atom is defined as follows:

Atomic charge = number of protons – number of electrons

As will be discussed in more detail later in this chapter, atoms (and molecules) typically acquire charge by gaining or losing electrons. An atom that gains one or more electrons will exhibit a negative charge and is called an anion. Positively charged atoms called cations are formed when an atom loses one or more electrons. For example, a neutral sodium atom (Z = 11) has 11 electrons. If this atom loses one electron, it will become a cation with a 1+ charge (11 – 10 = 1+). A neutral oxygen atom (Z = 8) has eight electrons, and if it gains two electrons it will become an anion with a 2– charge (8 – 10 = 2–).

## Example 8.7.1: Composition of an Atom

Iodine is an essential trace element in our diet; it is needed to produce thyroid hormone. Insufficient iodine in the diet can lead to the development of a goiter, an enlargement of the thyroid gland (Figure 8.7.2).



(a)




(b)

Figure 8.7.2: (a) Insufficient iodine in the diet can cause an enlargement of the thyroid gland called a goiter. (b) The addition of small amounts of iodine to salt, which prevents the formation of goiters, has helped eliminate this concern in the US where

salt consumption is high. (credit a: modification of work by “Almazi”/Wikimedia Commons; credit b: modification of work by Mike Mozart) Figure A shows a photo of a person who has a very swollen thyroid in his or her neck. Figure B shows a photo of a canister of Morton iodized salt.

The addition of small amounts of iodine to table salt (iodized salt) has essentially eliminated this health concern in the United States, but as much as 40% of the world’s population is still at risk of iodine deficiency. The iodine atoms are added as anions, and each has a 1– charge and a mass number of 127. Determine the numbers of protons, neutrons, and electrons in one of these iodine anions.

**Solution**

The atomic number of iodine (53) tells us that a neutral iodine atom contains 53 protons in its nucleus and 53 electrons outside its nucleus. Because the sum of the numbers of protons and neutrons equals the mass number, 127, the number of neutrons is 74 (127 – 53 = 74). Since the iodine is added as a 1– anion, the number of electrons is 54 [53 – (1–) = 54].

**Exercise 8.7.1**

An ion of platinum has a mass number of 195 and contains 74 electrons. How many protons and neutrons does it contain, and what is its charge?

*Answer available at the end of this section.*

## Chemical Symbols

A chemical symbol is an abbreviation that we use to indicate an element or an atom of an element. For example, the symbol for mercury is Hg (Figure 8.7.3). We use the same symbol to indicate one atom of mercury (microscopic domain) or to label a container of many atoms of the element mercury (macroscopic domain).

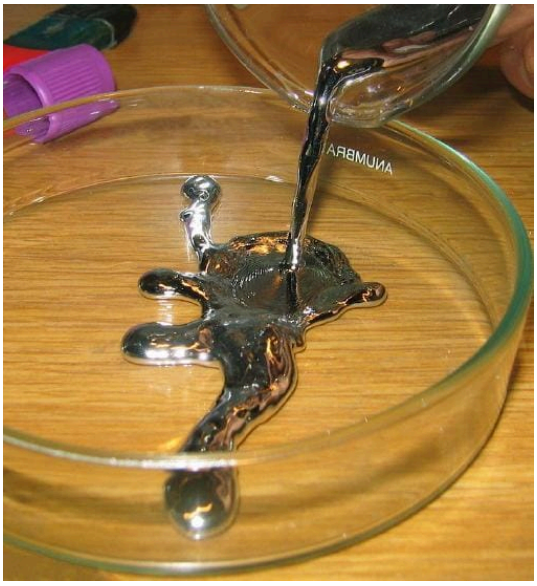


Figure 8.7.3: The symbol Hg represents the element mercury regardless of the amount; it could represent one atom of mercury or a large amount of mercury. from Wikipedia (user: Materialscientist).

The symbols for several common elements and their atoms are listed in Table 8.7.2. Some symbols are derived from the common name of the element; others are abbreviations of the name in another language. Symbols have one or two letters, for example, H for hydrogen and Cl for chlorine. To avoid confusion with other notations, only the first letter of a symbol is capitalized. For example, Co is the symbol for the element cobalt, but CO is the notation for the compound carbon monoxide, which contains atoms of the elements carbon (C) and oxygen (O). All known elements and their symbols are in the periodic table.

*Table 8.7.2: Some Common Elements and Their Symbols*

| Element | Symbol | Element | Symbol |
|---|---|---|---|
| aluminum | Al | iron | Fe (from *ferrum*) |
| bromine | Br | lead | Pb (from *plumbum*) |
| calcium | Ca | magnesium | Mg |
| carbon | C | mercury | Hg (from *hydrargyrum*) |
| chlorine | Cl | nitrogen | N |
| chromium | Cr | oxygen | O |
| cobalt | Co | potassium | K (from *kalium*) |
| copper | Cu (from *cuprum*) | silicon | Si |
| fluorine | F | silver | Ag (from *argentum*) |
| gold | Au (from *aurum*) | sodium | Na (from *natrium*) |
| helium | He | sulfur | S |
| hydrogen | H | tin | Sn (from *stannum*) |
| iodine | I | zinc | Zn |

Traditionally, the discoverer (or discoverers) of a new element names the element. However, until the name is recognized by the International Union of Pure and Applied Chemistry (IUPAC), the recommended name of the new element is based on the Latin word(s) for its atomic number. For example, element 106 was called unnilhexium (Unh), element 107 was called unnilseptium (Uns), and element 108 was called unniloctium (Uno) for several years. These elements are now named after scientists or locations; for example, element 106 is now known as *seaborgium* (Sg) in honor of Glenn Seaborg, a Nobel Prize winner who was active in the discovery of several heavy elements.

## Isotopes

The symbol for a specific isotope of any element is written by placing the mass number as a superscript to the left of the element symbol (Figure 8.7.4). The atomic number is sometimes written as a subscript preceding the symbol, but since this number defines the element's identity, as does its symbol, it is often omitted. For example, magnesium exists as a mixture of three isotopes, each with an atomic number of 12 and with mass numbers of 24, 25, and 26, respectively. These isotopes can be identified as $^{24}$Mg, $^{25}$Mg, and $^{26}$Mg. These isotope symbols are read as "element, mass number" and can be symbolized consistent with this reading. For instance, $^{24}$Mg is read as "magnesium 24," and can be written as "magnesium-24" or "Mg-24." $^{25}$Mg is read as "magnesium 25," and can be written as "magnesium-25" or "Mg-25." All magnesium atoms have 12 protons in their nucleus. They differ only because a $^{24}$Mg atom has 12 neutrons in its nucleus, a $^{25}$Mg atom has 13 neutrons, and a $^{26}$Mg has 14 neutrons.

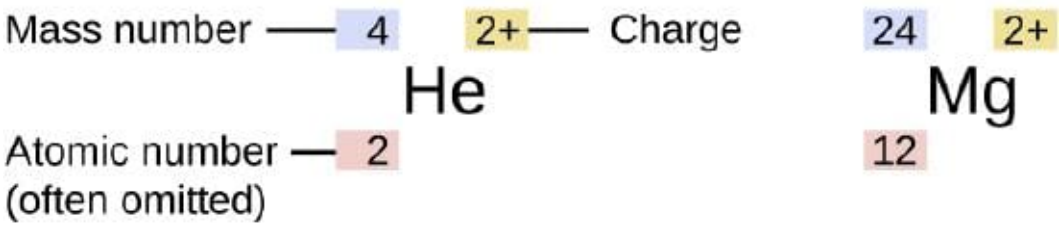



Figure 8.7.4: The symbol for an atom indicates the element via its usual two-letter symbol, the mass number as a left superscript, the atomic number as a left subscript (sometimes omitted), and the charge as a right superscript. This diagram shows the symbol for helium, "H e." The number to the upper left of the symbol is the mass number, which is 4. The number to the upper right of the symbol is the charge which is positive 2. The number to the lower left of the symbol is the atomic number, which is 2. This number is often omitted. Also shown is "M g" which stands for magnesium It has a mass number of 24, a charge of positive 2, and an atomic number of 12.

Information about the naturally occurring isotopes of elements with atomic numbers 1 through 10 is given in Table 8.7.2. Note that in addition to standard names and symbols, the isotopes of hydrogen are often referred to using common names and accompanying symbols. Hydrogen-2, symbolized $^{2}H$, is also called deuterium and sometimes symbolized D. Hydrogen-3, symbolized $^{3}H$, is also called tritium and sometimes symbolized T.

*Table 8.7.2: Nuclear Compositions of Atoms of the Very Light Elements*

| Element | Symbol | Atomic Number | Number of Protons | Number of Neutrons | Mass (amu) | % Natural Abundance |
|---|---|---|---|---|---|---|
| hydrogen | $^{1}_{1}H$ (protium) | 1 | 1 | 0 | 1.0078 | 99.989 |
| | $^{2}_{1}H$ (deuterium) | 1 | 1 | 1 | 2.0141 | 0.0115 |
| | $^{3}_{1}H$ (tritium) | 1 | 1 | 2 | 3.01605 | — (trace) |
| helium | $^{3}_{2}He$ | 2 | 2 | 1 | 3.01603 | 0.00013 |
| | $^{4}_{2}He$ | 2 | 2 | 2 | 4.0026 | 100 |
| lithium | $^{6}_{3}Li$ | 3 | 3 | 3 | 6.0151 | 7.59 |
| | $^{7}_{3}Li$ | 3 | 3 | 4 | 7.0160 | 92.41 |
| beryllium | $^{9}_{4}Be$ | 4 | 4 | 5 | 9.0122 | 100 |
| boron | $^{10}_{5}B$ | 5 | 5 | 5 | 10.0129 | 19.9 |
| | $^{11}_{5}B$ | 5 | 5 | 6 | 11.0093 | 80.1 |
| carbon | $^{12}_{6}C$ | 6 | 6 | 6 | 12.0000 | 98.89 |
| | $^{13}_{6}C$ | 6 | 6 | 7 | 13.0034 | 1.11 |
| | $^{14}_{6}C$ | 6 | 6 | 8 | 14.0032 | — (trace) |
| nitrogen | $^{14}_{7}N$ | 7 | 7 | 7 | 14.0031 | 99.63 |
| | $^{15}_{7}N$ | 7 | 7 | 8 | 15.0001 | 0.37 |
| oxygen | $^{16}_{8}O$ | 8 | 8 | 8 | 15.9949 | 99.757 |
| | $^{17}_{8}O$ | 8 | 8 | 9 | 16.9991 | 0.038 |
| | $^{18}_{8}O$ | 8 | 8 | 10 | 17.9992 | 0.205 |
| fluorine | $^{19}_{9}F$ | 9 | 9 | 10 | 18.9984 | 100 |
| neon | $^{20}_{10}Ne$ | 10 | 10 | 10 | 19.9924 | 90.48 |
| | $^{21}_{10}Ne$ | 10 | 10 | 11 | 20.9938 | 0.27 |
| | $^{22}_{10}Ne$ | 10 | 10 | 12 | 21.9914 | 9.25 |

## Atomic Mass

Because each proton and each neutron contribute approximately one amu to the mass of an atom, and each electron contributes far less, the atomic mass of a single atom is approximately equal to its mass number (a whole number). However, the average masses of atoms of most elements are not whole numbers because most elements exist naturally as mixtures of two or more isotopes.

The mass of an element shown in a periodic table or listed in a table of atomic masses is a weighted, average mass of all the isotopes present in a naturally occurring sample of that element. This is equal to the sum of each individual isotope's mass multiplied by its fractional abundance.

$$\text{average mass} = \sum_i (\text{fractional abundance} \times \text{isotopic mass})_i$$

For example, the element boron is composed of two isotopes: About 19.9% of all boron atoms are $^{10}$B with a mass of 10.0129 amu, and the remaining 80.1% are $^{11}$B with a mass of 11.0093 amu. The average atomic mass for boron is calculated to be:

$$\begin{aligned}\text{boron average mass} &= (0.199 \times 10.0129\ \text{amu}) + (0.801 \times 11.0093\ \text{amu}) \\ &= 1.99\ \text{amu} + 8.82\ \text{amu} \\ &= 10.81\ \text{amu}\end{aligned}$$

It is important to understand that no single boron atom weighs exactly 10.8 amu; 10.8 amu is the average mass of all boron atoms, and individual boron atoms weigh either approximately 10 amu or 11 amu.

## Example 8.7.2: Calculation of Average Atomic Mass

A meteorite found in central Indiana contains traces of the noble gas neon picked up from the solar wind during the meteorite's trip through the solar system. Analysis of a sample of the gas showed that it consisted of 91.84% $^{20}$Ne (mass 19.9924 amu), 0.47% $^{21}$Ne (mass 20.9940 amu), and 7.69% $^{22}$Ne (mass 21.9914 amu). What is the average mass of the neon in the solar wind?

**Solution**

$$\begin{aligned}\text{average mass} &= (0.9184 \times 19.9924\ \text{amu}) + (0.0047 \times 20.9940\ \text{amu}) + (0.0769 \times 21.9914\ \text{amu}) \\ &= (18.36 + 0.099 + 1.69)\ \text{amu} \\ &= 20.15\ \text{amu}\end{aligned}$$

The average mass of a neon atom in the solar wind is 20.15 amu. (The average mass of a terrestrial neon atom is 20.1796 amu. This result demonstrates that we may find slight differences in the natural abundance of isotopes, depending on their origin.)

## Exercise 8.7.2

A sample of magnesium is found to contain 78.70% of $^{24}$Mg atoms (mass 23.98 amu), 10.13% of $^{25}$Mg atoms (mass 24.99 amu), and 11.17% of $^{26}$Mg atoms (mass 25.98 amu). Calculate the average mass of a Mg atom.

*Answer available at the end of this section.*

We can also do variations of this type of calculation, as shown in the next example.

## Example 8.7.3: Calculation of Percent Abundance

Naturally occurring chlorine consists of $^{35}$Cl (mass 34.96885 amu) and $^{37}$Cl (mass 36.96590 amu), with an average mass of 35.453 amu. What is the percent composition of Cl in terms of these two isotopes?

**Solution**

The average mass of chlorine is the fraction that is $^{35}$Cl times the mass of $^{35}$Cl plus the fraction that is $^{37}$Cl times the mass of $^{37}$Cl.

$$\text{average mass} = (\text{fraction of } ^{35}\text{Cl} \times \text{mass of } ^{35}\text{Cl}) + (\text{fraction of } ^{37}\text{Cl} \times \text{mass of } ^{37}\text{Cl})$$

If we let *x* represent the fraction that is $^{35}$Cl, then the fraction that is $^{37}$Cl is represented by 1.00 – *x*.

(The fraction that is $^{35}$Cl + the fraction that is $^{37}$Cl must add up to 1, so the fraction of $^{37}$Cl must equal 1.00 – the fraction of $^{35}$Cl.)

Substituting this into the average mass equation, we have:

$$\begin{aligned} 35.453\ \text{amu} &= (x \times 34.96885\ \text{amu}) + [(1.00 - x) \times 36.96590\ \text{amu}] \\ 35.453 &= 34.96885x + 36.96590 - 36.96590x \\ 1.99705x &= 1.513 \\ x &= \frac{1.513}{1.99705} = 0.7576 \end{aligned}$$

So solving yields: $x$ = 0.7576, which means that 1.00 – 0.7576 = 0.2424. Therefore, chlorine consists of 75.76% $^{35}$Cl and 24.24% $^{37}$Cl.

## Exercise 8.7.3

Naturally occurring copper consists of $^{63}$Cu (mass 62.9296 amu) and $^{65}$Cu (mass 64.9278 amu), with an average mass of 63.546 amu. What is the percent composition of Cu in terms of these two isotopes?

*Answer available at the end of this section.*

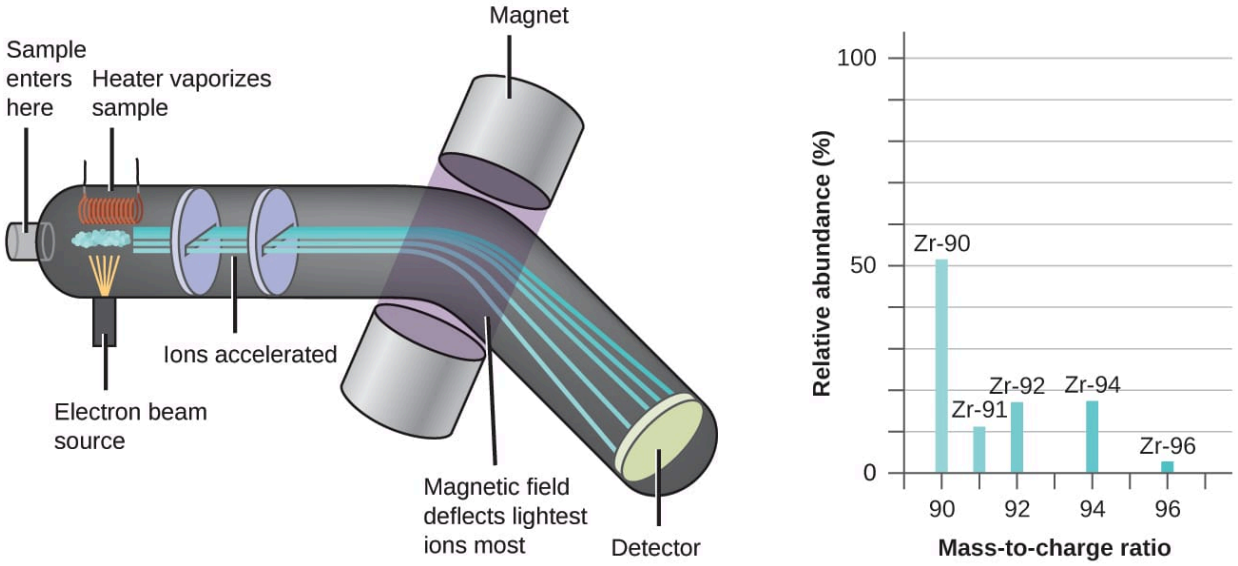



Figure 8.7.5: Analysis of zirconium in a mass spectrometer produces a mass spectrum with peaks showing the different isotopes of Zr. The left diagram shows how a mass spectrometer works. The graph to the right of the spectrometer shows a mass spectrum of zirconium. The relative abundance, as a percentage from 0 to 100, is graphed on the y axis, and the mass to charge ratio is graphed on the x axis. The sample contains five different isomers of zirconium.

The occurrence and natural abundances of isotopes can be experimentally determined using an instrument called a mass spectrometer. Mass spectrometry (MS) is widely used in chemistry, forensics, medicine, environmental science, and many other fields to analyze and help identify the substances in a sample of material. In a typical mass spectrometer (Figure 8.7.5), the sample is vaporized and exposed to a high-energy electron beam that causes the sample's atoms (or molecules) to become electrically charged, typically by losing one or more electrons. These cations then pass through a (variable) electric or magnetic field that deflects each cation's path to an extent that depends on both its mass and charge (similar to how the path of a large steel ball bearing rolling past a magnet is deflected to a lesser extent that that of a small steel BB). The ions are detected, and a plot of the relative number of ions generated versus their mass-to-charge ratios (a *mass spectrum*) is made. The height of each vertical feature or peak in a mass spectrum is proportional to the fraction of cations with the specified mass-to-charge ratio. Since its initial use during the development of modern atomic theory, MS has evolved to become a powerful tool for chemical analysis in a wide range of applications.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

***Video*** 8.7.1*: Watch this video from the Royal Society for Chemistry for a brief description of the rudiments of mass spectrometry.*

## Summary

An atom consists of a small, positively charged nucleus surrounded by electrons. The nucleus contains protons and neutrons; its diameter is about 100,000 times smaller than that of the atom. The mass of one atom is usually expressed in atomic mass units (amu), which is referred to as the atomic mass. An amu is defined as exactly $1/12$ of the mass of a carbon-12 atom and is equal to $1.6605 \times 10^{-24}$ g.

Protons are relatively heavy particles with a charge of 1+ and a mass of 1.0073 amu. Neutrons are relatively heavy particles with no charge and a mass of 1.0087 amu. Electrons are light particles with a charge of 1– and a mass of 0.00055 amu. The number of protons in the nucleus is called the atomic number (Z) and is the property that defines an atom's elemental identity. The sum of the numbers of protons and neutrons in the nucleus is called the mass number and, expressed in amu, is approximately equal to the mass of the atom. An atom is neutral when it contains equal numbers of electrons and protons.

Isotopes of an element are atoms with the same atomic number but different mass numbers; isotopes of an element, therefore, differ from each other only in the number of neutrons within the nucleus. When a naturally occurring element is composed of several isotopes, the atomic mass of the element represents the average of the masses of the isotopes involved. A chemical symbol identifies the atoms in a substance using symbols, which are one-, two-, or three-letter abbreviations for the atoms.

## Key Equations

- $\text{average mass} = \sum_i (\text{fractional abundance} \times \text{isotopic mass})_i$

## Glossary

**anion**
: negatively charged atom or molecule (contains more electrons than protons)

**atomic mass**
: average mass of atoms of an element, expressed in amu

**atomic mass unit (amu)**
: (also, unified atomic mass unit, u, or Dalton, Da) unit of mass equal to $\frac{1}{12}$ of the mass of a $^{12}C$ atom

**atomic number (Z)**
: number of protons in the nucleus of an atom

**cation**
: positively charged atom or molecule (contains fewer electrons than protons)

**chemical symbol**
: one-, two-, or three-letter abbreviation used to represent an element or its atoms

**Dalton (Da)**
: alternative unit equivalent to the atomic mass unit

**fundamental unit of charge**
: (also called the elementary charge) equals the magnitude of the charge of an electron (e) with $e = 1.602 \times 10^{-19}$ C

**ion**
: electrically charged atom or molecule (contains unequal numbers of protons and electrons)

**mass number (A)**
: sum of the numbers of neutrons and protons in the nucleus of an atom

**unified atomic mass unit (u)**
: alternative unit equivalent to the atomic mass unit

- Paul Flowers (University of North Carolina - Pembroke), Klaus Theopold (University of Delaware) and Richard Langley (Stephen F. Austin State University) with contributing authors. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/85abf193-2bd...a7ac8df6@9.110).

## Answers

**Exercise** 8.7.1

78 protons; 117 neutrons; charge is 4+

**Exercise** 8.7.2

24.31 amu

**Exercise** 8.7.3

69.15% Cu-63 and 30.85% Cu-65

This page titled 8.7: Atomic Structure and Symbolism is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.8: Chemical Formulas

## Learning Objectives

- Symbolize the composition of molecules using molecular formulas and empirical formulas
- Represent the bonding arrangement of atoms within molecules using structural formulas

A molecular formula is a representation of a molecule that uses chemical symbols to indicate the types of atoms followed by subscripts to show the number of atoms of each type in the molecule. (A subscript is used only when more than one atom of a given type is present.) Molecular formulas are also used as abbreviations for the names of compounds.

The structural formula for a compound gives the same information as its molecular formula (the types and numbers of atoms in the molecule) but also shows how the atoms are connected in the molecule. The structural formula for methane contains symbols for one C atom and four H atoms, indicating the number of atoms in the molecule (Figure 8.8.1). The lines represent bonds that hold the atoms together. (A chemical bond is an attraction between atoms or ions that holds them together in a molecule or a crystal.) We will discuss chemical bonds and see how to predict the arrangement of atoms in a molecule later. For now, simply know that the lines are an indication of how the atoms are connected in a molecule. A ball-and-stick model shows the geometric arrangement of the atoms with atomic sizes not to scale, and a space-filling model shows the relative sizes of the atoms.

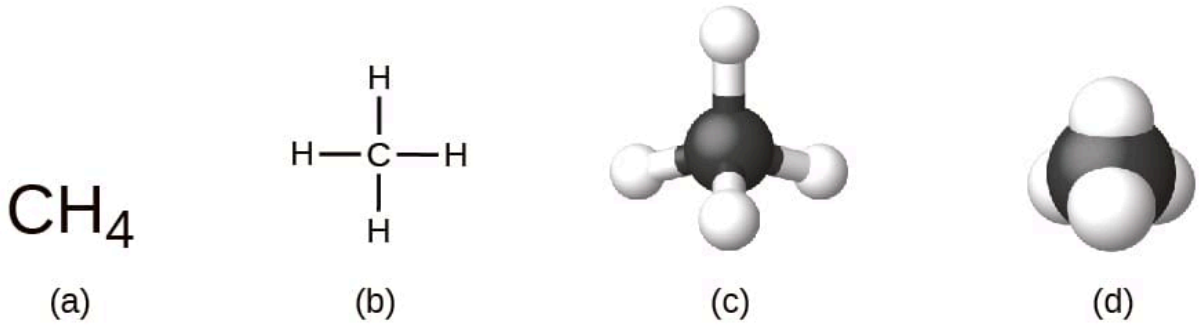



Figure 8.8.1: A methane molecule can be represented as (a) a molecular formula, (b) a structural formula, (c) a ball-and-stick model, and (d) a space-filling model. Carbon and hydrogen atoms are represented by black and white spheres, respectively. Figure A shows C H subscript 4. Figure B shows a carbon atom that is bonded to four hydrogen atoms at right angles: one above, one to the left, one to the right, and one below. Figure C shows a 3-D, ball-and-stick model of the carbon atom bonded to four hydrogen atoms. Figure D shows a space-filling model of a carbon atom with hydrogen atoms partially embedded into the surface of the carbon atom.

Although many elements consist of discrete, individual atoms, some exist as molecules made up of two or more atoms of the element chemically bonded together. For example, most samples of the elements hydrogen, oxygen, and nitrogen are composed of molecules that contain two atoms each (called diatomic molecules) and thus have the molecular formulas $H_2$, $O_2$, and $N_2$, respectively. Other elements commonly found as diatomic molecules are fluorine ($F_2$), chlorine ($Cl_2$), bromine ($Br_2$), and iodine ($I_2$). The most common form of the element sulfur is composed of molecules that consist of eight atoms of sulfur; its molecular formula is $S_8$ (Figure 8.8.2).

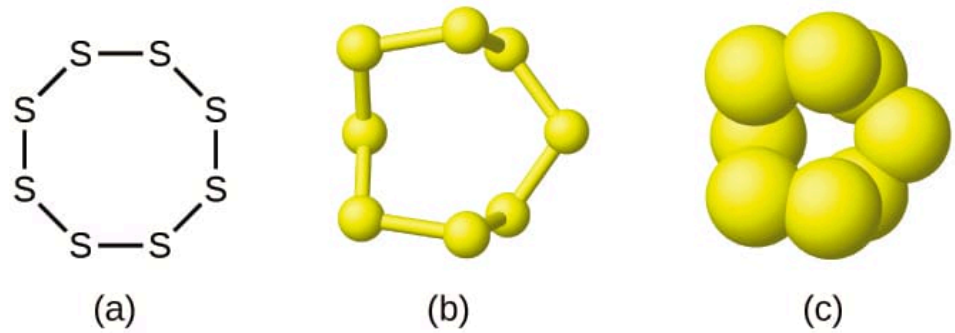



Figure 8.8.2: A molecule of sulfur is composed of eight sulfur atoms and is therefore written as $S_8$. It can be represented as (a) a structural formula, (b) a ball-and-stick model, and (c) a space-filling model. Sulfur atoms are represented by yellow spheres. Figure A shows eight sulfur atoms, symbolized with the letter S, that are bonded to each other to form an octagon. Figure B shows a 3-D, ball-and-stick model of the arrangement of the sulfur atoms. The shape is clearly not octagonal as it is represented in the structural formula. Figure C is a space-filling model that shows each sulfur atom is partially embedded into the sulfur atom it bonds with.

It is important to note that a subscript following a symbol and a number in front of a symbol do not represent the same thing; for example, $H_2$ and 2H represent distinctly different species. $H_2$ is a molecular formula; it represents a diatomic molecule of hydrogen, consisting of two atoms of the element that are chemically bonded together. The expression 2H, on the other hand, indicates two separate hydrogen atoms that are not combined as a unit. The expression $2H_2$ represents two molecules of diatomic hydrogen (Figure 8.8.3).

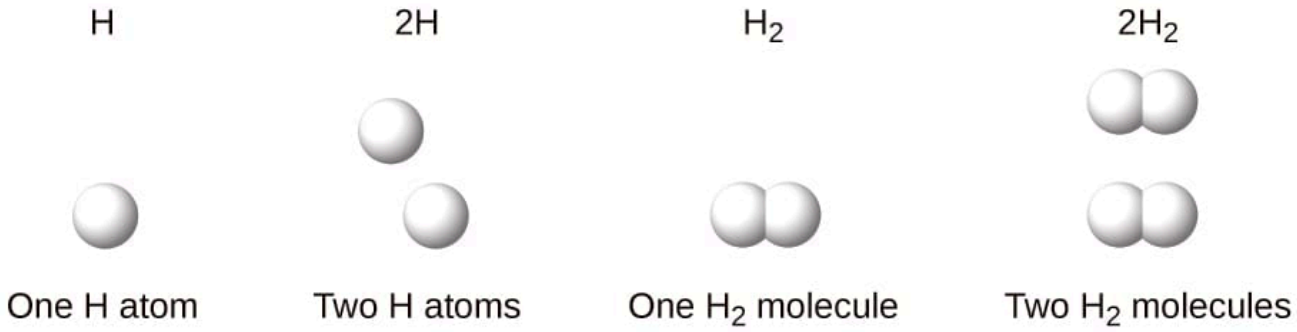



Figure 8.8.3: The symbols H, 2H, $H_2$, and $2H_2$ represent very different entities. This figure shows four diagrams. The diagram for H shows a single, white sphere and is labeled one H atom. The diagram for 2 H shows two white spheres that are not bonded together. It is labeled 2 H atoms. The diagram for H subscript 2 shows two white spheres bonded together. It is labeled one H subscript 2 molecule. The diagram for 2 H subscript 2 shows two sets of bonded, white spheres. It is labeled 2 H subscript 2 molecules.

Compounds are formed when two or more elements chemically combine, resulting in the formation of bonds. For example, hydrogen and oxygen can react to form water, and sodium and chlorine can react to form table salt. We sometimes describe the composition of these compounds with an empirical formula, which indicates the types of atoms present and *the simplest whole-number ratio of the number of atoms (or ions) in the compound*. For example, titanium dioxide (used as pigment in white paint and in the thick, white, blocking type of sunscreen) has an empirical formula of $TiO_2$. This identifies the elements titanium (Ti) and oxygen (O) as the constituents of titanium dioxide, and indicates the presence of twice as many atoms of the element oxygen as atoms of the element titanium (Figure 8.8.4).

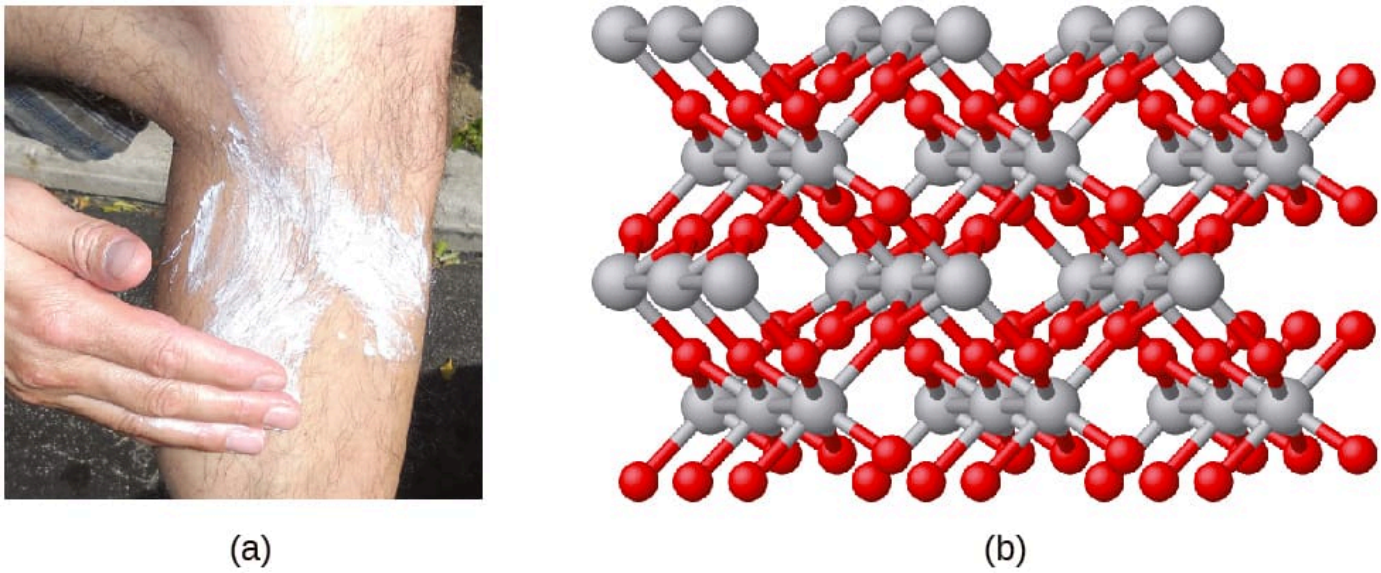



Figure 8.8.4: (a) The white compound titanium dioxide provides effective protection from the sun. (b) A crystal of titanium dioxide, $TiO_2$, contains titanium and oxygen in a ratio of 1 to 2. The titanium atoms are gray and the oxygen atoms are red. (credit a: modification of work by "osseous"/Flickr). Figure A shows a photo of a person applying suntan lotion to his or her lower leg. Figure B shows a 3-D ball-and-stick model of the molecule titanium dioxide, which involves a complicated interlocking of many titanium and oxygen atoms. The titanium atoms in the molecule are shown as silver spheres and the oxygen atoms are shown as red spheres. There are twice as many oxygen atoms as titanium atoms in the molecule.

As discussed previously, we can describe a compound with a molecular formula, in which the subscripts indicate the *actual numbers of atoms* of each element in a molecule of the compound. In many cases, the molecular formula of a substance is derived from experimental determination of both its empirical formula and its molecular mass (the sum of atomic masses for all atoms composing the molecule). For example, it can be determined experimentally that benzene contains two elements, carbon (C) and hydrogen (H), and that for every carbon atom in benzene, there is one hydrogen atom. Thus, the empirical formula is CH. An experimental determination of the molecular mass reveals that a molecule of benzene contains six carbon atoms and six hydrogen atoms, so the molecular formula for benzene is $C_6H_6$ (Figure 8.8.5).

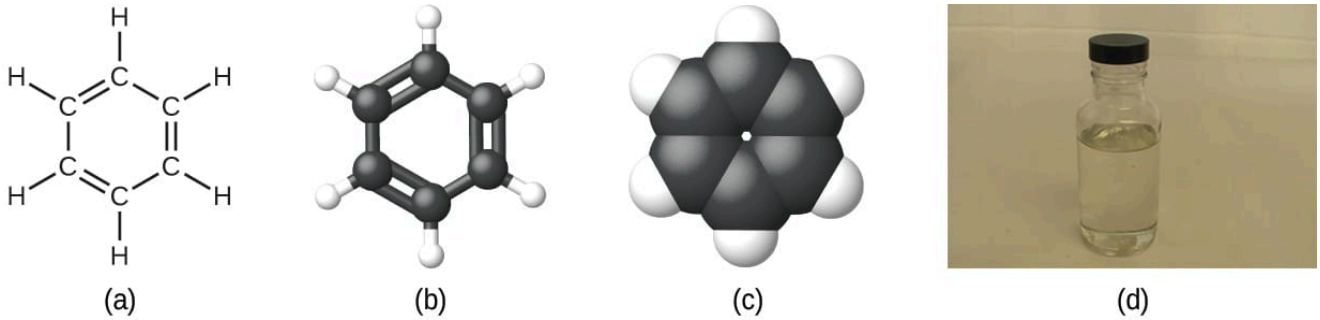



Figure 8.8.5: Benzene, $C_6H_6$, is produced during oil refining and has many industrial uses. A benzene molecule can be represented as (a) a structural formula, (b) a ball-and-stick model, and (c) a space-filling model. (d) Benzene is a clear liquid. (credit d: modification of work by Sahar Atwa). Figure A shows that benzene is composed of six carbons shaped like a hexagon. Every other bond between the carbon atoms is a double bond. Each carbon also has a single bonded hydrogen atom. Figure B shows a 3-D, ball-and-stick drawing of benzene. The six carbon atoms are black spheres while the six hydrogen atoms are smaller, white spheres. Figure C is a space-filling model of benzene which shows that most of the interior space is occupied by the carbon atoms. The hydrogen atoms are embedded in the outside surface of the carbon atoms. Figure d shows a small vial filled with benzene which appears to be clear.

If we know a compound's formula, we can easily determine the empirical formula. (This is somewhat of an academic exercise; the reverse chronology is generally followed in actual practice.) For example, the molecular formula for acetic acid, the component that gives vinegar its sharp taste, is $C_2H_4O_2$. This formula indicates that a molecule of acetic acid (Figure 8.8.6) contains two carbon atoms, four hydrogen atoms, and two oxygen atoms. The ratio of atoms is 2:4:2. Dividing by the lowest common denominator (2) gives the simplest, whole-number ratio of atoms, 1:2:1, so the empirical formula is $CH_2O$. Note that a molecular

formula is always a whole-number multiple of an empirical formula.

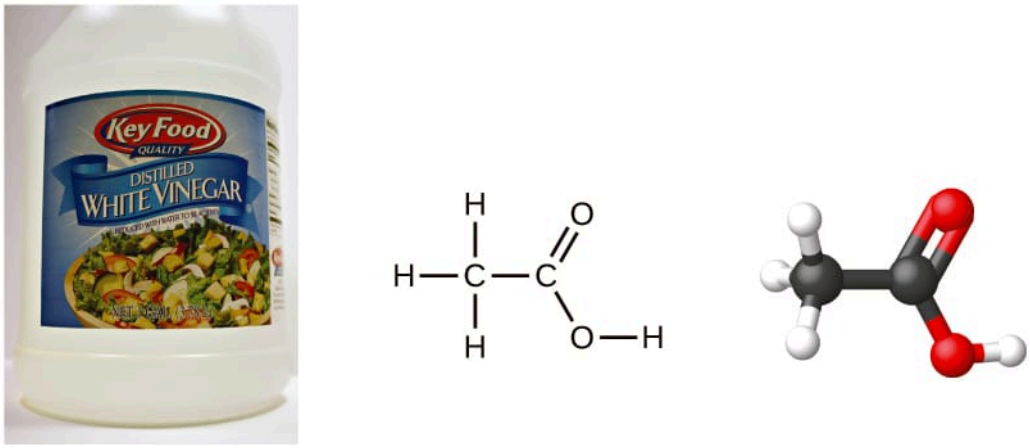



Figure 8.8.6: (a) Vinegar contains acetic acid, $C_2H_4O_2$, which has an empirical formula of $CH_2O$. It can be represented as (b) a structural formula and (c) as a ball-and-stick model. (credit a: modification of work by "HomeSpot HQ"/Flickr) Figure A shows a jug of distilled, white vinegar. Figure B shows a structural formula for acetic acid which contains two carbon atoms connected by a single bond. The left carbon atom forms single bonds with three hydrogen atoms. The right carbon atom forms a double bond with an oxygen atom. The right carbon atom also forms a single bond with an oxygen atom. This oxygen forms a single bond with a hydrogen atom. Figure C shows a 3-D ball-and-stick model of acetic acid.

## Example 8.8.1: Empirical and Molecular Formulas

Molecules of glucose (blood sugar) contain 6 carbon atoms, 12 hydrogen atoms, and 6 oxygen atoms. What are the molecular and empirical formulas of glucose?

**Solution**

The molecular formula is $C_6H_{12}O_6$ because one molecule actually contains 6 C, 12 H, and 6 O atoms. The simplest whole-number ratio of C to H to O atoms in glucose is 1:2:1, so the empirical formula is $CH_2O$.

## Exercise 8.8.1

A molecule of metaldehyde (a pesticide used for snails and slugs) contains 8 carbon atoms, 16 hydrogen atoms, and 4 oxygen atoms. What are the molecular and empirical formulas of metaldehyde?

*Answer available at the end of this section.*

It is important to be aware that it may be possible for the same atoms to be arranged in different ways: Compounds with the same molecular formula may have different atom-to-atom bonding and therefore different structures. For example, could there be another compound with the same formula as acetic acid, $C_2H_4O_2$? And if so, what would be the structure of its molecules?

If you predict that another compound with the formula $C_2H_4O_2$ could exist, then you demonstrated good chemical insight and are correct. Two C atoms, four H atoms, and two O atoms can also be arranged to form methyl formate, which is used in manufacturing, as an insecticide, and for quick-drying finishes. Methyl formate molecules have one of the oxygen atoms between the two carbon atoms, differing from the arrangement in acetic acid molecules. Acetic acid and methyl formate are examples of isomers—compounds with the same chemical formula but different molecular structures (Figure 8.8.7). Note that this small difference in the arrangement of the atoms has a major effect on their respective chemical properties. You would certainly not want to use a solution of methyl formate as a substitute for a solution of acetic acid (vinegar) when you make salad dressing.

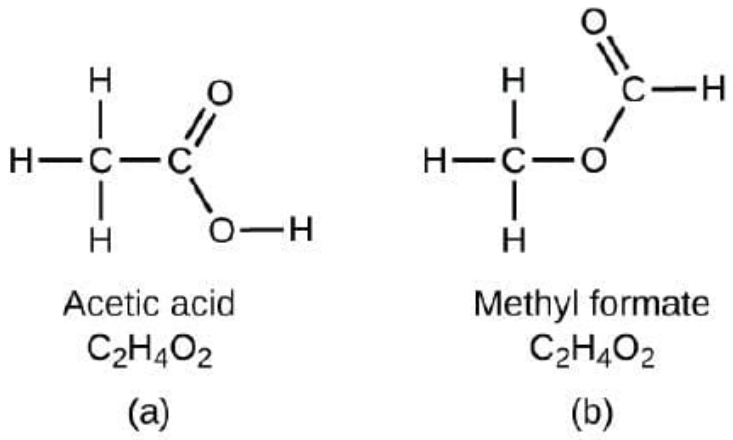



Figure 8.8.7: Molecules of (a) acetic acid and methyl formate (b) are structural isomers; they have the same formula ($C_2H_4O_2$) but different structures (and therefore different chemical properties). Figure A shows a structural diagram of acetic acid, C subscript 2 H subscript 4 O subscript 2. Acetic acid contains two carbon atoms connected by a single bond. The left carbon atom forms single bonds with three hydrogen atoms. The carbon on the right forms a double bond with an oxygen atom. The right carbon atom also forms a single bond to an oxygen atom which forms a single bond with a hydrogen atom. Figure B shows a structural diagram of methyl formate, C subscript 2 H subscript 4 O subscript 2. This molecule contains a carbon atom which forms single bonds with three hydrogen atoms, and a single bond with an oxygen atom. The oxygen atom forms a single bond with another carbon atom which forms a double bond with another oxygen atom and a single bond with a hydrogen atom.

Many types of isomers exist (Figure 8.8.8). Acetic acid and methyl formate are structural isomers, compounds in which the molecules differ in how the atoms are connected to each other. There are also various types of spatial isomers, in which the relative orientations of the atoms in space can be different. For example, the compound carvone (found in caraway seeds, spearmint, and mandarin orange peels) consists of two isomers that are mirror images of each other. S-(+)-carvone smells like caraway, and *R*-(−)-carvone smells like spearmint.

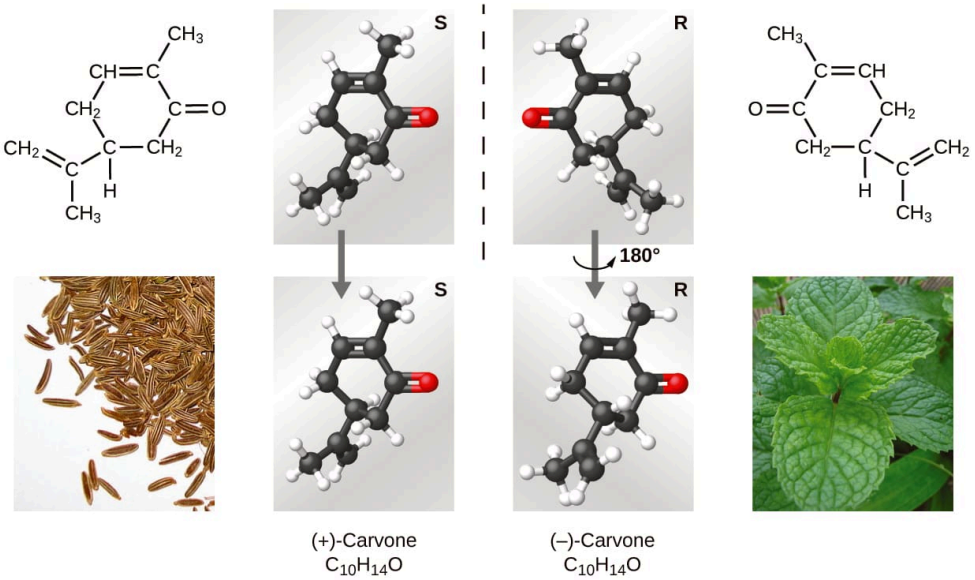



Figure 8.8.8: Molecules of carvone are spatial isomers; they only differ in the relative orientations of the atoms in space. (credit bottom left: modification of work by “Miansari66”/Wikimedia Commons; credit bottom right: modification of work by Forest & Kim Starr)

## Summary

A molecular formula uses chemical symbols and subscripts to indicate the exact numbers of different atoms in a molecule or compound. An empirical formula gives the simplest, whole-number ratio of atoms in a compound. A structural formula indicates the bonding arrangement of the atoms in the molecule. Ball-and-stick and space-filling models show the geometric arrangement of atoms in a molecule. Isomers are compounds with the same molecular formula but different arrangements of atoms.

# Glossary

**empirical formula**
: formula showing the composition of a compound given as the simplest whole-number ratio of atoms

**isomers**
: compounds with the same chemical formula but different structures

**molecular formula**
: formula indicating the composition of a molecule of a compound and giving the actual number of atoms of each element in a molecule of the compound.

**spatial isomers**
: compounds in which the relative orientations of the atoms in space differ

**structural isomer**
: one of two substances that have the same molecular formula but different physical and chemical properties because their atoms are bonded differently

**structural formula**
: shows the atoms in a molecule and how they are connected

## Answers

**Exercise** 8.8.1

Molecular formula, $C_8H_{16}O_4$; empirical formula, $C_2H_4O$

This page titled 8.8: Chemical Formulas is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

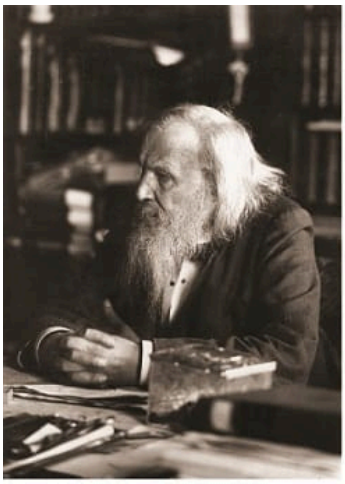
# 8.9: The Periodic Table

## Learning Objectives

- State the periodic law and explain the organization of elements in the periodic table
- Predict the general properties of elements based on their location within the periodic table
- Identify metals, nonmetals, and metalloids by their properties and/or location on the periodic table

As early chemists worked to purify ores and discovered more elements, they realized that various elements could be grouped together by their similar chemical behaviors. One such grouping includes lithium (Li), sodium (Na), and potassium (K): These elements all are shiny, conduct heat and electricity well, and have similar chemical properties. A second grouping includes calcium (Ca), strontium (Sr), and barium (Ba), which also are shiny, good conductors of heat and electricity, and have chemical properties in common. However, the specific properties of these two groupings are notably different from each other. For example: Li, Na, and K are much more reactive than are Ca, Sr, and Ba; Li, Na, and K form compounds with oxygen in a ratio of two of their atoms to one oxygen atom, whereas Ca, Sr, and Ba form compounds with one of their atoms to one oxygen atom. Fluorine (F), chlorine (Cl), bromine (Br), and iodine (I) also exhibit similar properties to each other, but these properties are drastically different from those of any of the elements above.

Dimitri Mendeleev in Russia (1869) and Lothar Meyer in Germany (1870) independently recognized that there was a periodic relationship among the properties of the elements known at that time. Both published tables with the elements arranged according to increasing atomic mass. But Mendeleev went one step further than Meyer: He used his table to predict the existence of elements that would have the properties similar to aluminum and silicon, but were yet unknown. The discoveries of gallium (1875) and germanium (1886) provided great support for Mendeleev's work. Although Mendeleev and Meyer had a long dispute over priority, Mendeleev's contributions to the development of the periodic table are now more widely recognized (Figure 8.9.1).



(a)

| Reihen | Gruppo I. — $R^1O$ | Gruppo II. — RO | Gruppo III. — $R^1O^3$ | Gruppo IV. $RH^4$ $RO^2$ | Gruppo V. $RH^3$ $R^1O^5$ | Gruppo VI. $RH^2$ $RO^3$ | Gruppo VII. RH $R^1O^7$ | Gruppo VIII. — $RO^4$ |
|---|---|---|---|---|---|---|---|---|
| 1 | H=1 | | | | | | | |
| 2 | Li=7 | Be=9,4 | B=11 | C=12 | N=14 | O=16 | F=19 | |
| 3 | Na=23 | Mg=24 | Al=27,3 | Si=28 | P=31 | S=32 | Cl=35,5 | |
| 4 | K=39 | Ca=40 | —=44 | Ti=48 | V=51 | Cr=52 | Mn=55 | Fe=56, Co=59, Ni=59, Cu=63. |
| 5 | (Cu=63) | Zn=65 | —=68 | —=72 | As=75 | Se=78 | Br=80 | |
| 6 | Rb=85 | Sr=87 | ?Yt=88 | Zr=90 | Nb=94 | Mo=96 | —=100 | Ru=104, Rh=104, Pd=106, Ag=108. |
| 7 | (Ag=108) | Cd=112 | In=113 | Sn=118 | Sb=122 | Te=125 | J=127 | |
| 8 | Cs=133 | Ba=137 | ?Di=138 | ?Ce=140 | — | — | — | — — — — |
| 9 | (—) | — | — | — | — | — | — | |
| 10 | — | — | ?Er=178 | ?La=180 | Ta=182 | W=184 | — | Os=195, Ir=197, Pt=198, Au=199. |
| 11 | (Au=199) | Hg=200 | Tl=204 | Pb=207 | Bi=208 | — | — | |
| 12 | — | — | — | Th=231 | — | U=240 | — | — — — — |

(b)

Figure 8.9.1: (a) Dimitri Mendeleev is widely credited with creating (b) the first periodic table of the elements. (credit a: modification of work by Serge Lachinov; credit b: modification of work by "Den fjättrade ankan"/Wikimedia Commons) Figure A shows a photograph of Dimitri Mendeleev. Figure B shows the first periodic table developed by Mendeleev, which had eight groups and twelve periods. In the first group (—, R superscript plus sign 0) is the following information: H = 1, L i = 7, N a = 23, K = 39, (C u = 63), R b = 85, (A g = 108), C a = 183, (—),—, (A u = 199) —. Note that each of these entries corresponds to one of the twelve periods respectively. The second group (—, R 0) contains the following information: (not entry for period 1) B o = 9, 4, M g = 24, C a = 40, Z n = 65, S r = 87, C d = 112, B a = 187, —, —, H g = 200, —. Note the ach of these entries corresponds to one of the twelve periods respectively. Group three (—, R superscript one 0 superscript nine) contains the information: (no entry for period 1), B = 11, A l = 27, 8. — = 44, — = 68, ? Y t = 88, I n = 113, ? D I = 138, —, ? E r = 178, T l = 204, —. Note that each of these entries corresponds to one of the twelve periods respectively. Group four (RH superscript four, R0 superscript eight) contains the following information: (no entry for period 1), C = 12, B i = 28, T i = 48, — = 72, Z r = 90, S n = 118, ? C o = 140, ? L a = 180, P b = 207, T h = 231. Note that each of these entries corresponds to one of the twelve periods respectively. Group five (R H superscript

two, R superscript two 0 superscript five) contains the following information: (no entry for period 1), N = 14, P = 31, V = 51, A s = 75, N b = 94, S b = 122, —, —, T a = 182, B i = 208, —. Note that each of these entries corresponds to one of the twelve periods respectively. Group six (R H superscript two, R 0 superscript three) contains the following information: (no entry for period 1), O = 16, S = 32, C r = 52, S o = 78, M o = 96, T o = 125, —, —, W = 184, —, U = 240. Note that each of these entries corresponds to one of the twelve periods respectively. Group seven (R H , R superscript plus sing, 0 superscript 7) contains the following information: (no entry for period 1), F = 19, C l = 35, 5, M n = 55, B r = 80, — = 100, J = 127, —, —, —, —, —. Note that each of these entries corresponds to one of the twelve periods respectively. Group 8 (—, R 0 superscript four) contains the following information: (no entry for periods 1, 2, 3), in period 4: F o = 56, C o = 59, N i = 59, C u = 63, no entry for period five, in period 6: R u = 104, R h = 104, P d = 106, A g = 108, no entries for periods 7, 8 , or 9, in period 10: O s = 195, I r = 197, P t = 198, A u = 199, no entries for periods 11 or 12.

By the twentieth century, it became apparent that the periodic relationship involved atomic numbers rather than atomic masses. The modern statement of this relationship, the periodic law, is as follows: *the properties of the elements are periodic functions of their atomic numbers*. A modern periodic table arranges the elements in increasing order of their atomic numbers and groups atoms with similar properties in the same vertical column (Figure 8.9.2). Each box represents an element and contains its atomic number, symbol, average atomic mass, and (sometimes) name. The elements are arranged in seven horizontal rows, called periods or series, and 18 vertical columns, called groups. Groups are labeled at the top of each column. In the United States, the labels traditionally were numerals with capital letters. However, IUPAC recommends that the numbers 1 through 18 be used, and these labels are more common. For the table to fit on a single page, parts of two of the rows, a total of 14 columns, are usually written below the main body of the table.

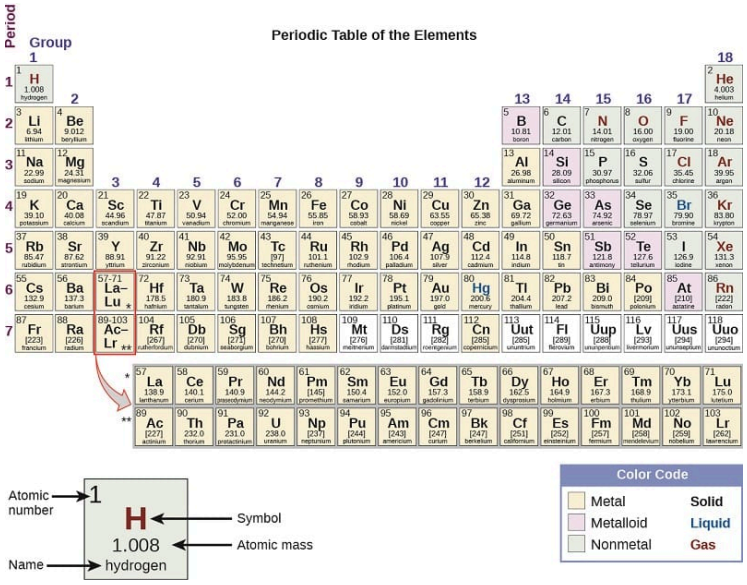



Figure 8.9.2: Elements in the periodic table are organized according to their properties. The Periodic Table of Elements is shown. The 18 columns are labeled “Group” and the 7 rows are labeled “Period.” Below the table to the right is a box labeled “Color Code” with different colors for metals, metalloids, and nonmetals, as well as solids, liquids, and gases. To the left of this box is an enlarged picture of the upper-left most box on the table. The number 1 is in its upper-left hand corner and is labeled “Atomic number.” The letter “H” is in the middle in red indicating that it is a gas. It is labeled “Symbol.” Below that is the number 1.008 which is labeled “Atomic Mass.” Below that is the word hydrogen which is labeled “name.” The color of the box indicates that it is a nonmetal. Each element will be described in this order: atomic number; name; symbol; whether it is a metal, metalloid, or nonmetal; whether it is a solid, liquid, or gas; and atomic mass. Beginning at the top left of the table, or period 1, group 1, is a box containing “1; hydrogen; H; nonmetal; gas; and 1.008.” There is only one other element box in period 1, group 18, which contains “2; helium; H e; nonmetal; gas; and 4.003.” Period 2, group 1 contains “3; lithium; L i; metal; solid; and 6.94” Group 2 contains “4; beryllium; B e; metal; solid; and 9.012.” Groups 3 through 12 are skipped and group 13 contains “5; boron; B; metalloid; solid; 10.81.” Group 14 contains “6; carbon; C; nonmetal; solid; and 12.01.” Group 15 contains “7; nitrogen; N; nonmetal; gas; and 14.01.” Group 16 contains “8; oxygen; O; nonmetal; gas; and 16.00.” Group 17 contains “9; fluorine; F; nonmetal; gas; and 19.00.” Group 18 contains “10; neon; N e; nonmetal; gas; and 20.18.” Period 3, group 1 contains “11; sodium; N a; metal; solid; and 22.99.” Group 2 contains “12; magnesium; M g; metal; solid; and 24.31.” Groups 3 through 12 are skipped again in period 3 and group 13 contains “13; aluminum; A l; metal; solid; and 26.98.” Group 14 contains “14; silicon; S i; metalloid; solid; and 28.09.” Group 15 contains “15; phosphorous; P; nonmetal; solid; and 30.97.” Group 16 contains “16; sulfur; S; nonmetal; solid; and 32.06.” Group 17 contains “17; chlorine; C l; nonmetal; gas; and 35.45.” Group 18 contains “18; argon; A r; nonmetal; gas; and 39.95.” Period 4, group 1 contains “19; potassium; K; metal; solid; and 39.10.” Group

2 contains “20; calcium; C a; metal; solid; and 40.08.” Group 3 contains “21; scandium; S c; metal; solid; and 44.96.” Group 4 contains “22; titanium; T i; metal; solid; and 47.87.” Group 5 contains “23; vanadium; V; metal; solid; and 50.94.” Group 6 contains “24; chromium; C r; metal; solid; and 52.00.” Group 7 contains “25; manganese; M n; metal; solid; and 54.94.” Group 8 contains “26; iron; F e; metal; solid; and 55.85.” Group 9 contains “27; cobalt; C o; metal; solid; and 58.93.” Group 10 contains “28; nickel; N i; metal; solid; and 58.69.” Group 11 contains “29; copper; C u; metal; solid; and 63.55.” Group 12 contains “30; zinc; Z n; metal; solid; and 65.38.” Group 13 contains “31; gallium; G a; metal; solid; and 69.72.” Group 14 contains “32; germanium; G e; metalloid; solid; and 72.63.” Group 15 contains “33; arsenic; A s; metalloid; solid; and 74.92.” Group 16 contains “34; selenium; S e; nonmetal; solid; and 78.97.” Group 17 contains “35; bromine; B r; nonmetal; liquid; and 79.90.” Group 18 contains “36; krypton; K r; nonmetal; gas; and 83.80.” Period 5, group 1 contains “37; rubidium; R b; metal; solid; and 85.47.” Group 2 contains “38; strontium; S r; metal; solid; and 87.62.” Group 3 contains “39; yttrium; Y; metal; solid; and 88.91.” Group 4 contains “40; zirconium; Z r; metal; solid; and 91.22.” Group 5 contains “41; niobium; N b; metal; solid; and 92.91.” Group 6 contains “42; molybdenum; M o; metal; solid; and 95.95.” Group 7 contains “43; technetium; T c; metal; solid; and 97.” Group 8 contains “44; ruthenium; R u; metal; solid; and 101.1.” Group 9 contains “45; rhodium; R h; metal; solid; and 102.9.” Group 10 contains “46; palladium; P d; metal; solid; and 106.4.” Group 11 contains “47; silver; A g; metal; solid; and 107.9.” Group 12 contains “48; cadmium; C d; metal; solid; and 112.4.” Group 13 contains “49; indium; I n; metal; solid; and 114.8.” Group 14 contains “50; tin; S n; metal; solid; and 118.7.” Group 15 contains “51; antimony; S b; metalloid; solid; and 121.8.” Group 16 contains “52; tellurium; T e; metalloid; solid; and 127.6.” Group 17 contains “53; iodine; I; nonmetal; solid; and 126.9.” Group 18 contains “54; xenon; X e; nonmetal; gas; and 131.3.” Period 6, group 1 contains “55; cesium; C s; metal; solid; and 132.9.” Group 2 contains “56; barium; B a; metal; solid; and 137.3.” Group 3 breaks the pattern. The box has a large arrow pointing to a row of elements below the table with atomic numbers ranging from 57-71. In sequential order by atomic number, the first box in this row contains “57; lanthanum; L a; metal; solid; and 138.9.” To its right, the next is “58; cerium; C e; metal; solid; and 140.1.” Next is “59; praseodymium; P r; metal; solid; and 140.9.” Next is “60; neodymium; N d; metal; solid; and 144.2.” Next is “61; promethium; P m; metal; solid; and 145.” Next is “62; samarium; S m; metal; solid; and 150.4.” Next is “63; europium; E u; metal; solid; and 152.0.” Next is “64; gadolinium; G d; metal; solid; and 157.3.” Next is “65; terbium; T b; metal; solid; and 158.9.” Next is “66; dysprosium; D y; metal; solid; and 162.5.” Next is “67; holmium; H o; metal; solid; and 164.9.” Next is “68; erbium; E r; metal; solid; and 167.3.” Next is “69; thulium; T m; metal; solid; and 168.9.” Next is “70; ytterbium; Y b; metal; solid; and 173.1.” The last in this special row is “71; lutetium; L u; metal; solid; and 175.0.” Continuing in period 6, group 4 contains “72; hafnium; H f; metal; solid; and 178.5.” Group 5 contains “73; tantalum; T a; metal; solid; and 180.9.” Group 6 contains “74; tungsten; W; metal; solid; and 183.8.” Group 7 contains “75; rhenium; R e; metal; solid; and 186.2.” Group 8 contains “76; osmium; O s; metal; solid; and 190.2.” Group 9 contains “77; iridium; I r; metal; solid; and 192.2.” Group 10 contains “78; platinum; P t; metal; solid; and 195.1.” Group 11 contains “79; gold; A u; metal; solid; and 197.0.” Group 12 contains “80; mercury; H g; metal; liquid; and 200.6.” Group 13 contains “81; thallium; T l; metal; solid; and 204.4.” Group 14 contains “82; lead; P b; metal; solid; and 207.2.” Group 15 contains “83; bismuth; B i; metal; solid; and 209.0.” Group 16 contains “84; polonium; P o; metal; solid; and 209.” Group 17 contains “85; astatine; A t; metalloid; solid; and 210.” Group 18 contains “86; radon; R n; nonmetal; gas; and 222.” Period 7, group 1 contains “87; francium; F r; metal; solid; and 223.” Group 2 contains “88; radium; R a; metal; solid; and 226.” Group 3 breaks the pattern much like what occurs in period 6. A large arrow points from the box in period 7, group 3 to a special row containing the elements with atomic numbers ranging from 89-103, just below the row which contains atomic numbers 57-71. In sequential order by atomic number, the first box in this row contains “89; actinium; A c; metal; solid; and 227.” To its right, the next is “90; thorium; T h; metal; solid; and 232.0.” Next is “91; protactinium; P a; metal; solid; and 231.0.” Next is “92; uranium; U; metal; solid; and 238.0.” Next is “93; neptunium; N p; metal; solid; and N p.” Next is “94; plutonium; P u; metal; solid; and 244.” Next is “95; americium; A m; metal; solid; and 243.” Next is “96; curium; C m; metal; solid; and 247.” Next is “97; berkelium; B k; metal; solid; and 247.” Next is “98; californium; C f; metal; solid; and 251.” Next is “99; einsteinium; E s; metal; solid; and 252.” Next is “100; fermium; F m; metal; solid; and 257.” Next is “101; mendelevium; M d; metal; solid; and 258.” Next is “102; nobelium; N o; metal; solid; and 259.” The last in this special row is “103; lawrencium; L r; metal; solid; and 262.” Continuing in period 7, group 4 contains “104; rutherfordium; R f; metal; solid; and 267.” Group 5 contains “105; dubnium; D b; metal; solid; and 270.” Group 6 contains “106; seaborgium; S g; metal; solid; and 271.” Group 7 contains “107; bohrium; B h; metal; solid; and 270.” Group 8 contains “108; hassium; H s; metal; solid; and 277.” Group 9 contains “109; meitnerium; M t; not indicated; solid; and 276.” Group 10 contains “110; darmstadtium; D s; not indicated; solid; and 281.” Group 11 contains “111; roentgenium; R g; not indicated; solid; and 282.” Group 12 contains “112; copernicium; C n; metal; liquid; and 285.” Group 13 contains “113; ununtrium; U u t; not indicated; solid; and 285.” Group 14 contains “114; flerovium; F l; not indicated; solid; and 289.” Group 15 contains “115; ununpentium; U u p; not indicated; solid; and 288.” Group 16 contains “116; livermorium; L v; not indicated; solid; and 293.” Group 17 contains “117; ununseptium; U u s; not indicated; solid; and 294.” Group 18 contains “118; ununoctium; U u o; not indicated; solid; and 294.”

Many elements differ dramatically in their chemical and physical properties, but some elements are similar in their behaviors. For example, many elements appear shiny, are malleable (able to be deformed without breaking) and ductile (can be drawn into wires), and conduct heat and electricity well. Other elements are not shiny, malleable, or ductile, and are poor conductors of heat and electricity. We can sort the elements into large classes with common properties: metals (elements that are shiny, malleable, good conductors of heat and electricity—shaded yellow); nonmetals (elements that appear dull, poor conductors of heat

and electricity—shaded green); and metalloids (elements that conduct heat and electricity moderately well, and possess some properties of metals and some properties of nonmetals—shaded purple).

The elements can also be classified into the main-group elements (or representative elements) in the columns labeled 1, 2, and 13–18; the transition metals in the columns labeled 3–12; and inner transition metals in the two rows at the bottom of the table (the top-row elements are called lanthanides and the bottom-row elements are actinides; Figure 8.9.3). The elements can be subdivided further by more specific properties, such as the composition of the compounds they form. For example, the elements in group 1 (the first column) form compounds that consist of one atom of the element and one atom of hydrogen. These elements (except hydrogen) are known as alkali metals, and they all have similar chemical properties. The elements in group 2 (the second column) form compounds consisting of one atom of the element and two atoms of hydrogen: These are called alkaline earth metals, with similar properties among members of that group. Other groups with specific names are the pnictogens (group 15), chalcogens (group 16), halogens (group 17), and the noble gases (group 18, also known as inert gases). The groups can also be referred to by the first element of the group: For example, the chalcogens can be called the oxygen group or oxygen family. Hydrogen is a unique, nonmetallic element with properties similar to both group 1 and group 17 elements. For that reason, hydrogen may be shown at the top of both groups, or by itself.

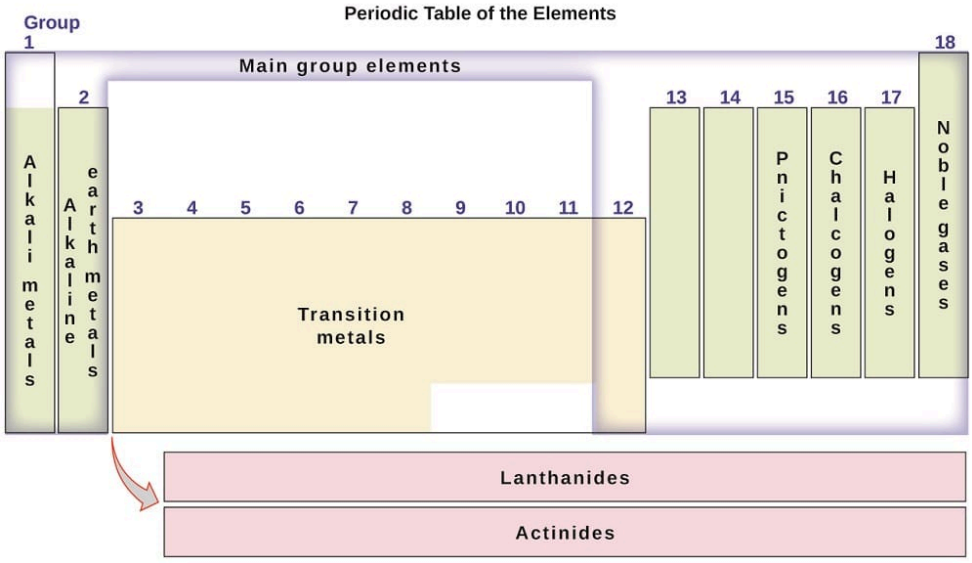



Figure 8.9.3: The periodic table organizes elements with similar properties into groups. This diagram combines the groups and periods of the periodic table based on their similar properties. Group 1 contains the alkali metals, group 2 contains the earth alkaline metals, group 15 contains the pnictogens, group 16 contains the chalcogens, group 17 contains the halogens and group 18 contains the noble gases. The main group elements consist of groups 1, 2, and 12 through 18. Therefore, most of the transition metals, which are contained in groups 3 through 11, are not main group elements. The lanthanides and actinides are called out at the bottom of the periodic table.

## Example 8.9.1: Naming Groups of Elements

Atoms of each of the following elements are essential for life. Give the group name for the following elements:

a. chlorine
b. calcium
c. sodium
d. sulfur

**Solution**

The family names are as follows:

a. halogen
b. alkaline earth metal

c. alkali metal
d. chalcogen

**Exercise 8.9.1**

Give the group name for each of the following elements:

a. krypton
b. selenium
c. barium
d. lithium

*Answer a available at the end of this section.*

In studying the periodic table, you might have noticed something about the atomic masses of some of the elements. Element 43 (technetium), element 61 (promethium), and most of the elements with atomic number 84 (polonium) and higher have their atomic mass given in square brackets. This is done for elements that consist entirely of unstable, radioactive isotopes (you will learn more about radioactivity in the nuclear chemistry chapter). An average atomic weight cannot be determined for these elements because their radioisotopes may vary significantly in relative abundance, depending on the source, or may not even exist in nature. The number in square brackets is the atomic mass number (and approximate atomic mass) of the most stable isotope of that element.

## Summary

The discovery of the periodic recurrence of similar properties among the elements led to the formulation of the periodic table, in which the elements are arranged in order of increasing atomic number in rows known as periods and columns known as groups. Elements in the same group of the periodic table have similar chemical properties. Elements can be classified as metals, metalloids, and nonmetals, or as a main-group elements, transition metals, and inner transition metals. Groups are numbered 1–18 from left to right. The elements in group 1 are known as the alkali metals; those in group 2 are the alkaline earth metals; those in 15 are the pnictogens; those in 16 are the chalcogens; those in 17 are the halogens; and those in 18 are the noble gases.

## Glossary

**actinide**
: inner transition metal in the bottom of the bottom two rows of the periodic table

**alkali metal**
: element in group 1

**alkaline earth metal**
: element in group 2

**chalcogen**
: element in group 16

**group**
: vertical column of the periodic table

**halogen**
: element in group 17

**inert gas**
: (also, noble gas) element in group 18

**inner transition metal**
: (also, lanthanide or actinide) element in the bottom two rows; if in the first row, also called lanthanide, or if in the second row, also called actinide

**lanthanide**
inner transition metal in the top of the bottom two rows of the periodic table

**main-group element**
(also, representative element) element in columns 1, 2, and 12–18

**metal**
element that is shiny, malleable, good conductor of heat and electricity

**metalloid**
element that conducts heat and electricity moderately well, and possesses some properties of metals and some properties of nonmetals

**noble gas**
(also, inert gas) element in group 18

**nonmetal**
element that appears dull, poor conductor of heat and electricity

**period**
(also, series) horizontal row of the periodic table

**periodic law**
properties of the elements are periodic function of their atomic numbers.

**periodic table**
table of the elements that places elements with similar chemical properties close together

**pnictogen**
element in group 15

**representative element**
(also, main-group element) element in columns 1, 2, and 12–18

**transition metal**
element in columns 3–11

**series**
(also, period) horizontal row of the period table

## Answers

**Exercise** 8.9.1

noble gas
chalcogen
alkaline earth metal
alkali metal

This page titled 8.9: The Periodic Table is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.10: ORGANIZING THE ELEMENTS

8.10.1: Mendeleev's Periodic Table
8.10.2: Periodic Law
8.10.3: Modern Periodic Table- Periods and Groups

This page titled 8.10: Organizing the Elements is shared under a CK-12 license and was authored, remixed, and/or curated by LibreTexts.

# 8.10.1: Mendeleev's Periodic Table

**Learning Objectives**

- Know the patterns that emerged in the discovery of the periodic table.




Figure 8.10.1.1 (Credit: Laura Guerin; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

## When you study for a test, how do you approach the task?

One useful way to study for a test is to use flashcards. To make them, write down vocabulary words, foreign language terms, math formulas, chemistry reactions—anything you want to learn. Then, sort these cards into categories and topics that go together. This organization of information helps you see patterns in the material so that you can tie different ideas together and make better sense of them.

The periodic table was first built using a set of cards. With this strategy, Dmitri Mendeleev, the creator of the periodic table, could organize and rearrange material until patterns emerged. And he did this while trying to find an easy way to help his students learn the material.

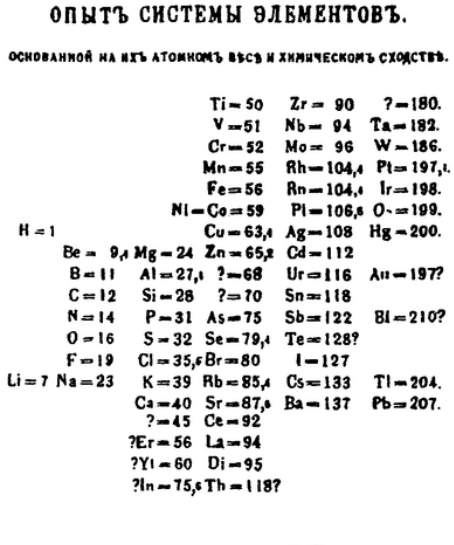
## Mendeleev's Periodic Table

In 1869, Russian chemist and teacher Dmitri Mendeleev (1836-1907) published a periodic table of the elements. The following year, German chemist Lothar Meyer independently published a very similar table. Mendeleev is generally given more credit than Meyer because his table was published first, and because of several key insights that he made regarding the table.



Figure 8.10.1.2: Dmitri Mendeleev. (Credit: Unknown; Source: wikimedia; License: Public Domain)

Mendeleev was writing a chemistry textbook for his students and wanted to organize all of the known elements at that time according to their chemical properties. He famously organized the information for each element onto separate note cards that were then easy to rearrange as needed. He discovered that when he placed them in order of increasing atomic mass, certain similarities in chemical behavior repeated at regular intervals. This type of a repeating pattern is called "periodic". A pendulum that swings back and forth in a given time interval is periodic, as is the movement of the moon around the Earth.



ОПЫТЪ СИСТЕМЫ ЭЛЕМЕНТОВЪ.

ОСНОВАННОЙ НА ИХЪ АТОМНОМЪ ВѢСѢ И ХИМИЧЕСКОМЪ СХОДСТВѢ.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | Ti = 50 | Zr = 90 | ? = 180. | |
| | | | V = 51 | Nb = 94 | Ta = 182. | |
| | | | Cr = 52 | Mo = 96 | W = 186. | |
| | | | Mn = 55 | Rh = 104,4 | Pt = 197,1. | |
| | | | Fe = 56 | Ru = 104,4 | Ir = 198. | |
| | | | Ni = Co = 59 | Pl = 106,6 | O = 199. | |
| H = 1 | | | Cu = 63,4 | Ag = 108 | Hg = 200. | |
| | Be = 9,4 | Mg = 24 | Zn = 65,2 | Cd = 112 | | |
| | B = 11 | Al = 27,4 | ? = 68 | Ur = 116 | Au = 197? | |
| | C = 12 | Si = 28 | ? = 70 | Sn = 118 | | |
| | N = 14 | P = 31 | As = 75 | Sb = 122 | Bi = 210? | |
| | O = 16 | S = 32 | Se = 79,4 | Te = 128? | | |
| | F = 19 | Cl = 35,6 | Br = 80 | I = 127 | | |
| Li = 7 | Na = 23 | K = 39 | Rb = 85,4 | Cs = 133 | Tl = 204. | |
| | | Ca = 40 | Sr = 87,6 | Ba = 137 | Pb = 207. | |
| | | ? = 45 | Ce = 92 | | | |
| | | ?Er = 56 | La = 94 | | | |
| | | ?Yt = 60 | Di = 95 | | | |
| | | ?In = 75,6 | Th = 118? | | | |

Д. Менделѣевъ

Figure 8.10.1.3: Mendeleev's 1869 periodic table. (Credit: Dmitri Mendeleev; Source: wikimedia; License: Public Domain)

In the figure above, atomic mass increases from top to bottom of vertical columns, with successive columns going left to right. As a result, elements that are in the same horizontal row are groups of elements that were known to exhibit similar chemical properties. One of Mendeleev's insights is illustrated by the elements tellurium $(\mathrm{Te})$ and iodine $(\mathrm{I})$. Notice that tellurium is listed before iodine even though its atomic mass is higher. Mendeleev reversed the order because he knew that the properties of iodine were much more similar to those of fluorine $(\mathrm{F})$, chlorine $(\mathrm{Cl})$, and bromine $(\mathrm{Br})$ than they were to oxygen $(\mathrm{O})$, sulfur $(\mathrm{S})$, and selenium $(\mathrm{Se})$. He simply assumed that there was an error in the determination of one or both of the atomic masses. This turned out not to be the case, but Mendeleev was indeed correct to group these two elements as he did.

Notice that there are several places in the table that have no chemical symbol, but are instead labeled with a question mark. Between zinc $(Zn)$ and arsenic $(\mathrm{As})$ are two such missing elements. Mendeleev believed that elements with atomic masses of 68 and 70 would eventually be discovered and that they would fit chemically into each of those spaces. Listed in the table below are other properties that Mendeleev predicted for the first of these two missing elements, which he called "eka-aluminum", compared with the element gallium.

cK-12 LICENSED UNDER CK-12 License © CK-12 Foundation • Visit us at ck12.org

*Mendeleev's predicted properties for Eka-aluminum and gallium.*

| | Eka-Aluminum (Ea) | Gallium (Ga) |
|---|---|---|
| Atomic Mass | $68\ \text{amu}$ | $69.9\ \text{amu}$ |
| Melting Point | Low | $30.15^{\circ}\text{C}$ |
| Density | $5.9\ \text{g/cm}^3$ | $5.94\ \text{g/cm}^3$ |
| Formula of Oxide | $\text{Ea}_2\text{O}_3$ | $\text{Ga}_2\text{O}_3$ |

The element gallium was discovered four years after the publication of Mendeleev's table, and its properties matched up remarkably well with eka-aluminum, fitting into the table exactly where he had predicted. This was also the case with the element that followed gallium, which was eventually named germanium.

Mendeleev's periodic table gained wide acceptance with the scientific community and earned him credit as the discoverer of the periodic law. Element number 101, synthesized in 1955, is named **mendelevium** after the founder of the periodic table. It was, however, several years after Mendeleev died before the several discrepancies with the atomic masses could be explained, and before the reasons behind the repetition of chemical properties could be fully explained.

The following video contains some of the story of Mendeleev's background and career.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- Mendeleev published his periodic table in 1869.
- His organization of elements was based on atomic mass.
- Mendeleev's periodic table made it possible to predict properties of elements that had not yet been discovered.

This page titled 8.10.1: Mendeleev's Periodic Table is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 10.32: Nuclear Reactions (Conceptual Questions Q10.4.2) by OpenStax is licensed CC BY 4.0.

# 8.10.2: Periodic Law

**Learning Objectives**

- Apply the periodic law when referencing the periodic table of elements.



Figure 8.10.2.1: Copy and Paste Caption here. (Credit: Christopher Fynn; Source: Flickr, Mysore - Fruit vendor(opens in new window) [www.flickr.com]; License: CC-BY 2.0)

## How are these items related to one another?

We have all enjoyed looking around a market for delicious foods to eat later at home. When you get to the market you know you need to get fruits, vegetables and grains to ensure you eat a balanced diet. In the market, these items are all grouped together to make it easier to find the type of fruit, vegetable or grain you are looking for; it would be inconvenient if the different fruits were all scattered in different places. The periodic table is organized in a similar way, ensuring similar elements are found in the same group or period.

## The Periodic Law

When Mendeleev put his periodic table together, nobody knew about the existence of the atomic nucleus. Later experiments showed that the periodic relationship shown in the periodic table was not based on the atomic mass, but on the atomic number. (The atomic number is a property of the atomic nucleus which will be discussed in greater detail in the next chapter.) As it turns out, the atomic mass increases as the atomic number increases. For most of the periodic table the increases are on a similar scale, but there were some exceptions which were the cause of the discrepancies in Mendeleev's original periodic table.

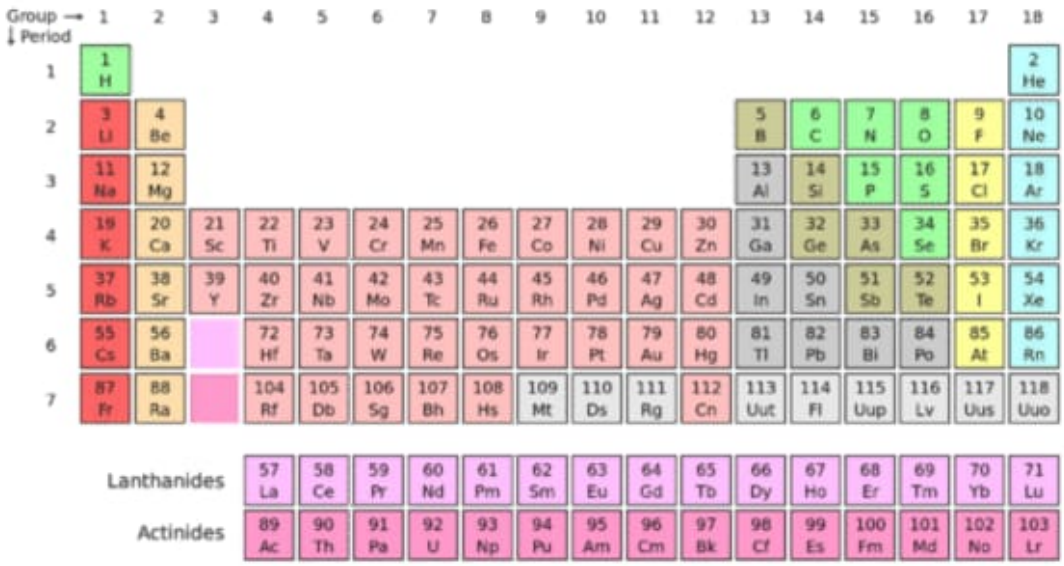



Figure 8.10.2.2 (Credit: User:Cepheus/Wikimedia Commons; Source: Commons Wikimedia, Periodic Table(opens in new window) [commons.wikimedia.org]; License: Public Domain)

When ordered by atomic number, the discrepancies within Mendeleev's table disappeared. Tellurium has an atomic number of 52, while iodine has an atomic number of 53. So even though tellurium does indeed have a greater atomic mass than iodine, it is properly placed before iodine in the periodic table. Mendeleev and Moseley are credited with being most responsible for the modern **periodic law**: When elements are arranged in order of increasing atomic number, there is a periodic repetition of their chemical and physical properties. The result is the periodic table as we know it today. Each new horizontal row of the periodic table corresponds to the beginning of a new **period** because a new principal energy level is being filled with electrons. Elements with similar chemical properties appear at regular intervals, within the vertical columns called **groups**.

### PERIODIC LAW

When elements are arranged in order of increasing atomic number, there is a periodic repetition of their chemical and physical properties

### YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- Elements of the periodic table are arranged in order of increasing atomic number.
- The periodic law states: "When elements are arranged in order of increasing atomic number, there is a periodic repetition of their chemical and physical properties."

---

This page titled 8.10.2: Periodic Law is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 10.34: Nuclear Reactions (Conceptual Questions Q10.4.4) by OpenStax is licensed CC BY 4.0.

# 8.10.3: Modern Periodic Table- Periods and Groups

## Learning Objectives

- Know how groups and periods are used to organize the modern periodic table, and how this relationship expresses the similarities among the elements.

The periodic table has undergone extensive changes in the time since it was originally developed by Mendeleev and Moseley. Many new elements have been discovered, while others have been artificially synthesized. Each fits properly into a **group** of elements with similar properties. The periodic table is an arrangement of the elements in order of their atomic numbers, so that elements with similar properties appear in the same vertical column or group.

The figure below shows the most commonly used form of the periodic table. Each square shows the chemical symbol of the element along with its name. Notice that several of the symbols seem to be unrelated to the name of the element: $\mathrm{Fe}$ for iron, $\mathrm{Pb}$ for lead, etc. Most of these are the elements that have been known since ancient times and have symbols based on their Latin names. The atomic number of each element is written above the symbol.

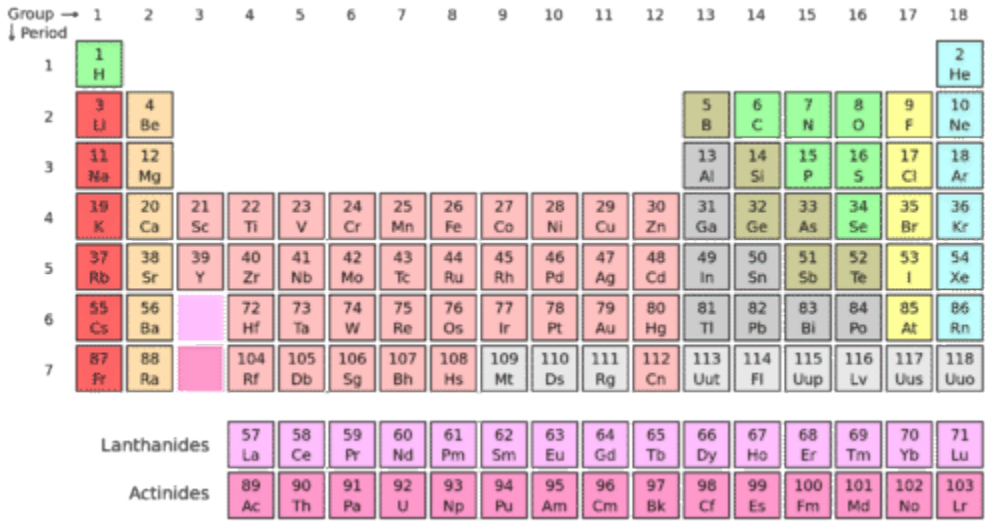



Figure 8.10.3.1 (Credit: User:Cepheus/Wikimedia Commons; Source: wikimedia; License: Public Domain)

A **period** is a horizontal row of the periodic table. There are seven periods in the periodic table, with each one beginning at the far left. A new period begins when a new principal energy level begins filling with electrons. Period 1 has only two elements (hydrogen and helium), while periods 2 and 3 have 8 elements. Periods 4 and 5 have 18 elements. Periods 6 and 7 have 32 elements, because the two bottom rows that are separate from the rest of the table belong to those periods. These two rows are pulled out in order to make the table itself fit more easily onto a single page.

A group is a vertical column of the periodic table, based on the organization of the outer shell electrons. There are a total of 18 groups. There are two different numbering systems that are commonly used to designate groups, and you should be familiar with both. The traditional system used in the United States involves the use of the letters A and B. The first two groups are 1A and 2A, while the last six groups are 3A through 8A. The middle groups use B in their titles. Unfortunately, there was a slightly different system in place in Europe. To eliminate confusion, the International Union of Pure and Applied Chemistry (IUPAC) decided that the official system for numbering groups would be a simple 1 through 18 from left to right. Many periodic tables show both systems simultaneously.

You might notice two rows of elements towards the bottom of the periodic table that seem unconnected to the rest of it. These elements do fit with the rest of the elements, but the width of the periodic table would have to be much greater to fit them in. You can look at the numbers to see where they fit in. These rows of elements are called series, and are named for whichever element

is first in the series: lanthanides and actinides.

Most recent depictions of the periodic table show an incomplete seventh period. In fact, many of those seventh period elements were not known before the early twentieth century. Many elements have been synthesized by bombarding known elements with sub-atomic particles such as neutrons or alpha particles. Uranium has been used in this manner to produce elements 93-100. Larger elements (atomic numbers 101 and above) are formed by fusing nuclei of smaller elements together. These synthetic elements tend to be very unstable, often existing for less than a second, so little is known about them.

The video below reviews groups and trends in the periodic table.

## Section Summary

- The periodic table is arranged in order of atomic number.
- A period is a horizontal row of the periodic table.
- A group is a vertical row of the periodic table.

## Glossary

**period**
: a horizontal row on the periodic table whose properties are repeated in adjacent rows.

**group**
: a vertical column of the periodic table, within which are similar properties.

This page titled 8.10.3: Modern Periodic Table- Periods and Groups is shared under a license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 10.35: Nuclear Reactions (Conceptual Questions Q10.4.5) by OpenStax is licensed CC BY 4.0.
- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 8.11: Chemical Nomenclature

**Learning Objectives**

- Derive names for common types of inorganic compounds using a systematic approach.
- Describe how to name binary covalent compounds including acids and oxyacids.

Nomenclature, a collection of rules for naming things, is important in science and in many other situations. This module describes an approach that is used to name simple ionic and molecular compounds, such as NaCl, $CaCO_3$, and $N_2O_4$. The simplest of these are binary compounds, those containing only two elements, but we will also consider how to name ionic compounds containing polyatomic ions, and one specific, very important class of compounds known as acids (subsequent chapters in this text will focus on these compounds in great detail). We will limit our attention here to inorganic compounds, compounds that are composed principally of elements other than carbon, and will follow the nomenclature guidelines proposed by IUPAC. The rules for organic compounds, in which carbon is the principle element, will be treated in a later chapter on organic chemistry.

## Ionic Compounds

To name an inorganic compound, we need to consider the answers to several questions. First, is the compound ionic or molecular? If the compound is ionic, does the metal form ions of only one type (fixed charge) or more than one type (variable charge)? Are the ions monatomic or polyatomic? If the compound is molecular, does it contain hydrogen? If so, does it also contain oxygen? From the answers we derive, we place the compound in an appropriate category and then name it accordingly. We will begin with the nomenclature rules for ionic compounds.

## Compounds Containing Only Monatomic Ions

The name of a binary compound containing monatomic ions consists of the name of the cation (the name of the metal) followed by the name of the anion (the name of the nonmetallic element with its ending replaced by the suffix *–ide*). Some examples are given in Table 8.11.2.

*Table 8.11.1: Names of Some Ionic Compounds*

| NaCl, sodium chloride | $Na_2O$, sodium oxide |
|---|---|
| KBr, potassium bromide | CdS, cadmium sulfide |
| $CaI_2$, calcium iodide | $Mg_3N_2$, magnesium nitride |
| CsF, cesium fluoride | $Ca_3P_2$, calcium phosphide |
| LiCl, lithium chloride | $Al_4C_3$, aluminum carbide |

## Compounds Containing Polyatomic Ions

Compounds containing polyatomic ions are named similarly to those containing only monatomic ions, except there is no need to change to an *–ide* ending, since the suffix is already present in the name of the anion. Examples are shown in Table 8.11.2.

CL, ammonium chloride, C a S O subscript 4 calcium sulfate, and M g subscript 3 ( P O subscript 4 ) subscript 2 magnesium phosphate." data-quail-id="54" data-mt-width="1246">

*Table 8.11.2: Names of Some Polyatomic Ionic Compounds*

| $KC_2H_3O_2$, potassium acetate | $(NH_4)Cl$, ammonium chloride |
|---|---|
| $NaHCO_3$, sodium bicarbonate | $CaSO_4$, calcium sulfate |

| $KC_2H_3O_2$, potassium acetate | $(NH_4)Cl$, ammonium chloride |
|---|---|
| $Al_2(CO_3)_3$, aluminum carbonate | $Mg_3(PO_4)_2$, magnesium phosphate |

### Ionic Compounds in Your Cabinets

Every day you encounter and use a large number of ionic compounds. Some of these compounds, where they are found, and what they are used for are listed in Table 8.11.3. Look at the label or ingredients list on the various products that you use during the next few days, and see if you run into any of those in this table, or find other ionic compounds that you could now name or write as a formula.

*Table 8.11.3: Everyday Ionic Compounds*

| Ionic Compound | Name | Use |
|---|---|---|
| NaCl | sodium chloride | ordinary table salt |
| KI | potassium iodide | added to “iodized” salt for thyroid health |
| NaF | sodium fluoride | ingredient in toothpaste |
| $NaHCO_3$ | sodium bicarbonate | baking soda; used in cooking (and in antacids) |
| $Na_2CO_3$ | sodium carbonate | washing soda; used in cleaning agents |
| NaOCl | sodium hypochlorite | active ingredient in household bleach |
| $CaCO_3$ | calcium carbonate | ingredient in antacids |
| $Mg(OH)_2$ | magnesium hydroxide | ingredient in antacids |
| $Al(OH)_3$ | aluminum hydroxide | ingredient in antacids |
| NaOH | sodium hydroxide | lye; used as drain cleaner |
| $K_3PO_4$ | potassium phosphate | food additive (many purposes) |
| $MgSO_4$ | magnesium sulfate | added to purified water |
| $Na_2HPO_4$ | sodium hydrogen phosphate | anti-caking agent; used in powdered products |
| $Na_2SO_3$ | sodium sulfite | preservative |

## Compounds Containing a Metal Ion with a Variable Charge

Most of the transition metals can form two or more cations with different charges. Compounds of these metals with nonmetals are named with the same method as compounds in the first category, except the charge of the metal ion is specified by a Roman numeral in parentheses after the name of the metal. The charge of the metal ion is determined from the formula of the compound and the charge of the anion. For example, consider binary ionic compounds of iron and chlorine. Iron typically exhibits a charge of either 2+ or 3+, and the two corresponding compound formulas are $FeCl_2$ and $FeCl_3$. The simplest name, “iron chloride,” will, in this case, be ambiguous, as it does not distinguish between these two compounds. In cases like this, the charge of the metal ion is included as a Roman numeral in parentheses immediately following the metal name. These two compounds are then unambiguously named iron(II) chloride and iron(III) chloride, respectively. Other examples are provided in Table 8.11.4.

*Table 8.11.4: Names of Some Transition Metal Ionic Compounds*

| Transition Metal Ionic Compound | Name |
|---|---|
| $FeCl_3$ | iron(III) chloride |
| $Hg_2O$ | mercury(I) oxide |
| HgO | mercury(II) oxide |

| Transition Metal Ionic Compound | Name |
|---|---|
| $Cu_3(PO_4)_2$ | copper(II) phosphate |

Out-of-date nomenclature used the suffixes *–ic* and *–ous* to designate metals with higher and lower charges, respectively: Iron(III) chloride, $FeCl_3$, was previously called ferric chloride, and iron(II) chloride, $FeCl_2$, was known as ferrous chloride. Though this naming convention has been largely abandoned by the scientific community, it remains in use by some segments of industry. For example, you may see the words *stannous fluoride* on a tube of toothpaste. This represents the formula $SnF_2$, which is more properly named tin(II) fluoride. The other fluoride of tin is $SnF_4$, which was previously called stannic fluoride but is now named tin(IV) fluoride.

## Example 8.11.1: Naming Ionic Compounds

Name the following ionic compounds, which contain a metal that can have more than one ionic charge:

a. $Fe_2S_3$
b. CuSe
c. GaN
d. $CrCl_3$
e. $Ti_2(SO_4)_3$

**Solution**

The anions in these compounds have a fixed negative charge ($S^{2-}$, $Se^{2-}$, $N^{3-}$, $Cl^-$, and $SO_4{}^{2-}$), and the compounds must be neutral. Because the total number of positive charges in each compound must equal the total number of negative charges, the positive ions must be $Fe^{3+}$, $Cu^{2+}$, $Ga^{3+}$, $Cr^{3+}$, and $Ti^{3+}$. These charges are used in the names of the metal ions:

a. iron(III) sulfide
b. copper(II) selenide
c. gallium(III) nitride
d. chromium(III) chloride
e. titanium(III) sulfate

## Exercise 8.11.1

Write the formulas of the following ionic compounds:

a. chromium(III) phosphide
b. mercury(II) sulfide
c. manganese(II) phosphate
d. copper(I) oxide
e. chromium(VI) fluoride

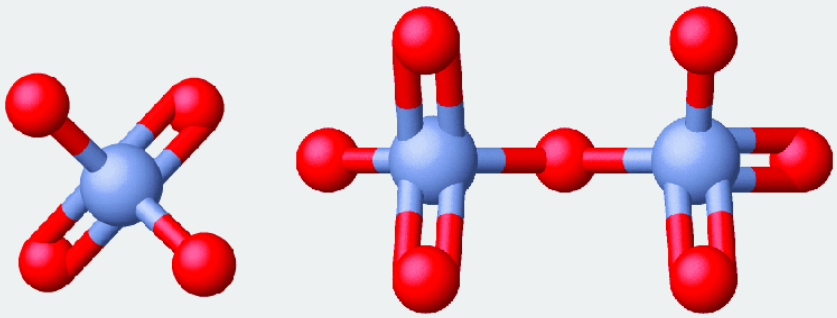
*Answer a available at the end of this section.*

## Erin Brokovich and Chromium Contamination

In the early 1990s, legal file clerk Erin Brockovich (Figure 8.11.2) discovered a high rate of serious illnesses in the small town of Hinckley, California. Her investigation eventually linked the illnesses to groundwater contaminated by Cr(VI) used by Pacific Gas & Electric (PG&E) to fight corrosion in a nearby natural gas pipeline. As dramatized in the film *Erin Brokovich* (for which Julia Roberts won an Oscar), Erin and lawyer Edward Masry sued PG&E for contaminating the water near Hinckley in 1993. The settlement they won in 1996—$333 million—was the largest amount ever awarded for a direct-action lawsuit in the US at that time.



(a)



(b)

Figure 8.11.2: (a) Erin Brockovich found that Cr(VI), used by PG&E, had contaminated the Hinckley, California, water supply. (b) The Cr(VI) ion is often present in water as the polyatomic ions chromate, $CrO_4^{2-}$ (left), and dichromate, $Cr_2O_7^{2-}$ (right).

Chromium compounds are widely used in industry, such as for chrome plating, in dye-making, as preservatives, and to prevent corrosion in cooling tower water, as occurred near Hinckley. In the environment, chromium exists primarily in either the Cr(III) or Cr(VI) forms. Cr(III), an ingredient of many vitamin and nutritional supplements, forms compounds that are not very soluble in water, and it has low toxicity. Cr(VI), on the other hand, is much more toxic and forms compounds that are reasonably soluble in water. Exposure to small amounts of Cr(VI) can lead to damage of the respiratory, gastrointestinal, and immune systems, as well as the kidneys, liver, blood, and skin.

Despite cleanup efforts, Cr(VI) groundwater contamination remains a problem in Hinckley and other locations across the globe. A 2010 study by the Environmental Working Group found that of 35 US cities tested, 31 had higher levels of Cr(VI) in their tap water than the public health goal of 0.02 parts per billion set by the California Environmental Protection Agency.

## Molecular (Covalent) Compounds

The bonding characteristics of inorganic molecular compounds are different from ionic compounds, and they are named using a different system as well. The charges of cations and anions dictate their ratios in ionic compounds, so specifying the names of the ions provides sufficient information to determine chemical formulas. However, because covalent bonding allows for significant variation in the combination ratios of the atoms in a molecule, the names for molecular compounds must explicitly identify these ratios.

## Compounds Composed of Two Elements

When two nonmetallic elements form a molecular compound, several combination ratios are often possible. For example, carbon and oxygen can form the compounds CO and $CO_2$. Since these are different substances with different properties, they cannot both have the same name (they cannot both be called carbon oxide). To deal with this situation, we use a naming method that is somewhat similar to that used for ionic compounds, but with added prefixes to specify the numbers of atoms of each element. The name of the more metallic element (the one farther to the left and/or bottom of the periodic table) is first, followed by the name of the more nonmetallic element (the one farther to the right and/or top) with its ending changed to the suffix *–ide*. The numbers of atoms of each element are designated by the Greek prefixes shown in Table 8.11.5.

*Table 8.11.5: Nomenclature Prefixes*

| Number | Prefix | Number | Prefix |
|---|---|---|---|
| 1 (sometimes omitted) | mono- | 6 | hexa- |
| 2 | di- | 7 | hepta- |
| 3 | tri- | 8 | octa- |

| Number | Prefix | Number | Prefix |
|---|---|---|---|
| 4 | tetra- | 9 | nona- |
| 5 | penta- | 10 | deca- |

When only one atom of the first element is present, the prefix *mono-* is usually deleted from that part. Thus, CO is named carbon monoxide, and $CO_2$ is called carbon dioxide. When two vowels are adjacent, the *a* in the Greek prefix is usually dropped. Some other examples are shown in Table 8.11.6.

*Table 8.11.6: Names of Some Molecular Compounds Composed of Two Elements*

| Compound | Name | Compound | Name |
|---|---|---|---|
| $SO_2$ | sulfur dioxide | $BCl_3$ | boron trichloride |
| $SO_3$ | sulfur trioxide | $SF_6$ | sulfur hexafluoride |
| $NO_2$ | nitrogen dioxide | $PF_5$ | phosphorus pentafluoride |
| $N_2O_4$ | dinitrogen tetroxide | $P_4O_{10}$ | tetraphosphorus decaoxide |
| $N_2O_5$ | dinitrogen pentoxide | $IF_7$ | iodine heptafluoride |

There are a few common names that you will encounter as you continue your study of chemistry. For example, although NO is often called nitric oxide, its proper name is nitrogen monoxide. Similarly, $N_2O$ is known as nitrous oxide even though our rules would specify the name dinitrogen monoxide. (And $H_2O$ is usually called water, not dihydrogen monoxide.) You should commit to memory the common names of compounds as you encounter them.

## Example 8.11.2: Naming Covalent Compounds

Name the following covalent compounds:

a. $SF_6$
b. $N_2O_3$
c. $Cl_2O_7$
d. $P_4O_6$

**Solution**

Because these compounds consist solely of nonmetals, we use prefixes to designate the number of atoms of each element:

a. sulfur hexafluoride
b. dinitrogen trioxide
c. dichlorine heptoxide
d. tetraphosphorus hexoxide

## Exercise 8.11.2

Write the formulas for the following compounds:

a. phosphorus pentachloride
b. dinitrogen monoxide
c. iodine heptafluoride
d. carbon tetrachloride

*Answer a available at the end of this section.*

## Binary Acids

Some compounds containing hydrogen are members of an important class of substances known as acids. The chemistry of these compounds is explored in more detail in later chapters of this text, but for now, it will suffice to note that many acids release hydrogen ions, $H^+$, when dissolved in water. To denote this distinct chemical property, a mixture of water with an acid is given a name derived from the compound's name. If the compound is a binary acid (comprised of hydrogen and one other nonmetallic element):

1. The word "hydrogen" is changed to the prefix *hydro-*
2. The other nonmetallic element name is modified by adding the suffix *-ic*
3. The word "acid" is added as a second word

For example, when the gas $HCl$ (hydrogen chloride) is dissolved in water, the solution is called *hydrochloric acid*. Several other examples of this nomenclature are shown in Table 8.11.7.

*Table 8.11.7: Names of Some Simple Acids*

| Name of Gas | Name of Acid |
|---|---|
| HF(*g*), hydrogen fluoride | HF(*aq*), hydrofluoric acid |
| HCl(*g*), hydrogen chloride | HCl(*aq*), hydrochloric acid |
| HBr(*g*), hydrogen bromide | HBr(*aq*), hydrobromic acid |
| HI(*g*), hydrogen iodide | HI(*aq*), hydroiodic acid |
| $H_2S$(*g*), hydrogen sulfide | $H_2S$(*aq*), hydrosulfuric acid |

## Oxyacids

Many compounds containing three or more elements (such as organic compounds or coordination compounds) are subject to specialized nomenclature rules that you will learn later. However, we will briefly discuss the important compounds known as oxyacids, compounds that contain hydrogen, oxygen, and at least one other element, and are bonded in such a way as to impart acidic properties to the compound (you will learn the details of this in a later chapter). Typical oxyacids consist of hydrogen combined with a polyatomic, oxygen-containing ion. To name oxyacids:

1. Omit "hydrogen"
2. Start with the root name of the anion
3. Replace *–ate* with *–ic*, or *–ite* with *–ous*
4. Add "acid"

For example, consider $H_2CO_3$ (which you might be tempted to call "hydrogen carbonate"). To name this correctly, "hydrogen" is omitted; the *–ate* of carbonate is replace with *–ic*; and acid is added—so its name is carbonic acid. Other examples are given in Table 8.11.8. There are some exceptions to the general naming method (e.g., $H_2SO_4$ is called sulfuric acid, not sulfic acid, and $H_2SO_3$ is sulfurous, not sulfous, acid).

*Table 8.11.8: Names of Common Oxyacids*

| Formula | Anion Name | Acid Name |
|---|---|---|
| $HC_2H_3O_2$ | acetate | acetic acid |
| $HNO_3$ | nitrate | nitric acid |
| $HNO_2$ | nitrite | nitrous acid |
| $HClO_4$ | perchlorate | perchloric acid |
| $H_2CO_3$ | carbonate | carbonic acid |
| $H_2SO_4$ | sulfate | sulfuric acid |

| Formula | Anion Name | Acid Name |
|---|---|---|
| $H_2SO_3$ | sulfite | sulfurous acid |
| $H_3PO_4$ | phosphate | phosphoric acid |

## Summary

Chemists use nomenclature rules to clearly name compounds. Ionic and molecular compounds are named using somewhat-different methods. Binary ionic compounds typically consist of a metal and a nonmetal. The name of the metal is written first, followed by the name of the nonmetal with its ending changed to *–ide*. For example, $K_2O$ is called potassium oxide. If the metal can form ions with different charges, a Roman numeral in parentheses follows the name of the metal to specify its charge. Thus, $FeCl_2$ is iron(II) chloride and $FeCl_3$ is iron(III) chloride. Some compounds contain polyatomic ions; the names of common polyatomic ions should be memorized. Molecular compounds can form compounds with different ratios of their elements, so prefixes are used to specify the numbers of atoms of each element in a molecule of the compound. Examples include $SF_6$, sulfur hexafluoride, and $N_2O_4$, dinitrogen tetroxide. Acids are an important class of compounds containing hydrogen and having special nomenclature rules. Binary acids are named using the prefix *hydro-*, changing the *–ide* suffix to *–ic*, and adding "acid;" HCl is hydrochloric acid. Oxyacids are named by changing the ending of the anion (*-ate* to *–ic*, and *-ite* to *-ous*) and adding "acid;" $H_2CO_3$ is carbonic acid.

## Glossary

**binary acid**
: compound that contains hydrogen and one other element, bonded in a way that imparts acidic properties to the compound (ability to release $H^+$ ions when dissolved in water)

**binary compound**
: compound containing two different elements.

**oxyacid**
: compound that contains hydrogen, oxygen, and one other element, bonded in a way that imparts acidic properties to the compound (ability to release $H^+$ ions when dissolved in water)

**nomenclature**
: system of rules for naming objects of interest

## Answers

**Exercise** 8.11.1

CrP

HgS

$Mn_3(PO_4)_2$

$Cu_2O$

$CrF_6$

**Exercise** 8.11.2

$PCl_5$

$N_2O$

$IF_7$

$CCl_4$

This page titled 8.11: Chemical Nomenclature is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 8.02: Phases and Classification of Matter

## Learning Objectives

- Describe the basic properties of each physical state of matter: solid, liquid, and gas.
- Define and give examples of atoms and molecules.
- Classify matter as an element, compound, homogeneous mixture, or heterogeneous mixture with regard to its physical state and composition.
- Use symbolic, particulate, or macroscopic representations to describe or classify the different types of matter.
- Distinguish between mass and weight.
- Apply the law of conservation of matter.

Matter is defined as anything that occupies space and has mass, and it is all around us. Solids and liquids are more obviously matter: We can see that they take up space, and their weight tells us that they have mass. Gases are also matter; if gases did not take up space, a balloon would stay collapsed rather than inflate when filled with gas.

Solids, liquids, and gases are the three states of matter commonly found on earth (Figure 8.2.1). A solid is rigid and possesses a definite shape. A liquid flows and takes the shape of a container, except that it forms a flat or slightly curved upper surface when acted upon by gravity. (In zero gravity, liquids assume a spherical shape.) Both liquid and solid samples have volumes that are very nearly independent of pressure. A gas takes both the shape and volume of its container.

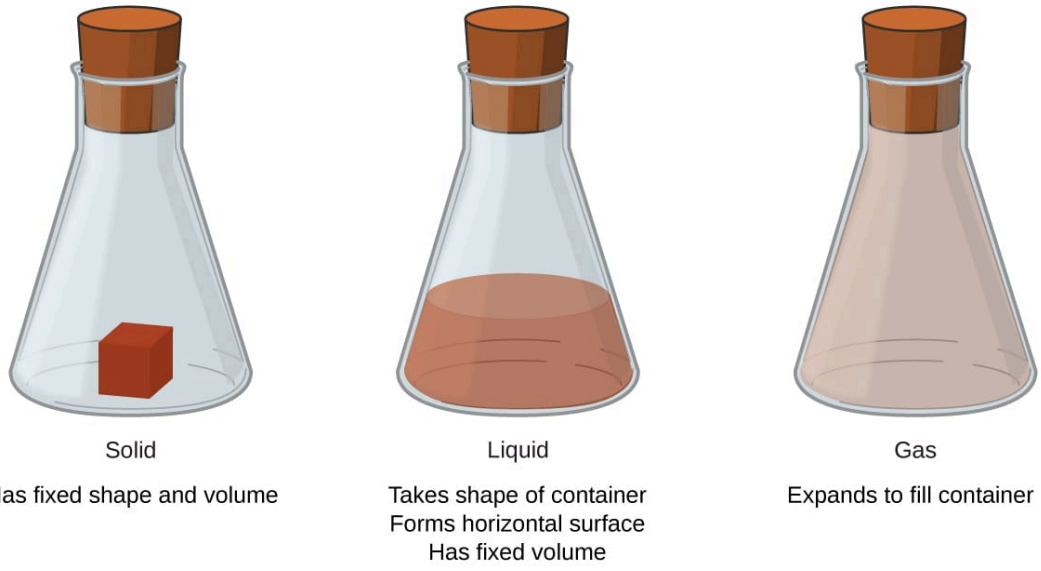



Figure 8.2.1: The three most common states or phases of matter are solid, liquid, and gas. A beaker labeled solid contains a cube of red matter and says has fixed shape and volume. A beaker labeled liquid contains a brownish-red colored liquid. This beaker says takes shape of container, forms horizontal surfaces, has fixed volume. The beaker labeled gas is filled with a light brown gas. This beaker says expands to fill container.

A fourth state of matter, plasma, occurs naturally in the interiors of stars. A plasma is a gaseous state of matter that contains appreciable numbers of electrically charged particles (Figure 8.2.2). The presence of these charged particles imparts unique properties to plasmas that justify their classification as a state of matter distinct from gases. In addition to stars, plasmas are found in some other high-temperature environments (both natural and man-made), such as lightning strikes, certain television screens, and specialized analytical instruments used to detect trace amounts of metals.



Figure 8.2.2: A plasma torch can be used to cut metal. (credit: "Hypertherm"/Wikimedia Commons) A cutting torch is being used to cut a piece of metal. Bright, white colored plasma can be seen near the tip of the torch, where it is contacting the metal.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

***Video*** 8.2.1*: In a tiny cell in a plasma television, the plasma emits ultraviolet light, which in turn causes the display at that location to appear a specific color. The composite of these tiny dots of color makes up the image that you see. Watch this video to learn more about plasma and the places you encounter it.*

Some samples of matter appear to have properties of solids, liquids, and/or gases at the same time. This can occur when the sample is composed of many small pieces. For example, we can pour sand as if it were a liquid because it is composed of many small grains of solid sand. Matter can also have properties of more than one state when it is a mixture, such as with clouds. Clouds appear to behave somewhat like gases, but they are actually mixtures of air (gas) and tiny particles of water (liquid or solid).

The mass of an object is a measure of the amount of matter in it. One way to measure an object's mass is to measure the force it takes to accelerate the object. It takes much more force to accelerate a car than a bicycle because the car has much more mass. A more common way to determine the mass of an object is to use a balance to compare its mass with a standard mass.

Although weight is related to mass, it is not the same thing. Weight refers to the force that gravity exerts on an object. This force is directly proportional to the mass of the object. The weight of an object changes as the force of gravity changes, but its mass does not. An astronaut's mass does not change just because she goes to the moon. But her weight on the moon is only one-sixth her earth-bound weight because the moon's gravity is only one-sixth that of the earth's. She may feel "weightless" during her trip when she experiences negligible external forces (gravitational or any other), although she is, of course, never "massless."

The law of conservation of matter summarizes many scientific observations about matter: It states that *there is no detectable change in the total quantity of matter present when matter converts from one type to another (a chemical change) or changes among solid, liquid, or gaseous states (a physical change)*. Brewing beer and the operation of batteries provide examples of the conservation of matter (Figure 8.2.4). During the brewing of beer, the ingredients (water, yeast, grains, malt, hops, and sugar) are converted into beer (water, alcohol, carbonation, and flavoring substances) with no actual loss of substance. This is most clearly seen during the bottling process, when glucose turns into ethanol and carbon dioxide, and the total mass of the substances does not change. This can also be seen in a lead-acid car battery: The original substances (lead, lead oxide, and sulfuric acid), which

are capable of producing electricity, are changed into other substances (lead sulfate and water) that do not produce electricity, with no change in the actual amount of matter.

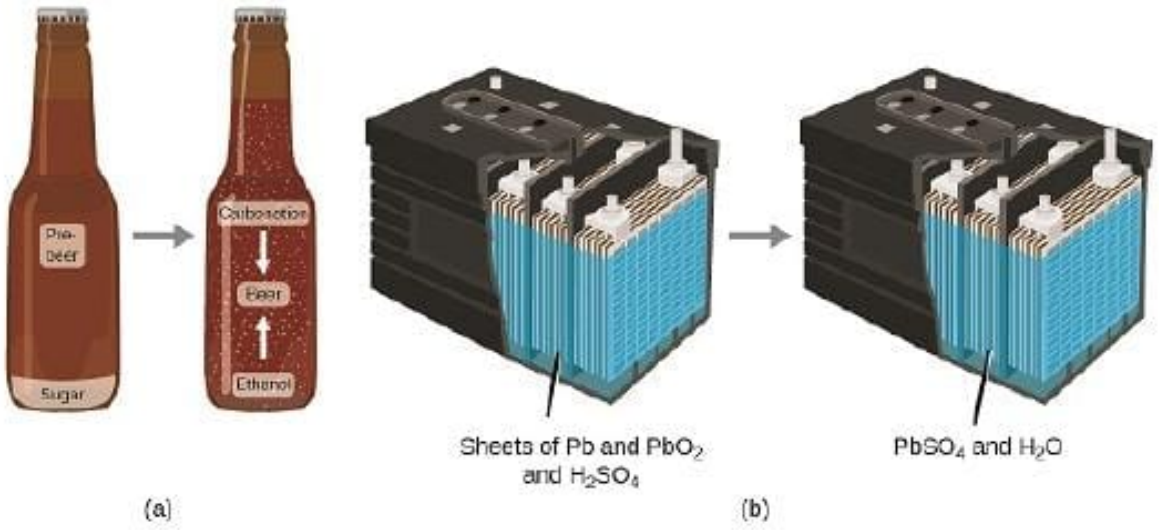



Figure 8.2.3: (a) The mass of beer precursor materials is the same as the mass of beer produced: Sugar has become alcohol and carbonation. (b) The mass of the lead, lead oxide plates, and sulfuric acid that goes into the production of electricity is exactly equal to the mass of lead sulfate and water that is formed. Diagram A shows a beer bottle containing pre-beer and sugar. An arrow points from this bottle to a second bottle. This second bottle contains the same volume of liquid, however, the sugar has been converted into ethanol and carbonation as beer was made. Diagram B shows a car battery that contains sheets of P B and P B O subscript 2 along with H subscript 2 S O subscript 4. After the battery is used, it contains an equal mass of P B S O subscript 4 and H subscript 2 O.

Although this conservation law holds true for all conversions of matter, convincing examples are few and far between because, outside of the controlled conditions in a laboratory, we seldom collect all of the material that is produced during a particular conversion. For example, when you eat, digest, and assimilate food, all of the matter in the original food is preserved. But because some of the matter is incorporated into your body, and much is excreted as various types of waste, it is challenging to verify by measurement.

## Atoms and Molecules

An atom is the smallest particle of an element that has the properties of that element and can enter into a chemical combination. Consider the element gold, for example. Imagine cutting a gold nugget in half, then cutting one of the halves in half, and repeating this process until a piece of gold remained that was so small that it could not be cut in half (regardless of how tiny your knife may be). This minimally sized piece of gold is an atom (from the Greek *atomos*, meaning “indivisible”) (Figure 1.2.4). This atom would no longer be gold if it were divided any further.

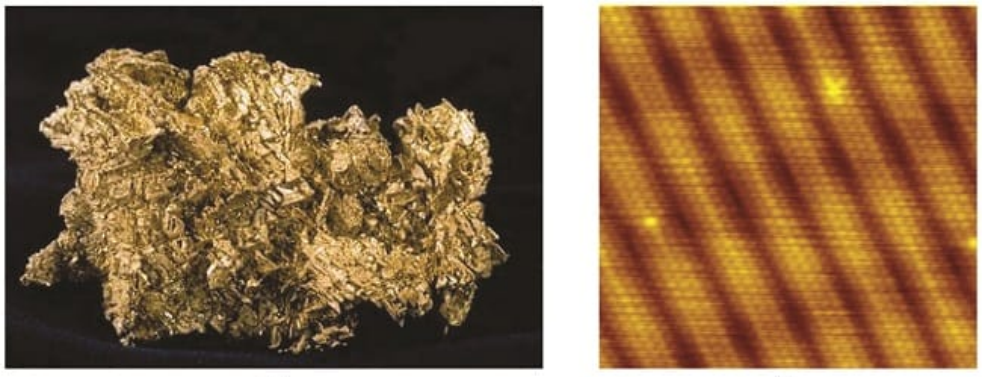



Figure 8.2.4: (a) This photograph

shows a gold nugget. (b) A scanning-tunneling microscope (STM) can generate views of the surfaces of solids, such as this image of a gold crystal. Each sphere represents one gold atom. (credit a: modification of work by United States Geological Survey; credit b: modification of work by "Erwinrossen"/Wikimedia Commons) Figure A shows a gold nugget as it would appear to the naked eye. The gold nugget is very irregular, with many sharp edges. It appears gold in color. The microscope image of a gold crystal shows many similarly sized gold stripes that are separated by dark areas. Looking closely, one can see that the gold stripes are made of many, tiny, circular atoms.

The first suggestion that matter is composed of atoms is attributed to the Greek philosophers Leucippus and Democritus, who developed their ideas in the 5th century BCE. However, it was not until the early nineteenth century that John Dalton (1766–1844), a British schoolteacher with a keen interest in science, supported this hypothesis with quantitative measurements. Since that time, repeated experiments have confirmed many aspects of this hypothesis, and it has become one of the central theories of chemistry. Other aspects of Dalton's atomic theory are still used but with minor revisions (details of Dalton's theory are provided in the chapter on atoms and molecules).

An atom is so small that its size is difficult to imagine. One of the smallest things we can see with our unaided eye is a single thread of a spider web: These strands are about 1/10,000 of a centimeter (0.0001 cm) in diameter. Although the cross-section of one strand is almost impossible to see without a microscope, it is huge on an atomic scale. A single carbon atom in the web has a diameter of about 0.000000015 centimeter, and it would take about 7000 carbon atoms to span the diameter of the strand. To put this in perspective, if a carbon atom were the size of a dime, the cross-section of one strand would be larger than a football field, which would require about 150 million carbon atom "dimes" to cover it. (Figure 8.2.5) shows increasingly close microscopic and atomic-level views of ordinary cotton.




Figure 8.2.5: These images provide an increasingly closer view: (a) a cotton boll, (b) a single cotton fiber viewed under an optical microscope (magnified 40 times), (c) an image of a cotton fiber obtained with an electron microscope (much higher magnification than with the optical microscope); and (d and e) atomic-level models of the fiber (spheres of different colors represent atoms of different elements). (credit c: modification of work by "Featheredtar"/Wikimedia Commons) Figure A shows a puffy white cotton boll growing on a brown twig. Figure B shows a magnified cotton strand. The strand appears transparent but contains dark areas within its interior. Figure C shows the surface of several crisscrossing and overlapping cotton fibers. Its surface is rough along the edges but smooth near the center of each strand. Figure D shows three strands of molecules connected into three vertical chains. Each strand contains about five molecules. Figure E shows that the cotton molecule contains about a dozen atoms. The black carbon atoms form rings that are connected by red oxygen atoms. Many of the carbon atoms are also bonded to hydrogen atoms, shown as white balls, or other oxygen atoms.

An atom is so light that its mass is also difficult to imagine. A billion lead atoms (1,000,000,000 atoms) weigh about $3 \times 10^{-13}$ grams, a mass that is far too light to be weighed on even the world's most sensitive balances. It would require over 300,000,000,000,000 lead atoms (300 trillion, or $3 \times 10^{14}$) to be weighed, and they would weigh only 0.0000001 gram.

It is rare to find collections of individual atoms. Only a few elements, such as the gases helium, neon, and argon, consist of a collection of individual atoms that move about independently of one another. Other elements, such as the gases hydrogen, nitrogen, oxygen, and chlorine, are composed of units that consist of pairs of atoms (Figure 8.2.6). One form of the element phosphorus consists of units composed of four phosphorus atoms. The element sulfur exists in various forms, one of which consists of units composed of eight sulfur atoms. These units are called molecules. A molecule consists of two or more atoms joined by strong forces called chemical bonds. The atoms in a molecule move around as a unit, much like the cans of soda in a six-pack or a bunch of keys joined together on a single key ring. A molecule may consist of two or more identical atoms, as in the molecules found in the elements hydrogen, oxygen, and sulfur, or it may consist of two or more different atoms, as in the molecules found in water. Each water molecule is a unit that contains two hydrogen atoms and one oxygen atom. Each glucose molecule is a unit that contains 6 carbon atoms, 12 hydrogen atoms, and 6 oxygen atoms. Like atoms, molecules are incredibly

small and light. If an ordinary glass of water were enlarged to the size of the earth, the water molecules inside it would be about the size of golf balls.

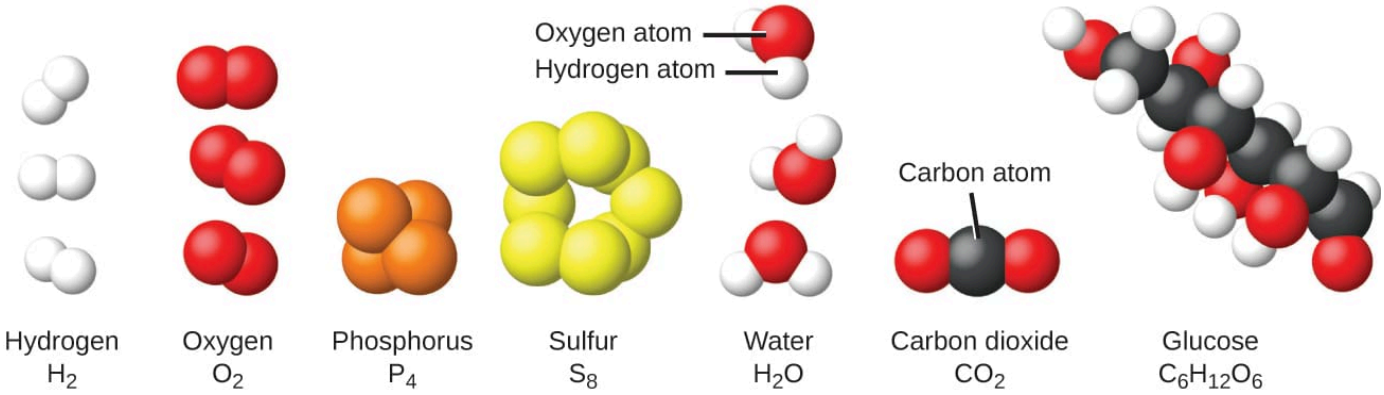



Figure 8.2.6: The elements hydrogen, oxygen, phosphorus, and sulfur form molecules consisting of two or more atoms of the same element. The compounds water, carbon dioxide, and glucose consist of combinations of atoms of different elements. The hydrogen molecule, H subscript 2, is shown as two small, white balls bonded together. The oxygen molecule O subscript 2, is shown as two red balls bonded together. The phosphorous molecule, P subscript 4, is shown as four orange balls bonded tightly together. The sulfur molecule, S subscript 8, is shown as 8 yellow balls linked together. Water molecules, H subscript 2 O, consist of one red oxygen atom bonded to two smaller white hydrogen atoms. The hydrogen atoms are at an angle on the oxygen molecule. Carbon dioxide, C O subscript 2, consists of one carbon atom and two oxygen atoms. One oxygen atom is bonded to the carbon's right side and the other oxygen is bonded to the carbon's left side. Glucose, C subscript 6 H subscript 12 O subscript 6, contains a chain of carbon atoms that have attached oxygen or hydrogen atoms.

## Classifying Matter

We can classify matter into several categories. Two broad categories are mixtures and pure substances. A pure substance has a constant composition. All specimens of a pure substance have exactly the same makeup and properties. Any sample of sucrose (table sugar) consists of 42.1% carbon, 6.5% hydrogen, and 51.4% oxygen by mass. Any sample of sucrose also has the same physical properties, such as melting point, color, and sweetness, regardless of the source from which it is isolated.

We can divide pure substances into two classes: elements and compounds. Pure substances that cannot be broken down into simpler substances by chemical changes are called elements. Iron, silver, gold, aluminum, sulfur, oxygen, and copper are familiar examples of the more than 100 known elements, of which about 90 occur naturally on the earth, and two dozen or so have been created in laboratories.

Pure substances that can be broken down by chemical changes are called compounds. This breakdown may produce either elements or other compounds, or both. Mercury(II) oxide, an orange, crystalline solid, can be broken down by heat into the elements mercury and oxygen (Figure 8.2.7). When heated in the absence of air, the compound sucrose is broken down into the element carbon and the compound water. (The initial stage of this process, when the sugar is turning brown, is known as caramelization—this is what imparts the characteristic sweet and nutty flavor to caramel apples, caramelized onions, and caramel). Silver(I) chloride is a white solid that can be broken down into its elements, silver and chlorine, by absorption of light. This property is the basis for the use of this compound in photographic films and photochromic eyeglasses (those with lenses that darken when exposed to light).

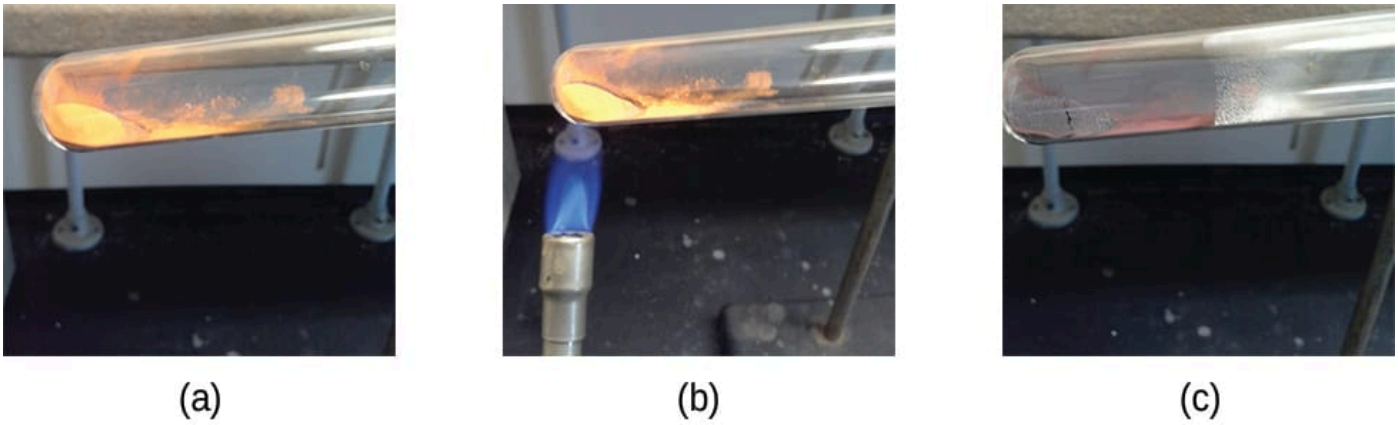



Figure 8.2.7: (a)The compound mercury(II) oxide, (b)when heated, (c) decomposes into silvery droplets of liquid mercury and invisible oxygen gas. (credit: modification of work by Paul Flowers) This figure shows a series of three photos labeled a, b, and c. Photo a shows the bottom of a test tube that is filled with an orange-red substance. A slight amount of a silver substance is also visible. Photo b shows the substance in the test tube being heated over a flame. Photo c shows a test tube that is not longer being heated. The orange-red substance is almost completely gone, and small, silver droplets of a substance are left.

The properties of combined elements are different from those in the free, or uncombined, state. For example, white crystalline sugar (sucrose) is a compound resulting from the chemical combination of the element carbon, which is a black solid in one of its uncombined forms, and the two elements hydrogen and oxygen, which are colorless gases when uncombined. Free sodium, an element that is a soft, shiny, metallic solid, and free chlorine, an element that is a yellow-green gas, combine to form sodium chloride (table salt), a compound that is a white, crystalline solid.

A mixture is composed of two or more types of matter that can be present in varying amounts and can be separated by physical changes, such as evaporation (you will learn more about this later). A mixture with a composition that varies from point to point is called a heterogeneous mixture. Italian dressing is an example of a heterogeneous mixture (Figure 8.2.1*a*). Its composition can vary because we can make it from varying amounts of oil, vinegar, and herbs. It is not the same from point to point throughout the mixture—one drop may be mostly vinegar, whereas a different drop may be mostly oil or herbs because the oil and vinegar separate and the herbs settle. Other examples of heterogeneous mixtures are chocolate chip cookies (we can see the separate bits of chocolate, nuts, and cookie dough) and granite (we can see the quartz, mica, feldspar, and more).

A homogeneous mixture, also called a solution, exhibits a uniform composition and appears visually the same throughout. An example of a solution is a sports drink, consisting of water, sugar, coloring, flavoring, and electrolytes mixed together uniformly (Figure 8.2.1*b*). Each drop of a sports drink tastes the same because each drop contains the same amounts of water, sugar, and other components. Note that the composition of a sports drink can vary—it could be made with somewhat more or less sugar, flavoring, or other components, and still be a sports drink. Other examples of homogeneous mixtures include air, maple syrup, gasoline, and a solution of salt in water.




Figure 8.2.7: (a) Oil and vinegar salad dressing is a heterogeneous mixture because its composition is not uniform throughout. (b) A commercial sports drink is a homogeneous mixture because its composition is uniform throughout. (credit a “left”: modification of work by John Mayer; credit a “right”: modification of work by Umberto Salvagnin; credit b “left: modification of work by Jeff Bedford) Diagram A shows a glass containing a red liquid with a layer of yellow oil floating on the surface of the red liquid. A

zoom in box is magnifying a portion of the red liquid that contains some of the yellow oil. The zoomed in image shows that oil is forming round droplets within the red liquid. Diagram B shows a photo of Gatorade G 2. A zoom in box is magnifying a portion of the Gatorade, which is uniformly red.

Although there are just over 100 elements, tens of millions of chemical compounds result from different combinations of these elements. Each compound has a specific composition and possesses definite chemical and physical properties by which we can distinguish it from all other compounds. And, of course, there are innumerable ways to combine elements and compounds to form different mixtures. A summary of how to distinguish between the various major classifications of matter is shown in (Figure 1.2.8).

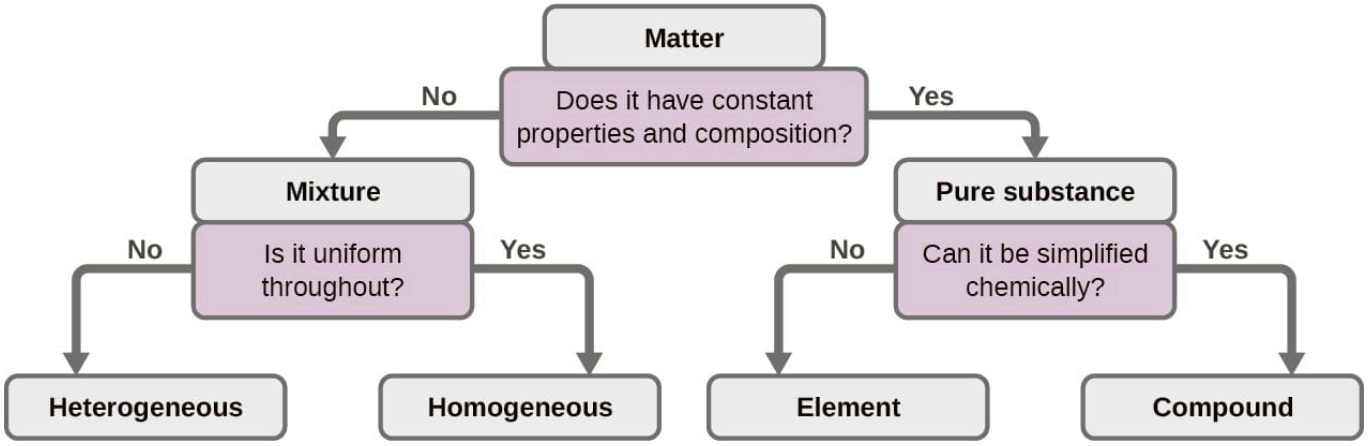



Figure 8.2.8: Depending on its properties, a given substance can be classified as a homogeneous mixture, a heterogeneous mixture, a compound, or an element. This flow chart begins with matter at the top and the question: does the matter have constant properties and composition? If no, then it is a mixture. This leads to the next question: is it uniform throughout? If no, it is heterogeneous. If yes, it is homogenous. If the matter does have constant properties and composition, it is a pure substance. This leads to the next question: can it be simplified chemically? If no, it is an element. If yes, then it is a compound.

Eleven elements make up about 99% of the earth's crust and atmosphere (Table 8.2.1). Oxygen constitutes nearly one-half and silicon about one-quarter of the total quantity of these elements. A majority of elements on earth are found in chemical combinations with other elements; about one-quarter of the elements are also found in the free state.

*Table 8.2.1: Elemental Composition of Earth*

| Element | Symbol | Percent Mass | | Element | Symbol | Percent Mass |
|---|---|---|---|---|---|---|
| oxygen | O | 49.20 | | chlorine | Cl | 0.19 |
| silicon | Si | 25.67 | | phosphorus | P | 0.11 |
| aluminum | Al | 7.50 | | manganese | Mn | 0.09 |
| iron | Fe | 4.71 | | carbon | C | 0.08 |
| calcium | Ca | 3.39 | | sulfur | S | 0.06 |
| sodium | Na | 2.63 | | barium | Ba | 0.04 |
| potassium | K | 2.40 | | nitrogen | N | 0.03 |
| magnesium | Mg | 1.93 | | fluorine | F | 0.03 |
| hydrogen | H | 0.87 | | strontium | Sr | 0.02 |
| titanium | Ti | 0.58 | | all others | - | 0.47 |

## Decomposition of Water / Production of Hydrogen

Water consists of the elements hydrogen and oxygen combined in a 2 to 1 ratio. Water can be broken down into hydrogen and oxygen gases by the addition of energy. One way to do this is with a battery or power supply, as shown in (Figure 8.2.9).

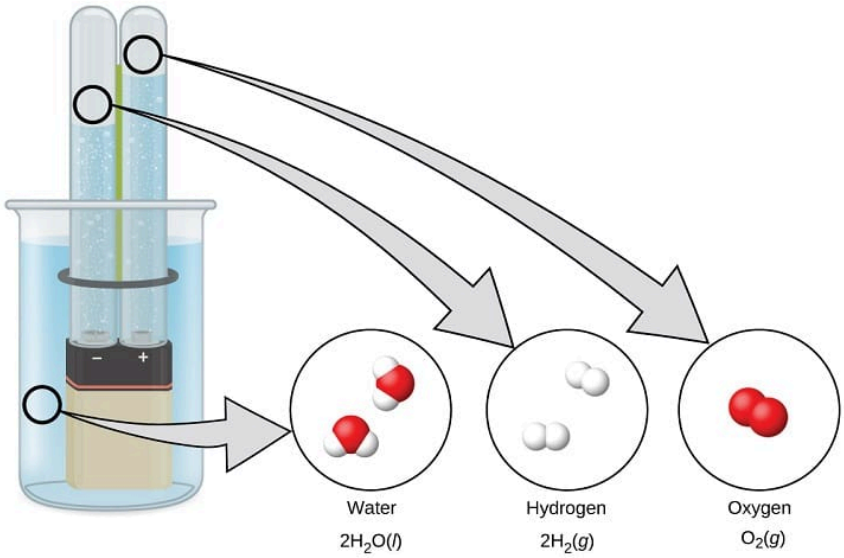



Figure 8.2.9: The decomposition of water is shown at the macroscopic, microscopic, and symbolic levels. The battery provides an electric current (microscopic) that decomposes water. At the macroscopic level, the liquid separates into the gases hydrogen (on the left) and oxygen (on the right). Symbolically, this change is presented by showing how liquid $H_2O$ separates into $H_2$ and $O_2$ gases. A rectangular battery is immersed in a beaker filled with liquid. Each of the battery terminals are covered by an overturned test tube. The test tubes each contain a bubbling liquid. Zoom in areas indicate that the liquid in the beaker is water, 2 H subscript 2 O liquid. The bubbles in the test tube over the negative terminal are hydrogen gas, 2 H subscript 2 gas. The bubbles in the test tube over the positive terminal are oxygen gas, O subscript 2 gas.

The breakdown of water involves a rearrangement of the atoms in water molecules into different molecules, each composed of two hydrogen atoms and two oxygen atoms, respectively. Two water molecules form one oxygen molecule and two hydrogen molecules. The representation for what occurs, $2\,\mathrm{H_2O}(l) \rightarrow 2\,\mathrm{H_2}(g) + \mathrm{O_2}(g)$, will be explored in more depth in later chapters.

The two gases produced have distinctly different properties. Oxygen is not flammable but is required for combustion of a fuel, and hydrogen is highly flammable and a potent energy source. How might this knowledge be applied in our world? One application involves research into more fuel-efficient transportation. Fuel-cell vehicles (FCV) run on hydrogen instead of gasoline (Figure 8.2.10). They are more efficient than vehicles with internal combustion engines, are nonpolluting, and reduce greenhouse gas emissions, making us less dependent on fossil fuels. FCVs are not yet economically viable, however, and current hydrogen production depends on natural gas. If we can develop a process to economically decompose water, or produce hydrogen in another environmentally sound way, FCVs may be the way of the future.

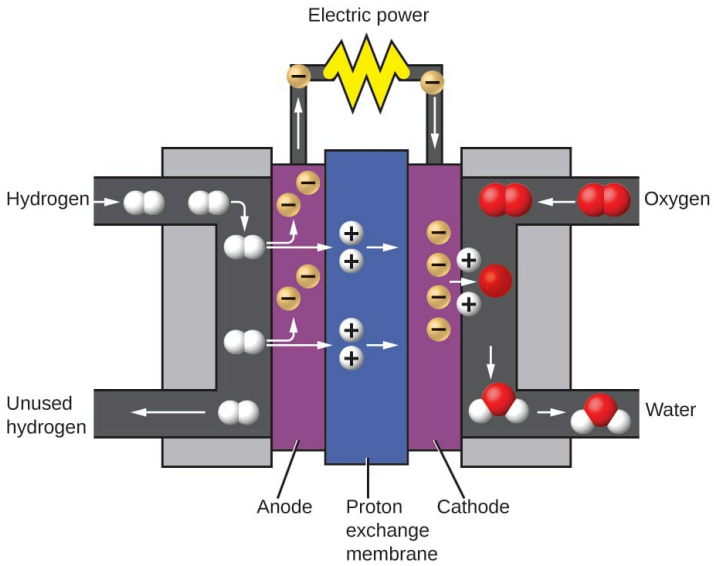



Figure 8.2.10: A fuel cell generates electrical energy from

hydrogen and oxygen via an electrochemical process and produces only water as the waste product. The fuel cell consists of a proton exchange membrane sandwiched between an anode and a cathode. Hydrogen gas enters the battery near the anode. Oxygen gas enters the battery near the cathode. The entering hydrogen gas is broken up into single white spheres that each have a positive charge. These are protons. The protons repel negatively-charged electrons within the anode. These electrons travel through a circuit, providing electricity to anything attached to the battery. The protons continue through the proton exchange membrane and through the cathode to reach the oxygen gas molecules at the opposite end of the battery. There, the oxygen atoms split up into single red spheres. Each oxygen atom takes on two of the incoming protons to form a water molecule.

## Chemistry of Cell Phones

Imagine how different your life would be without cell phones (Figure 8.2.11) and other smart devices. Cell phones are made from numerous chemical substances, which are extracted, refined, purified, and assembled using an extensive and in-depth understanding of chemical principles. About 30% of the elements that are found in nature are found within a typical smart phone. The case/body/frame consists of a combination of sturdy, durable polymers comprised primarily of carbon, hydrogen, oxygen, and nitrogen [acrylonitrile butadiene styrene (ABS) and polycarbonate thermoplastics], and light, strong, structural metals, such as aluminum, magnesium, and iron. The display screen is made from a specially toughened glass (silica glass strengthened by the addition of aluminum, sodium, and potassium) and coated with a material to make it conductive (such as indium tin oxide). The circuit board uses a semiconductor material (usually silicon); commonly used metals like copper, tin, silver, and gold; and more unfamiliar elements such as yttrium, praseodymium, and gadolinium. The battery relies upon lithium ions and a variety of other materials, including iron, cobalt, copper, polyethylene oxide, and polyacrylonitrile.




Figure 8.2.11: Almost one-third of naturally occurring elements are used to make a modern cell phone. (credit: modification of work by John Taylor) A cell phone is labeled to show what its components are made of. The case components are made of polymers such as A B S and or metals such as aluminum, iron, and magnesium. The processor components are made of silicon, common metals such as copper, tin and gold, and uncommon elements such as yttrium and gadolinium. The screen components are made of silicon oxide, also known as glass. The glass is strengthened by the addition of aluminum, sodium, and potassium. The battery components contain lithium combined with other metals such as cobalt, iron, and copper.

## Summary

Matter is anything that occupies space and has mass. The basic building block of matter is the atom, the smallest unit of an element that can enter into combinations with atoms of the same or other elements. In many substances, atoms are combined into molecules. On earth, matter commonly exists in three states: solids, of fixed shape and volume; liquids, of variable shape but fixed volume; and gases, of variable shape and volume. Under high-temperature conditions, matter also can exist as a plasma. Most matter is a mixture: It is composed of two or more types of matter that can be present in varying amounts and can be separated by physical means. Heterogeneous mixtures vary in composition from point to point; homogeneous mixtures have the same composition from point to point. Pure substances consist of only one type of matter. A pure substance can be an element, which consists of only one type of atom and cannot be broken down by a chemical change, or a compound, which consists of two

or more types of atoms.

## Glossary

**atom**
: smallest particle of an element that can enter into a chemical combination

**compound**
: pure substance that can be decomposed into two or more elements

**element**
: substance that is composed of a single type of atom; a substance that cannot be decomposed by a chemical change

**gas**
: state in which matter has neither definite volume nor shape

**heterogeneous mixture**
: combination of substances with a composition that varies from point to point

**homogeneous mixture**
: (also, solution) combination of substances with a composition that is uniform throughout

**liquid**
: state of matter that has a definite volume but indefinite shape

**law of conservation of matter**
: when matter converts from one type to another or changes form, there is no detectable change in the total amount of matter present

**mass**
: fundamental property indicating amount of matter

**matter**
: anything that occupies space and has mass

**mixture**
: matter that can be separated into its components by physical means

**molecule**
: bonded collection of two or more atoms of the same or different elements

**plasma**
: gaseous state of matter containing a large number of electrically charged atoms and/or molecules

**pure substance**
: homogeneous substance that has a constant composition

**solid**
: state of matter that is rigid, has a definite shape, and has a fairly constant volume

**weight**
: force that gravity exerts on an object

---

This page titled 8.02: Phases and Classification of Matter is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 11.28: Particle Accelerators and Detectors (Conceptual Questions Q11.4.2) by OpenStax is licensed CC BY 4.0.

# 8.03: Physical and Chemical Properties

## Learning Objectives

- Describe the difference between physical and chemical properties or changed.
- Identify a property or transformation as either physical or chemical using symbolic, particulate, or macroscopic representations.
- Identify the properties of matter as extensive or intensive.
- Recognize and describe the parts of the NFPA hazard diamond.

The characteristics that enable us to distinguish one substance from another are called properties. A physical property is a characteristic of matter that is not associated with a change in its chemical composition. Familiar examples of physical properties include density, color, hardness, melting and boiling points, and electrical conductivity. We can observe some physical properties, such as density and color, without changing the physical state of the matter observed. Other physical properties, such as the melting temperature of iron or the freezing temperature of water, can only be observed as matter undergoes a physical change. A physical change is a change in the state or properties of matter without any accompanying change in its chemical composition (the identities of the substances contained in the matter). We observe a physical change when wax melts, when sugar dissolves in coffee, and when steam condenses into liquid water (Figure 8.3.1). Other examples of physical changes include magnetizing and demagnetizing metals (as is done with common antitheft security tags) and grinding solids into powders (which can sometimes yield noticeable changes in color). In each of these examples, there is a change in the physical state, form, or properties of the substance, but no change in its chemical composition.

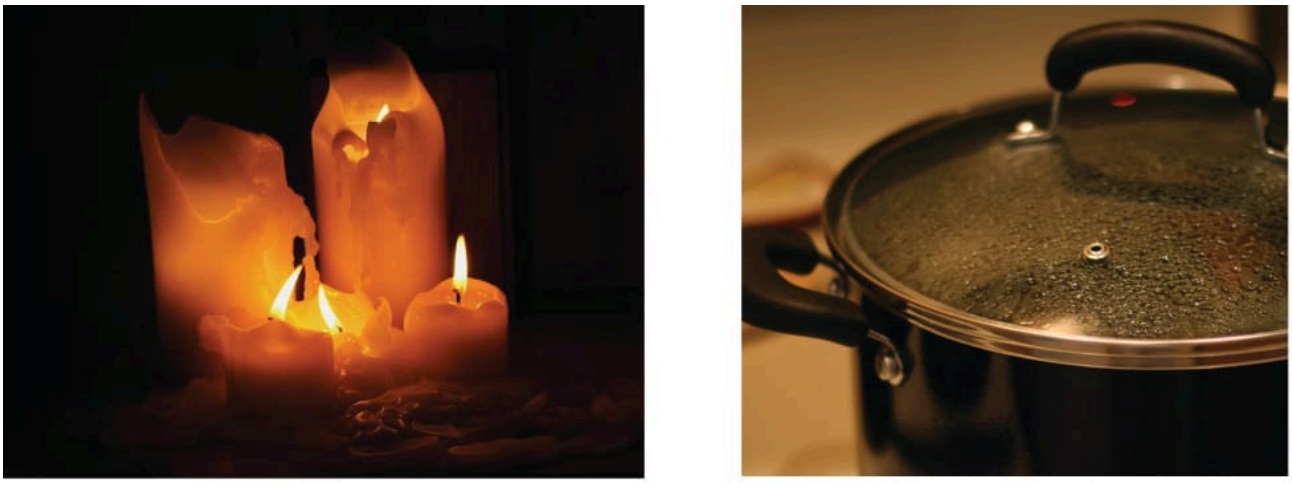



Figure 8.3.1: (a) Wax undergoes a physical change when solid wax is heated and forms liquid wax. (b) Steam condensing inside a cooking pot is a physical change, as water vapor is changed into liquid water. (credit a: modification of work by “95jb14”/Wikimedia Commons; credit b: modification of work by “mjneuby”/Flickr).

The change of one type of matter into another type (or the inability to change) is a chemical property. Examples of chemical properties include flammability, toxicity, acidity, reactivity (many types), and heat of combustion. Iron, for example, combines with oxygen in the presence of water to form rust; chromium does not oxidize (Figure 8.3.2). Nitroglycerin is very dangerous because it explodes easily; neon poses almost no hazard because it is very unreactive.

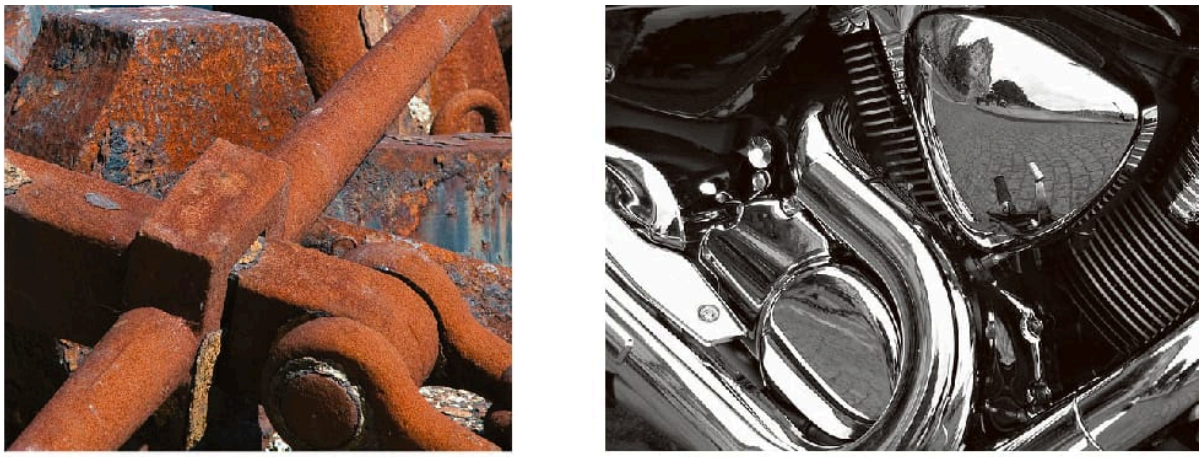



Figure 8.3.2: (a) One of the chemical properties of iron is that it rusts; (b) one of the chemical properties of chromium is that it does not. (credit a: modification of work by Tony Hisgett; credit b: modification of work by "Atoma"/Wikimedia Commons) Comparison of two metal equipments. The first image shows a rusted brown appearance and the following image shows a shiny metallic appearance.

To identify a chemical property, we look for a chemical change. A chemical change always produces one or more types of matter that differ from the matter present before the change. The formation of rust is a chemical change because rust is a different kind of matter than the iron, oxygen, and water present before the rust formed. The explosion of nitroglycerin is a chemical change because the gases produced are very different kinds of matter from the original substance. Other examples of chemical changes include reactions that are performed in a lab (such as copper reacting with nitric acid), all forms of combustion (burning), and food being cooked, digested, or rotting (Figure 8.3.3).

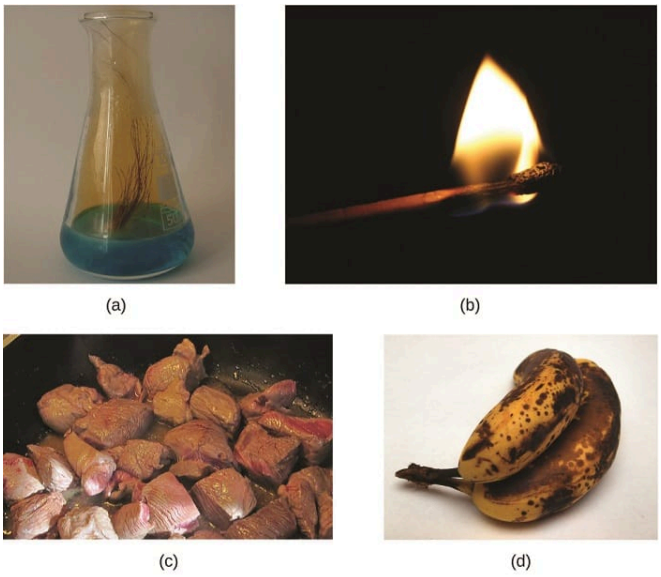



Figure 8.3.3: (a) Copper and nitric acid undergo a chemical change to form copper nitrate and brown, gaseous nitrogen dioxide. (b) During the combustion of a match, cellulose in the match and oxygen from the air undergo a chemical change to form carbon dioxide and water vapor. (c) Cooking red meat causes a number of chemical changes, including the oxidation of iron in myoglobin that results in the familiar red-to-brown color change. (d) A banana turning brown is a chemical change as new, darker (and less tasty) substances form. (credit b: modification of work by Jeff Turner; credit c: modification of work by Gloria Cabada-Leman; credit d: modification of work by Roberto Verzo) Four pictures. A. a flask with copper wires placed in a blue solution and also the presence of a gas. B. A lighted match. C. meat cubes being cooked on a pan. D. Bananas forming black spots on its peel.

Properties of matter fall into one of two categories. If the property depends on the amount of matter present, it is an extensive property. The mass and volume of a substance are examples of extensive properties; for instance, a gallon of milk has a larger mass and volume than a cup of milk. The value of an extensive property is directly proportional to the amount of matter in question. If the property of a sample of matter does not depend on the amount of matter present, it is an intensive property.

Temperature is an example of an intensive property. If the gallon and cup of milk are each at 20 °C (room temperature), when they are combined, the temperature remains at 20 °C. As another example, consider the distinct but related properties of heat and temperature. A drop of hot cooking oil spattered on your arm causes brief, minor discomfort, whereas a pot of hot oil yields severe burns. Both the drop and the pot of oil are at the same temperature (an intensive property), but the pot clearly contains much more heat (extensive property).

## Hazard Diamond

You may have seen the symbol shown in Figure 8.3.4 on containers of chemicals in a laboratory or workplace. Sometimes called a “fire diamond” or “hazard diamond,” this chemical hazard diamond provides valuable information that briefly summarizes the various dangers of which to be aware when working with a particular substance.

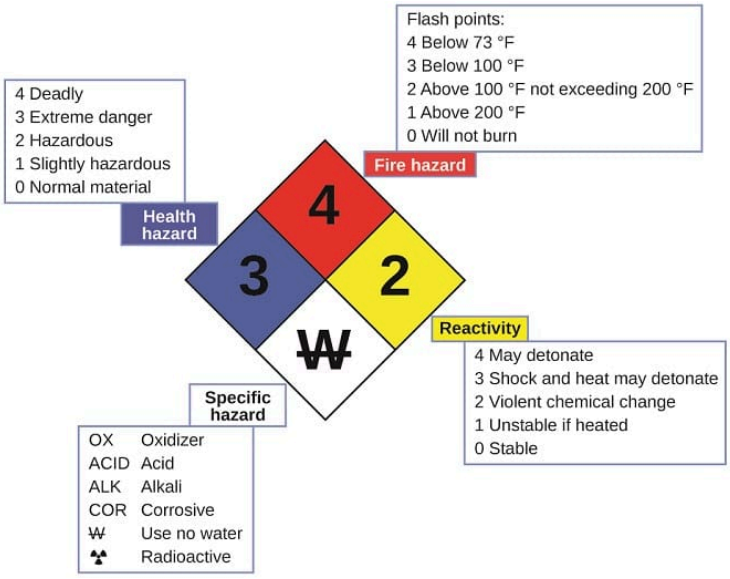



Figure 8.3.4: The National Fire Protection Agency (NFPA) hazard diamond summarizes the major hazards of a chemical substance.

The National Fire Protection Agency (NFPA) 704 Hazard Identification System was developed by NFPA to provide safety information about certain substances. The system details flammability, reactivity, health, and other hazards. Within the overall diamond symbol, the top (red) diamond specifies the level of fire hazard (temperature range for flash point). The blue (left) diamond indicates the level of health hazard. The yellow (right) diamond describes reactivity hazards, such as how readily the substance will undergo detonation or a violent chemical change. The white (bottom) diamond points out special hazards, such as if it is an oxidizer (which allows the substance to burn in the absence of air/oxygen), undergoes an unusual or dangerous reaction with water, is corrosive, acidic, alkaline, a biological hazard, radioactive, and so on. Each hazard is rated on a scale from 0 to 4, with 0 being no hazard and 4 being extremely hazardous.

While many elements differ dramatically in their chemical and physical properties, some elements have similar properties. We can identify sets of elements that exhibit common behaviors. For example, many elements conduct heat and electricity well, whereas others are poor conductors. These properties can be used to sort the elements into three classes: metals (elements that conduct well), nonmetals (elements that conduct poorly), and metalloids (elements that have properties of both metals and nonmetals).

The periodic table is a table of elements that places elements with similar properties close together (Figure 8.3.5). You will learn more about the periodic table as you continue your study of chemistry.

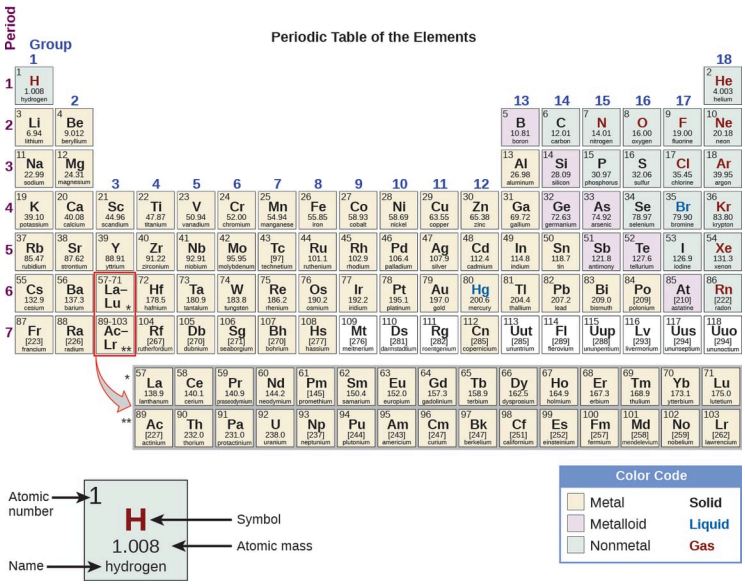



Figure 8.3.5: The periodic table shows how elements may be grouped according to certain similar properties. Note the background color denotes whether an element is a metal, metalloid, or nonmetal, whereas the element symbol color indicates whether it is a solid, liquid, or gas.

## Summary

All substances have distinct physical and chemical properties, and may undergo physical or chemical changes. Physical properties, such as hardness and boiling point, and physical changes, such as melting or freezing, do not involve a change in the composition of matter. Chemical properties, such flammability and acidity, and chemical changes, such as rusting, involve production of matter that differs from that present beforehand.

Measurable properties fall into one of two categories. Extensive properties depend on the amount of matter present, for example, the mass of gold. Intensive properties do not depend on the amount of matter present, for example, the density of gold. Heat is an example of an extensive property, and temperature is an example of an intensive property.

## Glossary

**chemical change**
: change producing a different kind of matter from the original kind of matter

**chemical property**
: behavior that is related to the change of one kind of matter into another kind of matter

**extensive property**
: property of a substance that depends on the amount of the substance

**intensive property**
: property of a substance that is independent of the amount of the substance

**physical change**
: change in the state or properties of matter that does not involve a change in its chemical composition

**physical property**
: characteristic of matter that is not associated with any change in its chemical composition

---

This page titled 8.03: Physical and Chemical Properties is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- 11.29: Particle Accelerators and Detectors (Conceptual Questions Q11.4.3) by OpenStax is licensed CC BY 4.0.

# 9: MODULE 7-ATOMIC STRUCTURES AND PERIODICITY


9.1: Introduction to Atoms and Elements

9.2: Modern Understandings of Elements

**9.3: The Development of Atomic Theory**

9.3.1: Law of Conservation of Mass

9.3.2: Law of Definite Proportions

9.3.3: Law of Multiple Proportions

9.3.4: Dalton's Atomic Theory

9.4: Prelude to Chemical Bonding and Molecular Geometry

9.5: Ionic Bonding

9.6: Ionic Bonding (1)

9.7: Covalent Bonds and Molecules

9.8: Covalent Bonding

9.9: Prelude to Liquids and Solids

9.10: Intermolecular Forces

9.11: Intermolecular Forces (1)



This page titled 9: Module 7-Atomic Structures and Periodicity is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Arunava Roy.

# 9.1: Introduction to Atoms and Elements

In this chapter we will be exploring the concepts of atoms and elements. These two words make most people think of chemistry. They are related, but are not exactly the same thing. As we read through this chapter we will see the nuances of these two words. To introduce these ideas, we will first consider how they were thought about before people began to investigate them scientifically.



Figure 9.1.1 (Credit: Justin Norris (Flickr:JustinMN); Source: flickr)(opens in new window); License: CC by 2.0(opens in new window))

## The Philosophers' Views on Atoms

People enjoy getting together to discuss things, whether it is how your favorite sports team is doing, what the best new movie is, the current politics, or any number of other topics. Often the question is raised about who is right and who is wrong. If the football game is to be played this coming weekend, all we can do is offer opinions as to its outcome. The game has not been played yet, so we do not know who will actually win.

The ancient Greek philosophers did a lot of discussing, with part of their conversations concerning the physical world and its composition. There were different opinions about what made up matter. Some felt one thing was true while others believed another set of ideas. Since these scholars did not have laboratories and had not developed the idea of the experiment, they were left to debate. Whoever could offer the best argument was considered right. However, often the best argument had little to do with reality.

One of the on-going debates had to do with sand. The question posed was: into how small of pieces can you divide a grain of sand? The prevailing thought at the time, pushed by Aristotle, was that the grain of sand could be divided indefinitely, that you could always get a smaller particle by dividing a larger one and there was no limit to how small the resulting particle could be.



Figure 9.1.2: Into how small of pieces can you divide a grain of sand? (Credit: User:(WT-shared) Shoestring/WTS.Wikivoyage; Source: wikimedia; License: Public Domain)

Since Aristotle was such an influential philosopher, very few people disagreed with him. However, there were some philosophers who believed that there was a limit to how small a grain of sand could be divided. One of these philosophers was Democritus (~460-370 B.C.E.), often referred to as the "laughing philosopher" because of his emphasis on cheerfulness. He taught that there were substances called **atoms** and that these atoms made up all material things. The atoms were unchangeable, indestructible, and always existed.

The word "**atom**" comes from the Greek $\alpha\tau o\mu o\sigma$ and means "indivisible". The atomists of the time (Democritus being one of the leading atomists) believed there were two realities that made up the physical world: atoms and void. There were an infinite number of atoms, but different types of atoms had different sizes and shapes. The void was the empty space in which the atoms moved and collided with one another. When these atoms collided with one another, they might repel each other or they might connect in clusters, held together by tiny hooks and barbs on the surfaces of the atoms.

Aristotle disagreed with Democritus and offered his own idea of the composition of matter. According to Aristotle, everything was composed of four elements: earth, air, fire, and water. Democritus' theory better explained things, but Aristotle was more influential, so his ideas prevailed. It took almost two thousand years before scientists came around to seeing the atom as Democritus did.

It is very interesting that Democritus had the basic idea of atoms, even though he had no experimental evidence to support his thinking. We now know more about how atoms hold together in "clusters" (compounds), but the basic concept existed over two thousand years ago. We also know that atoms can be further subdivided, but there is still a lower limit to how small we can break up that grain of sand.

## The Philosophers' Views on Elements

Greek philosophers proposed many ideas for what things were made of as well. Some proposed 5 elements, with distinct shapes: **octahedra**, **tetrahedra**, **cubes**, etc. For a long while, the four element model (earth, air, fire, water) was popular. This model, which Plato and Aristotle also used, suggested that all matter was composed of these four elements in different ratios. For instance, when wood is burned, you get smoke (air), ash (earth), pitch (a viscous liquid, here identified with water), and fire, so wood is made of all these things. Other models suggested everything was made of just water, or just air.

Despite the vast array of models for the elements that philosophers came up with, none of them got close to the current understanding of elements. It would take the work of science, not mere chance in a game of conjectures, to find that.

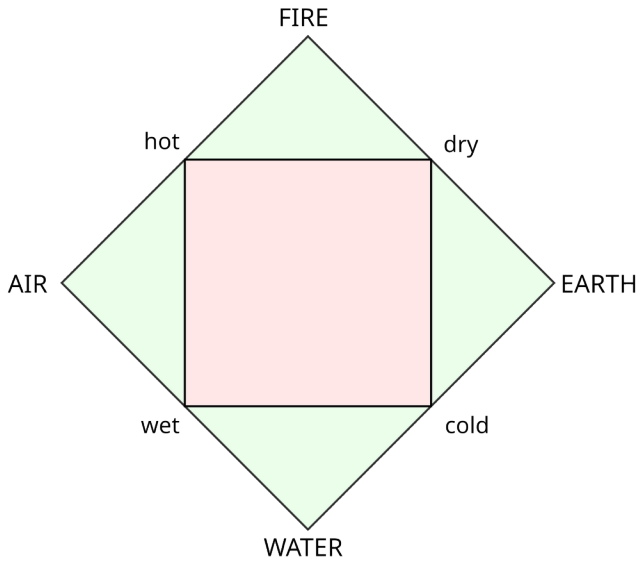



Figure 9.1.3: representation of an ancient idea of the elements. By en:User:Heron - Transwikied from en:. Originally uploaded by en:User:Heron. Recompressed with OptiPNG by Michael. Converted to SVG by tiZom, Public Domain, wikimedia

## Contributors and Attributions

- Emily V Eames(opens in new window) (City College of San Francisco) wrote language adopted for the discussion of the ancient view of elements in this text.

This page titled 9.1: Introduction to Atoms and Elements is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 3.32: Instantaneous Velocity and Speed (Q3.2.7) by OpenStax is licensed CC BY 4.0.
- 10.31: Nuclear Reactions (Conceptual Questions Q10.4.1) by OpenStax is licensed CC BY 4.0.

# 9.2: Modern Understandings of Elements

**Learning Objectives**

- Know the modern definition of an element.

## The Modern Definition of an Element

Early chemical experimenters toyed with the many ideas of the ancients as possibilities for what was an element. Chemistry was a bit later than other sciences to experience a true scientific revolution, because much of the practice of chemical experimentation occurred in the field of alchemy, where there was much secrecy and linking of experimental practices to metaphysical interpretations.

It was Antoine Lavoisier (1743-1794) who finally came up with the modern concept of an element as something which can not be broken down into anything simpler by a chemical reaction. Lavoisier also proposed the Law of Conservation of Mass (which we shall discuss shortly). He also proposed many other important ideas in chemistry during his short time of experimentation. His work in defining elements and discovering this first law of chemistry paved the way for the work of Dalton and the others. He certainly had the most impact of any chemist of the 18th century, and in fact some consider him to be the first true chemist.

The idea that an element was something that could not be broken down any further was not necessarily new. What was new in the work of Lavoisier was that these substances were identified empirically instead of philosophically. Lavoisier proposed about 30 substances as elements which could not be broken down any further. Later experiments showed that some of these could be broken down further and others were not forms of matter at all. However, it was a huge step to list both lead and gold as elements which could not be broken down further when transmuting lead into gold had long been the goal of alchemists. The reliance on experimentation instead of reasoning was again a key to understanding the natural world.

Watch this video to learn more about the work of Lavoisier and some of the other earlier chemists:

**YouTube Video**

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Glossary

**element**
: something which can't be broken down into anything simpler by a chemical reaction.

This page titled 9.2: Modern Understandings of Elements is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 10.31: Nuclear Reactions (Conceptual Questions Q10.4.1) by OpenStax is licensed CC BY 4.0.
- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 9.3: THE DEVELOPMENT OF ATOMIC THEORY


9.3.1: Law of Conservation of Mass

9.3.2: Law of Definite Proportions

9.3.3: Law of Multiple Proportions

9.3.4: Dalton's Atomic Theory



This page titled 9.3: The Development of Atomic Theory is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

# 9.3.1: Law of Conservation of Mass

**Learning Objectives**

- Recognize the law of conservation of mass.



Figure 9.3.1.1 (Credit: Douglas Heriot; Source:flickr; License: CC by 2.0(opens in new window))

## Have you ever lost a screw?

The following situation happens all too often. You have taken apart a piece of equipment to clean it up. When you put the equipment back together, somehow you have an extra screw or two. Or you find out that a screw is missing that was a part of the original equipment. In either case, you know something is wrong. You expect to end up with the same amount of material that you started with, not with more or less than what you had originally.

## Law of Conservation of Mass

By the late 1700s, chemists accepted the definition of an element as a substance that cannot be broken down into a simpler substance by ordinary chemical means. It was also clear that elements combine with one another to form more complex substances called compounds. The chemical and physical properties of these compounds are different than the properties of the elements from which they were formed. There were questions about the details of these processes.

In the 1790s, a greater emphasis began to be placed on the quantitative analysis of chemical reactions. Accurate and reproducible measurements of the masses of reacting elements and the compounds they form led to the formulation of several basic **laws**. One of these is called the law of **conservation of mass**, which states that during a chemical reaction, the total mass of the **products** must be equal to the total mass of the **reactants**. In other words, mass cannot be created or destroyed during a chemical reaction, but is always conserved. This is another of the rare conservation laws we are being introduced to.

**LAW OF CONSERVATION OF MASS**

During a chemical reaction, the total mass of all products must be equal to the total mass of all reactants.

As an example, consider the reaction between silver nitrate and sodium chloride. These two compounds will dissolve in water to

form silver chloride and sodium nitrate. The silver chloride does not dissolve in water, so it forms a solid that we can filter off. When we evaporate the water, we can recover the sodium nitrate formed. If we react 58.5 grams of sodium chloride with 169.9 grams of silver nitrate, we start with 228.4 grams of materials. After the reaction is complete and the materials separated, we find that we have formed 143.4 grams of silver chloride and 85.0 grams of sodium nitrate, giving us a total mass of 228.4 grams for the products. So, the total mass of reactants equals the total mass of products, a proof of the law of conservation of mass.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- The law of conservation of mass states that, during a chemical reaction, the total mass of the products must be equal to the total mass of the reactants.

## Glossary

**reactants**
substances which react during a chemical reaction.

**products**
substances which are formed during a chemical reaction.

**law of conservation of mass**
mass cannot be created or destroyed during a chemical reaction, but is always conserved.

This page titled 9.3.1: Law of Conservation of Mass is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 9.3.2: Law of Definite Proportions

## Learning Objectives

- Recognize the Law of Definite Proportions

The discovery that mass was always conserved in chemical reactions was soon followed by the **law of definite proportions**, which states that a given chemical compound always contains the same elements in the exact same proportions by mass. As an example, any sample of pure water contains $11.19\%$ hydrogen and $88.81\%$ oxygen by mass. It does not matter where the sample of water came from or how it was prepared. Its composition, like that of every other compound, is fixed.

## LAW OF DEFINITE PROPORTIONS

Any sample of a chemical compound will always contain the same elements in the exact same proportion by mass.



Figure $9.3.2.1$: Water. (Credit: Jenn Durfey; Source: Flickr Photos, Water Faucet(opens in new window) [www.flickr.com]; License: CC by 2.0(opens in new window))

Another example is carbon dioxide. This gas is produced from a variety of reactions, often by the burning of materials. The structure of the gas consists of one atom of carbon and two atoms of oxygen. Carbon dioxide production is of interest in many areas, from the amount we breathe to the amount of the gas produced by burning wood or fossil fuels. By knowing the exact composition of carbon dioxide, we can make predictions as to the effects of different chemical processes.

cK-12 LICENSED UNDER CK-12 License
© CK-12 Foundation • Visit us at ck12.org



Figure 9.3.2.2: Carbon dioxide is produced during the burning process. (Credit: Courtesy of Patsy Lynch, FEMA; Source: Commons Wikimedia, FEMA Photograph by Patsy Lynch taken on 04-17-2006 in Missouri(opens in new window) [commons.wikimedia.org]; License: Public Domain)

## Section Summary

- The law of definite proportions states that a given chemical compound always contains the same elements in the exact same proportions by mass.

This page titled 9.3.2: Law of Definite Proportions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 9.11: Molecular Spectra (Conceptual Questions Q9.2.2) by OpenStax is licensed CC BY 4.0.

# 9.3.3: Law of Multiple Proportions

**Learning Objectives**

- Recognize the Law of Multiple Proportions

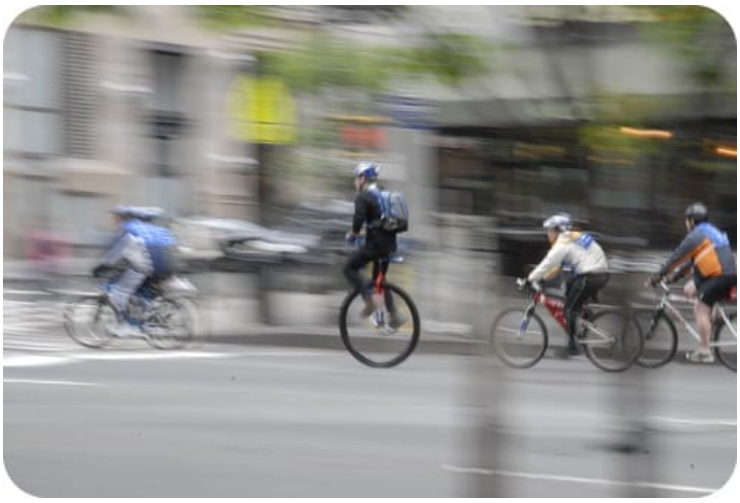


Figure $9.3.3.1$ (Credit: Flickr: Sutibu; Source: Commons WIkimedia, Five Boro Bike Tour Unicycle(opens in new window) [commons.wikimedia.org]; License: CC by 2.0(opens in new window))

## What are the similarities and differences between a unicycle and a bicycle?

Just from the words themselves, the astute Latin-speaking scholar can tell that, whatever it is made of, the unicycle has one of them (*uni* = "one") and the bicycle has two (*bi* = "two"). From pictures, we get additional information that helps us tell the two apart. The unicycle has one wheel and the bicycle has two. In particular, they are made up of the same materials, and the only significant difference is the number of wheels on the two vehicles. Now—how many wheels are on a tricycle?

## Law of Multiple Proportions

Once the idea that elements combined in definite proportions to form compounds was established, experiments also began to demonstrate that the same pairs of certain elements could combine to form more than one compound. Consider the elements carbon and oxygen. Combined in one way, they form the familiar compound carbon dioxide. In every sample of carbon dioxide, there are $32.0\ \text{g}$ of oxygen present for every $12.0\ \text{g}$ of carbon. By dividing $32.0$ by $12.0$, this simplifies to a mass ratio of oxygen to carbon of 2.66 to 1. There is another compound that forms from the combination of carbon and oxygen called carbon monoxide. Every sample of carbon monoxide contains $16.0\ \text{g}$ of oxygen for every $12.0\ \text{g}$ of carbon. This is a mass ratio of oxygen to carbon of 1.33 to 1. In the carbon dioxide, there is exactly twice as much oxygen present as there is in the carbon monoxide. This example illustrates the **law of multiple proportions**: whenever the same two elements form more than one compound, the different masses of one element that combine with the same mass of the other element are in the ratio of small whole numbers.

### LAW OF MULTIPLE PROPORTIONS

Whenever the same two elements form more than one compound, the different masses of one element that combine with the same mass of the other element are in the ratio of small whole numbers.

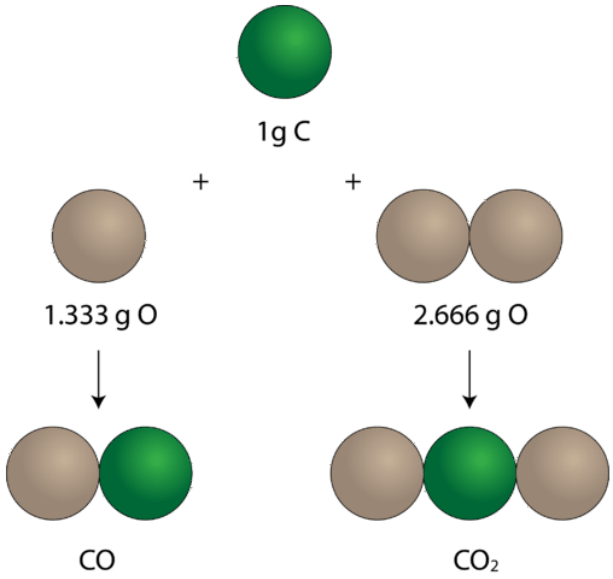



Figure $9.3.3.2$: Carbon can form two different compounds with oxygen. (Credit: Christopher Auyeung; Source: CK-12 Foundation; License: CC BY-NC-SA 3.0(opens in new window))

In carbon monoxide, on the left, there is $1.333\ \text{g}$ of oxygen for every $1\ \text{g}$ of carbon. In carbon dioxide, on the right, there is $2.666\ \text{g}$ of oxygen for every gram of carbon. So the ratio of oxygen in the two compounds is 1:2, a small whole number ratio.

The difference between carbon monoxide and carbon dioxide is significant. Carbon monoxide is a deadly gas, formed from the incomplete combustion of some carbon-containing materials (such as wood and gasoline). This compound will attach to hemoglobin in the red blood cells and block the binding of oxygen to those cells. If oxygen does not bind, it cannot be carried to the cells of the body where it is needed, and death can occur. Carbon dioxide, on the other hand, is not toxic like carbon monoxide is. However, it can displace oxygen in systems since it is heavier. Carbon dioxide fire extinguishers cut off the flow of oxygen in a fire, putting out the fire.

## Section Summary

- The law of multiple proportions states that whenever the same two elements form more than one compound, the different masses of one element that combine with the same mass of the other element are in the ratio of small whole numbers.

This page titled 9.3.3: Law of Multiple Proportions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 9.12: Molecular Spectra (Conceptual Questions Q9.2.3) by OpenStax is licensed CC BY 4.0.

# 9.3.4: Dalton's Atomic Theory

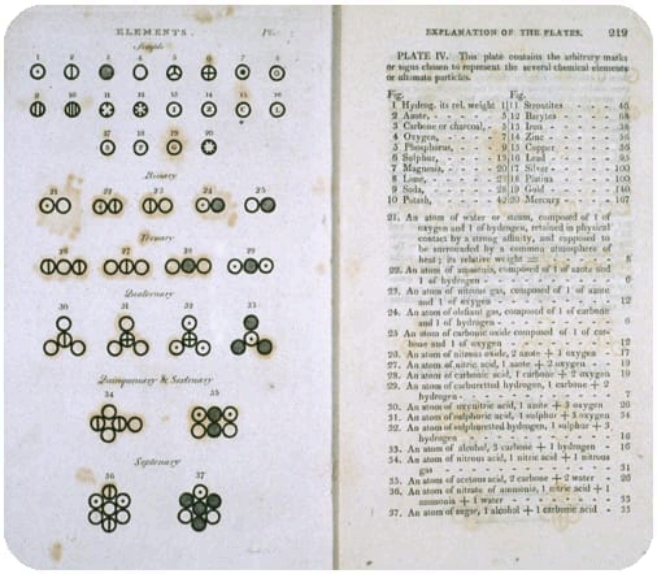
## Learning Objectives

- Recognize the 5 tenants of Dalton's Atomic Theory.

Though we have illustrated our current knowledge of the behavior of atoms in our discussion of the laws of definite proportions and multiple proportions, this information was not available to Dalton when he formulated atomic theory. In fact, it was the information in the laws themselves which naturally led to the conclusion that atoms must exist. Using only the three laws we have introduced in this section, Dalton devised the **atomic theory** as an explanation for the observations which led to these laws. It was a scientific way of suggesting that atoms existed, as opposed to the non-scientific debates that the Greek philosophers engaged in. The general tenets of this theory are as follows:

- All matter is composed of extremely small particles called atoms.
- Atoms of a given element are identical in size, mass, and other properties. Atoms of different elements differ in size, mass, and other properties.
- Atoms cannot be subdivided, created, or destroyed.
- Atoms of different elements can combine in simple whole number ratios to form chemical compounds.
- In chemical reactions, atoms are combined, separated, or rearranged.

Dalton's atomic theory has been largely accepted by the scientific community, with the exception of three changes. We now know that (1) an atom can be further subdivided, (2) all atoms of an element are not identical in mass, and (3) using nuclear fission and fusion techniques, we can create or destroy atoms by changing them into other atoms. We will explore each of these ideas later in this text. However, for any chemical reaction the ideas held within the atomic theory would correctly predict the results of that reaction.


ELEMENTS.

EXPLANATION OF THE PLATES. 219

PLATE IV. This plate contains the arbitrary marks or signs chosen to represent the several chemical elements or ultimate particles.

Figure 9.3.4.3: Dalton's symbols. (Credit: John Dalton; Source: wikimedia; License: Public Domain)

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Summary

- The general tenets of Dalton's atomic theory are:
  - All matter is composed of extremely small particles called atoms
  - Atoms cannot be subdivided, created, or destroyed.
  - Atoms of a given element are identical in size, mass, and other properties. Atoms of different elements differ in size, mass, and other properties.
  - Atoms of different elements can combine in whole number ratios to form chemical compounds;
  - In chemical reaction, atoms are rearranged, combined, or separated.

This page titled 9.3.4: Dalton's Atomic Theory is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 9.14: Molecular Spectra (Problems Q9.2.2) by OpenStax is licensed CC BY 4.0.

# 9.4: Prelude to Chemical Bonding and Molecular Geometry

It has long been known that pure carbon occurs in different forms (allotropes) including graphite and diamonds. But it was not until 1985 that a new form of carbon was recognized: buckminsterfullerene, commonly known as a “buckyball.” This molecule was named after the architect and inventor R. Buckminster Fuller (1895–1983), whose signature architectural design was the geodesic dome, characterized by a lattice shell structure supporting a spherical surface. Experimental evidence revealed the formula, $C_{60}$, and then scientists determined how 60 carbon atoms could form one symmetric, stable molecule. They were guided by bonding theory—the topic of this chapter—which explains how individual atoms connect to form more complex structures.



Figure 9.4.1: Nicknamed “buckyballs,” buckminsterfullerene molecules ($C_{60}$) contain only carbon atoms. Here they are shown in a ball-and-stick model (left). These molecules have single and double carbon-carbon bonds arranged to form a geometric framework of hexagons and pentagons, similar to the pattern on a soccer ball (center). This unconventional molecular structure is named after architect R. Buckminster Fuller, whose innovative designs combined simple geometric shapes to create large, strong structures such as this weather radar dome near Tucson, Arizona (right). (credit middle: modification of work by “Petey21”/Wikimedia Commons; credit right: modification of work by Bill Morrow). Three figures are shown. The left figure is a many-sides spherical ball composed of hexagonal rings which have carbon atoms at each corner. The center picture shows a soccer ball. The right picture shown as water tower with sides shaped like hexagonal rings.

This page titled 9.4: Prelude to Chemical Bonding and Molecular Geometry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 9.5: Ionic Bonding

**Learning Objectives**

- Predict the formulas of the ionic compounds that form according to the Octet Rule.
- Describe the structure of ionic compounds.



Figure $9.5.1$ (Credit: Leon Brooks; Source: Commons Wikimedia, Waneroo Beach(opens in new window) [commons.wikimedia.org]; License: Public Domain)

## Does the sea really have salt in it?

We can get common table salt from several sources. It can be mined in the solid form in salt mines, or found as a solid in deposits. We can also get salt from the ocean, but it really does not exist as a salt when in solution. The sodium ions and chloride ions are dissolved, but not combined into a structure until all the water is removed.

Most of the rocks and minerals that make up the Earth's crust are composed of positive and negative ions held together by ionic bonding. An ionic compound is an electrically neutral compound consisting of positive and negative ions. You are very familiar with some ionic compounds, such as sodium chloride $(\mathrm{NaCl})$. A sodium chloride crystal consists of equal numbers of positive sodium ions $(\mathrm{Na}^+)$ and negative chloride ions $(\mathrm{Cl}^-)$.

## Ionic Bonds

Oppositely charged particles attract each other. This attractive force is often referred to as an **electrostatic force**. An **ionic bond** is the electrostatic force that holds ions together in an **ionic compound**. The strength of the ionic bond is directly dependent upon the quantity of the charges and inversely dependent on the distance between the charged particles. A cation with a $2+$ charge will make a stronger ionic bond than a cation with a $1+$ charge. A larger ion makes a weaker ionic bond because of the greater distance between its electrons and the nucleus of the oppositely charged ion.

## Electron Dot Diagrams

We will use sodium chloride as an example to demonstrate the nature of the ionic bond and how it forms. As you know, sodium is a metal and loses its one valence electron to become a cation. Chlorine is a nonmetal and gains one electron in becoming an anion. Both achieve a noble-gas electron configuration. However, electrons cannot be simply "lost" to nowhere in particular. A more accurate way to describe what is happening is that a single electron is transferred from the sodium atom to the chlorine

atom, as shown below.

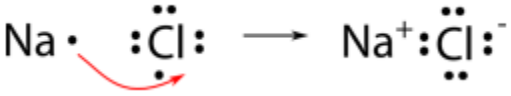



Figure 9.5.2 (Credit: CK-12 Foundation; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

The ionic bond is the attraction of the $\mathrm{Na}^{+}$ ion for the $\mathrm{Cl}^{-}$ ion. It is conventional to show the cation without dots around the symbol to emphasize that the original energy level that contained the valence electron is now empty. The anion is now shown with a complete octet of electrons.

For a compound such as magnesium chloride, it is not quite as simple. Because magnesium has two valence electrons, it needs to lose both to achieve the noble-gas configuration. Therefore, two chlorine atoms will be needed.

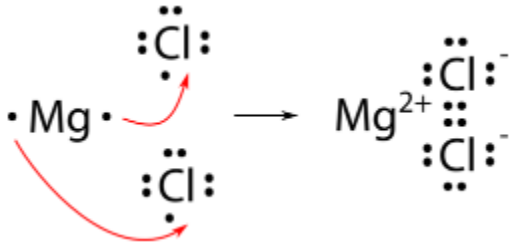



Figure 9.5.3 (Credit: CK-12 Foundation; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

The final formula for magnesium chloride is $\mathrm{MgCl_2}$.

## Ionic Crystal Structure

Electron dot diagrams show the nature of the electron transfer that takes place between metal and nonmetal atoms. However, ionic compounds do not exist as discrete molecules, as the dot diagrams may suggest. In order to minimize the potential energy of the system, ionic compounds take on the form of an extended three-dimensional array of alternating cations and anions. This maximizes the attractive forces between the oppositely charged ions. The figure below shows two different ways of representing the ionic crystal lattice. A ball and stick model makes it easier to see how individual ions are oriented with respect to one another. A space filling diagram is a more accurate representation of how the ions pack together in the **crystal**.

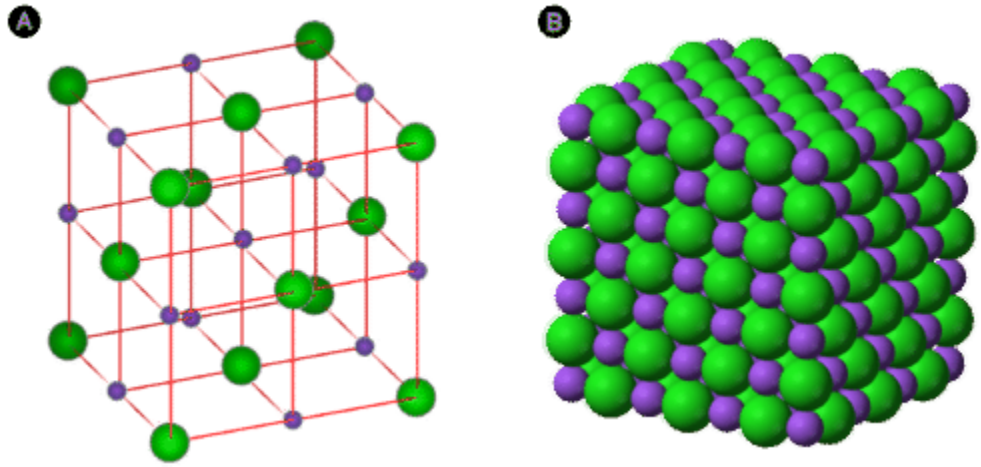



Figure 9.5.2: Two models of a sodium chloride crystal are shown. The purple spheres represent the $\mathrm{Na}^{+}$ ions, while the green spheres represent the $\mathrm{Cl}^{-}$ ions. (A) In an expanded view, the distances between ions are exaggerated, more easily showing the coordination numbers of each ion. (B) In a space filling model, the electron clouds of the ions are in contact with each other. (Credit: (A) Eloy; (B) Ben Mills (Wikimedia: Benjah-bmm27); Source: (A) wikimedia; (B) wikimedia; License: Public Domain)

Naturally occurring sodium chloride **(halite)** does not look at first glance like the neat diagrams shown above. It is only when we

use modern techniques to analyze the crystal structure at the atomic level that we can see the true regularity of the organized ions.



Figure 9.5.3: Halite crystals. (Credit: Ingo Wölbern (Wikimedia: Iwoelbern); Source: wikimedia [commons.wikimedia.org]; License: Public Domain)

## Section Summary

- An ionic compound contains positive and negative ions.
- An ionic bond is electrostatic in nature.
- Electron dot diagrams can be used to illustrate electron movements and ion formation.
- Ionic compounds take on the form of extended three-dimensional arrays of cations and anions.
- The arrangement maximizes the attractive force between oppositely-charged ions.

## Glossary

**ionic bond**
: the electrostatic force that holds ions together in an ionic compound.

**ionic compound**
: a compound formed by the attraction between ions.

**crystal**
: a structure of matter that is an extended three-dimensional array of repeating patterns of particles that make up that form of matter.

This page titled 9.5: Ionic Bonding is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 9.6: Ionic Bonding (1)

## Learning Objectives

- Explain the formation of cations, anions, and ionic compounds
- Predict the charge of common metallic and nonmetallic elements, and write their electron configurations

As you have learned, ions are atoms or molecules bearing an electrical charge. A cation (a positive ion) forms when a neutral atom loses one or more electrons from its valence shell, and an anion (a negative ion) forms when a neutral atom gains one or more electrons in its valence shell.

Compounds composed of ions are called ionic compounds (or salts), and their constituent ions are held together by ionic bonds: electrostatic forces of attraction between oppositely charged cations and anions. The properties of ionic compounds shed some light on the nature of ionic bonds. Ionic solids exhibit a crystalline structure and tend to be rigid and brittle; they also tend to have high melting and boiling points, which suggests that ionic bonds are very strong. Ionic solids are also poor conductors of electricity for the same reason—the strength of ionic bonds prevents ions from moving freely in the solid state. Most ionic solids, however, dissolve readily in water. Once dissolved or melted, ionic compounds are excellent conductors of electricity and heat because the ions can move about freely.

Neutral atoms and their associated ions have very different physical and chemical properties. Sodium *atoms* form sodium metal, a soft, silvery-white metal that burns vigorously in air and reacts explosively with water. Chlorine *atoms* form chlorine gas, $Cl_2$, a yellow-green gas that is extremely corrosive to most metals and very poisonous to animals and plants. The vigorous reaction between the elements sodium and chlorine forms the white, crystalline compound sodium chloride, common table salt, which contains sodium *cations* and chloride *anions* (Figure Figure $9.6.1$). The compound composed of these ions exhibits properties entirely different from the properties of the elements sodium and chlorine. Chlorine is poisonous, but sodium chloride is essential to life; sodium atoms react vigorously with water, but sodium chloride simply dissolves in water.



(a) (b) (c)

Figure $9.6.1$: (a) Sodium is a soft metal that must be stored in mineral oil to prevent reaction with air or water. (b) Chlorine is a pale yellow-green gas. (c) When combined, they form white crystals of sodium chloride (table salt). (credit a: modification of work by "Jurii"/Wikimedia Commons)

## The Formation of Ionic Compounds

Binary ionic compounds are composed of just two elements: a metal (which forms the cations) and a nonmetal (which forms the anions). For example, NaCl is a binary ionic compound. We can think about the formation of such compounds in terms of the periodic properties of the elements. Many metallic elements have relatively low ionization potentials and lose electrons easily. These elements lie to the left in a period or near the bottom of a group on the periodic table. Nonmetal atoms have relatively high electron affinities and thus readily gain electrons lost by metal atoms, thereby filling their valence shells. Nonmetallic elements are found in the upper-right corner of the periodic table.

As all substances must be electrically neutral, the total number of positive charges on the cations of an ionic compound must equal the total number of negative charges on its anions. The formula of an ionic compound represents the simplest ratio of the numbers of ions necessary to give identical numbers of positive and negative charges. For example, the formula for aluminum oxide, $Al_2O_3$, indicates that this ionic compound contains two aluminum cations, $Al^{3+}$, for every three oxide anions, $O^{2-}$ [thus, (2 × +3) + (3 × −2) = 0].

It is important to note, however, that the formula for an ionic compound does *not* represent the physical arrangement of its ions. It is incorrect to refer to a sodium chloride (NaCl) "molecule" because there is not a single ionic bond, per se, between any specific pair of sodium and chloride ions. The attractive forces between ions are isotropic—the same in all directions—meaning that any particular ion is equally attracted to all of the nearby ions of opposite charge. This results in the ions arranging themselves into a tightly bound, three-dimensional lattice structure. Sodium chloride, for example, consists of a regular arrangement of equal numbers of $Na^+$ cations and $Cl^-$ anions (Figure Figure 9.6.2).

Figure 9.6.2: The atoms in sodium chloride (common table salt) are arranged to (a) maximize opposite charges interacting. The smaller spheres represent sodium ions, the larger ones represent chloride ions. In the expanded view (b), the geometry can be seen more clearly. Note that each ion is "bonded" to all of the surrounding ions—six in this case. Two diagrams are shown and labeled "a" and "b." Diagram a shows a cube made up of twenty-seven alternating purple and green spheres. The purple spheres are smaller than the green spheres. Diagram b shows the same spheres, but this time, they are spread out and connected in three dimensions by white rods. The purple spheres are labeled "N superscript postive sign" while the green are labeled "C l superscript negative sign."

The strong electrostatic attraction between $Na^+$ and $Cl^-$ ions holds them tightly together in solid NaCl. It requires 769 kJ of energy to dissociate one mole of solid NaCl into separate gaseous $Na^+$ and $Cl^-$ ions:

$$\mathrm{NaCl}(s) \longrightarrow \mathrm{Na}^+(g) + \mathrm{Cl}^-(g) \quad \Delta H = 769\ \mathrm{kJ}$$

## Electronic Structures of Cations

When forming a cation, an atom of a main group element tends to lose all of its valence electrons, thus assuming the electronic structure of the noble gas that precedes it in the periodic table. For groups 1 (the alkali metals) and 2 (the alkaline earth metals), the group numbers are equal to the numbers of valence shell electrons and, consequently, to the charges of the cations formed from atoms of these elements when all valence shell electrons are removed. For example, calcium is a group 2 element whose neutral atoms have 20 electrons and a ground state electron configuration of $1s^22s^22p^63s^23p^64s^2$. When a Ca atom loses both of its valence electrons, the result is a cation with 18 electrons, a 2+ charge, and an electron configuration of $1s^22s^22p^63s^23p^6$. The $Ca^{2+}$ ion is therefore isoelectronic with the noble gas Ar.

For groups 12–17, the group numbers exceed the number of valence electrons by 10 (accounting for the possibility of full *d* subshells in atoms of elements in the fourth and greater periods). Thus, the charge of a cation formed by the loss of all valence electrons is equal to the group number minus 10. For example, aluminum (in group 13) forms 3+ ions ($Al^{3+}$).

Exceptions to the expected behavior involve elements toward the bottom of the groups. In addition to the expected ions $Tl^{3+}$, $Sn^{4+}$, $Pb^{4+}$, and $Bi^{5+}$, a partial loss of these atoms' valence shell electrons can also lead to the formation of $Tl^+$, $Sn^{2+}$, $Pb^{2+}$, and $Bi^{3+}$ ions. The formation of these 1+, 2+, and 3+ cations is ascribed to the inert pair effect, which reflects the relatively low energy of the valence s-electron pair for atoms of the heavy elements of groups 13, 14, and 15. Mercury (group 12) also exhibits an unexpected behavior: it forms a diatomic ion, $Hg_2^{2+}$ (an ion formed from two mercury atoms, with an Hg-Hg bond), in addition to the expected monatomic ion $Hg^{2+}$ (formed from only one mercury atom).

Transition and inner transition metal elements behave differently than main group elements. Most transition metal cations have 2+ or 3+ charges that result from the loss of their outermost *s* electron(s) first, sometimes followed by the loss of one or two *d* electrons from the next-to-outermost shell. For example, iron ($1s^22s^22p^63s^23p^63d^64s^2$) forms the ion $Fe^{2+}$ ($1s^22s^22p^63s^23p^63d^6$) by the loss of the 4*s* electrons and the ion $Fe^{3+}$ ($1s^22s^22p^63s^23p^63d^5$) by the loss of the 4*s* electrons and one of the 3*d* electrons. Although the *d* orbitals of the transition elements are—according to the Aufbau principle—the last to fill when building up electron configurations, the outermost *s* electrons are the first to be lost when these atoms ionize. When the inner transition metals form ions, they usually have a 3+ charge, resulting from the loss of their outermost *s* electrons and a *d* or *f* electron.

### Example 9.6.1: Determining the Electronic Structures of Cations

There are at least 14 elements categorized as “essential trace elements” for the human body. They are called “essential” because they are required for healthy bodily functions, “trace” because they are required only in small amounts, and “elements” in spite of the fact that they are really ions. Two of these essential trace elements, chromium and zinc, are required as $Cr^{3+}$ and $Zn^{2+}$. Write the electron configurations of these cations.

**Solution**

First, write the electron configuration for the neutral atoms:

- Zn: $[Ar]3d^{10}4s^2$
- Cr: $[Ar]3d^5 4s^1$

Next, remove electrons from the highest energy orbital. For the transition metals, electrons are removed from the s orbital first and then from the *d* orbital. For the *p*-block elements, electrons are removed from the *p* orbitals and then from the s orbital. Zinc is a member of group 12, so it should have a charge of 2+, and thus loses only the two electrons in its *s* orbital. Chromium is a transition element and should lose its *s* electrons and then its *d* electrons when forming a cation. Thus, we find the following electron configurations of the ions:

- $Zn^{2+}$: $[Ar]3d^{10}$
- $Cr^{3+}$: $[Ar]3d^3$

### Exercise 9.6.1

Potassium and magnesium are required in our diet. Write the electron configurations of the ions expected from these elements.

*Answer available at the end of this section.*

## Electronic Structures of Anions

Most monatomic anions form when a neutral nonmetal atom gains enough electrons to completely fill its outer s and *p* orbitals, thereby reaching the electron configuration of the next noble gas. Thus, it is simple to determine the charge on such a negative ion: The charge is equal to the number of electrons that must be gained to fill the s and *p* orbitals of the parent atom. Oxygen, for example, has the electron configuration $1s^2 2s^2 2p^4$, whereas the oxygen anion has the electron configuration of the noble gas neon (Ne), $1s^2 2s^2 2p^6$. The two additional electrons required to fill the valence orbitals give the oxide ion the charge of 2– ($O^{2-}$).

### Example 9.6.2: Determining the Electronic Structure of Anions

Selenium and iodine are two essential trace elements that form anions. Write the electron configurations of the anions.

**Solution**

$Se^{2-}$: $[Ar]3d^{10}4s^2 4p^6$

$I^-$: $[Kr]4d^{10}5s^2 5p^6$

**Exercise 9.6.2**

Write the electron configurations of a phosphorus atom and its negative ion. Give the charge on the anion.

*Answer available at the end of this section.*

## Summary

Atoms gain or lose electrons to form ions with particularly stable electron configurations. The charges of cations formed by the representative metals may be determined readily because, with few exceptions, the electronic structures of these ions have either a noble gas configuration or a completely filled electron shell. The charges of anions formed by the nonmetals may also be readily determined because these ions form when nonmetal atoms gain enough electrons to fill their valence shells.

## Glossary

**inert pair effect**
: tendency of heavy atoms to form ions in which their valence s electrons are not lost

**ionic bond**
: strong electrostatic force of attraction between cations and anions in an ionic compound

## Answers

**Exercise 9.6.1**

$K^+$: [Ar], $Mg^{2+}$: [Ne]

**Exercise 9.6.2**

P: $[Ne]3s^2 3p^3$

$P^{3-}$: $[Ne]3s^2 3p^6$

This page titled 9.6: Ionic Bonding (1) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 9.7: Covalent Bonds and Molecules

**Learning Objectives**

- Explain how covalent bonds form by sharing electrons.
- Describe the characteristic properties of molecules.



Figure 9.7.1 (Credit: CK-12 Foundation; Source: CK-12 Foundation; License: CK-12 Curriculum Materials license)

In a tennis match, two players keep hitting the ball back and forth. The ball bounces from one player to the other, over and over again. The ball keeps the players moving together on the court. What if the two players represented the nuclei of two atoms and the ball represented valence electrons? What would the back and forth movement of the ball represent? The answer is a covalent bond.

## Covalent Bonds: Sharing Electrons

A covalent bond is the force of attraction that holds together two atoms that share a pair of valence electrons. The shared electrons are attracted to the nuclei of both atoms. This forms a molecule consisting of two or more atoms. Covalent bonds form only between atoms of nonmetals.

The two atoms that are held together by a covalent bond may be atoms of the same element or different elements. When atoms of different elements form covalent bonds, a new substance, called a covalent compound, results. All covalent compounds are also molecules. Water is an example of a covalent compound. A water molecule is modeled in the figure below. A molecule is the smallest particle of a covalent compound that still has the properties of the compound.

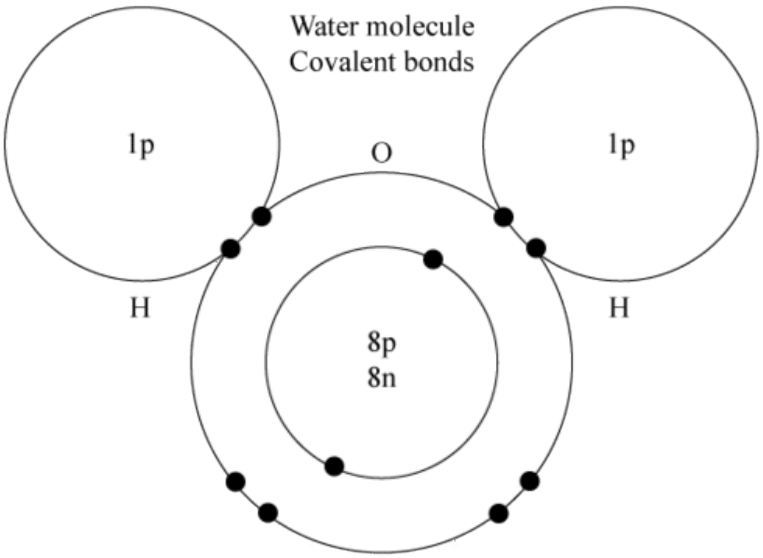



Figure 9.7.2 (Credit: Zachary Wilson; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

## Exercise 9.7.1

How many valence electrons does the oxygen atom (O) share with each hydrogen atom (H)? How many covalent bonds hold the water molecule together? (Refer to Figure 9.7.2)

*Answer available at the end of this section.*

The diagram in the figure below shows an example of covalent bonds between two atoms of the same element, in this case two atoms of oxygen. The diagram represents an oxygen molecule, so it's not a new compound. Oxygen normally occurs in diatomic ("two-atom") molecules. Several other elements also occur as diatomic molecules: hydrogen, nitrogen, and all but one of the halogens (fluorine, chlorine, bromine, and iodine).

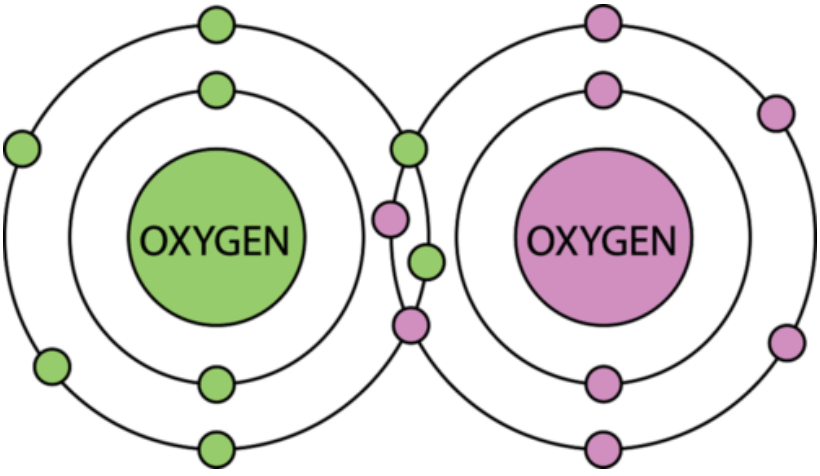



Figure 9.7.3 (Credit: Christopher AuYeung; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

## Exercise 9.7.2

How many electrons do these two oxygen atoms share? How many covalent bonds hold the oxygen molecule together?(Refer to Figure 9.7.3)

*Answer available at the end of this section.*

## Why Covalent Bonds Form

Covalent bonds form because they give atoms a more stable arrangement of electrons. Nonmetals in molecules follow the Octet Rule by sharing electrons in the formation of a covalent bond in such a way that each of the atoms involved in the bond can attain a noble-gas electron configuration. The shared electrons are "counted" for each of the atoms involved in the sharing. For hydrogen (H2)(H2), the shared pair of electrons means that each of the atoms is able to attain the electron configuration of helium, the noble gas with two electrons. For atoms other than hydrogen, the sharing of electrons will usually provide each of the atoms with eight valence electrons

## Multiple Covalent Bonds

Some molecules are not able to satisfy the octet rule by making only single covalent bonds between the atoms. Consider the compound ethene, which has a molecular formula of $\mathrm{C_2H_4}$. The carbon atoms are bonded together, with each carbon also bonded to two hydrogen atoms.

$$\text{two } \mathrm{C} \text{ atoms } = 2 \times 4 = 8 \text{ valence electrons}$$

$$\text{four } \mathrm{H} \text{ atoms } = 4 \times 1 = 4 \text{ valence electrons}$$

$$= \text{total of 12 valence electrons in the molecule}$$

If the Lewis electron dot structure was drawn with a single bond between the carbon atoms and with the octet rule followed, it would look like this:

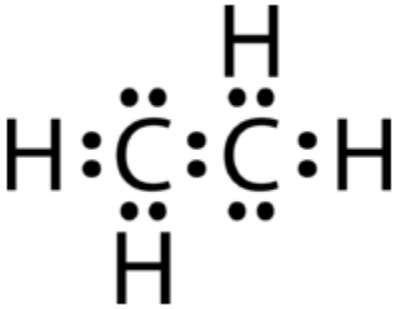


Figure $9.7.2$: Incorrect dot structure of ethene. (Credit: Joy Sheng; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

This Lewis structure is incorrect because it contains a total of 14 electrons. However, the Lewis structure can be changed by eliminating the lone pairs on the carbon atoms and having the carbon atoms share two pairs, instead of only one pair, of electrons.

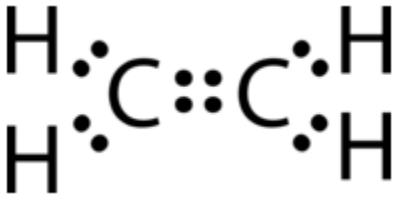


Figure $9.7.3$:Correct dot structure for ethene. (Credit: Joy Sheng; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

A **double covalent bond** is a covalent bond formed by atoms that share two pairs of electrons. The double covalent bond that occurs between the two carbon atoms in ethane can also be represented by a structural formula and with a molecular model, as shown in the figure below.

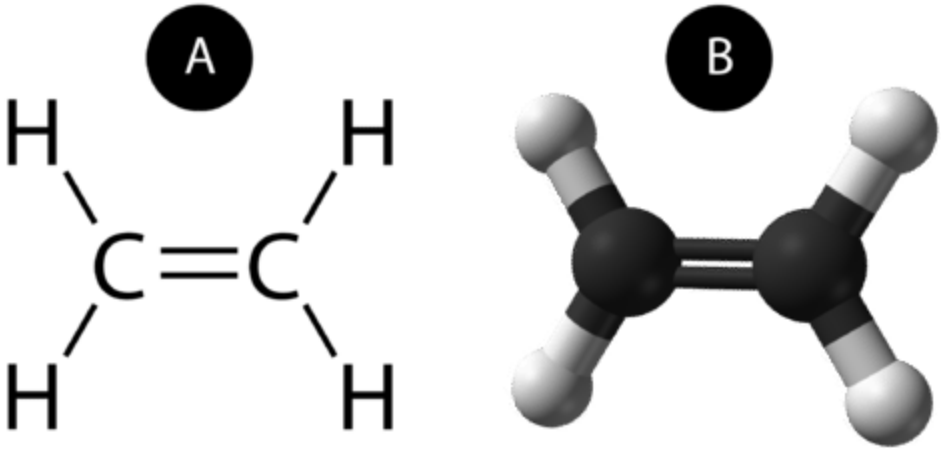



Figure 9.7.4: (A) The structural model for $C_2H_4$ consists of a double covalent bond between the two carbon atoms and single bonds to the hydrogen atoms. (B) Molecular model of $C_2H_4$. (Credit: Ben Mills (Wikimedia: Benjah-bmm27); Source: (A) wikimedia; (B) wikimedia; License: Public Domain)

A **triple covalent bond** is a covalent bond formed by atoms that share three pairs of electrons. The element nitrogen is a gas that composes the majority of Earth's atmosphere. A nitrogen atom has five valence electrons, which can be shown as one pair and three single electrons. When combining with another nitrogen atom to form a diatomic molecule, the three single electrons on each atom combine to form three shared pairs of electrons.

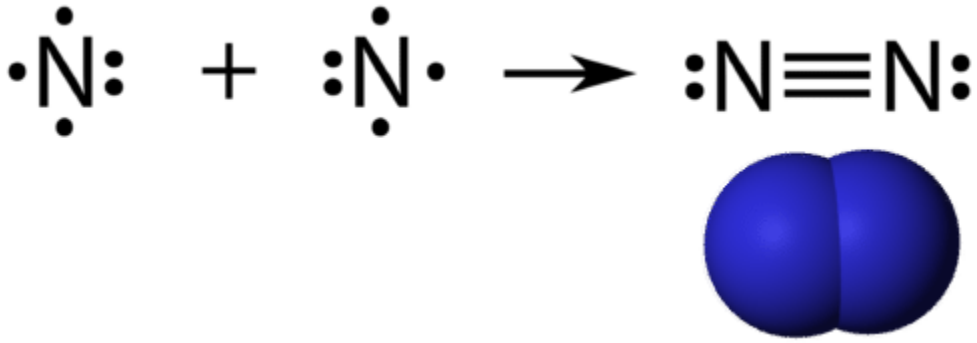


Figure 9.7.5: Triple bond in $N_2$. (Credit: CK-12 Foundation - Joy Sheng, using 3D molecular structure by Ben Mills (Wikimedia: Benjah-bmm27); Source: Molecular structure: wikimedia ; License: CC BY-NC-SA 3.0(opens in new window); Molecule Public Domain)

Each nitrogen atom follows the octet rule with one lone pair of electrons, and six electrons that are shared between the atoms.

## Properties of Molecules

The covalent bonds of molecules are responsible for many of their properties, regardless of whether they are elements or compounds. Because valence electrons are shared in molecules, rather than transferred between atoms as they are in ionic compounds, molecules have very different properties than ionic compounds.

- Many covalent compounds, especially those containing carbon and hydrogen, burn easily. In contrast, many ionic compounds do not burn.
- Many covalent compounds do not dissolve in water, whereas most ionic compounds dissolve well in water.
- Unlike ionic compounds, covalent compounds do not have freely moving electrons, so they cannot conduct electricity.
- The individual molecules of covalent compounds are more easily separated than the ions in a crystal, so most covalent compounds have relatively low boiling points. This explains why many of them are liquids or gases at room temperature.

## Section Summary

- A covalent bond is the force of attraction that holds together two atoms that share a pair of valence electrons. Covalent bonds form only between atoms of nonmetals.

- The two atoms that are held together in a covalent bond may be atoms of the same element or different elements. When atoms of different elements bond together, it forms a covalent compound.
- Covalent bonds form because the shared electrons fill each atom's outer energy level and this is the most stable arrangement of electrons.
- Lewis electron-dot structures show the bonding in covalent molecules.
- Covalent compounds contain two or more nonmetallic elements held together by covalent bonds, in which atoms share pairs of valence electrons. A molecule is the smallest particle of a covalent compound that still has the properties of the compound.
- Covalent bonds are responsible for many of the properties of covalent compounds. Covalent compounds have relatively low boiling points, cannot conduct electricity, and may not dissolve in water.
- Covalent bonds between atoms can be indicated either with dots (:) or a dash (–).
- Lewis structures can be drawn for molecules that share multiple pairs of electrons.
- A double covalent bond is a covalent bond formed by atoms that share two pairs of electrons.
- A triple covalent bond is a covalent bond formed by atoms that share three pairs of electrons.

## Answers

**Exercise** 9.7.1

The oxygen atom shares one pair of valence electrons with each hydrogen atom. Each pair of shared electrons represents one covalent bond, so two covalent bonds hold the water molecule together.

**Exercise** 9.7.2

The two oxygen atoms share two pairs of electrons, so two covalent bonds hold the oxygen molecule together.

This page titled 9.7: Covalent Bonds and Molecules is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 9.8: Covalent Bonding

**Learning Objectives**

- Describe the formation of covalent bonds
- Define electronegativity and assess the polarity of covalent bonds

In ionic compounds, electrons are transferred between atoms of different elements to form ions. But this is not the only way that compounds can be formed. Atoms can also make chemical bonds by sharing electrons between each other. Such bonds are called covalent bonds. Covalent bonds are formed between two atoms when both have similar tendencies to attract electrons to themselves (i.e., when both atoms have identical or fairly similar ionization energies and electron affinities). For example, two hydrogen atoms bond covalently to form an $H_2$ molecule; each hydrogen atom in the $H_2$ molecule has two electrons stabilizing it, giving each atom the same number of valence electrons as the noble gas He.

Compounds that contain covalent bonds exhibit different physical properties than ionic compounds. Because the attraction between molecules, which are electrically neutral, is weaker than that between electrically charged ions, covalent compounds generally have much lower melting and boiling points than ionic compounds. In fact, many covalent compounds are liquids or gases at room temperature, and, in their solid states, they are typically much softer than ionic solids. Furthermore, whereas ionic compounds are good conductors of electricity when dissolved in water, most covalent compounds, being electrically neutral, are poor conductors of electricity in any state.

## Formation of Covalent Bonds

Nonmetal atoms frequently form covalent bonds with other nonmetal atoms. For example, the hydrogen molecule, $H_2$, contains a covalent bond between its two hydrogen atoms. Figure 9.8.1 illustrates why this bond is formed. Starting on the far right, we have two separate hydrogen atoms with a particular potential energy, indicated by the red line. Along the *x*-axis is the distance between the two atoms. As the two atoms approach each other (moving left along the *x*-axis), their valence orbitals (1s) begin to overlap. The single electrons on each hydrogen atom then interact with both atomic nuclei, occupying the space around both atoms. The strong attraction of each shared electron to both nuclei stabilizes the system, and the potential energy decreases as the bond distance decreases. If the atoms continue to approach each other, the positive charges in the two nuclei begin to repel each other, and the potential energy increases. The bond length is determined by the distance at which the lowest potential energy is achieved.

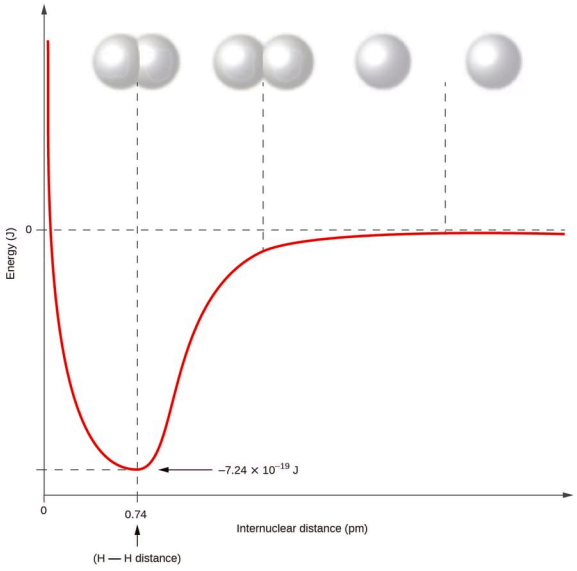



Figure 9.8.1: *The potential energy of two separate hydrogen atoms (right) decreases as they approach each other, and the single electrons on each atom are shared to form a covalent bond. The bond*

*length is the internuclear distance at which the lowest potential energy is achieved.*

It is essential to remember that energy must be added to break chemical bonds (an endothermic process), whereas forming chemical bonds releases energy (an exothermic process). In the case of $H_2$, the covalent bond is very strong; a large amount of energy, 436 kJ, must be added to break the bonds in one mole of hydrogen molecules and cause the atoms to separate:

$$\mathrm{H_2}(g) \longrightarrow 2\,\mathrm{H}(g) \quad \Delta H = 436\ \mathrm{kJ}$$

Conversely, the same amount of energy is released when one mole of $H_2$ molecules forms from two moles of H atoms:

$$2\,\mathrm{H}(g) \longrightarrow \mathrm{H_2}(g) \quad \Delta H = -436\ \mathrm{kJ}$$

## Pure vs. Polar Covalent Bonds

If the atoms that form a covalent bond are identical, as in $H_2$, $Cl_2$, and other diatomic molecules, then the electrons in the bond must be shared equally. We refer to this as a pure covalent bond. Electrons shared in pure covalent bonds have an equal probability of being near each nucleus. In the case of $Cl_2$, each atom starts off with seven valence electrons, and each Cl shares one electron with the other, forming one covalent bond:

$$\mathrm{Cl} + \mathrm{Cl} \longrightarrow \mathrm{Cl_2}$$

The total number of electrons around each individual atom consists of six nonbonding electrons and two shared (i.e., bonding) electrons for eight total electrons, matching the number of valence electrons in the noble gas argon. Since the bonding atoms are identical, $Cl_2$ also features a pure covalent bond.

When the atoms linked by a covalent bond are different, the bonding electrons are shared, but no longer equally. Instead, the bonding electrons are more attracted to one atom than the other, giving rise to a shift of electron density toward that atom. This unequal distribution of electrons is known as a polar covalent bond, characterized by a partial positive charge on one atom and a partial negative charge on the other. The atom that attracts the electrons more strongly acquires the partial negative charge and vice versa. For example, the electrons in the H–Cl bond of a hydrogen chloride molecule spend more time near the chlorine atom than near the hydrogen atom. Thus, in an HCl molecule, the chlorine atom carries a partial negative charge and the hydrogen atom has a partial positive charge. Figure $9.8.2$ shows the distribution of electrons in the H–Cl bond. Note that the shaded area around Cl is much larger than it is around H. Compare this to Figure $9.8.1$, which shows the even distribution of electrons in the $H_2$ nonpolar bond.

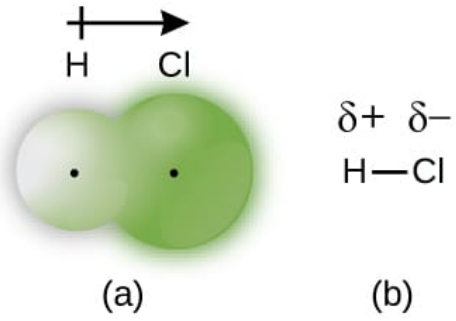



Figure $9.8.2$: (a) The distribution of electron density in the HCl molecule is uneven. The electron density is greater around the chlorine nucleus. The small, black dots indicate the location of the hydrogen and chlorine nuclei in the molecule. (b) Symbols δ+ and δ– indicate the polarity of the H–Cl bond.

We sometimes designate the positive and negative atoms in a polar covalent bond using a lowercase Greek letter “delta,” δ, with a plus sign or minus sign to indicate whether the atom has a partial positive charge (δ+) or a partial negative charge (δ–). This symbolism is shown for the H–Cl molecule in Figure $9.8.2b$.

## Electronegativity

Whether a bond is nonpolar or polar covalent is determined by a property of the bonding atoms called electronegativity. Electronegativity is a measure of the tendency of an atom to attract electrons (or electron density) towards itself. It determines how the shared electrons are distributed between the two atoms in a bond. The more strongly an atom attracts the electrons in

its bonds, the larger its electronegativity. Electrons in a polar covalent bond are shifted toward the more electronegative atom; thus, the more electronegative atom is the one with the partial negative charge. The greater the difference in electronegativity, the more polarized the electron distribution and the larger the partial charges of the atoms.

Figure 9.8.3 shows the electronegativity values of the elements as proposed by one of the most famous chemists of the twentieth century: Linus Pauling. In general, electronegativity increases from left to right across a period in the periodic table and decreases down a group. Thus, the nonmetals, which lie in the upper right, tend to have the highest electronegativities, with fluorine the most electronegative element of all (EN = 4.0). Metals tend to be less electronegative elements, and the group 1 metals have the lowest electronegativities. Note that noble gases are excluded from this figure because these atoms usually do not share electrons with others atoms since they have a full valence shell. (While noble gas compounds such as $XeO_2$ do exist, they can only be formed under extreme conditions, and thus they do not fit neatly into the general model of electronegativity.)

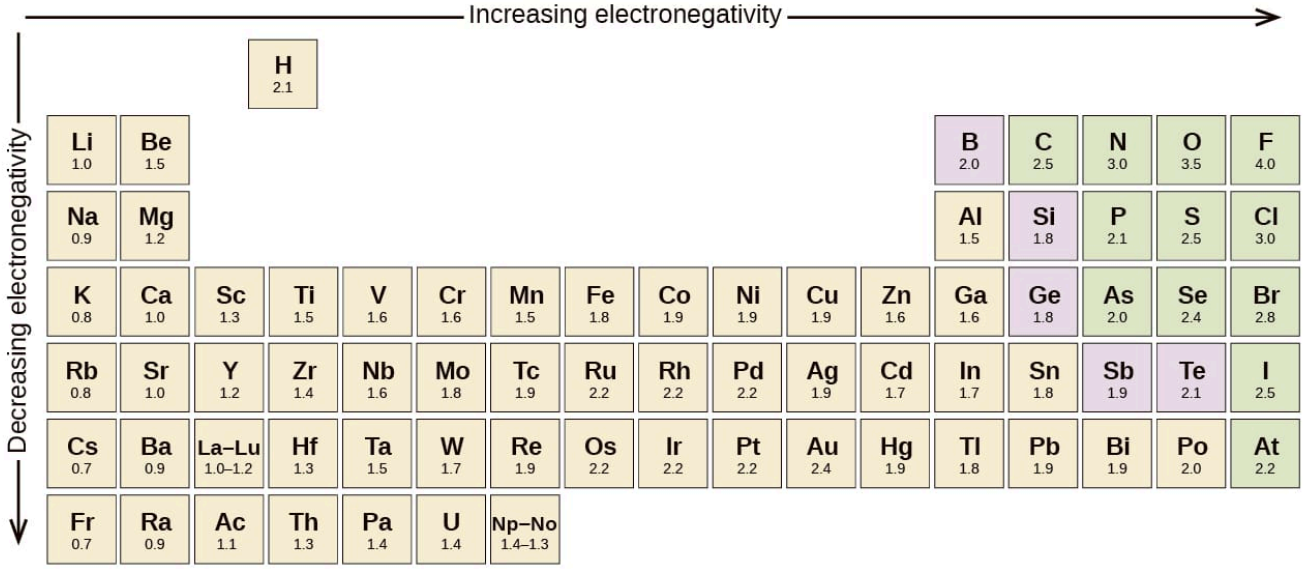
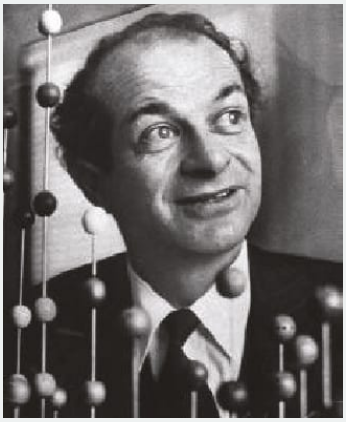



Figure 9.8.3: The electronegativity values derived by Pauling follow predictable periodic trends with the higher electronegativities toward the upper right of the periodic table.

## Linus Pauling

Linus Pauling is the only person to have received two unshared (individual) Nobel Prizes: one for chemistry in 1954 for his work on the nature of chemical bonds and one for peace in 1962 for his opposition to weapons of mass destruction. He developed many of the theories and concepts that are foundational to our current understanding of chemistry, including electronegativity and resonance structures.



Linus Pauling (1901–1994) made many important contributions to the field of chemistry. He was also a prominent activist, publicizing issues related to health and nuclear weapons.

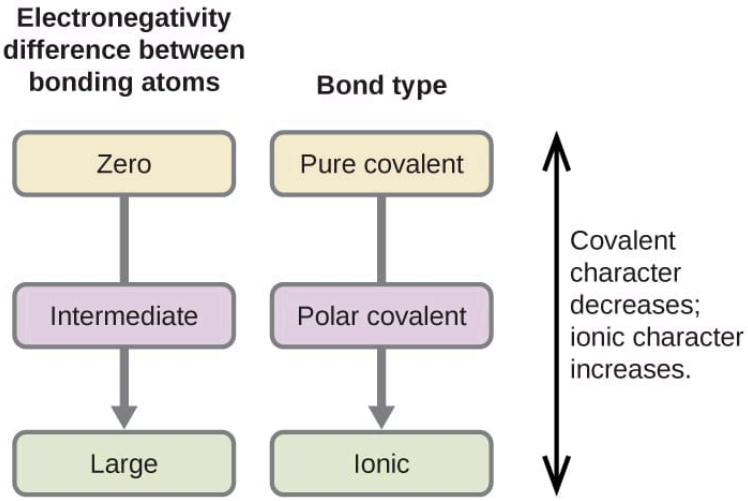
Pauling also contributed to many other fields besides chemistry. His research on sickle cell anemia revealed the cause of the disease—the presence of a genetically inherited abnormal protein in the blood—and paved the way for the field of molecular genetics. His work was also pivotal in curbing the testing of nuclear weapons; he proved that radioactive fallout from nuclear testing posed a public health risk.

## Electronegativity versus Electron Affinity

We must be careful not to confuse electronegativity and electron affinity. The electron affinity of an element is a measurable physical quantity, namely, the energy released or absorbed when an isolated gas-phase atom acquires an electron, measured in kJ/mol. Electronegativity, on the other hand, describes how tightly an atom attracts electrons in a bond. It is a dimensionless quantity that is calculated, not measured. Pauling derived the first electronegativity values by comparing the amounts of energy required to break different types of bonds. He chose an arbitrary relative scale ranging from 0 to 4.

## Electronegativity and Bond Type

The absolute value of the difference in electronegativity (ΔEN) of two bonded atoms provides a rough measure of the polarity to be expected in the bond and, thus, the bond type. When the difference is very small or zero, the bond is covalent and nonpolar. When it is large, the bond is polar covalent or ionic. The absolute values of the electronegativity differences between the atoms in the bonds H–H, H–Cl, and Na–Cl are 0 (nonpolar), 0.9 (polar covalent), and 2.1 (ionic), respectively. The degree to which electrons are shared between atoms varies from completely equal (pure covalent bonding) to not at all (ionic bonding). Figure 9.8.4 shows the relationship between electronegativity difference and bond type.




| Bond Type | Electronegativity Difference |
|---|---|
| pure covalent | < 0.4 |
| polar covalent | between 0.4 and 1.8 |
| ionic | > 1.8 |

Figure 9.8.4: As the electronegativity difference increases between two atoms, the bond becomes more ionic.

A rough approximation of the electronegativity differences associated with covalent, polar covalent, and ionic bonds is shown in Figure 9.8.4. This table is just a general guide, however, with many exceptions. For example, the H and F atoms in HF have an electronegativity difference of 1.9, and the N and H atoms in $NH_3$ a difference of 0.9, yet both of these compounds form bonds that are considered polar covalent. Likewise, the Na and Cl atoms in NaCl have an electronegativity difference of 2.1, and the Mn and I atoms in $MnI_2$ have a difference of 1.0, yet both of these substances form ionic compounds.

The best guide to the covalent or ionic character of a bond is to consider the types of atoms involved and their relative positions in the periodic table. Bonds between two nonmetals are generally covalent; bonding between a metal and a nonmetal is often ionic.

Some compounds contain both covalent and ionic bonds. The atoms in polyatomic ions, such as $OH^-$, $NO_3^-$, and $NH_4^+$, are held together by polar covalent bonds. However, these polyatomic ions form ionic compounds by combining with ions of opposite charge. For example, potassium nitrate, $KNO_3$, contains the $K^+$ cation and the polyatomic $NO_3^-$ anion. Thus, bonding in potassium nitrate is ionic, resulting from the electrostatic attraction between the ions $K^+$ and $NO_3^-$, as well as covalent between the nitrogen

and oxygen atoms in $NO_3^-$.

## Example 9.8.1: Electronegativity and Bond Polarity

Bond polarities play an important role in determining the structure of proteins. Using the electronegativity values in Table A2, arrange the following covalent bonds—all commonly found in amino acids—in order of increasing polarity. Then designate the positive and negative atoms using the symbols δ+ and δ–:

C–H, C–N, C–O, N–H, O–H, S–H

**Solution**

The polarity of these bonds increases as the absolute value of the electronegativity difference increases. The atom with the δ– designation is the more electronegative of the two. Table 9.8.1 shows these bonds in order of increasing polarity.

*Table 9.8.1: Bond Polarity and Electronegativity Difference*

| Bond | ΔEN | Polarity |
|---|---|---|
| C–H | 0.4 | $\overset{\delta-}{\mathrm{C}}-\overset{\delta+}{\mathrm{H}}$ |
| S–H | 0.4 | $\overset{\delta-}{\mathrm{S}}-\overset{\delta+}{\mathrm{H}}$ |
| C–N | 0.5 | $\overset{\delta+}{\mathrm{C}}-\overset{\delta-}{\mathrm{N}}$ |
| N–H | 0.9 | $\overset{\delta-}{\mathrm{N}}-\overset{\delta+}{\mathrm{H}}$ |
| C–O | 1.0 | $\overset{\delta+}{\mathrm{C}}-\overset{\delta-}{\mathrm{O}}$ |
| O–H | 1.4 | $\overset{\delta-}{\mathrm{O}}-\overset{\delta+}{\mathrm{H}}$ |

## Exercise 9.8.1

Silicones are polymeric compounds containing, among others, the following types of covalent bonds: Si–O, Si–C, C–H, and C–C. Using the electronegativity values in Figure 9.8.3, arrange the bonds in order of increasing polarity and designate the positive and negative atoms using the symbols δ+ and δ–.

**Answer**

*Answer to Exercise 7.2.1*

| Bond | Electronegativity Difference | Polarity |
|---|---|---|
| C–C | 0.0 | nonpolar |
| C–H | 0.4 | $\overset{\delta-}{\mathrm{C}}-\overset{\delta+}{\mathrm{H}}$ |
| Si–C | 0.7 | $\overset{\delta+}{\mathrm{Si}}-\overset{\delta-}{\mathrm{C}}$ |
| Si–O | 1.7 | $\overset{\delta+}{\mathrm{Si}}-\overset{\delta-}{\mathrm{O}}$ |

# Summary

Covalent bonds form when electrons are shared between atoms and are attracted by the nuclei of both atoms. In pure covalent bonds, the electrons are shared equally. In polar covalent bonds, the electrons are shared unequally, as one atom exerts a stronger force of attraction on the electrons than the other. The ability of an atom to attract a pair of electrons in a chemical bond is called its electronegativity. The difference in electronegativity between two atoms determines how polar a bond will be. In a diatomic

molecule with two identical atoms, there is no difference in electronegativity, so the bond is nonpolar or pure covalent. When the electronegativity difference is very large, as is the case between metals and nonmetals, the bonding is characterized as ionic.

## Glossary

**bond length**
: distance between the nuclei of two bonded atoms at which the lowest potential energy is achieved

**covalent bond**
: bond formed when electrons are shared between atoms

**electronegativity**
: tendency of an atom to attract electrons in a bond to itself

**polar covalent bond**
: covalent bond between atoms of different electronegativities; a covalent bond with a positive end and a negative end

**pure covalent bond**
: (also, nonpolar covalent bond) covalent bond between atoms of identical electronegativities

This page titled 9.8: Covalent Bonding is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 9.9: Prelude to Liquids and Solids

The great distances between atoms and molecules in a gaseous phase, and the corresponding absence of any significant interactions between them, allows for simple descriptions of many physical properties that are the same for all gases, regardless of their chemical identities. As described in the final module of the chapter on gases, this situation changes at high pressures and low temperatures—conditions that permit the atoms and molecules to interact to a much greater extent.




Figure 9.9.1: Solid carbon dioxide (“dry ice”, left) sublimes vigorously when placed in a liquid (right), cooling the liquid and generating a fog of condensed water vapor above the cylinder. (credit: modification of work by Paul Flowers) This figure shows pieces of a white substance which appear to be sublimating. To the right of these pieces are three graduated cylinders. Each cylinder holds a different color liquid, and above the liquid, the cylinders are filled with a fog-like substance. This fog-like substance swirls out of the top and around the outside of the cylinders.

In the liquid and solid states, these interactions are of considerable strength and play an important role in determining a number of physical properties that *do* depend on the chemical identity of the substance. In this chapter, the nature of these interactions and their effects on various physical properties of liquid and solid phases will be examined.

This page titled 9.9: Prelude to Liquids and Solids is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 9.10: Intermolecular Forces

**Learning Objectives**

- Explain the difference between dipole-dipole forces, dispersion forces, and hydrogen bonds

Molecules are attracted to each other, but not as strongly as a chemical bond such as an ionic bond. However these attractions are important as we will see. In this section we will discuss some different forms of these attractions, which are called **intermolecular forces.**

## Dipole-Dipole Forces

**Dipole-dipole forces** are the attractive forces that occur between polar molecules. A molecule of hydrogen chloride has a partially positive hydrogen atom and a partially negative chlorine atom. In a collection of many hydrogen chloride molecules, the molecules will align themselves so that the oppositely charged regions of neighboring molecules are near each other.

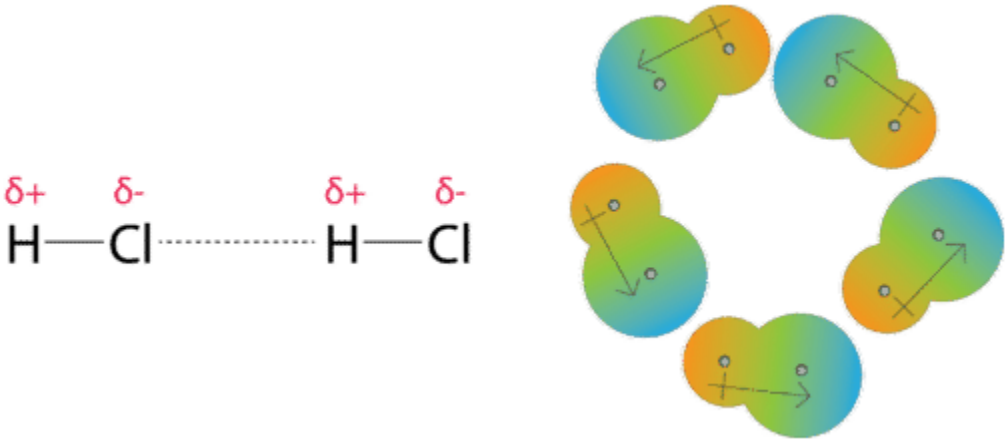



Figure 9.10.1: Dipole-dipole forces are a result of the attraction of the positive end of one dipole to the negative end of a neighboring dipole. (Credit: Joy Sheng; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

Dipole-dipole forces are similar in nature to ionic bonds, but much weaker.

## London Dispersion Forces

Dispersion forces are also considered a type of van der Waals force and are the weakest of all intermolecular forces. They are often called London dispersion forces after Fritz London (1900-1954), who first proposed their existence in 1930. **London dispersion forces** are the intermolecular forces that occur between atoms, and between nonpolar molecules as a result of the motion of electrons.

The electron cloud of a helium atom contains two electrons, which can normally be expected to be equally distributed spatially around the nucleus. However, at any given moment the electron distribution may be uneven, resulting in an instantaneous dipole. This weak and temporary dipole subsequently influences neighboring helium atoms through electrostatic attraction and repulsion. It induces a dipole on nearby helium atoms (see figure below).

CK-12 LICENSED UNDER CK-12 License
© CK-12 Foundation • Visit us at ck12.org

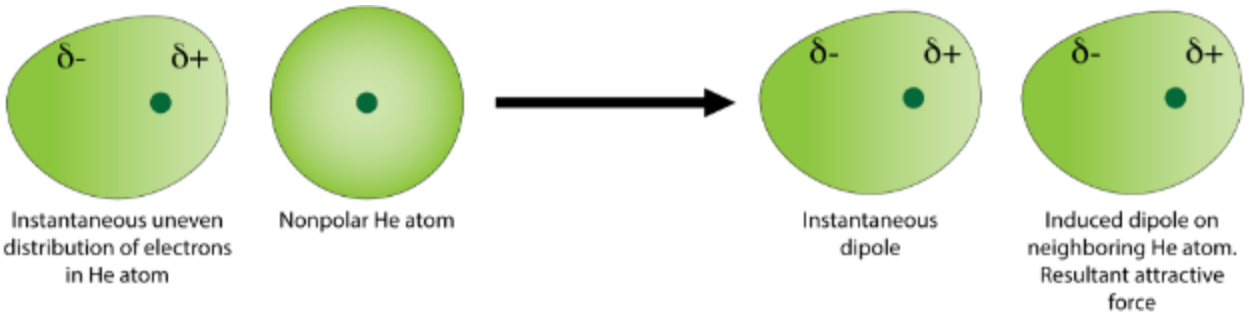



Figure 9.10.3: A short-lived or instantaneous dipole in a helium atom. (Credit: Zachary Wilson; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

The instantaneous and induced dipoles are weakly attracted to one another. The strength of dispersion forces increases as the number of electrons in the atoms or nonpolar molecules increases.

The halogen group consists of four elements that all take the form of nonpolar diatomic molecules. The table below shows a comparison of the melting and boiling points for each.

*Melting and Boiling Points of Halogens*

**Table 9.10.1: Melting and Boiling Points of Halogens**

| Molecule | Total Number of Electrons | Melting Point (°C) | Boiling Point (°C) | Physical State at Room Temperature |
|---|---|---|---|---|
| $F_2$ | 18 | -220 | -188 | gas |
| $Cl_2$ | 34 | -102 | -34 | gas |
| $Br_2$ | 70 | -7 | 59 | liquid |
| $I_2$ | 106 | 114 | 184 | solid |

The dispersion forces are strongest for iodine molecules because they have the greatest number of electrons. The relatively stronger forces result in melting and boiling points that are the highest of the halogen group. These forces are strong enough to hold iodine molecules close together in the solid state at room temperature. The dispersion forces are progressively weaker for bromine, chloride, and fluorine; this is illustrated in their steadily lower melting and boiling points. Bromine is a liquid at room temperature, while chlorine and fluorine are gases whose molecules are much further apart from one another. Intermolecular forces are nearly nonexistent in the gas state, and so the dispersion forces in chlorine and fluorine only become measurable as the temperature decreases and they condense into the liquid state.

## Hydrogen Bonding

The attractive force between water molecules is a dipole interaction. The hydrogen atoms are bound to the highly electronegative oxygen atom (which also possesses two lone pair sets of electrons, making for a very polar bond). The partially positive hydrogen atom of one molecule is then attracted to the oxygen atom of a nearby water molecule (see figure below).

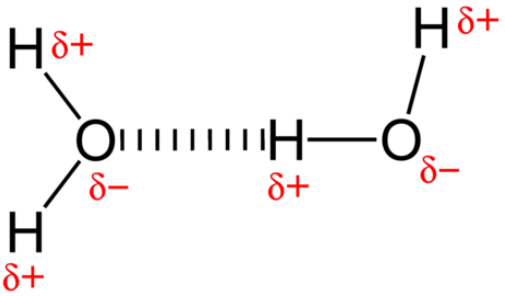



Figure 9.10.3: A hydrogen bond in water occurs between the hydrogen atom of one water molecule and the lone pair of electrons on an oxygen atom of a neighboring water molecule. (Credit: Ben Mills (Wikimedia: Benjah-bmm27); Source: wikimedia; License: Public Domain)

A **hydrogen bond** is an intermolecular attractive force in which a hydrogen atom that is covalently bonded to a small, highly electronegative atom is attracted to a lone pair of electrons on an atom in a neighboring molecule. Hydrogen bonds are very strong compared to other dipole interactions. The strength of a typical hydrogen bond is about $5\%$ of that of a covalent bond.

Hydrogen bonding occurs only in molecules where hydrogen is covalently bonded to one of three elements: fluorine, oxygen, or nitrogen. These three elements are so electronegative that they withdraw the majority of the electron density in the covalent bond with hydrogen, leaving the $\mathrm{H}$ atom very electron-deficient. The $\mathrm{H}$ atom nearly acts as a bare proton, leaving it very attracted to lone pair electrons on a nearby atom.

The hydrogen bonding that occurs in water leads to some unusual, but very important, properties. Most molecular compounds that have a mass similar to water are gases at room temperature. Because of the strong hydrogen bonds, water molecules are able to stay condensed in the liquid state. The figure below shows how the bent shape, and two hydrogen atoms per molecule, allows each water molecule to be able to hydrogen bond to two other molecules.

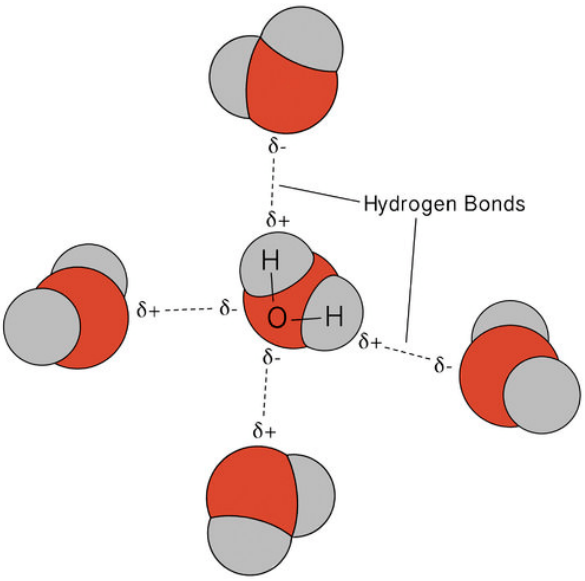



Figure $9.10.4$: Multiple hydrogen bonds occur simultaneously in water because of its bent shape and the presence of two hydrogen atoms per molecule. (Credit: Laura Guerin; Source: CK-12 Foundation; License: CC BY-NC 3.0(opens in new window))

In the liquid state, the hydrogen bonds of water can break and reform as the molecules flow from one place to another. When water is cooled, the molecules begin to slow down. Eventually, when water is frozen to ice, the hydrogen bonds become permanent and form a very specific network (see figure below).

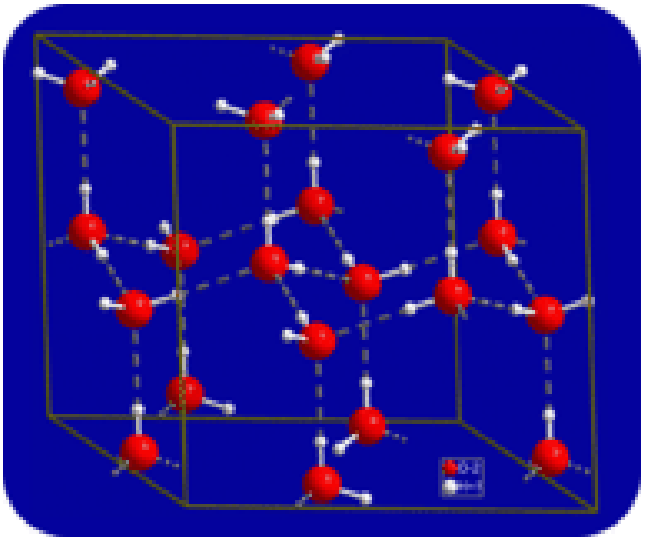


Figure $9.10.5$: When water freezes to ice, the hydrogen bonding

network becomes permanent. Each oxygen atom has an approximately tetrahedral geometry—two covalent bonds and two hydrogen bonds. (Credit: User:Materialscientist/Wikimedia Commons; Source: wikimedia; License: Public Domain)

The bent shape of the molecules leads to gaps in the hydrogen bonding network of ice. Ice has a very unusual property—its solid state is less dense than its liquid state. Ice floats in water. Virtually all other substances are denser in the solid state than in the liquid state. Hydrogen bonds play a very important biological role in the physical structures of proteins and nucleic acids.

## Section Summary

- Van der Waals forces are weak interactions between molecules that involve dipoles.
- Polar molecules have permanent dipole-dipole interactions.
- Nonpolar molecules can interact by way of London dispersion forces.
- Hydrogen bonds form when a $\mathrm{H}$ attached to a $\mathrm{N}$, $\mathrm{O}$, or $\mathrm{F}$ atom interacts with another $\mathrm{N}$, $\mathrm{O}$, or $\mathrm{F}$ atom.

## Glossary

**dipole-dipole forces**
: The attractive forces that occur between polar molecules.

**dispersion forces**
: The intermolecular forces that occur between atoms, and between nonpolar molecules as a result of the motion of electrons.

**hydrogen bond**
: an intermolecular attractive force in which a hydrogen atom that is covalently bonded to a small, highly electronegative atom is attracted to a lone pair of electrons on an atom in a neighboring molecule.

**intermolecular forces**
: Attractions between molecules which are not as strong as a chemical bond.

This page titled 9.10: Intermolecular Forces is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.4: Time, Velocity, and Speed by Andrew Park is licensed CC BY 4.0.

# 9.11: Intermolecular Forces (1)

## Learning Objectives

- Describe the types of intermolecular forces possible between atoms or molecules in condensed phases (dispersion forces, dipole-dipole attractions, and hydrogen bonding)
- Identify the types of intermolecular forces experienced by specific molecules based on their structures
- Explain the relation between the intermolecular forces present within a substance and the temperatures associated with changes in its physical state

As was the case for gaseous substances, the kinetic molecular theory may be used to explain the behavior of solids and liquids. In the following description, the term *particle* will be used to refer to an atom, molecule, or ion. Note that we will use the popular phrase "intermolecular attraction" to refer to attractive forces between the particles of a substance, regardless of whether these particles are molecules, atoms, or ions.

Consider these two aspects of the molecular-level environments in solid, liquid, and gaseous matter:

- Particles in a solid are tightly packed together and often arranged in a regular pattern; in a liquid, they are close together with no regular arrangement; in a gas, they are far apart with no regular arrangement.
- Particles in a solid vibrate about fixed positions and do not generally move in relation to one another; in a liquid, they move past each other but remain in essentially constant contact; in a gas, they move independently of one another except when they collide.

The differences in the properties of a solid, liquid, or gas reflect the strengths of the attractive forces between the atoms, molecules, or ions that make up each phase. The phase in which a substance exists depends on the relative extents of its intermolecular forces (IMFs) and the kinetic energies (KE) of its molecules. IMFs are the various forces of attraction that may exist between the atoms and molecules of a substance due to electrostatic phenomena, as will be detailed in this module. These forces serve to hold particles close together, whereas the particles' KE provides the energy required to overcome the attractive forces and thus increase the distance between particles. Figure 9.11.1 illustrates how changes in physical state may be induced by changing the temperature, hence, the average KE, of a given substance.

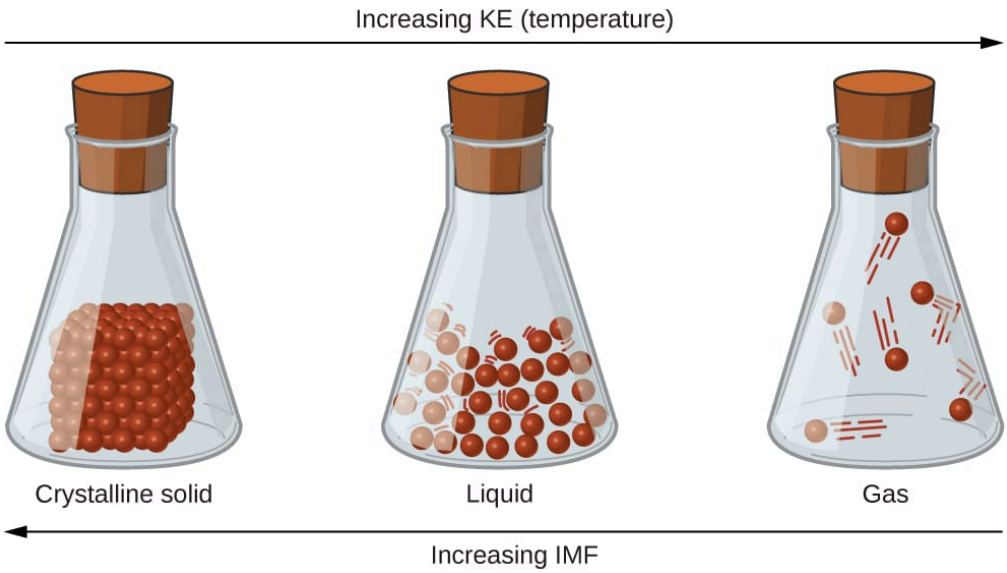



Figure 9.11.1: Transitions between solid, liquid, and gaseous states of a substance occur when conditions of temperature or pressure favor the associated changes in intermolecular forces. (Note: The space between particles in the gas phase is much greater than shown.) Three sealed flasks are labeled, "Crystalline solid," "Liquid," and "Gas," from left to right. The first flask holds a cube composed of small spheres sitting on the bottom while the second flask shows a lot of small spheres in the bottom that are spaced a small distance apart from one

another and have lines around them to indicate motion. The third flask shows a few spheres spread far from one another with larger lines to indicate motion. There is a right-facing arrow that spans the top of all three flasks. The arrow is labeled, “Increasing K E ( temperature ).” There is a left-facing arrow that spans the bottom of all three flasks. The arrow is labeled, “Increasing I M F.”

As an example of the processes depicted in this figure, consider a sample of water. When gaseous water is cooled sufficiently, the attractions between $H_2O$ molecules will be capable of holding them together when they come into contact with each other; the gas condenses, forming liquid $H_2O$. For example, liquid water forms on the outside of a cold glass as the water vapor in the air is cooled by the cold glass, as seen in Figure 9.11.2.




Figure 9.11.2: Condensation forms when water vapor in the air is cooled enough to form liquid water, such as (a) on the outside of a cold beverage glass or (b) in the form of fog. (credit a: modification of work by Jenny Downing; credit b: modification of work by Cory Zanker) Image a shows a brown colored beverage in a glass with condensation on the outside. Image b shows a body of water with fog hovering above the surface of the water.

We can also liquefy many gases by compressing them, if the temperature is not too high. The increased pressure brings the molecules of a gas closer together, such that the attractions between the molecules become strong relative to their KE. Consequently, they form liquids. Butane, $C_4H_{10}$, is the fuel used in disposable lighters and is a gas at standard temperature and pressure. Inside the lighter’s fuel compartment, the butane is compressed to a pressure that results in its condensation to the liquid state, as shown in Figure 9.11.3.



Figure 9.11.3: Gaseous butane is compressed within the storage compartment of a disposable lighter,

resulting in its condensation to the liquid state. (credit: modification of work by “Sam-Cat”/Flickr)

Finally, if the temperature of a liquid becomes sufficiently low, or the pressure on the liquid becomes sufficiently high, the molecules of the liquid no longer have enough KE to overcome the IMF between them, and a solid forms. A more thorough discussion of these and other changes of state, or phase transitions, is provided in a later module of this chapter.

## Forces between Molecules

Under appropriate conditions, the attractions between all gas molecules will cause them to form liquids or solids. This is due to intermolecular forces, not *intra*molecular forces. *Intra*molecular forces are those *within* the molecule that keep the molecule together, for example, the bonds between the atoms. *Inter*molecular forces are the attractions *between* molecules, which determine many of the physical properties of a substance. Figure 9.11.4 illustrates these different molecular forces. The strengths of these attractive forces vary widely, though usually the IMFs between small molecules are weak compared to the intramolecular forces that bond atoms together within a molecule. For example, to overcome the IMFs in one mole of liquid HCl and convert it into gaseous HCl requires only about 17 kilojoules. However, to break the covalent bonds between the hydrogen and chlorine atoms in one mole of HCl requires about 25 times more energy—430 kilojoules.

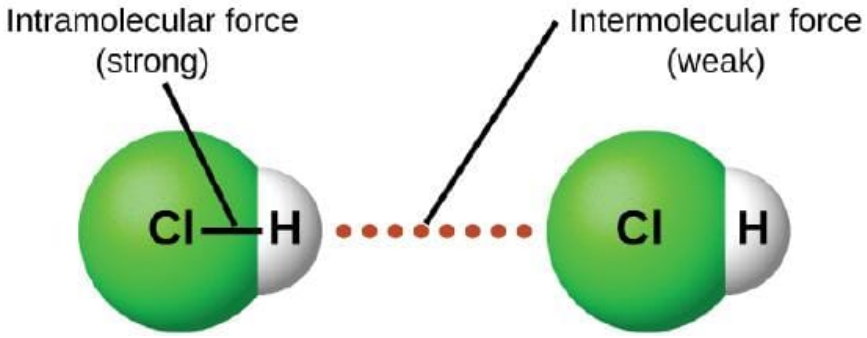



Figure 9.11.4: Intramolecular forces keep a molecule intact. Intermolecular forces hold multiple molecules together and determine many of a substance’s properties. An image is shown in which two molecules composed of a green sphere labeled “C l” connected on the right to a white sphere labeled “H” are near one another with a dotted line labeled “Intermolecular force ( weak )” drawn between them. A line connects the two spheres in each molecule and the line is labeled “Intramolecular force ( strong ).”

All of the attractive forces between neutral atoms and molecules are known as van der Waals forces, although they are usually referred to more informally as intermolecular attraction. We will consider the various types of IMFs in the next three sections of this module.

## Dispersion Forces

One of the three van der Waals forces is present in all condensed phases, regardless of the nature of the atoms or molecules composing the substance. This attractive force is called the London dispersion force in honor of German-born American physicist Fritz London who, in 1928, first explained it. This force is often referred to as simply the dispersion force. Because the electrons of an atom or molecule are in constant motion (or, alternatively, the electron’s location is subject to quantum-mechanical variability), at any moment in time, an atom or molecule can develop a temporary, instantaneous dipole if its electrons are distributed asymmetrically. The presence of this dipole can, in turn, distort the electrons of a neighboring atom or molecule, producing an induced dipole. These two rapidly fluctuating, temporary dipoles thus result in a relatively weak electrostatic attraction between the species—a so-called dispersion force like that illustrated in Figure 9.11.5.

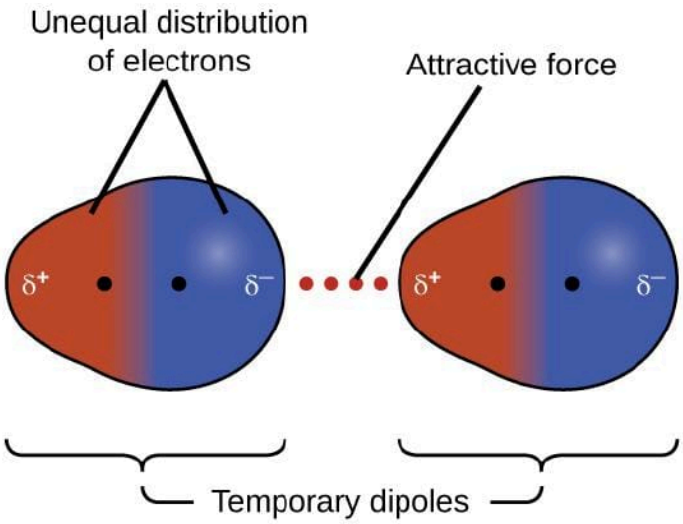



Figure 9.11.5: Dispersion forces result from the formation of temporary dipoles, as illustrated here for two nonpolar diatomic molecules. Two pairs of molecules are shown where each molecule has one larger blue side labeled “delta sign, negative sign” and a smaller red side labeled “delta sign, positive sign.” Toward the middle of the both molecules, but still on each distinct side, is a black dot. Between the two images is a dotted line labeled, “Attractive force.” In the first image, the red and blue sides are labeled, “Unequal distribution of electrons.” Below both images are brackets. The brackets are labeled, “Temporary dipoles.”

Dispersion forces that develop between atoms in different molecules can attract the two molecules to each other. The forces are relatively weak, however, and become significant only when the molecules are very close. Larger and heavier atoms and molecules exhibit stronger dispersion forces than do smaller and lighter atoms and molecules. $F_2$ and $Cl_2$ are gases at room temperature (reflecting weaker attractive forces); $Br_2$ is a liquid, and $I_2$ is a solid (reflecting stronger attractive forces). Trends in observed melting and boiling points for the halogens clearly demonstrate this effect, as seen in Table 9.11.1.

*Table 9.11.1: Melting and Boiling Points of the Halogens*

| Halogen | Molar Mass | Atomic Radius | Melting Point | Boiling Point |
|---|---|---|---|---|
| fluorine, $F_2$ | 38 g/mol | 72 pm | 53 K | 85 K |
| chlorine, $Cl_2$ | 71 g/mol | 99 pm | 172 K | 238 K |
| bromine, $Br_2$ | 160 g/mol | 114 pm | 266 K | 332 K |
| iodine, $I_2$ | 254 g/mol | 133 pm | 387 K | 457 K |
| astatine, $At_2$ | 420 g/mol | 150 pm | 575 K | 610 K |

The increase in melting and boiling points with increasing atomic/molecular size may be rationalized by considering how the strength of dispersion forces is affected by the electronic structure of the atoms or molecules in the substance. In a larger atom, the valence electrons are, on average, farther from the nuclei than in a smaller atom. Thus, they are less tightly held and can more easily form the temporary dipoles that produce the attraction. The measure of how easy or difficult it is for another electrostatic charge (for example, a nearby ion or polar molecule) to distort a molecule’s charge distribution (its electron cloud) is known as polarizability. A molecule that has a charge cloud that is easily distorted is said to be very polarizable and will have large dispersion forces; one with a charge cloud that is difficult to distort is not very polarizable and will have small dispersion forces.

## Example 9.11.1: London Forces and Their Effects

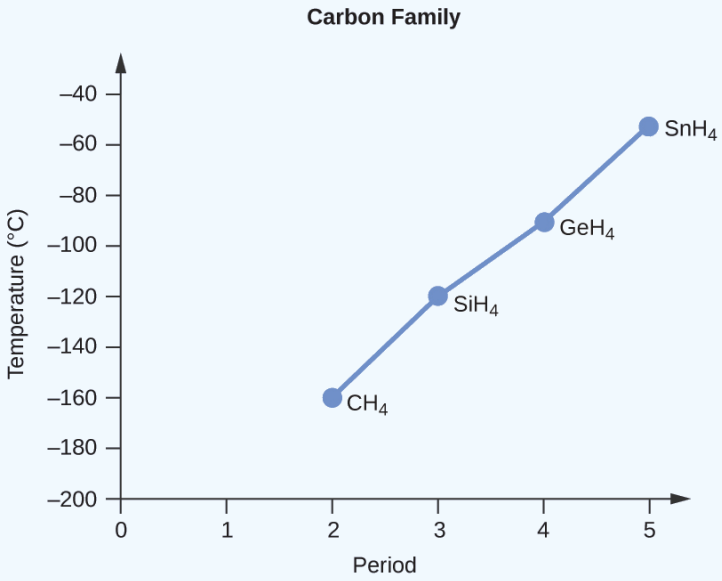
Order the following compounds of a group 14 element and hydrogen from lowest to highest boiling point: $CH_4$, $SiH_4$, $GeH_4$, and $SnH_4$. Explain your reasoning.

**Solution**

Applying the skills acquired in the chapter on chemical bonding and molecular geometry, all of these compounds are predicted to be nonpolar, so they may experience only dispersion forces: the smaller the molecule, the less polarizable and the weaker the dispersion forces; the larger the molecule, the larger the dispersion forces. The molar masses of $CH_4$, $SiH_4$, $GeH_4$, and $SnH_4$ are approximately 16 g/mol, 32 g/mol, 77 g/mol, and 123 g/mol, respectively. Therefore, $CH_4$ is expected to have the lowest boiling point and $SnH_4$ the highest boiling point. The ordering from lowest to highest boiling point is expected to be

$$CH_4 < SiH_4 < GeH_4 < SnH_4$$

A graph of the actual boiling points of these compounds versus the period of the group 14 elements shows this prediction to be correct:




A line graph, titled “Carbon Family,” is shown where the y-axis is labeled “Temperature, ( degree sign C )” and has values of “negative 200” to “negative 40” from bottom to top in increments of 20. The x-axis is labeled “Period” and has values of “0” to “5” in increments of 1. The first point on the graph is labeled “C H subscript 4” and is at point “2, negative 160.” The second point on the graph is labeled “S i H subscript 4” and is at point “3, negative 120” while the third point on the graph is labeled “G e H subscript 4” and is at point “4, negative 100.” The fourth point on the graph is labeled “S n H subscript 4” and is at point “5, negative 60.”

## Exercise 9.11.1

Order the following hydrocarbons from lowest to highest boiling point: $C_2H_6$, $C_3H_8$, and $C_4H_{10}$.

*Answer available at the end of this section.*

The shapes of molecules also affect the magnitudes of the dispersion forces between them. For example, boiling points for the isomers *n*-pentane, isopentane, and neopentane (shown in Figure 9.11.6) are 36 °C, 27 °C, and 9.5 °C, respectively. Even though these compounds are composed of molecules with the same chemical formula, $C_5H_{12}$, the difference in boiling points suggests that dispersion forces in the liquid phase are different, being greatest for *n*-pentane and least for neopentane. The elongated shape of *n*-pentane provides a greater surface area available for contact between molecules, resulting in correspondingly stronger dispersion forces. The more compact shape of isopentane offers a smaller surface area available for intermolecular contact and, therefore, weaker dispersion forces. Neopentane molecules are the most compact of the three, offering the least available surface area for intermolecular contact and, hence, the weakest dispersion forces. This behavior is analogous to the connections that may be formed between strips of VELCRO brand fasteners: the greater the area of the strip’s contact, the stronger the connection.

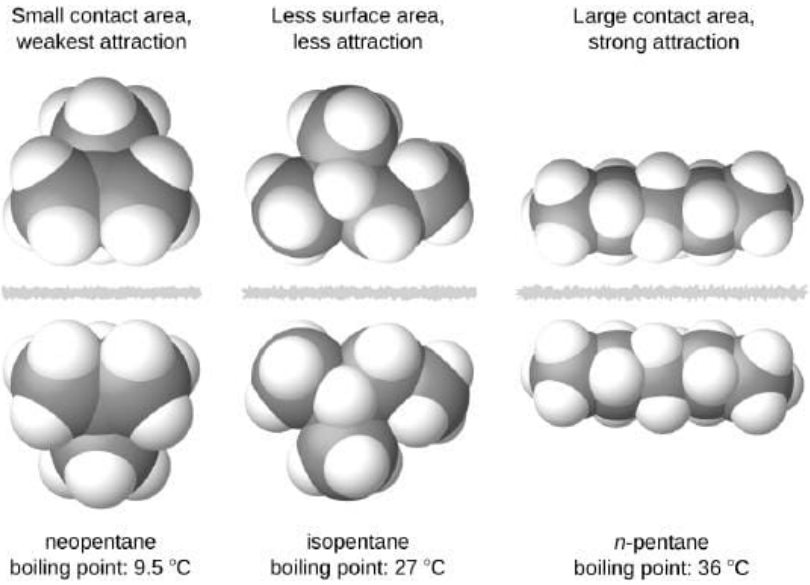



Figure 9.11.6: The strength of the dispersion forces increases with the contact area between molecules, as demonstrated by the boiling points of these pentane isomers. Three images of molecules are shown. The first shows a cluster of large, gray spheres each bonded together and to several smaller, white spheres. There is a gray, jagged line and then the mirror image of the first cluster of spheres is shown. Above these two clusters is the label, “Small contact area, weakest attraction,” and below is the label, “neopentane boiling point: 9.5 degrees C.” The second shows a chain of three gray spheres bonded by the middle sphere to a fourth gray sphere. Each gray sphere is bonded to several smaller, white spheres. There is a jagged, gray line and then the mirror image of the first chain appears. Above these two chains is the label, “Less surface area, less attraction,” and below is the label, “isopentane boiling point: 27 degrees C.” The third image shows a chain of five gray spheres bonded together and to several smaller, white spheres. There is a jagged gray line and then the mirror image of the first chain appears. Above these chains is the label, “Large contact area, strong attraction,” and below is the label, “n-pentane boiling point 36 degrees C.”

## Applications: Geckos and Intermolecular Forces

Geckos have an amazing ability to adhere to most surfaces. They can quickly run up smooth walls and across ceilings that have no toe-holds, and they do this without having suction cups or a sticky substance on their toes. And while a gecko can lift its feet easily as it walks along a surface, if you attempt to pick it up, it sticks to the surface. How are geckos (as well as spiders and some other insects) able to do this? Although this phenomenon has been investigated for hundreds of years, scientists only recently uncovered the details of the process that allows geckos’ feet to behave this way.

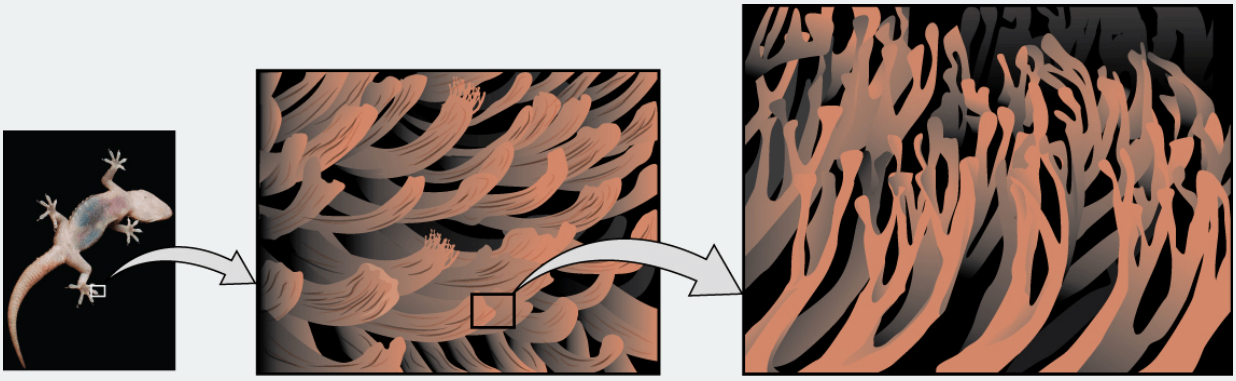



Figure 9.11.7: Geckos’ toes contain large numbers of tiny hairs (setae), which branch into many triangular tips (spatulae). Geckos adhere to surfaces because of van der Waals attractions between the surface and a gecko’s millions of spatulae. By changing how the spatulae contact the surface, geckos can turn their stickiness “on” and “off.” (credit photo: modification of work by “JC*+A!”/Flickr) Three figures are shown. The first is a photo of the bottom of a gecko’s foot. The second is bigger version which shows the setae. The third is a bigger version of the setae and shows the spatulae.

Geckos' toes are covered with hundreds of thousands of tiny hairs known as *setae*, with each seta, in turn, branching into hundreds of tiny, flat, triangular tips called *spatulae*. The huge numbers of spatulae on its setae provide a gecko, shown in Figure 9.11.7, with a large total surface area for sticking to a surface. In 2000, Kellar Autumn, who leads a multi-institutional gecko research team, found that geckos adhered equally well to both polar silicon dioxide and nonpolar gallium arsenide. This proved that geckos stick to surfaces because of dispersion forces—weak intermolecular attractions arising from temporary, synchronized charge distributions between adjacent molecules. Although dispersion forces are very weak, the total attraction over millions of spatulae is large enough to support many times the gecko's weight.

In 2014, two scientists developed a model to explain how geckos can rapidly transition from "sticky" to "non-sticky." Alex Greaney and Congcong Hu at Oregon State University described how geckos can achieve this by changing the angle between their spatulae and the surface. Geckos' feet, which are normally nonsticky, become sticky when a small shear force is applied. By curling and uncurling their toes, geckos can alternate between sticking and unsticking from a surface, and thus easily move across it. Further investigations may eventually lead to the development of better adhesives and other applications.

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

*Watch this video to learn more about Kellar Autumn's research that determined that van der Waals forces are responsible for a gecko's ability to cling and climb.*

## Dipole-Dipole Attractions

Recall from the chapter on chemical bonding and molecular geometry that *polar* molecules have a partial positive charge on one side and a partial negative charge on the other side of the molecule—a separation of charge called a *dipole*. Consider a polar molecule such as hydrogen chloride, HCl. In the HCl molecule, the more electronegative Cl atom bears the partial negative charge, whereas the less electronegative H atom bears the partial positive charge. An attractive force between HCl molecules results from the attraction between the positive end of one HCl molecule and the negative end of another. This attractive force is called a dipole-dipole attraction—the electrostatic force between the partially positive end of one polar molecule and the partially negative end of another, as illustrated in Figure 9.11.8.

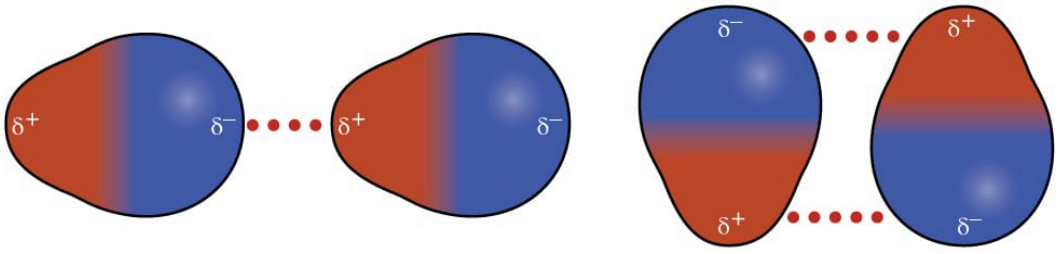



Figure 9.11.8: This image shows two arrangements of polar molecules, such as HCl, that allow an attraction between the partial negative end of one molecule and the partial positive end of another. Two pairs of molecules are shown where each molecule has one larger blue side labeled "delta sign, negative sign" and a smaller red side labeled "delta sign, positive sign. In the first pair, the red sides of the two molecules both face to the left and the blue side to the right. A horizontal dotted line lies in between the two. In the second pair, the molecules face up and down, with the red and blue ends aligning. A horizontal dotted line lies between the red and blue ends facing upward and another lies between the red and blue ends facing downward.

The effect of a dipole-dipole attraction is apparent when we compare the properties of HCl molecules to nonpolar $F_2$ molecules. Both HCl and $F_2$ consist of the same number of atoms and have approximately the same molecular mass. At a temperature of 150 K, molecules of both substances would have the same average KE. However, the dipole-dipole attractions between HCl molecules are sufficient to cause them to "stick together" to form a liquid, whereas the relatively weaker dispersion forces between nonpolar $F_2$ molecules are not, and so this substance is gaseous at this temperature. The higher normal boiling point of HCl (188 K) compared to $F_2$ (85 K) is a reflection of the greater strength of dipole-dipole attractions between HCl molecules, compared

to the attractions between nonpolar $F_2$ molecules. We will often use values such as boiling or freezing points, or enthalpies of vaporization or fusion, as indicators of the relative strengths of IMFs of attraction present within different substances.

### Example 9.11.2: Dipole-Dipole Forces and Their Effects

Predict which will have the higher boiling point: $N_2$ or CO. Explain your reasoning.

**Solution**

CO and $N_2$ are both diatomic molecules with masses of about 28 amu, so they experience similar London dispersion forces. Because CO is a polar molecule, it experiences dipole-dipole attractions. Because $N_2$ is nonpolar, its molecules cannot exhibit dipole-dipole attractions. The dipole-dipole attractions between CO molecules are comparably stronger than the dispersion forces between nonpolar $N_2$ molecules, so CO is expected to have the higher boiling point.

A common method for preparing oxygen is the decomposition

### Exercise 9.11.2

Predict which will have the higher boiling point: $\mathrm{ICl}$ or $\mathrm{Br_2}$. Explain your reasoning.

*Answer available at the end of this section.*

## Hydrogen Bonding

Nitrosyl fluoride (ONF, molecular mass 49 amu) is a gas at room temperature. Water ($H_2O$, molecular mass 18 amu) is a liquid, even though it has a lower molecular mass. We clearly cannot attribute this difference between the two compounds to dispersion forces. Both molecules have about the same shape and ONF is the heavier and larger molecule. It is, therefore, expected to experience more significant dispersion forces. Additionally, we cannot attribute this difference in boiling points to differences in the dipole moments of the molecules. Both molecules are polar and exhibit comparable dipole moments. The large difference between the boiling points is due to a particularly strong dipole-dipole attraction that may occur when a molecule contains a hydrogen atom bonded to a fluorine, oxygen, or nitrogen atom (the three most electronegative elements). The very large difference in electronegativity between the H atom (2.1) and the atom to which it is bonded (4.0 for an F atom, 3.5 for an O atom, or 3.0 for a N atom), combined with the very small size of a H atom and the relatively small sizes of F, O, or N atoms, leads to *highly concentrated partial charges* with these atoms. Molecules with F-H, O-H, or N-H moieties are very strongly attracted to similar moieties in nearby molecules, a particularly strong type of dipole-dipole attraction called hydrogen bonding. Examples of hydrogen bonds include HF···HF, $H_2O$···HOH, and $H_3N$···$HNH_2$, in which the hydrogen bonds are denoted by dots. Figure 9.11.9 illustrates hydrogen bonding between water molecules.

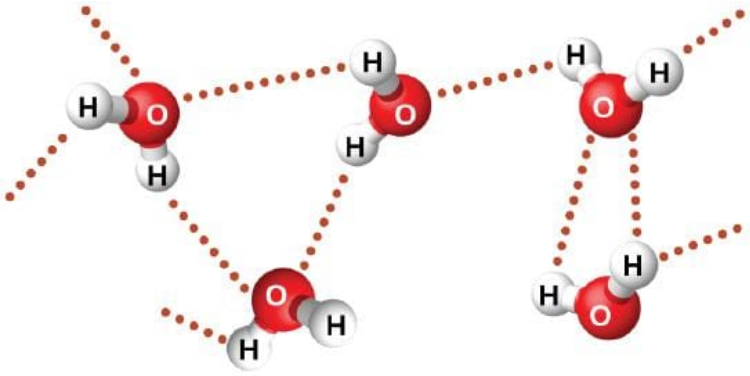



Figure 9.11.9: Water molecules participate in multiple hydrogen-bonding interactions with nearby water molecules. Five water molecules are shown near one another, but not touching. A dotted line lies between many of the hydrogen atoms on one molecule and the oxygen atom on another molecule.

Despite use of the word “bond,” keep in mind that hydrogen bonds are *intermolecular* attractive forces, not *intramolecular*

attractive forces (covalent bonds). Hydrogen bonds are much weaker than covalent bonds, only about 5 to 10% as strong, but are generally much stronger than other dipole-dipole attractions and dispersion forces.

Hydrogen bonds have a pronounced effect on the properties of condensed phases (liquids and solids). For example, consider the trends in boiling points for the binary hydrides of group 15 ($NH_3$, $PH_3$, $AsH_3$, and $SbH_3$), group 16 hydrides ($H_2O$, $H_2S$, $H_2Se$, and $H_2Te$), and group 17 hydrides (HF, HCl, HBr, and HI). The boiling points of the heaviest three hydrides for each group are plotted in Figure 9.11.10. As we progress down any of these groups, the polarities of the molecules decrease slightly, whereas the sizes of the molecules increase substantially. The effect of increasingly stronger dispersion forces dominates that of increasingly weaker dipole-dipole attractions, and the boiling points are observed to increase steadily. For the group 15, 16, and 17 hydrides, the boiling points for each class of compounds increase with increasing molecular mass for elements in periods 3, 4, and 5.

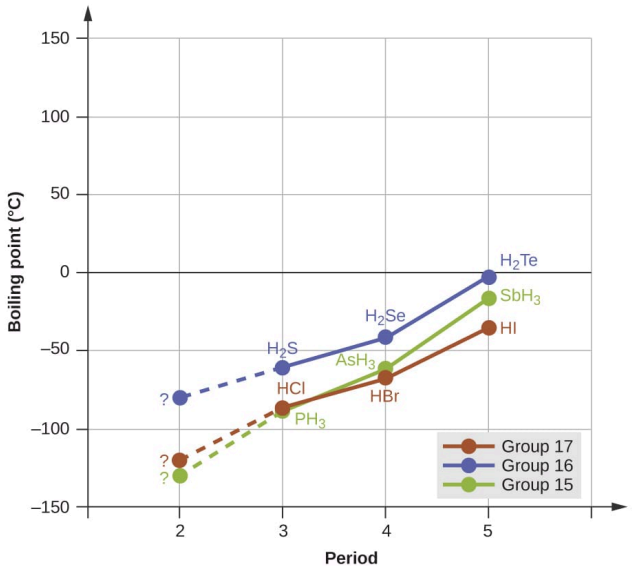



Figure 9.11.10: In comparison to periods 3–5, the binary hydrides of period 2 elements in groups 17, 16 and 15 (F, O and N, respectively) exhibit anomalously high boiling points due to hydrogen bonding. A line graph is shown where the y-axis is labeled “Boiling point (, degree sign, C )” and has values of “ negative 150” to “150” from bottom to top in increments of 50. The x-axis is labeled “Period” and has values of “0” to “5” in increments of 1. Three lines are shown on the graph and are labeled in the legend. The red line is labeled as “halogen family,” the blue is “oxygen family” and the green is “nitrogen family.” The first point on the red line is labeled “question mark” and is at point “2, negative 120”. The second point on the line is labeled “H C l” and is at point “3, negative 80” while the third point on the line is labeled “H B r” and is at point “4, negative 60”. The fourth point on the line is labeled “H I” and is at point “5, negative 40.” The first point on the green line is labeled “question mark” and is at point “2, negative 125.” The second point on the line is labeled “P H, subscript 3” and is at point “3, negative 80” while the third point on the line is labeled “A s H, subscript 3” and is at point “4, negative 55.” The fourth point on the line is labeled “S b H, subscript 3” and is at point “5, negative 10.” The first point on the blue line is labeled “question mark” and is at point “2, negative 80.” The second point on the line is labeled “H, subscript 2, S” and is at point “3, negative 55” while the third point on the line is labeled “H, subscript 2, S e” and is at point “4, negative 45.” The fourth point on the line is labeled “H, subscript 2, T e” and is at point “5, negative 3.”

If we use this trend to predict the boiling points for the lightest hydride for each group, we would expect $NH_3$ to boil at about −120 °C, $H_2O$ to boil at about −80 °C, and HF to boil at about −110 °C. However, when we measure the boiling points for these compounds, we find that they are dramatically higher than the trends would predict, as shown in Figure 9.11.10. The stark contrast between our naïve predictions and reality provides compelling evidence for the strength of hydrogen bonding.

## Example 9.11.3: Effect of Hydrogen Bonding on Boiling Points

Consider the compounds dimethylether ($CH_3OCH_3$), ethanol ($CH_3CH_2OH$), and propane ($CH_3CH_2CH_3$). Their boiling points, not necessarily in order, are −42.1 °C, −24.8 °C, and 78.4 °C. Match each compound with its boiling point. Explain your reasoning.

**Solution**

The VSEPR-predicted shapes of $CH_3OCH_3$, $CH_3CH_2OH$, and $CH_3CH_2CH_3$ are similar, as are their molar masses (46 g/mol, 46 g/mol, and 44 g/mol, respectively), so they will exhibit similar dispersion forces. Since $CH_3CH_2CH_3$ is nonpolar, it may exhibit *only* dispersion forces. Because $CH_3OCH_3$ is polar, it will also experience dipole-dipole attractions. Finally, $CH_3CH_2OH$ has an –OH group, and so it will experience the uniquely strong dipole-dipole attraction known as hydrogen bonding. So the ordering in terms of strength of IMFs, and thus boiling points, is $CH_3CH_2CH_3 < CH_3OCH_3 < CH_3CH_2OH$. The boiling point of propane is –42.1 °C, the boiling point of dimethylether is –24.8 °C, and the boiling point of ethanol is 78.5 °C.

**Exercise 9.11.3**

Ethane ($CH_3CH_3$) has a melting point of –183 °C and a boiling point of –89 °C. Predict the melting and boiling points for methylamine ($CH_3NH_2$). Explain your reasoning.

*Answer available at the end of this section.*

## Hydrogen Bonding and DNA

Deoxyribonucleic acid (DNA) is found in every living organism and contains the genetic information that determines the organism's characteristics, provides the blueprint for making the proteins necessary for life, and serves as a template to pass this information on to the organism's offspring. A DNA molecule consists of two (anti-)parallel chains of repeating nucleotides, which form its well-known double helical structure, as shown in Figure 9.11.10.

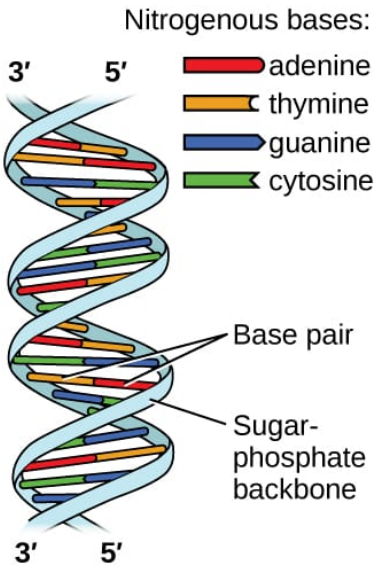



Figure 9.11.11: Two separate DNA molecules form a double-stranded helix in which the molecules are held together via hydrogen bonding. (credit: modification of work by Jerome Walker, Dennis Myts) Two images are shown. The first lies on the left side of the page and shows a helical structure like a twisted ladder where the rungs of the ladder, labeled "Base pair" are red, yellow, green and blue paired bars. The red and yellow bars, which are always paired together, are labeled in the legend, which is titled "Nitrogenous bases" as "adenine" and "thymine," respectively. The blue and green bars, which are always paired together, are labeled in the legend as "guanine" and "cytosine," respectively. At the top of the helical structure, the left-hand side rail, or "Sugar, dash, phosphate backbone," is labeled as "3, prime" while the right is labeled as "5, prime." These labels are reversed at the bottom of the helix. To the right of the page is a large Lewis structure. The top left corner of this structure, labeled "5, prime," shows a phosphorus atom single bonded to three oxygen atoms, one of which has a superscripted negative charge, and double bonded to a fourth oxygen atom. One of the single bonded oxygen atoms is single bonded to the left corner of a five-membered ring with an oxygen atom at its top point and which is single bonded to an oxygen atom on the bottom left. This oxygen atom is single bonded to a phosphorus atom that is single bonded to two other hydrogen atoms and double bonded to a fourth oxygen atom. The lower left of these oxygen atoms is single bonded to another oxygen atom that is single bonded to a five-membered ring with an oxygen in the upper bonding site. The bottom left of this ring has a hydroxyl group attached to it while the upper right carbon is single bonded to a nitrogen atom that is part of a five-membered ring bonded to a six-membered ring. Both of these rings have points of unsaturation and nitrogen atoms bonded into their structures.

On the right side of the six-membered ring are two single bonded amine groups and a double bonded oxygen. Three separate dotted lines extend from these sites to corresponding sites on a second six-membered ring. This ring has points of unsaturation and a nitrogen atom in the bottom right bonding position that is single bonded to a five-membered ring on the right side of the image. This ring is single bonded to a carbon that is single bonded to an oxygen that is single bonded to a phosphorus. The phosphorus is single bonded to two other oxygen atoms and double bonded to a fourth oxygen atom. This group is labeled “5, prime.” The five-membered ring is also bonded on the top side to an oxygen that is bonded to a phosphorus single bonded to two other oxygen atoms and double bonded to a fourth oxygen atom. The upper left oxygen of this group is single bonded to a carbon that is single bonded to a five-membered ring with an oxygen in the bottom bonding position. This ring has a hydroxyl group on its upper right side that is labeled “3, prime” and is bonded on the left side to a nitrogen that is a member of a five-membered ring. This ring is bonded to a six-membered ring and both have points of unsaturation. This ring has a nitrogen on the left side, as well as an amine group, that have two dotted lines leading from them to an oxygen and amine group on a six membered ring. These dotted lines are labeled “Hydrogen bonds.” The six membered ring also has a double bonded oxygen on its lower side and a nitrogen atom on its left side that is single bonded to a five-membered ring. This ring connects to the two phosphate groups mentioned at the start of this to form a large circle. The name “guanine” is written below the lower left side of this image while the name “cytosine” is written on the lower right. The name “thymine” is written above the right side of the image and “adenine” is written on the top right. Three sections are indicated below the images where the left is labeled “Sugar, dash, phosphate backbone,” the middle is labeled “Bases” and the right is labeled “Sugar, dash, phosphate backbone.”

Each nucleotide contains a (deoxyribose) sugar bound to a phosphate group on one side, and one of four nitrogenous bases on the other. Two of the bases, cytosine (C) and thymine (T), are single-ringed structures known as pyrimidines. The other two, adenine (A) and guanine (G), are double-ringed structures called purines. These bases form complementary base pairs consisting of one purine and one pyrimidine, with adenine pairing with thymine, and cytosine with guanine. Each base pair is held together by hydrogen bonding. A and T share two hydrogen bonds, C and G share three, and both pairings have a similar shape and structure Figure 9.11.12

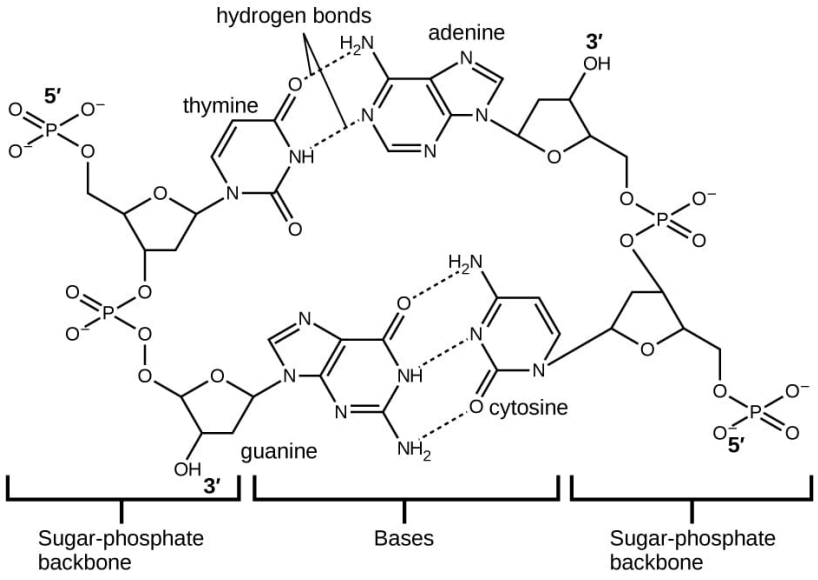



Figure 9.11.12: The geometries of the base molecules result in maximum hydrogen bonding between adenine and thymine (AT) and between guanine and cytosine (GC), so-called “complementary base pairs.” A large Lewis structure is shown. The top left corner of this structure, labeled “5, prime,” shows a phosphorus atom single bonded to three oxygen atoms, one of which has a superscripted negative charge, and double bonded to a fourth oxygen atom. One of the single bonded oxygen atoms is single bonded to the left corner of a five-membered ring with an oxygen atom at its top point and which is single bonded to an oxygen atom on the bottom left. This oxygen atom is single bonded to a phosphorus atom that is single bonded to two other hydrogen atoms and double bonded to a fourth oxygen atom. The lower left of these oxygen atoms is single bonded to another oxygen atom that is single bonded to a five-membered ring with an oxygen in the upper bonding site. The bottom left of this ring has a hydroxyl group attached to it while the upper right carbon is single bonded to a nitrogen atom that is part of a five-membered ring bonded to a six-membered ring. Both of these rings have points of unsaturation and nitrogen atoms bonded into their structures. On the right side of the six-membered ring are two single bonded amine groups and a double bonded oxygen. Three separate dotted lines extend from these sites to corresponding sites on a second six-membered ring. This ring has points of unsaturation and a nitrogen atom in the bottom right bonding position that is single bonded to a five-membered ring on the right side of the image. This ring is single bonded to a carbon that is single bonded to an oxygen that is single bonded to a phosphorus. The phosphorus is single bonded to two

other oxygen atoms and double bonded to a fourth oxygen atom. This group is labeled “5, prime.” The five-membered ring is also bonded on the top side to an oxygen that is bonded to a phosphorus single bonded to two other oxygen atoms and double bonded to a fourth oxygen atom. The upper left oxygen of this group is single bonded to a carbon that is single bonded to a five-membered ring with an oxygen in the bottom bonding position. This ring has a hydroxyl group on its upper right side that is labeled “3, prime” and is bonded on the left side to a nitrogen that is a member of a five-membered ring. This ring is bonded to a six-membered ring and both have points of unsaturation. This ring has a nitrogen on the left side, as well as an amine group, that have two dotted lines leading from them to an oxygen and amine group on a six membered ring. These dotted lines are labeled “Hydrogen bonds.” The six membered ring also has a double bonded oxygen on its lower side and a nitrogen atom on its left side that is single bonded to a five-membered ring. This ring connects to the two phosphate groups mentioned at the start of this to form a large circle. The name “guanine” is written below the lower left side of this image while the name “cytosine” is written on the lower right. The name “thymine” is written above the right side of the image and “adenine” is written on the top right. Three sections are indicated below the images where the left is labeled “Sugar, dash, phosphate backbone,” the middle is labeled “Bases” and the right is labeled “Sugar, dash, phosphate backbone.”

The cumulative effect of millions of hydrogen bonds effectively holds the two strands of DNA together. Importantly, the two strands of DNA can relatively easily “unzip” down the middle since hydrogen bonds are relatively weak compared to the covalent bonds that hold the atoms of the individual DNA molecules together. This allows both strands to function as a template for replication.

## Summary

The physical properties of condensed matter (liquids and solids) can be explained in terms of the kinetic molecular theory. In a liquid, intermolecular attractive forces hold the molecules in contact, although they still have sufficient kinetic energy to move past each other. Intermolecular attractive forces, collectively referred to as van der Waals forces, are responsible for the behavior of liquids and solids and are electrostatic in nature. Dipole-dipole attractions result from the electrostatic attraction of the partial negative end of one dipolar molecule for the partial positive end of another. The temporary dipole that results from the motion of the electrons in an atom can induce a dipole in an adjacent atom and give rise to the London dispersion force. London forces increase with increasing molecular size. Hydrogen bonds are a special type of dipole-dipole attraction that results when hydrogen is bonded to one of the three most electronegative elements: F, O, or N.

## Glossary

**dipole-dipole attraction**
: intermolecular attraction between two permanent dipoles

**dispersion force**
: (also, London dispersion force) attraction between two rapidly fluctuating, temporary dipoles; significant only when particles are very close together

**hydrogen bonding**
: occurs when exceptionally strong dipoles attract; bonding that exists when hydrogen is bonded to one of the three most electronegative elements: F, O, or N

**induced dipole**
: temporary dipole formed when the electrons of an atom or molecule are distorted by the instantaneous dipole of a neighboring atom or molecule

**instantaneous dipole**
: temporary dipole that occurs for a brief moment in time when the electrons of an atom or molecule are distributed asymmetrically

**intermolecular force**
: noncovalent attractive force between atoms, molecules, and/or ions

**polarizability**
: measure of the ability of a charge to distort a molecule’s charge distribution (electron cloud)

**van der Waals force**
: attractive or repulsive force between molecules, including dipole-dipole, dipole-induced dipole, and London dispersion forces; does not include forces due to covalent or ionic bonding, or the attraction between ions and

molecules

## Answers

**Exercise** 9.11.1

All of these compounds are nonpolar and only have London dispersion forces: the larger the molecule, the larger the dispersion forces and the higher the boiling point. The ordering from lowest to highest boiling point is therefore

$$C_2H_6 < C_3H_8 < C_4H_{10}.$$

**Exercise** 9.11.2

ICl. ICl and $Br_2$ have similar masses (~160 amu) and therefore experience similar London dispersion forces. ICl is polar and thus also exhibits dipole-dipole attractions; $Br_2$ is nonpolar and does not. The relatively stronger dipole-dipole attractions require more energy to overcome, so ICl will have the higher boiling point.

**Exercise** 9.11.3

The melting point and boiling point for methylamine are predicted to be significantly greater than those of ethane. $CH_3CH_3$ and $CH_3NH_2$ are similar in size and mass, but methylamine possesses an –NH group and therefore may exhibit hydrogen bonding. This greatly increases its IMFs, and therefore its melting and boiling points. It is difficult to predict values, but the known values are a melting point of –93 °C and a boiling point of –6 °C.

This page titled 9.11: Intermolecular Forces (1) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10: MODULE 8- CHEMICAL REACTIONS AND ORGANIC CHEMISTRY


10.1: Introduction to Chemical Changes

10.2: Aspects of a Chemical Change

10.3: Chemical Equations- Representing Chemical Changes

10.4: Formula Mass and the Mole Concept

10.5: Balancing Equations

10.6: Energy and Chemical Reactions

10.7: Classifying Reactions

10.8: Acid Base Reactions

10.9: Polymerization Reactions

10.10: Determining Empirical and Molecular Formulas

10.11: Prelude to Stoichiometry

10.12: Balancing Equations (1)

10.13: Writing and Balancing Chemical Equations

10.14: Prelude to Kinetics

10.15: Chemical Reaction Rates

10.16: Brønsted-Lowry Acids and Bases

10.17: pH and pOH

**10.18: Organic Chemistry**

10.18.1: Introduction

10.18.2: Hydrocarbons

10.18.3: Alcohols and Ethers

10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters

10.18.5: Amines and Amides



This page titled 10: Module 8- Chemical Reactions and Organic Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Arunava Roy.

# 10.1: Introduction to Chemical Changes



Figure 10.1.1 (Credit: Jon Sullivan; Source; License: Public Domain)

## Do you like to cook?

Cooking is a valuable skill that can benefit everyone. Whether it is fixing a simple grilled cheese sandwich or preparing an elaborate meal, cooking demonstrates some basic concepts of chemistry. When you bake bread, you mix flour, sugar, yeast, and water together. After baking, this mixture has been changed to form bread, another substance that has different characteristics and qualities from the original materials. The process of baking has produced chemical changes in the ingredients that result in bread being made.

## Chemical Change

Most of the elements we know about do not exist freely in nature. Sodium cannot be found by itself (unless we prepare it in the laboratory) because it interacts too easily with other materials. On the other hand, the element helium does not interact with other elements to any extent. We can isolate helium from natural gas during the process of drilling for oil.

A chemical change produces different materials than the ones we started with. One aspect of the science of chemistry is the study of the changes that matter undergoes. If chemistry was simply a study of elements that did nothing, life would be very boring (in fact, life would not exist since the elements are what make up our bodies and sustain us). But the processes of change that take place when different chemicals are combined produce all the materials that we use daily.

One type of chemical change (already mentioned) is when two elements combine to form a compound. Another type involves the breakdown of a compound to produce the elements that make it up. If we pass an electric current through bauxite (aluminum oxide, the raw material for aluminum metal), we get metallic aluminum as a product.



Figure 10.1.1: Electrolytic production of aluminum. (Credit: Ronny Jacques; Source; License: Public Domain)

However, the vast majority of chemical changes involve one compound being transformed into another compound. There are literally millions of possibilities when we take this approach to chemical change. New compounds can be made to produce better fabrics that are easier to clean and maintain; they can help preserve food so it doesn't spoil as quickly; we can make new medicines to treat diseases; even products, such as airbags (as seen in the simulation below), use a chemical change to improve the world around us. In this chapter we will learn more about these chemical changes: how to keep track of them with equations, how to balance those equations, how to categorize them, and how to think about the energy involved with them. We will end with some important examples of reaction types.

This page titled 10.1: Introduction to Chemical Changes is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 8.10: The Hydrogen Atom (Problems Q8.1.7) by OpenStax is licensed CC BY 4.0.

# 10.2: Aspects of a Chemical Change

**Learning Objectives**

- Identify chemical properties and chemical changes.
- Determine if a chemical change has occurred based on observations.

## Have you ever left your bicycle out in the rain?

It is not a good idea because the rain can cause the bicycle to rust. You start to get a reddish-orange build-up of a rough deposit on the metal. It may start with the chain, but can spread to other parts of the bicycle, especially if there are scratches that create a bare metal surface. The formation of rust is a chemical process that takes place when iron is exposed to water and oxygen. It is estimated that damage due to rust costs U.S. businesses, military, and government over 276 billion dollars a year—a very expensive chemical process.

## Chemical Properties

A **chemical property** describes the ability of a substance to undergo a specific chemical change. A chemical property of iron is that it is capable of combining with oxygen to form iron oxide, the chemical name of rust. The more general term for rusting and other similar processes is corrosion. Other terms that are commonly used to describe chemical changes are: burn, rot, explode, decompose, and ferment. Chemical properties are very useful in identifying substances. However, unlike physical properties, chemical properties can only be observed as the substance is in the process of being changed into a different substance.

## Chemical Change

A **chemical change** is also called a chemical reaction. A chemical reaction is a process that occurs when one or more substances are changed into one or more new substances. Zinc $(\mathrm{Zn})$ is a silver gray element that can be ground into a powder. If zinc is mixed at room temperature with powdered sulfur $(\mathrm{S})$, a bright yellow element, the result will simply be a mixture of zinc and sulfur. No chemical reaction occurs. However, if energy is provided to the mixture in the form of heat, the zinc will chemically react with the sulfur to form the compound zinc sulfide $(\mathrm{ZnS})$. The figure below shows the substances involved in this reaction.

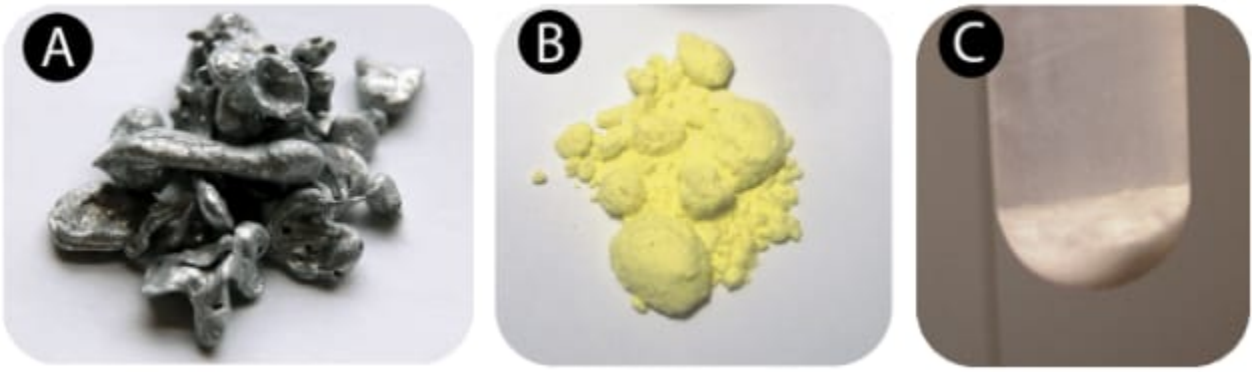



Figure $10.2.1$: Zinc and sulfur are two elements that undergo a chemical reaction when heated to form the compound zinc sulfide. (Credit: (A) Ben Mills (Benjah-bmm27); (B) User:Vineyard/Wikimedia Commons; (C) Courtesy of A. Eisen/NOAA; Source: (A); (B); (C); License: (A) Public Domain; (B) CC by 3.0(opens in new window); (C) Public Domain

## Recognizing Chemical Reactions

How can a scientist tell if a **chemical reaction** is taking place? There are four visual clues that indicate that a chemical reaction

is likely occurring:

1. A change in color occurs during the reaction.
2. A **gas** is produced during the reaction.
3. A solid product called a **precipitate** is produced in the reaction.
4. A transfer of energy occurs as a result of the reaction.

Mercury (II) oxide is a red solid. When it is heated to a temperature above $500^{\circ}\mathrm{C}$, it easily decomposes into mercury and oxygen gas. The red color of the mercury oxide reactant becomes the silver color of mercury. The color change is a sign that the reaction is occurring.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video $10.2.1$: Mercuric oxide decomposing into mercury and oxygen (Source)

When zinc reacts with hydrochloric acid, the reaction bubbles vigorously as hydrogen gas is produced. The production of a gas is also an indication that a chemical reaction is occurring.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video $10.2.2$: Zinc reacting with hydrochloric acid produces bubbles of hydrogen gas. (Source)

When a colorless solution of lead (II) nitrate is added to a colorless solution of potassium iodide, a yellow solid called a precipitate is instantly produced. A precipitate is a solid product that forms from a reaction and settles out of a liquid solution. The formation of a precipitate is an indication of a chemical reaction.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video$10.2.3$: A yellow precipitate of solid lead (II) iodide forms immediately when solutions of lead (II) nitrate and potassium iodide are mixed. (Source)

All chemical changes involve a transfer of energy. When zinc reacts with hydrochloric acid, the test tube becomes very warm as energy is released during the reaction. Some other reactions absorb energy. While energy changes are a potential sign of a

CK-12 LICENSED UNDER CC BY NC SA
© CK-12 Foundation • Visit us at ck12.org

chemical reaction, care must be taken to ensure that a chemical reaction is indeed taking place. Physical changes also involve a transfer of energy. The melting of a solid absorbs energy, while the condensation of a gas releases energy. The only way to be certain that a chemical reaction has taken place is to test the composition of the substances after the change has taken place, to see if they are different from the starting substances.

The following videos show some additional examples of chemical changes occurring.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- A chemical property describes the ability of a substance to undergo a specific chemical change.
- A chemical reaction is a process that occurs when one or more substances are changed into one or more new substances.
- There are four visual clues that indicate that a chemical reaction is likely occurring:
  - A change of color occurs during the reaction.
  - A gas is produced during the reaction.
  - A solid product called a precipitate is produced in the reaction.
  - A transfer of energy occurs as a result of the reaction.

## Glossary

**chemical property**
: Describes the ability of a substance to undergo a specific chemical change.

**chemical change (chemical reaction)**
: A process that occurs when one or more substances are changed into one or more new substances.

**precipitate**
: A solid product that occurs in some reactions.

This page titled 10.2: Aspects of a Chemical Change is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 8.12: The Hydrogen Atom (Problems Q8.1.9) by OpenStax is licensed CC BY 4.0.
- 8.14: The Hydrogen Atom (Problems Q8.1.11) by OpenStax is licensed CC BY 4.0.
- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.3: Chemical Equations- Representing Chemical Changes

**Learning Objectives**

- Represent chemical changes using both words and symbols.

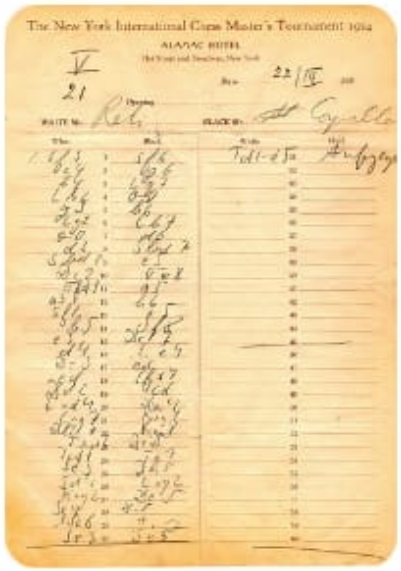



Figure 10.3.1 (Author Unknown; Source: Commons Wikimedia, Capablanca(opens in new window) [commons.wikimedia.org]; License: Public Domain)

## How do chess players monitor their moves in a game?

Suppose you were walking along and noticed a piece of paper on the ground with markings on it. You pick it up and see the paper in the picture above.

To most people, these notes are meaningless (maybe they're a secret spy code). But to a chess player, these symbols tell the story of a chess game. Each abbreviation describes a chess piece or a move during the game. The use of special symbols allows chess players to "see" the game without having to read a wordy and possibly incomplete description of what happened. We will learn similar tactics for analyzing the results of chemical equations.

## Chemical Symbols and Formulas

In order to illustrate chemical reactions and the elements and **compounds** involved in them, chemists use symbols and formulas. A **chemical symbol** is a one- or two-letter designation of an element. Compounds are combinations of two or more elements. A **chemical formula** is an expression that shows the elements in a compound and the relative proportions of those elements. Please review this topic before moving forward, as it will be important for understanding what comes next.

## Chemical Equation

The reaction between zinc and sulfur can be shown in what is called a **chemical equation**. In words, we could write the reaction as:

$$\text{zinc} + \text{sulfur} \rightarrow \text{zinc sulfide}$$

The more convenient way to express a chemical reaction is to use the symbols and formulas of the substances involved:

$$\mathrm{Zn} + \mathrm{S} \rightarrow \mathrm{ZnS}$$

The substance(s) to the left of the arrow in a chemical equation are called reactants. A **reactant** is a substance that is present at the start of a chemical reaction. The substance(s) to the right of the arrow are called **products**. A product is a substance that is present at the end of a chemical reaction. In the equation above, the zinc and sulfur are the reactants that chemically combine to form the zinc sulfide product.

There is a standard way of writing chemical equations. The reactants are all written on the left-hand side of the equation, with the products on the right-hand side. Unlike a math equation, a chemical equation does not use an equal sign. Instead, the arrow is called a yield sign and so the equation is described as "reactants yield products".

$$\text{reactants} \rightarrow \text{products}$$

The presence of the arrow also indicates that the reaction goes in one direction under the conditions indicated. There are reactions which can be easily reversed, but we will not take those up right now.

There are a wide variety of reactions possible: elements may form compounds (as seen in the reaction above), compounds may form elements (water will break down in the presence of an electric current to form hydrogen gas and oxygen gas) or compounds may combine, break apart, or rearrange to form new materials.

It is often important to know the physical states of the reactants and products taking part in a reaction. To do this, put the appropriate symbol in parentheses after each formula: $(s)$ for solid, $(l)$ for liquid, $(g)$ for gas, and $(aq)$ for an aqueous (water-based) solution. The previous reaction becomes:

$$\mathrm{CH_4}\,(g) + \mathrm{O_2}\,(g) \rightarrow \mathrm{CO_2}\,(g) + \mathrm{H_2O}\,(l)$$

The table below shows a listing of symbols used in chemical equations. Some, such as the double arrow which represents equilibrium, and the use of a catalyst in a reaction, will be detailed in other concepts.

*Table* 10.3.1*: Symbols Used in Chemical Equations*

| Symbols | Description |
|---|---|
| $+$ | Used to separate multiple reactants or products. |
| $\rightarrow$ | Yield sign; separates reactants from products. |
| $\rightleftharpoons$ | Replaces the yield sign for reversible reactions that reach equilibrium. |
| $(s)$ | Reactant or product in the solid state. |
| $(l)$ | Reactant or product in the liquid state. |
| $(g)$ | Reactant or product in the gaseous state. |
| $(aq)$ | Reactant or product in an aqueous solution (dissolved in water). |
| $\xrightarrow{\mathrm{Pt}}$ | Formula written above the arrow is used as a catalyst in the reaction. |
| $\xrightarrow{\Delta}$ | Triangle indicates that the reaction is being heated. |

## Word Equations

You can describe a chemical reaction by writing a **word equation**. When silver metal is exposed to sulfur, it reacts to form silver sulfide. Silver sulfide is commonly known as tarnish and turns the surface of silver objects dark and streaky black (see figure below). The sulfur that contributes to tarnish can come from traces of sulfur in the air, or from foods such as eggs. The word equation for the process is:

$$\text{Silver} + \text{sulfur} \rightarrow \text{Silver sulfide}$$

The silver and the sulfur are the reactants in the equation, while the silver sulfide is the product.

CK-12 LICENSED UNDER CC BY NC SA
© CK-12 Foundation • Visit us at ck12.org



Figure 10.3.2**:** You can see dark spots of tarnished (silver sulfide) forming on this ring as it reacts with sulfur compounds in the air. (Credit: CK-12 Foundation; Source: CK-12 Foundation; License: CK-12 Curriculum Materials license)

Another common chemical reaction is the burning of methane gas. Methane is the major component of natural gas and is commonly burned on a gas stove or in a Bunsen burner (figure below). Burning is a chemical reaction in which some type of fuel is reacted with oxygen gas. The products of the reaction in the burning of methane, as well as other fuels, are carbon dioxide and water. The word equation for this reaction is:

$$\text{Methane} + \text{oxygen} \rightarrow \text{carbon dioxide} + \text{water}$$

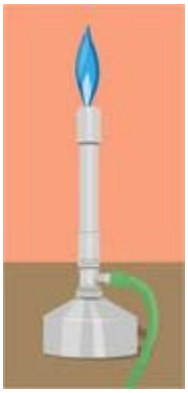


Figure 10.3.3: A Bunsen burner is commonly used to heat substances in a chemistry lab. Methane is reacted with oxygen to form carbon dioxide and water. (Credit: CK-12 Foundation; Source: CK-12 Foundation; License: CK-12 Curriculum Materials license)

Word equations can be very useful, but do have one major drawback—they cannot be used for any quantitative work. A word equation does not tell how many moles of each material are needed, or how many moles of product are formed. This is something we will consider in more detail in the next subsection.

The following video shows how cooking is a form of chemical reaction, as we stated at the introduction to this chapter.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

## Section Summary

- A chemical symbol is a one- or two-letter designation of an element.

- A chemical formula is an expression that shows the elements in a compound and the relative proportions of those elements.
- A chemical equation describes a chemical reaction.
- Reactants are starting materials and are written on the left-hand side of the equation.
- Products are the end result of the reaction and are written on the right-hand side of the equation.
- A chemical equation is representation a chemical reaction using chemical formulas reactants and products, as well as symbols.
- Symbols used in chemical equations are described and explained.
- Word equations are used to describe the conversion of reactants to products.

## Glossary

**product**
: A substance that is present at the end of a chemical reaction.

**reactant**
: A substance that is present at the start of a chemical reaction.

**chemical equation**
: The use of symbols and formulas to express a chemical change.

This page titled 10.3: Chemical Equations- Representing Chemical Changes is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 8.11: The Hydrogen Atom (Problems Q8.1.8) by OpenStax is licensed CC BY 4.0.
- 8.13: The Hydrogen Atom (Problems Q8.1.10) by OpenStax is licensed CC BY 4.0.
- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.4: Formula Mass and the Mole Concept

**Learning Objectives**

- Calculate formula masses for covalent and ionic compounds
- Define the amount unit mole and the related quantity Avogadro's number
- Explain the relation between mass, moles, and numbers of atoms or molecules, and perform calculations deriving these quantities from one another

We can argue that modern chemical science began when scientists started exploring the quantitative as well as the qualitative aspects of chemistry. For example, Dalton's atomic theory was an attempt to explain the results of measurements that allowed him to calculate the relative masses of elements combined in various compounds. Understanding the relationship between the masses of atoms and the chemical formulas of compounds allows us to quantitatively describe the composition of substances.

## Formula Mass

In an earlier chapter, we described the development of the atomic mass unit, the concept of average atomic masses, and the use of chemical formulas to represent the elemental makeup of substances. These ideas can be extended to calculate the formula mass of a substance by summing the average atomic masses of all the atoms represented in the substance's formula.

### Formula Mass for Covalent Substances

For covalent substances, the formula represents the numbers and types of atoms composing a single molecule of the substance; therefore, the formula mass may be correctly referred to as a molecular mass. Consider chloroform ($CHCl_3$), a covalent compound once used as a surgical anesthetic and now primarily used in the production of tetrafluoroethylene, the building block for the "anti-stick" polymer, Teflon. The molecular formula of chloroform indicates that a single molecule contains one carbon atom, one hydrogen atom, and three chlorine atoms. The average molecular mass of a chloroform molecule is therefore equal to the sum of the average atomic masses of these atoms. Figure 10.4.1 outlines the calculations used to derive the molecular mass of chloroform, which is 119.37 amu.

| Element | Quantity | | Average atomic mass (amu) | | Subtotal (amu) |
|---|---|---|---|---|---|
| C | 1 | × | 12.01 | = | 12.01 |
| H | 1 | × | 1.008 | = | 1.008 |
| Cl | 3 | × | 35.45 | = | 106.35 |
| | | | | **Molecular mass** | 119.37 |

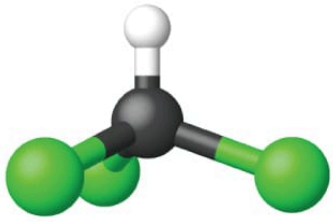
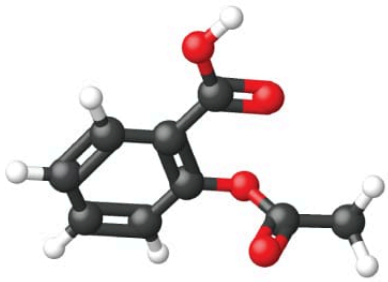


Figure 10.4.1: The average mass of a chloroform molecule, $CHCl_3$, is 119.37 amu, which is the sum of the average atomic masses of each of its constituent atoms. The model shows the molecular structure of chloroform. A table and diagram are shown. The table is made up of six columns and five rows. The header row reads: "Element," "Quantity," a blank space, "Average atomic mass (a m u)," a blank space, and "Subtotal (a m u)." The first column contains the symbols "C," "H," "C l" and a blank, merged cell that runs the width of the first five columns. The second column contains the numbers "1," "1," and "3" as well as the merged cell. The third column contains the multiplication symbol in each cell except for the last, merged cell. The fourth column contains the numbers "12.01," "1.008," and "35.45" as well as the merged cell. The fifth column contains the symbol "=" in each cell except for the last, merged cell. The sixth column contains the values "12.01," "1.008," "106.35," and "119.37." There is a thick black line below the number 106.35. The merged cell under the first five columns reads "Molecular mass." To the left of the table is a diagram of a molecule. Three green spheres are attached to a slightly smaller black sphere, which is also attached to a smaller white sphere. The green spheres lie beneath and to the sides of the black sphere while the white sphere is located straight up from the black

sphere.

Likewise, the molecular mass of an aspirin molecule, $C_9H_8O_4$, is the sum of the atomic masses of nine carbon atoms, eight hydrogen atoms, and four oxygen atoms, which amounts to 180.15 amu (Figure 10.4.2).

| Element | Quantity | | Average atomic mass (amu) | | Subtotal (amu) |
|---|---|---|---|---|---|
| C | 9 | × | 12.01 | = | 108.09 |
| H | 8 | × | 1.008 | = | 8.064 |
| O | 4 | × | 16.00 | = | 64.00 |
| | | | | Molecular mass | 180.15 |




| Element | Quantity | | Average atomic mass (amu) | | Subtotal (amu) |
|---|---|---|---|---|---|
| C | 9 | × | 12.01 | = | 108.09 |
| H | 8 | × | 1.008 | = | 8.064 |
| O | 4 | × | 16.00 | = | 64.00 |
| | | | | Molecular mass | 180.15 |

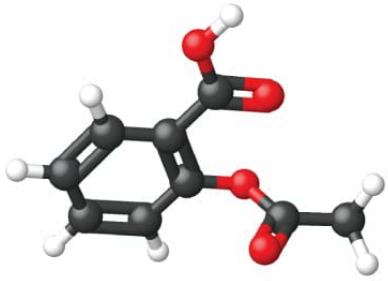



Figure 10.4.2: The average mass of an aspirin molecule is 180.15 amu. The model shows the molecular structure of aspirin, $C_9H_8O_4$. A table and diagram are shown. The table is made up of six columns and five rows. The header row reads: “Element,” “Quantity,” a blank space, “Average atomic mass (a m u),” a blank space, and “Subtotal (a m u).” The first column contains the symbols “C,” “H,” “O,” and a merged cell. The merged cell runs the length of the first five columns. The second column contains the numbers “9,” “8,” and “4” as well as the merged, cell. The third column contains the multiplication symbol in each cell except for the last, merged cell. The fourth column contains the numbers “12.01,” “1.008,” and “16.00” as well as the merged cell. The fifth column contains the symbol “=” in each cell except for the last, merged cell. The sixth column contains the values: “108.09,” “8.064,” “64.00,” and “180.15.” There is a thick black line below the number 64.00. The merged cell under the first five columns reads “Molecular mass.” To the left of the table is a diagram of a molecule. Six black spheres are located in a six-sided ring and connected by alternating double and single black bonds. Attached to each of the four black spheres is one smaller white sphere. Attached to the farthest right black sphere is a red sphere, connected to two more black spheres, all in a row. Attached to the last black sphere of that row are two more white spheres. Attached to the first black sphere of that row is another red sphere. A black sphere, attached to two red spheres and a white sphere is attached to the black sphere on the top right of the six-sided ring.

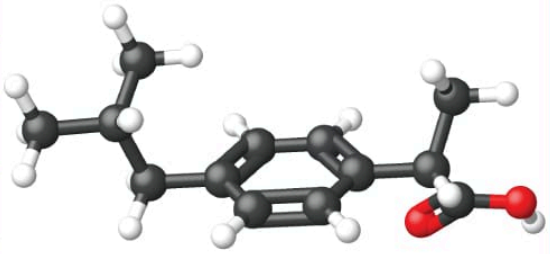
## Example 10.4.1: Computing Molecular Mass for a Covalent Compound

Ibuprofen, $C_{13}H_{18}O_2$, is a covalent compound and the active ingredient in several popular nonprescription pain medications, such as Advil and Motrin. What is the molecular mass (amu) for this compound?

**Solution**

Molecules of this compound are comprised of 13 carbon atoms, 18 hydrogen atoms, and 2 oxygen atoms. Following the approach described above, the average molecular mass for this compound is therefore:

| Element | Quantity | | Average atomic mass (amu) | | Subtotal (amu) |
|---|---|---|---|---|---|
| C | 13 | × | 12.01 | = | 156.13 |
| H | 18 | × | 1.008 | = | 18.114 |
| O | 2 | × | 16.00 | = | 32.00 |
| | | | | Molecular mass | 206.27 |



**Exercise 10.4.1**

Acetaminophen, $C_8H_9NO_2$, is a covalent compound and the active ingredient in several popular nonprescription pain medications, such as Tylenol. What is the molecular mass (amu) for this compound?

*Answer available at the end of this section.*

## Formula Mass for Ionic Compounds

Ionic compounds are composed of discrete cations and anions combined in ratios to yield electrically neutral bulk matter. The formula mass for an ionic compound is calculated in the same way as the formula mass for covalent compounds: by summing the average atomic masses of all the atoms in the compound's formula. Keep in mind, however, that the formula for an ionic compound does not represent the composition of a discrete molecule, so it may not correctly be referred to as the "molecular mass."

As an example, consider sodium chloride, NaCl, the chemical name for common table salt. Sodium chloride is an ionic compound composed of sodium cations, $Na^+$, and chloride anions, $Cl^-$, combined in a 1:1 ratio. The formula mass for this compound is computed as 58.44 amu (Figure 10.4.3).

| Element | Quantity | | Average atomic mass (amu) | | Subtotal |
|---|---|---|---|---|---|
| Na | 1 | × | 22.99 | = | 22.99 |
| Cl | 1 | × | 35.45 | = | 35.45 |
| | | | | Formula mass | 58.44 |

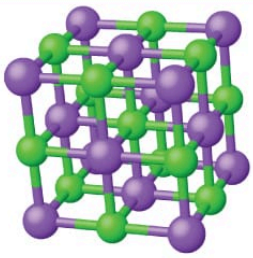


Figure 10.4.3: Table salt, NaCl, contains an array of sodium and chloride ions combined in a 1:1 ratio. Its formula mass is 58.44 amu. A table and diagram are shown. The table is made up of six columns and four rows. The header row reads: "Element," "Quantity," a blank space, "Average atomic mass (a m u)," a blank space and "Subtotal (a m u)." The first column contains the symbols "N a", "C l," and a merged cell. The merged cell runs the length of the first five columns. The second column contains the numbers "1" and "1" as well as the merged cell. The third column contains the multiplication symbol in each cell except for the last, merged cell. The fourth column contains the numbers "22.99" and "35.45" as well as the merged cell. The fifth column contains the symbol "=" in each cell except for the last, merged cell. The sixth column contains the values "22.99," "35.45," and "58.44." There is a thick black line below the number "35.45." The merged cell under the first five columns reads "Formula mass." To the left of the table is a diagram of a chemical structure. The diagram shows green and purple spheres placed in an alternating pattern, making up the corners of eight stacked cubes to form one larger cube. The green spheres are slightly smaller than the purple spheres.

Note that the average masses of neutral sodium and chlorine atoms were used in this computation, rather than the masses for sodium cations and chlorine anions. This approach is perfectly acceptable when computing the formula mass of an ionic

compound. Even though a sodium cation has a slightly smaller mass than a sodium atom (since it is missing an electron), this difference will be offset by the fact that a chloride anion is slightly more massive than a chloride atom (due to the extra electron). Moreover, the mass of an electron is negligibly small with respect to the mass of a typical atom. Even when calculating the mass of an isolated ion, the missing or additional electrons can generally be ignored, since their contribution to the overall mass is negligible, reflected only in the nonsignificant digits that will be lost when the computed mass is properly rounded. The few exceptions to this guideline are very light ions derived from elements with precisely known atomic masses.

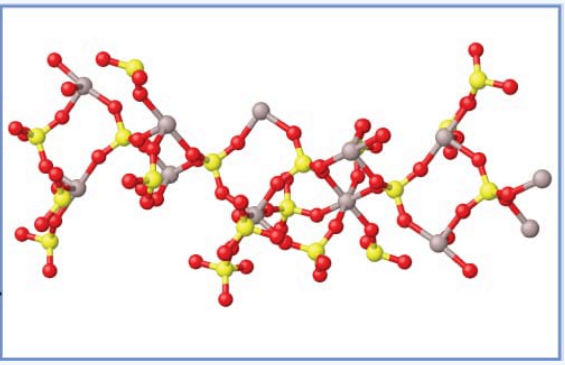
### Example 10.4.2: Computing Formula Mass for an Ionic Compound

Aluminum sulfate, $Al_2(SO_4)_3$, is an ionic compound that is used in the manufacture of paper and in various water purification processes. What is the formula mass (amu) of this compound?

**Solution**

The formula for this compound indicates it contains $Al^{3+}$ and $SO_4^{2-}$ ions combined in a 2:3 ratio. For purposes of computing a formula mass, it is helpful to rewrite the formula in the simpler format, $Al_2S_3O_{12}$. Following the approach outlined above, the formula mass for this compound is calculated as follows:

| Element | Quantity | | Average atomic mass (amu) | | Subtotal (amu) |
|---|---|---|---|---|---|
| Al | 2 | × | 26.98 | = | 53.96 |
| S | 3 | × | 32.06 | = | 96.18 |
| O | 12 | × | 16.00 | = | 192.00 |
| | | | | **Molecular mass** | 342.14 |



### Exercise 10.4.2

Calcium phosphate, $Ca_3(PO_4)_2$, is an ionic compound and a common anti-caking agent added to food products. What is the formula mass (amu) of calcium phosphate?

*Answer available at the end of this section.*

## The Mole

The identity of a substance is defined not only by the types of atoms or ions it contains, but by the quantity of each type of atom or ion. For example, water, $H_2O$, and hydrogen peroxide, $H_2O_2$, are alike in that their respective molecules are composed of hydrogen and oxygen atoms. However, because a hydrogen peroxide molecule contains two oxygen atoms, as opposed to the water molecule, which has only one, the two substances exhibit very different properties. Today, we possess sophisticated instruments that allow the direct measurement of these defining microscopic traits; however, the same traits were originally derived from the measurement of macroscopic properties (the masses and volumes of bulk quantities of matter) using relatively simple tools (balances and volumetric glassware). This experimental approach required the introduction of a new unit for amount of substances, the *mole*, which remains indispensable in modern chemical science.

The mole is an amount unit similar to familiar units like pair, dozen, gross, etc. It provides a specific measure of *the number* of atoms or molecules in a bulk sample of matter. A mole is defined as the amount of substance containing the same number of discrete entities (such as atoms, molecules, and ions) as the number of atoms in a sample of pure $^{12}C$ weighing exactly 12 g. One Latin connotation for the word “mole” is “large mass” or “bulk,” which is consistent with its use as the name for this unit. The mole provides a link between an easily measured macroscopic property, bulk mass, and an extremely important fundamental property,

number of atoms, molecules, and so forth.

The number of entities composing a mole has been experimentally determined to be $6.02214179 \times 10^{23}$, a fundamental constant named **Avogadro's number** ($N_A$) or the Avogadro constant in honor of Italian scientist Amedeo Avogadro. This constant is properly reported with an explicit unit of "per mole," a conveniently rounded version being $6.022 \times 10^{23}/\text{mol}$.

Consistent with its definition as an amount unit, 1 mole of any element contains the same number of atoms as 1 mole of any other element. The masses of 1 mole of different elements, however, are different, since the masses of the individual atoms are drastically different. The molar mass of an element (or compound) is the mass in grams of 1 mole of that substance, a property expressed in units of grams per mole (g/mol) (Figure 10.4.4).

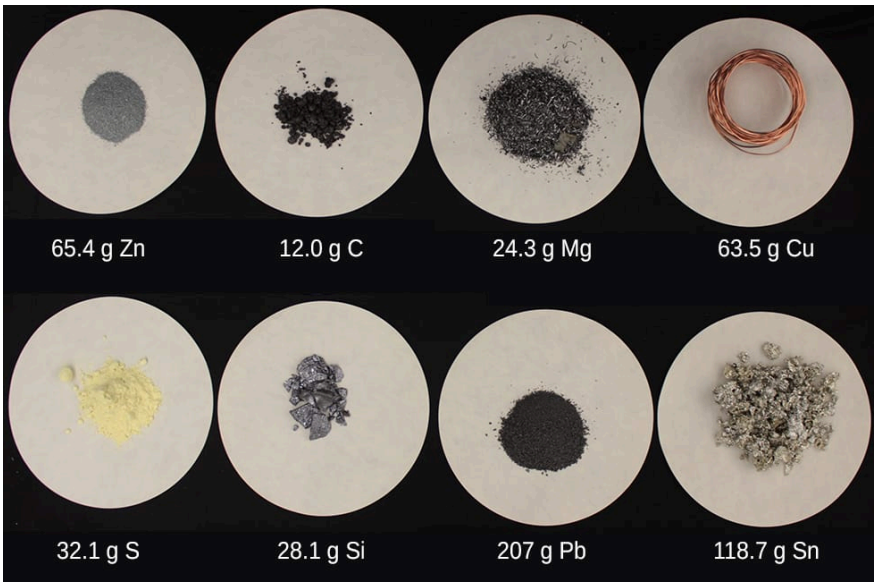



Figure 10.4.4: Each sample contains $6.022 \times 10^{23}$ atoms —1.00 mol of atoms. From left to right (top row): 65.4 g zinc, 12.0 g carbon, 24.3 g magnesium, and 63.5 g copper. From left to right (bottom row): 32.1 g sulfur, 28.1 g silicon, 207 g lead, and 118.7 g tin. (credit: modification of work by Mark Ott). This figure contains eight different substances displayed on white circles. The amount of each substance is visibly different.

Because the definitions of both the mole and the atomic mass unit are based on the same reference substance, $^{12}C$, the molar mass of any substance is numerically equivalent to its atomic or formula weight in amu. Per the amu definition, a single $^{12}C$ atom weighs 12 amu (its atomic mass is 12 amu). According to the definition of the mole, 12 g of $^{12}C$ contains 1 mole of $^{12}C$ atoms (its molar mass is 12 g/mol). This relationship holds for all elements, since their atomic masses are measured relative to that of the amu-reference substance, $^{12}C$. Extending this principle, the molar mass of a compound in grams is likewise numerically equivalent to its formula mass in amu (Figure 10.4.5).

<div data-mt-source="1""This" height="292" width="437" src="/@api/deki/files/56173/CNX_Chem_03_02_compound.jpg"> Figure 10.4.5: Each sample contains $6.022 \times 10^{23}$ molecules or formula units—1.00 mol of the compound or element. Clock-wise from the upper left: 130.2 g of $C_8H_{17}OH$ (1-octanol, formula mass 130.2 amu), 454.4 g of $HgI_2$ (mercury(II) iodide, formula mass 454.4 amu), 32.0 g of $CH_3OH$ (methanol, formula mass 32.0 amu) and 256.5 g of $S_8$ (sulfur, formula mass 256.5 amu). (credit: Sahar Atwa).

*Table 10.4.1: Mass of one mole of elements*

| Element | Average Atomic Mass (amu) | Molar Mass (g/mol) | Atoms/Mole |
|---|---|---|---|
| C | 12.01 | 12.01 | $6.022 \times 10^{23}$ |
| H | 1.008 | 1.008 | $6.022 \times 10^{23}$ |
| O | 16.00 | 16.00 | $6.022 \times 10^{23}$ |
| Na | 22.99 | 22.99 | $6.022 \times 10^{23}$ |
| Cl | 33.45 | 35.45 | $6.022 \times 10^{23}$ |

While atomic mass and molar mass are numerically equivalent, keep in mind that they are vastly different in terms of scale, as represented by the vast difference in the magnitudes of their respective units (amu versus g). To appreciate the enormity of the mole, consider a small drop of water after a rainfall. Although this represents just a tiny fraction of 1 mole of water (~18 g), it contains more water molecules than can be clearly imagined. If the molecules were distributed equally among the roughly seven billion people on earth, each person would receive more than 100 billion molecules.

**YouTube Video**

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video 10.4.1: The mole is used in chemistry to represent $6.022 \times 10^{23}$ of something, but it can be difficult to conceptualize such a large number. Watch this video and then complete the "Think" questions that follow. Explore more about the mole by reviewing the information under "Dig Deeper."

The relationships between formula mass, the mole, and Avogadro's number can be applied to compute various quantities that describe the composition of substances and compounds. For example, if we know the mass and chemical composition of a substance, we can determine the number of moles and calculate number of atoms or molecules in the sample. Likewise, if we know the number of moles of a substance, we can derive the number of atoms or molecules and calculate the substance's mass.

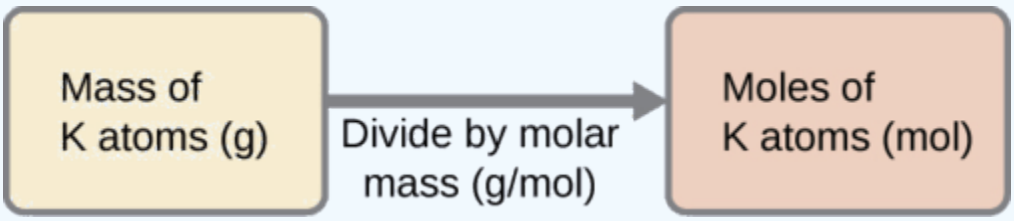
### Example 10.4.3: Deriving Moles from Grams for an Element

According to nutritional guidelines from the US Department of Agriculture, the estimated average requirement for dietary potassium is 4.7 g. What is the estimated average requirement of potassium in moles?

**Solution**

The mass of K is provided, and the corresponding amount of K in moles is requested. Referring to the periodic table, the atomic mass of K is 39.10 amu, and so its molar mass is 39.10 g/mol. The given mass of K (4.7 g) is a bit more than one-tenth the molar mass (39.10 g), so a reasonable "ballpark" estimate of the number of moles would be slightly greater than 0.1 mol.

The molar amount of a substance may be calculated by dividing its mass (g) by its molar mass (g/mol):




The factor-label method supports this mathematical approach since the unit "g" cancels and the answer has units of "mol:"

$$4.7 \cancel{\text{g}}\,\text{K}\left(\frac{\text{mol K}}{39.10 \cancel{\text{g}}}\right) = 0.12 \text{ mol K}$$

The calculated magnitude (0.12 mol K) is consistent with our ballpark expectation, since it is a bit greater than 0.1 mol.

## Exercise 10.4.3: Beryllium

Beryllium is a light metal used to fabricate transparent X-ray windows for medical imaging instruments. How many moles of Be are in a thin-foil window weighing 3.24 g?

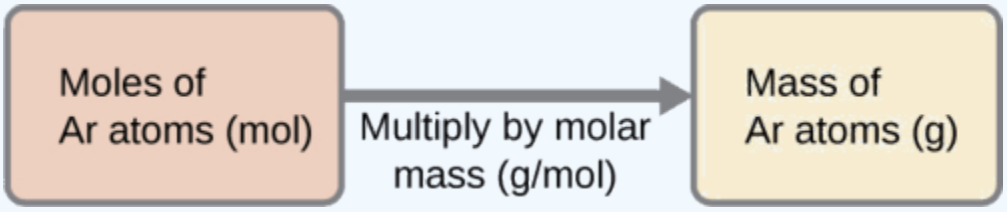
*Answer available at the end of this section.*

## Example 10.4.4: Deriving Grams from Moles for an Element

A liter of air contains $9.2 \times 10^{-4}$ mol argon. What is the mass of Ar in a liter of air?

**Solution**

The molar amount of Ar is provided and must be used to derive the corresponding mass in grams. Since the amount of Ar is less than 1 mole, the mass will be less than the mass of 1 mole of Ar, approximately 40 g. The molar amount in question is approximately one-one thousandth (~$10^{-3}$) of a mole, and so the corresponding mass should be roughly one-one thousandth of the molar mass (~0.04 g):




In this case, logic dictates (and the factor-label method supports) multiplying the provided amount (mol) by the molar mass (g/mol):

$$9.2 \times 10^{-4} \cancel{\text{mol}}\ \text{Ar} \left( \frac{39.95\ \text{g}}{\cancel{\text{mol}}\ \text{Ar}} \right) = 0.037\ \text{g Ar}$$

The result is in agreement with our expectations, around 0.04 g Ar.

## Exercise 10.4.4

What is the mass of 2.561 mol of gold?

*Answer available at the end of this section.*

## Example 10.4.6: Deriving Number of Atoms from Mass for an Element

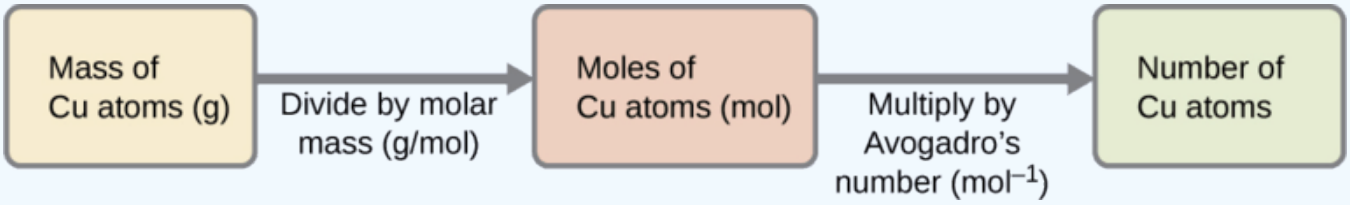
Copper is commonly used to fabricate electrical wire (Figure 10.4.6). How many copper atoms are in 5.00 g of copper wire?



Figure 10.4.6: Copper wire is composed of many, many atoms of Cu. (credit: Emilian Robert Vicol)

**Solution**

The number of Cu atoms in the wire may be conveniently derived from its mass by a two-step computation: first calculating the molar amount of Cu, and then using Avogadro's number ($N_A$) to convert this molar amount to number of Cu atoms:




Considering that the provided sample mass (5.00 g) is a little less than one-tenth the mass of 1 mole of Cu (~64 g), a reasonable estimate for the number of atoms in the sample would be on the order of one-tenth $N_A$, or approximately $10^{22}$ Cu atoms. Carrying out the two-step computation yields:

$$5.00 \cancel{\text{ g}}\ \text{Cu}\left(\frac{\cancel{\text{mol}}\ \text{Cu}}{63.55 \cancel{\text{ g}}}\right)\left(\frac{6.022 \times 10^{23}\ \text{atoms}}{\cancel{\text{mol}}}\right) = 4.74 \times 10^{22}\ \text{atoms of copper}$$

The factor-label method yields the desired cancellation of units, and the computed result is on the order of $10^{22}$ as expected.

## Exercise 10.4.6

A prospector panning for gold in a river collects 15.00 g of pure gold. How many Au atoms are in this quantity of gold?

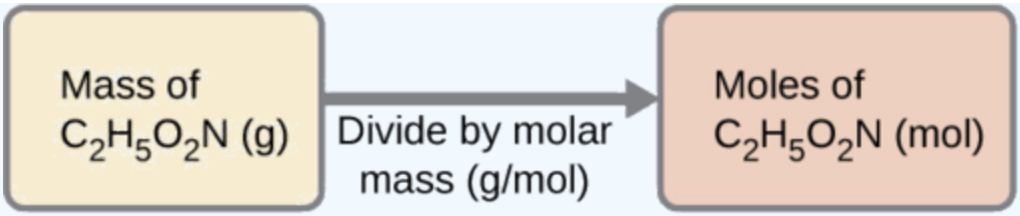
*Answer available at the end of this section.*

## Example 10.4.7: Deriving Moles from Grams for a Compound

Our bodies synthesize protein from amino acids. One of these amino acids is glycine, which has the molecular formula $C_2H_5O_2N$. How many moles of glycine molecules are contained in 28.35 g of glycine?

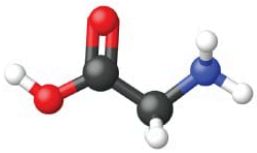
**Solution**

We can derive the number of moles of a compound from its mass following the same procedure we used for an element in Example 10.4.6:




The molar mass of glycine is required for this calculation, and it is computed in the same fashion as its molecular mass. One mole of glycine, $C_2H_5O_2N$, contains 2 moles of carbon, 5 moles of hydrogen, 2 moles of oxygen, and 1 mole of nitrogen:

| Element | Quantity (mol element/ mol compound) | | Molar mass (g/mol element) | | Subtotal (g/mol compound) |
|---|---|---|---|---|---|
| C | 2 | × | 12.01 | = | 24.02 |
| H | 5 | × | 1.008 | = | 5.040 |
| O | 2 | × | 16.00 | = | 32.00 |
| N | 1 | × | 14.007 | = | 14.007 |
| Molecular mass (g/mol compound) | | | | | 75.07 |



A table is shown that is made up of six columns and six rows. The header row reads: “Element,” “Quantity (mol element / mol compound,” a blank space, “Molar mass (g / mol element),” a blank space, and “Subtotal (a m u).” The first column contains the symbols “C,” “H,” “O,” “N,” and a merged cell. The merged cell runs the width of the first five columns. The second column contains the numbers “2,” “5,” “2,” and “1” as well as the merged cell. The third column contains the multiplication symbol in each cell except for the last, merged cell. The fourth column contains the numbers “12.01,” “1.008,” “16.00,” and “14.007” as well as the merged cell. The fifth column contains the symbol “=” in each cell except for the last, merged cell. The sixth column contains the values “24.02,” “5.040,” “32.00,” “14.007,” and “75.07.” There is a thick black line under the number 14.007. The merged cell under the first five columns reads “Molar mass (g / mol compound). There is a ball-and-stick drawing to the right of this table. It shows a black sphere that forms a double bond with a slightly smaller red sphere, a single bond with another red sphere, and a single bond with another black sphere. The red sphere that forms a single bond with the black sphere also forms a single bond with a smaller, white sphere. The second black sphere forms a single bond with a smaller, white sphere and a smaller blue sphere. The blue sphere forms a single bond with two smaller, white spheres each.

The provided mass of glycine (~28 g) is a bit more than one-third the molar mass (~75 g/mol), so we would expect the computed result to be a bit greater than one-third of a mole (~0.33 mol). Dividing the compound’s mass by its molar mass yields:

$$28.35 \cancel{g}\ \text{glycine}\left(\frac{\text{mol glycine}}{75.07 \cancel{g}}\right) = 0.378\ \text{mol glycine}$$

This result is consistent with our rough estimate.

## Exercise 10.4.7

How many moles of sucrose, $C_{12}H_{22}O_{11}$, are in a 25-g sample of sucrose?

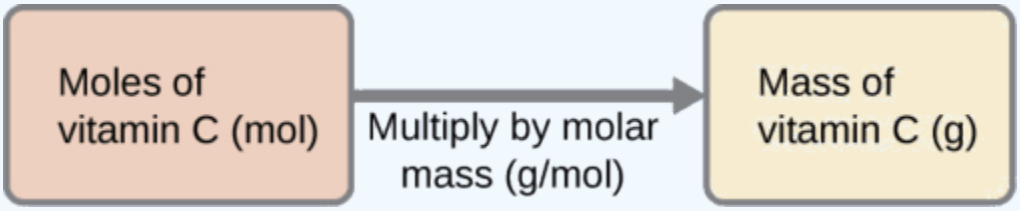
*Answer available at the end of this section.*

## Example 10.4.8: Deriving Grams from Moles for a Compound

Vitamin C is a covalent compound with the molecular formula $C_6H_8O_6$. The recommended daily dietary allowance of vitamin C for children aged 4–8 years is $1.42 \times 10^{-4}$ mol. What is the mass of this allowance in grams?

**Solution**

As for elements, the mass of a compound can be derived from its molar amount as shown:




The molar mass for this compound is computed to be 176.124 g/mol. The given number of moles is a very small fraction of a mole (~$10^{-4}$ or one-ten thousandth); therefore, we would expect the corresponding mass to be about one-ten thousandth of the molar mass (~0.02 g). Performing the calculation, we get:

$$1.42 \times 10^{-4} \cancel{\text{mol}}\ \text{vitamin C} \left( \frac{176.124\ \text{g}}{\cancel{\text{mol}}\ \text{vitamin C}} \right) = 0.0250\ \text{g vitamin C}$$

This is consistent with the anticipated result.

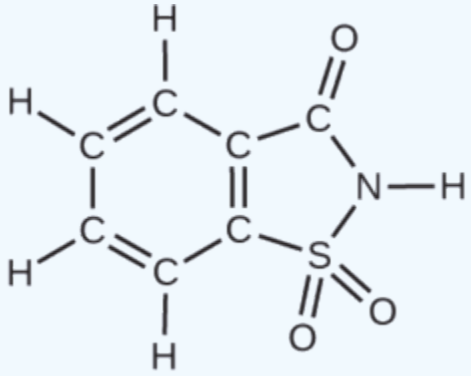
## Exercise 10.4.8

What is the mass of 0.443 mol of hydrazine, $N_2H_4$?

*Answer available at the end of this section.*

## Example 10.4.9: Deriving the Number of Molecules from the Compound Mass

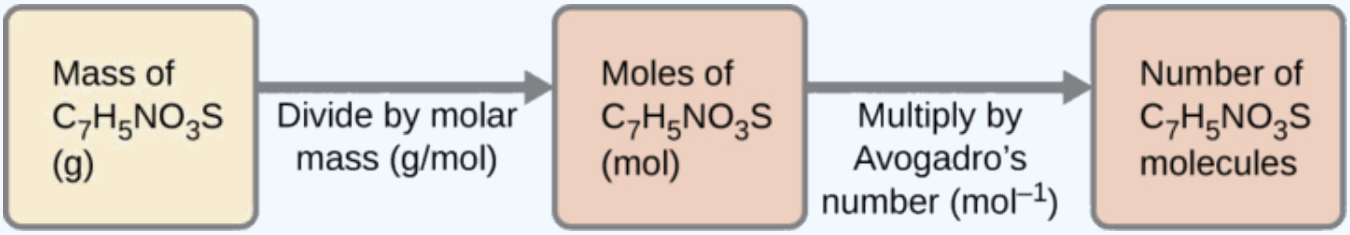
A packet of an artificial sweetener contains 40.0 mg of saccharin ($C_7H_5NO_3S$), which has the structural formula:



Given that saccharin has a molar mass of 183.18 g/mol, how many saccharin molecules are in a 40.0-mg (0.0400-g) sample of saccharin? How many carbon atoms are in the same sample?

**Solution**

The number of molecules in a given mass of compound is computed by first deriving the number of moles, as demonstrated in Example 10.4.8, and then multiplying by Avogadro's number:




Using the provided mass and molar mass for saccharin yields:

$$0.0400 \cancel{g}\ C_7H_5NO_3S \left(\frac{\cancel{mol}\ C_7H_5NO_3S}{183.18 \cancel{g}\ C_7H_5NO_3S}\right)\left(\frac{6.022 \times 10^{23}\ C_7H_5NO_3S\ molecules}{1\ \cancel{mol}\ C_7H_5NO_3S}\right)$$
$$= 1.31 \times 10^{20}\ C_7H_5NO_3S\ molecules$$

The compound's formula shows that each molecule contains seven carbon atoms, and so the number of C atoms in the provided sample is:

$$1.31 \times 10^{20}\ C_7H_5NO_3S\ molecules \left(\frac{7\ C\ atoms}{1\ C_7H_5NO_3S\ molecule}\right) = 9.20 \times 10^{21}\ C\ atoms$$

**Exercise 10.4.9**

How many $C_4H_{10}$ molecules are contained in 9.213 g of this compound? How many hydrogen atoms?

*Answer available at the end of this section.*

## Summary

The formula mass of a substance is the sum of the average atomic masses of each atom represented in the chemical formula and is expressed in atomic mass units. The formula mass of a covalent compound is also called the molecular mass. A convenient amount unit for expressing very large numbers of atoms or molecules is the mole. Experimental measurements have determined the number of entities composing 1 mole of substance to be 6.022 × $10^{23}$, a quantity called Avogadro's number. The mass in grams of 1 mole of substance is its molar mass. Due to the use of the same reference substance in defining the atomic mass unit and the mole, the formula mass (amu) and molar mass (g/mol) for any substance are numerically equivalent (for example, one $H_2O$ molecule weighs approximately18 amu and 1 mole of $H_2O$ molecules weighs approximately 18 g).

## Footnotes

1. 1 Omiatek, Donna M., Amanda J. Bressler, Ann-Sofie Cans, Anne M. Andrews, Michael L. Heien, and Andrew G. Ewing. "The Real Catecholamine Content of Secretory Vesicles in the CNS Revealed by Electrochemical Cytometry." *Scientific Report* 3 (2013): 1447, accessed January 14, 2015, doi:10.1038/srep01447.

## Glossary

**Avogadro's number ($N_A$)**

experimentally determined value of the number of entities comprising 1 mole of substance, equal to 6.022 × $10^{23}$ $mol^{-1}$

**formula mass**
: sum of the average masses for all atoms represented in a chemical formula; for covalent compounds, this is also the molecular mass

**mole**
: amount of substance containing the same number of atoms, molecules, ions, or other entities as the number of atoms in exactly 12 grams of $^{12}C$

**molar mass**
: mass in grams of 1 mole of a substance

## Answers

**Exercise** 10.4.1
151.16 amu

**Exercise** 10.4.2
310.18 amu

**Exercise** 10.4.3**: Beryllium**
0.360 mol

**Exercise** 10.4.4
504.4 g

**Exercise** 10.4.6
$4.586 \times 10^{22}$ $Au$ atoms

**Exercise** 10.4.7
0.073 mol

**Exercise** 10.4.8
14.2 g

**Exercise** 10.4.9

- $9.545 \times 10^{22}$ molecules $C_4H_{10}$
- $9.545 \times 10^{23}$ atoms $H$

This page titled 10.4: Formula Mass and the Mole Concept is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.5: Balancing Equations

**Learning Objectives**

- Balance a chemical equation.

## Any leftovers?

When you cook a meal, quite often there are leftovers because you prepared more than people would eat at one sitting. Sometimes when you repair a piece of equipment, you end up with what are called "pocket parts"—small pieces you put in your pocket because you're not sure where they belong. Chemistry tries to avoid leftovers and pocket parts. In normal chemical processes, we cannot create or destroy matter (law of conservation of mass). If we start out with ten carbon atoms, we need to end up with ten carbon atoms. John Dalton's atomic theory said that chemical reactions basically involve the rearrangement of atoms. Chemical equations need to follow these principles in order to be correct.

## Balancing Chemical Equations

A **balanced equation** is a chemical equation in which mass is conserved and there are equal numbers of atoms of each element on both sides of the equation. We can write a chemical equation for the reaction of carbon with hydrogen gas to form methane $(\mathrm{CH_4})$:

$$\underset{2\text{ C atoms}}{\mathrm{C}\,(s)} \quad + \quad \underset{2\text{ H atoms}}{\mathrm{H_2}\,(g)} \quad \rightarrow \quad \underset{1\text{ C atom, }4\text{ H atoms}}{\mathrm{CH_4}\,(g)}$$

In order to write a correct equation, you must first write the correct skeleton equation with the correct chemical formulas. Recall that hydrogen is a diatomic molecule and so is written as $\mathrm{H_2}$.

When we count the number of atoms of both elements, shown under the equation, we see that the equation is not balanced. There are only 2 atoms of hydrogen on the reactant side of the equation, while there are 4 atoms of hydrogen on the product side. We can balance the above equation by adding a coefficient of 2 in front of the formula for hydrogen.

$$\mathrm{C}\,(s) + 2\mathrm{H_2}\,(g) \rightarrow \mathrm{CH_4}\,(g)$$

A **coefficient** is a small in magnitude whole number placed in front of a formula in an equation to signify the quantity needed to balance it. The 2 in front of the $\mathrm{H_2}$ means that there are a total of $2 \times 2 = 4$ atoms of hydrogen as reactants. Visually, the reaction looks like the figure below.

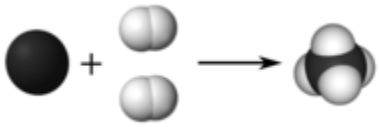


Figure $10.5.2$: Reaction between carbon and hydrogen to form methane. (Credit: Ben Mills (Wikimedia: Benjah-bmm27; Source, and Source); License: Public Domain)

In the balanced equation, there is one atom of carbon and four atoms of hydrogen on both sides of the arrow. Below are guidelines for writing and balancing chemical equations.

1. Determine the correct chemical formulas for each reactant and product.
2. Write the skeleton equation.
3. Count the number of atoms of each element that appears as a reactant and as a product. If a polyatomic ion is unchanged on both sides of the equation, count it as a unit.
4. Balance each element one at a time by placing coefficients in front of the formulas.
   a. It is best to begin by balancing elements that only appear in one chemical formula on each side of the equation.
   b. It is best to end with an element by itself (if possible).

CK-12 LICENSED UNDER CK-12 License
© CK-12 Foundation • Visit us at ck12.org

c. No coefficient is written for a 1.
d. NEVER change the subscripts in a chemical formula—you can only balance equations by using coefficients.

5. Check each atom or polyatomic ion to be sure that they are equal on both sides of the equation.
6. Make sure that all coefficients are in the lowest possible ratio. If necessary, reduce to the lowest ratio.

**Example 10.5.1: Balancing Chemical Equations**

Aqueous solutions of lead (II) nitrate and sodium chloride are mixed. The products of the reaction are an aqueous solution of sodium nitrate and a solid precipitate of lead (II) chloride. Write the balanced chemical equation for this reaction.

**Solution**

***Step 1: Plan the problem.***

Follow the steps for writing and balancing a chemical equation listed in the text.

***Step 2: Solve.***

Write the skeleton equation with the correct formulas.

$$\mathrm{Pb(NO_3)_2}\,(aq) + \mathrm{NaCl}\,(aq) \rightarrow \mathrm{NaNO_3}\,(aq) + \mathrm{PbCl_2}\,(s)$$

Count the number of each atom or polyatomic ion on both sides of the equation.

| Reactants | Products |
|---|---|
| 1 Pb atom | 1 Pb atom |
| 2 $\mathrm{NO_3}^-$ ions | 1 $\mathrm{NO_3}^-$ ions |
| 1 Na atom | 1 Na atom |
| 1 Cl atom | 2 Cl atoms |

The nitrate ions and the chlorine atoms are unbalanced. Start by placing a 2 in front of the $\mathrm{NaCl}$. This increases the reactant counts to 2 $\mathrm{Na}$ atoms and 2 $\mathrm{Cl}$ atoms. Then place a 2 in front of the $\mathrm{NaNO_3}$. The result is:

$$\mathrm{Pb(NO_3)_2}\,(aq) + 2\mathrm{NaCl}\,(aq) \rightarrow 2\mathrm{NaNO_3}\,(aq) + \mathrm{PbCl_2}\,(s)$$

The new count for each atom and polyatomic ion becomes:

| Reactants | Products |
|---|---|
| 1 Pb atom | 1 Pb atom |
| 2 $\mathrm{NO_3}^-$ ions | 2 $\mathrm{NO_3}^-$ ions |
| 2 Na atom | 2 Na atom |
| 2 Cl atom | 2 Cl atoms |

***Step 3: Think about your result.***

The equation is now balanced since there are equal numbers of atoms of each element on both sides of the equation.

**MISCONCEPTION ALERT: YOU CAN CHANGE ANY NUMBER YOU WANT TO BALANCE AN EQUATION**

The substances in a balanced equation are described by their chemical formula. The chemical formula will contain subscripts to indicate the number of atoms of a given element if that substance has more than one atom of that element in its formula. If we change subscripts in an attempt to balance an equation, we are actually changing the identity of the substances in the chemical reaction that the balanced equation is about. Imagine if you wanted a drink of water and when you put it in your mouth realized that it was actually hydrogen peroxide! The difference between water ($H_2O$) and hydrogen peroxide ($H_2O_2$) is significant, but the difference in the chemical formula is merely a subscript. Any time you change the subscript in a chemical equation you are making a similar change to the substances involved. This isn't balancing the equation, it is saying something else instead of answering the question.

CK-12 LICENSED UNDER CK-12 Science © CK-12 Foundation • Visit us at ck12.org

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Simulation

Practice balancing chemical equations with this simulation:

Can you balance a chemical equation?

## Section Summary

- The process of balancing chemical equations is described.

This page titled 10.5: Balancing Equations is shared under a CK-12 license and was authored, remixed, and/or curated by CK12 via source content that was edited to the style and standards of the LibreTexts platform.

- 11.2: Magnets by Andrew Park is licensed CC BY 4.0.

# 10.6: Energy and Chemical Reactions

**Learning Objectives**

- Describe how energy might interact with a chemical reaction, both graphically and in words.

Gunpowder was originally developed by the Chinese in the ninth century CE, primarily for rockets. This material is composed of charcoal, sulfur, and saltpeter (potassium nitrate). The explosive reaction that occurs involves the conversion of the charcoal to carbon dioxide, with the potassium nitrate providing the extra oxygen needed for a rapid reaction. Although sulfur was included to stabilize the product, gunpowder is still highly explosive.

## Chemical Potential Energy

As was discussed earlier in this text, energy is the capacity for doing work or supplying heat. When you fill your car with gasoline, you are providing it with potential energy. **Chemical potential energy** is the energy stored in the chemical bonds of a substance. The various chemicals that make up gasoline contain a large amount of chemical potential energy that is released when the gasoline is burned in a controlled way in the engine of the car. The release of that energy does two things: some of the potential energy is transformed into work, which is used to move the car; at the same time, some of the potential energy is converted to heat and makes the car's engine very hot. The energy changes of a system occur as either heat or work, or some combination of both.



Figure 10.6.1: A dragster is able to accelerate because of the chemical potential energy of its fuel. The burning of the fuel also produces large amounts of heat. (CC BY-NC; CK-12)

Dynamite is another example of chemical potential energy. The major component of dynamite is nitroglycerin, a very unstable material. By mixing it with diatomaceous earth, the stability is increased and it is less likely to explode if it receives a physical shock. When ignited, the nitroglycerin explodes rapidly, releasing large amounts of nitrogen and other gases along with a massive amount of heat.



Figure $10.6.2$: Dynamite explosion. (CC BY-NC; CK-12)

## Exothermic and Endothermic Processes

When physical or chemical changes occur, they are generally accompanied by a transfer of energy. The **law of conservation of energy** states that in any physical or chemical process, energy is neither created nor destroyed. In other words, the entire energy in the universe is conserved. In order to better understand the energy changes taking place during a reaction, we need to define two parts of the universe: *the system* and *the surroundings*. The **system** is the specific portion of matter in a given space that is being studied during an experiment or an observation. The **surroundings** is everything in the universe that is not part of the system. In practical terms for a laboratory chemist, the system is the particular chemicals being reacted, while the surroundings is the immediate vicinity within the room. During most processes, energy is exchanged between the system and the surroundings. If the system loses a certain amount of energy, that same amount of energy is gained by the surroundings. If the system gains a certain amount of energy, that energy is supplied by the surroundings.

A chemical reaction or physical change is **endothermic** if heat is absorbed by the system from the surroundings. In the course of an endothermic process, the system gains heat from the surroundings, and so the temperature of the surroundings decreases. The quantity of heat for a process is represented by the letter $q$. The sign of $q$ for an endothermic process is positive because the system is gaining heat. A chemical reaction or physical change is **exothermic** if heat is released by the system into the surroundings. Because the surroundings are gaining heat from the system, the temperature of the surroundings increases. The sign of $q$ for an exothermic process is negative because the system is losing heat.

Scientists have define a quantity called enthalpy to make it more useful to keep track of the flow of heat in a reaction from a thermodynamics perspective. The definition of enthalpy is beyond the scope of this text, but it is defined in such a way that the heat flow at constant pressure is equal to the change in enthalpy (represented as $(\Delta H)$. Most reactions of interest to chemists occur at constant pressure, so this creates a convenient way of taking a non-conservative property such a heat flow and equating it with a state function such as enthalpy, which is more convenient mathematically. Throughout the rest of this text we will refer to change in enthalpy, $(\Delta H)$, and heat flow somewhat interchangeably.

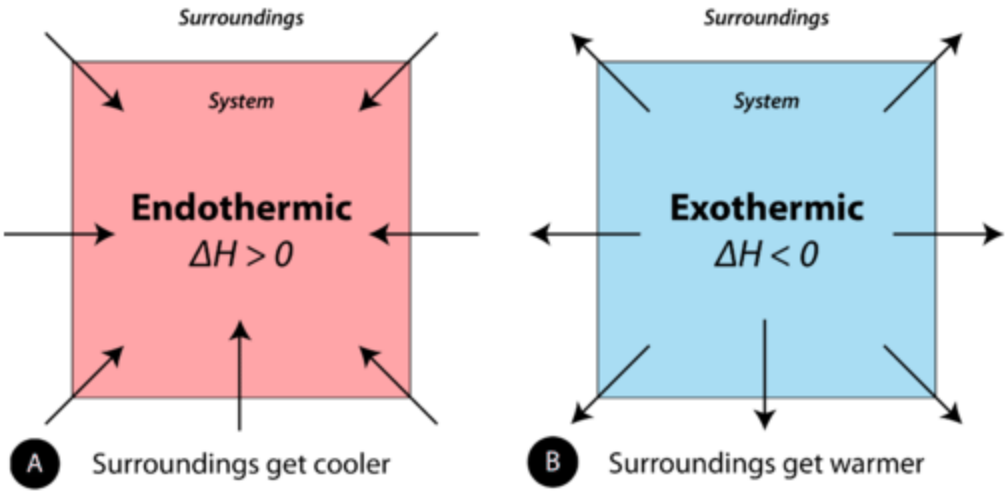



Figure 10.6.3: (A) Endothermic reaction. (B) Exothermic reaction. (CC BY-NC; CK-12)

## Activation Energy



Figure 10.6.4: Calcium metal stored in an argon atmosphere. (CC BY-NC; CK-12)

Why do some chemical reactions occur readily while others require input of heat in order to take place? If we mix metallic sodium with water, a reaction occurs immediately, releasing a great deal of heat (enough to ignite the hydrogen gas that is formed). Group II metals, such as calcium, react at a much slower rate. Unlike the extremely vigorous reaction with sodium, the reaction with calcium is slow enough that we can trap the hydrogen gas released.

Supplying reactant particles with energy causes the bonds between the atoms to vibrate with a greater frequency. This increase in vibrational energy makes a chemical bond more likely to break, and a chemical reaction more likely to occur, when those particles collide with other particles. The **activation energy** for a reaction is the minimum energy that colliding particles must have in order to undergo a reaction. Some reactions occur readily at room temperature because the reacting particles already have the requisite activation energy at that temperature. Other reactions only occur when heated because the particles do not have enough energy unless an external source of heat provides the particles with more kinetic energy.

## Potential Energy Diagrams

The energy changes that occur during a chemical reaction can be shown in a diagram called a potential energy diagram, or sometimes called a reaction progress curve. A **potential energy diagram** shows the change in potential energy of a system

as reactants are converted into products. The figure below shows basic potential energy diagrams for an endothermic (A) and an exothermic (B) reaction. Recall that the enthalpy change $(\Delta H)$ is positive for an endothermic reaction and negative for an exothermic reaction. This can be seen in the potential energy diagrams. The total potential energy of the system increases for the endothermic reaction as the system absorbs energy from the surroundings. The total potential energy of the system decreases for the exothermic reaction as the system releases energy to the surroundings.

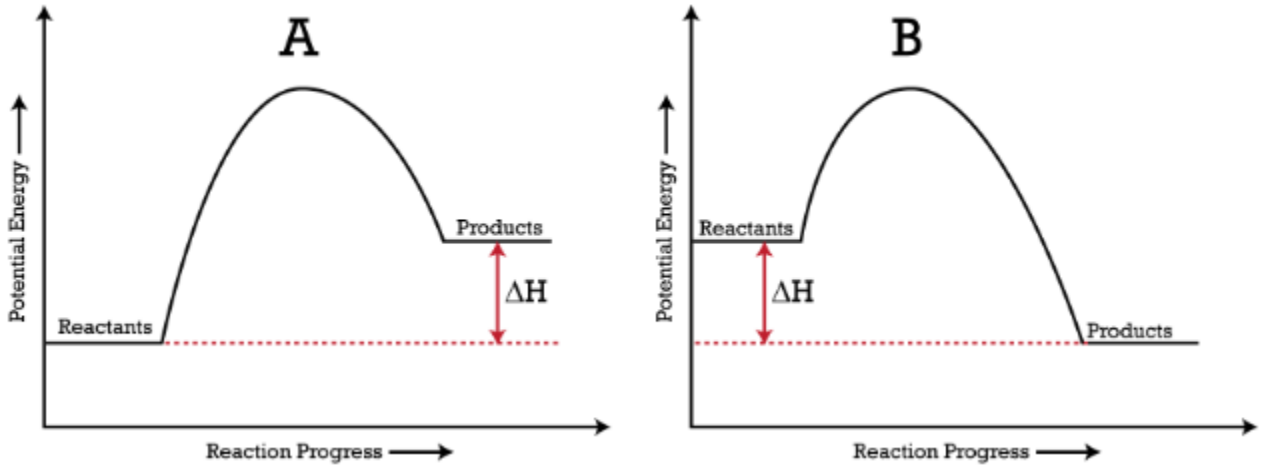



Figure 10.6.5: A potential energy diagram shows the total potential energy of a reacting system as the reaction proceeds. (A) In an endothermic reaction, the energy of the products is greater than the energy of the reactants and $\Delta H$ is positive. (B) In an exothermic reaction, the energy of the products is lower than the energy of the reactants and $\Delta H$ is negative. (CC BY-NC; CK-12)

The activation energy for a reaction is illustrated in the potential energy diagram by the height of the hill between the reactants and the products. For this reason, the activation energy of a reaction is sometimes referred to as the activation energy barrier. Reacting particles must have enough energy so that when they collide, they can overcome that barrier (see figure below).

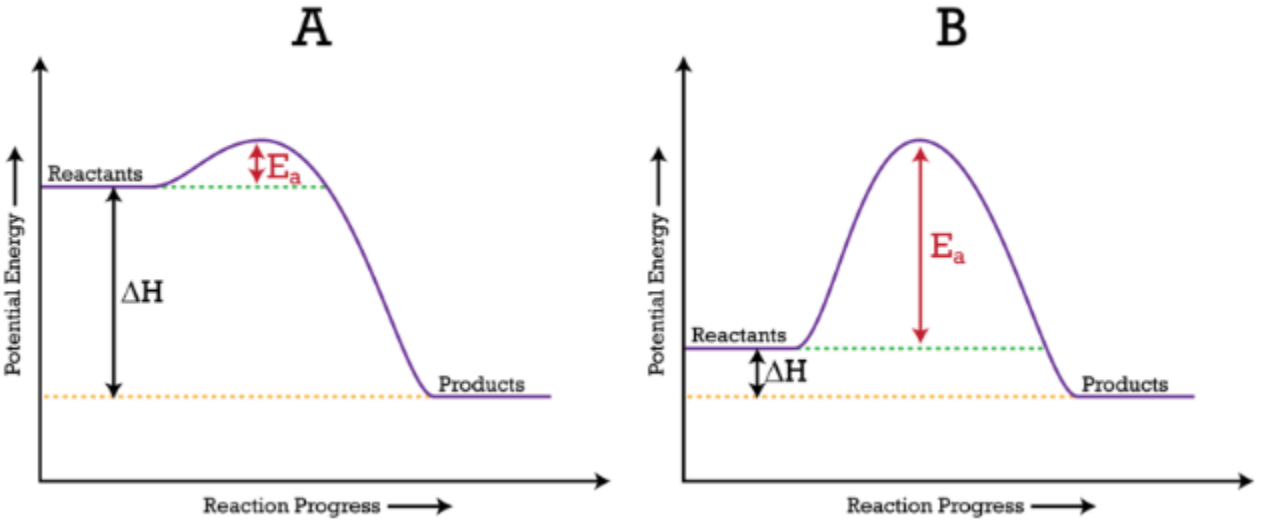



Figure 10.6.6: The activation energy $(E_a)$ of a reaction is the barrier that must be overcome for the reactants to be able to become products. (A) The activation energy is low, meaning that the reaction is likely to be fast. (B) The activation energy is high, meaning that the reaction is likely to be slow. (CC BY-NC; CK-12)

## Section Summary

- Chemical potential energy is energy available in the chemical bonds of a compound.
- A potential energy diagram shows the change in potential energy of a system as reactants are converted into products.
- Potential energy diagrams for endothermic and exothermic reactions are described.
- Diagrams of activation energy and reaction progress are given.

## Glossary

**chemical potential energy**
: The energy stored in the chemical bonds of a substance.

**system**
: The specific portion of matter in a given space that is being studied during an experiment or an observation.

**surroundings**
: Everything in the universe that is not part of the system.

**endothermic**
: A chemical or physical change in which heat is absorbed by the system from the surroundings.

**exothermic**
: A chemical or physical change in which heat is released by the system into the surroundings.

**activation energy**
: The minimum energy that colliding particles must have in order to undergo a reaction.

This page titled 10.6: Energy and Chemical Reactions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.7: Classifying Reactions

**Learning Objectives**

- Categorize a reaction as combination, decomposition, single replacement, or double replacement.

There are millions of chemical reactions that occur in the world. Just like we use the periodic table to help us remember properties of the elements, we can group chemical reactions into categories based on similar aspects. We will not look at all the possible reaction types in this section, but we will introduce some important categories for simple reactions. In the following sections we will discuss some more specific reaction types based on how they influence our lives.

## Combination Reactions

A **combination reaction** is a reaction in which two or more substances combine to form a single new substance. Combination reactions can also be called synthesis reactions. The general form of a combination reaction is:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{AB}$$

One combination reaction is two elements combining to form a compound. Solid sodium metal reacts with chlorine gas to produce solid sodium chloride.

$$2\mathrm{Na}\,(s) + \mathrm{Cl_2}\,(g) \rightarrow 2\mathrm{NaCl}\,(s)$$

**YouTube Video**

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video 10.7.1: Sodium reacting with chlorine to make table salt. (Source)

In order to write and balance the equation correctly, it is important to remember the seven elements that exist in nature as diatomic molecules: $\mathrm{H_2}$, $\mathrm{N_2}$, $\mathrm{O_2}$, $\mathrm{F_2}$, $\mathrm{Cl_2}$, $\mathrm{Br_2}$, and $\mathrm{I_2}$.

One sort of combination reaction that occurs frequently is the reaction of an element with oxygen to form an oxide. Metals and nonmetals both react readily with oxygen under most conditions. Magnesium reacts rapidly and dramatically when ignited, combining with oxygen from the air to produce a fine powder of magnesium oxide.

$$2\mathrm{Mg}\,(s) + \mathrm{O_2}\,(g) \rightarrow 2\mathrm{MgO}\,(s)$$

## Decomposition Reactions

A **decomposition reaction** is a reaction in which a compound breaks down into two or more simpler substances. The general form of a decomposition reaction is:

$$\mathrm{AB} \rightarrow \mathrm{A} + \mathrm{B}$$

Most decomposition reactions require an input of energy in the form of heat, light, or electricity.

CK-12 LICENSED UNDER CC BY NC SA
© CK-12 Foundation • Visit us at ck12.org

Binary compounds are compounds composed of just two elements. The simplest kind of decomposition reaction is when a binary compound decomposes into its elements. Mercury (II) oxide, a red solid, decomposes when heated to produce mercury and oxygen gas.

$$2\text{HgO}\,(s) \rightarrow 2\text{Hg}\,(l) + \text{O}_2\,(g)$$

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video 10.7.2: Mercuric oxide decomposing into mercury and oxygen (Source)

A reaction is also considered to be a decomposition reaction even when one or more of the products are still compounds. A metal carbonate decomposes into a metal oxide and carbon dioxide gas. For example, calcium carbonate decomposes into calcium oxide and carbon dioxide:

$$\text{CaCO}_3\,(s) \rightarrow \text{CaO}\,(s) + \text{CO}_2\,(g)$$

Metal hydroxides decompose on heating to yield metal oxides and water. Sodium hydroxide decomposes to produce sodium oxide and water:

$$2\text{NaOH}\,(s) \rightarrow \text{Na}_2\text{O}\,(s) + \text{H}_2\text{O}\,(g)$$

Some unstable acids decompose to produce nonmetal oxides and water. Carbonic acid decomposes easily at room temperature into carbon dioxide and water:

$$\text{H}_2\text{CO}_3\,(aq) \rightarrow \text{CO}_2\,(g) + \text{H}_2\text{O}\,(l)$$

# Single-Replacement Reactions

A **single-replacement reaction** is a reaction in which one element replaces a similar element in a compound. The general form of a single-replacement (also called single-displacement) reaction is:

$$\text{A} + \text{BC} \rightarrow \text{AC} + \text{B}$$

In this general reaction, element $\text{A}$ is a metal and replaces element $\text{B}$ (also a metal) in the compound. When the element that is doing the replacing is a nonmetal, it must replace another nonmetal in a compound, and the general equation becomes:

$$\text{Y} + \text{XZ} \rightarrow \text{XY} + \text{Z}$$

(Where $\text{Y}$ is a nonmetal and replaces the nonmetal $\text{Z}$ in the compound with $\text{X}$.)

## Metal Replacement

Magnesium is a more reactive metal than copper. When a strip of magnesium metal is placed in an aqueous solution of copper (II) nitrate, it replaces the copper. The products of the reaction are aqueous magnesium nitrate and solid copper metal.

$$\text{Mg}\,(s) + \text{Cu(NO}_3)_2\,(aq) \rightarrow \text{Mg(NO}_3)_2\,(aq) + \text{Cu}\,(s)$$

This subcategory of single-replacement reactions is called a metal replacement reaction because it is a metal that is being replaced (copper).

CK-12 LICENSED UNDER CC BY NC SA
© CK-12 Foundation • Visit us at ck12.org

## Hydrogen Replacement

Many metals react easily with acids and when they do so, one of the products of the reaction is hydrogen gas. Zinc reacts with hydrochloric acid to produce aqueous zinc chloride and hydrogen (figure below).

$$\mathrm{Zn}\,(s) + 2\mathrm{HCl}\,(aq) \rightarrow \mathrm{ZnCl_2}\,(aq) + \mathrm{H_2}\,(g)$$

In a hydrogen replacement reaction, the hydrogen in the acid is replaced by an active metal.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video 10.7.3: Zinc reacting with hydrochloric acid produces bubbles of hydrogen gas. (Source)

Some metals are so reactive that they are capable of replacing the hydrogen in water. The products of such a reaction are the metal hydroxide and hydrogen gas. All group 1 metals undergo this type of reaction. Sodium reacts vigorously with water to produce aqueous sodium hydroxide and hydrogen (see figure below).

$$2\mathrm{Na}\,(s) + 2\mathrm{H_2O}\,(l) \rightarrow 2\mathrm{NaOH}\,(aq) + \mathrm{H_2}\,(g)$$



Figure 10.7.1: Sodium metal reacts vigorously with water, giving off hydrogen gas. A large piece of sodium will often generate so much heat that the hydrogen will ignite. (Credit: User:Ajhalls/ Wikimedia Commons; Source; License: Public Domain)

# Double-Replacement Reactions

A **double-replacement reaction** is a reaction in which the positive and negative ions of two ionic compounds exchange places to form two new compounds. The general form of a double-replacement (also called double-displacement) reaction is:

$$\mathrm{AB} + \mathrm{CD} \rightarrow \mathrm{AD} + \mathrm{CB}$$

In this reaction, $\mathrm{A}$ and $\mathrm{C}$ are positively-charged cations, while $\mathrm{B}$ and $\mathrm{D}$ are negatively-charged anions. Double-replacement reactions generally occur between substances in aqueous solution. In order for a reaction to occur, one of the products is usually a solid precipitate, a gas, or a molecular compound such as water.

## Formation of a Precipitate

A precipitate forms in a double-replacement reaction when the cations from one of the reactants combine with the anions from the other reactant to form an insoluble ionic compound. When aqueous solutions of potassium iodide and lead (II) nitrate are mixed, the following reaction occurs:

$$2\text{KI}\,(aq) + \text{Pb}(\text{NO}_3)_2\,(aq) \rightarrow 2\text{KNO}_3\,(aq) + \text{PbI}_2\,(s)$$

There are very strong attractive forces that occur between $\text{Pb}^{2+}$ and $\text{I}^-$ ions and the result is a brilliant yellow precipitate (see figure below). The other product of the reaction, potassium nitrate, remains soluble.

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Video10.7.4: A yellow precipitate of solid lead (II) iodide forms immediately when solutions of lead (II) nitrate and potassium iodide are mixed. (Source)

## Section Summary

- Combination reactions occur when two or more substances combine to form a new substance.
- A definition of decomposition reaction and example reactions are given.
- The activity series describes the relative reactivities of metals and halogens.
- The double-replacement reaction generally takes the form of AB + CD → AD + CB where A and C are positively-charged cations, while B and D are negatively-charged anions.
- In a double replacement reactions, typically one of the products is a precipitate, a gas, or a molecular compound.

## Glossary

**combination reaction**
: A reaction in which two or more substances combine to form a single new substance.

**decomposition reaction**
: A reaction in which a compound breaks down into two or more simpler substances.

**single-replacement reaction**
: A reaction in which one element replaces a similar element in a compound.

**double-replacement reaction**
: A reaction in which the positive and negative ions of two ionic compounds exchange places to form two new compounds.

This page titled 10.7: Classifying Reactions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.8: Acid Base Reactions

**Learning Objectives**

- Identify acids and bases both in terms of their characteristic properties and in terms of their chemical reactivity.
- Identify acidic or basic solutions using the pH scale.

You are probably familiar with the terms acid, base, and pH. It turns out that acids and bases show up in our daily lives all the time: the foods we eat, the cleaning supplies we use, and even breathing involve acid base chemistry. In the section we will discuss these concepts in more detail as we explore acid base reactions.

## Acids

Acids are very common in some of the foods that we eat. Citrus fruits such as oranges and lemons contain citric acid and ascorbic acid, which is better known as vitamin C. Carbonated sodas contain phosphoric acid. Vinegar contains acetic acid. Your own stomach utilizes hydrochloric acid to digest food.

Acids are a distinct class of compounds because of the properties of their aqueous solutions. These properties are:

1. Aqueous solutions of acids are electrolytes, meaning that they conduct electrical current. Some acids are strong electrolytes because they ionize completely in water, yielding a great many ions. Other acids are weak electrolytes that exist primarily in a non-ionized form when dissolved in water.
2. Acids have a sour taste. Lemons, vinegar, and sour candies all contain acids.
3. Acids change the color of certain acid-base indicates. Two common indicators are litmus and phenolphthalein. Blue litmus turns red in the presence of an acid, while phenolphthalein turns colorless.
4. Acids react with active metals to yield hydrogen gas. Recall that an activity series is a list of metals in descending order of reactivity. Metals that are above hydrogen in the activity series will replace the hydrogen from an acid in a single-replacement reaction, as shown below:

$$\mathrm{Zn}\,(s) + \mathrm{H_2SO_4}\,(aq) \rightarrow \mathrm{ZnSO_4}\,(aq) + \mathrm{H_2}\,(g)$$

5. Acids react with bases to produce a salt compound and water. When equal moles of an acid and a base are combined, the acid is neutralized by the base. The products of this reaction are an ionic compound, which is labeled as a salt, and water.

## Bases

Bases have properties that mostly contrast with those of acids.

1. Aqueous solutions of bases are also electrolytes. Bases can be either strong or weak, just as acids can.
2. Bases often have a bitter taste and are found in foods less frequently than acids. Many bases, like soaps, are slippery to the touch.
3. Bases also change the color of indicators. Litmus turns blue in the presence of a base while phenolphthalein turns pink.
4. Bases do not react with metals in the way that acids do.
5. Bases react with acids to produce a salt and water.

CK-12 LICENSED UNDER CC BY NC SA
© CK-12 Foundation • Visit us at ck12.org



***Figure*** $10.8.1$***:*** *Phenolphthalein indicator in presence of base.*

Please note that tasting chemicals and touching them are NOT good lab practices and should be avoided—in other words, do not do this at home.

Bases are less common as foods, but they are nonetheless present in many household products. Many cleaners contain ammonia, a base. Sodium hydroxide is found in drain cleaner. **Antacids**, which combat excess stomach acid, are comprised of bases such as magnesium hydroxide or sodium hydrogen carbonate.

## Arrhenius Acids and Bases

Swedish chemist Svante Arrhenius (1859-1927) was the first to propose a theory to explain the observed behavior of acids and bases. Because of their ability to conduct a current, he knew that both acids and bases contained ions in solution. An **Arrhenius acid** is a compound that ionizes to yield hydrogen ions $(\mathrm{H^+})$ in aqueous solution. An **Arrhenius base** is a compound that ionizes to yield hydroxide ions $(\mathrm{OH^-})$ in aqueous solution.

Acids are molecular compounds with ionizable hydrogen atoms. Only hydrogen atoms that are part of a highly polar covalent bond are ionizable. Hydrogen chloride $(\mathrm{HCl})$ is a gas at room temperature and under normal pressure. The $\mathrm{H{-}Cl}$ bond in hydrogen chloride is a polar bond. The hydrogen atom is electron deficient because of the higher electronegativity of the chlorine atom. Consequently, the hydrogen atom is attracted to the lone pair of electrons in a water molecule when $\mathrm{HCl}$ is dissolved in water. The result is that the $\mathrm{H{-}Cl}$ bond breaks, with both bonding electrons remaining with the $\mathrm{Cl}$, forming a chloride ion. The $\mathrm{H^+}$ ion attaches to the water molecule, forming a polyatomic ion called the hydronium ion. The **hydronium ion** $(\mathrm{H_3O^+})$ can be thought of as a water molecule with an extra attached hydrogen ion.

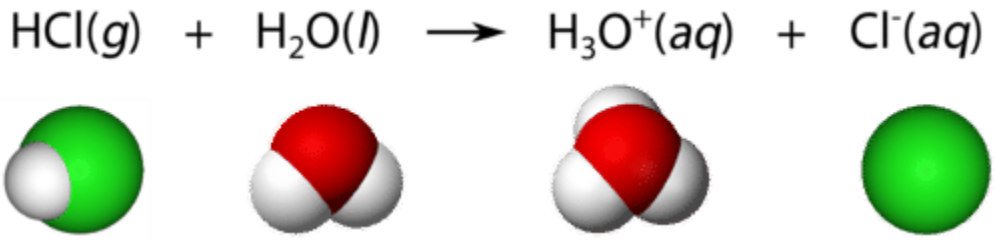



Figure $10.8.1$: Formation of a hydronium ion. (CC BY-NC; CK-12)

Equations showing the ionization of an acid in water are frequently simplified by omitting the water molecule:

$$\mathrm{HCl}\,(g) \rightarrow \mathrm{H^+}\,(aq) + \mathrm{Cl^-}\,(aq)$$

This is merely a simplification of the previous equation, but it is commonly used. Any hydrogen ions in an aqueous solution will be attached to water molecules as hydronium ions.

# The pH Scale

Expressing the acidity of a solution by using the molarity of the hydrogen ion is cumbersome because the quantities are generally very small. Danish scientist Søren Sørensen (1868-1939) proposed an easier system for indicating the concentration of $H^+$ called the pH scale. The letters pH stand for the power of the hydrogen ion. The **pH** of a solution is the negative logarithm of the hydrogen-ion concentration:

$$\text{pH} = -\log\ [\text{H}^+]$$

In pure water or a neutral solution, the $[\text{H}^+] = 1.0 \times 10^{-7}\ \text{M}$. Substituting into the pH expression:

$$\text{pH} = -\log\left[1.0 \times 10^{-7}\right] = -(-7.00) = 7.00$$

The pH of pure water or any neutral solution is thus 7.00. For recording purposes, the numbers to the right of the decimal point in the pH value are the significant figures. Since $1.0 \times 10^{-7}$ has two significant figures, the pH can be reported as 7.00.

A logarithmic scale condenses the range of acidity to numbers that are easy to use. Consider a solution with $[\text{H}^+] = 1.0 \times 10^{-4}\ \text{M}$. That is a hydrogen-ion concentration that is 1000 times higher than the concentration in pure water. The pH of such a solution is 4.00, a difference of just 3 pH units. Notice that when the $[\text{H}^+]$ is written in scientific notation and the coefficient is 1, the pH is simply the exponent with the sign changed. The pH of a solution with $[\text{H}^+] = 1 \times 10^{-2}\ \text{M}$ is 2 and the pH of a solution with $[\text{H}^+] = 1 \times 10^{-10}\ \text{M}$ is 10.

As we saw earlier, a solution with $[\text{H}^+]$ higher than $1.0 \times 10^{-7}$ is acidic, while a solution with $[\text{H}^+]$ lower than $1.0 \times 10^{-7}$ is basic. Consequently, solutions with a pH of less than 7 are acidic, while those with a pH higher than 7 are basic. Figure 10.8.1 illustrates this relationship, along with some examples of various solutions.

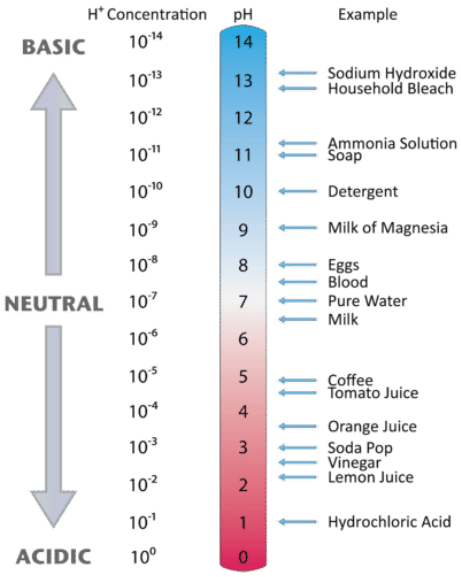



Figure 10.8.1: The pH values for several common materials. (CC BY-NC; CK-12)

# Section Summary

- Characteristic properties of acids and bases are provided.
- An Arrhenius acid is a compound that ionizes to yield hydrogen ions $(\text{H}^+)$ in aqueous solution.
- An Arrhenius base is a compound that ionizes to yield hydroxide ions (OH–)(OH–) in aqueous solution.
- The pH of a solution is the negative logarithm of the hydrogen-ion concentration.
- Solutions with a pH of less than 7 are acidic, while those with a pH higher than 7 are basic.
- pH values for several common materials are listed.

CK-12 LICENSED UNDER CC BY NC
© CK-12 Foundation • Visit us at ck12.org

# Glossary

**Arrhenius acid**
: A compound that ionizes to yield hydrogen ions $(H^+)$ in aqueous solution.

**Arrhenius base**
: A compound that ionizes to yield hydroxide ions $(OH^-)$ in aqueous solution.

**hydronium ion** $(H_3O^+)$
: A water molecule with an extra hydrogen ion attached to it.

**pH**
: The negative logarithm of the hydrogen-ion concentration.

---

This page titled 10.8: Acid Base Reactions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.33: Images Formed by Refraction (Problems Q2.3.1) by OpenStax is licensed CC BY 4.0.
- 2.34: Images Formed by Refraction (Problems Q2.3.2) by OpenStax is licensed CC BY 4.0.
- 2.35: Images Formed by Refraction (Problems Q2.3.3) by OpenStax is licensed CC BY 4.0.
- 2.41: Thin Lenses (Conceptual Questions Q2.4.3) by OpenStax is licensed CC BY 4.0.
- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.9: Polymerization Reactions

**Learning Objectives**

- Explain the difference between an addition polymer and a condensation polymer.

We reap the benefits of using Styrofoam containers, but do not often consider where they end up. Styrofoam materials do not break down quickly under exposure to the elements. When buried in a landfill, styrofoam will remain intact for a long time. The good news is that there is not a lot of this pollutant found in landfills (maybe about $0.5\%$ by weight of the total mass of garbage). There is no good way to recycle Styrofoam at present, but in the future, a creative scientist may change that. Styrofoam is an example of a polymer.

Polymers are very different from the other kinds of organic molecules that you have seen so far. Whereas other compounds are of relatively low molar mass, polymers are giant molecules of very high molar mass. Polymers are the primary components of all sorts of plastics and related compounds. A **polymer** is a large molecule formed of many smaller molecules covalently bonded in a repeating pattern. The small molecules which make up the polymer are called **monomers**. Polymers generally form either from an addition reaction or a condensation reaction. The "small molecules" that polymers are comprised of are often much more complicated than the compounds we have been exploring earlier in this chapter. We will not get into the details of these more advanced structures in this text, but will provide a few examples of types of polymer reactions.

## Addition Polymers

An **addition polymer** is a polymer formed by chain addition reactions between monomers that contain a double bond. Molecules of ethene can polymerize with each other under the right conditions to form the polymer called polyethylene.

$$n\mathrm{CH_2{=}CH_2} \rightarrow -(\mathrm{CH_2CH_2})_n-$$

The letter $n$ stands for the number of monomers that are joined in repeated fashion to make the polymer, and can have a value in the hundreds or even thousands.

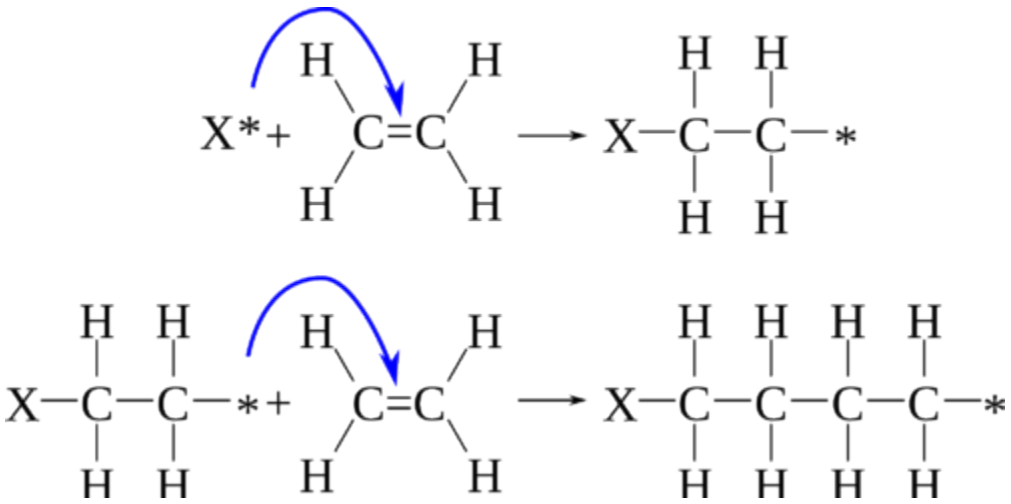


Figure $10.9.1$: Polyethylene synthesis.

Polyethylene can have different properties depending on the length of the polymer chains, and on how efficiently they pack together. Some common products made from different forms of polyethylene include plastic bottles, plastic bags, and harder plastic objects such as milk crates.

Several other kinds of unsaturated monomers can be polymerized, and are components in common household products. Polypropylene is stiffer than polyethylene, and is in plastic utensils and some other types of containers.

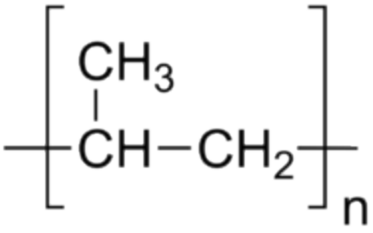



Figure 10.9.2: Polypropylene structure.

Polystyrene is used in insulation and in molded items such as coffee cups.

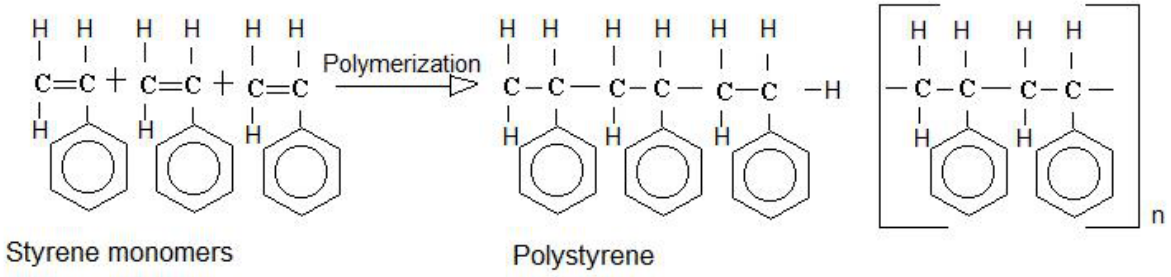



Figure 10.9.3: Polystyrene synthesis and structure.

Polyvinyl chloride (PVC) is extensively used for plumbing pipes.

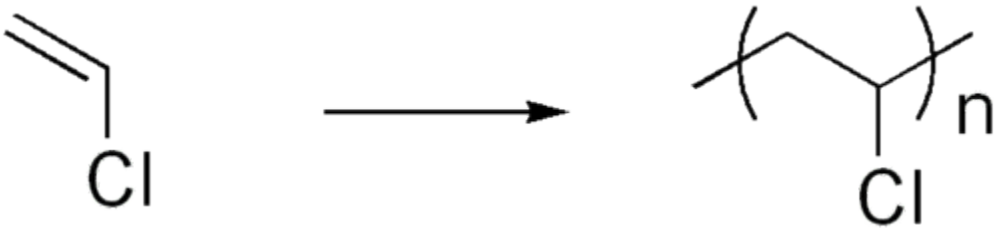



Figure 10.9.4: Polyvinyl chloride.

Polyisoprene is a polymer of isoprene and is better known as rubber. It is produced naturally by rubber trees, but several variants have been developed which demonstrate improvements on the properties of natural rubber.

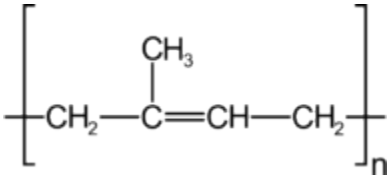



Figure 10.9.5: Polyisoprene.

## Condensation Polymers

A **condensation polymer** is a polymer formed by condensation reactions. Monomers of condensation polymers must contain two functional groups so that each monomer can link up with two other monomers. One type of condensation polymer is called a polyamide. An amide is characterized by the functional group shown below wherein the carbon of a carbonyl group is bonded to the nitrogen of an amine.

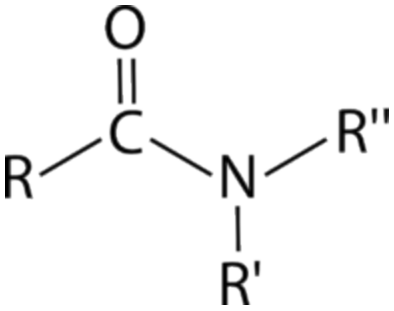



Figure 10.9.1: Amide.

One pair of monomers that can form a polyamide is that of adipic acid and hexanediamine. Adipic acid is a carboxylic acid with two carboxyl groups on either end of the molecule. Hexanediamine has amino groups on either end of a six-carbon chain. When these molecules react with each other, a molecule of water is eliminated, classifying it as a condensation reaction (see figure below).

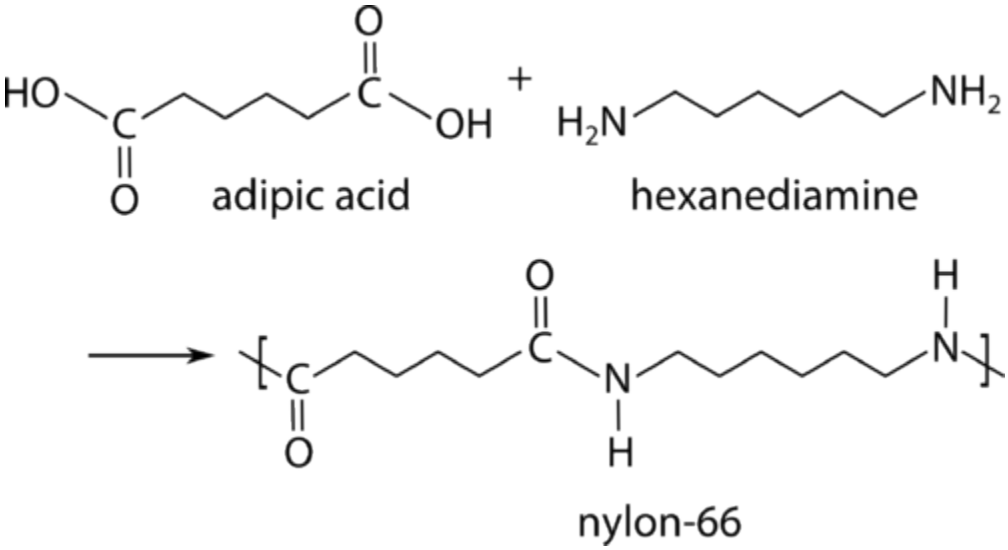


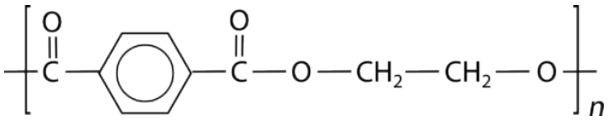

Figure 10.9.2: Nylon synthesis.

The polymer that results from the repetition of the condensation reaction is a polyamide called nylon-66. Nylon-66 was first invented in 1935 and has been used in all sorts of products. Polyamides, including Nylon-66, are commonly found in fibers and clothing, cooking utensils, fishing line, and carpeting—among many other applications.



Figure 10.9.3: Nylon spatula.

Polyester is another common type of condensation polymer. Recall that esters are formed from the reaction of an alcohol with a carboxylic acid. When both the acid and alcohol have two functional groups, the ester is capable of being polymerized. One such polyester is called polyethylene terephthalate (PET) and is formed from the reaction of ethylene glycol with terephthalic acid. The structure of PET is shown below.



PET is used in tires, photographic film, food packaging, and clothing. Polyester fabric is used in permanent-press clothing. Its resistance to wrinkling comes from the cross-linking of the polymer strands.

## Section Summary

- A polymer is a large molecule formed of many smaller molecules covalently bonded in a repeating pattern; they are the primary components of all sorts of plastics and related compounds.
- The small molecules which make up polymers are called monomers.
- Polymers generally form either from an addition reaction or a condensation reaction.
- An addition polymer is a polymer formed by chain addition reactions between monomers that contain a double bond.
- A condensation polymer is a polymer formed by condensation reactions which release a small molecule such as water.

## Glossary

**polymer**
: A large molecule formed of many smaller molecules covalently bonded in a repeating pattern.

**monomers**
: The small molecules which make up a polymer.

**addition polymer**
: A polymer formed by chain addition reactions between monomers that contain a double bond.

**condensation polymer**
: a polymer formed by condensation reactions, such as a reaction that forms water molecules as a side product.

This page titled 10.9: Polymerization Reactions is shared under a CK-12 license and was authored, remixed, and/or curated by Jamie MacArthur via source content that was edited to the style and standards of the LibreTexts platform.

- 2.5: Acceleration by Andrew Park is licensed CC BY 4.0.

# 10.10: Determining Empirical and Molecular Formulas

## Learning Objectives

- Compute the percent composition of a compound
- Determine the empirical formula of a compound
- Determine the molecular formula of a compound

In the previous section, we discussed the relationship between the bulk mass of a substance and the number of atoms or molecules it contains (moles). Given the chemical formula of the substance, we were able to determine the amount of the substance (moles) from its mass, and vice versa. But what if the chemical formula of a substance is unknown? In this section, we will explore how to apply these very same principles in order to derive the chemical formulas of unknown substances from experimental mass measurements.

## Percent Composition

The elemental makeup of a compound defines its chemical identity, and chemical formulas are the most succinct way of representing this elemental makeup. When a compound's formula is unknown, measuring the mass of each of its constituent elements is often the first step in the process of determining the formula experimentally. The results of these measurements permit the calculation of the compound's percent composition, defined as the percentage by mass of each element in the compound. For example, consider a gaseous compound composed solely of carbon and hydrogen. The percent composition of this compound could be represented as follows:

$$\%\mathrm{H} = \frac{\text{mass H}}{\text{mass compound}} \times 100\%$$

$$\%\mathrm{C} = \frac{\text{mass C}}{\text{mass compound}} \times 100\%$$

If analysis of a 10.0-g sample of this gas showed it to contain 2.5 g H and 7.5 g C, the percent composition would be calculated to be 25% H and 75% C:

$$\%\mathrm{H} = \frac{2.5\ \text{g H}}{10.0\ \text{g compound}} \times 100\% = 25\%$$

$$\%\mathrm{C} = \frac{7.5\ \text{g C}}{10.0\ \text{g compound}} \times 100\% = 75\%$$

### Example 10.10.1: Calculation of Percent Composition

Analysis of a 12.04-g sample of a liquid compound composed of carbon, hydrogen, and nitrogen showed it to contain 7.34 g C, 1.85 g H, and 2.85 g N. What is the percent composition of this compound?

**Solution**

To calculate percent composition, we divide the experimentally derived mass of each element by the overall mass of the compound, and then convert to a percentage:

$$\%\mathrm{C} = \frac{7.34\ \text{g C}}{12.04\ \text{g compound}} \times 100\% = 61.0\%$$

$$\%\mathrm{H} = \frac{1.85\ \text{g H}}{12.04\ \text{g compound}} \times 100\% = 15.4\%$$

$$\%\mathrm{N} = \frac{2.85\ \mathrm{g\ N}}{12.04\ \mathrm{g\ compound}} \times 100\% = 23.7\%$$

The analysis results indicate that the compound is 61.0% C, 15.4% H, and 23.7% N by mass.

**Exercise 10.10.1**

A 24.81-g sample of a gaseous compound containing only carbon, oxygen, and chlorine is determined to contain 3.01 g C, 4.00 g O, and 17.81 g Cl. What is this compound's percent composition?

*Answer available at the end of this section.*

## Determining Percent Composition from Formula Mass

Percent composition is also useful for evaluating the relative abundance of a given element in different compounds of known formulas. As one example, consider the common nitrogen-containing fertilizers ammonia ($NH_3$), ammonium nitrate ($NH_4NO_3$), and urea ($CH_4N_2O$). The element nitrogen is the active ingredient for agricultural purposes, so the mass percentage of nitrogen in the compound is a practical and economic concern for consumers choosing among these fertilizers. For these sorts of applications, the percent composition of a compound is easily derived from its formula mass and the atomic masses of its constituent elements. A molecule of $NH_3$ contains one N atom weighing 14.01 amu and three H atoms weighing a total of (3 × 1.008 amu) = 3.024 amu. The formula mass of ammonia is therefore (14.01 amu + 3.024 amu) = 17.03 amu, and its percent composition is:

$$\%\mathrm{N} = \frac{14.01\ \mathrm{amu\ N}}{17.03\ \mathrm{amu\ NH_3}} \times 100\% = 82.27\%$$

$$\%\mathrm{H} = \frac{3.024\ \mathrm{amu\ N}}{17.03\ \mathrm{amu\ NH_3}} \times 100\% = 17.76\%$$

This same approach may be taken considering a pair of molecules, a dozen molecules, or a mole of molecules, etc. The latter amount is most convenient and would simply involve the use of molar masses instead of atomic and formula masses, as demonstrated Example 10.10.2. As long as we know the chemical formula of the substance in question, we can easily derive percent composition from the formula mass or molar mass.

**Example 10.10.2: Determining Percent Composition from a Molecular Formula**

Aspirin is a compound with the molecular formula $C_9H_8O_4$. What is its percent composition?

**Solution**

To calculate the percent composition, we need to know the masses of C, H, and O in a known mass of $C_9H_8O_4$. It is convenient to consider 1 mol of $C_9H_8O_4$ and use its molar mass (180.159 g/mole, determined from the chemical formula) to calculate the percentages of each of its elements:

$$\%\mathrm{C} = \frac{9\ \mathrm{mol\ C} \times \mathrm{molar\ mass\ C}}{\mathrm{molar\ mass\ C_9H_{18}O_4}} \times 100 = \frac{9 \times 12.01\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100 = \frac{108.09\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100$$
$$\%\mathrm{C} = 60.00\,\%\,\mathrm{C}$$

$$\%\mathrm{H} = \frac{8\ \mathrm{mol\ H} \times \mathrm{molar\ mass\ H}}{\mathrm{molar\ mass\ C_9H_{18}O_4}} \times 100 = \frac{8 \times 1.008\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100 = \frac{8.064\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100$$
$$\%\mathrm{H} = 4.476\,\%\,\mathrm{H}$$

$$\%\mathrm{O} = \frac{4\ \mathrm{mol\ O} \times \mathrm{molar\ mass\ O}}{\mathrm{molar\ mass\ C_9H_{18}O_4}} \times 100 = \frac{4 \times 16.00\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100 = \frac{64.00\ \mathrm{g/mol}}{180.159\ \mathrm{g/mol}} \times 100$$
$$\%\mathrm{O} = 35.52\%$$

Note that these percentages sum to equal 100.00% when appropriately rounded.

**Exercise 10.10.2**

To three significant digits, what is the mass percentage of iron in the compound $Fe_2O_3$?

*Answer available at the end of this section.*

## Determination of Empirical Formulas

As previously mentioned, the most common approach to determining a compound's chemical formula is to first measure the masses of its constituent elements. However, we must keep in mind that chemical formulas represent the relative *numbers*, not masses, of atoms in the substance. Therefore, any experimentally derived data involving mass must be used to derive the corresponding numbers of atoms in the compound. To accomplish this, we can use molar masses to convert the mass of each element to a number of moles. We then consider the moles of each element relative to each other, converting these numbers into a whole-number ratio that can be used to derive the empirical formula of the substance. Consider a sample of compound determined to contain 1.71 g C and 0.287 g H. The corresponding numbers of atoms (in moles) are:

$$1.71\ \text{g C} \times \frac{1\ \text{mol C}}{12.01\ \text{g C}} = 0.142\ \text{mol C}$$

$$0.287\ \text{g H} \times \frac{1\ \text{mol H}}{1.008\ \text{g H}} = 0.284\ \text{mol H}$$

Thus, we can accurately represent this compound with the formula $C_{0.142}H_{0.284}$. Of course, per accepted convention, formulas contain whole-number subscripts, which can be achieved by dividing each subscript by the smaller subscript:

$$\text{C}_{\frac{0.142}{0.142}}\ \text{H}_{\frac{0.284}{0.142}}\ \text{or}\ \text{CH}_2$$

(Recall that subscripts of "1" are not written, but rather assumed if no other number is present.)

The empirical formula for this compound is thus $CH_2$. This may or not be the compound's *molecular formula* as well; however, we would need additional information to make that determination (as discussed later in this section).

Consider as another example a sample of compound determined to contain 5.31 g Cl and 8.40 g O. Following the same approach yields a tentative empirical formula of:

$$\text{Cl}_{0.150}\text{O}_{0.525} = \text{Cl}_{\frac{0.150}{0.150}}\ \text{O}_{\frac{0.525}{0.150}} = \text{ClO}_{3.5}$$

In this case, dividing by the smallest subscript still leaves us with a decimal subscript in the empirical formula. To convert this into a whole number, we must multiply each of the subscripts by two, retaining the same atom ratio and yielding $Cl_2O_7$ as the final empirical formula.

**Procedure**

In summary, empirical formulas are derived from experimentally measured element masses by:

1. Deriving the number of moles of each element from its mass
2. Dividing each element's molar amount by the smallest molar amount to yield subscripts for a tentative empirical formula
3. Multiplying all coefficients by an integer, if necessary, to ensure that the smallest whole-number ratio of subscripts is obtained

Figure 10.10.1 outlines this procedure in flow chart fashion for a substance containing elements A and X.

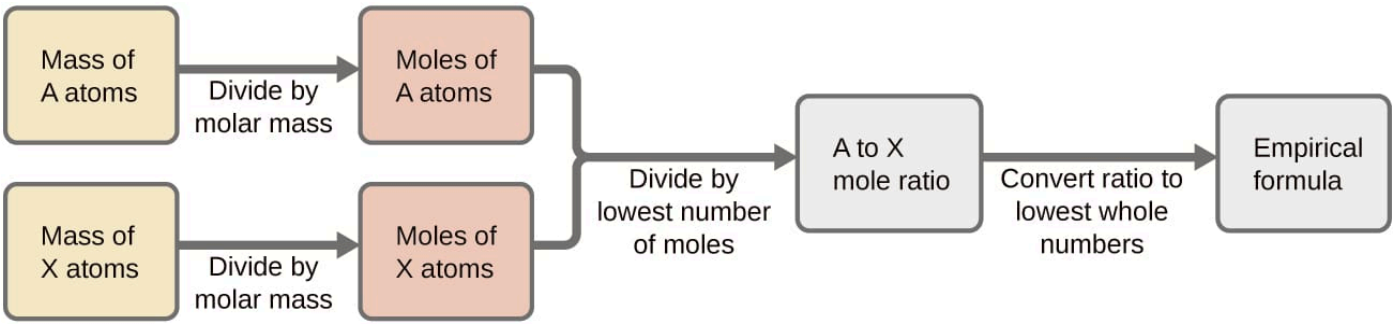



Figure 10.10.1: The empirical formula of a compound can be derived from the masses of all elements in the sample. A flow chart is shown that is composed of six boxes, two of which are connected together by a right facing arrow and located above two more that are also connected by a right-facing arrow. These two rows of boxes are connected vertically by a line that leads to a right-facing arrow and the last two boxes, connected by a final right facing arrow. The first two upper boxes have the phrases, “Mass of A atoms” and “Moles of A atoms” respectively, while the arrow that connects them has the phrase, “Divide by molar mass,” written below it. The second two bottom boxes have the phrases, “Mass of X atoms” and “Moles of X atoms” respectively, while the arrow that connects them has the phrase, “Divide by molar mass” written below it. The arrow that connects the upper and lower boxes to the last two boxes has the phrase “Divide by lowest number of moles” written below it. The last two boxes have the phrases, “A to X mole ratio” and “Empirical formula” respectively, while the arrow that connects them has the phrase, “Convert ratio to lowest whole numbers” written below it.

## Example 10.10.3: Determining an Empirical Formula from Masses of Elements

A sample of the black mineral hematite (Figure 10.10.2), an oxide of iron found in many iron ores, contains 34.97 g of iron and 15.03 g of oxygen. What is the empirical formula of hematite?



Figure 10.10.2: Hematite is an iron oxide that is used in jewelry. (credit: Mauro Cateb)

**Solution**

For this problem, we are given the mass in grams of each element. Begin by finding the moles of each:

$$34.97\ \text{g Fe}\left(\frac{\text{mol Fe}}{55.85\ \text{g}}\right) = 0.6261\ \text{mol Fe}$$

$$15.03\ \text{g O}\left(\frac{\text{mol O}}{16.00\ \text{g}}\right) = 0.9394\ \text{mol O}$$

Next, derive the iron-to-oxygen molar ratio by dividing by the lesser number of moles:

$$\frac{0.6261}{0.6261} = 1.000\ \text{mol Fe}$$

$$\frac{0.9394}{0.6261} = 1.500 \text{ mol O}$$

The ratio is 1.000 mol of iron to 1.500 mol of oxygen ($Fe_1O_{1.5}$). Finally, multiply the ratio by two to get the smallest possible whole number subscripts while still maintaining the correct iron-to-oxygen ratio:

$$2(Fe_1O_{1.5}) = Fe_2O_3$$

The empirical formula is $Fe_2O_3$.

### Exercise 10.10.3

What is the empirical formula of a compound if a sample contains 0.130 g of nitrogen and 0.370 g of oxygen?

*Answer available at the end of this section.*

### YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

***Video*** 10.10.1*: Additional worked examples illustrating the derivation of empirical formulas are presented in the brief video clip.*

## Deriving Empirical Formulas from Percent Composition

Finally, with regard to deriving empirical formulas, consider instances in which a compound's percent composition is available rather than the absolute masses of the compound's constituent elements. In such cases, the percent composition can be used to calculate the masses of elements present in any convenient mass of compound; these masses can then be used to derive the empirical formula in the usual fashion.

### Example 10.10.4: Determining an Empirical Formula from Percent Composition

The bacterial fermentation of grain to produce ethanol forms a gas with a percent composition of 27.29% C and 72.71% O (Figure 10.10.3). What is the empirical formula for this gas?



Figure 10.10.3: An oxide of carbon is removed from these fermentation tanks through the large copper pipes at the top. (credit: "Dual Freq"/Wikimedia Commons) A picture is shown of four copper-colored industrial containers with a large pipe connecting to the top of each one.

**Solution**

Since the scale for percentages is 100, it is most convenient to calculate the mass of elements present in a sample weighing 100 g. The calculation is "most convenient" because, per the definition for percent composition, the mass of a given element in grams is numerically equivalent to the element's mass percentage. This numerical equivalence results from the definition of the "percentage" unit, whose name is derived from the Latin phrase *per centum* meaning "by the hundred." Considering this definition, the mass percentages provided may be more conveniently expressed as fractions:

$$27.29\,\%\,\mathrm{C} = \frac{27.29\ \mathrm{g\ C}}{100\ \mathrm{g\ compound}}$$

$$72.71\,\%\,\mathrm{O} = \frac{72.71\ \mathrm{g\ O}}{100\ \mathrm{g\ compound}}$$

The molar amounts of carbon and hydrogen in a 100-g sample are calculated by dividing each element's mass by its molar mass:

$$27.29\ \mathrm{g\ C}\left(\frac{\mathrm{mol\ C}}{12.01\ \mathrm{g}}\right) = 2.272\ \mathrm{mol\ C}$$

$$72.71\ \mathrm{g\ O}\left(\frac{\mathrm{mol\ O}}{16.00\ \mathrm{g}}\right) = 4.544\ \mathrm{mol\ O}$$

Coefficients for the tentative empirical formula are derived by dividing each molar amount by the lesser of the two:

$$\frac{2.272\ \mathrm{mol\ C}}{2.272} = 1$$

$$\frac{4.544\ \mathrm{mol\ O}}{2.272} = 2$$

Since the resulting ratio is one carbon to two oxygen atoms, the empirical formula is $CO_2$.

## Exercise 10.10.4

What is the empirical formula of a compound containing 40.0% C, 6.71% H, and 53.28% O?

*Answer available at the end of this section.*

## Derivation of Molecular Formulas

Recall that empirical formulas are symbols representing the *relative* numbers of a compound's elements. Determining the *absolute* numbers of atoms that compose a single molecule of a covalent compound requires knowledge of both its empirical formula and its molecular mass or molar mass. These quantities may be determined experimentally by various measurement techniques. Molecular mass, for example, is often derived from the mass spectrum of the compound (see discussion of this technique in the previous chapter on atoms and molecules). Molar mass can be measured by a number of experimental methods, many of which will be introduced in later chapters of this text.

Molecular formulas are derived by comparing the compound's molecular or molar mass to its empirical formula mass. As the name suggests, an empirical formula mass is the sum of the average atomic masses of all the atoms represented in an empirical formula. If we know the molecular (or molar) mass of the substance, we can divide this by the empirical formula mass in order to identify the number of empirical formula units per molecule, which we designate as *n*:

$$\frac{\text{molecular or molar mass}\left(\text{amu or }\frac{\text{g}}{\text{mol}}\right)}{\text{empirical formula mass}\left(\text{amu or }\frac{\text{g}}{\text{mol}}\right)} = n\text{ formula units/molecule}$$

The molecular formula is then obtained by multiplying each subscript in the empirical formula by *n*, as shown by the generic empirical formula $A_xB_y$:

Double subscripts: use braces to clarify

For example, consider a covalent compound whose empirical formula is determined to be $CH_2O$. The empirical formula mass for this compound is approximately 30 amu (the sum of 12 amu for one C atom, 2 amu for two H atoms, and 16 amu for one O atom). If the compound's molecular mass is determined to be 180 amu, this indicates that molecules of this compound contain six times the number of atoms represented in the empirical formula:

$$\frac{180\text{ amu/molecule}}{30\,\frac{\text{amu}}{\text{formula unit}}} = 6\text{ formula units/molecule}$$

Molecules of this compound are then represented by molecular formulas whose subscripts are six times greater than those in the empirical formula:

$$(CH_2O)_6 = C_6H_{12}O_6$$

Note that this same approach may be used when the molar mass (g/mol) instead of the molecular mass (amu) is used. In this case, we are merely considering one mole of empirical formula units and molecules, as opposed to single units and molecules.

### Example 10.10.5: Determination of the Molecular Formula for Nicotine

Nicotine, an alkaloid in the nightshade family of plants that is mainly responsible for the addictive nature of cigarettes, contains 74.02% C, 8.710% H, and 17.27% N. If 40.57 g of nicotine contains 0.2500 mol nicotine, what is the molecular formula?

**Solution**

Determining the molecular formula from the provided data will require comparison of the compound's empirical formula mass to its molar mass. As the first step, use the percent composition to derive the compound's empirical formula. Assuming a convenient, a 100-g sample of nicotine yields the following molar amounts of its elements:

$$(74.02\ \text{g C})\left(\frac{1\ \text{mol C}}{12.01\ \text{g C}}\right) = 6.163\ \text{mol C} \tag{10.10.1}$$

$$(8.710\ \text{g H})\left(\frac{1\ \text{mol H}}{1.01\ \text{g H}}\right) = 8.624\ \text{mol H} \tag{10.10.2}$$

$$(17.27\ \text{g N})\left(\frac{1\ \text{mol N}}{14.01\ \text{g N}}\right) = 1.233\ \text{mol N} \tag{10.10.3}$$

Next, we calculate the molar ratios of these elements relative to the least abundant element, N.

$$6.163\ \text{mol C} / 1.233\ \text{mol N} = 5$$

$$8.264\ \text{mol H} / 1.233\ \text{mol N} = 7$$

$$1.233\ \text{mol N} / 1.233\ \text{mol N} = 1$$

$$1.233/1.233 = 1.000\ \text{mol N}$$

$$6.163/1.233 = 4.998\ \text{mol C}$$

$$8.624/1.233 = 6.994\ \text{mol H}$$

The C-to-N and H-to-N molar ratios are adequately close to whole numbers, and so the empirical formula is $C_5H_7N$. The empirical formula mass for this compound is therefore 81.13 amu/formula unit, or 81.13 g/mol formula unit.

We calculate the molar mass for nicotine from the given mass and molar amount of compound:

$$\frac{40.57\ \text{g nicotine}}{0.2500\ \text{mol nicotine}} = \frac{162.3\ \text{g}}{\text{mol}}$$

Comparing the molar mass and empirical formula mass indicates that each nicotine molecule contains two formula units:

$$\frac{162.3\ \text{g/mol}}{81.13\ \frac{\text{g}}{\text{formula unit}}} = 2\ \text{formula units/molecule}$$

Thus, we can derive the molecular formula for nicotine from the empirical formula by multiplying each subscript by two:

$$(C_5H_7N)_2 = C_{10}H_{14}N_2$$

### Exercise 10.10.5

What is the molecular formula of a compound with a percent composition of 49.47% C, 5.201% H, 28.84% N, and 16.48% O, and a molecular mass of 194.2 amu?

*Answer available at the end of this section.*

## Summary

The chemical identity of a substance is defined by the types and relative numbers of atoms composing its fundamental entities (molecules in the case of covalent compounds, ions in the case of ionic compounds). A compound's percent composition provides the mass percentage of each element in the compound, and it is often experimentally determined and used to derive the compound's empirical formula. The empirical formula mass of a covalent compound may be compared to the compound's molecular or molar mass to derive a molecular formula.

# Key Equations

- $\%\text{X} = \dfrac{\text{mass X}}{\text{mass compound}} \times 100\%$
- $\dfrac{\text{molecular or molar mass}\left(\text{amu or } \dfrac{\text{g}}{\text{mol}}\right)}{\text{empirical formula mass}\left(\text{amu or } \dfrac{\text{g}}{\text{mol}}\right)} = n \text{ formula units/molecule}$
- $(A_xB_y)_n = A_{nx}B_{ny}$

# Glossary

**percent composition**
: percentage by mass of the various elements in a compound

**empirical formula mass**
: sum of average atomic masses for all atoms represented in an empirical formula

## Answers

**Exercise** 10.10.1
12.1% C, 16.1% O, 71.8% Cl

**Exercise** 10.10.2
69.9% Fe

**Exercise** 10.10.3
$N_2O_5$

**Exercise** 10.10.4
$CH_2O$

**Exercise** 10.10.5
$C_8H_{10}N_4O_2$

This page titled 10.10: Determining Empirical and Molecular Formulas is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.11: Prelude to Stoichiometry

Solid-fuel rockets are a central feature in the world's space exploration programs, including the new Space Launch System being developed by the National Aeronautics and Space Administration (NASA) to replace the retired Space Shuttle fleet (Figure 10.11.1 ). The engines of these rockets rely on carefully prepared solid mixtures of chemicals combined in precisely measured amounts. Igniting the mixture initiates a vigorous chemical reaction that rapidly generates large amounts of gaseous products. These gases are ejected from the rocket engine through its nozzle, providing the thrust needed to propel heavy payloads into space. Both the nature of this chemical reaction and the relationships between the amounts of the substances being consumed and produced by the reaction are critically important considerations that determine the success of the technology. This chapter will describe how to symbolize chemical reactions using chemical equations, how to classify some common chemical reactions by identifying patterns of reactivity, and how to determine the quantitative relations between the amounts of substances involved in chemical reactions—that is, the reaction *stoichiometry*.



Figure 10.11.1: Many modern rocket fuels are solid mixtures of substances combined in carefully measured amounts and ignited to yield a thrust-generating chemical reaction. (credit: modification of work by NASA) An image is shown of a rocket that appears to have just passed through a layer of clouds as it travels skyward. A bright white light is seen in the upper right corner of the image. To the lower left appears the layer of clouds and the bottom of the rocket with fire projecting from the fuel cones at its base.

This page titled 10.11: Prelude to Stoichiometry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.12: Balancing Equations (1)

**Learning Objectives**

- Balance a chemical equation.

## Any leftovers?

When you cook a meal, quite often there are leftovers because you prepared more than people would eat at one sitting. Sometimes when you repair a piece of equipment, you end up with what are called "pocket parts"—small pieces you put in your pocket because you're not sure where they belong. Chemistry tries to avoid leftovers and pocket parts. In normal chemical processes, we cannot create or destroy matter (law of conservation of mass). If we start out with ten carbon atoms, we need to end up with ten carbon atoms. John Dalton's atomic theory said that chemical reactions basically involve the rearrangement of atoms. Chemical equations need to follow these principles in order to be correct.

## Balancing Chemical Equations

A **balanced equation** is a chemical equation in which mass is conserved and there are equal numbers of atoms of each element on both sides of the equation. We can write a chemical equation for the reaction of carbon with hydrogen gas to form methane $(\mathrm{CH_4})$:

$$\underset{2\text{ C atoms}}{\mathrm{C}\,(s)} \quad + \quad \underset{2\text{ H atoms}}{\mathrm{H_2}\,(g)} \quad \rightarrow \quad \underset{1\text{ C atom, }4\text{ H atoms}}{\mathrm{CH_4}\,(g)}$$

In order to write a correct equation, you must first write the correct skeleton equation with the correct chemical formulas. Recall that hydrogen is a diatomic molecule and so is written as $\mathrm{H_2}$.

When we count the number of atoms of both elements, shown under the equation, we see that the equation is not balanced. There are only 2 atoms of hydrogen on the reactant side of the equation, while there are 4 atoms of hydrogen on the product side. We can balance the above equation by adding a coefficient of 2 in front of the formula for hydrogen.

$$\mathrm{C}\,(s) + 2\mathrm{H_2}\,(g) \rightarrow \mathrm{CH_4}\,(g)$$

A **coefficient** is a small in magnitude whole number placed in front of a formula in an equation to signify the quantity needed to balance it. The 2 in front of the $\mathrm{H_2}$ means that there are a total of $2 \times 2 = 4$ atoms of hydrogen as reactants. Visually, the reaction looks like the figure below.

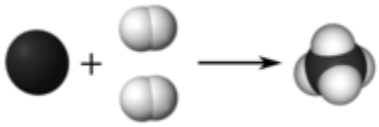


Figure $10.12.2$: Reaction between carbon and hydrogen to form methane. (Credit: Ben Mills (Wikimedia: Benjah-bmm27; Source, and Source); License: Public Domain)

In the balanced equation, there is one atom of carbon and four atoms of hydrogen on both sides of the arrow. Below are guidelines for writing and balancing chemical equations.

1. Determine the correct chemical formulas for each reactant and product.
2. Write the skeleton equation.
3. Count the number of atoms of each element that appears as a reactant and as a product. If a polyatomic ion is unchanged on both sides of the equation, count it as a unit.
4. Balance each element one at a time by placing coefficients in front of the formulas.
   a. It is best to begin by balancing elements that only appear in one chemical formula on each side of the equation.
   b. It is best to end with an element by itself (if possible).

CK-12 LICENSED UNDER CK-12 License
© CK-12 Foundation • Visit us at ck12.org

c. No coefficient is written for a 1.
d. NEVER change the subscripts in a chemical formula—you can only balance equations by using coefficients.

5. Check each atom or polyatomic ion to be sure that they are equal on both sides of the equation.
6. Make sure that all coefficients are in the lowest possible ratio. If necessary, reduce to the lowest ratio.

**Example** 10.12.1**: Balancing Chemical Equations**

Aqueous solutions of lead (II) nitrate and sodium chloride are mixed. The products of the reaction are an aqueous solution of sodium nitrate and a solid precipitate of lead (II) chloride. Write the balanced chemical equation for this reaction.

**Solution**

***Step 1: Plan the problem.***

Follow the steps for writing and balancing a chemical equation listed in the text.

***Step 2: Solve.***

Write the skeleton equation with the correct formulas.

$$\mathrm{Pb(NO_3)_2}\,(aq) + \mathrm{NaCl}\,(aq) \rightarrow \mathrm{NaNO_3}\,(aq) + \mathrm{PbCl_2}\,(s)$$

Count the number of each atom or polyatomic ion on both sides of the equation.

| Reactants | Products |
|---|---|
| 1 Pb atom | 1 Pb atom |
| 2 $\mathrm{NO_3}^-$ ions | 1 $\mathrm{NO_3}^-$ ions |
| 1 Na atom | 1 Na atom |
| 1 Cl atom | 2 Cl atoms |

The nitrate ions and the chlorine atoms are unbalanced. Start by placing a 2 in front of the $\mathrm{NaCl}$. This increases the reactant counts to 2 $\mathrm{Na}$ atoms and 2 $\mathrm{Cl}$ atoms. Then place a 2 in front of the $\mathrm{NaNO_3}$. The result is:

$$\mathrm{Pb(NO_3)_2}\,(aq) + 2\mathrm{NaCl}\,(aq) \rightarrow 2\mathrm{NaNO_3}\,(aq) + \mathrm{PbCl_2}\,(s)$$

The new count for each atom and polyatomic ion becomes:

| Reactants | Products |
|---|---|
| 1 Pb atom | 1 Pb atom |
| 2 $\mathrm{NO_3}^-$ ions | 2 $\mathrm{NO_3}^-$ ions |
| 2 Na atom | 2 Na atom |
| 2 Cl atom | 2 Cl atoms |

***Step 3: Think about your result.***

The equation is now balanced since there are equal numbers of atoms of each element on both sides of the equation.

**MISCONCEPTION ALERT: YOU CAN CHANGE ANY NUMBER YOU WANT TO BALANCE AN EQUATION**

The substances in a balanced equation are described by their chemical formula. The chemical formula will contain subscripts to indicate the number of atoms of a given element if that substance has more than one atom of that element in its formula. If we change subscripts in an attempt to balance an equation, we are actually changing the identity of the substances in the chemical reaction that the balanced equation is about. Imagine if you wanted a drink of water and when you put it in your mouth realized that it was actually hydrogen peroxide! The difference between water ($H_2O$) and hydrogen peroxide ($H_2O_2$) is significant, but the difference in the chemical formula is merely a subscript. Any time you change the subscript in a chemical equation you are making a similar change to the substances involved. This isn't balancing the equation, it is saying something else instead of answering the question.

CK-12 LICENSED UNDER CK-12 Science © CK-12 Foundation • Visit us at ck12.org

YouTube Video

*This interactive element is not available in this format. Visit the link or scan the QR code to access it online.*

View this YouTube Video online

Simulation

Practice balancing chemical equations with this simulation:

Can you balance a chemical equation?

## Section Summary

- The process of balancing chemical equations is described.

This page titled 10.12: Balancing Equations (1) is shared under a CK-12 license and was authored, remixed, and/or curated by CK12 via source content that was edited to the style and standards of the LibreTexts platform.

- 11.2: Magnets by Andrew Park is licensed CC BY 4.0.

# 10.13: Writing and Balancing Chemical Equations

## Learning Objectives

- Derive chemical equations from narrative descriptions of chemical reactions.
- Write and balance chemical equations in molecular, total ionic, and net ionic formats.

The preceding chapter introduced the use of element symbols to represent individual atoms. When atoms gain or lose electrons to yield ions, or combine with other atoms to form molecules, their symbols are modified or combined to generate chemical formulas that appropriately represent these species. Extending this symbolism to represent both the identities and the relative quantities of substances undergoing a chemical (or physical) change involves writing and balancing a chemical equation. Consider as an example the reaction between one methane molecule ($CH_4$) and two diatomic oxygen molecules ($O_2$) to produce one carbon dioxide molecule ($CO_2$) and two water molecules ($H_2O$). The chemical equation representing this process is provided in the upper half of Figure 10.13.1, with space-filling molecular models shown in the lower half of the figure.

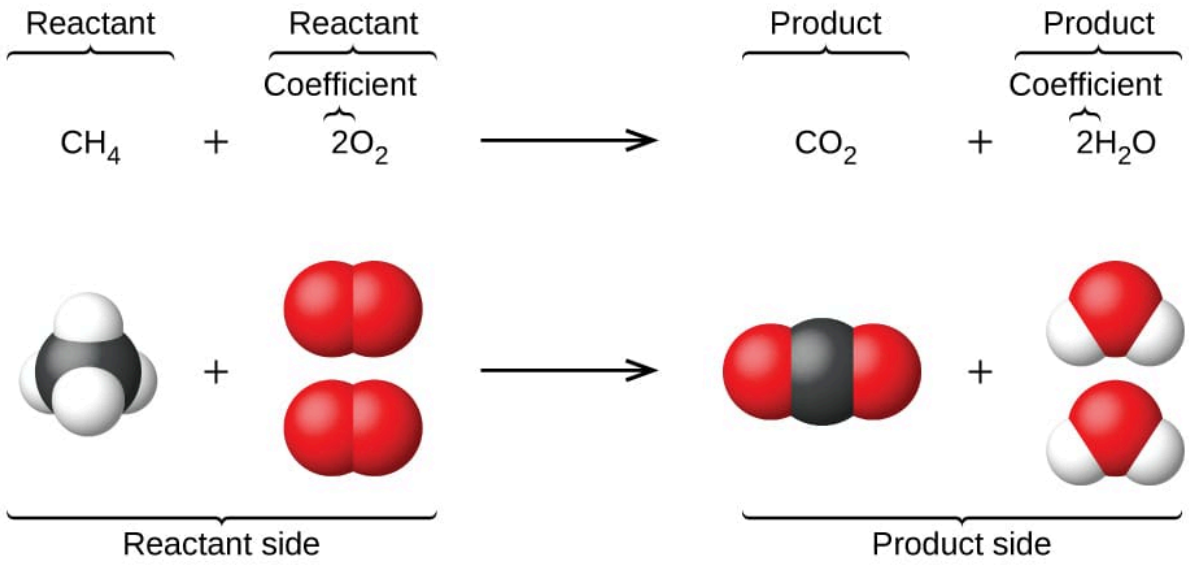



Figure 10.13.1: The reaction between methane and oxygen to yield carbon dioxide and water (shown at bottom) may be represented by a chemical equation using formulas (top). This figure shows a balanced chemical equation followed below by a representation of the equation using space-filling models. The equation reads C H subscript 4 plus 2 O subscript 2 arrow C O subscript 2 plus 2 H subscript 2 O. Under the C H subscript 4, the molecule is shown with a central black sphere, representing a C atom, to which 4 smaller white spheres, representing H atoms, are distributed evenly around. All four H atoms are bonded to the central black C atom. This is followed by a plus sign. Under the 2 O subscript 2, two molecules are shown. The molecules are each composed of two red spheres bonded together. The red spheres represent O atoms. To the right of an arrow and under the C O subscript 2, appears a single molecule with a black central sphere with two red spheres bonded to the left and right. Following a plus sign and under the 2 H subscript 2 O, are two molecules, each with a central red sphere and two smaller white spheres attached to the lower right and lower left sides of the central red sphere. Note that in space filling models of molecules, spheres appear slightly compressed in regions where there is a bond between two atoms.

This example illustrates the fundamental aspects of any chemical equation:

1. The substances undergoing reaction are called reactants, and their formulas are placed on the left side of the equation.
2. The substances generated by the reaction are called products, and their formulas are placed on the right sight of the equation.
3. Plus signs (+) separate individual reactant and product formulas, and an arrow (⟐) separates the reactant and product (left and right) sides of the equation.
4. The relative numbers of reactant and product species are represented by coefficients (numbers placed immediately to the

left of each formula). A coefficient of 1 is typically omitted.

It is common practice to use the smallest possible whole-number coefficients in a chemical equation, as is done in this example. Realize, however, that these coefficients represent the *relative* numbers of reactants and products, and, therefore, they may be correctly interpreted as ratios. Methane and oxygen react to yield carbon dioxide and water in a 1:2:1:2 ratio. This ratio is satisfied if the numbers of these molecules are, respectively, 1-2-1-2, or 2-4-2-4, or 3-6-3-6, and so on (Figure 10.13.2). Likewise, these coefficients may be interpreted with regard to any amount (number) unit, and so this equation may be correctly read in many ways, including:

- *One* methane molecule and *two* oxygen molecules react to yield *one* carbon dioxide molecule and *two* water molecules.
- *One dozen* methane molecules and *two dozen* oxygen molecules react to yield *one dozen* carbon dioxide molecules and *two dozen* water molecules.
- *One mole* of methane molecules and *2 moles* of oxygen molecules react to yield *1 mole* of carbon dioxide molecules and *2 moles* of water molecules.

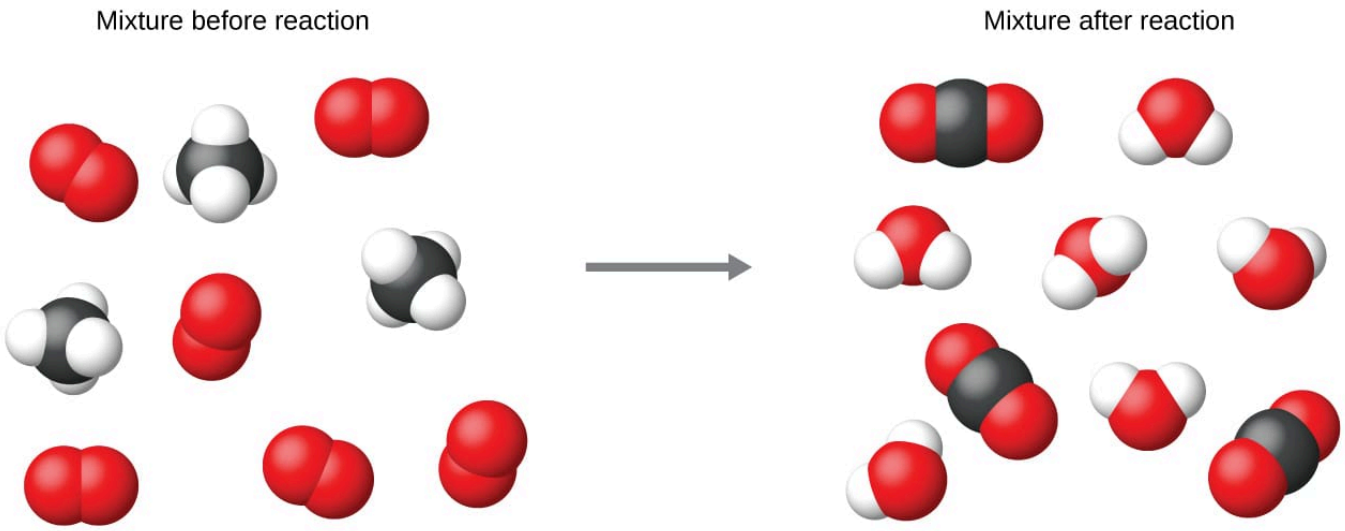



Figure 10.13.2: Regardless of the absolute number of molecules involved, the ratios between numbers of molecules of each species that react (the reactants) and molecules of each species that form (the products) are the same and are given in the chemical equation.

## Balancing Equations

When a chemical equation is balanced it means that equal numbers of atoms for each element involved in the reaction are represented on the reactant and product sides. This is a requirement the equation must satisfy to be consistent with the law of conservation of matter. It may be confirmed by simply summing the numbers of atoms on either side of the arrow and comparing these sums to ensure they are equal. Note that the number of atoms for a given element is calculated by multiplying the coefficient of any formula containing that element by the element's subscript in the formula. If an element appears in more than one formula on a given side of the equation, the number of atoms represented in each must be computed and then added together. For example, both product species in the example reaction, $\mathrm{CO_2}$ and $\mathrm{H_2O}$, contain the element oxygen, and so the number of oxygen atoms on the product side of the equation is

$$\left(1\ \cancel{\mathrm{CO_2\ molecule}} \times \frac{2\ \mathrm{O\ atoms}}{\cancel{\mathrm{CO_2\ molecule}}}\right) + \left(2\ \cancel{\mathrm{H_2O\ molecule}} \times \frac{1\ \mathrm{O\ atom}}{\cancel{\mathrm{H_2O\ molecule}}}\right) = 4\ \mathrm{O\ atoms}$$

The equation for the reaction between methane and oxygen to yield carbon dioxide and water is confirmed to be balanced per this approach, as shown here:

$$\mathrm{CH_4} + 2\,\mathrm{O_2} \rightarrow \mathrm{CO_2} + 2\,\mathrm{H_2O}$$

*Equation for the reaction*

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O$$

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| C | 1 × 1 = 1 | 1 × 1 = 1 | 1 = 1, yes |
| H | 4 × 1 = 4 | 2 × 2 = 4 | 4 = 4, yes |
| O | 2 × 2 = 4 | (1 × 2) + (2 × 1) = 4 | 4 = 4, yes |

A balanced chemical equation often may be derived from a qualitative description of some chemical reaction by a fairly simple approach known as balancing by inspection. Consider as an example the decomposition of water to yield molecular hydrogen and oxygen. This process is represented qualitatively by an *unbalanced* chemical equation:

$$H_2O \rightarrow H_2 + O_2 \qquad \text{(unbalanced)}$$

Comparing the number of H and O atoms on either side of this equation confirms its imbalance:

*Comparisons between H and O atoms*

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| H | 1 × 2 = 2 | 1 × 2 = 2 | 2 = 2, yes |
| O | 1 × 1 = 1 | 1 × 2 = 2 | 1 ≠ 2, no |

The numbers of H atoms on the reactant and product sides of the equation are equal, but the numbers of O atoms are not. To achieve balance, the *coefficients* of the equation may be changed as needed. Keep in mind, of course, that the *formula subscripts* define, in part, the identity of the substance, and so these cannot be changed without altering the qualitative meaning of the equation. For example, changing the reactant formula from $H_2O$ to $H_2O_2$ would yield balance in the number of atoms, but doing so also changes the reactant's identity (it's now hydrogen peroxide and not water). The O atom balance may be achieved by changing the coefficient for $H_2O$ to 2.

$$2\,H_2O \rightarrow H_2 + O_2 \qquad \text{(unbalanced)}$$

*O atom balance may be achieved by changing the coefficient for H2O to 2*

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| H | **2** × 2 = 4 | 1 × 2 = 2 | 4 ≠ 2, no |
| O | 2 × 1 = 2 | 1 × 2 = 2 | 2 = 2, yes |

The H atom balance was upset by this change, but it is easily reestablished by changing the coefficient for the $H_2$ product to 2.

$$2\,H_2O \rightarrow 2\,H_2 + O_2 \qquad \text{(balanced)}$$

*H atom balance upset but easily reestablished by changing the coefficient for the H2 product to 2.*

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| H | 2 × 2 = 4 | **2** × 2 = 2 | 4 = 4, yes |
| O | 2 × 1 = 2 | 1 × 2 = 2 | 2 = 2, yes |

These coefficients yield equal numbers of both H and O atoms on the reactant and product sides, and the balanced equation is, therefore:

$$2\,H_2O \rightarrow 2\,H_2 + O_2$$

## Example 10.13.1: Balancing Chemical Equations

Write a balanced equation for the reaction of molecular nitrogen ($N_2$) and oxygen ($O_2$) to form dinitrogen pentoxide.

**Solution**

First, write the unbalanced equation.

$$\mathrm{N_2 + O_2 \rightarrow N_2O_5} \qquad \text{(unbalanced)}$$

Next, count the number of each type of atom present in the unbalanced equation.

*Unbalanced Equation*

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| N | 1 × 2 = 2 | 1 × 2 = 2 | 2 = 2, yes |
| O | 1 × 2 = 2 | 1 × 5 = 5 | 2 ≠ 5, no |

Though nitrogen is balanced, changes in coefficients are needed to balance the number of oxygen atoms. To balance the number of oxygen atoms, a reasonable first attempt would be to change the coefficients for the $O_2$ and $N_2O_5$ to integers that will yield 10 O atoms (the least common multiple for the O atom subscripts in these two formulas).

$$\mathrm{N_2 + 5\,O_2 \rightarrow 2\,N_2O_5} \qquad \text{(unbalanced)}$$

$$\mathrm{N_2 + 5\,O_2 \rightarrow 2\,N_2O_5} \qquad \text{(unbalanced)}$$

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| N | 1 × 2 = 2 | **2** × 2 = 4 | 2 ≠ 4, no |
| O | **5** × 2 = 10 | **2** × 5 = 10 | 10 = 10, yes |

The N atom balance has been upset by this change; it is restored by changing the coefficient for the reactant $N_2$ to 2.

$$\mathrm{2\,N_2 + 5\,O_2 \rightarrow 2\,N_2O_5}$$

*N atom balance upset but restored by changing the coefficient for the reactant N2 to 2.*

| Element | Reactants | Products | Balanced? |
|---|---|---|---|
| N | **2** × 2 = 4 | 2 × 2 = 4 | 4 = 4, yes |
| O | 5 × 2 = 10 | 2 × 5 = 10 | 10 = 10, yes |

The numbers of N and O atoms on either side of the equation are now equal, and so the equation is balanced.

## Exercise 10.13.1

Write a balanced equation for the decomposition of ammonium nitrate to form molecular nitrogen, molecular oxygen, and water. (Hint: Balance oxygen last, since it is present in more than one molecule on the right side of the equation.)

*Answer available at the end of this section.*

### Balancing Reactions Which Contain Polyatomics: Balancing Reactions Which Contain Polyatomics(opens in new window) [youtu.be]

It is sometimes convenient to use fractions instead of integers as intermediate coefficients in the process of balancing a

chemical equation. When balance is achieved, all the equation's coefficients may then be multiplied by a whole number to convert the fractional coefficients to integers without upsetting the atom balance. For example, consider the reaction of ethane ($C_2H_6$) with oxygen to yield $H_2O$ and $CO_2$, represented by the unbalanced equation:

$$\mathrm{C_2H_6 + O_2 \rightarrow H_2O + CO_2} \quad \text{(unbalanced)}$$

Following the usual inspection approach, one might first balance C and H atoms by changing the coefficients for the two product species, as shown:

$$\mathrm{C_2H_6 + O_2 \rightarrow 3\,H_2O + 2\,CO_2} \quad \text{(unbalanced)}$$

This results in seven O atoms on the product side of the equation, an odd number—no integer coefficient can be used with the $O_2$ reactant to yield an odd number, so a fractional coefficient, $\frac{7}{2}$, is used instead to yield a provisional balanced equation:

$$\mathrm{C_2H_6 + \tfrac{7}{2}\,O_2 \rightarrow 3\,H_2O + 2\,CO_2}$$

A conventional balanced equation with integer-only coefficients is derived by multiplying each coefficient by 2:

$$\mathrm{2\,C_2H_6 + 7\,O_2 \rightarrow 6\,H_2O + 4\,CO_2}$$

Finally with regard to balanced equations, recall that convention dictates use of the *smallest whole-number coefficients*. Although the equation for the reaction between molecular nitrogen and molecular hydrogen to produce ammonia is, indeed, balanced,

$$\mathrm{3\,N_2 + 9\,H_2 \rightarrow 6\,NH_3}$$

the coefficients are not the smallest possible integers representing the relative numbers of reactant and product molecules. Dividing each coefficient by the greatest common factor, 3, gives the preferred equation:

$$\mathrm{N_2 + 3\,H_2 \rightarrow 2\,NH_3}$$

**Phet Simulation**

Use this interactive tutorial for additional practice balancing equations.

## Additional Information in Chemical Equations

The physical states of reactants and products in chemical equations very often are indicated with a parenthetical abbreviation following the formulas. Common abbreviations include *s* for solids, *l* for liquids, *g* for gases, and *aq* for substances dissolved in water (*aqueous solutions*, as introduced in the preceding chapter). These notations are illustrated in the example equation here:

$$\mathrm{2\,Na\,(s) + 2\,H_2O\,(l) \rightarrow 2\,NaOH\,(aq) + H_2\,(g)}$$

This equation represents the reaction that takes place when sodium metal is placed in water. The solid sodium reacts with liquid water to produce molecular hydrogen gas and the ionic compound sodium hydroxide (a solid in pure form, but readily dissolved in water).

Special conditions necessary for a reaction are sometimes designated by writing a word or symbol above or below the equation's arrow. For example, a reaction carried out by heating may be indicated by the uppercase Greek letter delta (Δ) over the arrow.

$$\mathrm{CaCO_3}(s) \xrightarrow{\Delta} \mathrm{CaO}(s) + \mathrm{CO_2}(g)$$

Other examples of these special conditions will be encountered in more depth in later chapters.

## Equations for Ionic Reactions

Given the abundance of water on earth, it stands to reason that a great many chemical reactions take place in aqueous media. When ions are involved in these reactions, the chemical equations may be written with various levels of detail appropriate to their

intended use. To illustrate this, consider a reaction between ionic compounds taking place in an aqueous solution. When aqueous solutions of $\mathrm{CaCl_2}$ and $\mathrm{AgNO_3}$ are mixed, a reaction takes place producing aqueous $\mathrm{Ca(NO_3)_2}$ and solid $\mathrm{AgCl}$:

$$\mathrm{CaCl_2}(aq) + 2\,\mathrm{AgNO_3}(aq) \rightarrow \mathrm{Ca(NO_3)_2}(aq) + 2\,\mathrm{AgCl}(s)$$

This balanced equation, derived in the usual fashion, is called a molecular equation because it doesn't explicitly represent the ionic species that are present in solution. When ionic compounds dissolve in water, they may *dissociate* into their constituent ions, which are subsequently dispersed homogenously throughout the resulting solution (a thorough discussion of this important process is provided in the chapter on solutions). Ionic compounds dissolved in water are, therefore, more realistically represented as dissociated ions, in this case:

$$\mathrm{CaCl_2}(aq) \rightarrow \blacksquare(aq) + 2\,\mathrm{Cl^-}(aq)$$

$$2\,\mathrm{AgNO_3}(aq) \rightarrow 2\,\mathrm{Ag^+}(aq) + 2\,\mathrm{NO_3}^-(aq)$$

$$\mathrm{Ca(NO_3)_2}(aq) \rightarrow \mathrm{Ca^{2+}}(aq) + 2\,\mathrm{NO_3}^-(aq)$$

Unlike these three ionic compounds, AgCl does not dissolve in water to a significant extent, as signified by its physical state notation, *(s)*.

Explicitly representing all dissolved ions results in a complete ionic equation. In this particular case, the formulas for the dissolved ionic compounds are replaced by formulas for their dissociated ions:

$$\blacksquare(aq) + 2\,\mathrm{Cl^-}(aq) + 2\,\mathrm{Ag^+}(aq) + 2\,\mathrm{NO_3}^-(aq) \rightarrow \blacksquare(aq) + 2\,\mathrm{NO_3}^-(aq) + 2\,\mathrm{AgCl}(s)$$

Examining this equation shows that two chemical species are present in identical form on both sides of the arrow, $\mathrm{Ca^{2+}\,(aq)}$ and $\mathrm{NO_3}^-(aq)$. These spectator ions—ions whose presence is required to maintain charge neutrality—are neither chemically nor physically changed by the process, and so they may be eliminated from the equation to yield a more succinct representation called a net ionic equation:

$$\cancel{\blacksquare(aq)} + 2\,\mathrm{Cl^-}(aq) + 2\,\mathrm{Ag^+}(aq) + \cancel{2\,\mathrm{NO_3}^-(aq)} \rightarrow \cancel{\blacksquare(aq)} + \cancel{2\,\mathrm{NO_3}^-(aq)} + 2\,\mathrm{AgCl}(s)$$

$$2\,\mathrm{Cl^-}(aq) + 2\,\mathrm{Ag^+}(aq) \rightarrow 2\,\mathrm{AgCl}(s)$$

Following the convention of using the smallest possible integers as coefficients, this equation is then written:

$$\mathrm{Cl^-}(aq) + \mathrm{Ag^+}(aq) \rightarrow \mathrm{AgCl}(s)$$

This net ionic equation indicates that solid silver chloride may be produced from dissolved chloride and silver(I) ions, regardless of the source of these ions. These molecular and complete ionic equations provide additional information, namely, the ionic compounds used as sources of $\mathrm{Cl^-}$ and $\mathrm{Ag^+}$.

### Example 10.13.2: Molecular and Ionic Equations

When carbon dioxide is dissolved in an aqueous solution of sodium hydroxide, the mixture reacts to yield aqueous sodium carbonate and liquid water. Write balanced molecular, complete ionic, and net ionic equations for this process.

**Solution**

Begin by identifying formulas for the reactants and products and arranging them properly in chemical equation form:

$$\mathrm{CO_2\,(aq) + NaOH\,(aq)} \rightarrow \mathrm{Na_2CO_3\,(aq) + H_2O\,(l)} \qquad \text{(unbalanced)}$$

Balance is achieved easily in this case by changing the coefficient for NaOH to 2, resulting in the molecular equation for this reaction:

$$\mathrm{CO_2(aq) + 2\,NaOH(aq)} \rightarrow \mathrm{Na_2CO_3\,(aq) + H_2O}(l)$$

The two dissolved ionic compounds, NaOH and $Na_2CO_3$, can be represented as dissociated ions to yield the complete ionic equation:

$$\mathrm{CO_2\,(aq) + 2\,Na^+\,(aq) + 2\,OH^-\,(aq) \rightarrow 2\,Na^+\,(aq) + {CO_3}^{2-}\,(aq) + H_2O\,(l)}$$

Finally, identify the spectator ion(s), in this case $Na^+(aq)$, and remove it from each side of the equation to generate the net ionic equation:

$$\mathrm{CO_2(aq) + \cancel{2\,Na^+(aq)} + 2\,OH^-(aq) \rightarrow \cancel{2\,Na^+(aq)} + {CO_3}^{2-}(aq) + H_2O(l)}$$
$$\mathrm{CO_2(aq) + 2\,OH^-(aq) \rightarrow {CO_3}^{2-}(aq) + H_2O(l)}$$

## Exercise 10.13.2

Diatomic chlorine and sodium hydroxide (lye) are commodity chemicals produced in large quantities, along with diatomic hydrogen, via the electrolysis of brine, according to the following unbalanced equation:

$$\mathrm{NaCl\,(aq) + H_2O\,(l) \xrightarrow{electricity} NaOH\,(aq) + H_2\,(g) + Cl_2\,(g)}$$

Write balanced molecular, complete ionic, and net ionic equations for this process.

*Answer available at the end of this section.*

# Key Concepts and Summary

Chemical equations are symbolic representations of chemical and physical changes. Formulas for the substances undergoing the change (reactants) and substances generated by the change (products) are separated by an arrow and preceded by integer coefficients indicating their relative numbers. Balanced equations are those whose coefficients result in equal numbers of atoms for each element in the reactants and products. Chemical reactions in aqueous solution that involve ionic reactants or products may be represented more realistically by complete ionic equations and, more succinctly, by net ionic equations.

# Glossary

**balanced equation**
: chemical equation with equal numbers of atoms for each element in the reactant and product

**chemical equation**
: symbolic representation of a chemical reaction

**coefficient**
: number placed in front of symbols or formulas in a chemical equation to indicate their relative amount

**complete ionic equation**
: chemical equation in which all dissolved ionic reactants and products, including spectator ions, are explicitly represented by formulas for their dissociated ions

**molecular equation**
: chemical equation in which all reactants and products are represented as neutral substances

**net ionic equation**
: chemical equation in which only those dissolved ionic reactants and products that undergo a chemical or physical change are represented (excludes spectator ions)

**product**
: substance formed by a chemical or physical change; shown on the right side of the arrow in a chemical equation

**reactant**
substance undergoing a chemical or physical change; shown on the left side of the arrow in a chemical equation

**spectator ion**
ion that does not undergo a chemical or physical change during a reaction, but its presence is required to maintain charge neutrality

## Answers

**Exercise** 10.13.1

$$2\,NH_4NO_3 \rightarrow 2\,N_2 + O_2 + 4\,H_2O$$

**Exercise** 10.13.2

Balanced molecular equation:

$$2\,NaCl\,(aq) + 2\,H_2O\,(l) \longrightarrow 2\,NaOH\,(aq) + H_2\,(g) + Cl_2\,(g)$$

Balanced ionic equation:

$$2\,Na^+\,(aq) + 2\,Cl^-\,(aq) + 2\,H_2O\,(l) \longrightarrow 2\,Na^+\,(aq) + 2\,OH^-\,(aq) + H_2\,(g) + Cl_2\,(g)$$

Balanced net ionic equation:

$$2\,Cl^-\,(aq) + 2\,H_2O\,(l) \longrightarrow 2\,OH^-\,(aq) + H_2\,(g) + Cl_2\,(g)$$

This page titled 10.13: Writing and Balancing Chemical Equations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.14: Prelude to Kinetics

The lizard in the photograph is not simply enjoying the sunshine or working on its tan. The heat from the sun's rays is critical to the lizard's survival. A warm lizard can move faster than a cold one because the chemical reactions that allow its muscles to move occur more rapidly at higher temperatures. In the absence of warmth, the lizard is an easy meal for predators.



Figure 10.14.1: *An agama lizard basks in the sun. As its body warms, the chemical reactions of its metabolism speed up.*

From baking a cake to determining the useful lifespan of a bridge, rates of chemical reactions play important roles in our understanding of processes that involve chemical changes. When planning to run a chemical reaction, we should ask at least two questions. The first is: "Will the reaction produce the desired products in useful quantities?" The second question is: "How rapidly will the reaction occur?" A reaction that takes 50 years to produce a product is about as useful as one that never gives a product at all. A third question is often asked when investigating reactions in greater detail: "What specific molecular-level processes take place as the reaction occurs?" Knowing the answer to this question is of practical importance when the yield or rate of a reaction needs to be controlled.

The study of chemical kinetics concerns the second and third questions—that is, the rate at which a reaction yields products and the molecular-scale means by which a reaction occurs. In this chapter, we will examine the factors that influence the rates of chemical reactions, the mechanisms by which reactions proceed, and the quantitative techniques used to determine and describe the rate at which reactions occur.

- Paul Flowers (University of North Carolina - Pembroke), Klaus Theopold (University of Delaware) and Richard Langley (Stephen F. Austin State University) with contributing authors. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/85abf193-2bd...a7ac8df6@9.110).

This page titled 10.14: Prelude to Kinetics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.15: Chemical Reaction Rates

## Learning Objectives

- Define chemical reaction rate
- Derive rate expressions from the balanced equation for a given chemical reaction
- Calculate reaction rates from experimental data

A *rate* is a measure of how some property varies with time. Speed is a familiar rate that expresses the distance traveled by an object in a given amount of time. Wage is a rate that represents the amount of money earned by a person working for a given amount of time. Likewise, the rate of a chemical reaction is a measure of how much reactant is consumed, or how much product is produced, by the reaction in a given amount of time.

The rate of reaction is the change in the amount of a reactant or product per unit time. Reaction rates are therefore determined by measuring the time dependence of some property that can be related to reactant or product amounts. Rates of reactions that consume or produce gaseous substances, for example, are conveniently determined by measuring changes in volume or pressure. For reactions involving one or more colored substances, rates may be monitored via measurements of light absorption. For reactions involving aqueous electrolytes, rates may be measured via changes in a solution's conductivity.

For reactants and products in solution, their relative amounts (concentrations) are conveniently used for purposes of expressing reaction rates. If we measure the concentration of hydrogen peroxide, $H_2O_2$, in an aqueous solution, we find that it changes slowly over time as the $H_2O_2$ decomposes, according to the equation:

$$2\,H_2O_2(aq) \longrightarrow 2\,H_2O(l)+O_2(g)$$

The rate at which the hydrogen peroxide decomposes can be expressed in terms of the rate of change of its concentration, as shown here:

$$\begin{aligned}\text{rate of decomposition of } H_2O_2 &= -\frac{\text{change in concentration of reactant}}{\text{time interval}}\\ &= -\frac{[H_2O_{2t_2}]-[H_2O_{2t_1}]}{t_2-t_1}\\ &= -\frac{\Delta[H_2O_2]}{\Delta t}\end{aligned}$$

This mathematical representation of the change in species concentration over time is the rate expression for the reaction. The brackets indicate molar concentrations, and the symbol delta (Δ) indicates "change in." Thus, $[H_2O_{2t_1}]$ represents the molar concentration of hydrogen peroxide at some time $t_1$; likewise, $[H_2O_{2t_2}]$ represents the molar concentration of hydrogen peroxide at a later time $t_2$; and $\Delta[H_2O_2]$ represents the change in molar concentration of hydrogen peroxide during the time interval $\Delta t$ (that is, $t_2 - t_1$). Since the reactant concentration decreases as the reaction proceeds, $\Delta[H_2O_2]$ is a negative quantity; we place a negative sign in front of the expression because reaction rates are, by convention, positive quantities. Figure 10.15.1 provides an example of data collected during the decomposition of $H_2O_2$.

| Time (h) | $[H_2O_2]$ (mol $L^{-1}$) | $\Delta[H_2O_2]$ (mol $L^{-1}$) | $\Delta t$ (h) | Rate of Decomposition, (mol/L/h) |
|---|---|---|---|---|
| 0.00 | 1.000 | | | |
| | | −0.500 | 6.00 | −0.0833 |
| 6.00 | 0.500 | | | |
| | | −0.250 | 6.00 | −0.0417 |
| 12.00 | 0.250 | | | |
| | | −0.125 | 6.00 | −0.0208 |
| 18.00 | 0.125 | | | |
| | | −0.062 | 6.00 | −0.0103 |
| 24.00 | 0.0625 | | | |

Figure 10.15.1: The rate of decomposition of $H_2O_2$ in an aqueous solution decreases as the concentration of $H_2O_2$ decreases.

To obtain the tabulated results for this decomposition, the concentration of hydrogen peroxide was measured every 6 hours over the course of a day at a constant temperature of 40 °C. Reaction rates were computed for each time interval by dividing the change in concentration by the corresponding time increment, as shown here for the first 6-hour period:

$$\frac{-\Delta[H_2O_2]}{\Delta t} = \frac{-(0.500\ \text{mol/L} - 1.000\ \text{mol/L})}{(6.00\ \text{h} - 0.00\ \text{h})} = 0.0833\ \text{mol}\ \text{L}^{-1}\ \text{h}^{-1}$$

Notice that the reaction rates vary with time, decreasing as the reaction proceeds. Results for the last 6-hour period yield a reaction rate of:

$$\frac{-\Delta[H_2O_2]}{\Delta t} = \frac{-(0.0625\ \text{mol/L} - 0.125\ \text{mol/L})}{(24.00\ \text{h} - 18.00\ \text{h})} = 0.0104\ \text{mol}\ \text{L}^{-1}\ \text{h}^{-1}$$

This behavior indicates the reaction continually slows with time. Using the concentrations at the beginning and end of a time period over which the reaction rate is changing results in the calculation of an average rate for the reaction over this time interval. At any specific time, the rate at which a reaction is proceeding is known as its instantaneous rate. The instantaneous rate of a reaction at "time zero," when the reaction commences, is its initial rate. Consider the analogy of a car slowing down as it approaches a stop sign. The vehicle's initial rate—analogous to the beginning of a chemical reaction—would be the speedometer reading at the moment the driver begins pressing the brakes ($t_0$). A few moments later, the instantaneous rate at a specific moment—call it $t_1$—would be somewhat slower, as indicated by the speedometer reading at that point in time. As time passes, the instantaneous rate will continue to fall until it reaches zero, when the car (or reaction) stops. Unlike instantaneous speed, the car's average speed is not indicated by the speedometer; but it can be calculated as the ratio of the distance traveled to the time required to bring the vehicle to a complete stop ($\Delta t$). Like the decelerating car, the average rate of a chemical reaction will fall somewhere between its initial and final rates.

The instantaneous rate of a reaction may be determined one of two ways. If experimental conditions permit the measurement of concentration changes over very short time intervals, then average rates computed as described earlier provide reasonably good approximations of instantaneous rates. Alternatively, a graphical procedure may be used that, in effect, yields the results that would be obtained if short time interval measurements were possible. If we plot the concentration of hydrogen peroxide against time, the instantaneous rate of decomposition of $H_2O_2$ at any time $t$ is given by the slope of a straight line that is tangent to the curve at that time (Figure 10.15.2). We can use calculus to evaluating the slopes of such tangent lines, but the procedure for doing so is beyond the scope of this chapter.

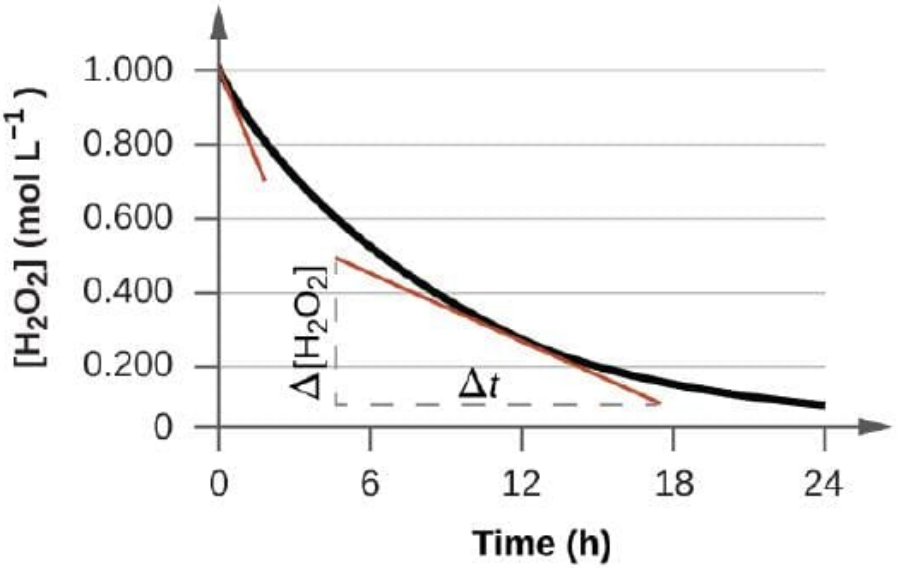



Figure 10.15.1: This graph shows a plot of concentration versus time for a 1.000 M solution of $H_2O_2$. The rate at any instant is equal to the opposite of the slope of a line tangential to this curve at that time. Tangents are shown at t = 0 h ("initial rate") and at t = 10 h ("instantaneous rate" at that particular time).

## Reaction Rates in Analysis: Test Strips for Urinalysis

Physicians often use disposable test strips to measure the amounts of various substances in a patient's urine (Figure 10.15.2 ). These test strips contain various chemical reagents, embedded in small pads at various locations along the strip, which undergo changes in color upon exposure to sufficient concentrations of specific substances. The usage instructions for test strips often stress that proper read time is critical for optimal results. This emphasis on read time suggests that kinetic aspects of the chemical reactions occurring on the test strip are important considerations.

The test for urinary glucose relies on a two-step process represented by the chemical equations shown here:

$$C_6H_{12}O_6 + O_2 \xrightarrow[\text{catalyst}]{} C_6H_{10}O_6 + H_2O_2 \tag{10.15.1}$$

$$2\,H_2O_2 + 2\,I^- \xrightarrow[\text{catalyst}]{} I_2 + 2\,H_2O + O_2 \tag{10.15.2}$$

Equation 10.15.1 depicts the oxidation of glucose in the urine to yield glucolactone and hydrogen peroxide. The hydrogen peroxide produced subsequently oxidizes colorless iodide ion to yield brown iodine (Equation 10.15.2), which may be visually detected. Some strips include an additional substance that reacts with iodine to produce a more distinct color change.

The two test reactions shown above are inherently very slow, but their rates are increased by special enzymes embedded in the test strip pad. This is an example of *catalysis*, a topic discussed later in this chapter. A typical glucose test strip for use with urine requires approximately 30 seconds for completion of the color-forming reactions. Reading the result too soon might lead one to conclude that the glucose concentration of the urine sample is lower than it actually is (a *false-negative* result). Waiting too long to assess the color change can lead to a *false positive* due to the slower (not catalyzed) oxidation of iodide ion by other substances found in urine.

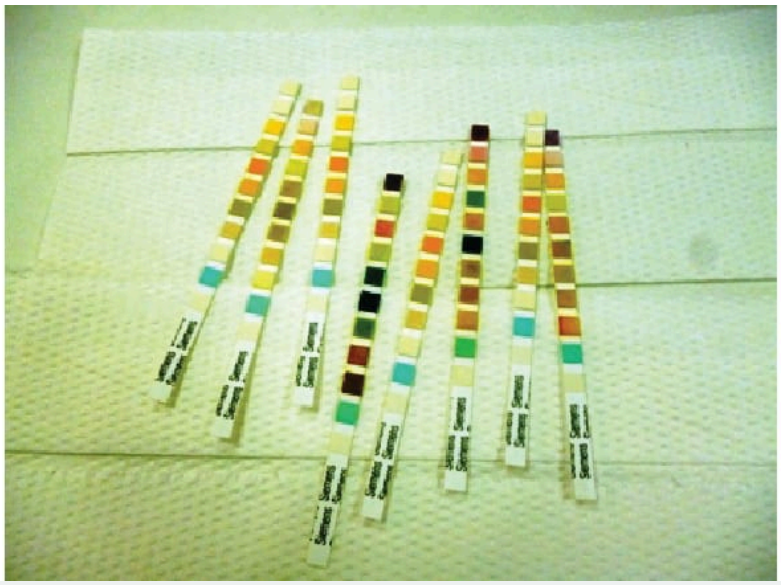


Figure 10.15.2: Test strips are commonly used to detect the presence of specific substances in a person's urine. Many test strips have several pads containing various reagents to permit the detection of multiple substances on a single strip. (credit: Iqbal Osman).

## Relative Rates of Reaction

The rate of a reaction may be expressed in terms of the change in the amount of any reactant or product, and may be simply derived from the stoichiometry of the reaction. Consider the reaction represented by the following equation:

$$2\,\mathrm{NH_3}(g) \longrightarrow \mathrm{N_2}(g) + 3\,\mathrm{H_2}(g)$$

The stoichiometric factors derived from this equation may be used to relate reaction rates in the same manner that they are used to related reactant and product amounts. The relation between the reaction rates expressed in terms of nitrogen production and ammonia consumption, for example, is:

$$-\frac{\Delta\text{mol NH}_3}{\Delta t} \times \frac{1\text{ mol N}_2}{2\text{ mol NH}_3} = \frac{\Delta\text{mol N}_2}{\Delta t}$$

We can express this more simply without showing the stoichiometric factor's units:

$$-\frac{1}{2}\frac{\Delta\text{mol NH}_3}{\Delta t} = \frac{\Delta\text{mol N}_2}{\Delta t}$$

Note that a negative sign has been added to account for the opposite signs of the two amount changes (the reactant amount is decreasing while the product amount is increasing). If the reactants and products are present in the same solution, the molar amounts may be replaced by concentrations:

$$-\frac{1}{2}\frac{\Delta[\mathrm{NH_3}]}{\Delta t} = \frac{\Delta[\mathrm{N_2}]}{\Delta t}$$

Similarly, the rate of formation of $H_2$ is three times the rate of formation of $N_2$ because three moles of $H_2$ form during the time required for the formation of one mole of $N_2$:

$$\frac{1}{3}\frac{\Delta[\mathrm{H_2}]}{\Delta t} = \frac{\Delta[\mathrm{N_2}]}{\Delta t}$$

Figure 10.15.3 illustrates the change in concentrations over time for the decomposition of ammonia into nitrogen and hydrogen at 1100 °C. We can see from the slopes of the tangents drawn at $t$ = 500 seconds that the instantaneous rates of change in the concentrations of the reactants and products are related by their stoichiometric factors. The rate of hydrogen production, for example, is observed to be three times greater than that for nitrogen production:

$$\frac{2.91 \times 10^{-6}\ M/\text{s}}{9.71 \times 10^{-6}\ M/\text{s}} \approx 3$$

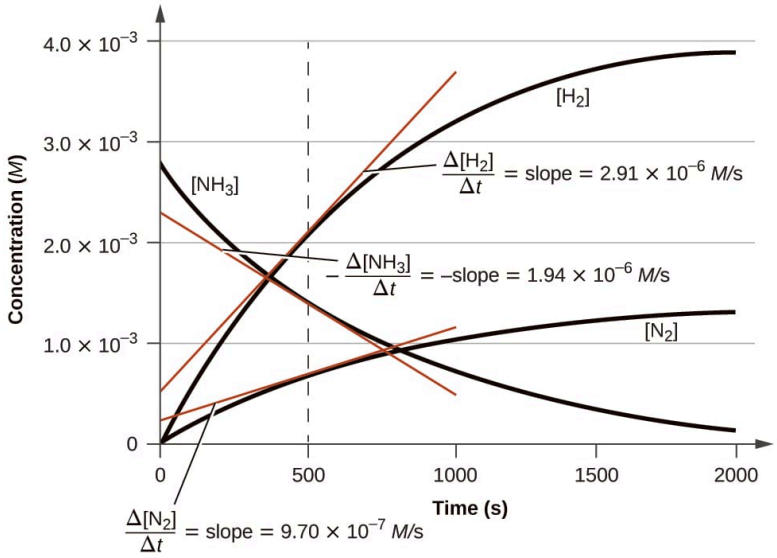



Figure 10.15.3: This graph shows the changes in concentrations of the reactants and products during the reaction $2\,\text{NH}_3 \longrightarrow \text{N}_2 + 3\,\text{H}_2$. *The rates of change of the three concentrations are related by their stoichiometric factors, as shown by the different slopes of the tangents at t = 500 s.*

## Example 10.15.1: Expressions for Relative Reaction Rates

The first step in the production of nitric acid is the combustion of ammonia:

$$4\,\text{NH}_3(g) + 5\,\text{O}_2(g) \longrightarrow 4\,\text{NO}(g) + 6\,\text{H}_2\text{O}(g)$$

Write the equations that relate the rates of consumption of the reactants and the rates of formation of the products.

**Solution**

Considering the stoichiometry of this homogeneous reaction, the rates for the consumption of reactants and formation of products are:

$$-\frac{1}{4}\frac{\Delta[\text{NH}_3]}{\Delta t} = -\frac{1}{5}\frac{\Delta[\text{O}_2]}{\Delta t} = \frac{1}{4}\frac{\Delta[\text{NO}]}{\Delta t} = \frac{1}{6}\frac{\Delta[\text{H}_2\text{O}]}{\Delta t}$$

## Exercise 10.15.1

The rate of formation of $Br_2$ is 6.0 × $10^{-6}$ mol/L/s in a reaction described by the following net ionic equation:

$$5\,\text{Br}^- + \text{BrO}_3{}^- + 6\,\text{H}^+ \longrightarrow 3\,\text{Br}_2 + 3\,\text{H}_2\text{O}$$

Write the equations that relate the rates of consumption of the reactants and the rates of formation of the products.

*Answer available at the end of this section.*

## Example 10.15.2: Reaction Rate Expressions for Decomposition of $H_2O_2$

The graph in Figure 10.15.3 shows the rate of the decomposition of $H_2O_2$ over time:

$$2\,H_2O_2 \longrightarrow 2\,H_2O + O_2$$

Based on these data, the instantaneous rate of decomposition of $H_2O_2$ at $t$ = 11.1 h is determined to be $3.20 \times 10^{-2}$ mol/L/h, that is:

$$-\frac{\Delta[H_2O_2]}{\Delta t} = 3.20 \times 10^{-2}\ \text{mol}\ \text{L}^{-1}\ \text{h}^{-1}$$

What is the instantaneous rate of production of $H_2O$ and $O_2$?

**Solution**

Using the stoichiometry of the reaction, we may determine that:

$$-\frac{1}{2}\frac{\Delta[H_2O_2]}{\Delta t} = \frac{1}{2}\frac{\Delta[H_2O]}{\Delta t} = \frac{\Delta[O_2]}{\Delta t}$$

Therefore:

$$\frac{1}{2} \times 3.20 \times 10^{-2}\ \text{mol}\ \text{L}^{-1}\ \text{h}^{-1} = \frac{\Delta[O_2]}{\Delta t}$$

and

$$\frac{\Delta[O_2]}{\Delta t} = 1.60 \times 10^{-2}\ \text{mol}\ \text{L}^{-1}\ \text{h}^{-1}$$

**Exercise 10.15.2**

If the rate of decomposition of ammonia, $NH_3$, at 1150 K is $2.10 \times 10^{-6}$ mol/L/s, what is the rate of production of nitrogen and hydrogen?

*Answer available at the end of this section.*

## Summary

The rate of a reaction can be expressed either in terms of the decrease in the amount of a reactant or the increase in the amount of a product per unit time. Relations between different rate expressions for a given reaction are derived directly from the stoichiometric coefficients of the equation representing the reaction.

## Glossary

**average rate**
: rate of a chemical reaction computed as the ratio of a measured change in amount or concentration of substance to the time interval over which the change occurred

**initial rate**
: instantaneous rate of a chemical reaction at $t$ = 0 s (immediately after the reaction has begun)

**instantaneous rate**
: rate of a chemical reaction at any instant in time, determined by the slope of the line tangential to a graph of concentration as a function of time

**rate of reaction**
: measure of the speed at which a chemical reaction takes place

**rate expression**
: mathematical representation relating reaction rate to changes in amount, concentration, or pressure of reactant or

product species per unit time

## Answers

**Exercise** 10.15.1

$$-\frac{1}{5}\frac{\Delta[\mathrm{Br}^-]}{\Delta t} = -\frac{\Delta[\mathrm{BrO_3}^-]}{\Delta t} = -\frac{1}{6}\frac{\Delta[\mathrm{H}^+]}{\Delta t} = \frac{1}{3}\frac{\Delta[\mathrm{Br_2}]}{\Delta t} = \frac{1}{3}\frac{\Delta[\mathrm{H_2O}]}{\Delta t}$$

**Exercise** 10.15.2

$1.05 \times 10^{-6}$ mol/L/s, $N_2$ and $3.15 \times 10^{-6}$ mol/L/s, $H_2$.

This page titled 10.15: Chemical Reaction Rates is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.16: Brønsted-Lowry Acids and Bases

## Learning Objectives

- Identify acids, bases, and conjugate acid-base pairs according to the Brønsted-Lowry definition
- Write equations for acid and base ionization reactions
- Use the ion-product constant for water to calculate hydronium and hydroxide ion concentrations
- Describe the acid-base behavior of amphiprotic substances

Acids and bases have been known for a long time. When Robert Boyle characterized them in 1680, he noted that acids dissolve many substances, change the color of certain natural dyes (for example, they change litmus from blue to red), and lose these characteristic properties after coming into contact with alkalis (bases). In the eighteenth century, it was recognized that acids have a sour taste, react with limestone to liberate a gaseous substance (now known to be $CO_2$), and interact with alkalis to form neutral substances. In 1815, Humphry Davy contributed greatly to the development of the modern acid-base concept by demonstrating that hydrogen is the essential constituent of acids. Around that same time, Joseph Louis Gay-Lussac concluded that acids are substances that can neutralize bases and that these two classes of substances can be defined only in terms of each other. The significance of hydrogen was reemphasized in 1884 when Svante Arrhenius defined an acid as a compound that dissolves in water to yield hydrogen cations (now recognized to be hydronium ions) and a base as a compound that dissolves in water to yield hydroxide anions.

In an earlier chapter on chemical reactions, we defined acids and bases as Arrhenius did: We identified an acid as a compound that dissolves in water to yield hydronium ions ($H_3O^+$) and a base as a compound that dissolves in water to yield hydroxide ions ( $OH^-$). This definition is not wrong; it is simply limited.

Later, we extended the definition of an acid or a base using the more general definition proposed in 1923 by the Danish chemist Johannes Brønsted and the English chemist Thomas Lowry. Their definition centers on the proton, $H^+$. A proton is what remains when a normal hydrogen atom, $^1_1H$, loses an electron. A compound that donates a proton to another compound is called a Brønsted-Lowry acid, and a compound that accepts a proton is called a Brønsted-Lowry base. An acid-base reaction is the transfer of a proton from a proton donor (acid) to a proton acceptor (base). In a subsequent chapter of this text we will introduce the most general model of acid-base behavior introduced by the American chemist G. N. Lewis.

Acids may be compounds such as HCl or $H_2SO_4$, organic acids like acetic acid ($CH_3COOH$) or ascorbic acid (vitamin C), or $H_2O$. Anions (such as $HSO_4^-$, $H_2PO_4^-$, $HS^-$, and $HCO_3^-$) and cations (such as $H_3O^+$, $NH_4^+$, and $[Al(H_2O)_6]^{3+}$) may also act as acids. Bases fall into the same three categories. Bases may be neutral molecules (such as $H_2O$, $NH_3$, and $CH_3NH_2$), anions (such as $OH^-$, $HS^-$, $HCO_3^-$, $CO_3^{2-}$, $F^-$, and $PO_4^{3-}$), or cations (such as $[Al(H_2O)_5OH]^{2+}$). The most familiar bases are ionic compounds such as $NaOH$ and $Ca(OH)_2$, which contain the hydroxide ion, $OH^-$. The hydroxide ion in these compounds accepts a proton from acids to form water:

$$H^+ + OH^- \rightarrow H_2O \tag{10.16.1}$$

We call the product that remains after an acid donates a proton the conjugate base of the acid. This species is a base because it can accept a proton (to re-form the acid):

$$\text{acid} \rightleftharpoons \text{proton} + \text{conjugate base} \tag{10.16.2}$$

$$HF \rightleftharpoons H^+ + F^- \tag{10.16.3}$$

$$H_2SO_4 \rightleftharpoons H^+ + HSO_4^- \tag{10.16.4}$$

$$H_2O \rightleftharpoons H^+ + OH^- \tag{10.16.5}$$

$$HSO_4^- \rightleftharpoons H^+ + SO_4^{2-} \tag{10.16.6}$$

$$NH_4^+ \rightleftharpoons H^+ + NH_3 \tag{10.16.7}$$

We call the product that results when a base accepts a proton the base's conjugate acid. This species is an acid because it can

give up a proton (and thus re-form the base):

$$\text{base} + \text{proton} \rightleftharpoons \text{conjugate acid} \tag{10.16.8}$$

$$\mathrm{OH^- + H^+ \rightleftharpoons H_2O} \tag{10.16.9}$$

$$\mathrm{H_2O + H^+ \rightleftharpoons H_3O^+} \tag{10.16.10}$$

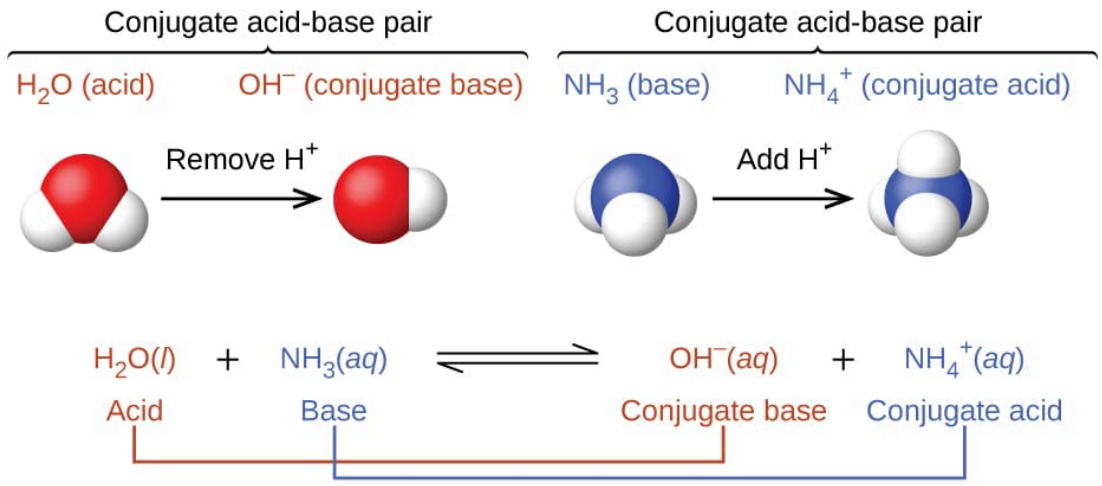
$$\mathrm{NH_3 + H^+ \rightleftharpoons NH_4^+} \tag{10.16.11}$$

$$\mathrm{S^{2-} + H^+ \rightleftharpoons HS^-} \tag{10.16.12}$$

$$\mathrm{CO_3^{2-} + H^+ \rightleftharpoons HCO_3^-} \tag{10.16.13}$$

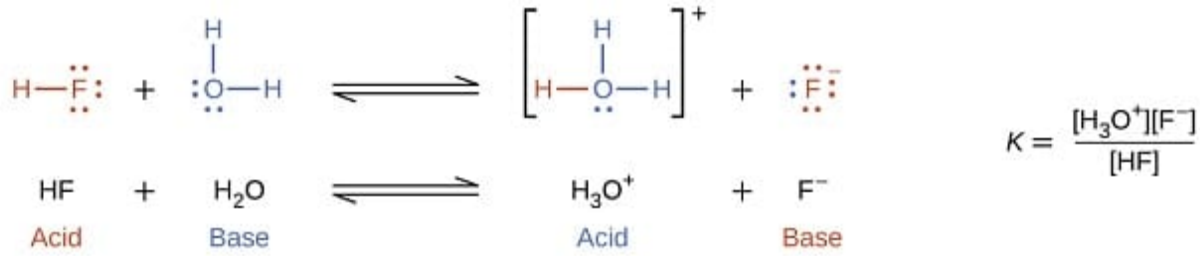
$$\mathrm{F^- + H^+ \rightleftharpoons HF} \tag{10.16.14}$$

In these two sets of equations, the behaviors of acids as proton donors and bases as proton acceptors are represented in isolation. In reality, all acid-base reactions involve the transfer of protons between acids and bases. For example, consider the acid-base reaction that takes place when ammonia is dissolved in water. A water molecule (functioning as an acid) transfers a proton to an ammonia molecule (functioning as a base), yielding the conjugate base of water, $\mathrm{OH^-}$, and the conjugate acid of ammonia, $\mathrm{NH_4^+}$:




This figure has three parts in two rows. In the first row, two diagrams of acid-base pairs are shown. On the left, a space filling model of H subscript 2 O is shown with a red O atom at the center and two smaller white H atoms attached in a bent shape. Above this model is the label “H subscript 2 O (acid)” in purple. An arrow points right, which is labeled “Remove H superscript plus.” To the right is another space filling model with a single red O atom to which a single smaller white H atom is attached. The label in purple above this model reads, “O H superscript negative (conjugate base).” Above both of these red and white models is an upward pointing bracket that is labeled “Conjugate acid-base pair.” To the right is a space filling model with a central blue N atom to which three smaller white H atoms are attached in a triangular pyramid arrangement. A label in green above reads “N H subscript 3 (base).” An arrow labeled “Add H superscript plus” points right. To the right of the arrow is another space filling model with a blue central N atom and four smaller white H atoms in a tetrahedral arrangement. The green label above reads “N H subscript 3 superscript plus (conjugate acid).” Above both of these blue and white models is an upward pointing bracket that is labeled “Conjugate acid-base pair.” The second row of the figure shows the chemical reaction, H subscript 2 O ( l ) is shown in purple, and is labeled below in purple as “acid,” plus N H subscript 3 (a q) in green, labeled below in green as “base,” followed by a double sided arrow arrow and O H superscript negative (a q) in purple, labeled in purple as “conjugate base,” plus N H subscript 4 superscript plus (a q)” in green, which is labeled in green as “conjugate acid.” The acid on the left side of the equation is connected to the conjugate base on the right with a purple line. Similarly, the base on the left is connected to the conjugate acid on the right side.

The reaction between a Brønsted-Lowry acid and water is called acid ionization. For example, when hydrogen fluoride dissolves in water and ionizes, protons are transferred from hydrogen fluoride molecules to water molecules, yielding hydronium ions and fluoride ions:

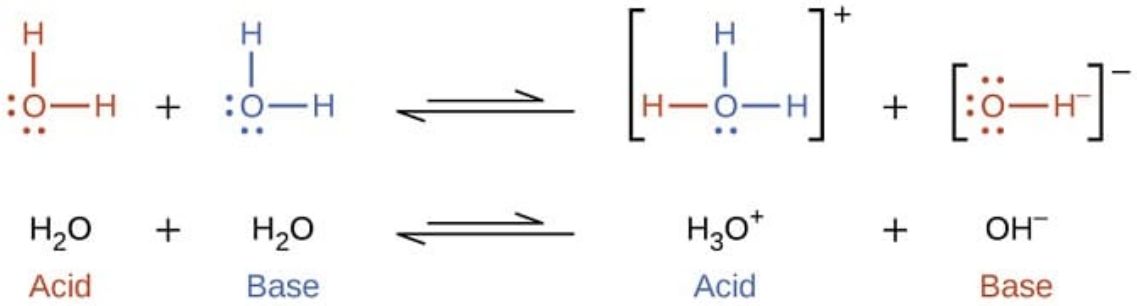



This figure has two rows. In both rows, a chemical reaction is shown. In the first, structural formulas are provided. In this model, in purple, an H atom is connected to an F atom with a single bond. The F atom has pairs of electron dots at the top, right, and bottom. This is followed by a plus sign, which is followed in green by an O atom which has H atoms singly bonded above and to the right. The O atom has pairs of electron dots on its left and lower sides. A double arrow follows. To the right, in brackets is a structure with a central O atom in green, with green H atoms singly bonded above and to the right. A pair of green electron dots is on the lower side of the O atom. To the left of the green O atom, a purple H atom is singly bonded. This is followed by a plus sign and an F atom in purple with pairs of electron dots above, right, below, and to the left. This atom also has a superscript negative sign. The reaction is written in symbolic form below. H F is labeled in purple below as “Acid subscript 1.” This is followed by plus H subscript 2 O, which is labeled in green below as “Base subscript 2.” A double sided arrow follows. To the right is H subscript 3 O superscript plus, which is labeled in green as below in as “Acid subscript 2.” This is followed by plus and F surrounded by 4 pairs of dots and superscript negative. The label below in purple reads, “Base subscript 1.” To the right of the reactions is the formula, K subscript a equals left bracket H subscript 3 O superscript plus right bracket left bracket F superscript negative right bracket all over left bracket H F right bracket.

When we add a base to water, a base ionization reaction occurs in which protons are transferred from water molecules to base molecules. For example, adding pyridine to water yields hydroxide ions and pyridinium ions:

<div data-mt-source="1"

Notice that both these ionization reactions are represented as equilibrium processes. The relative extent to which these acid and base ionization reactions proceed is an important topic treated in a later section of this chapter. In the preceding paragraphs we saw that water can function as either an acid or a base, depending on the nature of the solute dissolved in it. In fact, in pure water or in any aqueous solution, water acts both as an acid and a base. A very small fraction of water molecules donate protons to other water molecules to form hydronium ions and hydroxide ions:




This figure has two rows. In both rows, a chemical reaction is shown. In the first, structural formulas are provided. In this model, in purple, O atom which has H atoms singly bonded above and to the right. The O atom has pairs of electron dots on its left and lower sides. This is followed by a plus sign, which is followed in green by an O atom which has H atoms singly bonded above and to the right. The O atom has pairs of electron dots on its left and lower sides. A double arrow follows. To the right, in brackets is a structure with a central O atom in green, with green H atoms singly bonded above and to the right. A pair of green electron dots is on the lower side of the O atom. To the left of the green O atom, a purple H atom is singly bonded. Outside the brackets to the right is a superscript plus. This is followed by a plus sign and an O atom in purple with pairs of electron dots above, left, and below. An H atom is singly bonded to the right. This atom has a superscript negative sign. The reaction is written in symbolic form below. H subscript 2 O is labeled in purple below as “Acid subscript 1.” This is followed by plus H subscript 2 O, which is labeled in green below as “Base subscript 2.” A double sided arrow follows. To the right is H subscript 3 O superscript plus, which is labeled in green as below in as “Acid subscript 2.” This is followed by plus and O with pairs of dots above, below, and to the left with a singly bonded H on the right with a superscript negative. The label below in purple reads, “ Base subscript 1.”

This type of reaction, in which a substance ionizes when one molecule of the substance reacts with another molecule of the same

substance, is referred to as **autoionization**. Pure water undergoes autoionization to a very slight extent. Only about two out of every $10^9$ molecules in a sample of pure water are ionized at 25 °C. The equilibrium constant for the ionization of water is called the ion-product constant for water ($K_w$):

$$\mathrm{H_2O_{(\mathit{l})} + H_2O_{(\mathit{l})} \rightleftharpoons H_3O^+{}_{(\mathit{aq})} + OH^-{}_{(\mathit{aq})}} \quad K_w = [H_3O^+][OH^-] \tag{10.16.15}$$

The slight ionization of pure water is reflected in the small value of the equilibrium constant; at 25 °C, Kw has a value of $1.0 \times 10^{-14}$. The process is endothermic, and so the extent of ionization and the resulting concentrations of hydronium ion and hydroxide ion increase with temperature. For example, at 100 °C, the value for $K_w$ is approximately $5.1 \times 10^{-13}$, roughly 50-times larger than the value at 25 °C.

## Example 10.16.1: Ion Concentrations in Pure Water

What are the hydronium ion concentration and the hydroxide ion concentration in pure water at 25 °C?

**Solution**

The autoionization of water yields the same number of hydronium and hydroxide ions. Therefore, in pure water, $[H_3O^+] = [OH^-]$. At 25 °C:

$$K_w = [H_3O^+][OH^-] = [H_3O^+]^2 = [OH^-]^2 = 1.0 \times 10^{-14}$$

So:

$$[H_3O^+] = [OH^-] = \sqrt{1.0 \times 10^{-14}} = 1.0 \times 10^{-7}\ M$$

The hydronium ion concentration and the hydroxide ion concentration are the same, and we find that both equal $1.0 \times 10^{-7}\ M$.

## Exercise 10.16.1

The ion product of water at 80 °C is $2.4 \times 10^{-13}$. What are the concentrations of hydronium and hydroxide ions in pure water at 80 °C?

*Answer available at the end of this section.*

It is important to realize that the autoionization equilibrium for water is established in all aqueous solutions. Adding an acid or base to water will not change the position of the equilibrium. Example 14.12 demonstrates the quantitative aspects of this relation between hydronium and hydroxide ion concentrations.

## Example 10.16.2: The Inverse Proportionality of $[H_3O^+]$ and $[OH^-]$

The Inverse Proportionality of $[H_3O^+]$ and $[OH^-]$ A solution of carbon dioxide in water has a hydronium ion concentration of $2.0 \times 10^{-6}\ M$. What is the concentration of hydroxide ion at 25 °C?

**Solution**

We know the value of the ion-product constant for water at 25 °C:

$$\mathrm{2\,H_2O_{(\mathit{l})} \rightleftharpoons H_3O^+{}_{(\mathit{aq})} + OH^-{}_{(\mathit{aq})}}$$

$$K_w = [H_3O^+][OH^-] = 1.0 \times 10^{-14}$$

Thus, we can calculate the missing equilibrium concentration.

Rearrangement of the $K_w$ expression yields that $[OH^-]$ is directly proportional to the inverse of $[H_3O^+]$:

$$[OH^-] = \frac{K_w}{[H_3O^+]} = \frac{1.0 \times 10^{-14}}{2.0 \times 10^{-6}} = 5.0 \times 10^{-9}$$

The hydroxide ion concentration in water is reduced to $5.0 \times 10^{-9}\ M$ as the hydrogen ion concentration increases to $2.0 \times 10^{-6}\ M$. This is expected from Le Chatelier's principle; the autoionization reaction shifts to the left to reduce the stress of the increased hydronium ion concentration and the $[OH^-]$ is reduced relative to that in pure water.

A check of these concentrations confirms that our arithmetic is correct:

$$\begin{aligned} K_w &= [H_3O^+][OH^-] \\ &= (2.0 \times 10^{-6})(5.0 \times 10^{-9}) \\ &= 1.0 \times 10^{-14} \end{aligned}$$

### Exercise 10.16.2

What is the hydronium ion concentration in an aqueous solution with a hydroxide ion concentration of 0.001 M at 25 °C?

*Answer available at the end of this section.*

## Amphiprotic Species

Like water, many molecules and ions may either gain or lose a proton under the appropriate conditions. Such species are said to be amphiprotic. Another term used to describe such species is amphoteric, which is a more general term for a species that may act either as an acid or a base by any definition (not just the Brønsted-Lowry one). Consider for example the bicarbonate ion, which may either donate or accept a proton as shown here:

$$HCO^-_{3(aq)} + H_2O_{(l)} \rightleftharpoons [redacted]_{3(aq)} + H_3O^+_{(aq)} \quad (10.16.16)$$

$$HCO^-_{3(aq)} + H_2O_{(l)} \rightleftharpoons H_2CO_{3(aq)} + OH^-_{(aq)} \quad (10.16.17)$$

### Example 10.16.3: The Acid-Base Behavior of an Amphoteric Substance

Write separate equations representing the reaction of $HSO_3^-$

a. as an acid with $OH^-$
b. as a base with HI

**Solution**

a. $HSO_3^-(aq) + OH^-(aq) \rightleftharpoons SO_3^{2-}(aq) + H_2O_{(l)}$
b. $HSO_3^-(aq) + HI(aq) \rightleftharpoons H_2SO_3(aq) + I^-(aq)$

### Exercise 10.16.3

Write separate equations representing the reaction of $H_2PO_4^-$

a. as a base with HBr
b. as an acid with $OH^-$

*Answer a available at the end of this section.*

## Summary

A compound that can donate a proton (a hydrogen ion) to another compound is called a Brønsted-Lowry acid. The compound that accepts the proton is called a Brønsted-Lowry base. The species remaining after a Brønsted-Lowry acid has lost a proton is the conjugate base of the acid. The species formed when a Brønsted-Lowry base gains a proton is the conjugate acid of the base. Thus, an acid-base reaction occurs when a proton is transferred from an acid to a base, with formation of the conjugate base of the reactant acid and formation of the conjugate acid of the reactant base. Amphiprotic species can act as both proton donors and proton acceptors. Water is the most important amphiprotic species. It can form both the hydronium ion, $H_3O^+$, and the hydroxide ion, $OH^-$ when it undergoes autoionization:

$$2\,H_2O_{(l)} \rightleftharpoons H_3O^+(aq) + OH^-(aq)$$

The ion product of water, $K_w$ is the equilibrium constant for the autoionization reaction:

$$K_w = [H_3O^+][OH^-] = 1.0 \times 10^{-14} \text{ at } 25°C$$

## Key Equations

- $$K_w = [H_3O^+][OH^-\,] = 1.0 \times 10^{-14}\ (\text{at } 25\ °C)$$

## Glossary

**acid ionization**
: reaction involving the transfer of a proton from an acid to water, yielding hydronium ions and the conjugate base of the acid

**amphiprotic**
: species that may either gain or lose a proton in a reaction

**amphoteric**
: species that can act as either an acid or a base

**autoionization**
: reaction between identical species yielding ionic products; for water, this reaction involves transfer of protons to yield hydronium and hydroxide ions

**base ionization**
: reaction involving the transfer of a proton from water to a base, yielding hydroxide ions and the conjugate acid of the base

**Brønsted-Lowry acid**
: proton donor

**Brønsted-Lowry base**
: proton acceptor

**conjugate acid**
: substance formed when a base gains a proton

**conjugate base**
: substance formed when an acid loses a proton

**ion-product constant for water ($K_w$)**
: equilibrium constant for the autoionization of water

## Answers

**Exercise** 10.16.1

$[\mathrm{H_3O^+}] = [\mathrm{OH^-}] = 4.9 \times 10^{-7}\ M$

**Exercise** 10.16.2

$$[\mathrm{H_3O^+}] = 1 \times 10^{-11} M$$

**Exercise** 10.16.3

$\mathrm{H_2PO_4^-}(aq) + \mathrm{HBr}(aq) \rightleftharpoons \mathrm{H_3PO_4}(aq) + \mathrm{Br^-}(aq)$

$\mathrm{H_2PO_4^-}(aq) + \mathrm{OH^-}(aq) \rightleftharpoons \mathrm{HPO_4^{2-}}(aq) + \mathrm{H_2O}_{(l)}$

This page titled 10.16: Brønsted-Lowry Acids and Bases is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.17: pH and pOH

## Learning Objectives

- Explain the characterization of aqueous solutions as acidic, basic, or neutral
- Express hydronium and hydroxide ion concentrations on the pH and pOH scales
- Perform calculations relating pH and pOH

As discussed earlier, hydronium and hydroxide ions are present both in pure water and in all aqueous solutions, and their concentrations are inversely proportional as determined by the ion product of water ($K_w$). The concentrations of these ions in a solution are often critical determinants of the solution's properties and the chemical behaviors of its other solutes, and specific vocabulary has been developed to describe these concentrations in relative terms. A solution is neutral if it contains equal concentrations of hydronium and hydroxide ions; acidic if it contains a greater concentration of hydronium ions than hydroxide ions; and basic if it contains a lesser concentration of hydronium ions than hydroxide ions.

A common means of expressing quantities, the values of which may span many orders of magnitude, is to use a logarithmic scale. One such scale that is very popular for chemical concentrations and equilibrium constants is based on the p-function, defined as shown where "X" is the quantity of interest and "log" is the base-10 logarithm:

$$\mathrm{pX} = -\log \mathrm{X} \tag{10.17.1}$$

The pH of a solution is therefore defined as shown here, where $[H_3O^+]$ is the molar concentration of hydronium ion in the solution:

$$\mathrm{pH} = -\log[\mathrm{H_3O^+}] \tag{10.17.2}$$

Rearranging this equation to isolate the hydronium ion molarity yields the equivalent expression:

$$[\mathrm{H_3O^+}] = 10^{-\mathrm{pH}} \tag{10.17.3}$$

Likewise, the hydroxide ion molarity may be expressed as a p-function, or pOH:

$$\mathrm{pOH} = -\log[\mathrm{OH^-}] \tag{10.17.4}$$

or

$$[\mathrm{OH^-}] = 10^{-\mathrm{pOH}} \tag{10.17.5}$$

Finally, the relation between these two ion concentration expressed as p-functions is easily derived from the $K_w$ expression:

$$K_\mathrm{w} = [\mathrm{H_3O^+}][\mathrm{OH^-}] \tag{10.17.6}$$

$$-\log K_\mathrm{w} = -\log([\mathrm{H_3O^+}][\mathrm{OH^-}]) = -\log[\mathrm{H_3O^+}] + -\log[\mathrm{OH^-}] \tag{10.17.7}$$

$$\mathrm{p}K_\mathrm{w} = \mathrm{pH} + \mathrm{pOH} \tag{10.17.8}$$

At 25 °C, the value of $K_w$ is $1.0 \times 10^{-14}$, and so:

$$14.00 = \mathrm{pH} + \mathrm{pOH} \tag{10.17.9}$$

The hydronium ion molarity in pure water (or any neutral solution) is $1.0 \times 10^{-7}\ M$ at 25 °C. The pH and pOH of a neutral solution at this temperature are therefore:

$$\mathrm{pH} = -\log[\mathrm{H_3O^+}] = -\log(1.0 \times 10^{-7}) = 7.00 \tag{10.17.10}$$

$$\mathrm{pOH} = -\log[\mathrm{OH^-}] = -\log(1.0 \times 10^{-7}) = 7.00 \tag{10.17.11}$$

And so, at this temperature, acidic solutions are those with hydronium ion molarities greater than $1.0 \times 10^{-7}\ M$ and hydroxide

ion molarities less than $1.0 \times 10^{-7}\ M$ (corresponding to pH values less than 7.00 and pOH values greater than 7.00). Basic solutions are those with hydronium ion molarities less than $1.0 \times 10^{-7}\ M$ and hydroxide ion molarities greater than $1.0 \times 10^{-7}\ M$ (corresponding to pH values greater than 7.00 and pOH values less than 7.00).

## When $pH = 7$ Solutions are not Neutral

Since the autoionization constant $K_w$ is temperature dependent, these correlations between pH values and the acidic/neutral/basic adjectives will be different at temperatures other than 25 °C. For example, the hydronium molarity of pure water at 80 °C is 4.9 × 10$^{-7}$ *M*, which corresponds to pH and pOH values of:

$$\begin{aligned} pH &= -\log[\mathrm{H_3O^+}] \\ &= -\log(4.9 \times 10^{-7}) \\ &= 6.31 \end{aligned}$$

$$\begin{aligned} pOH &= -\log[\mathrm{OH^-}] \\ &= -\log(4.9 \times 10^{-7}) \\ &= 6.31 \end{aligned}$$

At this temperature, then, neutral solutions exhibit pH = pOH = 6.31, acidic solutions exhibit pH less than 6.31 and pOH greater than 6.31, whereas basic solutions exhibit pH greater than 6.31 and pOH less than 6.31. This distinction can be important when studying certain processes that occur at nonstandard temperatures, such as enzyme reactions in warm-blooded organisms. Unless otherwise noted, references to pH values are presumed to be those at standard temperature (25 °C) (Table 10.17.1).

*Table 10.17.1: Summary of Relations for Acidic, Basic and Neutral Solutions*

| Classification | Relative Ion Concentrations | pH at 25 °C |
|---|---|---|
| acidic | $[H_3O^+] > [OH^-]$ | pH < 7 |
| neutral | $[H_3O^+] = [OH^-]$ | pH = 7 |
| basic | $[H_3O^+] < [OH^-]$ | pH > 7 |

Figure 10.17.1 shows the relationships between $[H_3O^+]$, $[OH^-]$, pH, and pOH, and gives values for these properties at standard temperatures for some common substances.

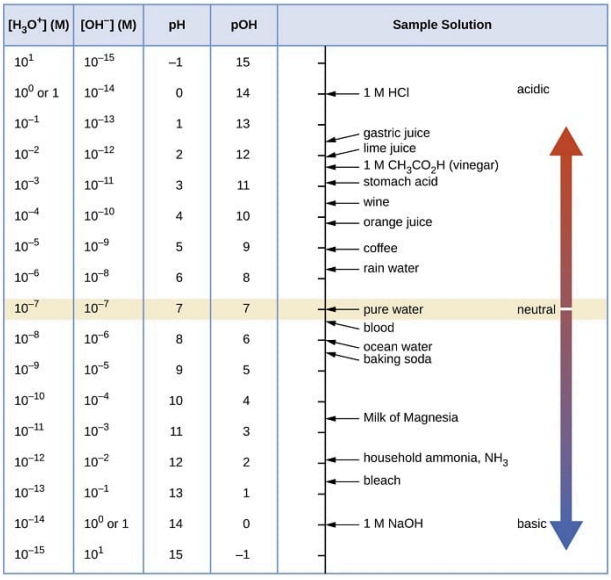



Figure 10.17.1: The pH and pOH scales represent concentrations of

[$H_3O^+$] and $OH^-$, respectively. The pH and pOH values of some common substances at standard temperature (25 °C) are shown in this chart. A table is provided with 5 columns. The first column is labeled “left bracket H subscript 3 O superscript plus right bracket (M).” Powers of ten are listed in the column beginning at 10 superscript 1, including 10 superscript 0 or 1, 10 superscript negative 1, decreasing by single powers of 10 to 10 superscript negative 15. The second column is labeled “left bracket O H superscript negative right bracket (M).” Powers of ten are listed in the column beginning at 10 superscript negative 15, increasing by single powers of 10 to including 10 superscript 0 or 1, and 10 superscript 1. The third column is labeled “p H.” Values listed in this column are integers beginning at negative 1, increasing by ones up to 14. The fourth column is labeled “p O H.” Values in this column are integers beginning at 15, decreasing by ones up to negative 1. The fifth column is labeled “Sample Solution.” A vertical line at the left of the column has tick marks corresponding to each p H level in the table. Substances are listed next to this line segment with line segments connecting them to the line to show approximate p H and p O H values. 1 M H C l is listed at a p H of 0. Gastric juices are listed at a p H of about 1.5. Lime juice is listed at a p H of about 2, followed by 1 M C H subscript 3 C O subscript 2 H, followed by stomach acid at a p H value of nearly 3. Wine is listed around 3.5. Coffee is listed just past 5. Pure water is listed at a p H of 7. Pure blood is just beyond 7. Milk of Magnesia is listed just past a p H of 10.5. Household ammonia is listed just before a pH of 12. 1 M N a O H is listed at a p H of 0. To the right of this labeled arrow is an arrow that points up and down through the height of the column. A beige strip passes through the table and to this double headed arrow at p H 7. To the left of the double headed arrow in this beige strip is the label “neutral.” A narrow beige strip runs through the arrow. Just above and below this region, the arrow is purple. It gradually turns to a bright red as it extends upward. At the top of the arrow, near the head of the arrow is the label “acidic.” Similarly, the lower region changes color from purple to blue moving to the bottom of the column. The head at this end of the arrow is labeled “basic.”

### Example 10.17.1: Calculation of pH from $[H_3O^+]$

What is the pH of stomach acid, a solution of HCl with a hydronium ion concentration of $1.2 \times 10^{-3}\ M$?

**Solution**

$$\begin{aligned} pH &= -\log[H_3O^+] \\ &= -\log(1.2 \times 10^{-3}) \\ &= -(-2.92) \\ &= 2.92 \end{aligned}$$

### Exercise 10.17.1

Water exposed to air contains carbonic acid, $H_2CO_3$, due to the reaction between carbon dioxide and water:

$$\mathrm{CO_2\,(aq) + H_2O\,(l) \rightleftharpoons H_2CO_3\,(aq)}$$

Air-saturated water has a hydronium ion concentration caused by the dissolved $CO_2$ of $2.0 \times 10^{-6}\ M$, about 20-times larger than that of pure water. Calculate the pH of the solution at 25 °C.

*Answer available at the end of this section.*

### Example 10.17.2: Calculation of Hydronium Ion Concentration from pH

Calculate the hydronium ion concentration of blood, the pH of which is 7.3 (slightly alkaline).

**Solution**

$$\mathrm{pH = -\log[H_3O^+] = 7.3}$$

$$\mathrm{\log[H_3O^+] = -7.3}$$

$$\mathrm{[H_3O^+] = 10^{-7.3}}$$

or

$$[H_3O^+] = \text{antilog of} - 7.3$$

$$[H_3O^+] = 5 \times 10^{-8}\ M$$

(On a calculator take the antilog, or the "inverse" log, of –7.3, or calculate $10^{-7.3}$.)

**Exercise 10.17.2**

Calculate the hydronium ion concentration of a solution with a pH of –1.07.

*Answer available at the end of this section.*

## Environmental Science

Normal rainwater has a pH between 5 and 6 due to the presence of dissolved $CO_2$ which forms carbonic acid:

$$H_2O\,(l) + CO_2\,(g) \longrightarrow H_2CO_3\,(aq) \tag{10.17.12}$$

$$H_2CO_3\,(aq) \rightleftharpoons H^+\,(aq) + HCO_3{}^-\,(aq) \tag{10.17.13}$$

Acid rain is rainwater that has a pH of less than 5, due to a variety of nonmetal oxides, including $CO_2$, $SO_2$, $SO_3$, NO, and $NO_2$ being dissolved in the water and reacting with it to form not only carbonic acid, but sulfuric acid and nitric acid. The formation and subsequent ionization of sulfuric acid are shown here:

$$H_2O\,(l) + SO_3\,(g) \longrightarrow H_2SO_4\,(aq) \tag{10.17.14}$$

$$H_2SO_4\,(aq) \longrightarrow H^+\,(aq) + HSO_4{}^-\,(aq) \tag{10.17.15}$$

Carbon dioxide is naturally present in the atmosphere because we and most other organisms produce it as a waste product of metabolism. Carbon dioxide is also formed when fires release carbon stored in vegetation or when we burn wood or fossil fuels. Sulfur trioxide in the atmosphere is naturally produced by volcanic activity, but it also stems from burning fossil fuels, which have traces of sulfur, and from the process of "roasting" ores of metal sulfides in metal-refining processes. Oxides of nitrogen are formed in internal combustion engines where the high temperatures make it possible for the nitrogen and oxygen in air to chemically combine.

Acid rain is a particular problem in industrial areas where the products of combustion and smelting are released into the air without being stripped of sulfur and nitrogen oxides. In North America and Europe until the 1980s, it was responsible for the destruction of forests and freshwater lakes, when the acidity of the rain actually killed trees, damaged soil, and made lakes uninhabitable for all but the most acid-tolerant species. Acid rain also corrodes statuary and building facades that are made of marble and limestone (Figure 10.17.2). Regulations limiting the amount of sulfur and nitrogen oxides that can be released into the atmosphere by industry and automobiles have reduced the severity of acid damage to both natural and manmade environments in North America and Europe. It is now a growing problem in industrial areas of China and India.



(a) (b)

Figure 10.17.2: (a) Acid rain makes trees more susceptible to drought and insect infestation, and depletes nutrients in the soil. (b) It also is corrodes statues that are carved from marble or limestone. (credit a: modification of work by Chris M Morris; credit b: modification of work by "Eden, Janine and Jim"/Flickr) Two photos are shown. Photograph a on the left shows the upper portion of trees against a bright blue sky. The tops of several trees at the center of the photograph have bare branches and appear to be dead. Image b shows a statue of a man that appears to from the revolutionary war era in either marble or limestone.

## Example 10.17.3: Calculation of pOH

What are the pOH and the pH of a 0.0125-M solution of potassium hydroxide, KOH?

**Solution**

Potassium hydroxide is a highly soluble ionic compound and completely dissociates when dissolved in dilute solution, yielding $[OH^-]$ = 0.0125 M:

$$\mathrm{pOH} = -\log[\mathrm{OH}^-] = -\log 0.0125$$

$$= -(-1.903) = 1.903$$

The pH can be found from the $\mathrm{pOH}$:

$$\mathrm{pH} + \mathrm{pOH} = 14.00$$

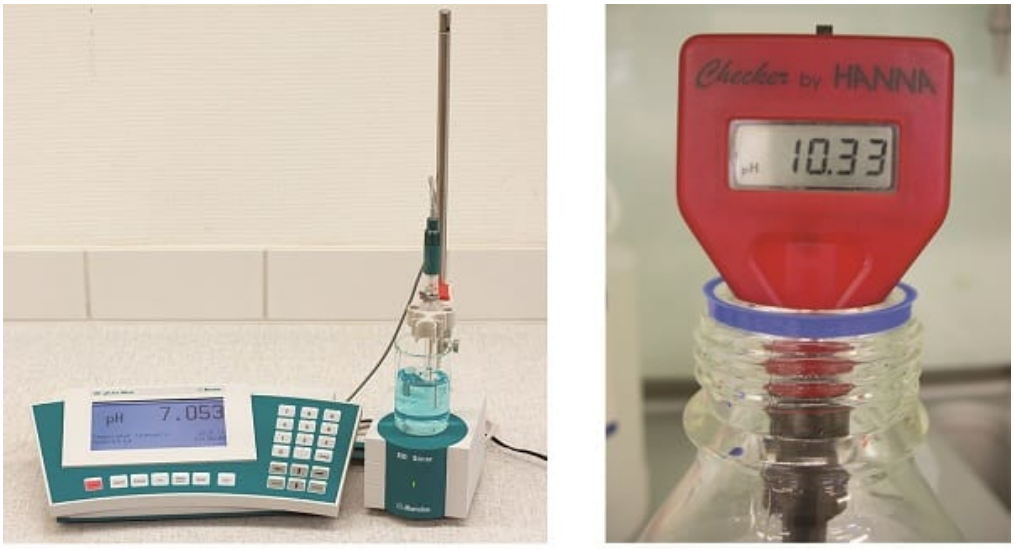
$$\mathrm{pH} = 14.00 - \mathrm{pOH} = 14.00 - 1.903 = 12.10$$

## Exercise 10.17.3

The hydronium ion concentration of vinegar is approximately $4 \times 10^{-3}\ M$. What are the corresponding values of pOH and pH?

*Answer available at the end of this section.*

The acidity of a solution is typically assessed experimentally by measurement of its pH. The pOH of a solution is not usually measured, as it is easily calculated from an experimentally determined pH value. The pH of a solution can be directly measured using a pH meter (Figure 10.17.3).




(a) (b)

Figure 10.17.3: (a) A research-grade pH meter used in a laboratory can have a resolution of 0.001 pH units, an accuracy of ± 0.002 pH units, and may cost in excess of $1000. (b) A portable pH meter has lower resolution (0.01 pH units), lower accuracy (± 0.2 pH units), and a far lower price tag. (credit b: modification of work by Jacopo Werther) This figure contains two images. The first, image a, is of an analytical digital p H meter on a laboratory counter. The second, image b, is of a portable hand held digital p H meter.

The pH of a solution may also be visually estimated using colored indicators (Figure 10.17.3).




(a) (b)

Figure 10.17.4: (a) A universal indicator assumes a different color in solutions of different pH values. Thus, it can be added to a solution to determine the pH of the solution. The eight vials each contain a universal indicator and 0.1-M solutions of progressively weaker acids: HCl (pH = l), $CH_3CO_2H$ (pH = 3), and $NH_4Cl$ (pH = 5), deionized water, a neutral substance (pH = 7); and 0.1-M solutions of the progressively stronger bases: KCl (pH = 7), aniline, $C_6H_5NH_2$ (pH = 9), $NH_3$ (pH = 11), and NaOH (pH = 13). (b) pH paper contains a mixture of indicators that give different colors in solutions of differing pH values. (credit: modification of work by Sahar Atwa). This figure contains two images. The first shows a variety of colors of solutions in labeled beakers. A red solution in a beaker is labeled “0.10 M H C l.” An orange solution is labeled “0.10 M C H subscript 3 C O O H.” A yellow-orange solution is labeled “0.1 M N H subscript 4 C l.” A yellow solution is labeled “deionized water.” A second solution beaker is labeled “0.10 M K C l.” A green solution is labeled “0.10 M aniline.” A blue solution is labeled “0.10 M N H subscript 4 C l (a q).” A final beaker containing a dark blue solution is labeled “0.10 M N a O H.” Image b shows pHydrion paper that is used for measuring pH in the range of p H from 1 to 12. The color scale for identifying p H based on color is shown along with several of the test strips used to evaluate p H.

## Summary

The concentration of hydronium ion in a solution of an acid in water is greater than $1.0 \times 10^{-7}\ M$ at 25 °C. The concentration of hydroxide ion in a solution of a base in water is greater than $1.0 \times 10^{-7}\ M$ at 25 °C. The concentration of $H_3O^+$ in a solution can be expressed as the pH of the solution; $\mathrm{pH} = -\log \mathrm{H_3O^+}$. The concentration of $OH^-$ can be expressed as the pOH of the solution: $\mathrm{pOH} = -\log[\mathrm{OH^-}]$. In pure water, pH = 7.00 and pOH = 7.00

## Key Equations

- $\mathrm{pH} = -\log[\mathrm{H_3O^+}]$
- $\mathrm{pOH} = -\log[\mathrm{OH^-}]$
- $[H_3O^+] = 10^{-pH}$
- $[OH^-] = 10^{-pOH}$
- pH + pOH = p$K_w$ = 14.00 at 25 °C

## Glossary

**acidic**
: describes a solution in which $[H_3O^+] > [OH^-]$

**basic**
: describes a solution in which $[H_3O^+] < [OH^-]$

**neutral**
: describes a solution in which $[H_3O^+] = [OH^-]$

**pH**
: logarithmic measure of the concentration of hydronium ions in a solution

**pOH**
: logarithmic measure of the concentration of hydroxide ions in a solution

### Answers

**Exercise** 10.17.1
5.70

**Exercise** 10.17.2
12 M

**Exercise** 10.17.3
pOH = 11.6,
pH = 14.00 - pOH = 2.4

This page titled 10.17: pH and pOH is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.18: ORGANIC CHEMISTRY

10.18.1: Introduction
10.18.2: Hydrocarbons
10.18.3: Alcohols and Ethers
10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters
10.18.5: Amines and Amides

This page titled 10.18: Organic Chemistry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.18.1: Introduction

All living things on earth are formed mostly of carbon compounds. The prevalence of carbon compounds in living things has led to the epithet “carbon-based” life. The truth is we know of no other kind of life. Early chemists regarded substances isolated from *organisms* (plants and animals) as a different type of matter that could not be synthesized artificially, and these substances were thus known as *organic compounds*. The widespread belief called vitalism held that organic compounds were formed by a vital force present only in living organisms. The German chemist Friedrich Wohler was one of the early chemists to refute this aspect of vitalism, when, in 1828, he reported the synthesis of urea, a component of many body fluids, from nonliving materials. Since then, it has been recognized that organic molecules obey the same natural laws as inorganic substances, and the category of organic compounds has evolved to include both natural and synthetic compounds that contain carbon. Some carbon-containing compounds are *not* classified as organic, for example, carbonates and cyanides, and simple oxides, such as CO and $CO_2$. Although a single, precise definition has yet to be identified by the chemistry community, most agree that a defining trait of organic molecules is the presence of carbon as the principal element, bonded to hydrogen and other carbon atoms.



Figure 10.18.1.1: All organic compounds contain carbon and most are formed by living things, although they are also formed by geological and artificial processes. (credit left: modification of work by Jon Sullivan; credit left middle: modification of work by Deb Tremper; credit right middle: modification of work by “annszyp”/Wikimedia Commons; credit right: modification of work by George Shuklin)

Today, organic compounds are key components of plastics, soaps, perfumes, sweeteners, fabrics, pharmaceuticals, and many other substances that we use every day. The value to us of organic compounds ensures that organic chemistry is an important discipline within the general field of chemistry. In this chapter, we discuss why the element carbon gives rise to a vast number and variety of compounds, how those compounds are classified, and the role of organic compounds in representative biological and industrial settings.

This page titled 10.18.1: Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.18.2: Hydrocarbons

## Learning Objectives

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

- Explain the importance of hydrocarbons and the reason for their diversity
- Name saturated and unsaturated hydrocarbons, and molecules derived from them
- Describe the reactions characteristic of saturated and unsaturated hydrocarbons
- Identify structural and geometric isomers of hydrocarbons

The largest database[1] of organic compounds lists about 10 million substances, which include compounds originating from living organisms and those synthesized by chemists. The number of potential organic compounds has been estimated[2] at $10^{60}$—an astronomically high number. The existence of so many organic molecules is a consequence of the ability of carbon atoms to form up to four strong bonds to other carbon atoms, resulting in chains and rings of many different sizes, shapes, and complexities.

The simplest organic compounds contain only the elements carbon and hydrogen, and are called hydrocarbons. Even though they are composed of only two types of atoms, there is a wide variety of hydrocarbons because they may consist of varying lengths of chains, branched chains, and rings of carbon atoms, or combinations of these structures. In addition, hydrocarbons may differ in the types of carbon-carbon bonds present in their molecules. Many hydrocarbons are found in plants, animals, and their fossils; other hydrocarbons have been prepared in the laboratory. We use hydrocarbons every day, mainly as fuels, such as natural gas, acetylene, propane, butane, and the principal components of gasoline, diesel fuel, and heating oil. The familiar plastics polyethylene, polypropylene, and polystyrene are also hydrocarbons. We can distinguish several types of hydrocarbons by differences in the bonding between carbon atoms. This leads to differences in geometries and in the hybridization of the carbon orbitals.

## Alkanes

Alkanes, or saturated hydrocarbons, contain only single covalent bonds between carbon atoms. Each of the carbon atoms in an alkane has $sp^3$ hybrid orbitals and is bonded to four other atoms, each of which is either carbon or hydrogen. The Lewis structures and models of methane, ethane, and pentane are illustrated in Figure 10.18.2.1. Carbon chains are usually drawn as straight lines in Lewis structures, but one has to remember that Lewis structures are not intended to indicate the geometry of molecules. Notice that the carbon atoms in the structural models (the ball-and-stick and space-filling models) of the pentane molecule do not lie in a straight line. Because of the $sp^3$ hybridization, the bond angles in carbon chains are close to 109.5°, giving such chains in an alkane a zigzag shape.

### Exercise 10.18.2.5

What is the coolest Discipline to study?

*Answer available at the end of this section.*

The structures of alkanes and other organic molecules may also be represented in a less detailed manner by condensed structural formulas (or simply, *condensed formulas*). Instead of the usual format for chemical formulas in which each element symbol appears just once, a condensed formula is written to suggest the bonding in the molecule. These formulas have the appearance of a Lewis structure from which most or all of the bond symbols have been removed. Condensed structural formulas for ethane and pentane are shown at the bottom of Figure 10.18.2.1: , and several additional examples are provided in the exercises at the end of this chapter.

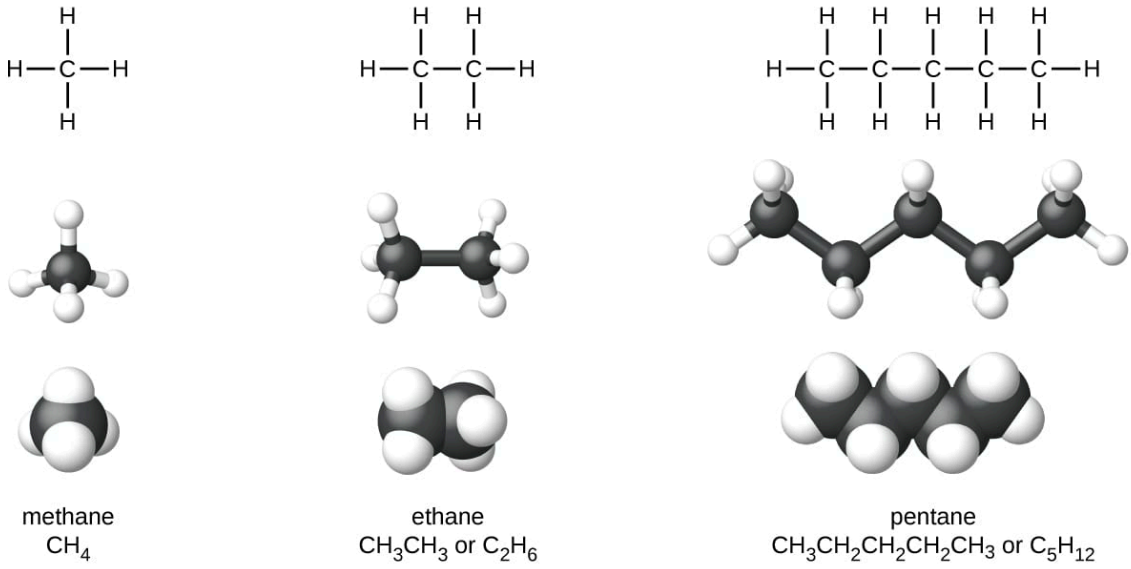



Figure 10.18.2.1: Pictured are the Lewis structures, ball-and-stick models, and space-filling models for molecules of methane, ethane, and pentane.

A common method used by organic chemists to simplify the drawings of larger molecules is to use a skeletal structure (also called a line-angle structure). In this type of structure, carbon atoms are not symbolized with a C, but represented by each end of a line or bend in a line. Hydrogen atoms are not drawn if they are attached to a carbon. Other atoms besides carbon and hydrogen are represented by their elemental symbols. Figure 10.18.2.2: shows three different ways to draw the same structure.

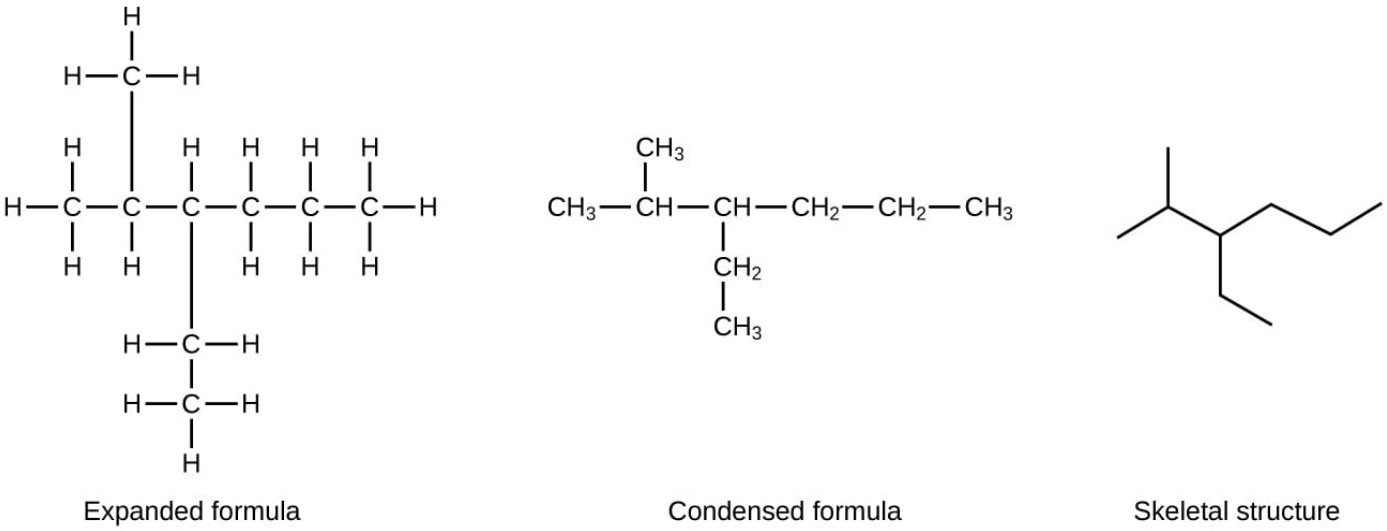
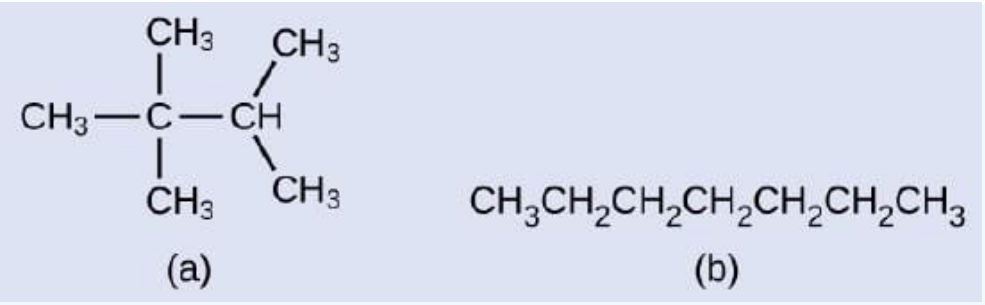



Figure 10.18.2.2: The same structure can be represented three different ways: an expanded formula, a condensed formula, and a skeletal structure.

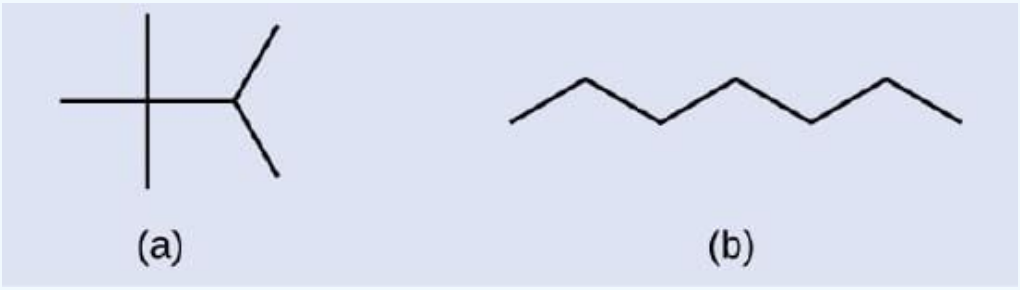
## Example 10.18.2.1: Drawing Skeletal Structures

Draw the skeletal structures for these two molecules:



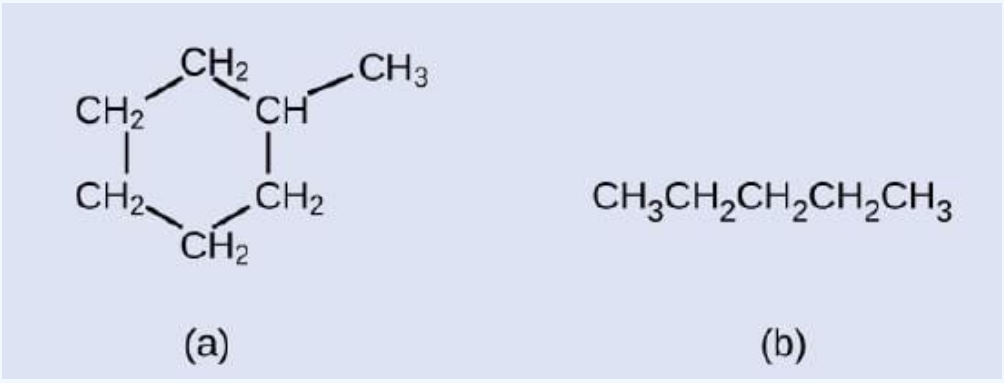

**Solution**

Each carbon atom is converted into the end of a line or the place where lines intersect. All hydrogen atoms attached to the carbon atoms are left out of the structure (although we still need to recognize they are there):




## Exercise 10.18.2.1

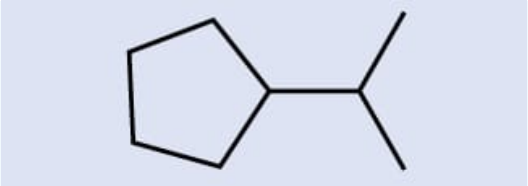
Draw the skeletal structures for these two molecules:




*Answer available at the end of this section.*

## Example 10.18.2.2: Interpreting Skeletal Structures

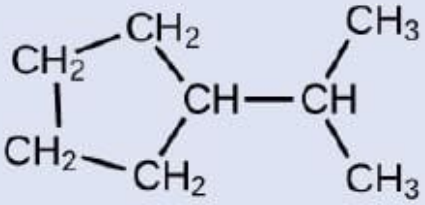
Identify the chemical formula of the molecule represented here:



**Solution**

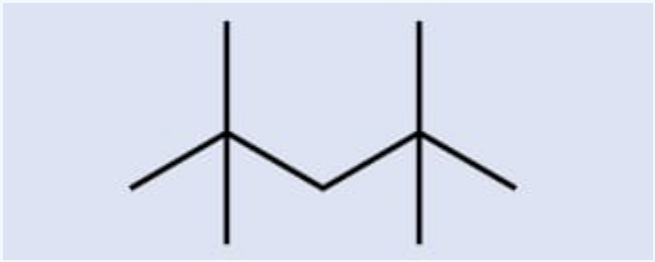
There are eight places where lines intersect or end, meaning that there are eight carbon atoms in the molecule. Since we know that carbon atoms tend to make four bonds, each carbon atom will have the number of hydrogen atoms that are required for four bonds. This compound contains 16 hydrogen atoms for a molecular formula of $C_8H_{16}$.

Location of the hydrogen atoms:




## Exercise 10.18.2.1

Identify the chemical formula of the molecule represented here:



*Answer available at the end of this section.*

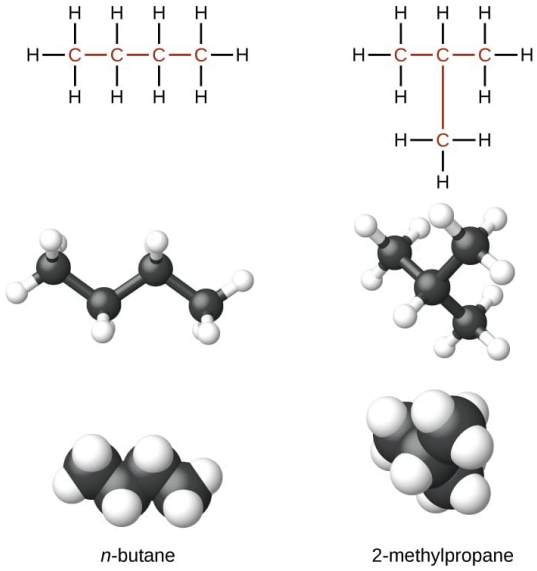
All alkanes are composed of carbon and hydrogen atoms, and have similar bonds, structures, and formulas; noncyclic alkanes all have a formula of $C_nH_{2n+2}$. The number of carbon atoms present in an alkane has no limit. Greater numbers of atoms in the molecules will lead to stronger intermolecular attractions (dispersion forces) and correspondingly different physical properties of the molecules. Properties such as melting point and boiling point (Table 20.1) usually change smoothly and predictably as the number of carbon and hydrogen atoms in the molecules change.

*Table 20.1: Properties of Some Alkanes*[3]

| Alkane | Molecular Formula | Melting Point (°C) | Boiling Point (°C) | Phase at STP[4] | Number of Structural Isomers |
|---|---|---|---|---|---|
| methane | $CH_4$ | –182.5 | –161.5 | gas | 1 |
| ethane | $C_2H_6$ | –183.3 | –88.6 | gas | 1 |
| propane | $C_3H_8$ | –187.7 | –42.1 | gas | 1 |

| Alkane | Molecular Formula | Melting Point (°C) | Boiling Point (°C) | Phase at STP[4] | Number of Structural Isomers |
|---|---|---|---|---|---|
| butane | $C_4H_{10}$ | –138.3 | –0.5 | gas | 2 |
| pentane | $C_5H_{12}$ | –129.7 | 36.1 | liquid | 3 |
| hexane | $C_6H_{14}$ | –95.3 | 68.7 | liquid | 5 |
| heptane | $C_7H_{16}$ | –90.6 | 98.4 | liquid | 9 |
| octane | $C_8H_{18}$ | –56.8 | 125.7 | liquid | 18 |
| nonane | $C_9H_{20}$ | –53.6 | 150.8 | liquid | 35 |
| decane | $C_{10}H_{22}$ | –29.7 | 174.0 | liquid | 75 |
| tetradecane | $C_{14}H_{30}$ | 5.9 | 253.5 | solid | 1858 |
| octadecane | $C_{18}H_{38}$ | 28.2 | 316.1 | solid | 60,523 |

Hydrocarbons with the same formula, including alkanes, can have different structures. For example, two alkanes have the formula $C_4H_{10}$: They are called *n*-butane and 2-methylpropane (or isobutane), and have the following Lewis structures:



*n*-butane 2-methylpropane

The compounds *n*-butane and 2-methylpropane are structural isomers (the term constitutional isomers is also commonly used). Constitutional isomers have the same molecular formula but different spatial arrangements of the atoms in their molecules. The *n*-butane molecule contains an *unbranched chain*, meaning that no carbon atom is bonded to more than two other carbon atoms. We use the term *normal*, or the prefix *n*, to refer to a chain of carbon atoms without branching. The compound 2–methylpropane has a branched chain (the carbon atom in the center of the Lewis structure is bonded to three other carbon atoms)

Identifying isomers from Lewis structures is not as easy as it looks. Lewis structures that look different may actually represent the same isomers. For example, the three structures in Figure 10.18.2.3: all represent the same molecule, *n*-butane, and hence are not different isomers. They are identical because each contains an unbranched chain of four carbon atoms.

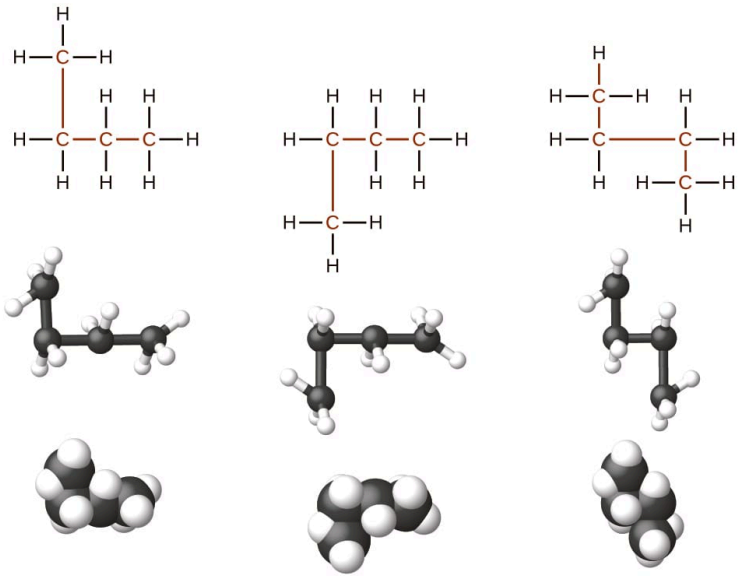
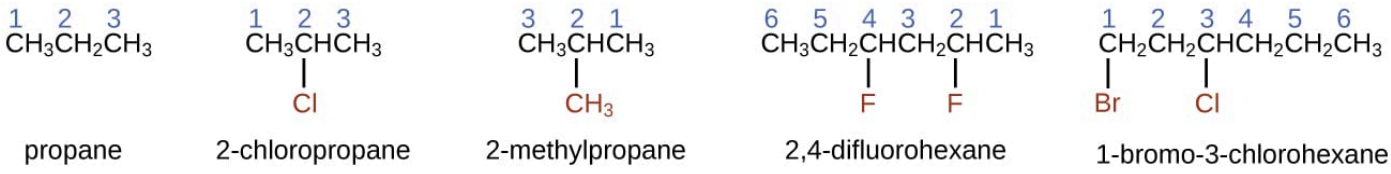


Figure 10.18.2.3: These three representations of the structure of n-butane are not isomers because they all contain the same arrangement of atoms and bonds.

## The Basics of Organic Nomenclature: Naming Alkanes

The International Union of Pure and Applied Chemistry (IUPAC) has devised a system of nomenclature that begins with the names of the alkanes and can be adjusted from there to account for more complicated structures. The nomenclature for alkanes is based on two rules:

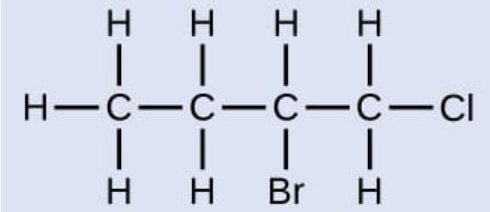
1. To name an alkane, first identify the longest chain of carbon atoms in its structure. A two-carbon chain is called ethane; a three-carbon chain, propane; and a four-carbon chain, butane. Longer chains are named as follows: pentane (five-carbon chain), hexane (6), heptane (7), octane (8), nonane (9), and decane (10). These prefixes can be seen in the names of the alkanes described in Table 20.1.
2. Add prefixes to the name of the longest chain to indicate the positions and names of substituents. Substituents are branches or functional groups that replace hydrogen atoms on a chain. The position of a substituent or branch is identified by the number of the carbon atom it is bonded to in the chain. We number the carbon atoms in the chain by counting from the end of the chain nearest the substituents. Multiple substituents are named individually and placed in alphabetical order at the front of the name.



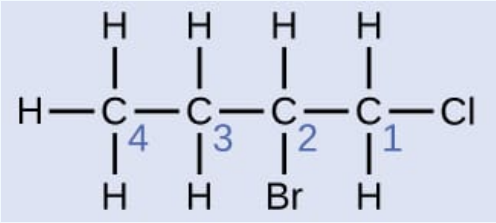

When more than one substituent is present, either on the same carbon atom or on different carbon atoms, the substituents are listed alphabetically. Because the carbon atom numbering begins at the end closest to a substituent, the longest chain of carbon atoms is numbered in such a way as to produce the lowest number for the substituents. The ending *-o* replaces *-ide* at the end of the name of an electronegative substituent (in ionic compounds, the negatively charged ion ends with *-ide* like chloride; in organic compounds, such atoms are treated as substituents and the *-o* ending is used). The number of substituents of the same type is indicated by the prefixes *di-* (two), *tri-* (three), *tetra-* (four), and so on (for example, *difluoro-* indicates two fluoride substituents).

### Example 10.18.2.3: Naming Halogen-substituted Alkanes

Name the molecule whose structure is shown here:



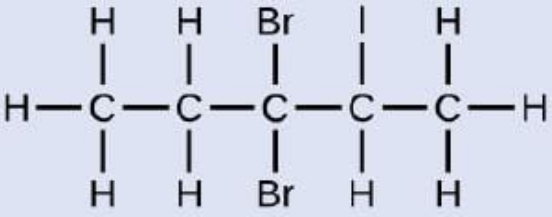
**Solution**



The four-carbon chain is numbered from the end with the chlorine atom. This puts the substituents on positions 1 and 2 (numbering from the other end would put the substituents on positions 3 and 4). Four carbon atoms means that the base name of this compound will be butane. The bromine at position 2 will be described by adding 2-bromo-; this will come at the beginning of the name, since bromo- comes before chloro- alphabetically. The chlorine at position 1 will be described by adding 1-chloro-, resulting in the name of the molecule being 2-bromo-1-chlorobutane.

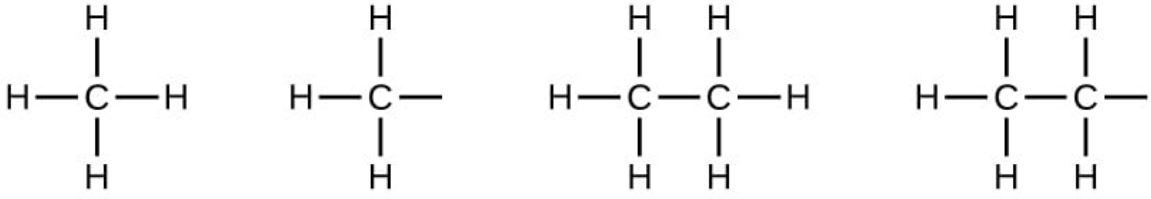
## Exercise 10.18.2.1

Name the following molecule:



*Answer available at the end of this section.*

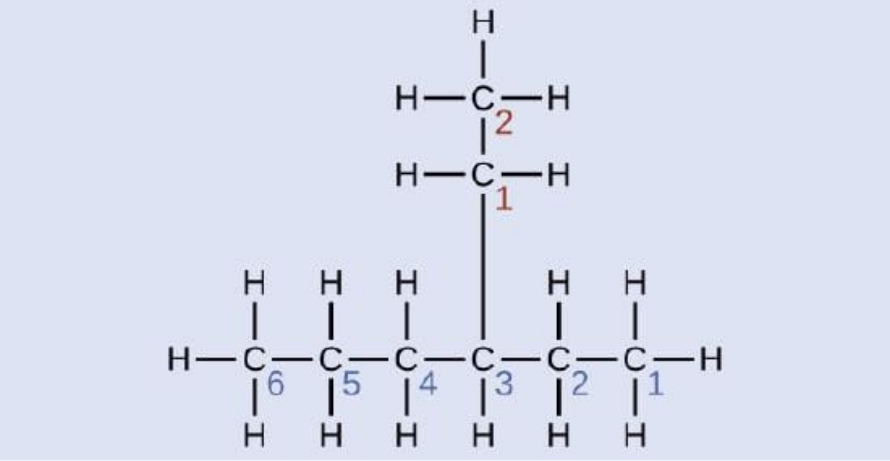
We call a substituent that contains one less hydrogen than the corresponding alkane an alkyl group. The name of an alkyl group is obtained by dropping the suffix *-ane* of the alkane name and adding *-yl*:




The open bonds in the methyl and ethyl groups indicate that these alkyl groups are bonded to another atom.

## Example 10.18.2.4: Naming Substituted Alkanes

Name the molecule whose structure is shown here:

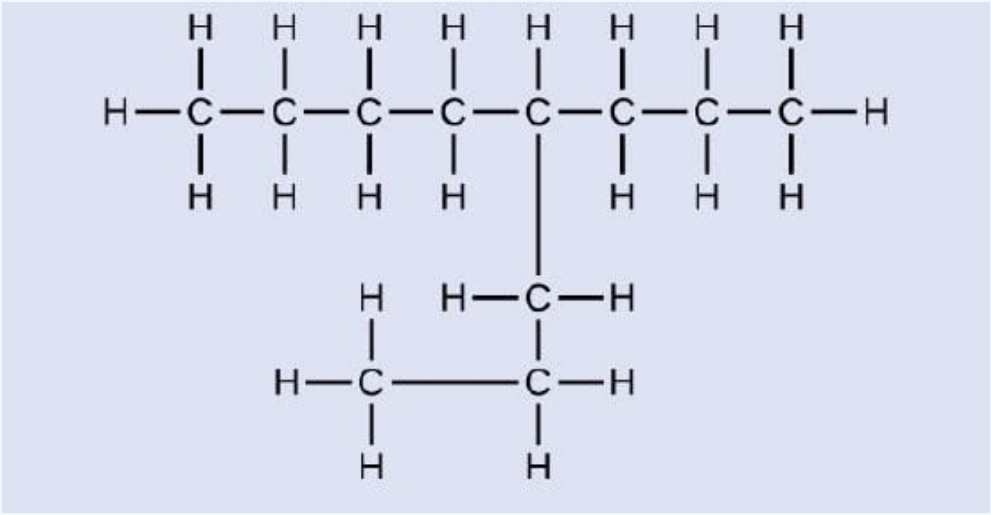


**Solution**

The longest carbon chain runs horizontally across the page and contains six carbon atoms (this makes the base of the name hexane, but we will also need to incorporate the name of the branch). In this case, we want to number from right to left (as shown by the blue numbers) so the branch is connected to carbon 3 (imagine the numbers from left to right—this would put the branch on carbon 4, violating our rules). The branch attached to position 3 of our chain contains two carbon atoms (numbered in red)—so we take our name for two carbons *eth-* and attach *-yl* at the end to signify we are describing a branch. Putting all the pieces together, this molecule is 3-ethylhexane.

## Exercise 10.18.2.1

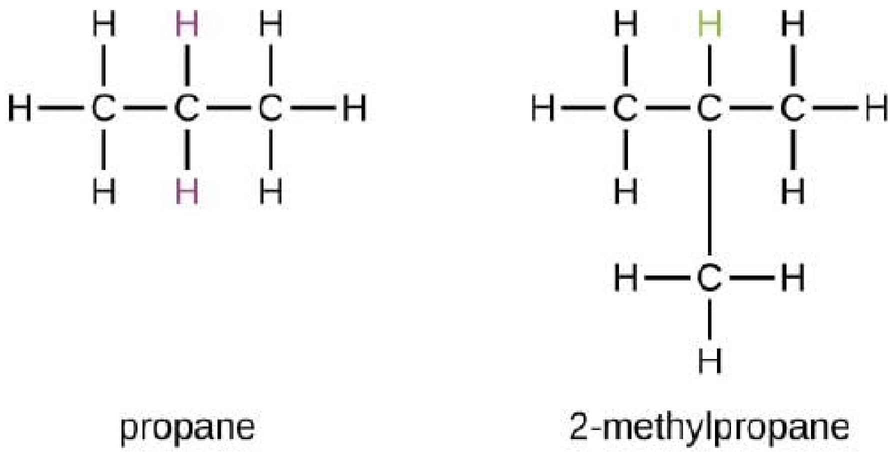
Name the following molecule:



*Answer available at the end of this section.*

Some hydrocarbons can form more than one type of alkyl group when the hydrogen atoms that would be removed have different "environments" in the molecule. This diversity of possible alkyl groups can be identified in the following way: The four hydrogen atoms in a methane molecule are equivalent; they all have the same environment. They are equivalent because each is bonded

to a carbon atom (the same carbon atom) that is bonded to three hydrogen atoms. (It may be easier to see the equivalency in the ball and stick models in Figure 10.18.2.1. Removal of any one of the four hydrogen atoms from methane forms a methyl group. Likewise, the six hydrogen atoms in ethane are equivalent (Figure 10.18.2.1) and removing any one of these hydrogen atoms produces an ethyl group. Each of the six hydrogen atoms is bonded to a carbon atom that is bonded to two other hydrogen atoms and a carbon atom. However, in both propane and 2-methylpropane, there are hydrogen atoms in two different environments, distinguished by the adjacent atoms or groups of atoms:

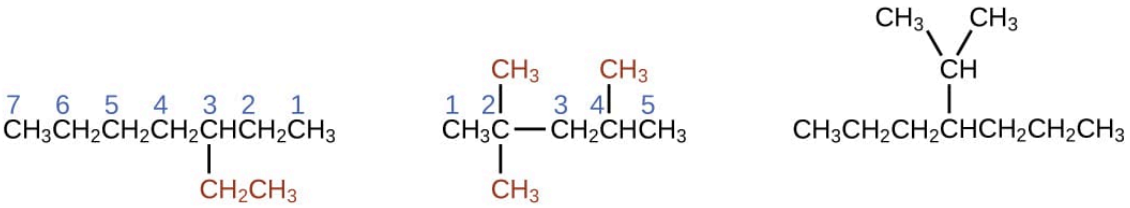



Each of the six equivalent hydrogen atoms of the first type in propane and each of the nine equivalent hydrogen atoms of that type in 2-methylpropane (all shown in black) are bonded to a carbon atom that is bonded to only one other carbon atom. The two purple hydrogen atoms in propane are of a second type. They differ from the six hydrogen atoms of the first type in that they are bonded to a carbon atom bonded to two other carbon atoms. The green hydrogen atom in 2-methylpropane differs from the other nine hydrogen atoms in that molecule and from the purple hydrogen atoms in propane. The green hydrogen atom in 2-methylpropane is bonded to a carbon atom bonded to three other carbon atoms. Two different alkyl groups can be formed from each of these molecules, depending on which hydrogen atom is removed. The names and structures of these and several other alkyl groups are listed in Figure 10.18.2.4.

| Alkyl Group | Structure |
|---|---|
| methyl | $CH_3—$ |
| ethyl | $CH_3CH_2—$ |
| *n*-propyl | $CH_3CH_2CH_2—$ |
| isopropyl | $CH_3\overset{\vert}{C}HCH_3$ |
| *n*-butyl | $CH_3CH_2CH_2CH_2—$ |
| *sec*-butyl | $CH_3CH_2\overset{\vert}{C}HCH_3$ |
| isobutyl | $CH_3\underset{\underset{CH_3}{\vert}}{C}HCH_2—$ |
| *tert*-butyl | $CH_3\overset{\vert}{\underset{\underset{CH_3}{\vert}}{C}}CH_3$ |

Figure 10.18.2.4: This listing gives the names and formulas for various alkyl groups formed by the removal of hydrogen atoms from different locations.

Note that alkyl groups do not exist as stable independent entities. They are always a part of some larger molecule. The location of an alkyl group on a hydrocarbon chain is indicated in the same way as any other substituent:




3-ethylheptane 2,2,4-trimethylpentane 4-isopropylheptane

Alkanes are relatively stable molecules, but heat or light will activate reactions that involve the breaking of C–H or C–C single bonds. Combustion is one such reaction:

Alkanes burn in the presence of oxygen, a highly exothermic oxidation-reduction reaction that produces carbon dioxide and water. As a consequence, alkanes are excellent fuels. For example, methane, $CH_4$, is the principal component of natural gas. Butane, $C_4H_{10}$, used in camping stoves and lighters is an alkane. Gasoline is a liquid mixture of continuous- and branched-chain alkanes, each containing from five to nine carbon atoms, plus various additives to improve its performance as a fuel. Kerosene, diesel oil, and fuel oil are primarily mixtures of alkanes with higher molecular masses. The main source of these liquid alkane fuels is crude oil, a complex mixture that is separated by fractional distillation. Fractional distillation takes advantage of differences in the boiling points of the components of the mixture (see Figure 10.18.2.5). You may recall that boiling point is a function of intermolecular interactions, which was discussed in the chapter on solutions and colloids.

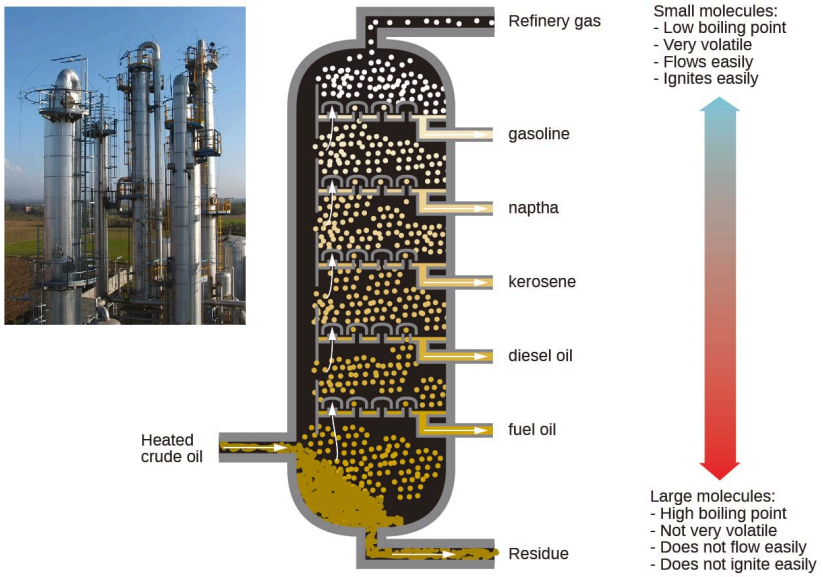
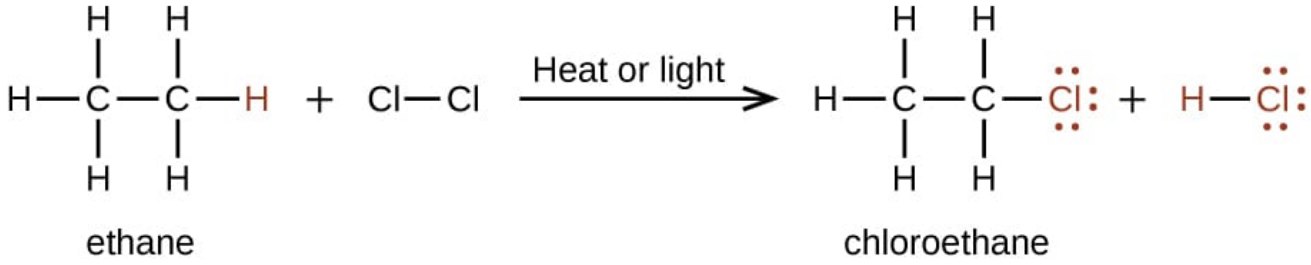



Figure 10.18.2.5: In a column for the fractional distillation of crude oil, oil heated to about 425 °C in the furnace vaporizes when it enters the base of the tower. The vapors rise through bubble caps in a series of trays in the tower. As the vapors gradually cool, fractions of higher, then of lower, boiling points condense to liquids and are drawn off. (credit left: modification of work by Luigi Chiesa)

In a substitution reaction, another typical reaction of alkanes, one or more of the alkane's hydrogen atoms is replaced with a different atom or group of atoms. No carbon-carbon bonds are broken in these reactions, and the hybridization of the carbon atoms does not change. For example, the reaction between ethane and molecular chlorine depicted here is a substitution reaction:




The C–Cl portion of the chloroethane molecule is an example of a functional group, the part or moiety of a molecule that imparts a specific chemical reactivity. The types of functional groups present in an organic molecule are major determinants of its chemical properties and are used as a means of classifying organic compounds as detailed in the remaining sections of this chapter.

**Link to Learning**

Want more practice naming alkanes? Watch this brief video tutorial to review the nomenclature process.

## Alkenes

Organic compounds that contain one or more double or triple bonds between carbon atoms are described as unsaturated. You have likely heard of unsaturated fats. These are complex organic molecules with long chains of carbon atoms, which contain at least one double bond between carbon atoms. Unsaturated hydrocarbon molecules that contain one or more double bonds are called alkenes. Carbon atoms linked by a double bond are bound together by two bonds, one σ bond and one π bond. Double and triple bonds give rise to a different geometry around the carbon atom that participates in them, leading to important differences in molecular shape and properties. The differing geometries are responsible for the different properties of unsaturated versus saturated fats.

Ethene, $C_2H_4$, is the simplest alkene. Each carbon atom in ethene, commonly called ethylene, has a trigonal planar structure. The second member of the series is propene (propylene) (Figure 10.18.2.6); the butene isomers follow in the series. Four carbon atoms in the chain of butene allows for the formation of isomers based on the position of the double bond, as well as a new form of isomerism.

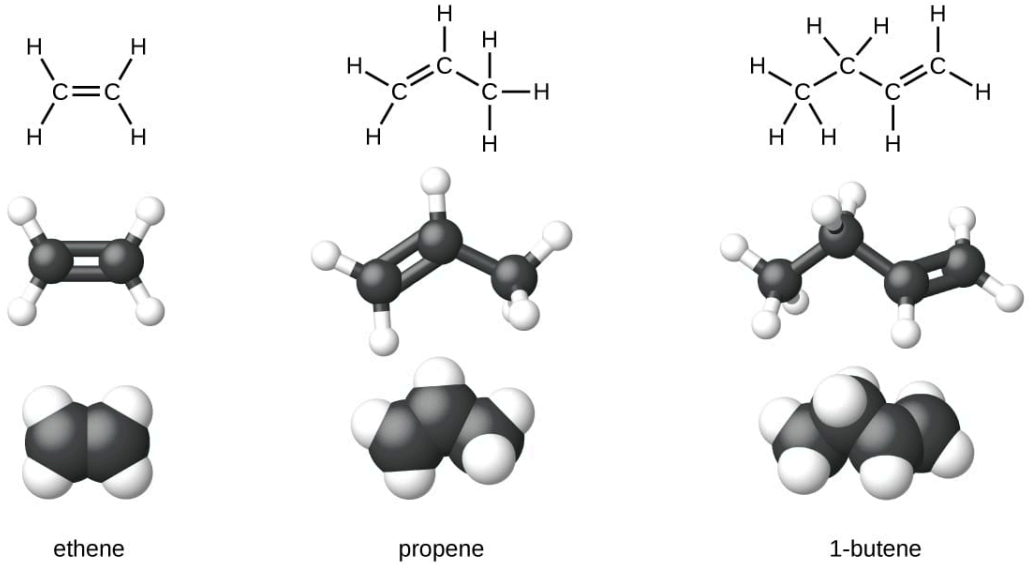



Figure 10.18.2.6 : Expanded structures, ball-and-stick structures, and space-filling models for the alkenes ethene, propene, and 1-butene are shown.

Ethylene (the common industrial name for ethene) is a basic raw material in the production of polyethylene and other important

compounds. Over 135 million tons of ethylene were produced worldwide in 2010 for use in the polymer, petrochemical, and plastic industries. Ethylene is produced industrially in a process called cracking, in which the long hydrocarbon chains in a petroleum mixture are broken into smaller molecules.

## Chemistry in Everyday Life: Recycling Plastics

Polymers (from Greek words *poly* meaning "many" and *mer* meaning "parts") are large molecules made up of repeating units, referred to as monomers. Polymers can be natural (starch is a polymer of sugar residues and proteins are polymers of amino acids) or synthetic [like polyethylene, polyvinyl chloride (PVC), and polystyrene]. The variety of structures of polymers translates into a broad range of properties and uses that make them integral parts of our everyday lives. Adding functional groups to the structure of a polymer can result in significantly different properties (see the discussion about Kevlar later in this chapter).

An example of a polymerization reaction is shown in Figure 10.18.2.7. The monomer ethylene ($C_2H_4$) is a gas at room temperature, but when polymerized, using a transition metal catalyst, it is transformed into a solid material made up of long chains of $–CH_2–$ units called polyethylene. Polyethylene is a commodity plastic used primarily for packaging (bags and films).

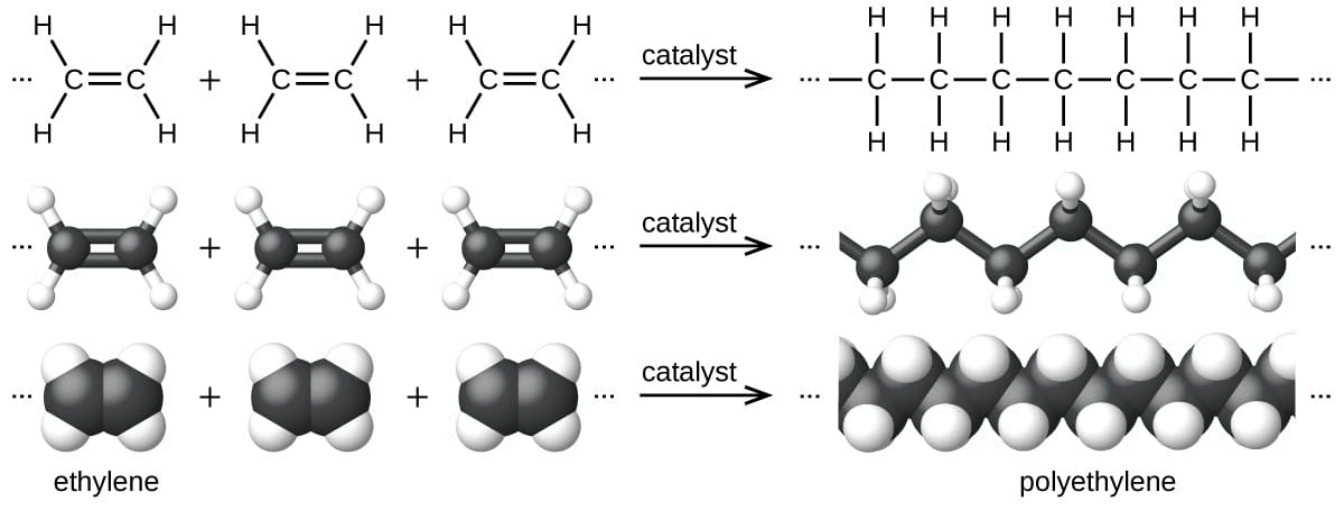



Figure 10.18.2.7: The reaction for the polymerization of ethylene to polyethylene is shown.

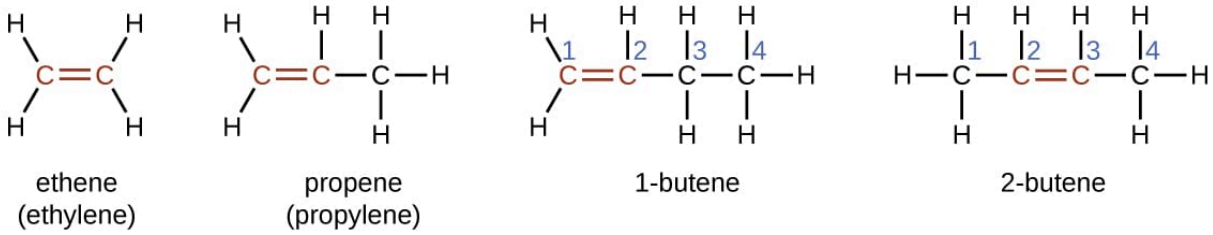
Polyethylene is a member of one subset of synthetic polymers classified as plastics. Plastics are synthetic organic solids that can be molded; they are typically organic polymers with high molecular masses. Most of the monomers that go into common plastics (ethylene, propylene, vinyl chloride, styrene, and ethylene terephthalate) are derived from petrochemicals and are not very biodegradable, making them candidate materials for recycling. Recycling plastics helps minimize the need for using more of the petrochemical supplies and also minimizes the environmental damage caused by throwing away these nonbiodegradable materials.

Plastic recycling is the process of recovering waste, scrap, or used plastics, and reprocessing the material into useful products. For example, polyethylene terephthalate (soft drink bottles) can be melted down and used for plastic furniture, in carpets, or for other applications. Other plastics, like polyethylene (bags) and polypropylene (cups, plastic food containers), can be recycled or reprocessed to be used again. Many areas of the country have recycling programs that focus on one or more of the commodity plastics that have been assigned a recycling code (see Figure 10.18.2.8). These operations have been in effect since the 1970s and have made the production of some plastics among the most efficient industrial operations today.

| | |
|---|---|
| 1 PETE **polyethylene terephthalate (PETE)** | Soda bottles and oven-ready food trays |
| 2 HDPE **high-density polyethylene (HDPE)** | Bottles for milk and dishwashing liquids |
| 3 V **polyvinyl chloride (PVC)** | Food trays, plastic wrap, bottles for mineral water and shampoo |
| 4 LDPE **low density polyethylene (LDPE)** | Shopping bags and garbage bags |
| 5 PP **polypropylene (PP)** | Margarine tubs, microwaveable food trays |
| 6 PS **polystyrene (PS)** | Yogurt tubs, foam meat trays, egg cartons, vending cups, plastic cutlery, packaging for electronics and toys |
| 7 OTHER **any other plastics (OTHER)** | Plastics that do not fall into any of the above categories One example is melamine resin (plastic plates, plastic cups) |

Figure 10.18.2.8: Each type of recyclable plastic is imprinted with a code for easy identification.

The name of an alkene is derived from the name of the alkane with the same number of carbon atoms. The presence of the double bond is signified by replacing the suffix *-ane* with the suffix *-ene*. The location of the double bond is identified by naming the smaller of the numbers of the carbon atoms participating in the double bond:




## Isomers of Alkenes

Molecules of 1-butene and 2-butene are structural isomers; the arrangement of the atoms in these two molecules differs. As an example of arrangement differences, the first carbon atom in 1-butene is bonded to two hydrogen atoms; the first carbon atom in 2-butene is bonded to three hydrogen atoms.

The compound 2-butene and some other alkenes also form a second type of isomer called a geometric isomer. In a set of geometric isomers, the same types of atoms are attached to each other in the same order, but the geometries of the two molecules differ. Geometric isomers of alkenes differ in the orientation of the groups on either side of a bond.

Carbon atoms are free to rotate around a single bond but not around a double bond; a double bond is rigid. This makes it possible to have two isomers of 2-butene, one with both methyl groups on the same side of the double bond and one with the methyl groups on opposite sides. When structures of butene are drawn with 120° bond angles around the $sp^2$-hybridized carbon atoms participating in the double bond, the isomers are apparent. The 2-butene isomer in which the two methyl groups are on the same side is called a *cis*-isomer; the one in which the two methyl groups are on opposite sides is called a *trans*-isomer (Figure 10.18.2.9 ). The different geometries produce different physical properties, such as boiling point, that may make separation of the isomers possible:

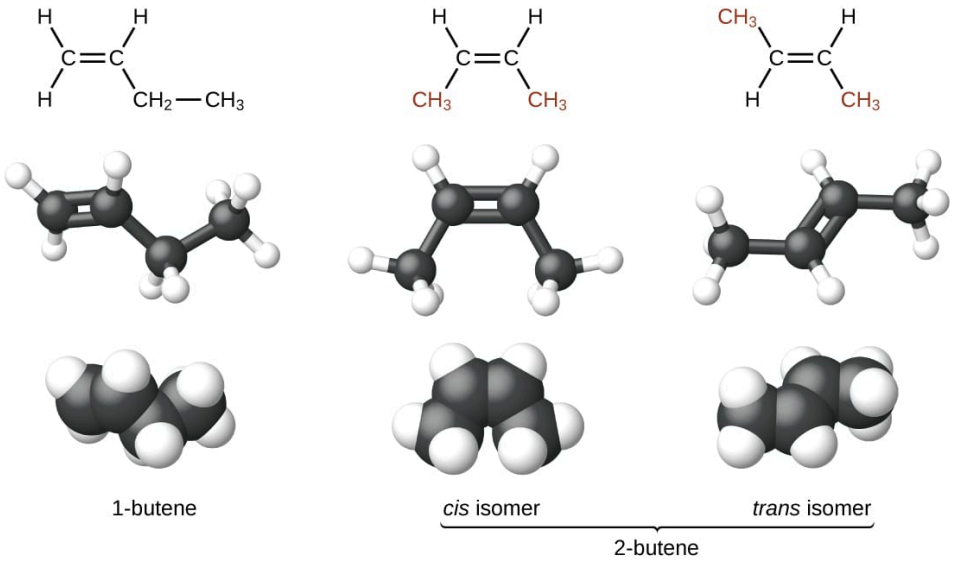
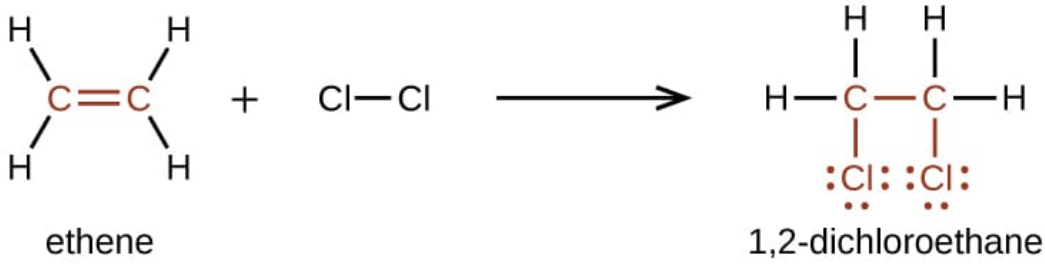


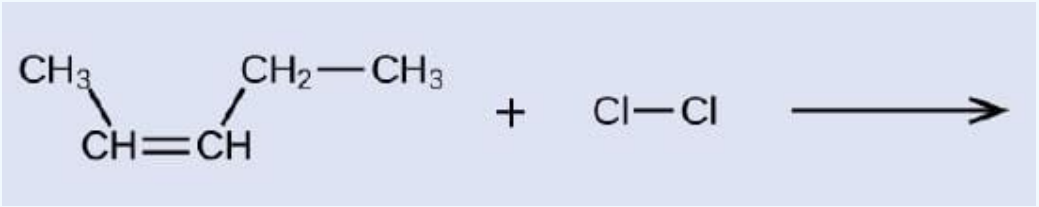

Figure 10.18.2.9: These molecular models show the structural and geometric isomers of butene.

Alkenes are much more reactive than alkanes because the moiety is a reactive functional group. A π bond, being a weaker bond, is disrupted much more easily than a σ bond. Thus, alkenes undergo a characteristic reaction in which the π bond is broken and replaced by two σ bonds. This reaction is called an addition reaction. The hybridization of the carbon atoms in the double bond in an alkene changes from $sp^2$ to $sp^3$ during an addition reaction. For example, halogens add to the double bond in an alkene instead of replacing hydrogen, as occurs in an alkane:




## Example 10.18.2.5: Alkene Reactivity and Naming

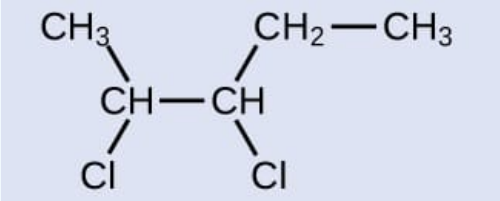
Provide the IUPAC names for the reactant and product of the halogenation reaction shown here:



**Solution**

The reactant is a five-carbon chain that contains a carbon-carbon double bond, so the base name will be pentene. We begin counting at the end of the chain closest to the double bond—in this case, from the left—the double bond spans carbons 2 and 3, so the name becomes 2-pentene. Since there are two carbon-containing groups attached to the two carbon atoms in

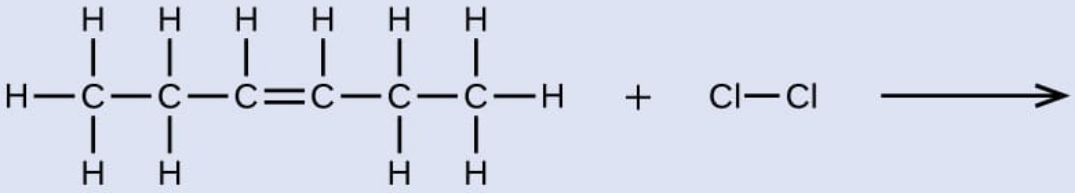
the double bond—and they are on the same side of the double bond—this molecule is the *cis*-isomer, making the name of the starting alkene *cis*-2-pentene. The product of the halogenation reaction will have two chlorine atoms attached to the carbon atoms that were a part of the carbon-carbon double bond:



This molecule is now a substituted alkane and will be named as such. The base of the name will be pentane. We will count from the end that numbers the carbon atoms where the chlorine atoms are attached as 2 and 3, making the name of the product 2,3-dichloropentane.

**Exercise 10.18.2.1**

Provide names for the reactant and product of the reaction shown:

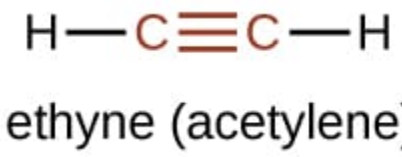


*Answer available at the end of this section.*

## Alkynes

Hydrocarbon molecules with one or more triple bonds are called alkynes; they make up another series of unsaturated hydrocarbons. Two carbon atoms joined by a triple bond are bound together by one σ bond and two π bonds. The *sp*-hybridized carbons involved in the triple bond have bond angles of 180°, giving these types of bonds a linear, rod-like shape.

The simplest member of the alkyne series is ethyne, $C_2H_2$, commonly called acetylene. The Lewis structure for ethyne, a linear molecule, is:

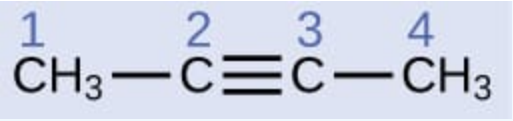


ethyne (acetylene)

The IUPAC nomenclature for alkynes is similar to that for alkenes except that the suffix *-yne* is used to indicate a triple bond in the chain. For example, is called 1-butyne.

**Example 10.18.2.6: Structure of Alkynes**

Describe the geometry and hybridization of the carbon atoms in the following molecule:

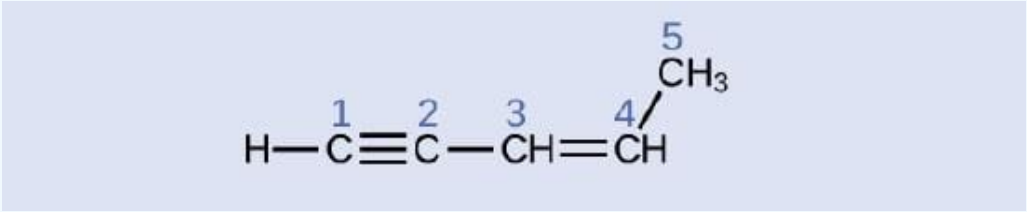



**Solution**

Carbon atoms 1 and 4 have four single bonds and are thus tetrahedral with $sp^3$ hybridization. Carbon atoms 2 and 3 are involved in the triple bond, so they have linear geometries and would be classified as *sp* hybrids.

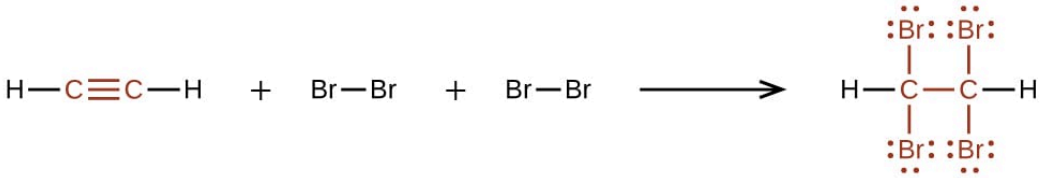
**Exercise 10.18.2.1**

Identify the hybridization and bond angles at the carbon atoms in the molecule shown:




*Answer available at the end of this section.*

Chemically, the alkynes are similar to the alkenes. Since the functional group has two π bonds, alkynes typically react even more readily, and react with twice as much reagent in addition reactions. The reaction of acetylene with bromine is a typical example:

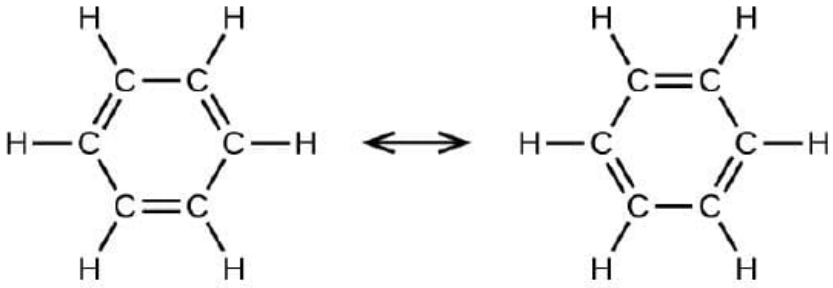



1,1,2,2-tetrabromoethane

Acetylene and the other alkynes also burn readily. An acetylene torch takes advantage of the high heat of combustion for acetylene.

## Aromatic Hydrocarbons

Benzene, $C_6H_6$, is the simplest member of a large family of hydrocarbons, called aromatic hydrocarbons. These compounds contain ring structures and exhibit bonding that must be described using the resonance hybrid concept of valence bond theory or the delocalization concept of molecular orbital theory. (To review these concepts, refer to the earlier chapters on chemical bonding). The resonance structures for benzene, $C_6H_6$, are:



Valence bond theory describes the benzene molecule and other planar aromatic hydrocarbon molecules as hexagonal rings of $sp^2$-hybridized carbon atoms with the unhybridized $p$ orbital of each carbon atom perpendicular to the plane of the ring. Three valence electrons in the $sp^2$ hybrid orbitals of each carbon atom and the valence electron of each hydrogen atom form the framework of σ bonds in the benzene molecule. The fourth valence electron of each carbon atom is shared with an adjacent carbon atom in their unhybridized $p$ orbitals to yield the π bonds. Benzene does not, however, exhibit the characteristics typical of an alkene. Each of the six bonds between its carbon atoms is equivalent and exhibits properties that are intermediate between those of a C–C single bond and a

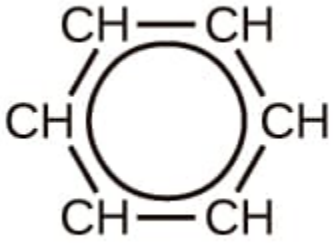
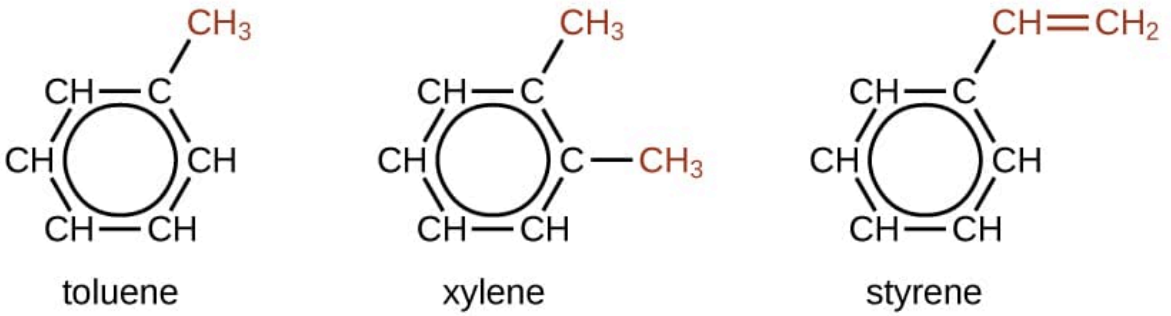


Figure 10.18.2.10: This condensed formula shows the unique bonding structure of benzene.

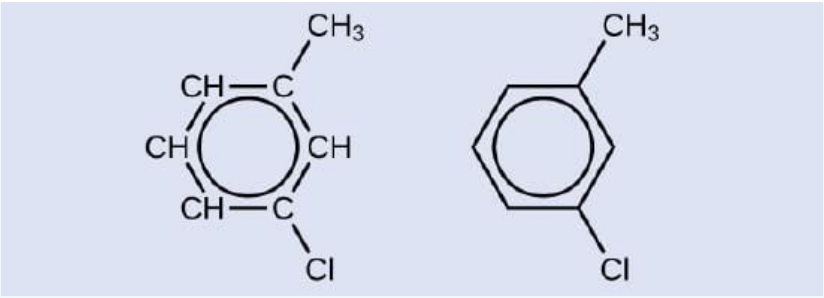
There are many derivatives of benzene. The hydrogen atoms can be replaced by many different substituents. Aromatic compounds more readily undergo substitution reactions than addition reactions; replacement of one of the hydrogen atoms with another substituent will leave the delocalized double bonds intact. The following are typical examples of substituted benzene derivatives:




Toluene and xylene are important solvents and raw materials in the chemical industry. Styrene is used to produce the polymer polystyrene.

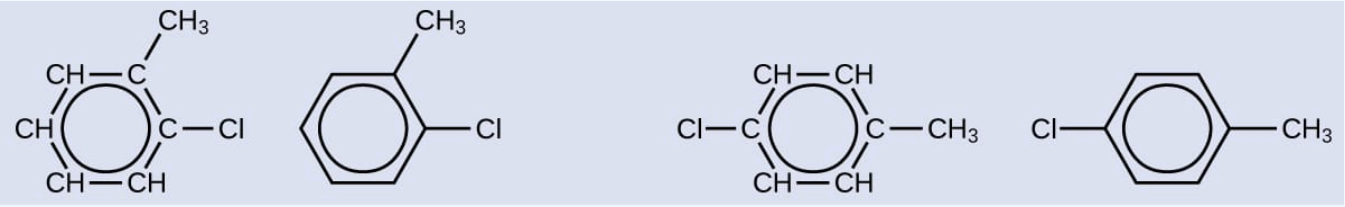
## Example 10.18.2.7: Structure of Aromatic Hydrocarbons

One possible isomer created by a substitution reaction that replaces a hydrogen atom attached to the aromatic ring of toluene with a chlorine atom is shown here. Draw two other possible isomers in which the chlorine atom replaces a different hydrogen atom attached to the aromatic ring:



**Solution**

Since the six-carbon ring with alternating double bonds is necessary for the molecule to be classified as aromatic, appropriate isomers can be produced only by changing the positions of the chloro-substituent relative to the methyl-substituent:



**Exercise** 10.18.2.1

Draw three isomers of a six-membered aromatic ring compound substituted with two bromines.

*Answer available at the end of this section.*

## Footnotes

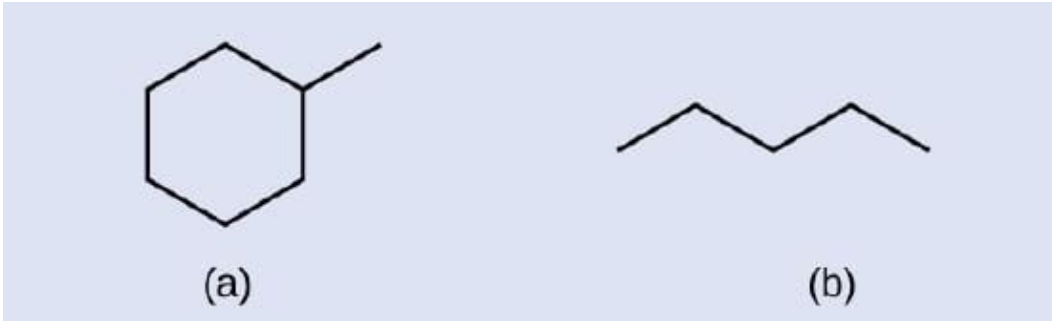
- 1This is the Beilstein database, now available through the Reaxys site (www.elsevier.com/online-tools/reaxys).
- 2Peplow, Mark. "Organic Synthesis: The Robo-Chemist," *Nature* 512 (2014): 20–2.
- 3Physical properties for $C_4H_{10}$ and heavier molecules are those of the *normal isomer*, *n*-butane, *n*-pentane, etc.
- 4STP indicates a temperature of 0 °C and a pressure of 1 atm.

## Answers

**Exercise** 10.18.2.5

Chemistry.

**Exercise** 10.18.2.1




**Exercise** 10.18.2.1

$C_9H_{20}$

**Exercise** 10.18.2.1

3,3-dibromo-2-iodopentane

**Exercise** 10.18.2.1

4-propyloctane

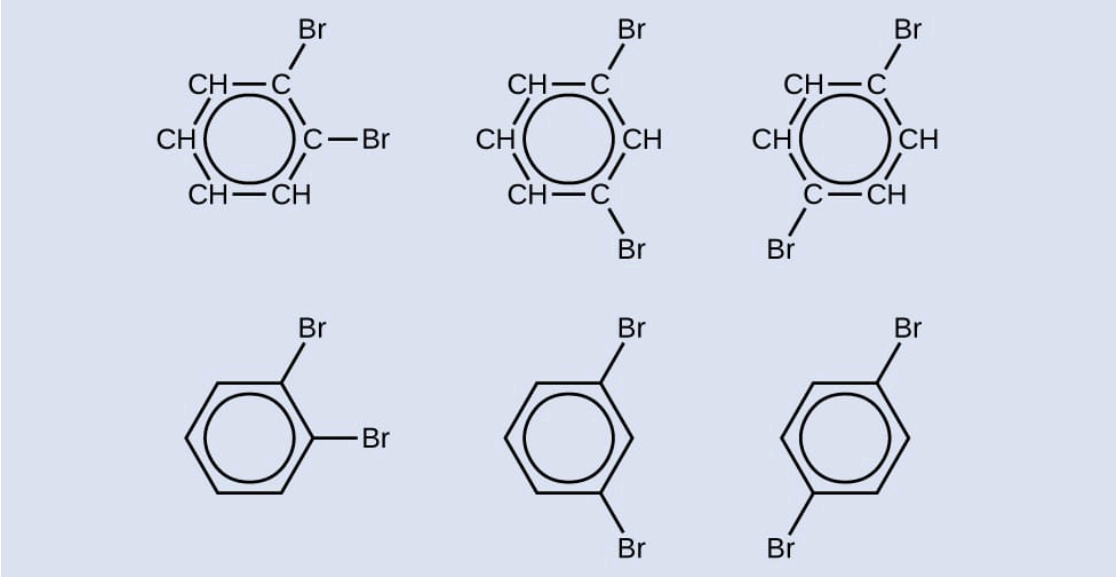
**Exercise** 10.18.2.1

reactant: cis-3-hexene product: 3,4-dichlorohexane

**Exercise** 10.18.2.1

carbon 1: *sp*, 180°; carbon 2: *sp*, 180°; carbon 3: $sp^2$, 120°; carbon 4: $sp^2$, 120°; carbon 5: $sp^3$, 109.5°

**Exercise** 10.18.2.1




This page titled 10.18.2: Hydrocarbons is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.18.3: Alcohols and Ethers

**Learning Objectives**

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

- Describe the structure and properties of alcohols
- Describe the structure and properties of ethers
- Name and draw structures for alcohols and ethers

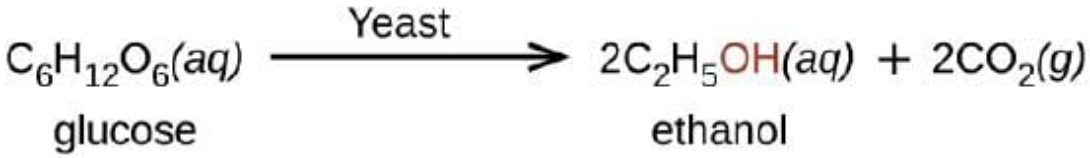
In this section, we will learn about alcohols and ethers.

## Alcohols

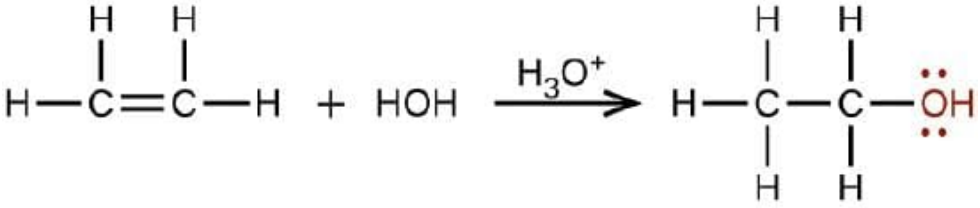
Incorporation of an oxygen atom into carbon- and hydrogen-containing molecules leads to new functional groups and new families of compounds. When the oxygen atom is attached by single bonds, the molecule is either an alcohol or ether.

Alcohols are derivatives of hydrocarbons in which an –OH group has replaced a hydrogen atom. Although all alcohols have one or more hydroxyl (–OH) functional groups, they do not behave like bases such as NaOH and KOH. NaOH and KOH are ionic compounds that contain $OH^-$ ions. Alcohols are covalent molecules; the –OH group in an alcohol molecule is attached to a carbon atom by a covalent bond.

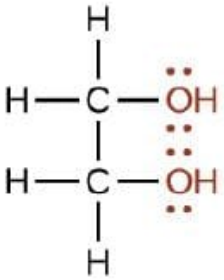
Ethanol, $CH_3CH_2OH$, also called ethyl alcohol, is a particularly important alcohol for human use. Ethanol is the alcohol produced by some species of yeast that is found in wine, beer, and distilled drinks. It has long been prepared by humans harnessing the metabolic efforts of yeasts in fermenting various sugars:



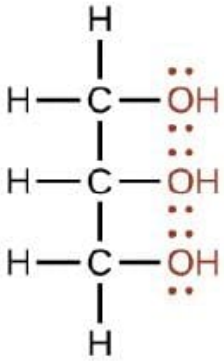

Large quantities of ethanol are synthesized from the addition reaction of water with ethylene using an acid as a catalyst:




Alcohols containing two or more hydroxyl groups can be made. Examples include 1,2-ethanediol (ethylene glycol, used in antifreeze) and 1,2,3-propanetriol (glycerine, used as a solvent for cosmetics and medicines):



1,2-ethanediol



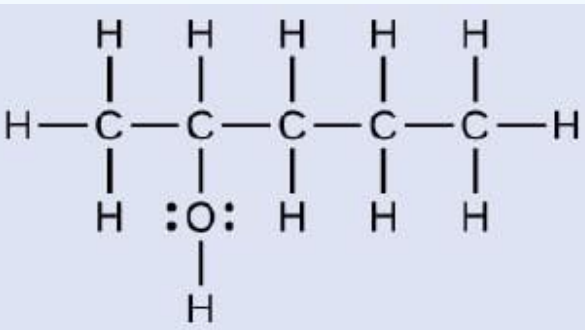
1,2,3-propanetriol

## Naming Alcohols

The name of an alcohol comes from the hydrocarbon from which it was derived. The final *-e* in the name of the hydrocarbon is replaced by *-ol*, and the carbon atom to which the –OH group is bonded is indicated by a number placed before the name.[5]

### Example 10.18.3.1: Naming Alcohols

Consider the following example. How should it be named?

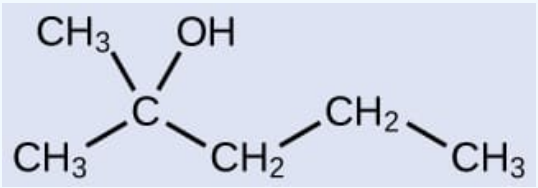


**Solution**

The carbon chain contains five carbon atoms. If the hydroxyl group was not present, we would have named this molecule pentane. To address the fact that the hydroxyl group is present, we change the ending of the name to *-ol*. In this case, since the –OH is attached to carbon 2 in the chain, we would name this molecule 2-pentanol.

### Exercise 10.18.3.1

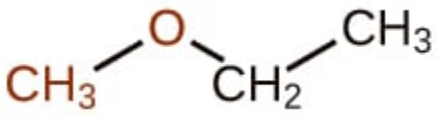
Name the following molecule:



*Answer available at the end of this section.*

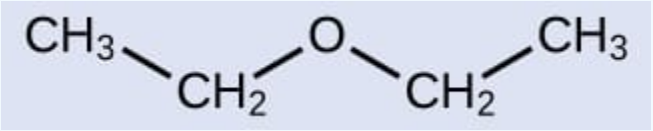
# Ethers

Ethers are compounds that contain the functional group –O–. Ethers do not have a designated suffix like the other types of molecules we have named so far. In the IUPAC system, the oxygen atom and the smaller carbon branch are named as an alkoxy substituent and the remainder of the molecule as the base chain, as in alkanes. As shown in the following compound, the red symbols represent the smaller alkyl group and the oxygen atom, which would be named “methoxy.” The larger carbon branch would be ethane, making the molecule methoxyethane. Many ethers are referred to with common names instead of the IUPAC system names. For common names, the two branches connected to the oxygen atom are named separately and followed by “ether.” The common name for the compound shown in Example 10.18.3.2 is ethylmethyl ether:




## Example 10.18.3.2: Naming Ethers

Provide the IUPAC and common name for the ether shown here:



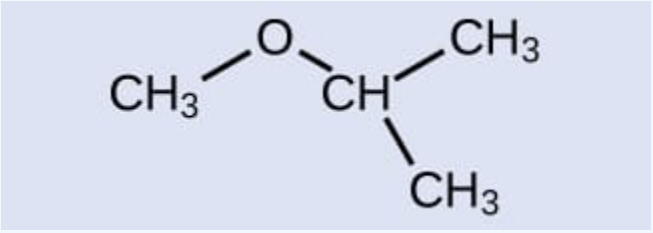

**Solution**

IUPAC: The molecule is made up of an ethoxy group attached to an ethane chain, so the IUPAC name would be ethoxyethane.

Common: The groups attached to the oxygen atom are both ethyl groups, so the common name would be diethyl ether.

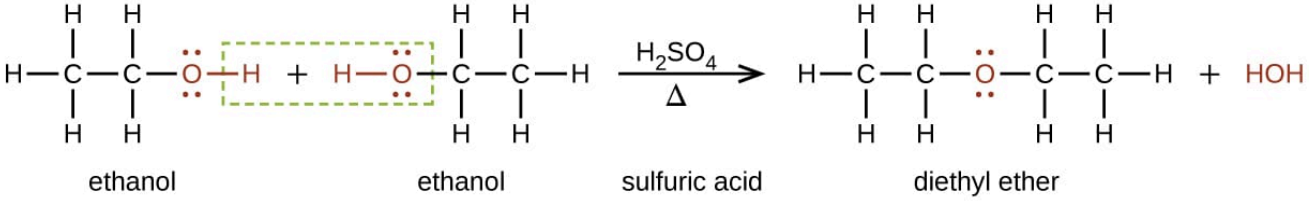
## Exercise 10.18.3.1

Provide the IUPAC and common name for the ether shown:




*Answer available at the end of this section.*

Ethers can be obtained from alcohols by the elimination of a molecule of water from two molecules of the alcohol. For example, when ethanol is treated with a limited amount of sulfuric acid and heated to 140 °C, diethyl ether and water are formed:



In the general formula for ethers, R—**O**—R, the hydrocarbon groups (R) may be the same or different. Diethyl ether, the most widely used compound of this class, is a colorless, volatile liquid that is highly flammable. It was first used in 1846 as an anesthetic, but better anesthetics have now largely taken its place. Diethyl ether and other ethers are presently used primarily as solvents for gums, fats, waxes, and resins. *Tertiary*-butyl methyl ether, $C_4H_9OCH_3$ (abbreviated MTBE—italicized portions of names are not counted when ranking the groups alphabetically—so butyl comes before methyl in the common name), is used as an additive for gasoline. MTBE belongs to a group of chemicals known as oxygenates due to their capacity to increase the oxygen content of gasoline.

## Link to Learning

Want more practice naming ethers? This brief video review summarizes the nomenclature for ethers.

## Chemistry in Everyday Life: Carbohydrates and Diabetes

Carbohydrates are large biomolecules made up of carbon, hydrogen, and oxygen. The dietary forms of carbohydrates are foods rich in these types of molecules, like pastas, bread, and candy. The name "carbohydrate" comes from the formula of the molecules, which can be described by the general formula $C_m(H_2O)_n$, which shows that they are in a sense "carbon and water" or "hydrates of carbon." In many cases, *m* and *n* have the same value, but they can be different. The smaller carbohydrates are generally referred to as "sugars," the biochemical term for this group of molecules is "saccharide" from the Greek word for sugar (Figure 20.12). Depending on the number of sugar units joined together, they may be classified as monosaccharides (one sugar unit), disaccharides (two sugar units), oligosaccharides (a few sugars), or polysaccharides (the polymeric version of sugars—polymers were described in the feature box earlier in this chapter on recycling plastics). The scientific names of sugars can be recognized by the suffix *-ose* at the end of the name (for instance, fruit sugar is a monosaccharide called "fructose" and milk sugar is a disaccharide called lactose composed of two monosaccharides, glucose and galactose, connected together). Sugars contain some of the functional groups we have discussed: Note the alcohol groups present in the structures and how monosaccharide units are linked to form a disaccharide by formation of an ether.

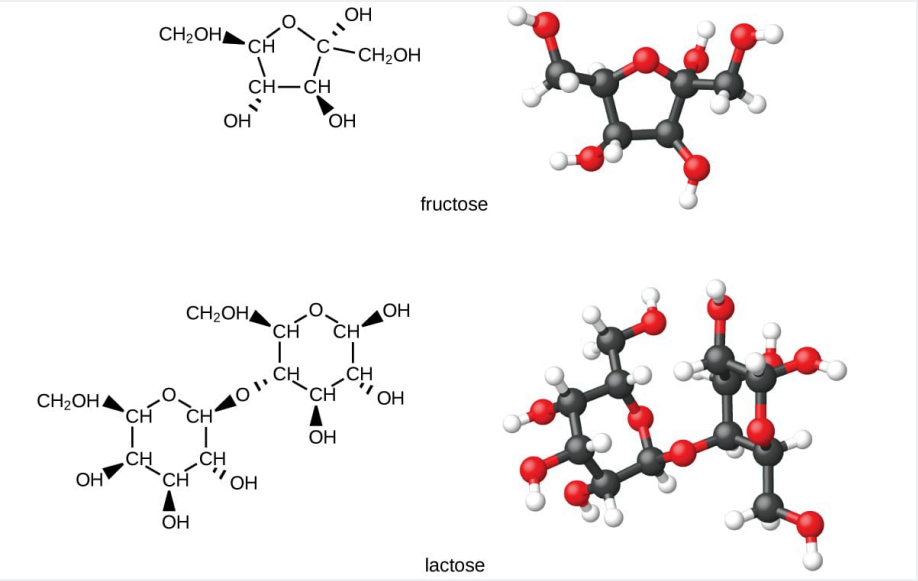



Figure 20.12 The illustrations show the molecular structures of fructose, a five-carbon monosaccharide, and of lactose, a disaccharide composed of two isomeric, six-carbon sugars.

Organisms use carbohydrates for a variety of functions. Carbohydrates can store energy, such as the polysaccharides glycogen in animals or starch in plants. They also provide structural support, such as the polysaccharide cellulose in plants and the modified polysaccharide chitin in fungi and animals. The sugars ribose and deoxyribose are components of the backbones of RNA and DNA, respectively. Other sugars play key roles in the function of the immune system, in cell-cell recognition, and in many other biological roles.

Diabetes is a group of metabolic diseases in which a person has a high sugar concentration in their blood (Figure 10.18.3.1 ). Diabetes may be caused by insufficient insulin production by the pancreas or by the body's cells not responding properly to the insulin that is produced. In a healthy person, insulin is produced when it is needed and functions to transport glucose from the blood into the cells where it can be used for energy. The long-term complications of diabetes can include loss of eyesight, heart disease, and kidney failure.

In 2013, it was estimated that approximately 3.3% of the world's population (~380 million people) suffered from diabetes, resulting in over a million deaths annually. Prevention involves eating a healthy diet, getting plenty of exercise, and maintaining a normal body weight. Treatment involves all of these lifestyle practices and may require injections of insulin.

Even after treatment protocols were introduced, the need to continually monitor their glucose levels posed a challenge for people with diabetes. The first tests required a doctor or lab, and therefore limited access and frequency. Eventually, researchers developed small tablets that would react to the presence of glucose in urine, but these still required a relatively complex process. Chemist Helen Free, who was working on improvements to the tablets, conceived a simpler device: a small test strip. With her husband and research partner, Alfred Free, she produced the first such product for measuring glucose; soon after, she expanded the technology to provide test strips for other compounds and conditions. While very recent advances (such as breath tests, discussed earlier in the text) have shown promise in replacing test strips, they have been widely used for decades and remain a primary method today.



Figure 10.18.3.1: Diabetes is a disease characterized by high concentrations of glucose in the blood. Treating diabetes involves making lifestyle changes, monitoring blood-sugar levels, and sometimes insulin injections. (credit: "Blausen Medical Communications"/Wikimedia Commons)

## Footnotes

- 5The IUPAC adopted new nomenclature guidelines in 2013 that require this number to be placed as an "infix" rather than a prefix. For example, the new name for 2-propanol would be propan-2-ol. Widespread adoption of this new nomenclature will take some time, and students are encouraged to be familiar with both the old and new naming protocols.

### Answers

**Exercise** 10.18.3.1

2-methyl-2-pentanol

**Exercise** 10.18.3.1

IUPAC: 2-methoxypropane; common: isopropylmethyl ether

This page titled 10.18.3: Alcohols and Ethers is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# 10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters

**Learning Objectives**

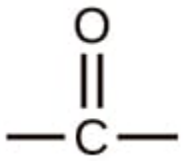
By the end of this section, you will be able to:

- Describe the structure and properties of aldehydes, ketones, carboxylic acids and esters

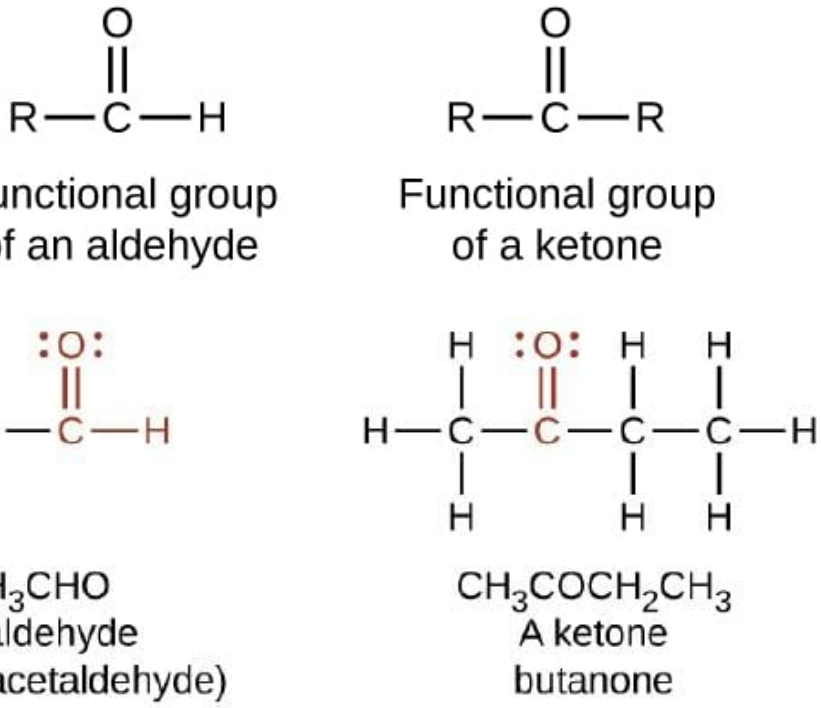
Another class of organic molecules contains a carbon atom connected to an oxygen atom by a double bond, commonly called a carbonyl group. The trigonal planar carbon in the carbonyl group can attach to two other substituents leading to several subfamilies (aldehydes, ketones, carboxylic acids and esters) described in this section.

## Aldehydes and Ketones

Both aldehydes and ketones contain a carbonyl group, a functional group with a carbon-oxygen double bond. The names for aldehyde and ketone compounds are derived using similar nomenclature rules as for alkanes and alcohols, and include the class-identifying suffixes *-al* and *-one*, respectively:



In an aldehyde, the carbonyl group is bonded to at least one hydrogen atom. In a ketone, the carbonyl group is bonded to two carbon atoms:




As text, an aldehyde group is represented as –CHO; a ketone is represented as –C(O)– or –CO–.

In both aldehydes and ketones, the geometry around the carbon atom in the carbonyl group is trigonal planar; the carbon atom exhibits $sp^2$ hybridization. Two of the $sp^2$ orbitals on the carbon atom in the carbonyl group are used to form σ bonds to the other carbon or hydrogen atoms in a molecule. The remaining $sp^2$ hybrid orbital forms a σ bond to the oxygen atom. The unhybridized *p* orbital on the carbon atom in the carbonyl group overlaps a *p* orbital on the oxygen atom to form the π bond in the double bond.

Like the

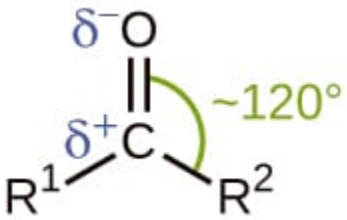
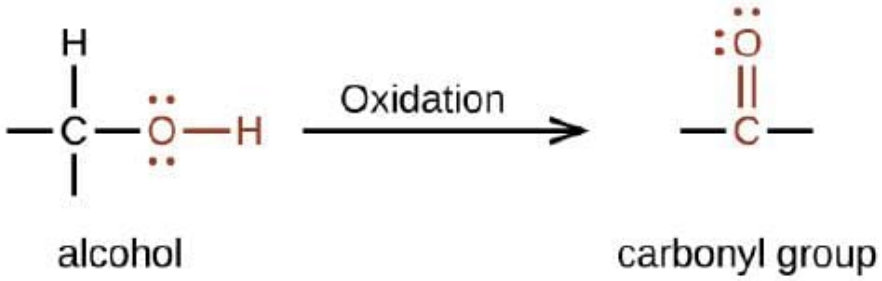



Figure 10.18.4.1: The carbonyl group is polar, and the geometry of the bonds around the central carbon is trigonal planar.

The importance of molecular structure in the reactivity of organic compounds is illustrated by the reactions that produce aldehydes and ketones. We can prepare a carbonyl group by oxidation of an alcohol—for organic molecules, oxidation of a carbon atom is said to occur when a carbon-hydrogen bond is replaced by a carbon-oxygen bond. The reverse reaction—replacing a carbon-oxygen bond by a carbon-hydrogen bond—is a reduction of that carbon atom. Recall that oxygen is generally assigned a –2 oxidation number unless it is elemental or attached to a fluorine. Hydrogen is generally assigned an oxidation number of +1 unless it is attached to a metal. Since carbon does not have a specific rule, its oxidation number is determined algebraically by factoring the atoms it is attached to and the overall charge of the molecule or ion. In general, a carbon atom attached to an oxygen atom will have a more positive oxidation number and a carbon atom attached to a hydrogen atom will have a more negative oxidation number. This should fit nicely with your understanding of the polarity of C–O and C–H bonds. The other reagents and possible products of these reactions are beyond the scope of this chapter, so we will focus only on the changes to the carbon atoms:




## Example 10.18.4.1: Oxidation and Reduction in Organic Chemistry

Methane represents the completely reduced form of an organic molecule that contains one carbon atom. Sequentially replacing each of the carbon-hydrogen bonds with a carbon-oxygen bond would lead to an alcohol, then an aldehyde, then a carboxylic acid (discussed later), and, finally, carbon dioxide:

What are the oxidation numbers for the carbon atoms in the molecules shown here?

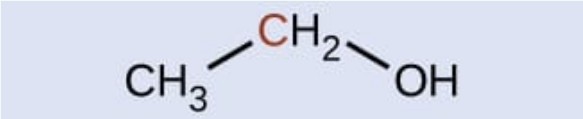
**Solution**

In this example, we can calculate the oxidation number (review the chapter on oxidation-reduction reactions if necessary) for the carbon atom in each case (note how this would become difficult for larger molecules with additional carbon atoms and hydrogen atoms, which is why organic chemists use the definition dealing with replacing C–H bonds with C–O bonds described). For $CH_4$, the carbon atom carries a –4 oxidation number (the hydrogen atoms are assigned oxidation numbers of +1 and the carbon atom balances that by having an oxidation number of –4). For the alcohol (in this case, methanol), the carbon atom has an oxidation number of –2 (the oxygen atom is assigned –2, the four hydrogen atoms each are assigned +1, and the carbon atom balances the sum by having an oxidation number of –2; note that compared to the carbon atom in $CH_4$, this carbon atom has lost two electrons so it was oxidized); for the aldehyde, the carbon atom's oxidation number is 0 (–2 for

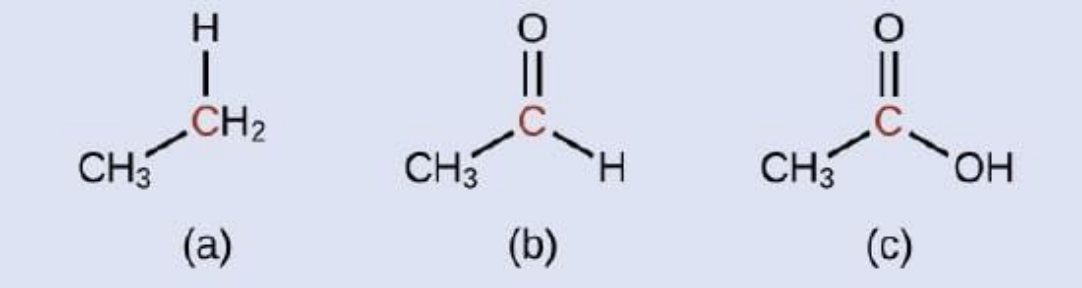
the oxygen atom and +1 for each hydrogen atom already balances to 0, so the oxidation number for the carbon atom is 0); for the carboxylic acid, the carbon atom's oxidation number is +2 (two oxygen atoms each at –2 and two hydrogen atoms at +1); and for carbon dioxide, the carbon atom's oxidation number is +4 (here, the carbon atom needs to balance the –4 sum from the two oxygen atoms).

**Exercise 10.18.4.1**

Indicate whether the marked carbon atoms in the three molecules here are oxidized or reduced relative to the marked carbon atom in ethanol:

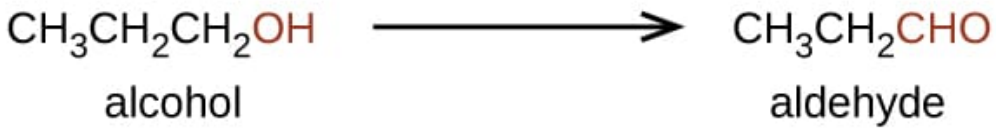



There is no need to calculate oxidation states in this case; instead, just compare the types of atoms bonded to the marked carbon atoms:




*Answer available at the end of this section.*

Aldehydes are commonly prepared by the oxidation of alcohols whose –OH functional group is located on the carbon atom at the end of the chain of carbon atoms in the alcohol:




Alcohols that have their –OH groups in the middle of the chain are necessary to synthesize a ketone, which requires the carbonyl group to be bonded to two other carbon atoms:




An alcohol with its –OH group bonded to a carbon atom that is bonded to no or one other carbon atom will form an aldehyde. An alcohol with its –OH group attached to two other carbon atoms will form a ketone. If three carbons are attached to the carbon bonded to the –OH, the molecule will not have a C–H bond to be replaced, so it will not be susceptible to oxidation.

Formaldehyde, an aldehyde with the formula HCHO, is a colorless gas with a pungent and irritating odor. It is sold in an aqueous solution called formalin, which contains about 37% formaldehyde by weight. Formaldehyde causes coagulation of proteins, so

it kills bacteria (and any other living organism) and stops many of the biological processes that cause tissue to decay. Thus, formaldehyde is used for preserving tissue specimens and embalming bodies. It is also used to sterilize soil or other materials. Formaldehyde is used in the manufacture of Bakelite, a hard plastic having high chemical and electrical resistance.

Dimethyl ketone, $CH_3COCH_3$, commonly called acetone, is the simplest ketone. It is made commercially by fermenting corn or molasses, or by oxidation of 2-propanol. Acetone is a colorless liquid. Among its many uses are as a solvent for lacquer (including fingernail polish), cellulose acetate, cellulose nitrate, acetylene, plastics, and varnishes; as a paint and varnish remover; and as a solvent in the manufacture of pharmaceuticals and chemicals.

## Carboxylic Acids and Esters

The odor of vinegar is caused by the presence of acetic acid, a carboxylic acid, in the vinegar. The odor of ripe bananas and many other fruits is due to the presence of esters, compounds that can be prepared by the reaction of a carboxylic acid with an alcohol. Because esters do not have hydrogen bonds between molecules, they have lower vapor pressures than the alcohols and carboxylic acids from which they are derived (see Figure 10.18.4.2).

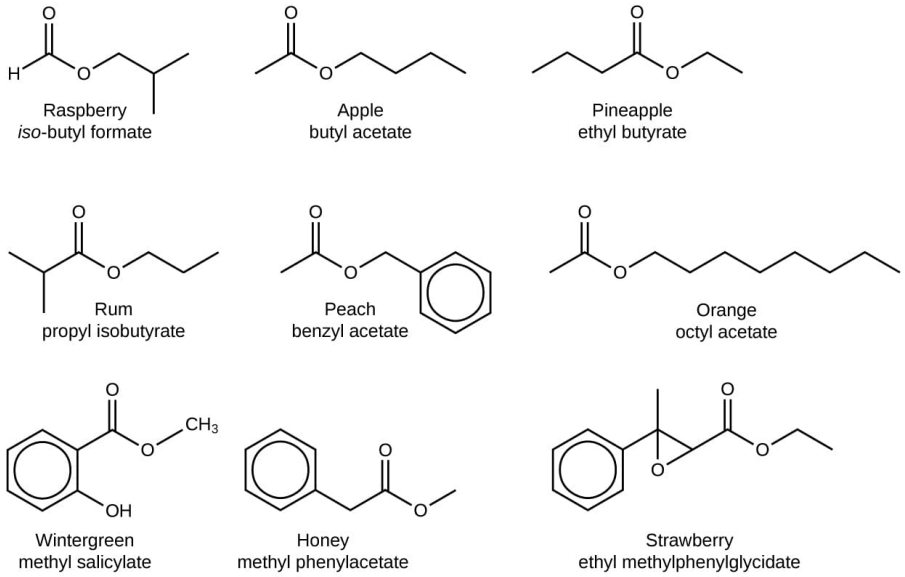
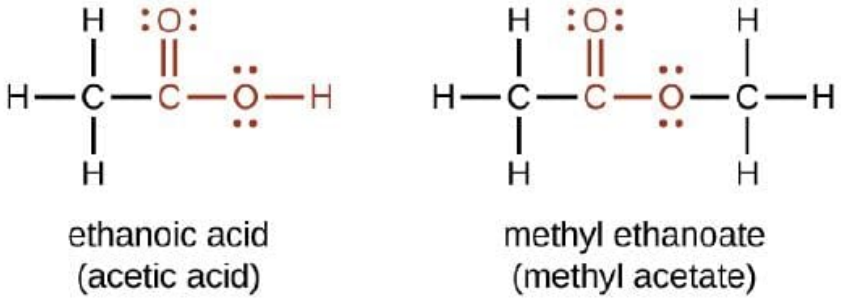


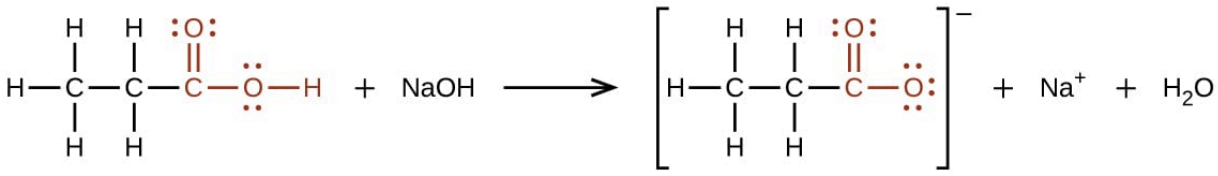

Figure 10.18.4.2: Esters are responsible for the odors associated with various plants and their fruits.

Both carboxylic acids and esters contain a carbonyl group with a second oxygen atom bonded to the carbon atom in the carbonyl group by a single bond. In a carboxylic acid, the second oxygen atom also bonds to a hydrogen atom. In an ester, the second oxygen atom bonds to another carbon atom. The names for carboxylic acids and esters include prefixes that denote the lengths of the carbon chains in the molecules and are derived following nomenclature rules similar to those for inorganic acids and salts (see these examples):



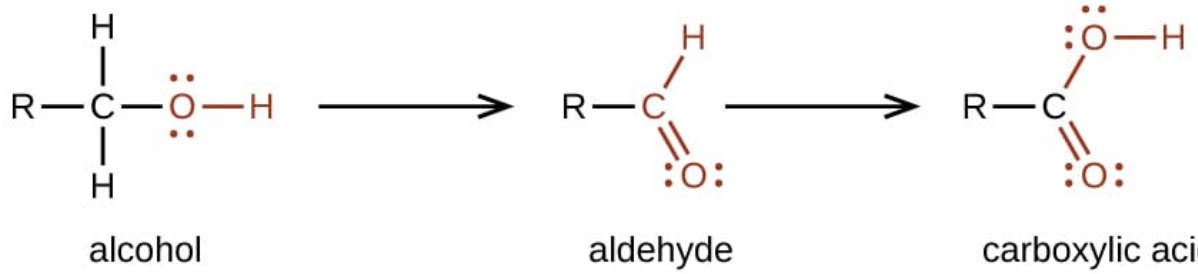

The functional groups for an acid and for an ester are shown in red in these formulas.

The hydrogen atom in the functional group of a carboxylic acid will react with a base to form an ionic salt:

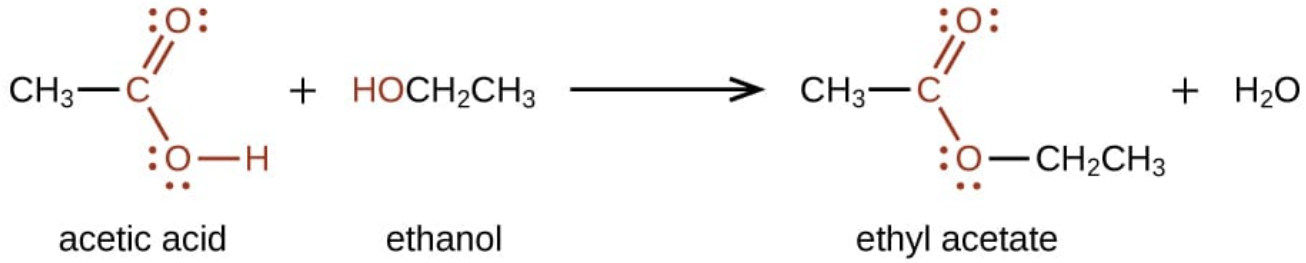



Carboxylic acids are weak acids (see the chapter on acids and bases), meaning they are not 100% ionized in water. Generally only about 1% of the molecules of a carboxylic acid dissolved in water are ionized at any given time. The remaining molecules are undissociated in solution.

We prepare carboxylic acids by the oxidation of aldehydes or alcohols whose –OH functional group is located on the carbon atom at the end of the chain of carbon atoms in the alcohol:




Esters are produced by the reaction of acids with alcohols. For example, the ester ethyl acetate, $CH_3CO_2CH_2CH_3$, is formed when acetic acid reacts with ethanol:




The simplest carboxylic acid is formic acid, $HCO_2H$, known since 1670. Its name comes from the Latin word *formicus*, which means “ant”; it was first isolated by the distillation of red ants. It is partially responsible for the pain and irritation of ant and wasp stings, and is responsible for a characteristic odor of ants that can be sometimes detected in their nests.

Acetic acid, $CH_3CO_2H$, constitutes 3–6% vinegar. Cider vinegar is produced by allowing apple juice to ferment without oxygen present. Yeast cells present in the juice carry out the fermentation reactions. The fermentation reactions change the sugar present in the juice to ethanol, then to acetic acid. Pure acetic acid has a penetrating odor and produces painful burns. It is an excellent solvent for many organic and some inorganic compounds, and it is essential in the production of cellulose acetate, a component of many synthetic fibers such as rayon.

The distinctive and attractive odors and flavors of many flowers, perfumes, and ripe fruits are due to the presence of one or more esters (Figure 10.18.4.3). Among the most important of the natural esters are fats (such as lard, tallow, and butter) and oils (such as linseed, cottonseed, and olive oils), which are esters of the trihydroxyl alcohol glycerine, $C_3H_5(OH)_3$, with large carboxylic acids, such as palmitic acid, $CH_3(CH_2)_{14}CO_2H$, stearic acid, $CH_3(CH_2)_{16}CO_2H$, and oleic acid, Oleic acid is an unsaturated acid; it contains a double bond. Palmitic and stearic acids are saturated acids that contain no double or triple bonds.



Figure 10.18.4.3: Over 350 different volatile molecules (many members of the ester family) have been identified in strawberries. (credit: Rebecca Siegel)

## Answers

**Exercise** 10.18.4.1

(a) reduced (bond to oxygen atom replaced by bond to hydrogen atom); (b) oxidized (one bond to hydrogen atom replaced by one bond to oxygen atom); (c) oxidized (2 bonds to hydrogen atoms have been replaced by bonds to an oxygen atom)

This page titled 10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

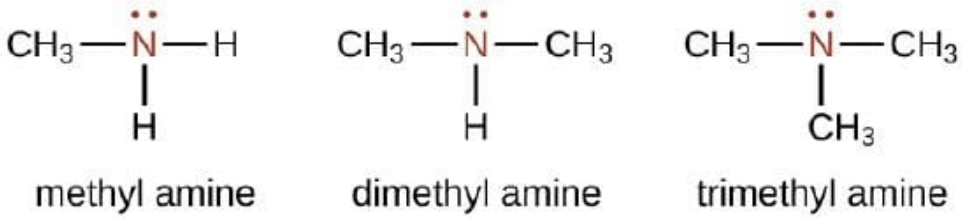
# 10.18.5: Amines and Amides

## Learning Objectives

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

- Describe the structure and properties of an amine
- Describe the structure and properties of an amide

Amines are molecules that contain carbon-nitrogen bonds. The nitrogen atom in an amine has a lone pair of electrons and three bonds to other atoms, either carbon or hydrogen. Various nomenclatures are used to derive names for amines, but all involve the class-identifying suffix *–ine* as illustrated here for a few simple examples:




In some amines, the nitrogen atom replaces a carbon atom in an aromatic hydrocarbon. Pyridine (Figure 10.18.5.1) is one such heterocyclic amine. A heterocyclic compound contains atoms of two or more different elements in its ring structure.

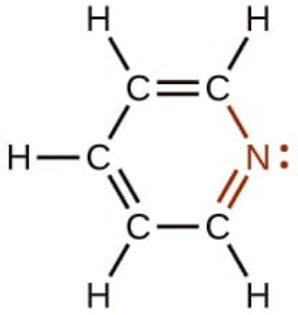


Figure 10.18.5.1: The illustration shows one of the resonance structures of pyridine.

## How Sciences Interconnect: DNA in Forensics and Paternity

The genetic material for all living things is a polymer of four different molecules, which are themselves a combination of three subunits. The genetic information, the code for developing an organism, is contained in the specific sequence of the four molecules, similar to the way the letters of the alphabet can be sequenced to form words that convey information. The information in a DNA sequence is used to form two other types of polymers, one of which are proteins. The proteins interact to form a specific type of organism with individual characteristics.

A genetic molecule is called DNA, which stands for deoxyribonucleic acid. The four molecules that make up DNA are called nucleotides. Each nucleotide consists of a single- or double-ringed molecule containing nitrogen, carbon, oxygen, and hydrogen called a nitrogenous base. Each base is bonded to a five-carbon sugar called deoxyribose. The sugar is in turn bonded to a phosphate group PO43–)Figure 10.18.5.2).

It probably makes sense that the sequence of nucleotides in the DNA of a cat differs from those of a dog. But it is also true that the sequences of the DNA in the cells of two individual pugs differ. Likewise, the sequences of DNA in you and a sibling differ (unless your sibling is an identical twin), as do those between you and an unrelated individual. However, the DNA

sequences of two related individuals are more similar than the sequences of two unrelated individuals, and these similarities in sequence can be observed in various ways. This is the principle behind DNA fingerprinting, which is a method used to determine whether two DNA samples came from related (or the same) individuals or unrelated individuals.

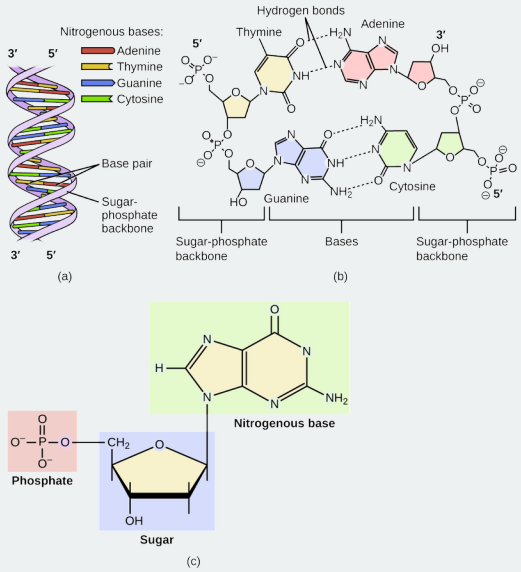


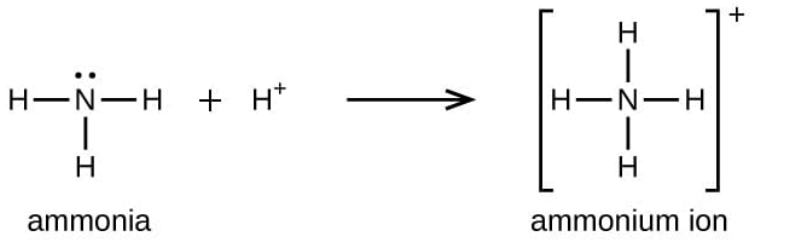

Figure 10.18.5.2: DNA is an organic molecule and the genetic material for all living organisms. (a) DNA is a double helix consisting of two single DNA strands hydrogen bonded together at each nitrogenous base. (b) This detail shows the hydrogen bonding (dotted lines) between nitrogenous bases on each DNA strand and the way in which each nucleotide is joined to the next, forming a backbone of sugars and phosphate groups along each strand. (c) This detail shows the structure of one of the four nucleotides that makes up the DNA polymer. Each nucleotide consists of a nitrogenous base (a double-ring molecule, in this case), a five-carbon sugar (deoxyribose), and a phosphate group.

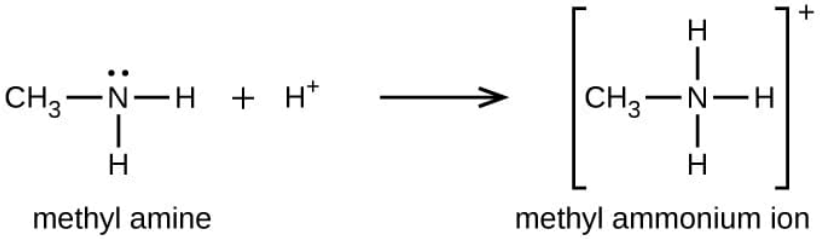
Using similarities in sequences, technicians can determine whether a man is the father of a child (the identity of the mother is rarely in doubt, except in the case of an adopted child and a potential birth mother). Likewise, forensic geneticists can determine whether a crime scene sample of human tissue, such as blood or skin cells, contains DNA that matches exactly the DNA of a suspect.

## Link to Learning

Watch this video animation of how DNA is packaged for a visual lesson in its structure.

Like ammonia, amines are weak bases due to the lone pair of electrons on their nitrogen atoms:




The basicity of an amine's nitrogen atom plays an important role in much of the compound's chemistry. Amine functional groups are found in a wide variety of compounds, including natural and synthetic dyes, polymers, vitamins, and medications such as penicillin and codeine. They are also found in many molecules essential to life, such as amino acids, hormones, neurotransmitters, and DNA.

## How Sciences Interconnect: Addictive Alkaloids

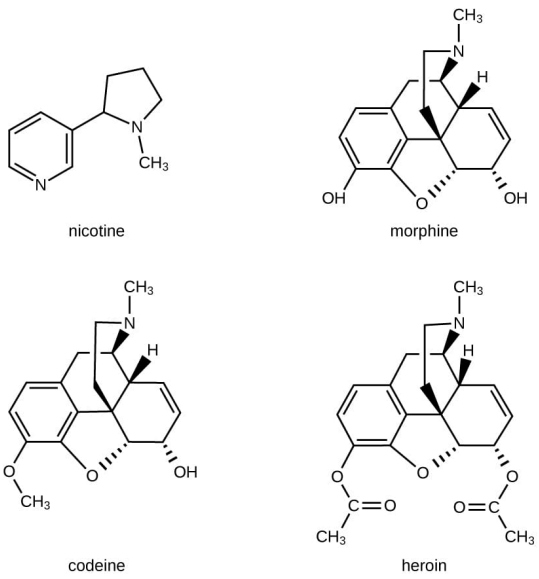
Since ancient times, plants have been used for medicinal purposes. One class of substances, called *alkaloids*, found in many of these plants has been isolated and found to contain cyclic molecules with an amine functional group. These amines are bases. They can react with $H_3O^+$ in a dilute acid to form an ammonium salt, and this property is used to extract them from the plant:

R3N+H3O++Cl–⟡[R3NH+]Cl–+H2OR3N+H3O++Cl–⟡[R3NH+]Cl–+H2O

The name alkaloid means "like an alkali." Thus, an alkaloid reacts with acid. The free compound can be recovered after extraction by reaction with a base:

R3NH+]Cl–+OH–⟡R3N+H2O+Cl–[R3NH+]Cl–+OH–⟡R3N+H2O+Cl–

The structures of many naturally occurring alkaloids have profound physiological and psychotropic effects in humans. Examples of these drugs include nicotine, morphine, codeine, and heroin. The plant produces these substances, collectively called secondary plant compounds, as chemical defenses against the numerous pests that attempt to feed on the plant:

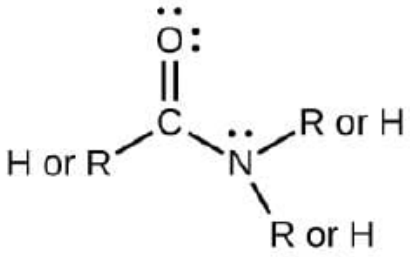


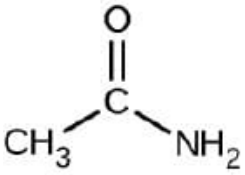

In these diagrams, as is common in representing structures of large organic compounds, carbon atoms in the rings and the hydrogen atoms bonded to them have been omitted for clarity. The solid wedges indicate bonds that extend out of the page. The dashed wedges indicate bonds that extend into the page. Notice that small changes to a part of the molecule change the properties of morphine, codeine, and heroin. Morphine, a strong narcotic used to relieve pain, contains two hydroxyl functional groups, located at the bottom of the molecule in this structural formula. Changing one of these hydroxyl groups to a methyl ether group forms codeine, a less potent drug used as a local anesthetic. If both hydroxyl groups are converted to esters of acetic acid, the powerfully addictive drug heroin results (Figure 10.18.5.3).



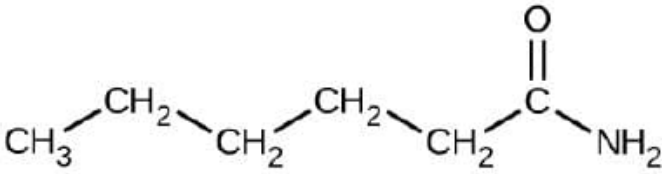
Figure 10.18.5.3: Poppies can be used in the production of opium, a plant latex that contains morphine from which other opiates, such as heroin, can be synthesized. (credit: Karen Roe)

Amides are molecules that contain nitrogen atoms connected to the carbon atom of a carbonyl group. Like amines, various nomenclature rules may be used to name amides, but all include use of the class-specific suffix *-amide*:



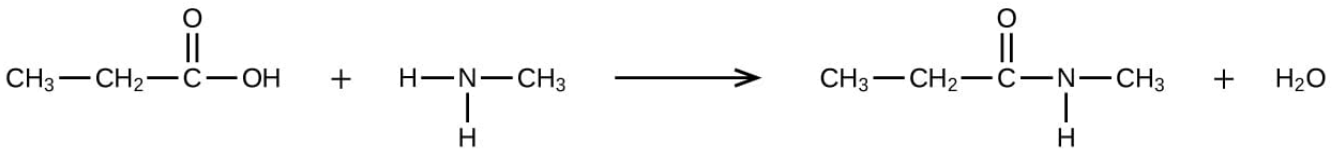




acetamide




hexanamide

Amides can be produced when carboxylic acids react with amines or ammonia in a process called amidation. A water molecule is eliminated from the reaction, and the amide is formed from the remaining pieces of the carboxylic acid and the amine (note the similarity to formation of an ester from a carboxylic acid and an alcohol discussed in the previous section):




carboxylic acid amine amide water

The reaction between amines and carboxylic acids to form amides is biologically important. It is through this reaction that amino acids (molecules containing both amine and carboxylic acid substituents) link together in a polymer to form proteins.

## How Sciences Interconnect: Proteins and Enzymes

Proteins are large biological molecules made up of long chains of smaller molecules called amino acids. Organisms rely on proteins for a variety of functions—proteins transport molecules across cell membranes, replicate DNA, and catalyze metabolic reactions, to name only a few of their functions. The properties of proteins are functions of the combination of amino acids that compose them and can vary greatly. Interactions between amino acid sequences in the chains of proteins result in the folding of the chain into specific, three-dimensional structures that determine the protein's activity.

Amino acids are organic molecules that contain an amine functional group ($-NH_2$), a carboxylic acid functional group (–COOH), and a side chain (that is specific to each individual amino acid). Most living things build proteins from the same 20 different amino acids. Amino acids connect by the formation of a peptide bond, which is a covalent bond formed between two amino acids when the carboxylic acid group of one amino acid reacts with the amine group of the other amino acid. The formation of the bond results in the production of a molecule of water (in general, reactions that result in the production of water when two other molecules combine are referred to as condensation reactions). The resulting bond—between the carbonyl group carbon atom and the amine nitrogen atom is called a peptide link or peptide bond. Since each of the original amino acids has an unreacted group (one has an unreacted amine and the other an unreacted carboxylic acid), more peptide bonds can form to other amino acids, extending the structure. (Figure 10.18.5.4) A chain of connected amino acids is called a polypeptide. Proteins contain at least one long polypeptide chain.

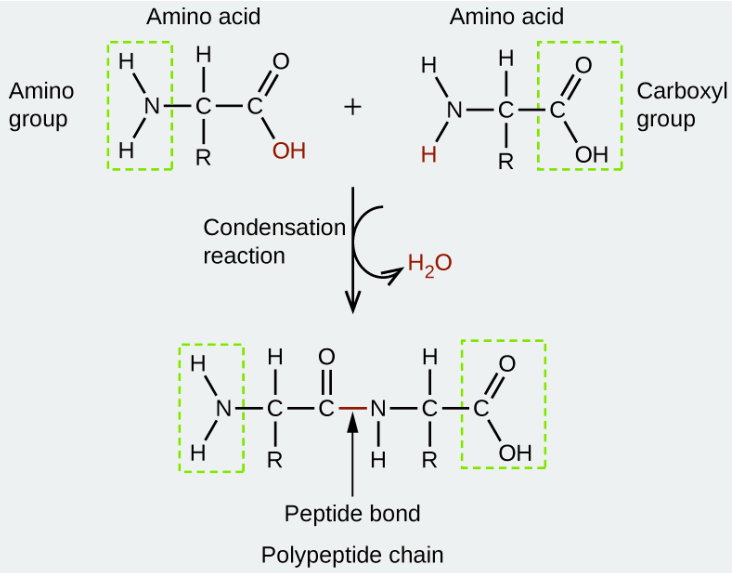



Figure 10.18.5.4: This condensation reaction forms a dipeptide from two amino acids and leads to the formation of water.

Enzymes are large biological molecules, mostly composed of proteins, which are responsible for the thousands of metabolic processes that occur in living organisms. Enzymes are highly specific catalysts; they speed up the rates of certain reactions. Enzymes function by lowering the activation energy of the reaction they are catalyzing, which can dramatically increase the rate of the reaction. Most reactions catalyzed by enzymes have rates that are millions of times faster than the noncatalyzed version. Like all catalysts, enzymes are not consumed during the reactions that they catalyze. Enzymes do differ from other catalysts in how specific they are for their substrates (the molecules that an enzyme will convert into a different product). Each enzyme is only capable of speeding up one or a few very specific reactions or types of reactions. Since the function of enzymes is so specific, the lack or malfunctioning of an enzyme can lead to serious health consequences. One disease that is the result of an enzyme malfunction is phenylketonuria. In this disease, the enzyme that catalyzes the first step in the degradation of the amino acid phenylalanine is not functional (Figure 10.18.5.5). Untreated, this can lead to an accumulation of phenylalanine, which can lead to intellectual disabilities.

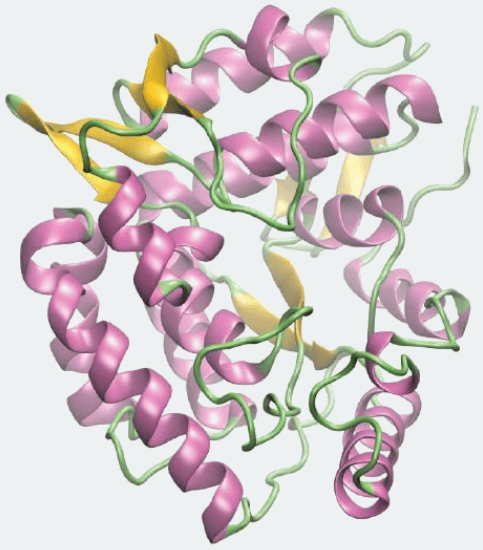


Figure 10.18.5.5: A computer rendering shows the three-dimensional structure of the enzyme phenylalanine hydroxylase. In the disease phenylketonuria, a defect in the shape of phenylalanine hydroxylase causes it to lose its function in breaking down phenylalanine.

## Chemistry in Everyday Life: Kevlar

Kevlar (Figure 10.18.5.6) is a synthetic polymer made from two monomers 1,4-phenylene-diamine and terephthaloyl chloride (Kevlar is a registered trademark of DuPont). The material was developed by Susan Kwolek while she worked to find a replacement for steel in tires. Kwolek's work involved synthesizing polyamides and dissolving them in solvents, then spinning the resulting solution into fibers. One of her solutions proved to be quite different in initial appearance and structure. And once spun, the resulting fibers were particularly strong. From this initial discovery, Kevlar was created. The material has a high tensile strength-to-weight ratio (it is about 5 times stronger than an equal weight of steel), making it useful for many applications from bicycle tires to sails to body armor.

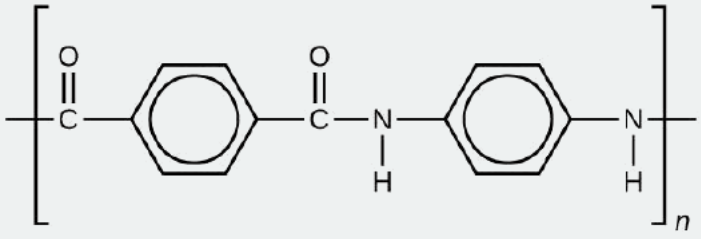



Figure 10.18.5.6: This illustration shows the formula for polymeric Kevlar.

The material owes much of its strength to hydrogen bonds between polymer chains (refer back to the chapter on intermolecular interactions). These bonds form between the carbonyl group oxygen atom (which has a partial negative charge due to oxygen's electronegativity) on one monomer and the partially positively charged hydrogen atom in the N–H bond of an adjacent monomer in the polymer structure (see dashed lines in Figure 10.18.5.7). There is additional strength derived from the interaction between the unhybridized *p* orbitals in the six-membered rings, called aromatic stacking.

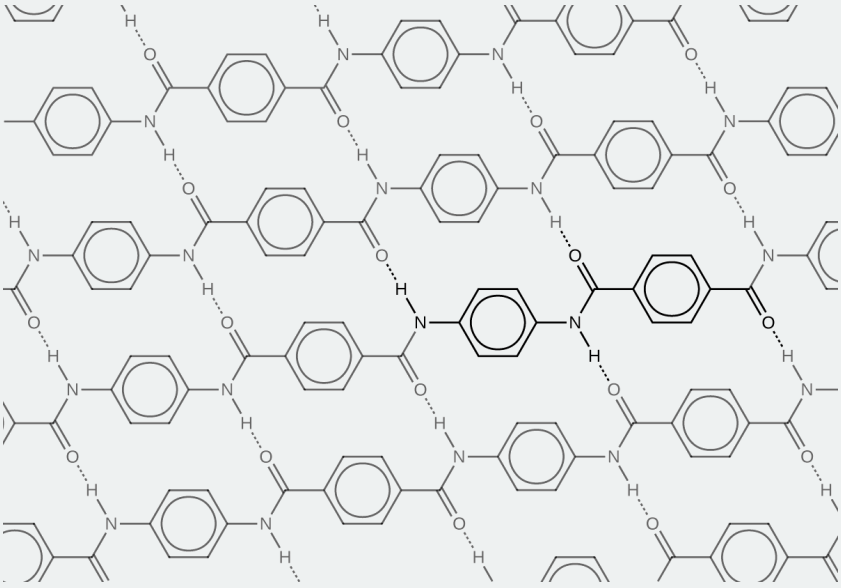


Figure 10.18.5.7: The diagram shows the polymer structure of Kevlar, with hydrogen bonds between polymer chains represented by dotted lines.

Kevlar may be best known as a component of body armor, combat helmets, and face masks. Since the 1980s, the US military has used Kevlar as a component of the PASGT (personal armor system for ground troops) helmet and vest. Kevlar is also used to protect armored fighting vehicles and aircraft carriers. Civilian applications include protective gear for emergency service personnel such as body armor for police officers and heat-resistant clothing for fire fighters. Kevlar based clothing is considerably lighter and thinner than equivalent gear made from other materials (Figure 10.18.5.8). Beyond Kevlar, Susan Kwolek was instrumental in the development of Nomex, a fireproof material, and was also involved in the creation of Lycra. She became just the fourth woman inducted into the National Inventors Hall of Fame, and received a number of other awards for her significant contributions to science and society.



(a) (b) (c)

Figure 10.18.5.8: (a) These soldiers are sorting through pieces of a Kevlar helmet that helped absorb a grenade blast. Kevlar is also used to make (b) canoes and (c) marine mooring lines. (credit a: modification of work by “Cla68”/Wikimedia Commons; credit b: modification of work by “OakleyOriginals”/Flickr; credit c: modification of work by Casey H. Kyhl)

In addition to its better-known uses, Kevlar is also often used in cryogenics for its very low thermal conductivity (along with its high strength). Kevlar maintains its high strength when cooled to the temperature of liquid nitrogen (–196 °C).

The table here summarizes the structures discussed in this chapter:

| Compound Name | Structure of Compound and Functional Group (red) | Example | |
|---|---|---|---|
| | | Formula | Name |
| alkene | C=C | $C_2H_4$ | ethene |
| alkyne | C≡C | $C_2H_2$ | ethyne |
| alcohol | R—O—H | $CH_3CH_2OH$ | ethanol |
| ether | R—O—R′ | $(C_2H_5)_2O$ | diethyl ether |
| aldehyde | R—C(=O)—H | $CH_3CHO$ | ethanal |
| ketone | R—C(=O)—R′ | $CH_3COCH_2CH_3$ | methyl ethyl ketone |
| carboxylic acid | R—C(=O)—O—H | $CH_3COOH$ | acetic acid |
| ester | R—C(=O)—O—R′ | $CH_3CO_2CH_2CH_3$ | ethyl acetate |
| amine | R—N(—H)—H; R—N(—R′)—H; R—N(—R″)—R′ | $C_2H_5NH_2$ | ethylamine |
| amide | R—C(=O)—N(—H)—R′ | $CH_3CONH_2$ | acetamide |

This page titled 10.18.5: Amines and Amides is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

# INDEX


**A**

Arunava Roy
*2: Module 1b-Motion*

**S**

scientific method
*1.4.1: Laws and Theories in Science*


This page titled Index is shared under a Public Domain license and was authored, remixed, and/or curated by Arunava Roy.

# GLOSSARY

## S

**Sample Word 1** | Sample Definition 1

This page titled Glossary is shared under a license and was authored, remixed, and/or curated by Arunava Roy.

# DETAILED LICENSING

## Overview

**Title:** Introduction to Interdisciplinary Physical Science

**Webpages:** 165

**Applicable Restrictions:** Noncommercial

**All licenses found:**

- CC BY 4.0: 68.5% (113 pages)
- CK-12 License: 13.3% (22 pages)
- CC BY-NC-SA 4.0: 12.1% (20 pages)
- Unknown License: 3% (5 pages)
- Public Domain: 2.4% (4 pages)
- Undeclared: 0.6% (1 page)
- Unknown License: 0% (0 page)
- Unknown License: 0% (0 page)
- Unknown License: 0% (0 page)
- Unknown License: 0% (0 page)

## By Page


- Introduction to Interdisciplinary Physical Science - *Unknown License*
  - Front Matter - *Unknown License*
    - TitlePage - *CC BY-NC-SA 4.0*
    - InfoPage - *Public Domain*
    - Table of Contents - *Public Domain*
    - Licensing - *Public Domain*
  - 1: Module 1a-Scientific Method - *CC BY-NC-SA 4.0*
    - 1.1: Prelude to Science and the Realm of Physics, Physical Quantities, and Units - *CC BY 4.0*
    - 1.2: Observing the World - *CC BY-NC-SA 4.0*
    - 1.3: Developing the Scientific Method - *CC BY-NC-SA 4.0*
      - 1.3.1: Careful Observations- Documentation, Measurements, and Patterns - *CC BY-NC-SA 4.0*
      - 1.3.2: Experiments - *CC BY-NC-SA 4.0*
      - 1.3.3: A Scientific Method - *CC BY-NC-SA 4.0*
    - 1.4: Processes of Science - *CC BY-NC-SA 4.0*
      - 1.4.1: Laws and Theories in Science - *CC BY-NC-SA 4.0*
      - 1.4.2: The Scientific Process - *CC BY-NC-SA 4.0*
  - 2: Module 1b-Motion - *CC BY-NC-SA 4.0*
    - 2.1: Physics- An Introduction - *CC BY 4.0*
    - 2.2: Prelude to One-Dimensional Kinematics - *CC BY 4.0*
    - 2.3: Displacement - *CC BY 4.0*
    - 2.4: Vectors, Scalars, and Coordinate Systems - *CC BY 4.0*
    - 2.5: Time, Velocity, and Speed - *CC BY 4.0*
    - 2.6: Acceleration - *CC BY 4.0*
    - 2.7: Falling Objects - *CC BY 4.0*
    - 2.8: Prelude to Two-Dimensional Kinematics - *CC BY 4.0*
    - 2.9: Kinematics in Two Dimensions - An Introduction - *CC BY 4.0*
    - 2.10: Projectile Motion - *CC BY 4.0*
    - 2.11: Prelude to Uniform Circular Motion and Gravitation - *CC BY 4.0*
    - 2.12: Rotation Angle and Angular Velocity - *CC BY 4.0*
    - 2.13: Centripetal Acceleration - *CC BY 4.0*
  - 3: Module 2-Motion and Forces - *CC BY-NC-SA 4.0*
    - 3.1: Prelude to Dynamics- Newton's Laws of Motion - *CC BY 4.0*
    - 3.2: Development of Force Concept - *CC BY 4.0*
    - 3.3: Newton's First Law of Motion - Inertia - *CC BY 4.0*
    - 3.4: Newton's Second Law of Motion- Concept of a System - *CC BY 4.0*
    - 3.5: Newton's Third Law of Motion- Symmetry in Forces - *CC BY 4.0*
    - 3.6: Summary of Newton's Laws - *CC BY 4.0*
    - 3.7: Newton's Universal Law of Gravitation - *CC BY 4.0*
  - 4: Module 3-Energy and Momentum - *CC BY-NC-SA 4.0*

- 4.1: Prelude to Work, Energy, and Energy Resources - *CC BY 4.0*
- 4.2: Work- The Scientific Definition - *CC BY 4.0*
- 4.3: Kinetic Energy and the Work-Energy Theorem - *CC BY 4.0*
- 4.4: Gravitational Potential Energy - *CC BY 4.0*
- 4.5: Conservative Forces and Potential Energy - *CC BY 4.0*
- 4.6: Conservation of Energy - *CC BY 4.0*
- 4.7: Power - *CC BY 4.0*
- 4.8: Work, Energy, and Power in Humans - *CC BY 4.0*
- 4.9: World Energy Use - *CC BY 4.0*
- 4.10: Prelude to Linear Momentum and Collisions - *CC BY 4.0*
- 4.11: Linear Momentum and Force - *CC BY 4.0*
- 4.12: Conservation of Momentum - *CC BY 4.0*
- 4.13: Energy and Momentum - *CC BY-NC-SA 4.0*

◦ 5: Module 4-Thermodynamics and Waves - *CC BY-NC-SA 4.0*
- 5.1: Prelude to Temperature, Kinetic Theory, and the Gas Laws - *CC BY 4.0*
- 5.2: Temperature - *CC BY 4.0*
- 5.3: Kinetic Theory- Atomic and Molecular Explanation of Pressure and Temperature - *CC BY 4.0*
- 5.4: Phase Changes - *CC BY 4.0*
- 5.5: Prelude to Heat and Heat Transfer Methods - *CC BY 4.0*
- 5.6: Heat - *CC BY 4.0*
- 5.7: Temperature Change and Heat Capacity - *CC BY 4.0*
- 5.8: Phase Change and Latent Heat - *CC BY 4.0*
- 5.9: Heat Transfer Methods - *CC BY 4.0*
- 5.10: Conduction - *CC BY 4.0*
- 5.11: Convection - *CC BY 4.0*
- 5.12: Radiation - *CC BY 4.0*
- 5.13: Prelude to Thermodynamics - *CC BY 4.0*
- 5.14: The First Law of Thermodynamics - *CC BY 4.0*
- 5.15: Introduction to the Second Law of Thermodynamics - Heat Engines and their Efficiency - *CC BY 4.0*
- 5.16: Waves - *CC BY 4.0*
- 5.17: Superposition and Interference - *CC BY 4.0*
- 5.18: Prelude to the Physics of Hearing - *CC BY 4.0*
- 5.19: Sound - *CC BY 4.0*
- 5.20: Speed of Sound, Frequency, and Wavelength - *CC BY 4.0*
- 5.21: Doppler Effect and Sonic Booms - *CC BY 4.0*
- 5.22: Prelude to Electromagnetic Waves - *CC BY 4.0*
- 5.23: The Electromagnetic Spectrum - *CC BY 4.0*

◦ 6: Module 5- Optics - *CC BY-NC-SA 4.0*
- 6.1: Prelude to Geometric Optics - *CC BY 4.0*
- 6.2: The Ray Aspect of Light - *CC BY 4.0*
- 6.3: The Law of Reflection - *CC BY 4.0*
- 6.4: The Law of Refraction - *CC BY 4.0*
- 6.5: Dispersion - Rainbows and Prisms - *CC BY 4.0*
- 6.6: Image Formation by Lenses - *CC BY 4.0*
- 6.7: Image Formation by Mirrors - *CC BY 4.0*

◦ 7: Module 6a- Electricity, Magnetism - *CC BY-NC-SA 4.0*
- 7.1: Prelude to Electric Charge and Electric Field - *CC BY 4.0*
- 7.2: Static Electricity and Charge - Conservation of Charge - *CC BY 4.0*
- 7.3: Conductors and Insulators - *CC BY 4.0*
- 7.4: Coulomb's Law - *CC BY 4.0*
- 7.5: Electric Field- Concept of a Field Revisited - *CC BY 4.0*
- 7.06: Introduction to Electric Potential and Electric Energy - *CC BY 4.0*
- 7.7: Electric Potential Energy- Potential Difference - *CC BY 4.0*
- 7.8: Prelude to Electric Current, Resistance, and Ohm's Law - *CC BY 4.0*
- 7.9: Current - *CC BY 4.0*
- 7.10: Ohm's Law - Resistance and Simple Circuits - *CC BY 4.0*
- 7.11: Resistance and Resistivity - *CC BY 4.0*
- 7.12: 20.4 Electric Power and Energy - *CC BY 4.0*
- 7.13: Alternating Current versus Direct Current - *CC BY 4.0*
- 7.14: Electric Hazards and the Human Body - *CC BY 4.0*
- 7.15: Prelude to Circuits and DC Instruments - *CC BY 4.0*
- 7.16: Resistors in Series and Parallel - *CC BY 4.0*
- 7.17: Prelude to Magnetism - *CC BY 4.0*
- 7.18: Magnets - *CC BY 4.0*
- 7.19: Ferromagnets and Electromagnets - *CC BY 4.0*
- 7.20: Magnetic Fields and Magnetic Field Lines - *CC BY 4.0*
- 7.21: Prelude to Electromagnetic Induction, AC Circuits and Electrical Technologies - *CC BY 4.0*

◦ 8: Module 6b-Introduction to Chemistry - *CC BY-NC-SA 4.0*
- 8.1: Chemistry in Context - *CC BY 4.0*
- 8.02: Phases and Classification of Matter - *CC BY 4.0*
- 8.03: Physical and Chemical Properties - *CC BY 4.0*
- 8.4: Prelude to Atoms - *CC BY 4.0*
- 8.5: Early Ideas in Atomic Theory - *CC BY 4.0*
- 8.6: Evolution of Atomic Theory - *CC BY 4.0*
- 8.7: Atomic Structure and Symbolism - *CC BY 4.0*

- 8.8: Chemical Formulas - *CC BY 4.0*
- 8.9: The Periodic Table - *CC BY 4.0*
- 8.10: Organizing the Elements - *CK-12 License*
  - 8.10.1: Mendeleev's Periodic Table - *CK-12 License*
  - 8.10.2: Periodic Law - *CK-12 License*
  - 8.10.3: Modern Periodic Table- Periods and Groups - *Unknown License*
- 8.11: Chemical Nomenclature - *CC BY 4.0*

◦ 9: Module 7-Atomic Structures and Periodicity - *CC BY-NC-SA 4.0*
- 9.1: Introduction to Atoms and Elements - *CK-12 License*
- 9.2: Modern Understandings of Elements - *CK-12 License*
- 9.3: The Development of Atomic Theory - *CK-12 License*
  - 9.3.1: Law of Conservation of Mass - *CK-12 License*
  - 9.3.2: Law of Definite Proportions - *CK-12 License*
  - 9.3.3: Law of Multiple Proportions - *CK-12 License*
  - 9.3.4: Dalton's Atomic Theory - *CK-12 License*
- 9.4: Prelude to Chemical Bonding and Molecular Geometry - *CC BY 4.0*
- 9.5: Ionic Bonding - *CK-12 License*
- 9.6: Ionic Bonding (1) - *CC BY 4.0*
- 9.7: Covalent Bonds and Molecules - *CK-12 License*
- 9.8: Covalent Bonding - *CC BY 4.0*
- 9.9: Prelude to Liquids and Solids - *CC BY 4.0*
- 9.10: Intermolecular Forces - *CK-12 License*
- 9.11: Intermolecular Forces (1) - *CC BY 4.0*

◦ 10: Module 8- Chemical Reactions and Organic Chemistry - *CC BY-NC-SA 4.0*
- 10.1: Introduction to Chemical Changes - *CK-12 License*
- 10.2: Aspects of a Chemical Change - *CK-12 License*
- 10.3: Chemical Equations- Representing Chemical Changes - *CK-12 License*
- 10.4: Formula Mass and the Mole Concept - *CC BY 4.0*
- 10.5: Balancing Equations - *CK-12 License*
- 10.6: Energy and Chemical Reactions - *CK-12 License*
- 10.7: Classifying Reactions - *CK-12 License*
- 10.8: Acid Base Reactions - *CK-12 License*
- 10.9: Polymerization Reactions - *CK-12 License*
- 10.10: Determining Empirical and Molecular Formulas - *CC BY 4.0*
- 10.11: Prelude to Stoichiometry - *CC BY 4.0*
- 10.12: Balancing Equations (1) - *CK-12 License*
- 10.13: Writing and Balancing Chemical Equations - *CC BY 4.0*
- 10.14: Prelude to Kinetics - *CC BY 4.0*
- 10.15: Chemical Reaction Rates - *CC BY 4.0*
- 10.16: Brønsted-Lowry Acids and Bases - *CC BY 4.0*
- 10.17: pH and pOH - *CC BY 4.0*
- 10.18: Organic Chemistry - *CC BY 4.0*
  - 10.18.1: Introduction - *CC BY 4.0*
  - 10.18.2: Hydrocarbons - *CC BY 4.0*
  - 10.18.3: Alcohols and Ethers - *CC BY 4.0*
  - 10.18.4: Aldehydes, Ketones, Carboxylic Acids, and Esters - *CC BY 4.0*
  - 10.18.5: Amines and Amides - *CC BY 4.0*

◦ Back Matter - *Undeclared*
- Index - *Public Domain*
- Glossary - *Unknown License*
- Detailed Licensing - *Unknown License*


This page titled Introduction to Interdisciplinary Physical Science is shared under a license and was authored, remixed, and/or curated by LibreTexts.