Physics 132: What is an Electron? What is Light?

Physics 132: What is an Electron? What is Light?

ROGER HINRICHS, PAUL PETER URONE, PAUL FLOWERS, EDWARD J. NETH, WILLIAM R. ROBINSON, KLAUS THEOPOLD, RICHARD LANGLEY, JULIANNE ZEDALIS, JOHN EGGEBRECHT, AND E.F. REDISH

KIM DIRKS, MANJULA SHARMA, JENNIFER LOOK, ALLISON SOULT, SIMON BOTT, DONALD CARPENETTI, ANDREW EKLUND, EMAD EL-GIAR, DON FRANTZ, PAUL HOOKER, CAROL MARTINEZ, TROY MILLIKEN, VICKI MORAVEC, GEORGE KAMINSKI, MARK BLASER, JASON POWELL, THOMAS SORENSEN, YAEL AVISSAR, JUNG CHOI, JEAN DESAIX, VLADIMIR JURUKOVSKI, CONNIE RYE, AND ROBERT WISE

UNIVERSITY OF MASSACHUSETTS AMHERST LIBRARIES
AMHERST, MA

Contents

Welcome to Physics 132 - Introduction to the Course

Hello, welcome to Physics 132 at University of Massachusetts, Amherst! This course is where we get to use the ideas from Physics 131 (forces, energy, etc.) to really understand two fundamental objects: electrons and light. These two fundamental objects are all around you. You can see this page due to light. How many electronic devices are you carrying right now? Moreover, understanding these two objects is key for understanding the physical original of biological processes. One cannot hope to understand molecular pathways within cells such as photosynthesis and neural activation without talking about electrons and light, but what *are* electrons? What *is* light? The goal of this course is to help you develop your own understanding of these questions.

How do you define what something is? Especially, as is the case for light and electrons, when the object you are trying to define is subatomic and so very far removed from our everyday experience? These are not scientific questions: we cannot design an experiment to test their answers. Thus, this physics course must, right out of the gate, go *beyond* physics to philosophy. Specifically, we must venture into **metaphysics***:* a branch of philosophy that explores the nature of being, existence, and reality. The word *physics* actually comes from a Greek word ΦΙΣΙΚ meaning "nature." META is a Greek word meaning "beyond." So metaphysics literally means *beyond nature*. In particular, to answer the question of "how do you define what something is?" we need a branch of metaphysics called **ontology**.

So how will we define things like electrons and light? In philosophy, we would ask, "what will be our ontological framework?"

We will construct our definitions of light and electrons in this course by:

- Listing what characteristics objects have
- Listing how objects interact with other constructs

Thus, our definition of an electron will be a list of its properties and interactions. In defining light and electrons in this way, we will see that we must actually look at two other objects: electric and magnetic fields to complete our picture. Does listing properties and interactions *really entirely define* what it means to be an electron or light? Probably not! There are certainly other possible ontological frameworks, but those are topics for a philosophy class. Science is a powerful way to understand the world, but its requirement of experimental falsifiability does have limitations which is why general education courses are so important for scientists!

How to Use This Book

This book has been specifically designed for this course out of free-and-open resources such as the OpenStax College Physics textbook[1], University of Maryland's UMD NEXUS Wikibook[2], as well as other resources from around the internet. While the text has all the information you need, some sections are also available as videos on our [course YouTube page](). These sections will have the link at the beginning with the section below.

Many students often complain that physics lectures, "spend all of their time on the easy problems and never have time to work through harder ones." In this course, we want to make sure that you have time to work through harder ideas and problems in-class where you can get help from the TAs and instructor. To make sure we have time to explore these concepts, you will need to have some familiarity with basic facts before coming to class. We will expect you to learn the relevant formulas, symbols, and terminology as well as review relevant material from 131 before coming to class using the resources in this book. We will NOT expect you to learn how to solve new types of advanced problems. This style of class, with the expectation of completing readings first is called a *flipped classroom* and is based on research[3].

This book, like the course, is divided into units. Each unit will begin with a *Unit on a Page* which summarizes the key points of the unit, both prep and in-class, on a single page. Due to the constraint of fitting on a single page, these pages are necessarily dense. The goal is to help you focus on the underlying themes, see what is really important. I would recommend you print these pages and keep them with you both during your prep and in-class to help you see the big picture.

After the *Unit on a Page*, you will find a short description of the context that we will be using to explore these physical concepts. We want physics to be relevant to *you*. Thus, each unit has one (or more!) biological applications that we will use as common themes throughout the unit. To make sure everyone is on the same page, we will explain the needed biology.

After the biology, there will be the physics material that you need to master *before* coming to class: readings, videos, flashcards to help you remember, and the occasional simulation for you to play with. To help you ensure that you have understood the material, there are homework problems linked from the text that will appear in special colored boxes that look like the box below to help you notice them. These problems are interspersed throughout the text, including on the biology material, to make it easy for you to know where to look for the information for each problem. To make sure you don't miss any homework problems, we will also make a list of all the problems at the end of the chapter. The homework problems are hosted in the Edfinity homework system. Follow the procedures outlined by your instructor on how to get an account, and check the syllabus for the grading policies.

Homework Problem:

1. https://openstax.org/details/college-physics
2. http://umdberg.pbworks.com/w/page/90716129/Working%20content%20I%20(2015) and http://umdberg.pbworks.com/w/page/104048687/Working%20content%20II%20(2016)
3. Stoltzfus, Matthew W. "Active Learning in the Flipped Classroom: Lessons Learned and Best Practices To Increase Student Engagement." In *The Flipped Classroom Volume 1: Background and Challenges*, 1223:105–22. ACS Symposium Series 1223. American Chemical Society, 2016.https://doi.org/10.1021/bk-2016-1223.ch008.

Your homework problems will appear as links in boxes like this one.

They are placed in-line with the text so that you know where to look for the information to solve each problem.

To further ensure that everyone is ready for what will be discussed in class, there will be quizzes on the material as outlined in your syllabus. We WANT you to be prepared for these quizzes; the entire purpose of this book is to prepare you. As such, there will be UMass-Amherst Instructor's Notes like the one below which will tell you exactly what you need to focus on to be ready for the quizzes.

Instructor's Notes

Be sure to pay attention to the text in this box! This is the material that will be on your quiz

We encourage your feedback on this book. If you have suggestions for comments or would like to report an error, please go to https://goo.gl/forms/VQaiFUtKD1PxJfH83 and complete the form.

We hope this book is helpful!

Editors:

Brokk Toggerson

Emily Hansen

Goals for The Course

The big questions: What is an electron? What is light?

These are the big questions that I am hoping to provide at least some answers to over the course of this semester. Light and electrons are two of the most fundamental building-blocks of our Universe (as far as we know they have no substructure!). Understanding these two basic elements and how they interact with each other will help you better understand many other fields of science and technology from chemistry to electronics.

Below are the fundamental goals that I would hope that you will take away from the “lecture” portion of this course. These are things that I hope that you will remember many years from now when you have forgotten all of the details. They are divided into two categories: Physics Goals and Skills Goals. There are a completely separate set of goals for the laboratory portion of the course which you will see in lab.

Physics Goals

These are the basic goals of any introductory physics course and are deeply connected to the material we will be covering. While these goals are generally similar to those from P131, I will be expecting that you will be developing a greater proficiency in these goals. Moreover, there are some changes for you to note.

1. **Physics is a set of principles and the fundamental ideas that relate them, NOT a list of equations…** This is quite possibly the most common misconception that people have about physics: they tend to think of physics as a list of formulae to be memorized and that all solving physics problems entails is finding the correct formula to get from where you are to where you want to be. This could not be further from the truth. Physics is a list of conceptual ideas expressed mathematically. These conceptual ideas form the “rules” that all of the other sciences, biology, chemistry, etc., have to follow. Thus, knowing the basic principles of physics is beneficial to any scientist!
2. **(More Emphasis!) These principles of physics can be expressed in multiple different ways…** When people think of physics they tend to think of equations. However, the ideas of physics can be represented in words (as was done in Europe in the days before Isaac Newton!) pictures, and graphs. As a budding scientist, it is important for you to be able to think of ideas in multiple formats and to be able to decide which format is the best for a given situation. Given the increased remoteness of the material in this course from your everyday experience, I am going to increase the emphasis on this goal relative to P131; being able to write clearly and succinctly about physics concepts using words will be important!
3. **Appreciate the value of the problem solving method used by the discipline of physics…** This is the goal most commonly given by non-physics faculty as to why they want their students to take physics, “problem solving.” However, the problem solving method used by physicists is a bit different than that to which you may be accustomed. In physics, we reason from the fundamental principles at play in a given situation. This requires you to look beyond the surface features in a given problem to see how a person throwing a ball, a block sliding down a ramp, and a building standing still are all similar in that they all rely on the same fundamental principle of . As a consequence of starting from fundamental principles, we physicists like to employ what is known as a “reductionist” approach and think about an idealized world first. This idealized world can seem quite bizarre to people new to physics as it is populated with point

masses moving across frictionless surfaces. The goal of this approach is to move to a problem where the fundamental principles are easier to see and understand. The complications that are present in the real world are then added back in later.

4. **(New relative to Physics 131) The fundamental principles of nature do not need to conform with "common sense"…** You already saw some of this in P131 where some results can seem counter-intuitive. In this course, however, we will encounter many more phenomena that may seem as though they do not make sense. Some of this will be due to the fact that you do not have as much experience with the world of electrons as you do with the world of friction and springs. Some of these topics' seemingly nonsensical nature, on the other hand, will be due to the fact that the topics represent fundamentally new concepts without any corresponding analog in your experience. For these ideas any analogy will inherently be imperfect and beginning from first principles (see Physics Goal #3) will be even more critical!
5. **Learn how to use fundamental principles to generalize from one specific situation to a class of similar ones…** Often, we will study a particular principle or idea within the context of a specific situation. The beauty of the laws of physics, however, is that these laws can be applied anywhere and the results of one analysis can often be applied to other related problems. For example, the motion of ANY object near the surface of the earth shares certain features. Thus, by studying the motion of a basketball, you can infer something about a skydiver. The trick is knowing what aspects of a given problem transfer to a new situation and which do not – your guide here are the fundamental principles (see Physics Goal #1).
6. **(More Emphasis!) Understand that the physics we study is connected to your everyday experience and the material in your other courses…** We want you to see the applications of physics all around you and to connect to your other courses. As such, we may ask you to pull information from your everyday experience or other knowledge to solve problems. Our hope is that physics will provide a new perspective on the material in your other classes. In this class, we will, in particular, be making a lot of connections to the subject of chemistry.

Skills Goals

In addition to physics content, there is a certain set of skills that I want you to take away from this class. These Skills Goals, however, are a bit different and more advanced than for P131 – as benefits a second semester course. These skills are just for the "lecture" portion of the course; the lab portion has additional goals emphasizing data analysis that will be discussed in your first lab sections.

1. **(More Emphasis!) Becoming more comfortable working in symbols…** This is a very important skill as every field becomes ever more quantitative; you need to be able to work in a wholly symbolic fashion and be able to read and interpret what the symbols in an equation mean. You began developing this skill in P131, but we will focus on it more heavily in this class.
2. **(New!) Be able to combine different ideas into a single analysis…** Most of the situations we looked at in P131 were single principle situations; we used either Newton's Laws or conservation of energy for our analysis. We almost never used more than one idea. However, almost all of modern science is what is called interdisciplinary: the result of combining ideas from multiple places. The trend of holistic health is a good example as it looks at not only biological factors but also sociological and psychological. The study of light and electrons is a great place to practice this skill as they are interesting multifaceted objects and we will need to be able to put multiple principles together to understand them.
3. **(New!) Interpretation of mathematical results…** In P131, when problems required a number or formula, you would solve it out and be done. P132 will force you to extend your problem solving skills beyond P131 to include one more step critical to more advanced analysis: interpretation of your result. In this class,

formulas will sometimes give you special results that require additional consideration. Other times, the result of a calculation may even be nonsense: infinity or imaginary! One of the amazing things about physics is that even these seemingly meaningless answers can tell you about how the world works. In this class, you will expected to develop the skills to interpret these results and learn what they can tell you.

Teamwork goals

I firmly believe in the critical role of teamwork to the success of the scientific enterprise. My own personal experience has confirmed this role many times and research shows that people learn better when they engage with ideas in conjunction with others. Throughout the course, regardless of if you are on a formal team or not, you will be furthering the scientific skills you developed in P131:

1. **Appreciate that the "solitary genius" image of a scientist which is so pervasive in our culture no longer exists (if they ever did)…** Science, all science, is now done in teams ranging in size from small teams of three to massive collaborations with memberships in the thousands. These skills are something you can put on your CV/Resume as having developed in this class!
2. **Appreciate that the work done by a team is usually better than the product of even its strongest member….** You will see over the course of the semester that this is true!. However, the environment in which we find ourselves is not as conducive to formalized teamwork as P131. However, I would strongly encourage you to work with other people both in-class on activities and out-of-class on homework and exam preparation.

Biology, Chemistry, Physics, and Mathematics

Editors' note: This section is based upon work from[1]

To become a biologist or health-care professional, you have to study a variety of scientific disciplines — biology, chemistry, physics, and math. You might have noted that the world doesn't actually divide itself in this way. Rather, the disciplines historically have been a way of choosing a sub-class of the phenomena that occur in the world and looking at a particular aspect of them with a particular purpose in mind. Different disciplines have different sets of tools and ways of knowing. Looking at something from different disciplinary perspectives adds a richness and depth to our understanding — like taking two 2-D pictures and merging them into a 3-D image.

Your introductory science and math classes often provide you with some basics — tools, concepts, and vocabulary — but may not give you a perspective on what each discipline adds to what you are learning and how they all fit together. Each discipline has its own orientation and perspective towards the development of a professional scientist. Here's a brief (and oversimplified) overview of the different disciplines that you encounter in studying biology.

Biology

Biology, as you well know, is the study of living organisms. The approach taken by biology is guided by and constrained by the fact that the subject is about living organisms.

- **A lot of biology is *complex*** — Because of the complexity, the first steps in biology (and in other sciences of the complex) are often about identification, classification, and description of phenomena. Whenever a science considers a complex phenomenon it does this — whether it's biology, organic chemistry, or plasma physics. In biology, it is important to describe the traits, structure, and behavior of a biological phenomenon before looking toward explanations of how it works. So it was important to do Linnaean classification and morphology before the ideas of evolution could be worked out; and an understanding of the nature of organic chemistry and biological molecules was necessary before the molecular functioning of biological systems could be disentangled. This results in biology having a huge vocabulary and many concepts to learn.
- **Biology depends on history** — By this, we don't mean the history of how the science of biology developed, but the history of how organisms developed. All biological organisms are connected through a common, unbroken, history – a chain or web of lifeforms – that affects how things are today. What has happened over time matters in biology and affects how things are today. This is like geology, and unlike chemistry, physics, or math. (Though when biology gets down to the mechanism of how things actually happen, it is very much like chemistry and physics, and uses math.) The properties of organisms that are currently alive and their relationships to their environments and to each other depend a lot on what happened to their ancestors in the distant past. The history of an organism is written in its genome. Knowledge of evolutionary processes is often an important tool to "explain" why a particular organism solves a biological

1. E.F. Redish, "The disciplines: Physics, Biology, Chemistry, and Math," in *Introductory Physics for the Life Sciences I - NEXUS Physics*, University of Maryland, College Park, 2013.

problem in a given way.

- **Biology looks for mechanism** — Biology is not just about "What is life?" It's also about "How does it work?" At one level, you might look at the organs and parts of either an animal or a cell and figure out what their function is for the organism. Today, using the tools of chemistry and physics (and using math), biology has gone down to the atomic and molecular level, figuring out the biochemistry of genes and proteins. Today, such quantitative measurements can be carried out simultaneously on thousands of genes or proteins in an organism. This has opened a new frontier of science, "Systems Biology", which aims to find mechanisms in these huge datasets and describe how thousands or millions of components work together in a biological system such as a cell, an organism, or a population.
- **Biology is multi-scaled** — an organism can be considered at many scales, for example, the atomic and molecular scale (biochemistry), in terms of the internal structure and functioning of its organs and parts (physiology), and as a part of a much larger system both in space (ecology) and time (evolution). The relation between these scales can be treated by reductionism or emergence — going to smaller scales to explain something (reductionism), or seeing new phenomena arise as one goes to a larger scale (emergence).
- **Biology is integrative**– Biological phenomena emerge from and must be consistent with the principles of chemistry, physics, and math. In other words, chemistry and physics constrain how an organism can behave or evolve. Therefore biologists must understand how physics and chemistry manifest themselves in biological organisms and higher-order systems. Increasingly, biologists searching for mechanisms of complex biological behavior are finding it valuable to use mathematical, physical, and chemical models in their research.

Chemistry

Chemistry starts with the idea that all matter is made up of certain fundamental pieces – atoms of about 100 different kinds (elements) – and is about the ways those elements combine to form more complex structures – molecules. But chemistry is not just about building molecules. It's about what you can do with that knowledge in our macroscopic world.

- **Chemistry is about how atoms interact to form molecules** — Understanding the basic principles of how atoms interact and combine is a fundamental starting point for chemistry.
- **Chemistry is about developing higher-level principles and heuristics** — Because there are so many different kinds of molecules possible, chemistry develops higher-level ideas that help you think about how complex reactions take place.
- **Chemistry frequently crosses scales** — connecting the microscopic with the macroscopic, trying to learn about molecular reactions from macroscopic observations and figuring our what is possible macroscopically from the way atoms behave. The connections are indirect, can be subtle, and may involve emergence.
- **Chemistry often assumes a macroscopic environment** — Much of what chemistry is about is not just idealized atoms interacting in a vacuum, but is about lots of atoms interacting in an environment, such as a liquid, gas, or crystal. In a water-based environment, the availability of H+ and OH- ions from the dissociation of water molecules in the environment plays an important role, while in a gas-based environment, the balance of partial pressures is critical.
- **Chemistry often simplifies** — In chemistry, you often select the dominant reactions to consider, idealize situations and processes in order to allow an understanding of the most important features.

For a chemist, most of what happens in biology is "macroscopic" – there are lots and lots of atoms involved

– even though you might need a microscope to study it. In introductory chemistry you often assume that reactions are taking place at standard temperature and pressure (300 K and 1 atm).

Physics

The goal of physics is to find the fundamental laws and principles that govern all matter — including biological organisms. Those laws and principles can lead to many types of complex and apparently different phenomena. Physics as traditionally taught at the introductory level tends to explicitly introduce four scientific skills that may seem different to what you see in introductory biology and chemistry classes, but these four skills will prove valuable for your career.

- **Physicists often spend a lot of time working out the simplest possible example ("toy model") that illustrates a principle** — even if that example appears not particularly interesting, relevant, or realistic. This lets you understand clearly and completely how the principle works. This understanding then can be woven into more complex situations to produce a better sense of what's going on (although the embedding of the simplicity in a realistic, relevant, and complex situation is often omitted in traditional introductory physics classes).
- **Physicists quantify their view of the real world** — Although there is a lot of conceptual and qualitative reasoning in physics, physicists tend not to be satisfied until they can quantify what they are talking about. This is because purely qualitative reasoning can sometimes be misleading. While you can come up with an argument that says A happens, if you think carefully, you might also come up with an argument that says something different happens — B. It's not until you can figure out that effect B is 1000 times bigger than effect A that you really know how to describe what's going on. This is just as true in biology and chemistry as physics, but physicists tend to introduce quantification sooner in the curriculum and more extensively than chemistry, which does it more in introductory classes than biology does.
- **Physicists think with equations** — This is more than just calculating numbers: physicists use equations to both organize their qualitative knowledge about what affects what and how, and to reason with in order to determine how things happen, what matters, and how much. Physicists go back and forth repeatedly between thinking conceptually about a problem and thinking mathematically about a problem, so that each of these ways of thinking sheds light on the other.
- **Physicists deal with realistic situations by modeling and approximating** — This means identifying what matters most in a complex situation and building up a fairly simple model that lets you get a good picture of what's happening. This is where the art lies in physics: in figuring out what can be ignored without losing what you want to look at. Einstein got it right when he said: "Physics should be as simple as possible, but not simpler." All sciences do this, but because physics is about "anything and everything", physicists often assume that they can get away in introductory classes with choosing systems that may seem to be simplified to the point of irrelevance. In this class, we'll try to be more explicit in modeling complex examples than in traditional physics classes.

This way of doing science is a bit different from the way biology is often done — but elements of this approach and the constraints imposed on biology by the laws of physics are becoming increasingly important both for research biologists and health-care professionals. For more discussion, see the page, What Physics Can do for Biologists.

Math

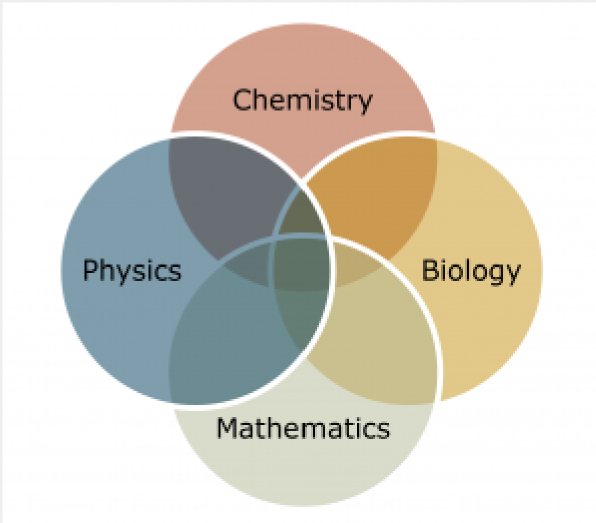
Math is a bit different from the sciences. In its essence math is about abstract relationships. Since math is about abstract relationships and how they behave, it's not "about" anything in the physical world. But it turns out that a lot of relationships in science can be modeled by relations that obey mathematical rules, often very accurately. (If you think this is surprising or strange, you aren't alone. For fun, take a look at the interesting article by the Nobel Prize-winning nuclear and mathematical physicist, Eugene Wigner, entitled, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences.")

Math as taught in math classes often is primarily about the abstract relationships — learning how to use the tools of math. Making the transition to using math in real-world situations may be quite jarring as there are now additional things to pay attention to other than the math itself — such as figuring out how the elements of the real-world system get translated into a mathematical model and worrying about whether the mathematical model is good enough or not. I like to think of it this way: in math class you learn the "grammar" of the language of science. Here, and in your other science courses, you need to start learning "vocabulary." With grammar and vocabulary together, you can begin to describe the Universe.

Bringing these disciplines together

Bringing these all together to permit coherent and productive thinking is a challenge! In this class we expect and encourage you to bring to bear knowledge you have from your other science classes — to try to see how they fit together, support each other, and to learn to identify when a particular disciplinary approach might be most appropriate and useful.

While these different scientific disciplines are all ultimately working to the same end: understanding the Universe. They did evolve semi-independently historically. As such, there are *cultural differences* between the sciences just as there are cultural differences between countries (driving on the right or left, for example). These cultural differences are not about "right" or "wrong" ways of doing things. They are just different. In fact, these differences in perspective are a strength! The different viewpoints between disciplines have often led to many important discoveries throughout history and are still where many of the most exciting advancements are being made. They can, however, be confusing. We have, therefore, worked with biology, chemistry, and mathematics instructors here at University of Massachusetts – Amherst. These discussions have resulted in some common language used in this book, which might, therefore, be different than in other physics text you may look at. Even so, there are sometimes places where we need to leverage the strength of a different viewpoint. We will point out these differences using boxes like the one below.



We will use boxes like this to point out important differences between disciplines. Again, these differences are not “right” or “wrong” ways of doing things. They are simply artifacts of the different sciences evolving independently for centuries with influences from different cultures from across the globe.

Copyright, Licenses & Sources

Published in the United States of America by
UMass Amherst Libraries
154 Hicks Way Amherst, MA 01003

What is an Electron? What is Light? is adapted from several openly licensed materials listed below.

OpenStax College Physics

Senior Contributing Authors

Paul Peter Urone, California State University, Sacramento

Roger Hinrichs, State University of New York, College at Oswego

Contributing Authors

Kim Dirks, University of Auckland

Manjula Sharma, University of Sydney

OpenStax Chemistry Atoms First 2e

Senior Contributing Authors

Paul Flowers, University of North Carolina at Pembroke

Edward J. Neth, University of Connecticut

William R. Robinson, PhD

Klaus Theopold, University of Delaware

Richard Langley, Stephen F. Austin State University

Contributing Authors

Jennifer Look, Mercer University

Allison Soult, University of Kentucky

Simon Bott, University of Houston

Donald Carpenetti, Craven Community College

Andrew Eklund, Alfred University

Emad El-Giar, University of Louisiana at Monroe

Don Frantz, Wilfrid Laurier University

Paul Hooker, Westminster College

Carol Martinez, Central New Mexico Community College

Troy Milliken, Jackson State University

Vicki Moravec, Trine University

George Kaminski, Worcester Polytechnic Institute

Mark Blaser, Shasta College

Jason Powell, Ferrum College

Thomas Sorensen, University of Wisconsin–Milwaukee

OpenStax AP Biology

Senior Contributing Authors

Julianne Zedalis, The Bishop's School in La Jolla, CA

John Eggebrecht, Cornell University

Contributing Authors

Yael Avissar, Rhode Island College

Jung Choi, Georgia Institute of Technology

Jean DeSaix, University of North Carolina at Chapel Hill

Vladimir Jurukovski, Suffolk County Community College

Connie Rye, East Mississippi Community College

Robert Wise, University of Wisconsin, Oshkosh

The disciplines: Physics, Biology, Chemistry, and Math, in Introductory Physics for the Life Sciences I – NEXUS Physics, University of Maryland, College Park and a few other sources from that Wikibook

Author

E.F. Redish

For other instructors who may wish to use this book

Abstract

This free-to-students open textbook is designed to be the preparatory reading and homework for a flipped second-semester IPLS course focused on the guiding questions of “What is an Electron?” and “What is Light?” The course has five units: quantum mechanics, geometric optics, electrostatics, circuits, and magnetostatics with some electrodynamics. Each unit is also has a guiding question from biology or chemistry. For example geometric optics uses the eye and circuits the neuron. Relevant biology and chemistry is reviewed throughout using biologically authentic language from textbooks and classroom videos.

Notes on the text and links to other resources

Hello,

This textbook is designed to supplement the Physics 132 – *Introductory Algebra-Based Physics for Life Sciences II: What* is *an Electron? What* is *Light?* course taught at University of Massachusetts Amherst. More details on the structure of the course can be found at https://physedgroup.umasscreate.net/course-materials/p132-second-semester-ipls/. Also at that link, you can request all of the course resources.

The homework for this course is hosted in the Edfinity homework system. The particular set of problems is available as a resource at https://edfinity.co/ipls2.

Please, do not hesitate to contact Brokk Toggerson, if you have questions or suggestions.

Hope you find this useful.

PART I

UNIT I

Unit I On-a-Page

Principles and Definitions

If you have had Dr. Toggerson or Dr. Bourgeois for Physics 131, you are familiar with a distinction made between *principles* and *definitions*. The **principles** are the fundamental rules of the Universe that describe how things work. Concepts which are **definitions**, on the other hand, simply describe a quantity. For example,

$\vec{p} = m\vec{v}$ p⃗=mv⃗p⃗=mv⃗widevec {p} =m widevec {v}">

is the *definition* of momentum for a massive particle; this equation offers no deep foundational insights on how the universe works. We physicists simply noted that the quantity $m\vec{v}$ came up a lot and we gave it a name $\vec{p}$. In order to describe how the Universe works, principles will often involve multiple definitions. Note, sometimes a principle or definition has an equation, other times it is just stated in words! This connects to Physics Goals 1 and 2 for this course.

To help get those of you who may not be used to this distinction acquainted and to help organize the huge amount of factual information in this particular unit, I will list the principles for this unit. You can quickly see how short this list is.

Principles for this Unit

Basic Properties of Light

- Light in a vacuum always travels at the speed of light $c = 299792458\,\frac{\mathrm{m}}{\mathrm{s}}$

Basic Properties of Waves

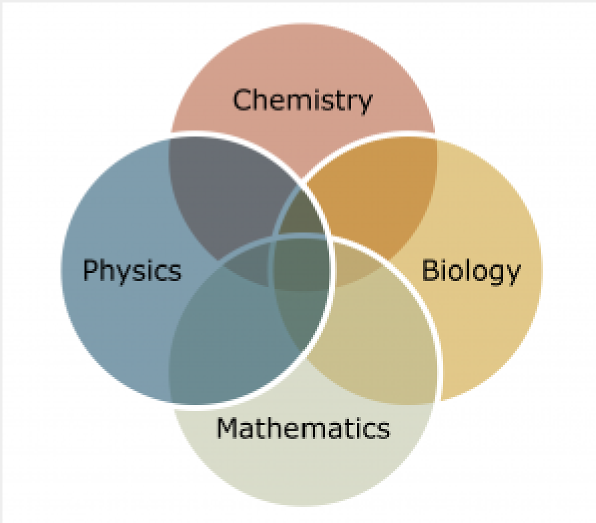
- Fundamental connection between **wavelength** λ, **frequency** ν, and wave speed v: $v = \lambda\nu$
- The **amplitude** A is independent of frequency ν

Basics of Energy

- Energy is conserved. The change in **energy** ΔE is caused by the exchange of energy through **heat** Q and **work** W: $\Delta E = Q + W$.

Wave Particle Duality

- You can convert from the wave picture to the particle picture through the de Broglie relation: $p = \frac{h}{\lambda}$ where p is the momentum of the particle.



In chemistry, you probably saw the conversion between energy and wavelength for photons done through the equation $E = \frac{hc}{\lambda}$. We will NOT be considering that as a principle to begin analyses / solving problems. The reason is that, in chemistry, you only considered converting between energy and wavelength for photons. We want to think about BOTH photons AND electrons. It turns out, that $E = \frac{hc}{\lambda}$ ONLY works for photons, while $p = h/\lambda$ works for both photons and electrons! Thus we consider $p = h/\lambda$ to be our principle. Many students get tripped up by applying $E = \frac{hc}{\lambda}$ to electrons. Don't fall into this trap!

- The probability of finding a particle in a given location is proportional to the square of the amplitude $P \propto A^2$.
  - Increase amplitude by 3, probability goes up by 9.
  - For light, we represent the amplitude not by A but by E. This is NOT the energy (confusing I know, but it is what it is). We will see why we use E later in the semester.

Standing Waves

- The wave must “fit” in the box or on the ring. For a box, this means that the box must be an integer number of 1/2 wavelengths $n\left(\frac{\lambda}{2}\right) = L$.

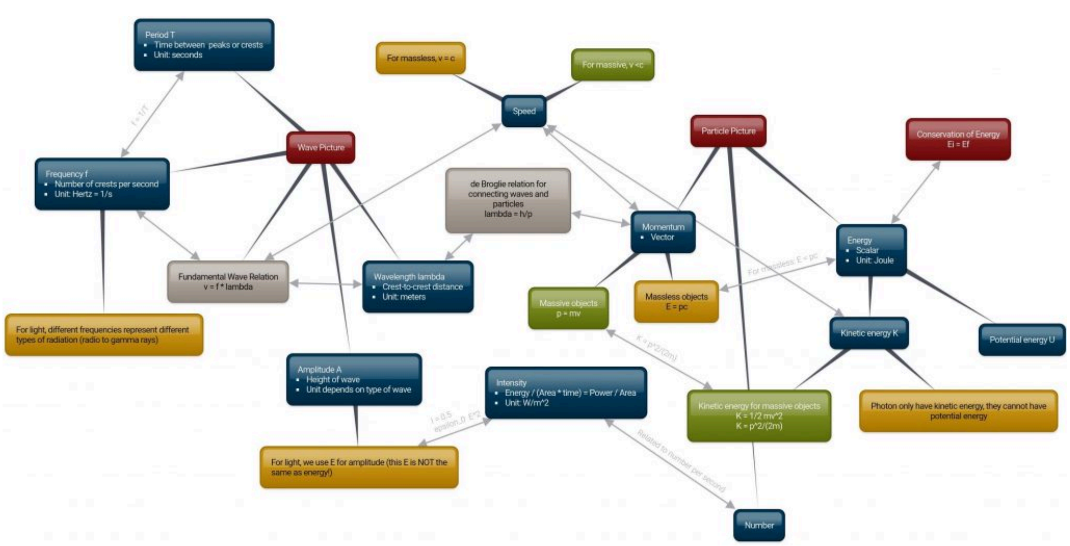
Instructor's Note

The ideas in this unit can be connected in many different ways. One possible useful way to organize such information is in a "concept map" like the one shown below. [The map is also available at this link].

In this map:

- UMass maroon bubbles are big ideas
- Yellow bubbles apply to massless particles like light
- Green bubbles apply to massive particles like electrons

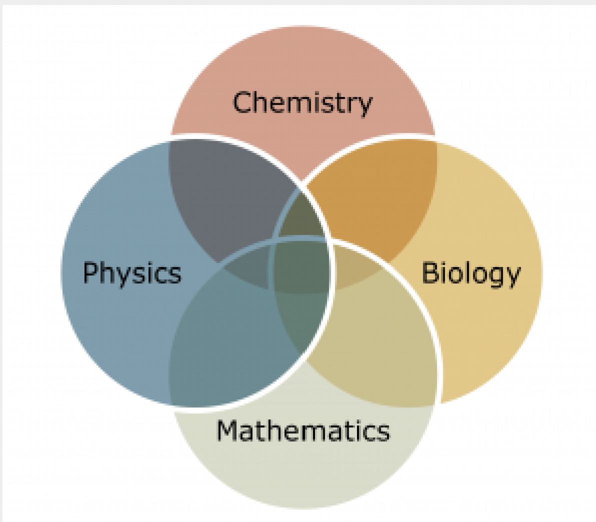
I would recommend printing a copy for use in class!



Unit 1 on a page!

1. Unit I - Introduction and Context for the Unit

Interdisciplinary questions we want to answer in this unit



The motivation for this unit is pretty straightforward: you know from Chemistry electrons and photons can behave as both particles and wave a property known as **wave-particle duality**.

- What does this wave-particle duality actually mean?
- WHY does what you learned in chemistry actually work the way it does?
  - How does this wave-particle duality result in some of the most fundamental properties of chemistry such as discrete energy levels in atoms?
  - Why is the wavelength of a photon emitted from an atom $\Delta E = \frac{hc}{\lambda}$?

Introduction to the Unit

In this unit, we will follow our ontological framework and begin exploring what light and electrons are by listing some of their basic properties. One of the most famous properties of both light and electrons is known as wave-particle duality: sometimes they behave like particles and sometimes they behave like waves. However, electrons and light are neither particles nor waves: they are a completely new type of object with properties of both. This duality is a reflection of the fact that both light and electrons do not obey the laws of Classical Physics that you learned in Physics 131, they are too small. Instead, electrons and light obey Quantum Mechanics.

The way I would recommend that you think about the relationship between classical physics and quantum mechanics is in the paradigm of physics striving for ever-more-accurate approximations to reality. Classical mechanics is a good enough approximation to get people to the moon (they did it!). When things get small, you need a better approximation: quantum mechanics. There is a principle, the correspondence principle, which states the quantum mechanics must reproduce classical mechanics for large objects.

In this unit we will explore both the wave and particle natures of light and electrons. First however, we must define to ourselves what waves and particles are! Following our ontological framework, we will therefore need to look at some of the basic properties that characterize waves and particles in general. Thus, we will begin with some review of particles from physics 131 and then a discussion of waves, which may be familiar to some of you. Once we have defined waves and particles through listing their properties we will explore how these properties manifest for electrons and light.

As you read, you MUST keep in mind that light and electrons are neither particles nor waves. They are something completely new (quantum mechanical objects) that you have zero previous experience with. Particles and waves are simply ways of visualizing these objects in ways our brains can understand. Neither picture is 100% correct. The correct approach is to jump back-and-forth between these two pictures

What we will do in class

This idea of electrons and light being neither particles no waves but having properties of both is a hard one to get used to. In class, we will spend a lot of time practicing jumping between the wave and particle pictures: seeing the benefits that each picture can bring in various situations. The goal is that, through practice, you become more comfortable with bouncing back and forth between pictures as the situation demands.

After a bit of practice with moving between the wave picture and particle pictures of light and electrons, we will combine this understanding with one of the most important ideas in physics: conservation of energy. Thinking about this fundamental principle in conjunction with the fundamentally quantum nature of light and electrons will allow us to understand many different phenomena such as, “Why do electrons in atoms have defined energy levels?” You know from chemistry courses that they do. Our goal is to explain why!

What you should get out of this prep

In order to explore these ideas in class, you need to have a grasp on the basic terminology of waves and particles. You need to know that particles are characterized by their energy, momentum, and number while waves are described by their wavelength and frequency/period, and amplitude. Both waves and particles can be characterized by a speed: a critical fact for converting between the wave and particle pictures. The following chapters will refresh energy and momentum from 131 and introduce the needed concepts for waves: amplitude, frequency, and wavelength. You need to know what all these terms mean, the basic formulas for them, and how they are connected

Also in this unit, some of the basic properties of light and electrons. We will establish electrons via a tour of the atom (probably review for most of you). We will also introduce anti-matter: a type of matter identical to the normal matter with which you are familiar in every respect with two exceptions. First, anti-matter has the opposite charge from normal matter (anti-electrons have positive charge). Second, when matter and anti-matter collide, the result is light. While there are other interesting questions about anti-matter (why isn’t it everywhere?) those two points are all you need to know from the prep. With regards to light, we will introduce the different kinds of light (radio, infrared, …) and the particle of light: the photon. Some of the basic properties of the photon will be introduced. The most important of which you need for the prep are the fact that the mass of the photon is zero and the fact that it always travels at the speed of light c. Finally, we will also explore de Broglie's relationship on how to connect wave and particle picture

The last topic in this unit's preparation you need to be familiar with is the idea of intensity: energy per area per time. The text will introduce the idea of power (energy per time) and its unit the Watt. This concept will also be important for converting between the wave and particle pictures.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

This is a lot of information. Remember, we are not expecting mastery of it all and we are certainly not expecting you to have a complete picture of how everything Just make sure you know the definitions, formulas, and units for everything. To help keep you focused on what is important, there is a summary tables below (only focusing on the facts for your quiz). Keep these tables handy as you go through the reading: the information will all be explained as you go.

Waves and Particles

This information applies to both electrons and light.

Intensity, detailed later, is the energy per time per area (or power $P = E/t$):

$$I = \frac{P}{A} = \frac{E}{t \cdot A}$$

	Properties	Convert between them
Waves	• Amplitude A, or E for light. • Wavelength λ, measured in meters. • Period T, measured in seconds. • Frequency $\nu = \frac{1}{T}$, measured in 1/s or Hz. • Speed $v = \lambda\nu$ measured in m/s. • Intensity: $I \propto A^2$	• de Broglie relation: $p = \frac{h}{\lambda}$ ◦ Do NOT use $E = \frac{hc}{\lambda}$ as that only works for light! • Amplitude: $A^2 \propto P$ where P is the *probability* of finding a particle. • Intensity: $A^2 \propto \frac{(\text{Number hitting per second})\cdot(\text{Energy of each})}{\text{Area}}$
Particles	• Energy E • Momentum $\vec{p}$ • Speed v • Number N • Intensity: $I = \frac{(\text{Number hitting per second})\cdot(\text{Energy of each})}{\text{Area}}$	

Electrons and Photons

Note, in general, if you see a c or an ϵ_0 in an equation, it applies to light only! The wave and particle information in the table still applies to both electrons and photons!

	Waves	Particles
Light	• Travel at c • Different frequencies → different kinds of radiation • Amplitude E (which, again, is NOT energy) is explicitly related to intensity: $I = \frac{1}{2}\epsilon_0 E^2$.	• Called photons γ • Mass: Zero • Electric charge: Zero • Speed: Always travel at c • Energy E and momentum p are related by $E = pc$ • Different energies → different types of radiation
Electrons	• Wavelength determined, as always, by de Broglie relation $p = h/\lambda$.	• Mass: $m_e = 9.11 \times 10^{-31}\,\mathrm{kg}$ • Electric charge $\lvert q_e \rvert = 1.602 \times 10^{-19}\,\mathrm{C}$ • Speed: Less than the speed of light • Momentum: $\vec{p} = m\vec{v}$ • Kinetic energy: $E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$ • There is also a particle called *anti-electron* or *positron* with the same mass and charge except it has positive charge while the familiar electron has negative charge. When the two meet, they destroy each other.

2. Basics of Matter

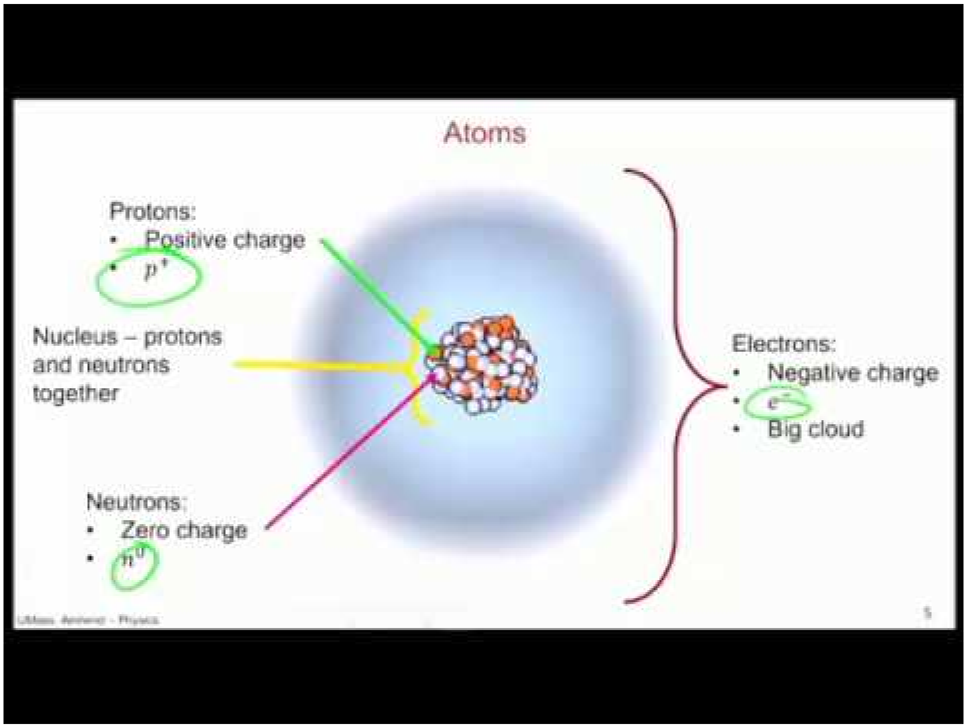
A Deeper Structure of the Atom

This section is available both as a video and as text. Below, you see the video as well as a text transcript. The content is the same: read or watch as is your preference.

Instructor's Note

The things you need to know for your homework and quizzes are:

- Electrons and protons are charged, neutrons are not; the size of the charge on the electron and proton is the same, but the signs are different, so same magnitude different sign
- Opposites attract and that is what holds the atom together
- Protons and neutrons have the same mass, and electrons are way lighter
- Protons and neutrons made of stuff, electrons are fundamental
- The nucleus is super tiny relative to atom



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=1650*](http://openbooks.library.umass.edu/toggerson-132/?p=1650)

You should be familiar with the basic structure of the atom, but as a review, in the middle of the atom the positively-charged protons and neutrons are huddled together in the nucleus.

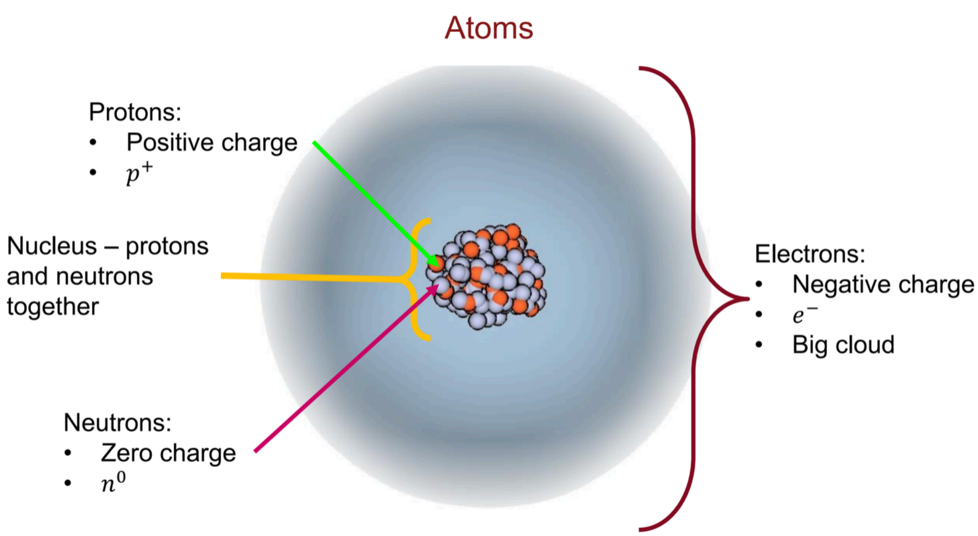


Figure 1: *The basic structure of the Atom.*

However, you might not be familiar with the related symbols. This symbol, p^+ means proton, *p* for proton, and then plus to remind us that it has a positive electrical charge (If you're wondering, can you have a negatively charged proton? Yes, it's called an **antimatter** proton, see below.) You also have neutrons. Neutron has zero charge, so it's symbol is n^0. Those are huddled together in the nucleus.

Surrounding the nucleus is a big cloud of negatively charged electrons, so we will use the symbol e^- for electron. It's the attraction between the positively charged protons and the negatively charged electrons that sort of hold the entire atom together, and how that all works will be the emphasis of Unit III.

Electrons are a big focus of this course, so it is worth discussing what they are made of. To our best of our knowledge they are not made up of anything. They are fundamental, we have been trying to smash them apart, but no luck. Maybe it's possible, but no one's been able to do it. If it is possible, we haven't hit it hard enough. That's very much the particle physics approach to everything- hit it harder and see if it breaks. So, electrons are fundamental building blocks- as far as we know they're not made up of anything.

Protons and neutrons on the other hand, are a lot more fun, because they are made up of smaller pieces called quarks. There are six kinds of quarks, we have 'up', 'down', 'strange', 'charm', 'top', and 'bottom'. Those are their official scientific names, I kid you not. In Figure 7 below, the sizing of the circles shows you the masses, how heavy this stuff is (but not their sizes – as far as we know the quarks have zero size!). The 'top' quark, the heaviest of the known quarks, actually has about the same mass as an entire atom. It's quite a heavy little thing. Three of these quarks, 'top', 'bottom', and 'charm', are actually heavier than protons.

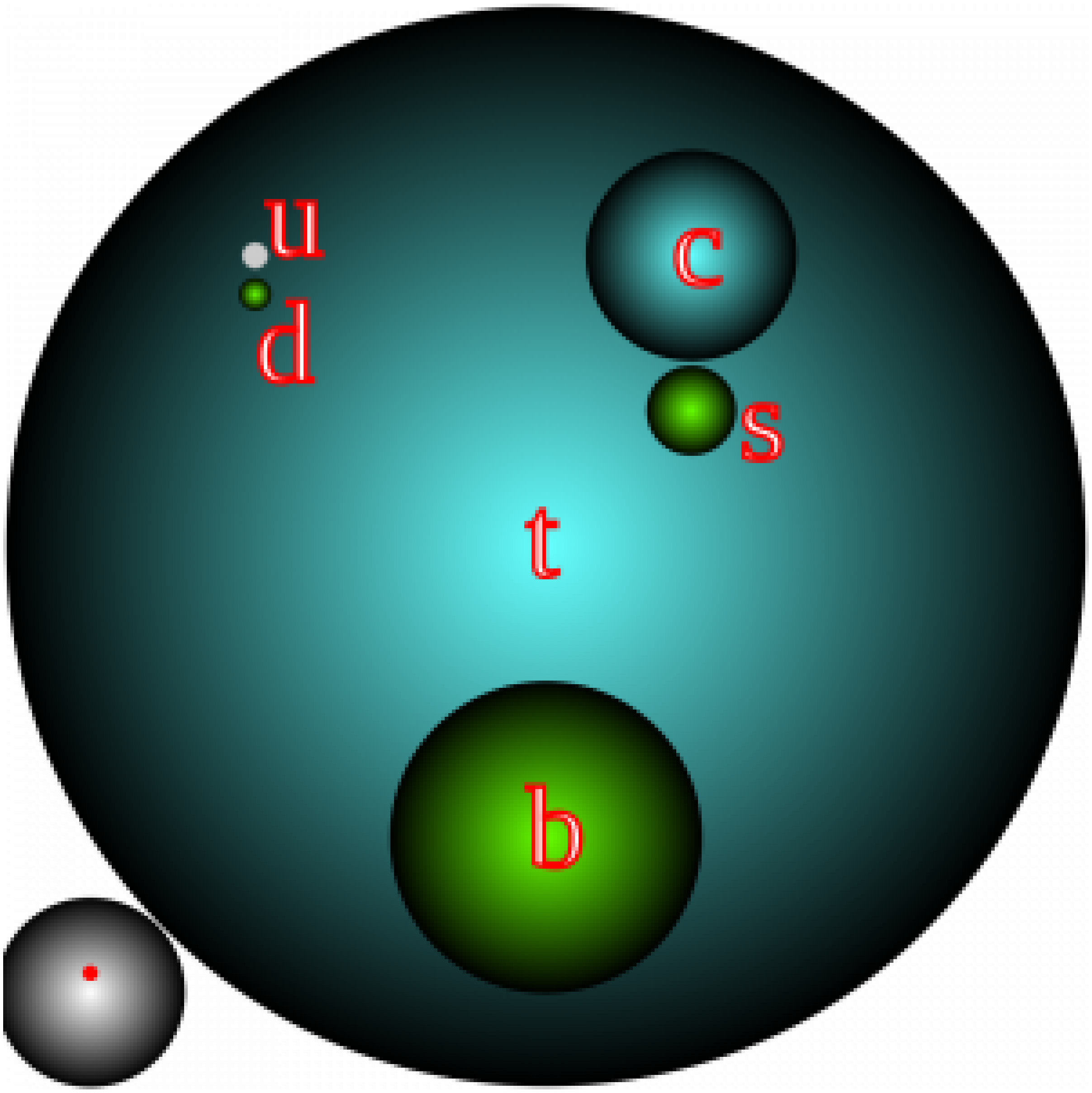


Figure 2: *The masses of the quarks: u for up, d for down, c for charm, s for strange, b for bottom, and t for top. Again, the size represents the mass, NOT the size; as far as we know all of these quarks have zero size! The grey ball in the lower left is a proton for scale. The small red dot inside the grey ball is an electron. (Credit: Incnis Mrsi [CC BY-SA (https://creativecommons.org/licenses/by-sa/3.0)])*

But, if charms, bottoms, and tops are all heavier than protons and neutrons, so what makes up a proton in a neutron? Protons and neutrons are made up of just these two 'ups' and 'downs'. So, a proton is made up of two 'up' quarks and one 'down' quark, a neutron on the other hand is two 'down' quarks and an 'up' quark as shown below in Figure 8.

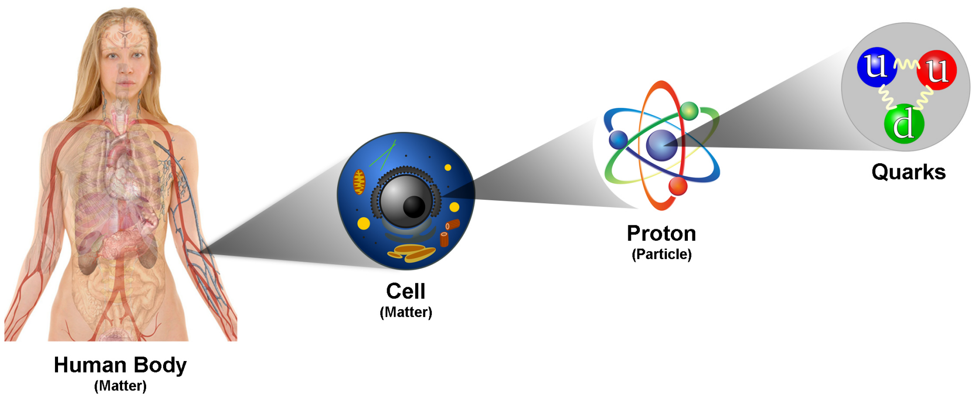


Figure 3: *Quarks make up both protons and neutrons. Here you can see the smaller-and-smaller steps all the way down to the two ups and a down which make up the proton. (Credit: [Finches & quarks [CC BY-SA (https://creativecommons.org/licenses/by-sa/4.0)]](https://creativecommons.org/licenses/by-sa/4.0))*

Now you start doing math, so you need the proton to have +1 charge, the neutron to have 0 charge, you have two 'ups' and a 'down', and two 'downs' and an 'up'. If you play with those numbers, what do you get? You have that 'up' quarks have a charge of 2/3 that of the proton, and down quarks are -1/3 that of the proton. And this works out: think 'up' 'up' 'down' so that's: (+2/3) + (+2/3) + (-1/3) = +1, which is the charge of a proton, and it works out.

Similarly, for the neutron think, (-1/3) + (-1/3) + (+2/3) = 0: the charge of a neutron, they work out.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

What's your big takeaway for this? Electrons are fundamental to our knowledge and cannot be broken apart. Protons and neutrons are made up of smaller stuff.

The other key things to know about atoms: protons and neutrons are very very close to the same mass, but neutrons are a tiny bit heavier, but not by much. Electrons on the other hand are way lighter than protons or neutrons. In fact, the electron is the lightest known particle to have electric charge with a mass of $9.11 \times 10^{-31}\ \mathrm{kg}$. Protons on the other hand, are much bigger, $1.67 \times 10^{-27}\ \mathrm{kg}$.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

What should you take away from this? Protons are way more massive than electrons, roughly 2,000 times (1836 times to be specific).

If you prefer to think about atoms instead of protons and neutrons, you can think about a Helium atom, you know there are two protons, two neutrons, and two electrons. The electrons make up 0.03% of the mass of helium. Electrons don't weigh squat. They don't really matter as far as mass goes. While the nucleus has most of the mass, it doesn't take up a lot of space. The standard analogy that people make is if you blow up the atom to the size of a large college football stadium, bigger than ours, the nucleus is roughly the size of a marble. Atoms are a whole bunch of empty nothing. The nucleus is about the size of a pea but it is 99.97% of the mass is in that marble comparatively speaking.

Figure 4: *If the atom were the size of a football stadium, the nucleus would be the size of a pea! (Credit: [MHarrison [CC BY-SA (https://creativecommons.org/licenses/by-sa/3.0)]](https://creativecommons.org/licenses/by-sa/3.0))*

Homework

Problem 1: What is the charge of the Δ?

Conservation of Mass is a lie! *Conservation of Energy* and *Conservation of Charge* are true! How this is connected to antimatter.

Instructor's Note

We will be occasionally dealing with antimatter in this class. You need to know that matter and antimatter are identical in mass, but opposite in charge: an anti-electron has a positive charge. You also need to know that when matter and anti-matter come together the result is pure energy.

In chemistry, mass is never created or destroyed. This is often described as the *conservation of mass*. However, this principle is actually NOT universally true! In more exotic situations, such as in particle accelerators, mass, Δm, can be created from energy in the amount using Einstein's famous relation:

$$E = mc^2 \rightarrow \Delta m = \frac{E}{c^2}.$$

This famous equation basically says that mass is just a particular form of energy: a highly concentrated form as even a small amount of mass, when multiplied by $c^2 = (3 \times 10^8)^2 = 9 \times 10^{16}$, yields a large amount of energy. In fact this is the energy you see everyday as sunlight as will be explored in your homework!

Homework

Problem 2: How much energy is produced in the sun from converting 2 protons and two neutrons into a helium nucleus?

In all observations to this point, every time a particle is created in this way another, having the *exact same* mass but *opposite* charge is always created along with it. The two particles are "matter-**antimatter**" counterparts. For example, an anti-electron would usually be created at the same time as an electron. The anti-electron has a positive charge (it is also called a positron as in positron emission tomography). Since the electron is negatively charged and the positron is positively charged, the total charge created is zero. This is a manifestation of one of the most important Laws of the Universe: *Conservation of Electric Charge*. Consider the reaction in Figure 4a: energy (with zero net charge produces an electron and a positron (also zero net charge!).

The reverse is also true: when matter and antimatter counterparts are brought together, they completely annihilate one another as seen in Figure 4b. By annihilate, we mean that the mass of the two particles is converted to energy *E*, again obeying the relationship

$$\Delta m = \frac{E}{c^2}.$$

Again, the total electric charge is conserved: we began with zero net charge (one negatively charged electron and one positively charged positron) and the result was energy with zero net electric charge.

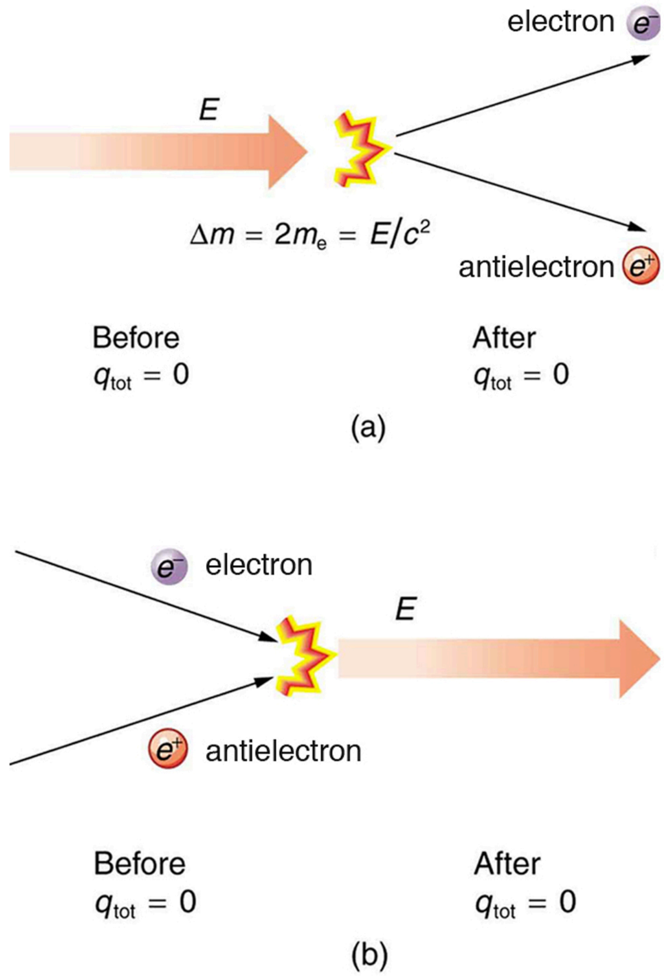


Figure 6: *a) When enough energy is present, it can be converted into matter. Here the matter created is an electron–antielectron pair. (m_e is the electron's mass.) The total charge before and after this event is zero. (b) When matter and antimatter collide, they annihilate each other; the total charge is conserved at zero before and after the annihilation.*

All particles have antiparticle counterparts: there are negatively charged anti-protons p^- made up of two anti-up quarks (charge -2/3) and one anti-down quark (+1/3). Similarly, there are even anti-neutrons made up of two anti-downs and one anti-up (note the anti-neutron also has charge zero, but would still annihilate if it met a regular neutron!).

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.

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.

Homework

Problem 3: How much energy is released when an electron and an anti-electron (positron) annihilate during positron emission tomography?

Where is all the anti-matter?

If every time an electron is produced, I also get a positron, why do we see electrons everywhere while positrons are so rare? In fact, why do we exist at all? In the early Universe after the Big Bang, there should have been equal amounts of matter and anti-matter which should have then annihilated itself leaving a Universe with no matter – just energy!

This is a good question! In fact, it is a question to which we do not know the answer! It may surprise you to learn that we don't know the answer to something as fundamental as "why is there something instead of nothing?" This is the fun of science! It means that there is still work to do!

3. Basics of Particles

What is a Particle?

What is a **particle**? The simplest image of a particle is probably just a ball. What properties apply to all particles? We talked about particles a lot in Physics 131 in the point mass approximation, but it's probably best that we flush out our definitions.

In its most generic sense, a particle is a chunk of stuff. It exists in a particular place and at a particular time and a particle doesn't go around corners. If I throw a ball at a door, it'll either go through the door or bounce back, it won't curve around it. Particles can, but do not necessarily have to, have mass, we will talk about a massless particle in a later section. But all particles can be thought of as having momentum, that quantity from 131 of mass times velocity. Particles can also be thought of as having energy.

Instructor's Note

In summary, you need to know that particles can be thought of as balls with defined position and speed and are characterized by:

- Their energy E
- Their momentum $\vec{p}$
- How many of them there are N

Linear Momentum and Force (Review from Physics 131)

This material is review from physics 131, but we will use these ideas in this unit, so here is a short refresher.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

You quiz will cover:

- Calculate the momentum for any object
- Recall that momentum is a vector
- From the change in momentum, compute the average force

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. In symbols, linear momentum is expressed as

$$\vec{p} = m\vec{v}.$$

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·m/s.

Example 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. In both parts of this example, the magnitude of momentum can be calculated directly from the definition of momentum given in the equation, 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.

$$p_{\text{player}} = (110\,\text{kg})(8.00\,\text{m/s}) = 880\,\text{kg}\cdot\text{m/s}$$

Solution for (b)

To determine the momentum of the ball, substitute the known values for the ball's mass and speed into the equation.

$$p_{\text{ball}} = (0.410\,\text{kg})(25.0\,\text{m/s}) = 10.3\,\text{kg}\cdot\text{m/s}$$

The ratio of the player's momentum to that of the ball is

$$\frac{p_{\text{player}}}{p_{\text{ball}}} = \frac{880}{10.3} = 85.9.$$

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.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

The example above is representative of what you will be asked to do on your homework and quizzes.

Homework

Problem 4: Compare the momenta of elephants, humans, and tranquilizer darts!

Momentum and Newton's 2nd Law (Optional)

All you need to know from this section is the definition of momentum. The following connection to Newton's 2nd Law is just to help you put this info into context.

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. Using symbols, this law is

$$\sum \vec{F} = \frac{\Delta \vec{p}}{\Delta t},$$

where $\sum \vec{F}$ is the net external force, $\Delta \vec{p}$ is the change in momentum, and Δt is the change in time.

Newton's 2nd Law in Terms of Momentum

The net external force equals the change in momentum of a system divided by the time over which it changes.

$$\sum \vec{F} = \frac{\Delta \vec{p}}{\Delta t},$$

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 $\sum \vec{F} = m\vec{a}$ as a special case. We can derive this form as follows. First, note that the change in momentum Δp is given by

$$\Delta p = \Delta(mv).$$

If the mass of the system is constant, then

$$\Delta(mv) = m\Delta v.$$

So that for constant mass, Newton's second law of motion becomes

$$\sum \vec{F} = \frac{\Delta \vec{p}}{\Delta t} = m\frac{\Delta \vec{v}}{\Delta t}.$$

Because

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

we get the familiar equation

$$\sum \vec{F} = m\vec{a}$$

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 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

$$\sum F = \frac{\Delta p}{\Delta t}.$$

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 Δp is calculated,

$$\sum F = \frac{\Delta p}{\Delta t}$$

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.

$$\Delta p = m(v_f - v_i)(0.057\,\mathrm{kg})(58\,\mathrm{m/s} - 0\,\mathrm{m/s}) = 3.306\,\mathrm{kg \cdot m/s} \approx 3.3\,\mathrm{kg \cdot m/s}$$

Now the magnitude of the net external force can determined by using $\sum F = \frac{\Delta p}{\Delta t}$:

$\sum F = \frac{\Delta p}{\Delta t} = \frac{3.306\,\mathrm{kg \cdot m/s}}{5.0 \times 10^{3}\,\mathrm{s}} = 661\,\mathrm{N} \approx 660\,\mathrm{N}$,

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 $\sum \vec{F} = m\vec{a}$, but one additional step would be required compared with the strategy used in this example.

Chapter Summary

- Linear momentum (momentum for brevity) is defined as the product of a system's mass multiplied by its velocity. In symbols, linear momentum $\vec{p}$ is defined to be $\vec{p} = m\vec{v}$, where m is the mass of the system and $\vec{v}$ is its velocity.
- The SI unit for momentum is kg·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 $\sum \vec{F} = \frac{\Delta p}{\Delta t}$, $\sum \vec{F}$ the net external force, $\Delta\vec{p}$ is the change in momentum, and Δt is the change time.

4. Review of Conservation of Energy

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

This unit, in fact this entire course, will spend a lot of time talking about energy: a topic covered extensively in Physics 131 as well as in your Biology and Chemistry courses. This chapter is therefore a bit different: we provide links to the relevant sections on energy from the Physics 131 textbook *Forces, Energy, Entropy* for your reference, with the key takeaways from each section. Just review what you need.

There are also a few homework problems at the end, just to make sure everyone is on the same page.

Relevant parts from *Physics 131: Forces, Energy, Entropy:*

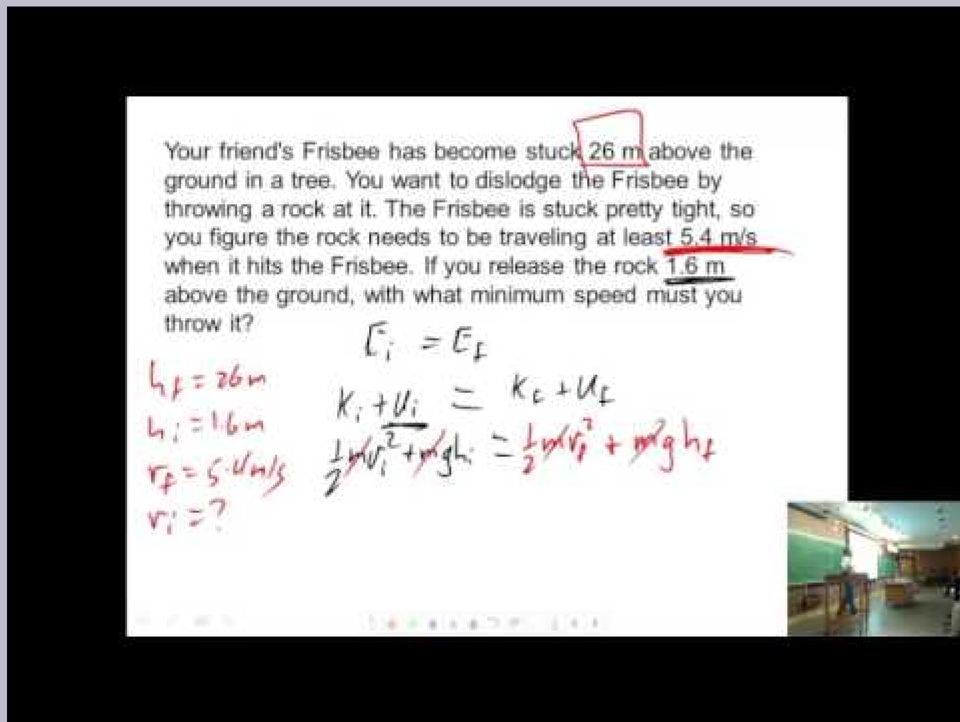
- Unit IV – Chapter 2 Introduction: Introduction to Energy
- Unit IV – Chapter 2.1: Units of Energy
    - Note, we will use the eV unit of energy described here much more in this class than in Physics 131.
- Unit IV – Chapter 2.2: Types of Energy and Scales of Energy
    - It is important to think about the fact that there are fundamentally only two kinds of energy: potential and kinetic.
- Unit IV – Chapter 2.3: Conservation of Energy
    - This is one of the fundamental **principles** of this unit (and all of physics). We will begin many problems with this idea.
- Unit IV – Chapter 2.4: Ways to Transfer Energy
    - The key idea here is that there are two ways to transfer energy: **heat** and **work**.

- Heat Q is the transfer of energy through collisions at the microscopic scale. This includes **photons**, an idea that will appear in a later section of this reading.
- Work is the application of a force for a distance $W = Fd\cos\theta$.

- Unit IV – Chapter 2.5: The Formal Statement of the Conservation of Energy as the First Law of Thermodynamics
  - The change in energy of a system ΔE is the amount going in or out as heat and work: $\Delta E = Q + W$. This is the principle from which we will begin many of our analyses.
- Unit IV – Chapter 3 Introduction: Energy of Objects as a Whole
  - This is where the ideas of kinetic and potential energy as you saw them in 131 are introduced.
- Unit IV – Chapter 3.1: Kinetic Energy of an Object
  - We will be looking a lot at the kinetic energies of electrons.
- Unit IV – Chapter 3.2: Examples Applying Conservation of Energy with only Kinetic Energy
  - If you need to refresh how to use the kinetic energy formula $K = \frac{1}{2}mv^2$
- Unit IV – Chapter 3.3: Macroscopic Potential Energy
  - This is where the idea of gravitational potential energy $U = mgh$ is discussed. We will use this expression in this unit.
  - More importantly, we will build an understanding of electrical potential energy in Unit III partially by analogy and comparing and contrasting with gravity.
  - The arbitrary nature of the definition of zero of potential energy discussed in this section will also be relevant, and may be worth review.
- Unit IV – Chapter 4.1: The Potential Energy of Molecules
  - The idea that two charges infinitely far apart have, by convention, zero potential energy will be relevant as will the idea that bonded electrons/atoms have negative potential energy.
- Unit IV – Chapter 4.2: Application of Bond Energies
  - This is an example problem looking at negative bond energies

A Video Reviewing Problem Solving with Conservation of Energy

This example can be either watched or read



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=201

With what minimum speed must you toss a $140\,\mathrm{g}$ ball straight up to hit the $14\,\mathrm{m}$ meter high roof of a gymnasium if you release the ball $1.3\,\mathrm{m}$ above the ground? With what speed does the ball hit the ground?

You can use conservation of energy to solve this problem.

What is the initial energy state of the ball? We have some kinetic energy and some potential energy, so we have both. How do we know we have kinetic energy? Because we throw the ball, if the ball has no initial Kinetic energy which means it's not moving that means it doesn't go up, it had to have had some kinetic energy for it to actually go up and had to have some initial velocity when we threw it. Does it have any initial potential energy? The ball starts $1.3\,\mathrm{m}$ above the ground initially, this tells me it started out with some potential energy, it's already above the ground.

What is its final energy state in the perfect world in physics land? Does it have Kinetic energy at the roof? No, we're assuming it just touches the roof and has zero velocity at the roof for that moment in time, so its kinetic final energy is actually zero. All we have left is potential final energy.

$$E_i = E_f$$

$$K_i + U_i = K_f + U_f$$

$$\frac{1}{2}mv^2 + mgh_i = 0 + mgh_f$$

$$\frac{1}{2}v^2 + gh_i = gh_f$$

What speed does it hit the ground? Energy initial equals energy final, what's the initial energy state? My initial is the ball at the top of the ceiling. My final is just before it hits the ground. How fast does it hit the ground? It started from the roof, falls down. What is the energy state at the roof? It's all potential. What's the energy state the moment before it hits the ground? It's lost all its potential energy, and its converted into kinetic energy.

$$E_i = E_f$$

$$U_i = K_f$$

$$mgh_i = \frac{1}{2}mv_f^2$$

$$gh_i = \frac{1}{2}v_f^2$$

$$v_f = \sqrt{2gh_i}$$

$$v_f = \sqrt{2 \cdot (9.8\,\mathrm{m/s^2}) \cdot (14\,\mathrm{m})} = 16.6\,\mathrm{m/s}$$

Your friends Frisbee has become stuck 26 meters above the ground in a tree. You want to dislodge the Frisbee by throwing a rock at it. The Frisbee is stuck pretty tight, so you figure the rock needs to be traveling at least 5.4m/s when it hits the Frisbee. If you release the rock 1.6 meters above the ground, with what minimum speed must you throw it?

Energy initial has to equal energy final, what is my initial state of affairs? When I'm throwing the rock, that's my initial state of affairs. Do I have kinetic energy in the beginning? I must have it. How do I know I must have kinetic energy? Because I'm throwing the rock, so the rock has to have some initial velocity. Do I have any initial potential energy? Yes, because I started 1.6 meters above the ground. What's my final state of affairs? Do I have any kinetic energy at the end? When the rocks up there at the frisbee, does it have any kinetic energy? I know that it had to have a velocity, 5.4m/s, I know that the moment before I hit the frisbee I had to have this velocity. Therefore, I know I had some kinetic energy up there. Do I have any final potential energy? Yes, because it is up in the tree.

$$E_i = E_f$$

$$K_i + U_i = K_f + U_f$$

$$\frac{1}{2}mv_i^2 + mgh_i = \frac{1}{2}mv_f^2 + mgh_f$$

$$\frac{1}{2}v_i^2 + gh_i = \frac{1}{2}v_f^2 + gh_f$$

$$v_i = \sqrt{2\left(\frac{1}{2}v_f^2 + gh_f - gh_i\right)}$$

Homework

Homework Problems

Problem 5: Assuming negligible air resistance, what is the final speed of a rock thrown from a bridge?

Problem 6: How many DNA molecules can a single electron from an old-fashioned TV break?

5. Some Energy-Related Ideas that Might be New or are Particularly Important

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 1.

Figure 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 or energy is converted.

Power

Power is the rate at which work is done.

$$P = W/t$$

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: 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 2.)

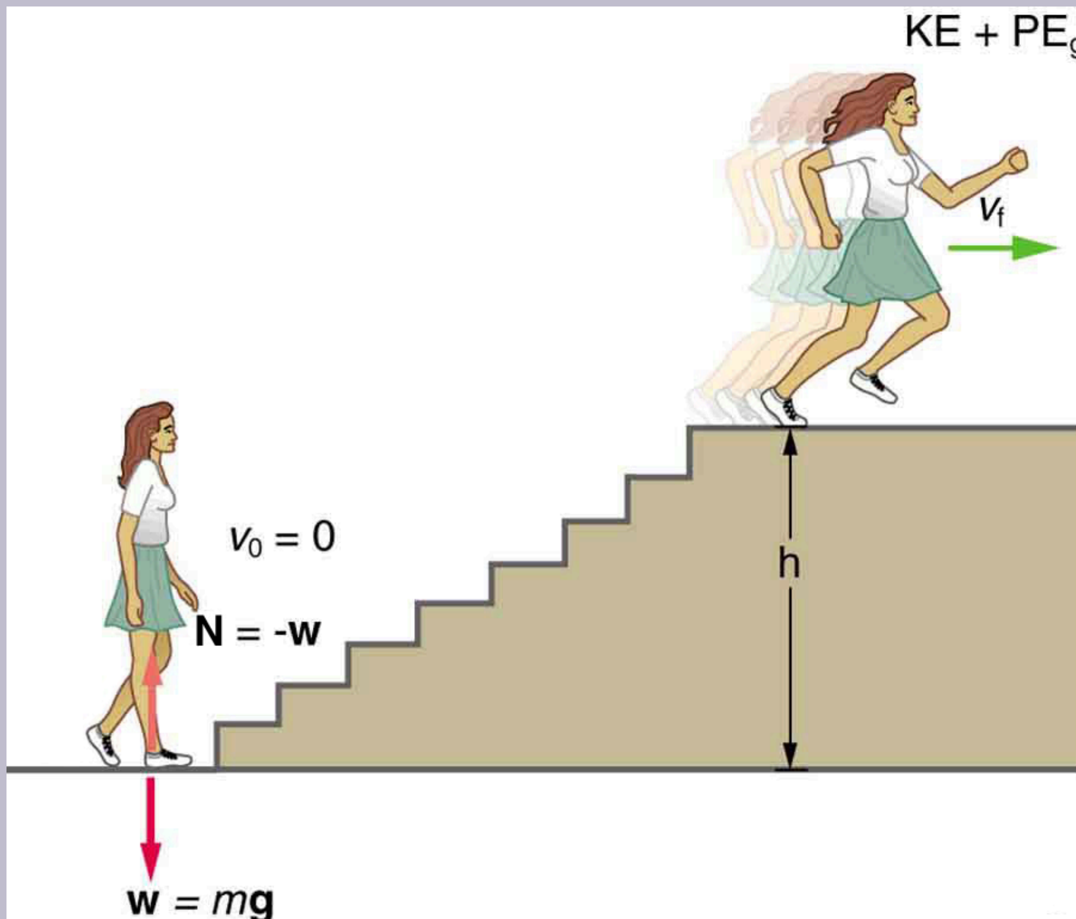


Figure 2: *When this woman runs upstairs starting from rest, she converts the chemical energy originally from food into kinetic energy and gravitational potential energy. Her power output depends on how fast she does this.*

Strategy and Concept

The work going into mechanical energy is

$$\Delta E = Q + W$$

There is no heat transfer in this situation, so $Q = 0$.

$$\Delta E = W$$

$$E_f - E_i = W$$

$$(U_f + K_f) - (U_i + K_i) = W$$

At the bottom of the stairs, we take both $K = 0$ and $U_g = 0$; thus,

$$(0) - (K_f + U_f) = W$$

$$-W = \frac{1}{2}mv^2 + mgh_f$$

where h is the vertical height of the stairs and the minus sign means the energy is leaving her body.Because all terms are given, we can calculate work and then divide it by time to get power.

Solution

Substituting the expression for W in the previous equation, $P = W/t$ yields

$$P = W/t$$

$$P = \frac{\frac{1}{2}mv_f^2 + mgh}{t}$$

Entering known values yields

$$P = \frac{0.5(60.0\,\mathrm{kg})(2.00\,\mathrm{m/s})^2 + (60.0\,\mathrm{kg})(9.80\,\mathrm{m/s^2})(3.00\,\mathrm{m})}{3.50\,\mathrm{s}} = \frac{120\,\mathrm{J} + 1764\,\mathrm{J}}{3.50\,\mathrm{s}} = 538\,\mathrm{W}$$

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 1 for some examples.) Sunlight reaching Earth's surface carries a maximum power of about 1.3 kilowatts per square meter (kW/m^2). This quantity of power per area is called **intensity**, and will be explored more in the chapter on Basics of Waves.

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 W 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 3.)

Figure 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)*

Table 1: Power Output or Consumption

Object or Phenomenon	Power in Watts
Milky Way galaxy	10^{37}
The Sun	4×10^{26}
Volcanic eruption (maximum)	4×10^{15}
Lightning bolt	2×10^{12}
Nuclear power plant (total electric and heat transfer)	3×10^{9}
Aircraft carrier (total useful and heat transfer)	10^{8}
Dragster (total useful and heat transfer)	2×10^{6}
Car (total useful and heat transfer)	8×10^{4}
Football player (total useful and heat transfer)	5×10^{3}
Clothes dryer	4×10^{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	33 size 12{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 $P = W/t = E/t$, where E is the energy supplied by the electricity company. So the energy consumed over a time t is

$$E = Pt$$

Electricity bills state the energy used in units of kilowatt-hours (kW·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.

Calculating Energy Costs

What is the cost of running a 0.200-kW computer 6.00 h per day for 30 days if the cost of electricity is $0.120 per kW·hkW·h?

Strategy

Cost is based on energy consumed; thus, we must find $from[latex]E = Pt$ and then calculate the cost. Because electrical energy is expressed in kW·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·h is

$$E = Pt$$

$$E = (0.200\,\mathrm{kW})(6.00\,\mathrm{h/d})(30\,\mathrm{d})$$

$$E = 36.0\,\mathrm{kW\cdot h}$$

and the cost is simply given by

cost=(36.0 kW·h)($0.120 per kW·h)=$4.32 per month

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.

Homework Problems

Problem 7: How long can you play tennis off a candy bar?

Problem 8: How long for a car of fixed power to get up to speed?

Units of Energy

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

While this is covered in the 131 material linked to in the previous chapter, it is of such importance to this class that we are including it again.

Your Quiz will Cover

- Converting between the different units of energy

In this course, we will be using Joules, electron-Volts exclusively, and kW-hrs exclusively. We are including these other units for your reference.

Note that the eV is significantly smaller than the joule; eV will generally be the smallest unit of energy used in this course.

If energy is defined as the ability to do work, then energy and work must have the same units. Thus, the SI unit of the energy is the Joule (recall 1J=1Nm=1kgms2). Energy, however, is one quantity where there are many other units in common use in scientific literature including electron-Volts (eV), kilowatt-hours (kW·hr), calories, and Calories.

Electron-Volts

A common quantity in chemistry is the electron-Volt or eV. One electron-Volt is the amount of energy gained by an electron as it travels between the two ends of a 1 Volt battery (a concept that will be discussed in more detail when you study electricity). Numerically, $1eV = 1.602\times10^{-19}J$. The reason this unit is common in chemistry is that the energies of atomic bonds are typically about 1eV as shown in the table below. The bond-dissociation energy is the energy released when the bond is formed.

From "Bond-Dissociation Energy – Wikipedia." Accessed August 1, 2017. https://en.wikipedia.org/wiki/Bond-dissociation_energy.

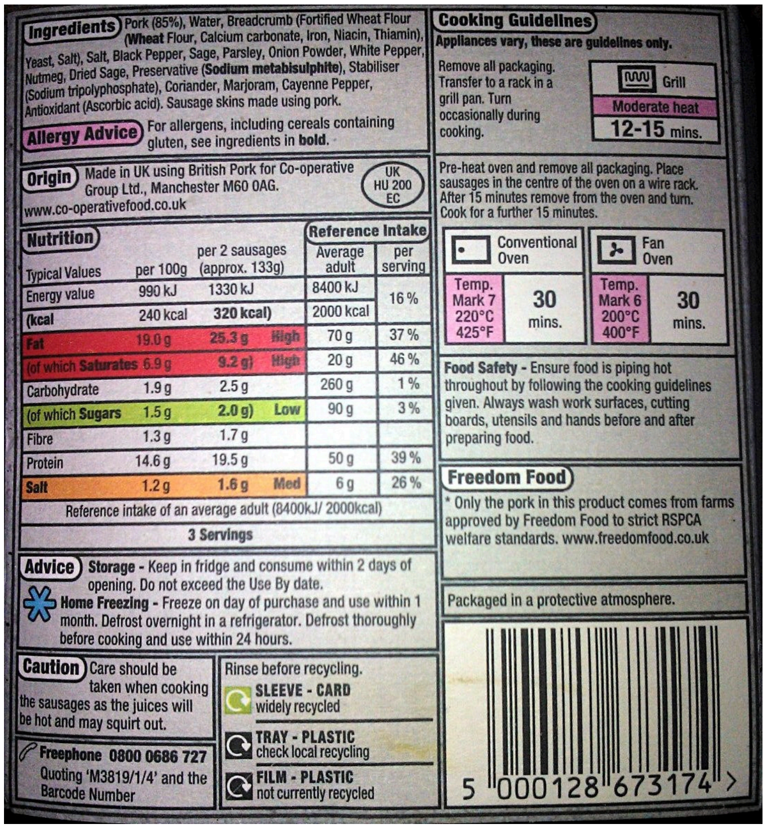
Bond	Bond-dissociation energy at 298K (eV/Bond)	Comment
C-C	3.60-3.69	Strong, but weaker than C–H bonds
Cl-Cl	2.51	Indicated by the yellowish colour of this gas
H-H	4.52	Strong, nonpolarizable bondCleaved only by metals and by strong oxidants
O-H	4.77	Slightly stronger than C–H bonds
OH-H	2.78	Far weaker than C–H bonds
C-O	11.16	Far stronger than C–H bonds
O-CO	5.51	Slightly stronger than C–H bonds
O=O	5.15	Stronger than single bondsWeaker than many other double bonds
N=N	9.79	One of the strongest bondsLarge activation energy in production of ammonia
H3C-H	4.550	One of the strongest aliphatic C–H bonds

Kilowatt Hours and Calories

Ad described above, when you buy electricity from the power company, the bill says how many kilowatt hours you have purchased. A Watt is a unit of a quantity called ***power*** and 1 Watt is equal to 1 Joule/second: 1W = 1 J/s. Thus, a kilowatt hour is therefore:

(1kW∗hr)(1000WkW)(1J/sW)(3600shr)=3.6106J=3.6MJs(1kW∗hr)(1000WkW)(1J/sW)(3600shr)=3.6106J=3.6MJsleft (1 kW * hr right) left ({1000 W} over {kW} right) left ({1 J/s} over {W} right) left ({3600 s} over {hr} right) =3.6106 J=3.6 MJ s">(1kW∗hr)(1000WkW)(1J/sW)(3600shr)=3.6106J=3.6MJs

The calorie is an imperial unit of energy that is still in common use in the nutritional sciences in the United States. One calorie (lowercase c) is the amount of energy needed to raise 1g of water 1^{O}C or 1 cal = 4.814J. On food labels, you will see energy listed in Calories (capital C). One Calorie is equal to 1kilocalorie; in other words, 1 Cal = 1000 cal. Thus, one 1 Cal = 4814 J. In other countries, you will see food labels in both Calories and Joules like the one shown in Figure 1.



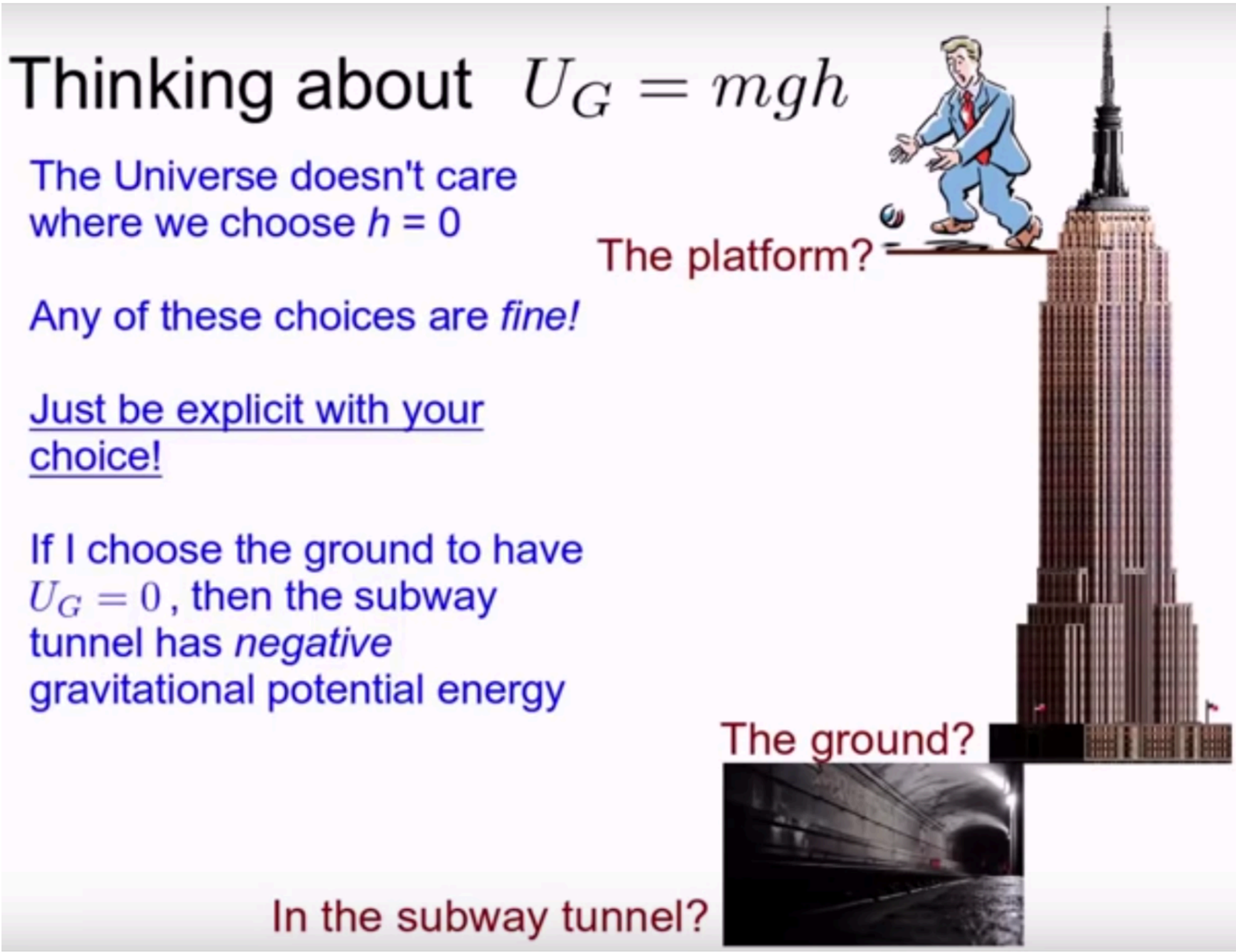
A food label from the UK showing the energy of the food in both Joules and kcal (or Calories).

The Potential Energy of Electrons in Atoms and Molecules

For this first unit, the primary source of microscopic potential energy with which we shall concern ourselves is the potential energy stored be electrons in atoms and in chemical bonds: U_{chem}. This potential energy is a result of the force of electrical attraction between different atoms (recall electricity and magnetism was one of our fundamental forces) or between an electron and its nucleus.

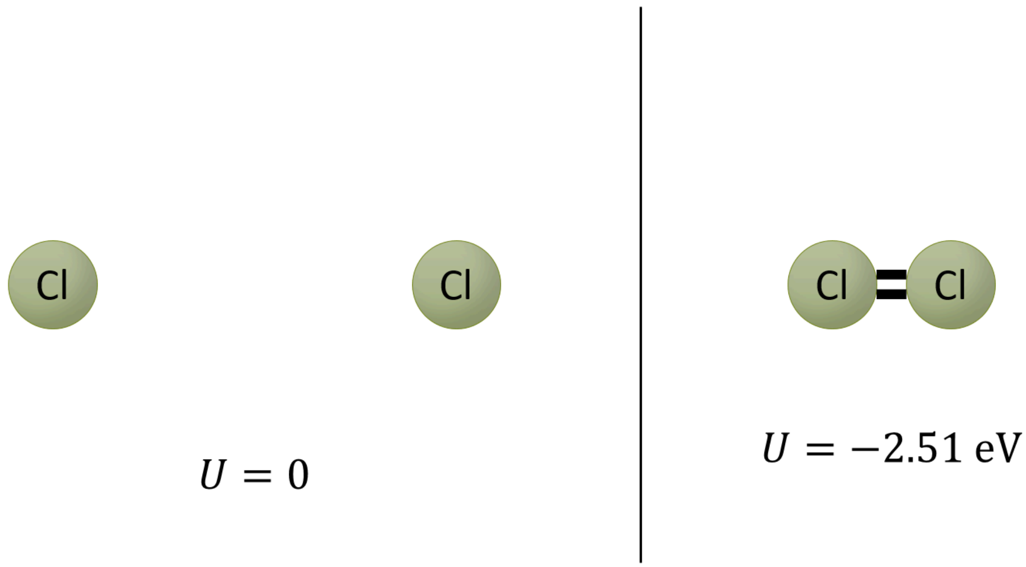
The strength of chemical bonds is typically quoted in one of two ways: either the energy in the bond, called the *bond dissociation energy*, is quoted directly (typically in eV) or the enthalpy per mole will be quoted. For example, the Cl-Cl bond has a bond dissociation energy of 2.51 eV/bond or a bond dissociation enthalpy per mole of ΔH=242kJ/molΔH=242kJ/molΔH=242 kJ/mol">ΔH=242kJ/mol. For atoms, the relevant energy is the ionization energy, which is the amount of energy needed to remove an electron: quoted directly in eV or, occasionally in kJ/mol. How do we interpret these numbers in terms of potential energy? We use the same freedom to choose the zero of potential energy discussed in [Unit IV – Chapter 3.3: Macroscopic Potential Energy] from the Physics 131 textbook which discusses potential energy at the macroscopic scale.

Thinking about gravity, we tend to put the zero of potential energy at ground level; objects above the ground then have positive potential energy while objects underground have negative potential energy. This use of negative potential energy makes sense, an object at ground level will fall to below ground level if allowed to do so and lose energy in the process.



Where to put the zero of gravitational potential energy? The top of a building? The ground? In the subway below? The choice is arbitrary.

For atoms and molecules, we have a similar freedom to choose where to put zero potential energy. The standard convention is to say that free atoms that are far apart have *zero* potential energy. Atoms in most bonds have lower potential energy than free atoms (that is why the bonds form!). Therefore the potential energy of the atoms is less than zero: the potential energy of atoms in bonds is *negative*. This may seem like a weird choice for the zero of potential energy, but it is the convention and it makes sense when you think about it!



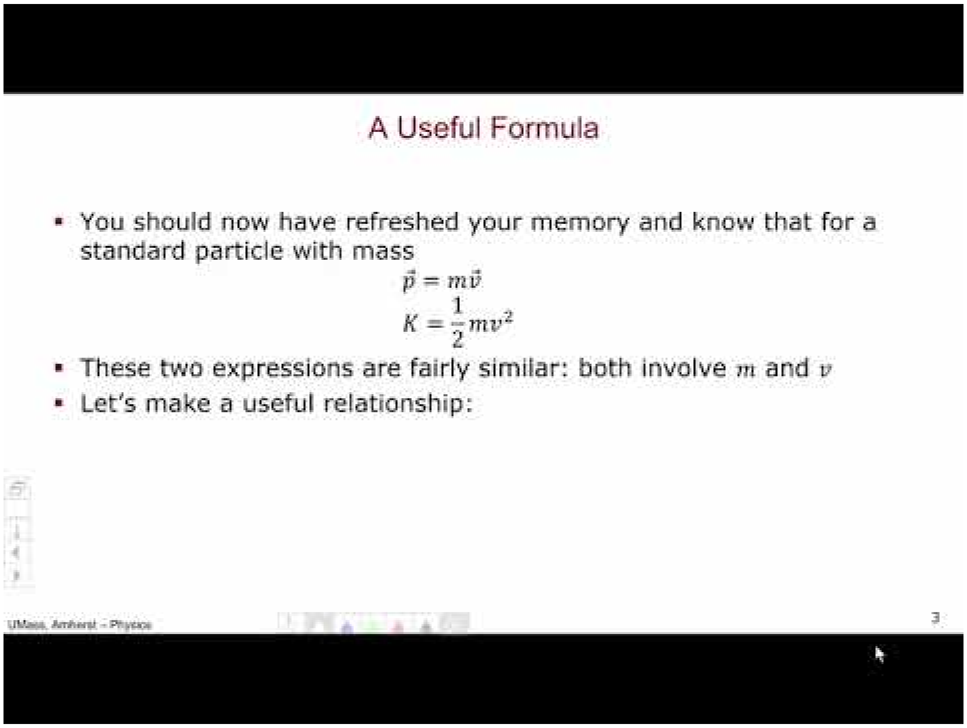
Two Cl atoms separated by a great distance have zero potential energy while two bonded Cl atoms have a potential energy of -2.51 eV. Remember, potential energy is the due to the relative position of two objects, so it does not make sense to ask which atom in the bonded pair has the potential energy. The potential energy is due to the two of them!

Let's return to the quoted Cl-Cl bond with dissociation energy of 2.51 eV/bond. What does this value mean? It means that two Cl atoms bonded together have a potential energy of 2.51 eV *less* than if they were free. Said another way, the potential energy of Cl atoms in Cl_2 is -2.51 eV, while the potential energy of free Cl atoms is 0 eV. This is consistent with what you probably already know about Chlorine: Cl_2 is the lower energy state than free Cl atoms. I would get 2.51 eV of energy for every Cl-Cl bond that is formed, as the atoms move from zero potential energy to -2.51 eV. Similarly, I would need to add 2.51 eV of energy to break a Cl-Cl bond and move the two atoms *up* to zero potential energy. The same idea holds for electrons in atoms. If you look up the ionization energy of hydrogen, you will see 13.6 eV. This value meas that the electron has 13.6 eV *less*potential energy than if it were free. Thus, a more accurate way to write this energy would be U = -13.6 eV.

Homework Problem

Problem 9: Which element has an outer electron with the lowest potential energy?

The Connection Between Kinetic Energy and Momentum



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=205

We will now develop a useful relationship between momentum and kinetic energy. This is a useful relationship that we will use throughout this course.

By now you should have refreshed your memory and know that for a standard particle with mass, such as an electron, the momentum of the particle is

$$\vec{p} = m\vec{v}$$

and the kinetic energy of the particle is

$$K = \frac{1}{2}mv^2.$$

If you look at these two expressions, they are fairly similar, both involve the mass of the particle m and its velocity v.

Now there are some important differences. The momentum is a vector, including the direction of motion, whereas the kinetic energy is a scalar and is independent of the particle's direction of motion. However, there's a useful way to relate these two.

Begin with the magnitude of the momentum, removing the vectors, in which case is just

$$p = mv.$$

Square both sides of this expression so you have

$$p^2 = m^2v^2.$$

Now divide both sides of the expression by $2m$, so now you have

$$\frac{p^2}{2m} = \frac{1}{2}mv^2,$$

Which is the kinetic energy.

The big punch line is that the kinetic energy of a particle with mass is
$K = \frac{p^2}{2m}$.
This is a useful expression that we'll be using throughout this course.

Homework Problem

Problem 10: Exploring the relationship between momentum and kinetic energy.

6. Basics of Waves

What is a Wave?

A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=203

Now, we have talked about particles. What about waves? Before we start talking about waves, it's probably best to give a few different examples of waves. If I ask you to think of a wave the first thing that probably would come to most your minds is a water wave, but we could also have waves on a string, or even sound waves. The most generic picture that a lot of you have, is probably some sort of sine or cosine shape traveling along, but this is not representative of all waves and we want our definition to be in terms of properties that apply to every possible wave that we can think of.

Let's go through a few questions and develop a definition of a wave.

Does the wave actually have to go anywhere, does a wave travel? No. Sure, most waves go somewhere, water waves travel across an ocean for example, but think of a guitar string, when you pluck it, certainly the string waves back and forth but the string doesn't go anywhere, the wave stands on the string. This is called a standing wave. So, traveling cannot be part of our definition of a wave.

Does a wave have to be a repeating pattern? Again, not really. While this might be the image that a lot of you have in mind when I say the word wave, remember we can have just a single pulse going back and forth on a string.

Does the wave have to have up and down motion? Well again, no. The standard picture of a wave that you have in your head might look like Figure 1, but I could also send a compression wave down the slinky like in figure 2 where the links of the slinky move back and forth in the same direction as the waves motion.

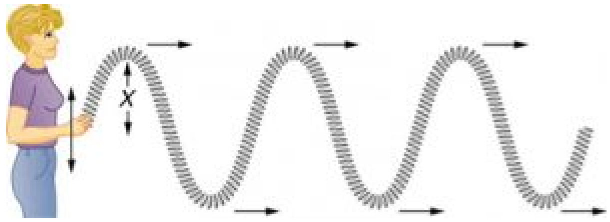


Figure 1: *A wave can go up and down – a transverse wave. This is probably what you are picturing when I say wave.*

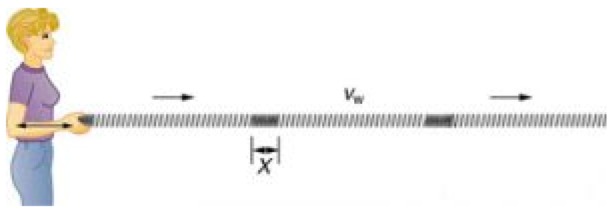


Figure 2: *A wave can go down the slinky as a compression – or longitudinal – wave.*

Now we need a little bit of terminology. Waves that do wiggle perpendicular to the direction of motion of the wave are called transverse waves. These are the waves that you probably have in mind and these are the ones that were mostly going to be interested in.

The basic terminology of transverse waves, we'll introduce some more later, are that waves have a peak and a trough, and then the distance from the zero line to either a peak or a trough is called the amplitude. This amplitude is labelled y in Figure 3.

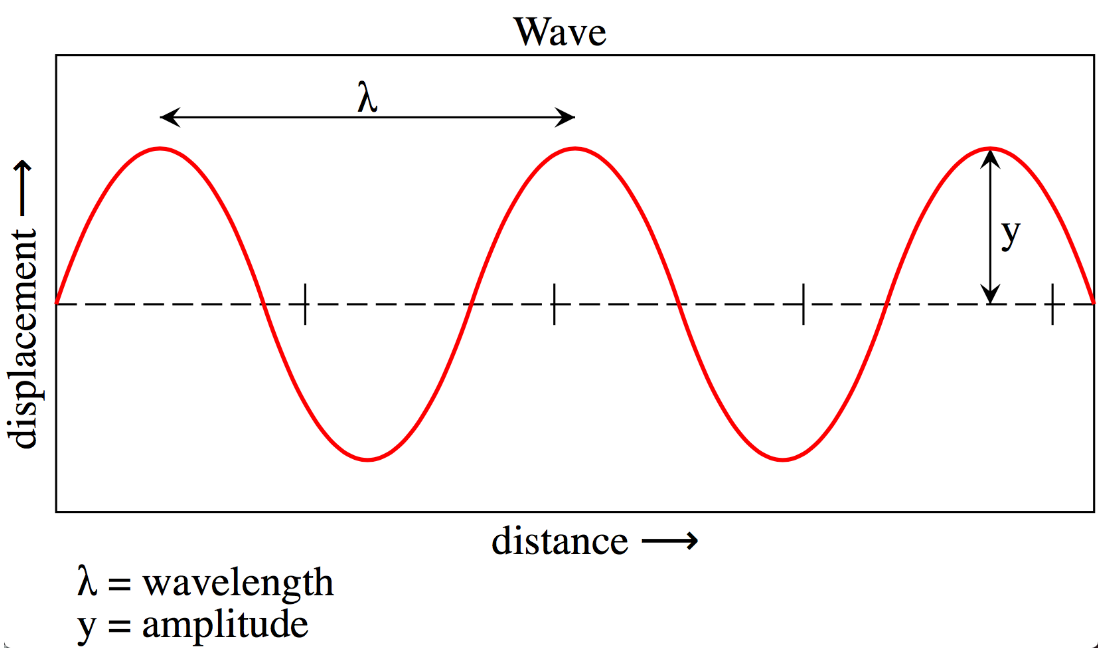


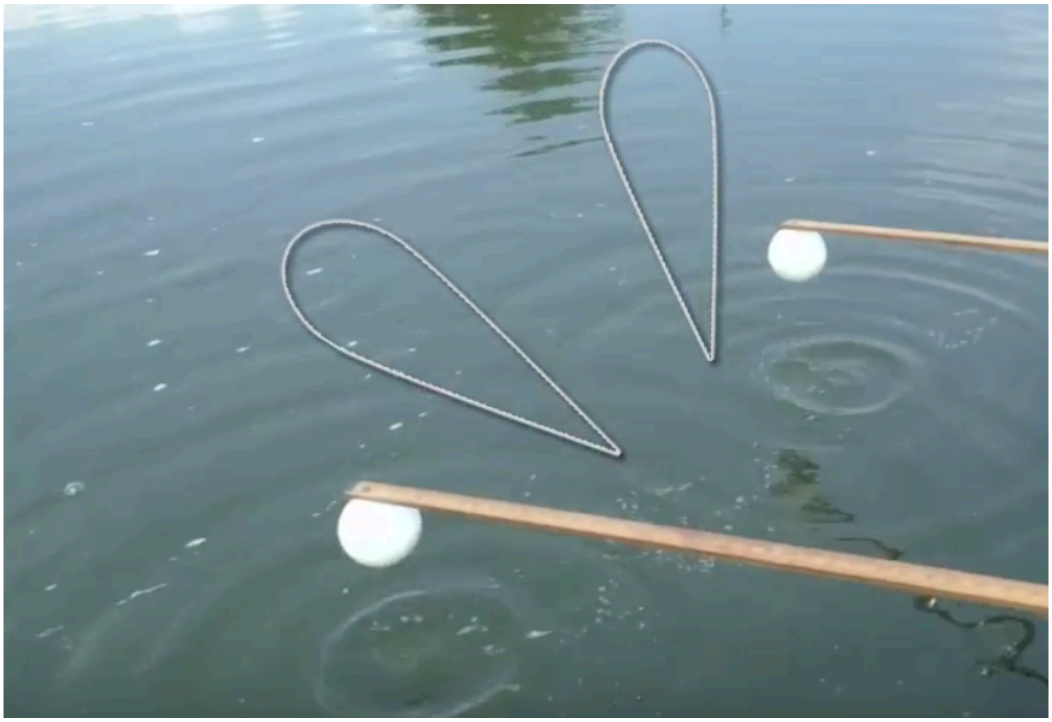
Figure 3: *Parts of a wave (Credit: Krishnavedala)*

What other properties of a wave could we perhaps use? Can a wave bend around corners? We know that particles don't bend around corners, what about waves? Well, it turns out that waves do bend around corners. Think of a water wave, it spreads out, bending around the corner. In the video on the next slide we'll see some other important properties of waves that all waves do share.

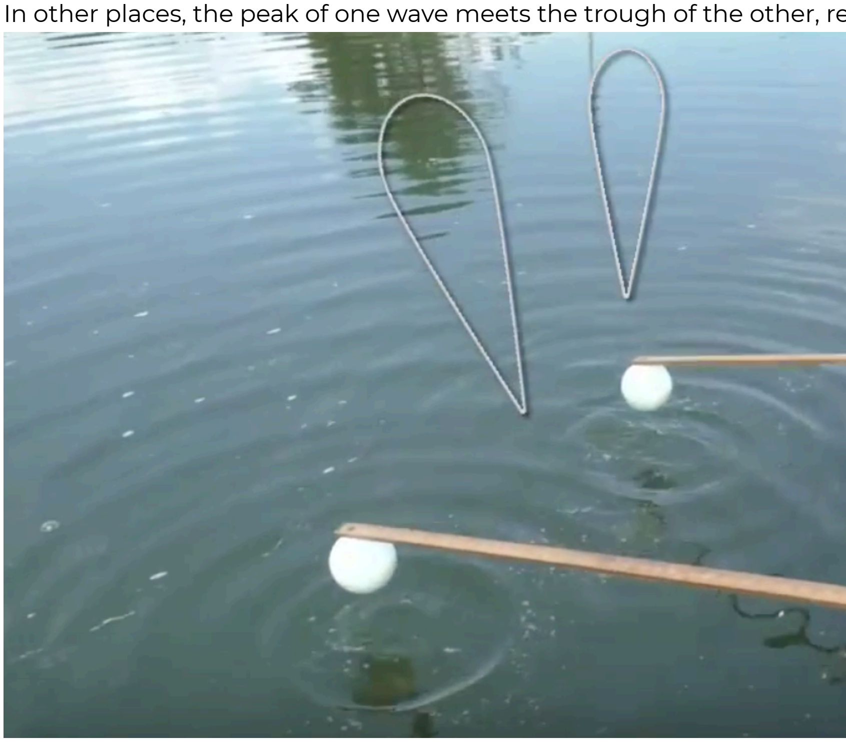
I go and just hit the water with a single stick, we see we get waves coming out in all directions, radiating away from the spot where the ball hits the water.

Things get a little bit more interesting, however, if we have two sources of waves going at the same time next to each other like so.

So now we get two waves, each radiating out from its source. In some places the waves line up peak to peak, or trough to trough, and add up, resulting in a larger wave at that point.



In other places, the peak of one wave meets the trough of the other, resulting in some cancellation.



This phenomenon of waves adding in some places and cancelling in others is known as interference and is a characteristic property of waves.

So, what is a wave? Well a wave is a disturbance that can, but doesn't necessarily have to, travel or it can just store energy and momentum. A traveling wave will carry energy from one position to another, think of the water wave that carries energy as it moves across the ocean and also momentum, as that wave hits you, you feel the momentum of the wave. For a standing wave, that energy is just being stored. When I pluck a guitar string, the energy is just being stored in the string and then ultimately releases this sound that we hear. A wave need not necessarily repeat, we can have simple pulse waves. But a wave can bend around corners, and waves of the same kind can interact with each other or with themselves, adding in some places and canceling in other places, through this idea of interference. These are the fundamental characteristics of waves. They don't exist at a particular place, they sort of spread out over a couple of different places and they can carry energy and momentum while bending around corners and interacting with themselves or other waves of the same kind.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

To Summarize:

- Particles are localized in space, they don't bend around corners, but can carry energy and momentum.
- Waves on the other hand, are spread out in space, they are some kind of disturbance that can transfer or store energy and momentum.
- However, waves, unlike particles, can bend around corners and also waves can interact with themselves or other waves of the same kind through this phenomenon of interference.

Period and Frequency in Oscillations

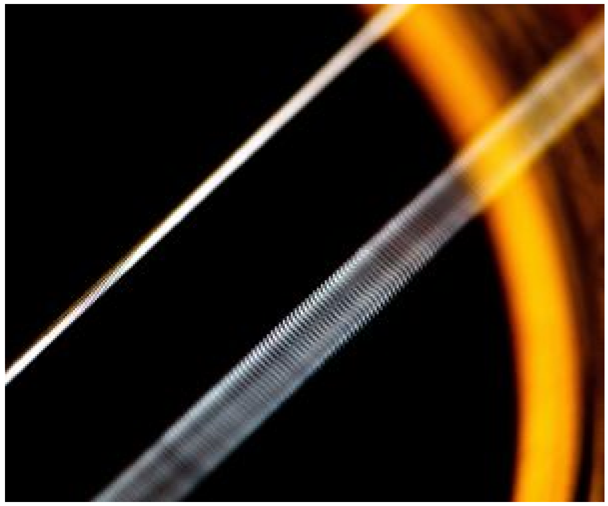


Figure 2: *The strings on this guitar vibrate at regular time intervals. (credit: JAR)*

When you pluck a guitar string, the resulting sound has a steady tone and lasts a long time. Each successive vibration of the string takes the same time as the previous one. We define periodic motion to be a motion that repeats itself at regular time intervals, such as exhibited by the guitar string or by an object on a spring moving up and down. The time to complete one oscillation remains constant and is called the period T. Its units are usually seconds, but may be any convenient unit of time. The word period refers to the time for some event whether repetitive or not; but we shall be primarily interested in periodic motion, which is by definition repetitive.

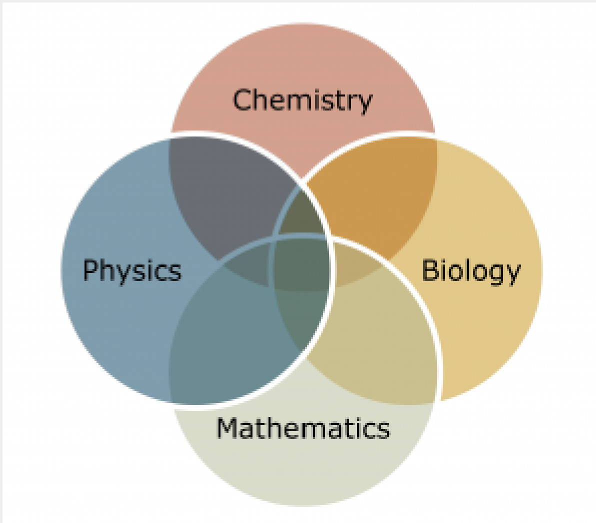
A concept closely related to period is the frequency of an event. For example, if you get a paycheck twice a month, the frequency of payment is two per month and the period between checks is half a month. Frequency ν is defined to be the number of events per unit time. For periodic motion, frequency is the number of oscillations per unit time. The relationship between frequency and period is

$$\nu = \frac{1}{T}.$$

The SI unit for frequency is the cycle per second, which is defined to be a hertz (Hz):

$$1\,\mathrm{Hz} = \frac{1\,\mathrm{cycle}}{\mathrm{s}} \text{ or } 1\,\mathrm{Hz} = \frac{1}{s}$$

A cycle is one complete oscillation. Note that a vibration can be a single or multiple event, whereas oscillations are usually repetitive for a significant number of cycles.



The different sciences use different symbols for frequency. If you have seen waves in a physics course before, they probably used the symbol f. However, in your chemistry classes, you probably used ν. Both are in use and mean the same quantity. We will, in general, stick to using the ν you used in chemistry, but don't worry if you see an f used somewhere in your homework or in the text, it means the same quantity.

Fun with the history of science, different disciplines discovered the same quantity and gave it different names!

Determine the Frequency of Two Oscillations: Medical Ultrasound and the Period of Middle C

We can use the formulas presented in this module to determine both the frequency based on known oscillations and the oscillation based on a known frequency. Let's try one example of each.

(a) A medical imaging device produces ultrasound by oscillating with a period of 0.400 µs. What is the frequency of this oscillation?

(b) The frequency of middle C on a typical musical instrument is 264 Hz. What is the time for one complete oscillation?

Strategy

Both questions (a) and (b) can be answered using the relationship between period and frequency. In question (a), the period T is given and we are asked to find frequencyν. In question (b), the frequencyν is given and we are asked to find the period T.

Solution (a):

Substitute 0.400µs for T in $\nu = \frac{1}{T}$:

$$\nu = \frac{1}{T} = \frac{1}{0.40010^{6}\ \mathrm{s}}.$$

Solve to find

$$\nu = 2.5010^6 \text{ Hz}$$

Discussion (a):

The frequency of sound found in (a) is much higher than the highest frequency that humans can hear and, therefore, is called ultrasound. Appropriate oscillations at this frequency generate ultrasound used for noninvasive medical diagnoses, such as observations of a fetus in the womb.

Solution (b):

Identify the known values:

The time for one complete oscillation is the period T:

$$\nu = \frac{1}{T}.$$

Solve for T:

$$T = \frac{1}{\nu}.$$

Substitute the given value for the frequency into the resulting expression:

$$T = \frac{1}{264 \text{ Hz}}$$

$$T = 3.7910^3 \text{ s} = 3.79 \text{ ms}.$$

Discussion (b)

The period found in (b) is the time per cycle, but this value is often quoted as simply the time in convenient units (ms or milliseconds in this case).

Everyday Periods and Frequencies

Identify an event in your life (such as receiving a paycheck) that occurs regularly. Identify both the period and frequency of this event.

Solution

I visit my parents for dinner every other Sunday. The frequency of my visits is 26 per calendar year. The period is two weeks.

Section Summary

- Periodic motion is a repetitious oscillation.
- The time for one oscillation is the period T
- The number of oscillations per unit time is the frequency ν (or sometimes f).

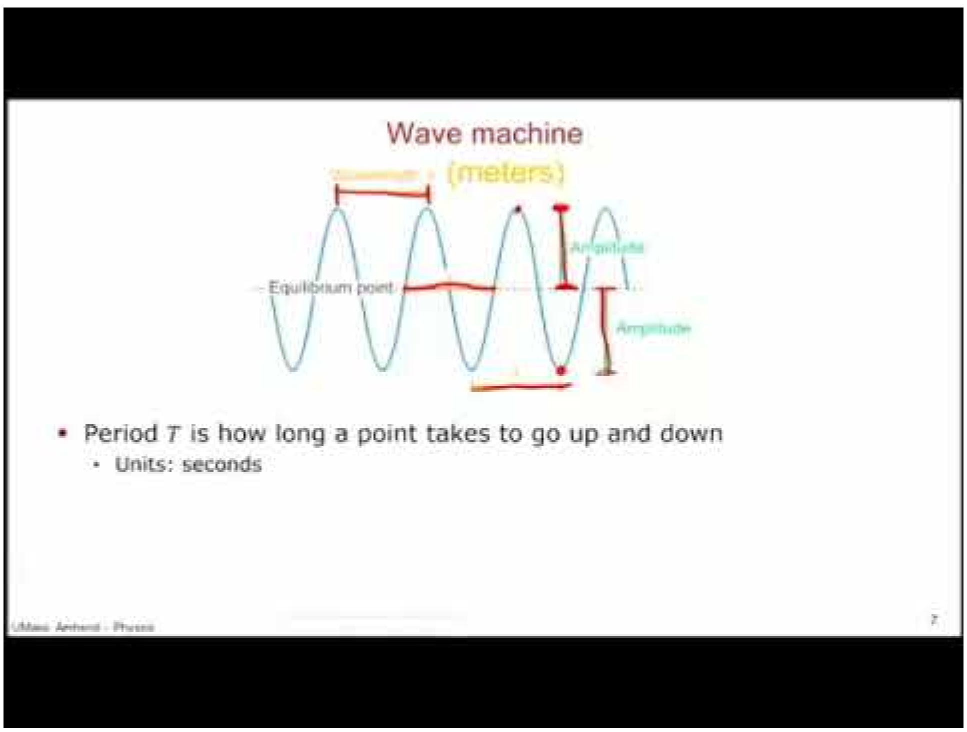
- These quantities are related by $\nu = \frac{1}{T}$.

Homework

Problem 11: What is the frequency of a stroboscope?

If you are curious about what a stroboscope is, check out [The Stroboscope Wikipedia page](https://en.wikipedia.org/wiki/Stroboscope).

Detailed description of a wave



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=203

Let's begin by thinking about the one thing we have already figured out, which is that light has some wave-like properties. In the figure below we have a wave, and we can talk about the wavelength, λ, which is the distance from one point on the wave through the same point as shown in Figure 4. Wavelength could be from peak to peak or from zero to zero or from trough to trough. All those distances are the same, and the wavelength is measured in meters.

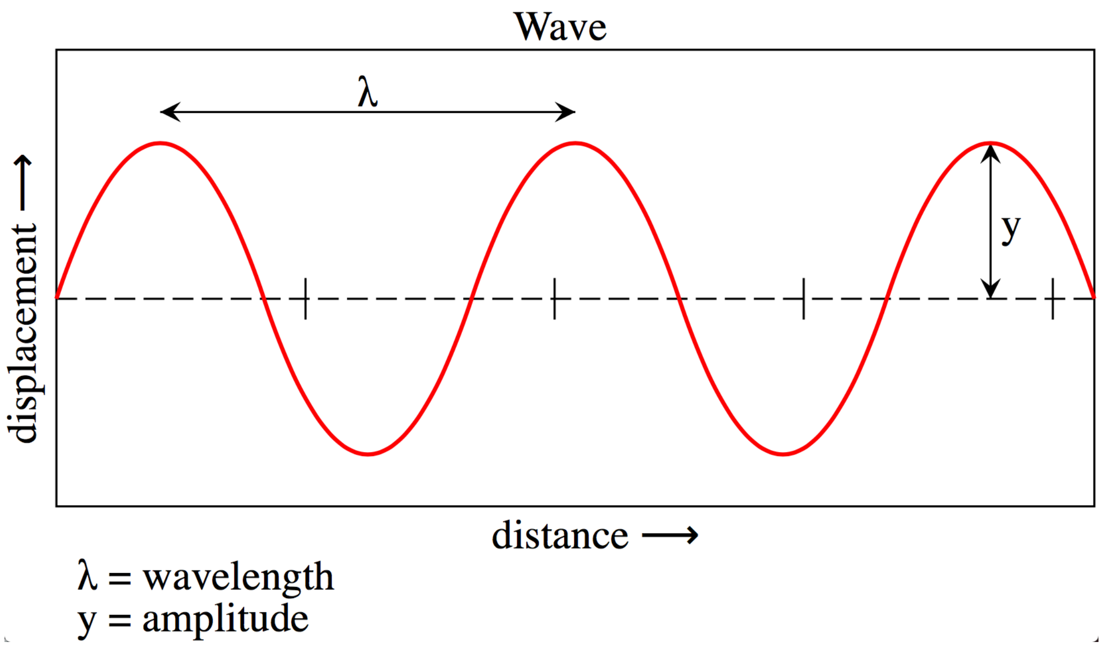


Figure 4: *Parts of a wave (Credit: Krishnavedala*

We also have the idea of the amplitude y in the figure. The difference from either peak to the middle average or from trough to the middle average is what we call the amplitude.

In addition to wavelength and amplitude, we can also talk about how long it takes for a point to go up and down. Think about one point on the wave, bouncing up and down. We can talk about how long it takes for that point on the wave to go from trough to peak to trough, the amount of time that takes is called the period, T, and, as discussed in the last section, will be measured in seconds.

For this class, we will work in SI, International System of Units, we will not work in, what I call barbaric units, there shall be no inches. Meters, kilograms, and seconds will be the norm. On the Moodle page there's a document of math I expect you to know, it's things like trigonometry and area of a circle, but I also expect you to know the SI prefixes Nano – Giga. I will not give you these, and on an exam if you come to the TA's and ask how big a micrometer is, we're going to have to say tough cookie, that's something you were supposed to know.

The period is the time it takes for a point on the wave to go up and down, measured in seconds, but I can also count how many that point oscillates in one second. I could say how many oscillations per second this point on this wave make? That quantity, again looking back to the last section, is known as the frequency ν and its value is

$$\nu = \frac{1}{T},$$

the unit of frequency is how many per second or one over seconds, which is called Hertz: Hz.

We will use these basic terms for all of the waves we discuss, electrons and light. We know that wavelength is measured in meters and we know that frequency is in Hertz, or 1 over seconds, so $\lambda\nu$ will be meters over seconds which is a velocity. What's the only velocity that we could have? The speed of the wave, it's the only speed we could talk about, so we have the speed of the wave is going to be

$$v = \lambda\nu$$

Exploring v = λν

If we know that the speed of a wave is fixed, for example, light travels at a fixed speed, then if the frequency goes up, what's the wavelength going to do?

Solution:

We know that

$v = \lambda\nu.$

The speed is fixed, if the frequency goes up, and λ times ν has to be the same thing, then that is going to tell me that the wavelength must go down.

Discussion:

This question is based on mathematical reasoning using symbols and not numbers. Remember one of the goals for this course was working in symbols and not numbers, so here's an example of that.

Homework Problems

Problem 12: Label the parts of a wave.

Problem 13: If the frequency of a wave is changed, which of the other properties must also change assuming the speed of the wave remains fixed?

Problem 14: Speed, wavelength, and frequency for sound.

Energy in Waves: Intensity

Instructors Notes

The key takeaways that you will be potentially quizzed on are:

- Intensity is power per area: $I = \frac{P}{A}$ with units W/m^2
- The intensity is related to the square of the amplitude $I \propto A^2$

The energy effects of a wave depend on time as well as amplitude. For example, the longer deep-heat ultrasound is applied, the more energy it transfers. Waves can also be concentrated or spread out. Sunlight, for example, can be focused to burn wood. Earthquakes spread out, so they do less damage the farther they get from the source. In both cases, changing the area the waves cover has important effects. All these pertinent factors are included in the definition of intensity I as power per unit area:

$$I = \frac{P}{A}$$

where P is the power carried by the wave through area A. The definition of intensity is valid for any energy in transit, including that carried by waves. The SI unit for intensity is watts per square meter (W/m^2). For example, infrared and visible energy from the Sun impinge on Earth at an intensity of 1300 W/m^2 just above the atmosphere.

Calculating intensity and power: How much energy is in a ray of sunlight?

The average intensity of sunlight on Earth's surface is about 700 W/m^2.

(a) Calculate the amount of energy that falls on a solar collector having an area of 0.500m^2 in four hours.

(b) What intensity would such sunlight have if concentrated by a magnifying glass onto an area 200 times smaller than its own?

Strategy a

Because power is energy per unit time or

$$P = E/t$$

the definition of intensity can be written as

$$I = \frac{P}{A} = \frac{E}{tA}$$

and this equation can be solved for E with the given information.

Solution a

1. Begin with the equation that states the definition of intensity:
$I = \frac{P}{A}$
2. Replace
P with its equivalent E/t:
$I = \frac{E}{tA}$
3. Solve for E:
$E = IAt$
4. Substitute known values into the equation:
$E = (700\,\mathrm{W/m^2})(0.5\,\mathrm{m^2})(4\,\mathrm{hr} \times 3600\,\mathrm{s/hr})$
5. Calculate to find *E:*
$E = 5.04 \times 10^6\,\mathrm{J} = 5.04\,\mathrm{MJ}$

Discussion a

The energy falling on the solar collector in 4 hours is enough to be useful—for example, for heating a significant amount of water.

Strategy b

Taking a ratio of new intensity to old intensity and using primes for the new quantities, we will find that it depends on the ratio of the areas. All other quantities will cancel.

Solution b

1. Take the ratio of intensities, which yields:
$\frac{I'}{I} = \frac{P'/A'}{P/A} = \frac{A}{A'}$
The powers cancel because, as the source is always the sun, the power (energy per time) is the same: $P = P'$
2. Identify the knowns:
$A = 200A'$ as the area of the magnifying glass is 200x smaller.
$\frac{I}{I'} = \frac{200A'}{A'} = 200$
3. Substitute known quantities:
$I' = 200I$
$I = 200(700\,\mathrm{W/m^2}$
4. Calculate to find:
$I' = 1.40 \times 10^5\,\mathrm{W/m^2}$

Discussion b

Decreasing the area increases the intensity considerably. The intensity of the concentrated sunlight could even start a fire.

Determine the combined intensity of two waves: Perfect constructive interference

If two identical waves, each having an intensity of $1.00 W/m^2$, line up perfectly peak-to-peak, what is the intensity of the resulting wave?

Strategy

If two waves, which have equal amplitudes A, line up exactly, the resulting wave has an amplitude of $2A$. Because a wave's intensity is proportional to amplitude squared, the intensity of the resulting wave is four times as great as in the individual waves.

Solution

1. Recall that intensity is proportional to amplitude squared.
2. Calculate the new amplitude:
   $I' \propto A'^2$
   $A' = 2A \leftarrow$ The new amplitude is twice the original
   $I' \propto 2A^2 = 4A$
3. Recall that the intensity of the old amplitude was: $I \propto A^2$
4. Take the ratio of new intensity to the old intensity. This gives:
   $\frac{I'}{I} = \frac{4A}{A} = 4$
5. Calculate to find I'
   $I' = 4I$
   $I' = 4(1\,\mathrm{W/m^2} = 4\,\mathrm{W/m^2}$

Discussion

The intensity goes up by a factor of 4 when the amplitude doubles. This answer is a little disquieting. The two individual waves each have intensities of $1.00 W/m^2$, yet their sum has an intensity of $4.00 W/m^2$, which may appear to violate conservation of energy. This violation, of course, cannot happen. What does happen is intriguing. The area over which the intensity is $4.00 W/m^2$ is much less than the area covered by the two waves before they interfered. For each spot where the waves line up peak-to-peak, there is somewhere else where they line up peak-to-trough and cancel. For example, if we have two stereo speakers putting out $1.00 W/m^2$ each, there will be places in the room where the intensity is $4.00 W/m^2$ and other places where the intensity is zero, and other places where the intensity is in between. Figure 5 shows what this interference might look like.

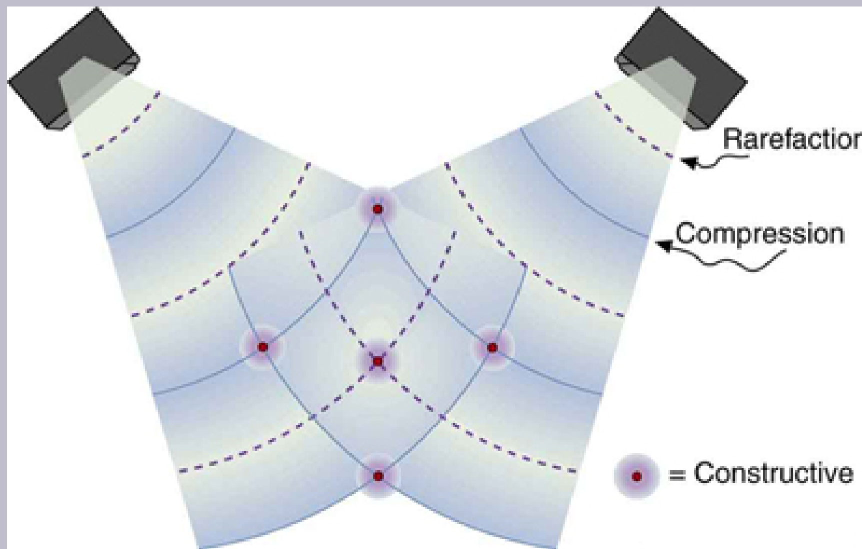


Figure 5: *These stereo speakers produce both constructive interference and destructive interference in the room, a property common to the superposition of all types of waves. The shading is proportional to intensity.*

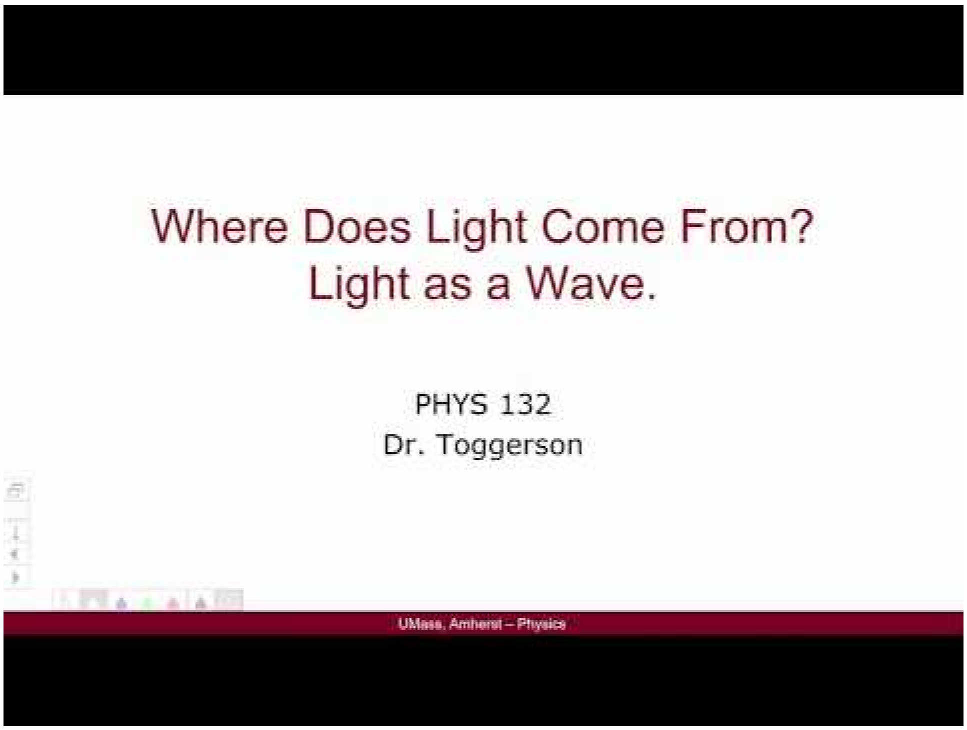
Homework

Problem 15: How long to collect a certain amount of sunlight?

Problem 16: What is the power output of an ultrasound machine for a given intensity and area?

7. Basics of Light

Where Does Light Come From?



A YouTube element has been excluded from this version of the text. You can view it online here: *http://openbooks.library.umass.edu/toggerson-132/?p=207*

Light is generated any time a charge undergoes acceleration; this is a connection to an idea from Physics 131. Just like in Physics 131 it's not the motion of the charge that matters, but its acceleration. Moving charges don't generate light only accelerating ones do. To expand upon this connection to 131 a little bit more, if a charge accelerates by slowing down, it is still accelerating then from Newton's second law,

$$\sum \vec{F} = m\vec{a}$$

, we know that a force has acted upon it. If it takes some distance for this slowing down to occur, then the force must have been applied for some distance and we know that work was therefore done on the charged particle. By the statement of conservation of energy, or equivalently the first law of thermodynamics, if work is done on a particle then the particles energy must change, that energy must go somewhere and where does it often go? It goes into light.

Here's an example with which you might be familiar from your chemistry class. An electron in an outer energy level of an atom falls to a lower energy level. There's a change in energy as the electron falls, that energy has to go somewhere. It goes into the release of light.

Electrons changing energy levels, however, is not the only way to produce light. Think about an old-school incandescent lamp with the filament in it that get hot as you turn them on, to understand why these incandescent lights give off light we have to understand a little bit about what temperature is.

Recall from Physics 131 that temperature is related to the average kinetic energies of particles moving around randomly on the atomic and subatomic scales. As these particles are bouncing around randomly, they're changing directions. From 131 we know that acceleration is a vector, so because velocity changes direction, then we know that there is acceleration. So once again, even any object with temperature will emit light due to the accelerating charges bouncing around on the atomic and subatomic scale.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

In summary

- Light is generated by charges accelerating.
- Every object with a temperature (i.e. everything) will emit some amount of light of some type.
- Our eyes, however, are only sensitive to certain kinds of light and we therefore cannot see this light from everyday objects such as you and I. We don't see light coming off of us because our eyes are not sensitive to the kind of light that we emit due to our temperature.
- However, we can build devices that can see the light given off by more everyday objects such as people by using technologies such as infrared cameras.

Homework

Problem 17: Which situations will create electromagnetic radiation?

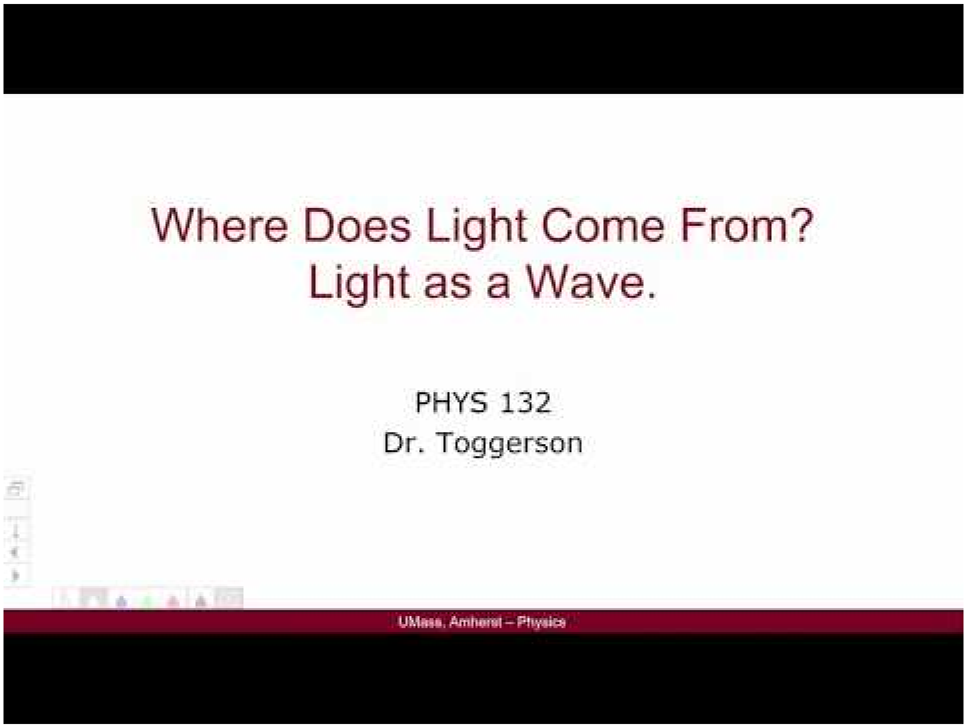
Properties of Light

University *of* Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

The video for this section uses f for frequency. The text, on the other hand uses ν. This is a good example of the fact that you need to get used to the idea that different disciplines use different letters for the same quantity!

On your equation sheet, in class, an on exams, we will use ν to be consistent with what you have used in chemistry.



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=207*](http://openbooks.library.umass.edu/toggerson-132/?p=207)

Like all waves, light waves are characterized by a **wavelength**, a **frequency**, a speed, which follows the usual relationship of $v = \lambda\nu$, and an **amplitude**. However, there are some important unique characteristics of light waves. For light, the wave in the vacuum speed is always the same, $c = 3 \times 10^8\,\mathrm{m/s}$. In a vacuum $v = \lambda\nu$, turns into $c = \lambda\nu$, because all light waves, regardless of their wavelength or frequency or amplitude, travel at this same fundamental speed.

For the amplitude of the light wave we will not use the symbol A we will instead use the symbol E and the amplitude of a light wave has the units of Newton's per Coulomb $\mathrm{N/C}$, Newton's are the unit of force and Coulomb, as you've already discussed elsewhere in your prep, is the unit of a charge. The amplitude of a light wave is a Newton per Coulomb. We will see why this is the unit of a light waves amplitude later in this particular course, but for right now you just need to know that those are the units.

There are many different kinds of light. Where do these different kinds of light come from? Well different wavelengths or frequencies represent different kinds of light. Light is also sometimes called electromagnetic radiation, and so the kinds of light are called the E/M spectrum. You'll see the terms 'electromagnetic spectrum' or 'E/M spectrum' used, which just means the kinds of light. You'll explore more of the different kinds of light in the next section.

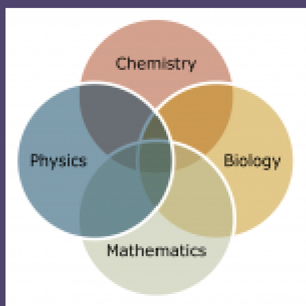
But this is giving you a bit of a hint on where this whole course is going and how light, electricity, and magnetism are all going to be deeply connected in some fundamental way, which will come to by the end of this course.

We've now seen that the frequency or wavelength of a light wave tells us what kind of light we are going to have. What does the amplitude of the light wave correspond to?

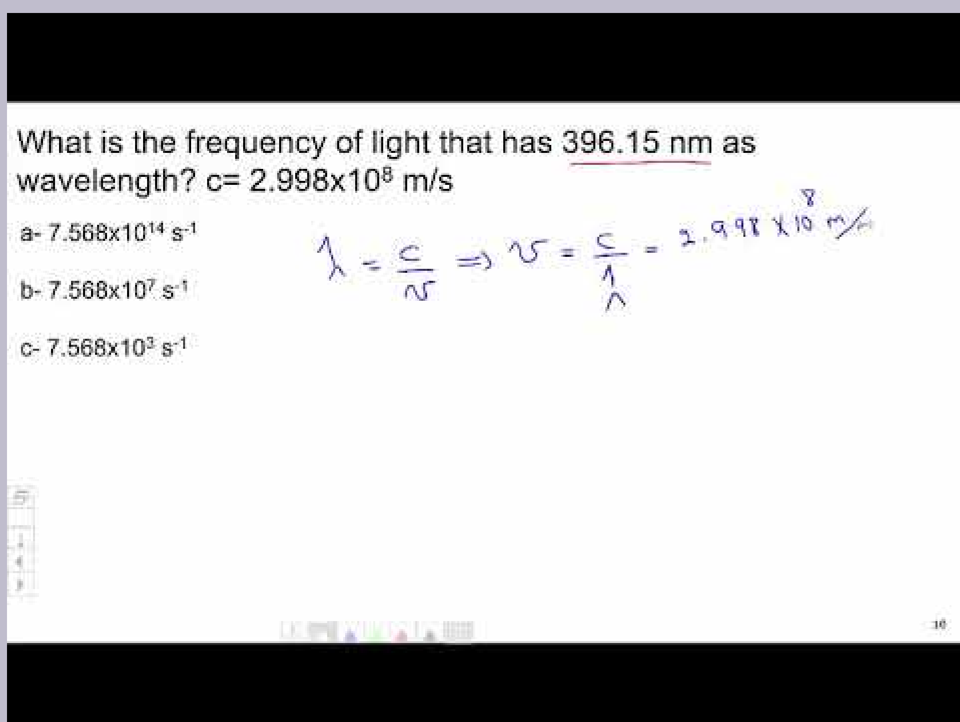
The amplitude, remember we're using E for the amplitude, is related to the intensity of the light, as in the watts per square meter, by this expression

$$I = \frac{1}{2} c \epsilon_0 E^2$$

where c is the usual speed of light $c = 3 \times 10^8\ \mathrm{m/s}$ and ϵ_0 is a property of just empty space. You might not think of empty space as having properties, but it does! The quantity ϵ_0 is a property of empty space called the permittivity of free space, and it has this value $\epsilon_0 = 8.85 \times 10^{-12}\ \frac{\mathrm{C}^2}{\mathrm{J \cdot m}}$. We will talk more about this number throughout this course, for now, you just need to know it's a property of empty space.



An example converting between wavelength and frequency for light (from Chemistry – so this should be familiar!)



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=207

Let's do some examples what is what is the frequency of light that has 396.15 nanometer as wavelength?

Solution:

Wavelength equals c over frequency: $\lambda = c/\nu$, meaning that frequency equals c over lambda $\nu = c/\lambda$. The speed of light in vacuum is given: $2.998 \times 10^8 \,\mathrm{m/s}$. For this question, the wavelength is in nanometers while the unit of the speed of light is in meters, so I know that I have to change the nanometer:

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

$$\nu = \frac{2.998 \times 10^8 \,\mathrm{m/s}}{(396.15 \,\mathrm{nm})\left(\frac{10^{-9} \,\mathrm{m}}{\mathrm{nm}}\right)}$$

$$\nu = 7.568 \times 10^{14} \,\mathrm{s}^{-1} = 7.568 \times 10^{14} \,\mathrm{Hz}$$

Discussion:

That means 7.568×10^{14} is how many waves will pass per one second.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

In summary

- Light is a wave with a: wavelength, frequency, speed, and amplitude.
- The speed of all light waves in vacuum is the same $c = 3 \times 10^8\ \mathrm{m/s}$
- The units of the amplitude of a light wave are Newtons/Coulomb
- We will use E for the amplitude of a light wave instead of A.
  - Keep in mind this is NOT the energy!
  - The amplitude has units Newtons/Coulomb
  - Newtons/Coulomb are not the same unit as the Joules we use for energy!
  - I know it is confusing, but we are running out of letters and there is a good reason for E which we will see later in the course
- While, in general, we know that intensity is proportional to amplitude squared $I \propto A^2$, for light we have exact equation:

$$I = \frac{1}{2} c \epsilon_0 E^2$$

- $\epsilon_0 = 8.85 \times 10^{-12}\ \frac{\mathrm{C}^2}{\mathrm{J \cdot m}}$ is a constant of the Universe, just like the speed of light. We will revisit this constant later.

Homework Problem

Problem 18: Speed dependencies for electromagnetic waves.

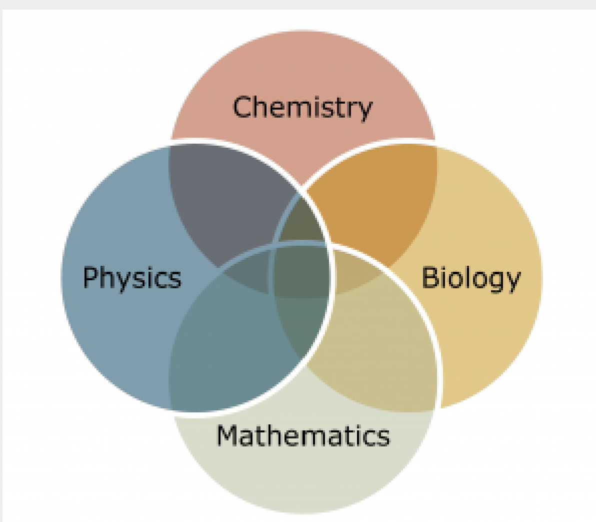
Problem 19: What is the frequency of a radio station given the wavelength?

The Main Parts of the Electromagnetic Spectrum

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

As scientifically trained people, you should have a basic familiarity with the electromagnetic spectrum. Thus, while this course is generally not about memorization, I will ask you to memorize the large basic divisions of the electromagnetic spectrum: radio, microwaves, infrared, visible, ultraviolet, x-rays, and gamma rays. You need to know that radio represents the longest wavelength and gamma rays represent the shortest wavelength. You should also know that, within visible light, red is the longest going through the rainbow to violet. You do NOT need to know the frequencies or wavelengths corresponding to each range. *The only exception to this rule* is that I do expect you to know that red is about 700nm wavelength while violet is about 350nm. The different types of radiation come up so frequently in scientific discussion that it is important to know some basic facts.



Below, you can find a video that summarizes the parts of the electromagnetic spectrum taken from General Chemistry I (Chem 111 at UMass-Amherst), prepared by Dr. Al-Hariri. Please use it to familiarize yourself with the parts of the spectrum if needed.

An additional graphic can be found below the video and its transcript.

A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=207*](http://openbooks.library.umass.edu/toggerson-132/?p=207)

Everyday we're bombarded with different types of radiation like the radio radiation from radio tower close to us, to microwave radiation, to the light radiation and so on; and if you look at the different wavelengths displayed in this picture you can see that the difference between between them is the length of that wave

Now, here are a couple of different types of electromagnetic radiation and the difference way and the different wavelengths of each:

- the infrared radiation with the wavelengths in the range of 10^{-5} meter, which is relatively the same size as pen tip.
- The microwave radiation with wavelengths of about 10^{-3} m, which is in the range of a dice.

- The radio the radio wave, which is the FM and AM: the wavelengths is in the range of 10^3 m, in the range of a mountain.
- The gamma radiationm which is harmful radiation for us, 10^{-12} m, about the atomic nucleus.
- The X-ray 10^{-10} m.
- The ultraviolet in the range of the DNA size and that would be 10^{-6} m.
- And lastly the visible light which is the light that we can see with our own eyes is in the range of 10^{-6} m and the same size as a bacteria.

Different electromagnetic radiations have different wavelengths.

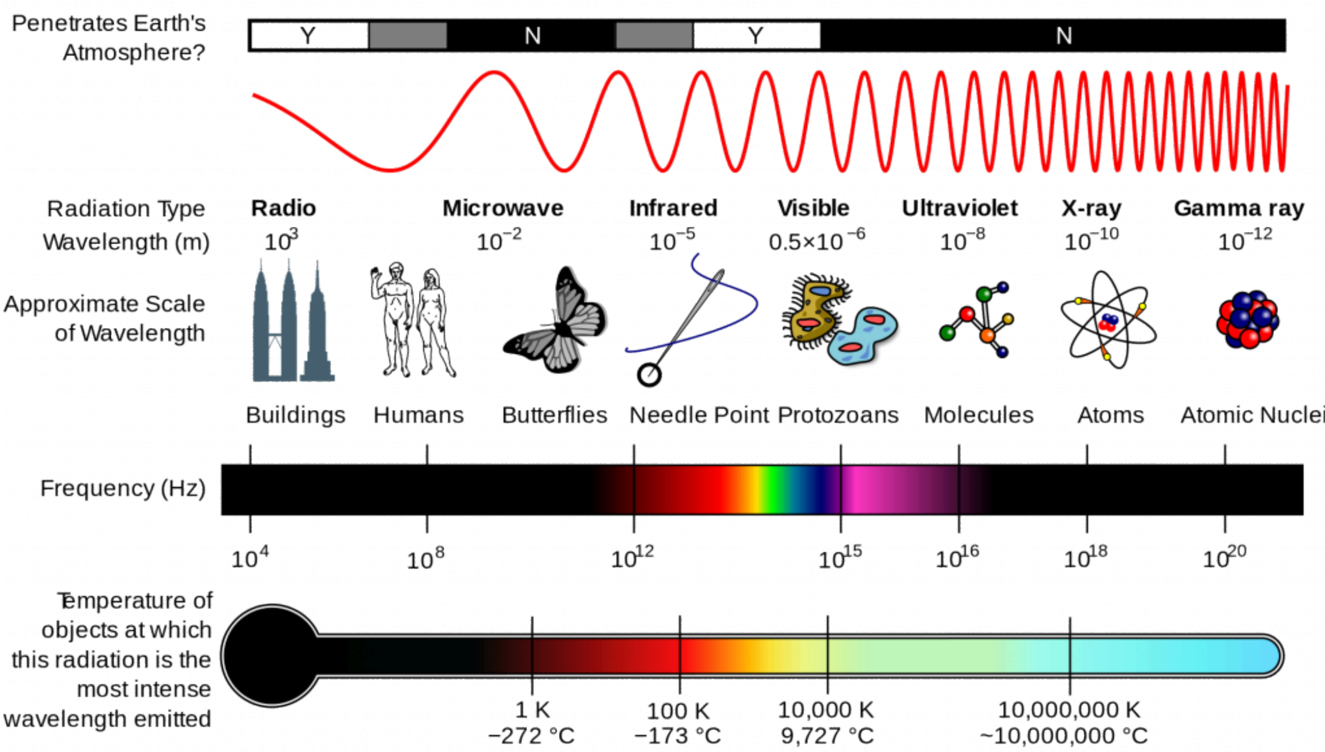


Figure 1: *A diagram of the EM spectrum, showing the type, wavelength (with examples), frequency, the black body emission temperature. (Inductiveload, NASA [CC BY-SA (http://creativecommons.org/licenses/by-sa/3.0/)])*

Homework Problem

Problem 20: Rank the types of waves in the EM spectrum by wavelength.

Introduction to the Photon

A YouTube element has been excluded from this version of the text. You can view it online here: *http://openbooks.library.umass.edu/toggerson-132/?p=207*

We've talked about light as a wave, we've talked about its frequency, its wavelength, its speed, its amplitude. We've talked about the wave properties of light, now we're going to move and think about the particle properties of light. What happens when we think of light as a particle as opposed to as a wave?

Let's say we have a laser, can I keep making this laser dot dimmer and dimmer and dimmer forever? This may seem like a very abstract philosophical question. I'm going to flip it on its head for you. Can I take a sample of water and keep reducing its amount forever? No, eventually I get down to one water molecule, and I'm done. This was the basis for the atomic theory. You can't separate matter forever. I'm just asking you the exact same question for a dot of light, can I keep having it forever? And it turns out the answer is no, I can't. At some point I reach the bottom, there's a smallest dimness, just like there's a smallest amount of water you can have, there's a smallest amount of light you can have, and we call this smallest amount of light we say it's a particle of light, and we call it a photon, and we are going to use this symbol γ, the Greek letter gamma, for photon.

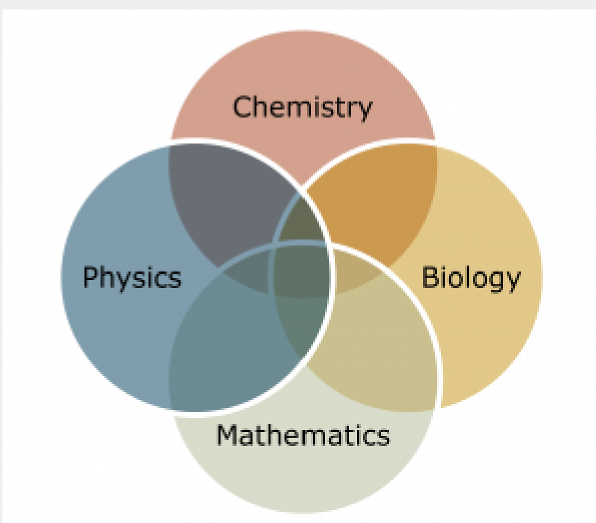
We can think of this laser as a light wave, where I change the amplitude to make it brighter or darker, or we can flip that on its head and say it's a bunch of photons flying along together and to make it brighter or darker I changed the number of photons. Already we're sort of bouncing back and forth between thinking of things as waves and particles. This photon image is really good when we think about light being absorbed by materials or emitted from materials; that's when thinking in terms of particles tends to be a good picture. Waves on the other hand tend to do really well when we're thinking about light flying through space.

Let's go through the properties of the photon. We are now imagining light to be made up of little balls, but we are imagining them to be made up of little massless spheres. Little massless particles that travel at the speed of light, c. But even though they are massless they still carry energy and momentum.

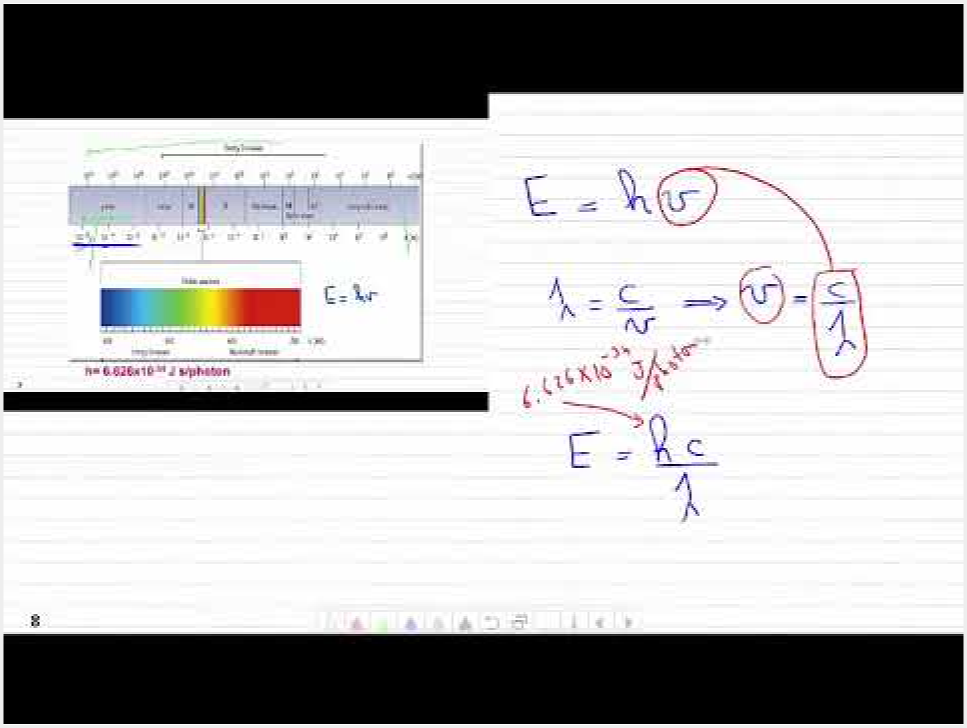
Thinking about detecting/absorbing light? Think particles!

Almost all detection systems talked about thus far—eyes, photographic plates, photomultiplier tubes in microscopes, and CCD cameras—rely on particle-like properties of photons interacting with a sensitive area. A change is caused and either the change is cascaded or zillions of points are recorded to form an image we detect. These detectors are used in biomedical imaging systems, and there is ongoing research into improving the efficiency of receiving photons, particularly by cooling detection systems and reducing thermal effects.

Photon Momentum – Relationship to Wavelength



In this part, we are explicitly trying to delve deeper into an equation you saw in Chemistry: $E = \frac{hc}{\lambda}$. We will see that this equation, while fine for chemistry, is NOT a fundamental principle and thus will NOT be a starting point for us in this class. If you wish to review the chemistry perspective, watch the video below. The video has captions. I did not include the transcript as this video is simply provided to review the chemistry perspective, not as a main focus for our course.



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=207*](http://openbooks.library.umass.edu/toggerson-132/?p=207)

The quantum of EM radiation we call a photon has properties analogous to those of particles we can see, such as grains of sand. A photon interacts as a unit in collisions or when absorbed, rather than as an extensive wave. Massive quanta, like electrons, also act like macroscopic particles—something we expect, because they are the smallest units of matter. Particles carry momentum as well as energy. Despite photons having no mass, there has long been evidence that EM radiation carries momentum. (Maxwell and others who studied EM waves predicted that they would carry momentum.) It is now a well-established fact that photons do have momentum. In fact, photon momentum is suggested by the photoelectric effect, where photons knock electrons out of a substance. Figure 2 shows macroscopic evidence of photon momentum.

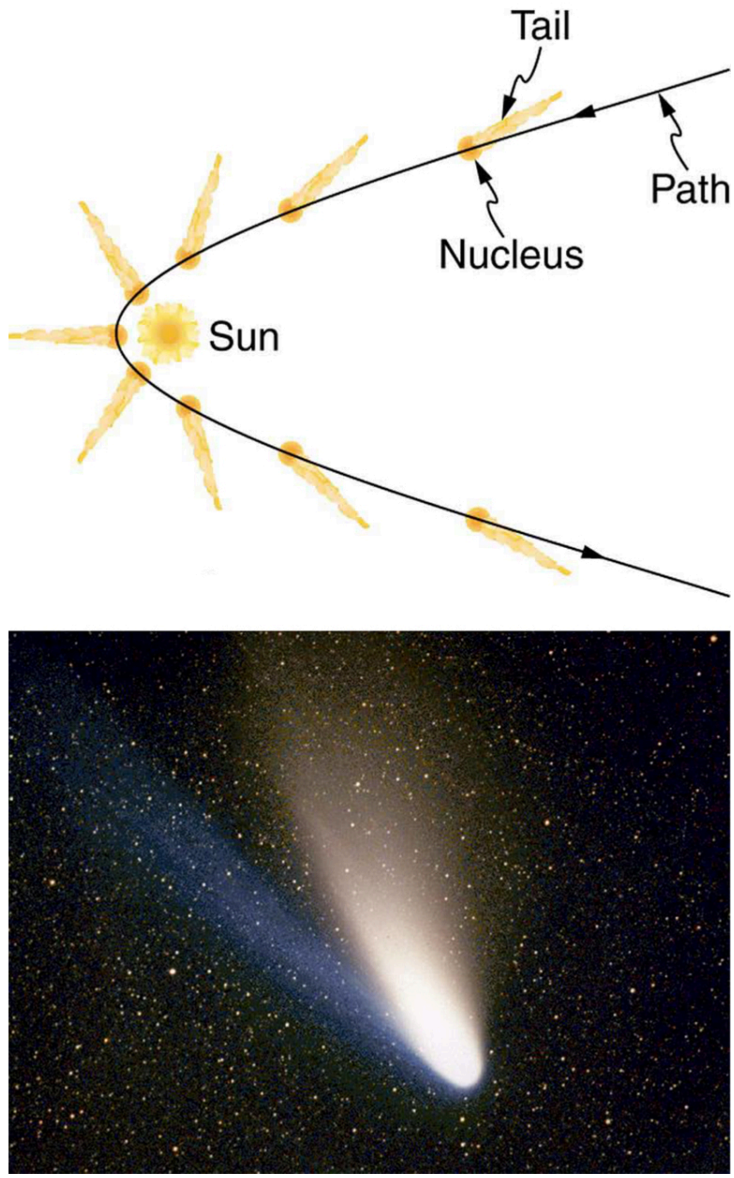


Figure 2: *The tails of the Hale-Bopp comet point away from the Sun, evidence that light has momentum. Dust emanating from the body of the comet forms this tail. Particles of dust are pushed away from the Sun by light reflecting from them. The blue ionized gas tail is also produced by photons interacting with atoms in the comet material. (credit: Geoff Chester, U.S. Navy, via Wikimedia Commons).*

Figure 2. shows a comet with two prominent tails. What most people do not know about the tails is that they always point *away* from the Sun rather than trailing behind the comet (like the tail of Bo Peep's sheep). Comet tails are composed of gases and dust evaporated from the body of the comet and ionized gas. The dust particles recoil away from the Sun when photons scatter from them. Evidently, photons carry momentum in the direction of their motion (away from the Sun), and some of this momentum is transferred to dust particles in collisions. Gas atoms and molecules in the blue tail are most affected by other particles of radiation, such as protons and electrons emanating from the Sun, rather than by the momentum of photons.

Conservation of Momentum

Not only is momentum conserved in all realms of physics, but all types of particles are found to have momentum. We expect particles with mass to have momentum, but now we see that massless particles including photons also carry momentum.

Some of the earliest direct experimental evidence of photon momentum came from scattering of X-ray photons by electrons in substances, named Compton scattering after the American physicist Arthur H. Compton (1892–1962). Around 1923, Compton observed that X-rays scattered from materials had a decreased energy and correctly analyzed this as being due to the scattering of photons from electrons. This phenomenon could be handled as a collision between two particles—a photon and an electron at rest in the material. Energy and momentum are conserved in the collision. (See Figure) He won a Nobel Prize in 1929 for the discovery of this scattering, now called the Compton effect, because it helped prove that photon momentum is given by the de Broglie relation

$$p = \frac{h}{\lambda}$$

where h is Planck's constant: a fundamental constant of the Universe (just like the speed of light c or ϵ_0). The value for Planck's constant is $h = 6.626 \times 10^{-34}\,\mathrm{J \cdot s}$, or in terms of electron volts eV (described in the review of energy) $h = 4.135 \times 10^{-15}\,\mathrm{eV \cdot s}$. This constant, like all constants, is provided on your equation sheet.

Instructor's Note

We will see in a later chapter on matter waves, that this same relation works for electrons as well. Thus, the de Broglie relation

$$p = \frac{h}{\lambda}$$

is one of the fundamental principles for this unit! It connects the particle nature of matter (p is a particle property) and matter's wave nature (λ is a wave property).

The Compton effect is the name given to the scattering of a photon by an electron shown in Figure 3. Energy and momentum are conserved, resulting in a reduction of both for the scattered photon. Studying this effect, Compton verified that photons have momentum. We can see that photon momentum is small, since p=h/λ, and h is very small. It is for this reason that we do not ordinarily observe photon momentum. Our mirrors do not recoil when light reflects from them (except perhaps in cartoons). Compton saw the effects of photon momentum because he was observing x rays, which have a small wavelength and a relatively large momentum, interacting with the lightest of particles, the electron. We will explore this particular phenomenon more in class.

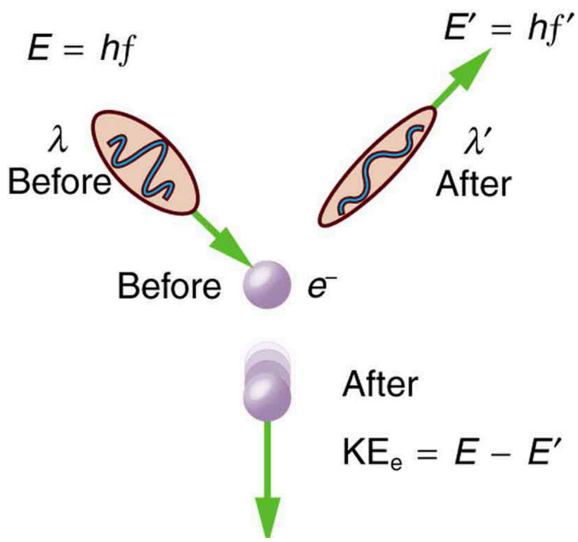


***Figure 3:** The Compton effect is the name given to the scattering of a photon by an electron. Energy and momentum are conserved, resulting in a reduction of both for the scattered photon. Studying this effect, Compton verified that photons have momentum.*

Electron and Photon Momentum Compared

(a) Calculate the momentum of a visible photon that has a wavelength of 500 nm. (b) Find the velocity of an electron having the same momentum.

Strategy

Finding the photon momentum is a straightforward application of its definition: $p = h/\lambda$. Then, we use the formulas we know from 131 to find the electron's momentum and velocity.

Solution for (a)

Photon momentum is given by the equation:

$$p = h/\lambda.$$

Entering the given photon wavelength yields

$$p = \frac{6.63\times10^{-34}\,\mathrm{J{\cdot}s}}{500\times10^{-9}\,\mathrm{m}} = 1.33\times10^{-27}\,\mathrm{kg\cdot m/s}.$$

Solution for (b)

Since this momentum is indeed small, we will use the classical expression $p = mv$ to find the velocity of an electron with this momentum. Solving for v and using the known value for the mass of an electron gives

$$v = \frac{p}{m} = \frac{1.33\times10^{-27}\,\mathrm{kg{\cdot}m/s}}{9.11\times10^{-31}\,\mathrm{kg}} = 1460\,\mathrm{m/s}.$$

Discussion

Photon momentum is indeed small. Even if we have huge numbers of them, the total momentum they carry is small. An electron with the same momentum has a 1460 m/s velocity, which is clearly nonrelativistic. A more massive particle with the same momentum would have an even smaller velocity. This is borne out by the fact that it takes far less energy to give an electron the same momentum as a photon. But on a quantum-mechanical scale, especially for high-energy photons interacting with small masses, photon momentum is significant. Even on a large scale, photon momentum can have an effect if there are enough of them and if there is nothing to prevent the slow recoil of matter. Comet tails are one example, but there are also proposals to build space sails that use huge low-mass mirrors (made of aluminized Mylar) to reflect sunlight. In the vacuum of space, the mirrors would gradually recoil and could actually take spacecraft from place to place in the solar system. (See Figure 4.)

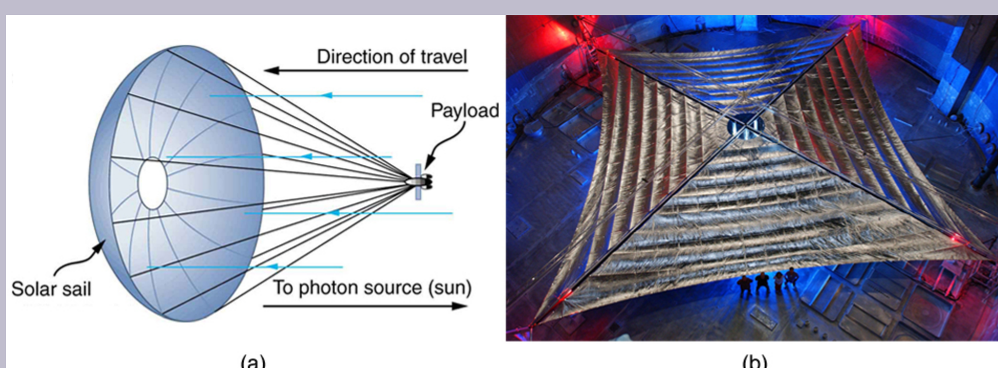


Figure 4: *(a) Space sails have been proposed that use the momentum of sunlight reflecting from gigantic low-mass sails to propel spacecraft about the solar system. A Russian test model of this (the Cosmos 1) was launched in 2005, but did not make it into orbit due to a rocket failure. (b) A U.S. version of this, labeled LightSail-1, is scheduled for trial launches in the first part of this decade. It will have a 40-m2 sail. (credit: Kim Newton/NASA)*

Homework Problems

Problem 21: Find the momentum of a microwave photon.

Photon Momentum – Relationship to Energy

Photons, in addition to having energy, also have momentum. This is the part that tends to get folks, because in 131. we told you that momentum was mass times velocity which is mostly true. It's true as long as you're not going too fast, once you start getting close to the speed of light this will actually break down on you. You need a new expression. But as long as you're going slow, this is fine. But clearly this does not work for photons because for photons mass is zero. Special relativity has an answer, it's the momentum of a photon is the energy divided by the speed of light,

$$p = \frac{E_\gamma}{c} \text{ or } E_\gamma = pc.$$

University *of* Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

If you look at the Unit I On-a-Page, you will see that this is one of the fundamental definitions of this unit: the definition of a photon's momentum in terms of its energy

$$p = \frac{E_\gamma}{c}.$$

A way to help keep all of these formula straight: if the formula contains a c then it only applies to light, if the formula contains an m, then it only applies to particles with mass (like electrons)!

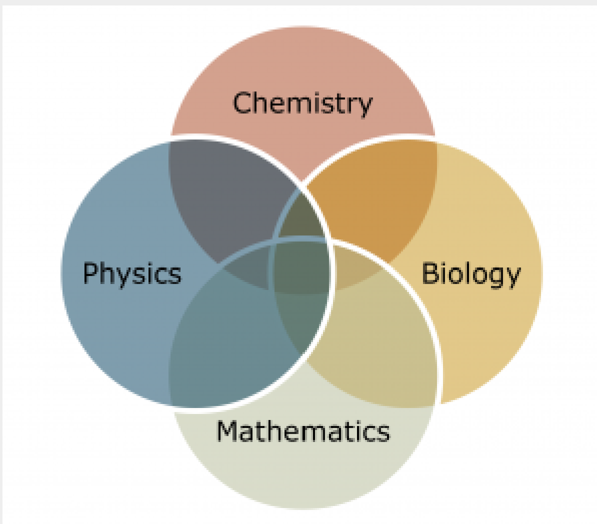
From the fundamental principle of this unit, the de Broglie relation $p = \frac{h}{\lambda}$, and this definition of a photon's momentum in terms of its energy $p = \frac{E}{c}$, we can derive a formula that was given to you in your chemistry classes. While I, in general, try to avoid derivations, I think this one is useful as it is short and shows you *why* what you learned in chemistry is the way it is. That is, after all, [one of the motivations for this unit: why does chemistry work?]

So we know, $p = \frac{h}{\lambda}$ and $p = \frac{E}{c}$. Since both equations are equal to p, we can set them equal to each other:

$$\frac{h}{\lambda} = \frac{E_\gamma}{c}$$

which, after some rearranging (move the c over) we get the familiar

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



You can start with this equation that you know from chemistry. However, keep in mind that it is NOT a fundamental relationship: it comes from combining:

- The fundamental *principle* of the de Broglie relation: $p = h/\lambda$ that connects the wave and particle natures for *all* matter.
- The definition of a photon's momentum in terms of its energy: $E_\gamma = pc$, which is *only specific to photons.*

Therefore, the relationship $E = \frac{hc}{\lambda}$ only applies to photons. I have seen many students make mistakes of trying to apply it to electrons!

Compare the Energies of the Photon and Electron from the Last Example

What is the energy of the 500 nm photon, and how does it compare with the energy of the electron with the same momentum?

The electron:

There are two ways of approaching this problem.

1. Use the fact that we know the electron's velocity to be $1460\,\mathrm{m/s}$, and the expression for kinetic energy from Physics 131: $K = \frac{1}{2}mv^2$:

$$K = \frac{1}{2}(9.11 \times 10^{-31}\,\mathrm{kg})(1460\,\mathrm{m/s})^2$$

$$K = 9.64 \times 10^{-25}\,\mathrm{J} = 6.02 \times 10^{-6}\,\mathrm{eV}$$

2. Directly use the fact that we already know the electron's momentum from the previous problem $p = 1.33 \times 10^{-27}\,\mathrm{kg \cdot m/s}$. Combine this knowledge and the idea of converting directly from momentum to energy for particles with mass using the formula derived in Some Energy-Related Ideas that Might be New: The Connection between Energy and Momentum:

$$K = \frac{p^2}{2m}$$

$$K = \frac{(1.33 \times 10^{-27}\,\mathrm{kg \cdot m/s})^2}{2(9.11 \times 10^{-31}\,\mathrm{kg})} = 9.64 \times 10^{-25}\,\mathrm{J} = 6.02 \times 10^{-6}\,\mathrm{eV}$$

Clearly, both approaches give the same response as they must.

The photon:

Again, there are two approaches:

1. Use the momentum of the photon to get the energy using $E_\gamma = pc$:

$$E_\gamma = pc$$

$$E_\gamma = ((1.33 \times 10^{-27}\,\mathrm{kg \cdot m/s})(3 \times 10^8\,\mathrm{m/s})$$

$$E_\gamma = 3.99 \times 10^{-19}\,\mathrm{J} = 2.49\,\mathrm{eV}.$$

Where eV are *electron volts* discussed in Units of Energy.

2. The second approach, is to use the wavelength, coupled with the expression we just derived / you learned in chemistry:

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

$$E = \frac{(6.626 \times 10^{-34}\,\mathrm{J \cdot s})(3 \times 10^8\,\mathrm{m/s})}{500 \times 10^{-9}\,\mathrm{m}}$$

$$E = 3.99 \times 10^{-19}\,\mathrm{J} = 2.49\,\mathrm{eV}.$$

Again, both approaches give the same value, *as they must.*

Homework Problem

Problem 22: From momentum, calculate the wavelength and energy of a photon.

Photon Energies and the Electromagnetic Spectrum

A photon is a quantum of EM radiation whose momentum is related to its wavelength by $p = \frac{h}{\lambda}$. Combined with the connection between a photon's energy and momentum $E_\gamma = pc$ yields the energy-wavelength relationship $E = \frac{hc}{\lambda}$.

All EM radiation is composed of photons. Figure 5 shows various divisions of the EM spectrum plotted against wavelength, frequency, and photon energy. Previously in this book, photon characteristics were alluded to in the discussion of some of the characteristics of UV, x rays, and γ-rays, the first of which start with frequencies just above violet in the visible spectrum. It was noted that these types of EM radiation have characteristics much different than visible light. We can now see that such properties arise because photon energy is larger at high frequencies.

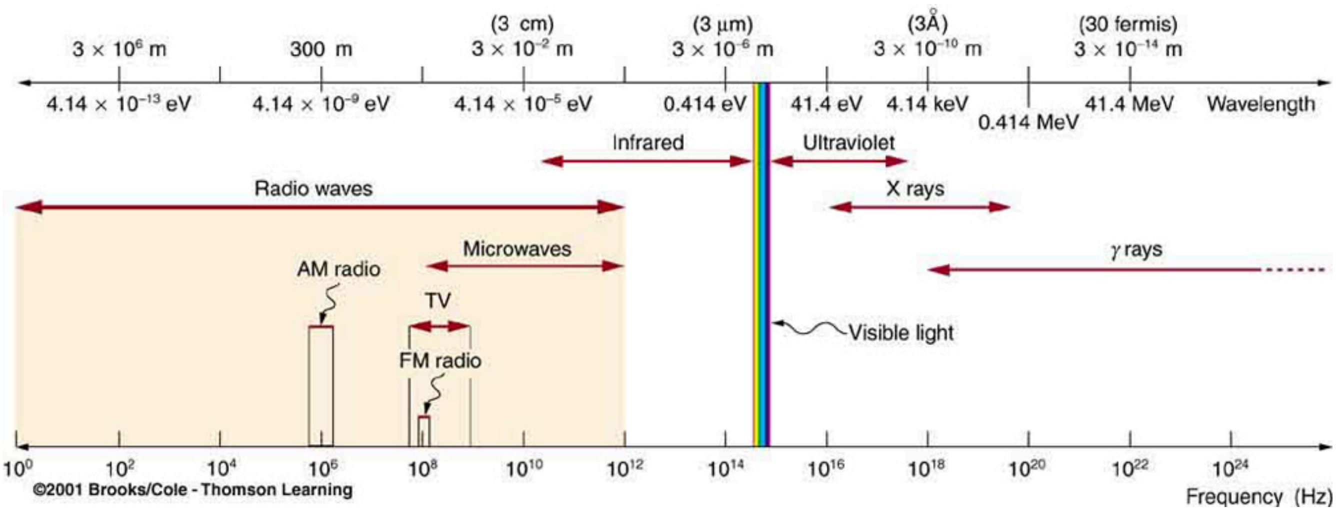


Figure 5: *The EM spectrum, showing major categories as a function of photon energy in eV, as well as wavelength and frequency. Certain characteristics of EM radiation are directly attributable to photon energy alone.*

Photons act as individual quanta and interact with individual electrons, atoms, molecules, and so on. The energy a photon carries is, thus, crucial to the effects it has. Table 1 lists representative submicroscopic energies in eV. When we compare photon energies from the EM spectrum in Figure 5 with energies in the table, we can see how effects vary with the type of EM radiation.

Table 1

Representative Energies for Submicroscopic Effects (Order of Magnitude Only)	
Energy between outer electron shells in atoms	1 eV
Binding energy of a weakly bound molecule	1 eV
Energy of red light	2 eV
Binding energy of a tightly bound molecule	10 eV
Energy to ionize atom or molecule	10 to 1000 eV

Ionizing Radiation

Gamma rays

A form of nuclear and cosmic EM radiation, can have the highest frequencies and, hence, the highest photon energies in the EM spectrum. For example, a γ-ray photon with $\nu = 10^{21}\,\mathrm{Hz}$ has an energy $E = h\nu = 6.63 \times 10^{-13}\,\mathrm{J} = 4.14\,\mathrm{MeV}$. This is sufficient energy to ionize thousands of atoms and molecules, since only 10 to 1000 eV are needed per ionization. In fact, γ rays are one type of ionizing radiation, as are x rays and UV, because they produce ionization in materials that absorb them. Because so much ionization can be produced, a single γ-ray photon can cause significant damage to biological tissue, killing cells or damaging their ability to properly reproduce. When cell reproduction is disrupted, the result can be cancer, one of the known effects of exposure to ionizing radiation. Since cancer cells are rapidly reproducing, they are exceptionally sensitive to the disruption produced by ionizing radiation. This means that ionizing radiation has positive uses in cancer treatment as well as risks in producing cancer. However, the high photon energy also enables γ rays to penetrate materials, since a collision with a single atom or molecule is unlikely to absorb all the γ ray's energy. This can make γ rays useful as a probe, and they are sometimes used in medical imaging.

X-rays

X-rays, as you can see in Figure 5, overlap with the low-frequency end of the γ ray range. Since x rays have energies of keV and up, individual x-ray photons also can produce large amounts of ionization. At lower photon energies, x rays are not as penetrating as γ rays and are slightly less hazardous. X-rays are ideal for medical imaging, their most common use, and a fact that was recognized immediately upon their discovery in 1895 by the German physicist W. C. Roentgen (1845–1923). (See Figure 6.) Within one year of their discovery, x rays (for a time called Roentgen rays) were used for medical diagnostics. Roentgen received the 1901 Nobel Prize for the discovery of x rays.

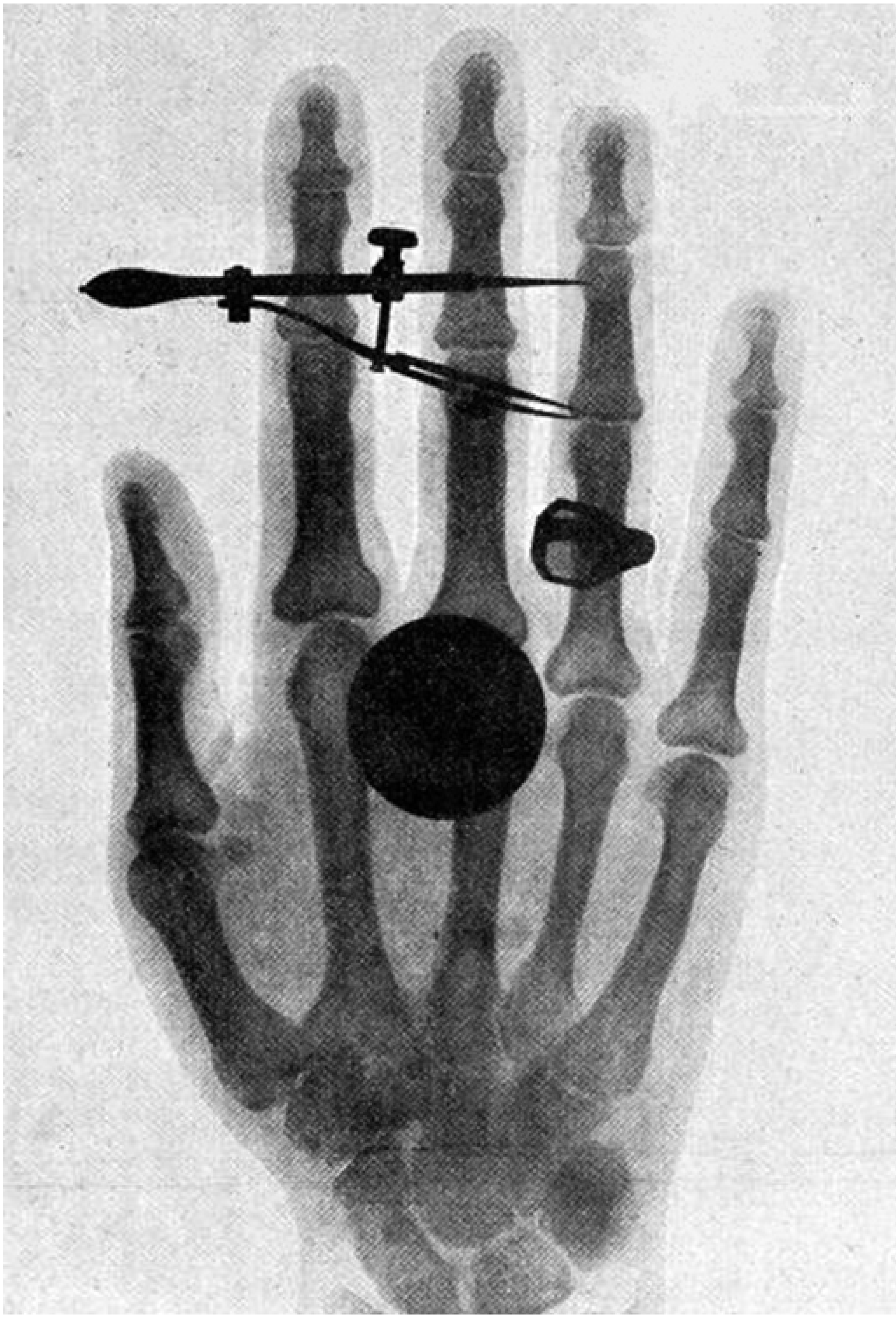


Figure 6: *One of the first x-ray images, taken by Röentgen himself. The hand belongs to Bertha Röentgen, his wife. (credit: Wilhelm Conrad Röntgen, via Wikimedia Commons)*

While γ rays originate in nuclear decay, x rays are produced by the process shown in Figure 7. Electrons ejected by thermal agitation from a hot filament in a vacuum tube are accelerated through a high voltage, gaining kinetic energy from the electrical potential energy. When they strike the anode, the electrons convert their kinetic energy to a variety of forms, including thermal energy. But since an accelerated charge radiates EM waves, and since the electrons act individually, photons are also produced. Some of these x-ray photons obtain the kinetic energy of the electron. The accelerated electrons originate at the cathode, so such a tube is called a cathode ray tube (CRT), and various versions of them are found in older TV and computer screens as well as in x-ray machines.

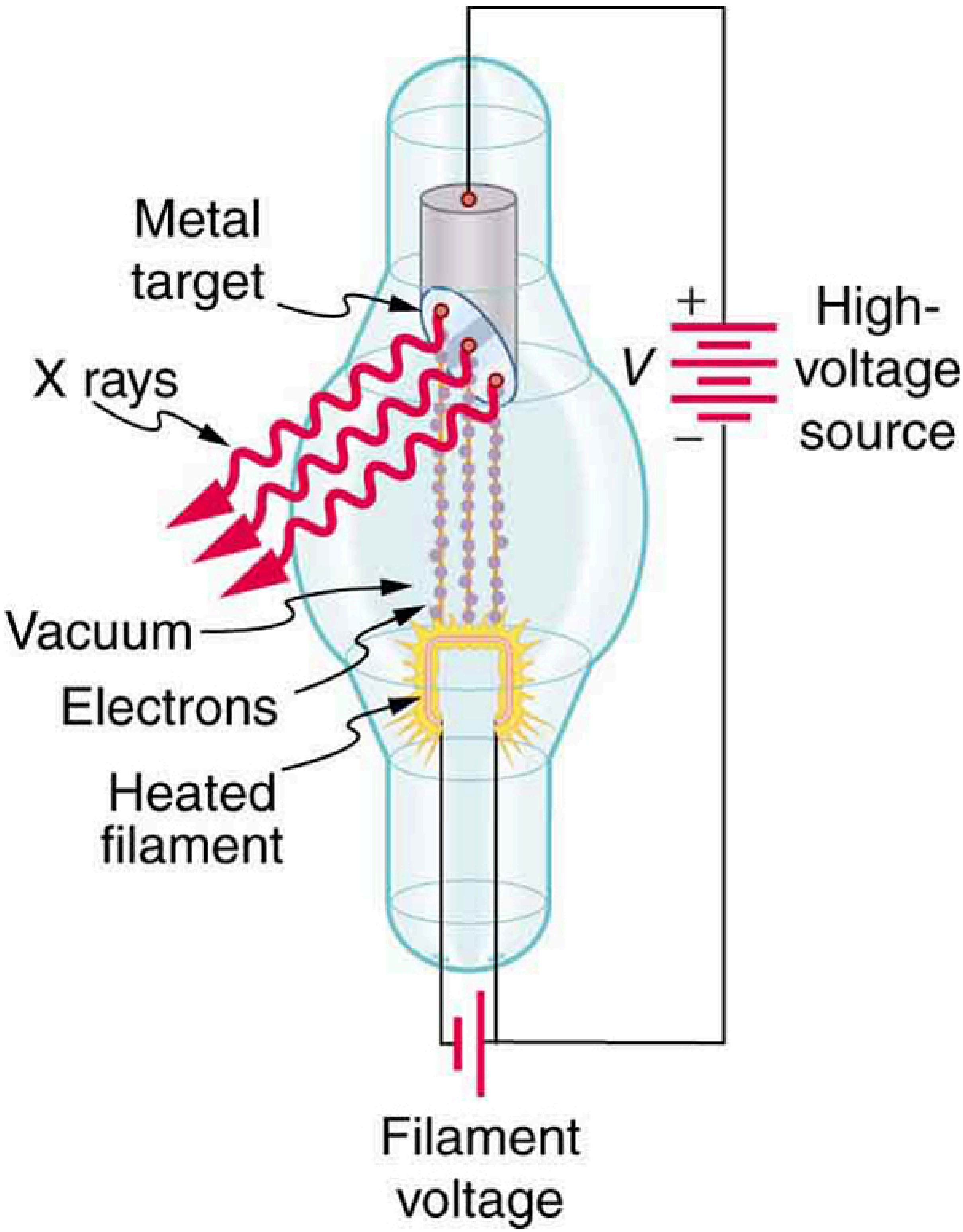


Figure 7: *X rays are produced when energetic electrons strike the copper anode of this cathode ray tube (CRT). Electrons (shown here as separate particles) interact individually with the material they strike, sometimes producing photons of EM radiation.*

Figure 8 shows the spectrum of x rays obtained from an x-ray tube. There are two distinct features to the spectrum. First, the smooth distribution results from electrons being decelerated in the anode material. A curve like this is obtained by detecting many photons, and it is apparent that the maximum energy is unlikely. This decelerating process produces radiation that is called bremsstrahlung (German for braking radiation). The second feature is the existence of sharp peaks in the spectrum; these are called characteristic x rays, since they are characteristic of the anode material. Characteristic x rays come from atomic excitations unique to a given type of anode material. They are akin to lines in atomic spectra, implying the energy levels of atoms are quantized.

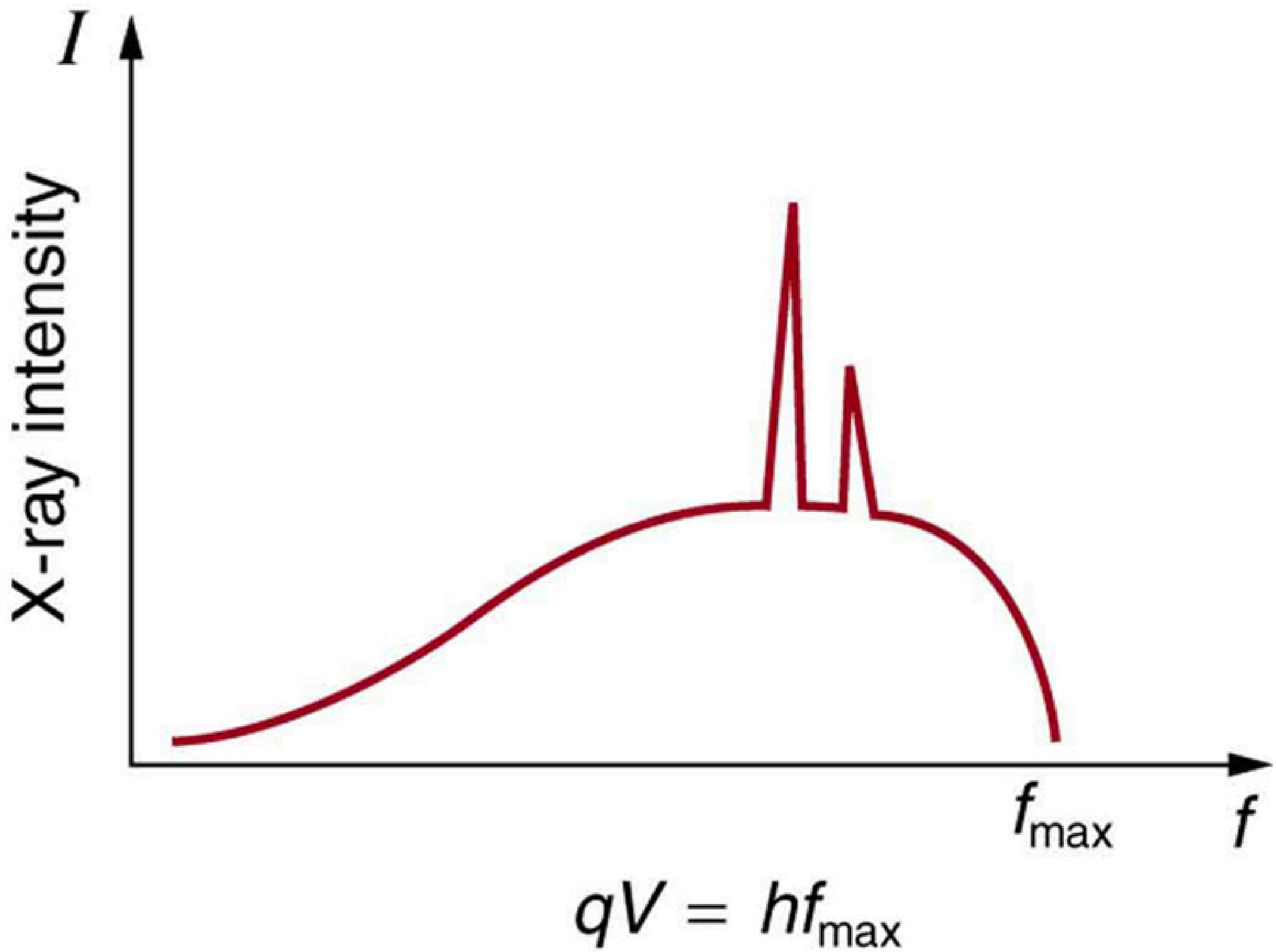


Figure 8: *X-ray spectrum obtained when energetic electrons strike a material. The smooth part of the spectrum is bremsstrahlung, while the peaks are characteristic of the anode material. Both are atomic processes that produce energetic photons known as x-ray photons.*

Connections: Conservation of Energy

Once again, we find that conservation of energy allows us to consider the initial and final forms that energy takes, without having to make detailed calculations of the intermediate steps.

Find the minimum wavelength of an x-ray photon produced by electrons accelerated through a potential energy difference of 50.0 keV in a CRT like the one in Figure 7.

Strategy

Electrons can give all of their kinetic energy to a single photon when they strike the anode of a CRT. The kinetic energy of the electron comes from electrical potential energy. Thus we can simply equate the maximum photon energy to the electrical potential energy

Solution

$$\Delta E = Q + W$$

$$E_f - E_i = Q + W$$

In the initial state, we have an electron with 50 keV of potential energy and no kinetic energy: $E_i = U_i = 50\,\mathrm{keV}$. At the end, all that energy is in the photon: $E_f = E_\gamma$. No other energy enters or leaves the system (the photon and electron are everything we care about!), so $Q = W = 0$

$$E_\gamma - U_i = 0$$

$$E_\gamma = U_i$$

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

$$\frac{1}{\lambda} = \frac{U_i}{hc}$$

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

$$\lambda = \frac{(6.626\times10^{-34}\ \mathrm{J{\cdot}s})(3\times10^{8}\ \mathrm{m/s})}{(50\times10^{3}\ \mathrm{eV})\left(\frac{1.602\times10^{-19}\ \mathrm{J}}{\mathrm{eV}}\right)}$$

$$\lambda = 2.48\times10^{-11}\,\mathrm{m} = 0.025\,\mathrm{nm}$$

Ultraviolet Radiation

Ultraviolet radiation (approximately 4 eV to 300 eV) overlaps with the low end of the energy range of x rays, but UV is typically lower in energy. UV comes from the de-excitation of atoms that may be part of a hot solid or gas. These atoms can be given energy that they later release as UV by numerous processes, including electric discharge, nuclear explosion, thermal agitation, and exposure to x rays. A UV photon has sufficient energy to ionize atoms and molecules, which makes its effects different from those of visible light. UV thus has some of the same biological effects as γ-rays and x-rays. For example, it can cause skin cancer and is used as a sterilizer. The major difference is that several UV photons are required to disrupt cell reproduction or kill a bacterium, whereas single γ-ray and x-ray photons can do the same damage. But since UV does have the energy to alter molecules, it can do what visible light cannot. One of the beneficial aspects of UV is that it triggers the production of vitamin D in the skin, whereas visible light has insufficient energy per photon to alter the molecules that trigger this production. Infantile jaundice is treated by exposing the baby to UV (with eye protection), called phototherapy, the beneficial effects of which are thought to be related to its ability to help prevent the buildup of potentially toxic bilirubin in the blood.

Photon Energy and Effects for UV

Short-wavelength UV is sometimes called vacuum UV, because it is strongly absorbed by air and must be studied in a vacuum. Calculate the photon energy in eV for 100-nm vacuum UV, and estimate the number of molecules it could ionize or break apart.

Strategy

Using the equation $E=\frac{hc}{\lambda}$ and appropriate constants, we can find the photon energy and compare it with energy information in Table 1.

Solution

The energy of a photon is given by

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

Using hc=1240 eV·nm,

we find that

E=hc/λ=(1240 eV·nm)/100 nm=12.4 eV.

Discussion

According to Table 1, this photon energy might be able to ionize an atom or molecule, and it is about what is needed to break up a tightly bound molecule, since they are bound by approximately 10 eV. This photon energy could destroy about a dozen weakly bound molecules. Because of its high photon energy, UV disrupts atoms and molecules it interacts with. One good consequence is that all but the longest-wavelength UV is strongly absorbed and is easily blocked by sunglasses. In fact, most of the Sun's UV is absorbed by a thin layer of ozone in the upper atmosphere, protecting sensitive organisms on Earth. Damage to our ozone layer by the addition of such chemicals as CFC's has reduced this protection for us.

Visible Light

The range of photon energies for visible light from red to violet is 1.63 to 3.26 eV, respectively. These energies are on the order of those between outer electron shells in atoms and molecules. This means that these photons can be absorbed by atoms and molecules. A single photon can actually stimulate the retina, for example, by altering a receptor molecule that then triggers a nerve impulse. As reviewed from chemistry in a future chapter, photons can be absorbed or emitted only by atoms and molecules that have precisely the correct quantized energy step to do so. For example, if a red photon of frequency ν encounters a molecule that has an energy step, $\Delta E=h\nu$, then the photon can be absorbed. Violet flowers absorb red and reflect violet; this implies there is no energy step between levels in the receptor molecule equal to the violet photon's energy, but there is an energy step for the red.

There are some noticeable differences in the characteristics of light between the two ends of the visible spectrum that are due to photon energies. Red light has insufficient photon energy to expose most black-and-white film, and it is thus used to illuminate darkrooms where such film is developed. Since violet light has a higher photon energy, dyes that absorb violet tend to fade more quickly than those that do not. (See Figure 9.)

Take a look at some faded color posters in a storefront some time, and you will notice that the blues and violets are the last to fade. This is because other dyes, such as red and green dyes, absorb blue and violet photons, the higher energies of which break up their weakly bound molecules. (Complex molecules such as those in dyes and DNA tend to be weakly bound.) Blue and violet dyes reflect those colors and, therefore, do not absorb these more energetic photons, thus suffering less molecular damage.

Figure 9: *Why do the reds, yellows, and greens fade before the blues and violets when exposed to the Sun, as with this poster? The answer is related to photon energy. (credit: Deb Collins, Flickr)*

Transparent materials, such as some glasses, do not absorb any visible light, because there is no energy step in the atoms or molecules that could absorb the light. Since individual photons interact with individual atoms, it is nearly impossible to have two photons absorbed simultaneously to reach a large energy step. Because of its lower photon energy, visible light can sometimes pass through many kilometers of a substance, while higher frequencies like UV, x-ray, and γγ size 12{γ} {}">γ-rays are absorbed, because they have sufficient photon energy to ionize the material.

How Many Photons per Second Does a Typical Light Bulb Produce?

Assuming that 10.0% of a 100-W light bulb's energy output is in the visible range (typical for incandescent bulbs) with an average wavelength of 580 nm, calculate the number of visible photons emitted per second.

Strategy

Power is energy per unit time, and so if we can find the energy per photon, we can determine the number of photons per second. This will best be done in Joules, since power is given in Watts, which are Joules per second.

Solution

The power in visible light production is 10.0% of 100 W, or 10.0 J/s. The energy of the average visible photon is found by substituting the given average wavelength into the formula

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

This produces

$$E = \frac{(6.63 \times 10^{-34}\ \mathrm{J{\cdot}s})(3.00 \times 10^{8} \mathrm{m/s})}{580 \times 10^{-9}\ \mathrm{m}} = 3.43 \times 10^{-19}\ \mathrm{J}.$$

The number of visible photons per second is thus

$$\frac{\mathrm{photon}}{\mathrm{s}} = \frac{10.0\ \mathrm{J/s}}{3.43 \times 10^{-19}\ \mathrm{J/photon}} = 2.92 \times 10^{19}\ \mathrm{photon/s}.$$

Discussion

This incredible number of photons per second is verification that individual photons are insignificant in ordinary human experience. It is also a verification of the correspondence principle—on the macroscopic scale, quantization becomes essentially continuous or classical. Finally, there are so many photons emitted by a 100-W lightbulb that it can be seen by the unaided eye many kilometers away.

Lower Energy Photons

Infrared Radiation (IR)

Infrared radiation (IR) has even lower photon energies than visible light and cannot significantly alter atoms and molecules. IR can be absorbed and emitted by atoms and molecules, particularly between closely spaced states. IR is extremely strongly absorbed by water, for example, because water molecules have many states separated by energies on the order of 10–5eV 10–5 eV size 12{" 10" rSup { size 8{"–5"} } " eV "} {}">10^{-5}eV to 10–2eV, 10–2 eV, size 12{" 10" rSup { size 8{"–2"} } " eV "} {}">10^{-2}eV, well within the IR and microwave energy ranges. This is why in the IR range, skin is almost jet black, with an emissivity near 1—there are many states in water molecules in the skin that can absorb a large range of IR photon energies. Not all molecules have this property. Air, for example, is nearly transparent to many IR frequencies.

Microwaves

Microwaves are the highest frequencies that can be produced by electronic circuits, although they are also produced naturally. Thus microwaves are similar to IR but do not extend to as high frequencies. There are states in water and other molecules that have the same frequency and energy as microwaves, typically about 10–5eV. 10–5 eV. size 12{" 10" rSup { size 8{"–5"} } " eV "} {}">10^{-5}eV. This is one reason why food absorbs microwaves more strongly than many other materials, making microwave ovens an efficient way of putting energy directly into food.

Photon energies for both IR and microwaves are so low that huge numbers of photons are involved in any significant energy transfer by IR or microwaves (such as warming yourself with a heat lamp or cooking pizza in the microwave). Visible light, IR, microwaves, and all lower frequencies cannot produce ionization with single photons and do not ordinarily have the hazards of higher frequencies. When visible, IR, or microwave radiation *is* hazardous, such as the inducement of cataracts by microwaves, the hazard is due to huge numbers of photons acting together (not to an accumulation of photons, such as sterilization by weak UV). The negative effects of visible, IR, or microwave radiation can be thermal effects, which could be produced by any heat source. But one difference is that at very high intensity, strong electric and magnetic fields can be produced by photons acting together. Such electromagnetic fields (EMF) can actually ionize materials.

Misconception Alert: High-Voltage Power Lines

Although some people think that living near high-voltage power lines is hazardous to one's health, ongoing studies of the transient field effects produced by these lines show their strengths to be insufficient to cause damage. Demographic studies also fail to show significant correlation of ill effects with high-voltage power lines. The American Physical Society issued a report over 10 years ago on power-line fields, which concluded that the scientific literature and reviews of panels show no consistent, significant link between cancer and power-line fields. They also felt that the "diversion of resources to eliminate a threat which has no persuasive scientific basis is disturbing."

Lower Energy than Microwaves

It is virtually impossible to detect individual photons having frequencies below microwave frequencies, because of their low photon energy. But the photons are there. A continuous EM wave can be modeled as photons. At low frequencies, EM waves are generally treated as time- and position-varying electric and magnetic fields with no discernible quantization. This is another example of the correspondence principle in situations involving huge numbers of photons.

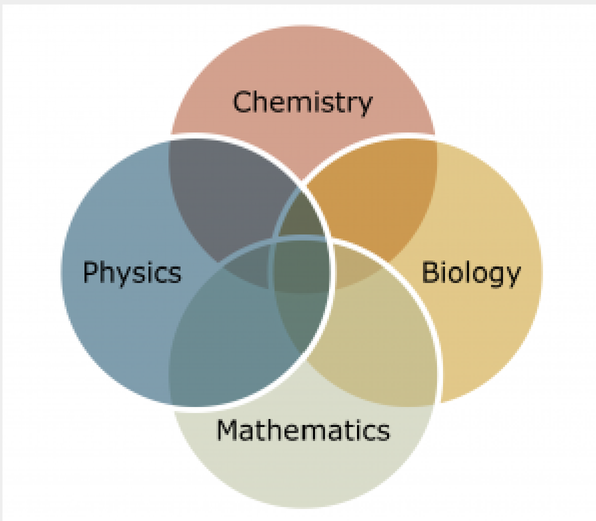
Homework Problems

Hint: Look carefully at the example above with the 100W light bulb!

Problem 23: An AM radio transmitter radiates some power at a given frequency. How many photons per second does the emitter emit?

Problem 24: If the brightness of a beam of light is increased, the ________ of the ______________ will also increase.

8. Review from Chemistry of Application of Conservation of Energy to Photons and Atoms



In your general chemistry courses, you already did some of what will be a big part of this unit: namely using the ideas of wave-particle duality in conjunction with conservation of energy. In chemistry, you did this in the context of looking at atomic transitions. In this chapter, you will review the ideas from chemistry, and then be exposed to some differences in how we will treat this same situation in a physics course. The reasons for the differences are three-fold. First, as described elsewhere in this book, each scientific discipline grew with its own history and conventions. The second reason for a different perspective is that the view taken by chemistry, while perfectly fine for all situations you encountered in that course, will break down for some of the situations we want to analyze in this course. The third, and arguably most important reason, mirrors the motivation for a variety of cultural perspectives in a humanities course: by exploring the same processes from different perspectives you gain a deeper and more holistic understanding of the material. Just as your understanding of history is incomplete if you only consider white men, so your understanding of conservation of energy is incomplete if you only look at it from a biology, chemistry, or physics perspective.

Review of Connecting Conservation of Energy to the Wave and Particle Natures of Light in the Context of the Hydrogen Atom from Chemistry[1]

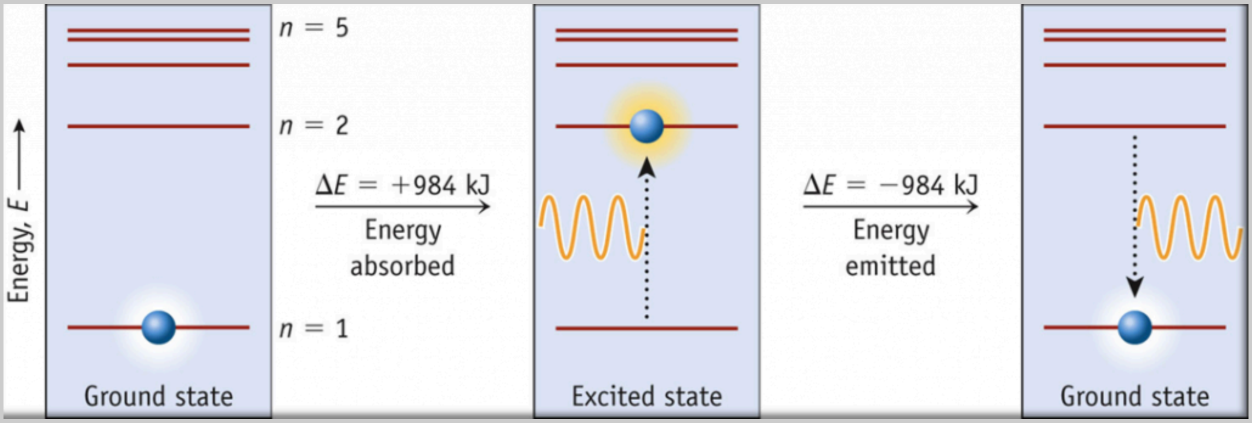
Instructor's Notes

NOTE: This is review. If you are familiar with this material, feel free to skip to the problems at the end!

What we expect you to know from this review:

- Electrons in atoms have discrete energy levels.
- In order for an electron to transition from one level to another it must absorb or emit a photon with the *exact* amount of energy that corresponds to the energy difference. If the energy does not *exactly* match, then the transition will not occur.
- If I know the energy difference, I can solve for the wavelength of the emitted photon.
- Be familiar with the idea of energy level diagrams such as the one shown below

1. Paul Flowers et al. *Chemistry: Atoms First 2e*. Open Stax, 2014.



- The difference in energy levels in a hydrogen atom is

$$\Delta E = -Rhc\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

where R is the Rydberg constant $R = 1.097 \times 10^7\,\mathrm{m}^{-1}$.

Following the work of Ernest Rutherford and his colleagues in the early twentieth century, the picture of atoms consisting of tiny dense nuclei surrounded by lighter and even tinier electrons continually moving about the nucleus was well established. This picture was called the planetary model, since it pictured the atom as a miniature "solar system" with the electrons orbiting the nucleus like planets orbiting the sun. The simplest atom is hydrogen, consisting of a single proton as the nucleus about which a single electron moves. The electrostatic force attracting the electron to the proton depends only on the distance between the two particles. This classical mechanics description of the atom is incomplete, however, since as described in Basics of Light: Where does Light Come From? an electron moving in an orbit would be accelerating (by changing direction) and, according to classical electromagnetism, it should continuously emit electromagnetic radiation. This loss in orbital energy should result in the electron's orbit getting continually smaller until it spirals into the nucleus, implying that atoms are inherently unstable.

In 1913, Niels Bohr attempted to resolve the atomic paradox by ignoring classical electromagnetism's prediction that the orbiting electron in hydrogen would continuously emit light. Instead, he incorporated into the classical mechanics description of the atom Planck's ideas of quantization and Einstein's finding that light consists of photons whose energy is proportional to their frequency (Basics of Light: Introduction to the Photon). Bohr assumed that the electron orbiting the nucleus would not normally emit any radiation (the stationary state hypothesis), but it would emit or absorb a photon if it moved to a different orbit. The energy absorbed or emitted would reflect differences in the orbital energies according to this equation:

$$|\Delta E| = |E_f - E_i| = \frac{hc}{\lambda}$$

In this equation, *h* is Planck's constant and E_i and E_f are the initial and final orbital energies, respectively. The absolute value of the energy difference is used, since frequencies and wavelengths are always positive. Instead of allowing for continuous values of energy, Bohr assumed the energies of these electron orbitals were quantized:

$$E_n = -\frac{Rhc}{n^2}$$

Where *R* is the so-called Rydberg constant that had been previously determined by experimental analysis of hydrogen spectra $R = 1.097 \times 10^7\,\mathrm{m}^{-1}$. Inserting this expression into the equation for ΔE gives

$$\Delta E = -Rhc\left(\frac{1}{n_f}^2 - \frac{1}{n_i}^2\right) = \frac{hc}{\lambda}.$$

The lowest few energy levels are shown in Figure 1. One of the fundamental laws of physics is that matter is most stable with the lowest possible energy. Thus, the electron in a hydrogen atom usually moves in the n = 1 orbit, the orbit in which it has the lowest energy. When the electron is in this lowest energy orbit, the atom is said to be in its ground electronic state (or simply ground state). If the atom receives energy from an outside source, it is possible for the electron to move to an orbit with a higher n value and the atom is now in an excited electronic state (or simply an excited state) with a higher energy. When an electron transitions from an excited state (higher energy orbit) to a less excited state, or ground state, the difference in energy is emitted as a photon. Similarly, if a photon is absorbed by an atom, the energy of the photon moves an electron from a lower energy orbit up to a more excited one. We can relate the energy of electrons in atoms to what we learned previously about energy. The law of conservation of energy says that we can neither create nor destroy energy. Thus, if a certain amount of external energy is required to excite an electron from one energy level to another, that same amount of energy will be liberated when the electron returns to its initial state (Figure 2).

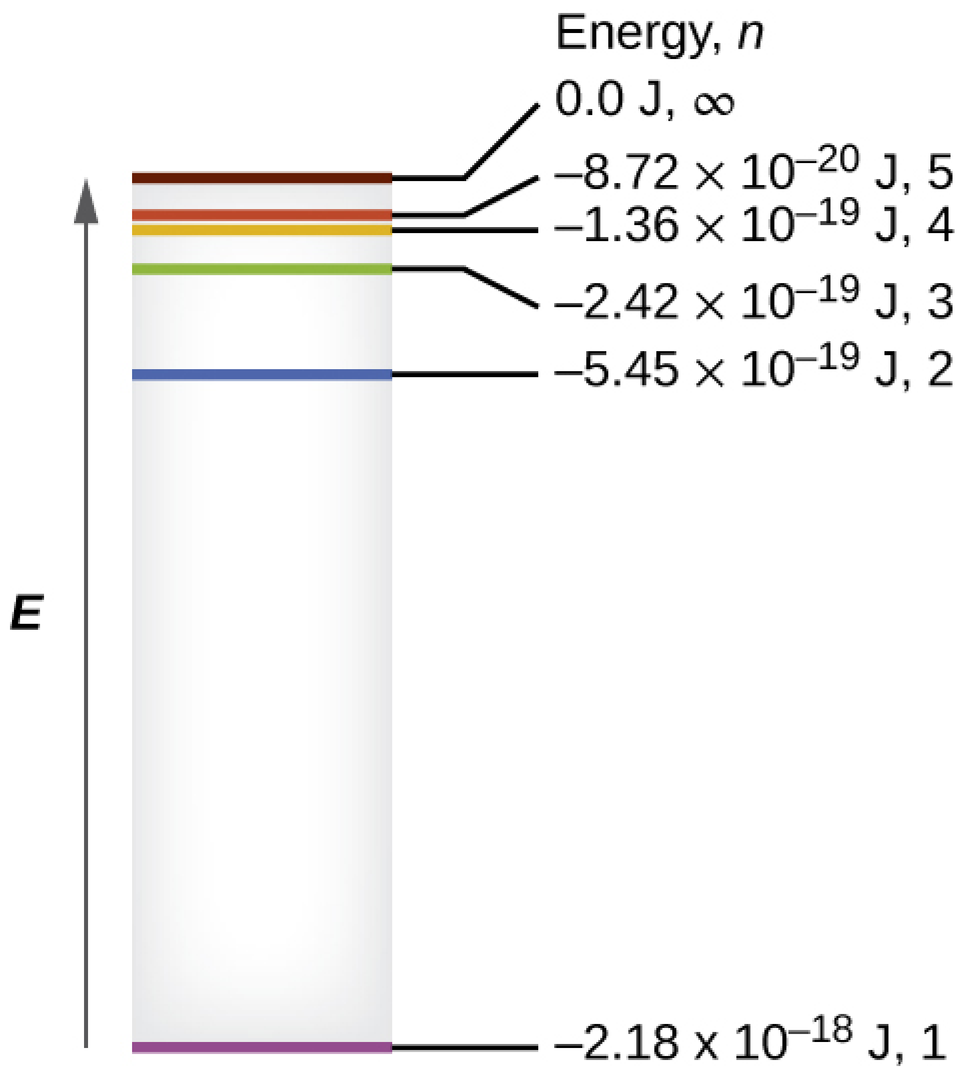


Figure 1: *Quantum numbers and energy levels in a hydrogen atom. The more negative the calculated value, the lower the energy.*

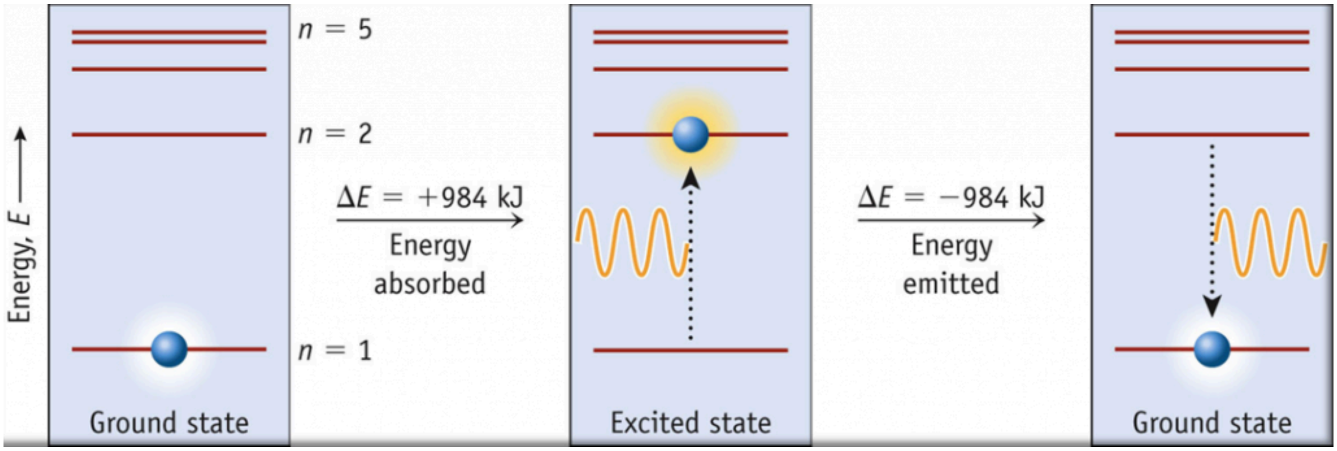


Figure 2: *An electron in hydrogen absorbing and then re-emitting a photon with an energy corresponding to the exact difference between two energy levels.*

Calculating the Energy and Wavelength of Electron Transitions in Hydrogen Using Bohr's Formula

What is the energy (in joules) and the wavelength (in meters) of the line in the spectrum of hydrogen that represents the movement of an electron from Bohr orbit with *n* = 6 to the orbit with *n* = 2? In what part of the electromagnetic spectrum do we find this radiation?

Solution

In this case, the electron starts out with *n* = 6, so $n_i = 6$. It comes to rest in the *n* = 2 orbit, so $n_f = 2$. The difference in energy between the two states is given by this expression:

$$\Delta E = E_f - E_i = Rhc\left(\frac{1}{n_f} - \frac{1}{n_i}\right)$$

$$\Delta E = (1.097 \times 10^7\,\mathrm{m}^{-1})(6.626 \times 10^{-34}\,\mathrm{J\cdot s})(3.00 \times 10^8\,\mathrm{m/s})\left(\frac{1}{2^2} - \frac{1}{6^2}\right)$$

$$\Delta E = -4.85 \times 10^{-19}\,\mathrm{J}$$

This energy difference is negative, indicating a photon leaves the system (is emitted) as the electron falls from the *n* = 6 orbit to the *n* = 2 orbit.

The wavelength of a photon with this energy is found by the expression $E = hc/$.

Rearrangement gives:

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

$$\lambda = \frac{(6.626\times10^{-34}\,\mathrm{J{\cdot}s})(3.00\times10^8\,\mathrm{m/s})}{4.85\times10^{-19}\,\mathrm{J}} = 4.10 \times 10^{-7}\,\mathrm{m}$$

From the illustration of the electromagnetic spectrum in [Basics of Light: The Main Parts of the Electromagnetic Spectrum](), we can see that this wavelength is found in the violet portion of the electromagnetic spectrum.

Thinking about Atomic Transitions from a Physics Perspective

In chemistry, the starting point for the analysis was generally

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

or if, you were looking at the hydrogen atom specifically,

$$\Delta E = -Rhc\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right).$$

Moreover, when solving the problem, you would probably just consider the absolute value of the change in energy and then figure out emission or absorption.

How would we look at this same problem in physics? In physics, we like to start with fundamental principles of the Universe and then apply definitions as described in Unit I On-A-Page: Principles and Definitions. In this case, the fundamental principle is Conservation of Energy: $\Delta E = Q + W$; this basic idea will, therefore be our starting point. To see how this works in practice, let's look at an example, the $n = 6 \rightarrow n = 2$ transition discussed in the video.

The $n = 6 \rightarrow n = 2$ hydrogen transition from a physics perspective

Problem: What wavelength of light is does an electron in the $n = 6$ state of the hydrogen atom emit if it falls to the $n = 2$ state?

Solution – Starting with Principles:

First we identify the relevant fundamental principle of the Universe. In this situation, we know the relevant principle will be conservation of energy as the problem describes an electron losing energy as it moves from one state to another. Thus, we begin with

$$\Delta E = Q + W.$$

Expanding out the Δ,

$$E_f - E_i = Q + W.$$

Now we need to consider the *physics of the situation*, or as I often call it, "what is the story that describes our phenomenon?" In this case, the story is:

> *I begin with an electron in a high energy state (n = 6). This electron then falls to a lower energy state (n=2). Since energy is conserved, that means that the energy lost by the electron has to go somewhere. In this case, the energy leaves the atom as light.*

Having this *story* helps us fill in the parts of our conservation of energy equation:

- E_f: is the energy of the electron at the end. In our case, the energy of the n = 2 state: $E_{(n=2)}$.
- E_i: is the energy of the electron at the beginning. In our case, the energy of the n = 6 state: $E_{(n=6)}$.
- Energy is leaving the atom through microscopic interactions, i.e. the emission of a photon. Thus, Q is the energy of the photon E_γ.
- There are no forces being applied to the atom. Since work is a force applied for some distance, $W = Fd\cos\theta$, we know that the work in this case is zero: $W = 0$.

Substituting our story into our equation ($E_f \rightarrow E_{(n=2)}$, $E_i \rightarrow E_{(n=6)}$, $Q \rightarrow E_\gamma$, and $W = 0$) then yields:

$$E_{(n=2)} - E_{(n=6)} = E_\gamma + 0.$$

Now add definitions:

From what we have covered, thus far, we can *now* put in definitions for the various quantities:

- We know, from reviewing chemistry in the section above, that the energy of an electron in the hydrogen atom is given by $E_n = \frac{-Rhc}{n^2}$.

We can now put these in definitions into our equation:

$$\left(\frac{-Rhc}{2^2}\right) - \left(\frac{-Rhc}{6^2}\right) = E_\gamma.$$

Notice, as always, we are working in symbols which is one of our goals of the course. Now, we have a math problem! We are looking for λ, which we can get from E_γ, and h, c, and R are all just constants of nature. We begin by noticing that we can factor out the $-Rhc$ on the left hand side:

$$-Rhc\left(\frac{1}{2^2} - \frac{1}{6^2}\right) = \frac{1}{\lambda}.$$

(This should now be starting to look like what you reviewed from chemistry!) We, of course, know 2^2 and 6^2, converting these and substituting in the values for $R = 1.097 \times 10^7\,\mathrm{m}^{-1}$, $c = 3 \times 10^8\,\mathrm{m/s}$, and $h = 6.626 \times 10^{-34}\,\mathrm{J \cdot s}$ we have:

$$\left(1.097 \times 10^7\,\mathrm{m}^{-1}\right)\left(6.626 \times 10^{-34}\,\mathrm{J \cdot s}\right)\left(c = 3 \times 10^8\,\mathrm{m/s}\right) \times \left(\frac{1}{4} - \frac{1}{36}\right) = E_\gamma$$

$$E_\gamma = -4.85 \times 10^{-19}\,\mathrm{J} = -3.02\,\mathrm{eV}$$

Notice, we never took any absolute values, and the result was negative. This negative sign has meaning! It tells us that the energy was lost to the atom! Of course, we already knew that, but this is a way to check our answers: we know energy is lost and we get a negative answer. The math takes care of itself!

Now we know from Basics of Light: Photon Momentum – Relationship to Energy that the energy of a photon is $E_\gamma = \frac{hc}{\lambda}$, and can get a wavelength:

$$E_\gamma = \frac{hc}{\lambda}$$

Here we *will* take an absolute value, because negative wavelengths, unlike negative energies, don't make sense. So we have

$$\lambda = \frac{hc}{E_\gamma}$$

$$\lambda = \frac{\left(6.626 \times 10^{-34}\,\mathrm{J \cdot s}\right)\left(c = 3 \times 10^8\,\mathrm{m/s}\right)}{4.85 \times 10^{-19}\,\mathrm{J}}$$

$$\lambda = 4.10 \times 10^{-7}\,\mathrm{m} = 410\,\mathrm{nm}$$

Exactly the blue line described in the video.

Why we do it this way

You may be saying to yourself, that the physics method seems a lot longer with a lot more steps. Why not just follow chemistry and start with

$$\frac{hc}{\lambda} = -Rhc\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)?$$

The reason for the physics approach is that it is a lot more versatile: it can be applied to loads of situations from x-ray emissions, to photo electric effects, to LEDs, to gravitational redshifts, to particle/anti-particle annihilations, to..., you get the idea. In contrast,

$$\frac{hc}{\lambda} = -Rhc\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

works ONLY for the wavelengths of photons and ONLY for hydrogen atoms. Learning to think in the physics perspective is one of the goals for this course!

Homework Problem

Problem 25: Determine the wavelength of the third Balmer line (transition from n = 5 to n = 2).

9. Matter as a Wave

De Broglie Wavelength

In 1923 a French physics graduate student named Prince Louis-Victor de Broglie (1892–1987) made a radical proposal based on the hope that nature is symmetric. If EM radiation has both particle and wave properties, then nature would be symmetric if matter also had both particle and wave properties. If what we once thought of as an unequivocal wave (EM radiation) is also a particle, then what we think of as an unequivocal particle (matter) may also be a wave. De Broglie's suggestion, made as part of his doctoral thesis, was so radical that it was greeted with some skepticism. A copy of his thesis was sent to Einstein, who said it was not only probably correct, but that it might be of fundamental importance. With the support of Einstein and a few other prominent physicists, de Broglie was awarded his doctorate.

De Broglie took both relativity and quantum mechanics into account to develop the proposal that *all particles have a wavelength*, given by

$p = \frac{h}{\lambda}$ (matter and photons)λ=hp(matter and photons),λ=hp(matter and photons), size 12{λ = { {h} over {p} } } {}">,

where hh size 12{h} {}">h is Planck's constant and pp size 12{p} {}">p is momentum. This is defined to be the de Broglie wavelength. (Note that we already have this for photons, from the equation p=h/λp=h/λ size 12{p = h/λ} {}">p=h/λ.) The hallmark of a wave is interference. If matter is a wave, then it must exhibit constructive and destructive interference. Why isn't this ordinarily observed? The answer is that in order to see significant interference effects, a wave must interact with an object about the same size as its wavelength. Since hh size 12{h} {}">h is very small, λλ size 12{λ} {}">λ is also small, especially for macroscopic objects. A 3-kg bowling ball moving at 10 m/s, for example, has

$$p = \frac{h}{\lambda}$$
$$\lambda = \frac{h}{p}$$
$$\lambda = \frac{h}{mv}$$

$\lambda = \frac{6.626\times10^{-34}\,\mathrm{J\cdot s}}{(3\,\mathrm{kg})(10\,\mathrm{m/s})} = 2\times10^{-35}\,\mathrm{m}$λ=h/p=(6.63×10–34J·s)/[(3 kg)(10 m/s)]=2×10–35m.λ=h/p= (6.63 × 10–34 J·s)/[(3 kg)(10 m/s)] = 2 × 10–35 m. size 12{λ = h/p"= " \(6 "." "63 " times " 10" rSup { size 8{"–34"} } " J·s" \) / \[\("3kg" \) \("10 m/s" \) " = 2 " times " 10" rSup { size 8{"–35"} } " m"} {}">.

This means that to see its wave characteristics, the bowling ball would have to interact with something about 10–35m 10–35 m size 12{" 10" rSup { size 8{"–35"} } " m"} {}">10^{-35}m in size—far smaller than anything known. When waves interact with objects much larger than their wavelength, they show negligible interference effects and move in straight lines (such as light rays in geometric optics). To get easily observed interference effects from particles of matter, the longest wavelength and hence smallest mass possible would be useful. Therefore, this effect was first observed with electrons.

Connections: Waves

All microscopic particles, whether massless, like photons, or having mass, like electrons, have wave properties. The relationship between momentum and wavelength is fundamental for all particles.

American physicists Clinton J. Davisson and Lester H. Germer in 1925 and, independently, British physicist G. P. Thomson (son of J. J. Thomson, discoverer of the electron) in 1926 scattered electrons from crystals and found diffraction patterns. These patterns are exactly consistent with interference of electrons having the de Broglie wavelength and are somewhat analogous to light interacting with a diffraction grating. (See Figure 1.)

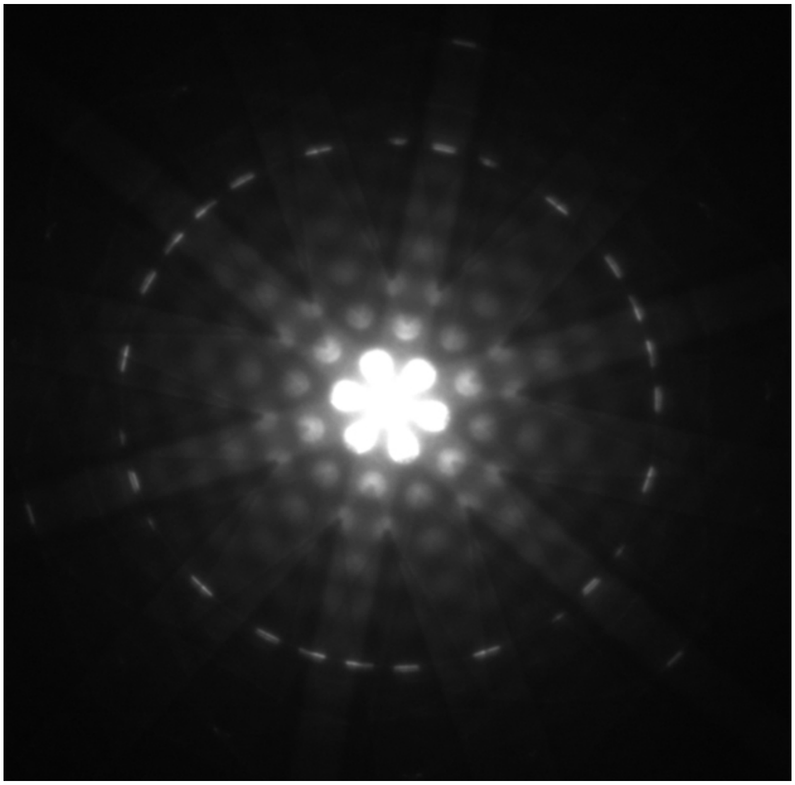


Figure 1: This diffraction pattern was obtained for electrons diffracted by crystalline silicon. Bright regions are those of constructive interference, while dark regions are those of destructive interference. (credit: Ndthe, Wikimedia Commons)

De Broglie's proposal of a wave nature for all particles initiated a remarkably productive era in which the foundations for quantum mechanics were laid. In 1926, the Austrian physicist Erwin Schrödinger (1887–1961) published four papers in which the wave nature of particles was treated explicitly with wave equations. At the same time, many others began important work. Among them was German physicist Werner Heisenberg (1901–1976) who, among many other contributions to quantum mechanics, formulated a mathematical treatment of the wave nature of matter that used matrices rather than wave equations. We will deal with some specifics in later sections, but it is worth noting that de Broglie's work was a watershed for the development of quantum mechanics. De Broglie was awarded the Nobel Prize in 1929 for his vision, as were Davisson and G. P. Thomson in 1937 for their experimental verification of de Broglie's hypothesis.

Electron Wavelength versus Velocity and Energy

For an electron having a de Broglie wavelength of 0.167 nm (appropriate for interacting with crystal lattice structures that are about this size): (a) Calculate the electron's velocity. (b) Calculate the electron's kinetic energy in eV.

Strategy

For part (a), since the de Broglie wavelength is given, the electron's velocity can be obtained from
λ=h/p

by using the nonrelativistic formula for momentum, p=mv. For part (b), once v is obtained (and it has been verified that v is nonrelativistic), the classical kinetic energy is simply (1/2)mv2.

Solution for (a)

Substituting the formula for momentum (p=mv) into the de Broglie wavelength gives

p=h/λ
mv=h/λ

Solving for v gives

v=h/mλ

Substituting known values yields

$$v=\frac{6.63\times10^{-34}\,\mathrm{J{\cdot}s}}{(9.11\times10^{-31}\,\mathrm{kg})(0.167\times10^{-9}\,\mathrm{m})}=4.36\times10^{6}\,\mathrm{m/s}.$$

Solution for (b)

While fast compared with a car, this electron's speed is not close to the speed of light, and so we can comfortably use the classical formula to find the electron's kinetic energy and convert it to eV as requested.

$$K=\frac{1}{2}mv^2$$
$$K=\frac{1}{2}(9.11\times10^{-31}\,\mathrm{kg})(4.36\times10^{6}\,\mathrm{m/s})^2$$
$$K=86.410^{-18}\,\mathrm{J}=54.0\,\mathrm{eV}$$

Electron Microscopes

One consequence or use of the wave nature of matter is found in the electron microscope. As we have discussed, there is a limit to the detail observed with any probe having a wavelength. Resolution, or observable detail, is limited to about one wavelength. Since a potential of only 54 V can produce electrons with sub-nanometer wavelengths, it is easy to get electrons with much smaller wavelengths than those of visible light (hundreds of nanometers). Electron microscopes can, thus, be constructed to detect much smaller details than optical microscopes. (See Figure 2.)

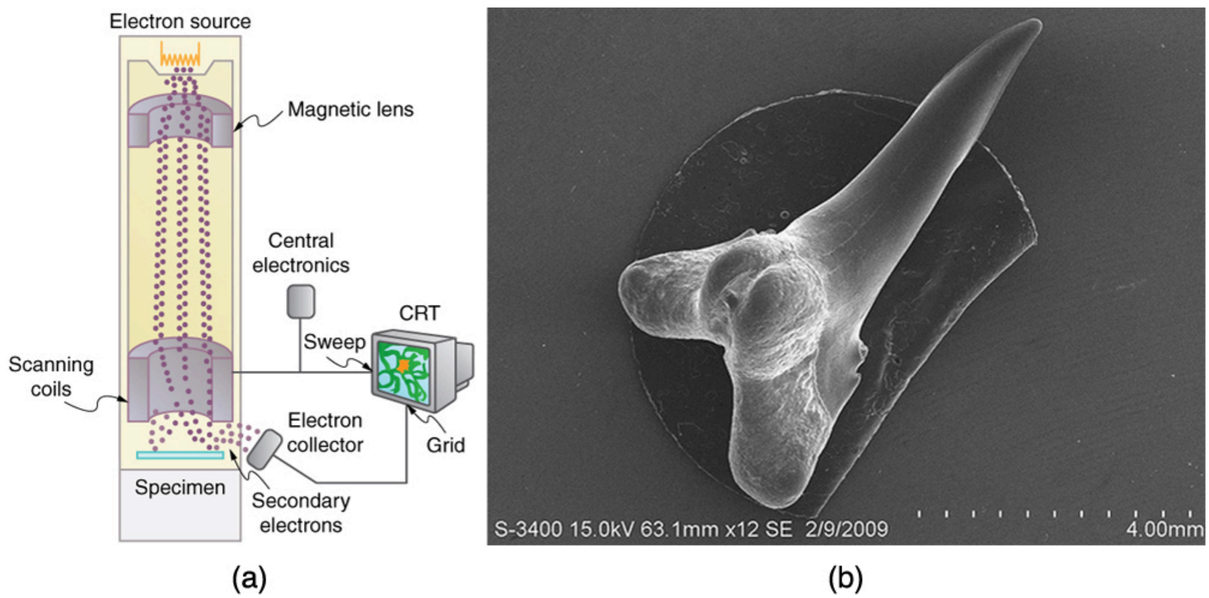


Figure 2: *Schematic of a scanning electron microscope (SEM) (a) used to observe small details, such as those seen in this image of a tooth of a Himipristis, a type of shark (b). (credit: Dallas Krentzel, Flickr)*

There are basically two types of electron microscopes. The transmission electron microscope (TEM) accelerates electrons that are emitted from a hot filament (the cathode). The beam is broadened and then passes through the sample. A magnetic lens focuses the beam image onto a fluorescent screen, a photographic plate, or (most probably) a CCD (light sensitive camera), from which it is transferred to a computer. The TEM is similar to the optical microscope, but it requires a thin sample examined in a vacuum. However it can resolve details as small as 0.1 nm (10−10m10−10m size 12{"10" rSup { size 8{ - "10"} } ` m} {}">10^{-10}m), providing magnifications of 100 million times the size of the original object. The TEM has allowed us to see individual atoms and structure of cell nuclei.

The scanning electron microscope (SEM) provides images by using secondary electrons produced by the primary beam interacting with the surface of the sample (see Figure 2). The SEM also uses magnetic lenses to focus the beam onto the sample. However, it moves the beam around electrically to “scan” the sample in the *x* and *y* directions. A CCD detector is used to process the data for each electron position, producing images like the one at the beginning of this chapter. The SEM has the advantage of not requiring a thin sample and of providing a 3-D view. However, its resolution is about ten times less than a TEM.

Electrons were the first particles with mass to be directly confirmed to have the wavelength proposed by de Broglie. Subsequently, protons, helium nuclei, neutrons, and many others have been observed to exhibit interference when they interact with objects having sizes similar to their de Broglie wavelength. The de Broglie wavelength for massless particles was well established in the 1920s for photons, and it has since been observed that all massless particles have a de Broglie wavelength pλ=h/p.λ=h/p. size 12{λ = h/p} {}">=h/λ.

The wave nature of all particles is a universal characteristic of nature. We shall see in following sections that implications of the de Broglie wavelength include the quantization of energy in atoms and molecules, and an alteration of our basic view of nature on the microscopic scale. The next section, for example, shows that there are limits to the precision with which we may make predictions, regardless of how hard we try. There are even limits to the precision with which we may measure an object’s location or energy.

Making Connections: A Submicroscopic Diffraction Grating

The wave nature of matter allows it to exhibit all the characteristics of other, more familiar, waves. Diffraction gratings, for example, produce diffraction patterns for light that depend on grating spacing and the wavelength of the light. This effect, as with most wave phenomena, is most pronounced when the wave interacts with objects having a size similar to its wavelength. For gratings, this is the spacing between multiple slits.) When electrons interact with a system having a spacing similar to the electron wavelength, they show the same types of interference patterns as light does for diffraction gratings, as shown at top left in Figure 3.

Atoms are spaced at regular intervals in a crystal as parallel planes, as shown in the bottom part of Figure 3. The spacings between these planes act like the openings in a diffraction grating. At certain incident angles, the paths of electrons scattering from successive planes differ by one wavelength and, thus, interfere constructively. At other angles, the path length differences are not an integral wavelength, and there is partial to total destructive interference. This type of scattering from a large crystal with well-defined lattice planes can produce dramatic interference patterns. It is called *Bragg reflection*, for the father-and-son team who first explored and analyzed it in some detail. The expanded view also shows the path-length differences and indicates how these depend on incident angle θθ size 12{θ} {}">θ in a manner similar to the diffraction patterns for x rays reflecting from a crystal.

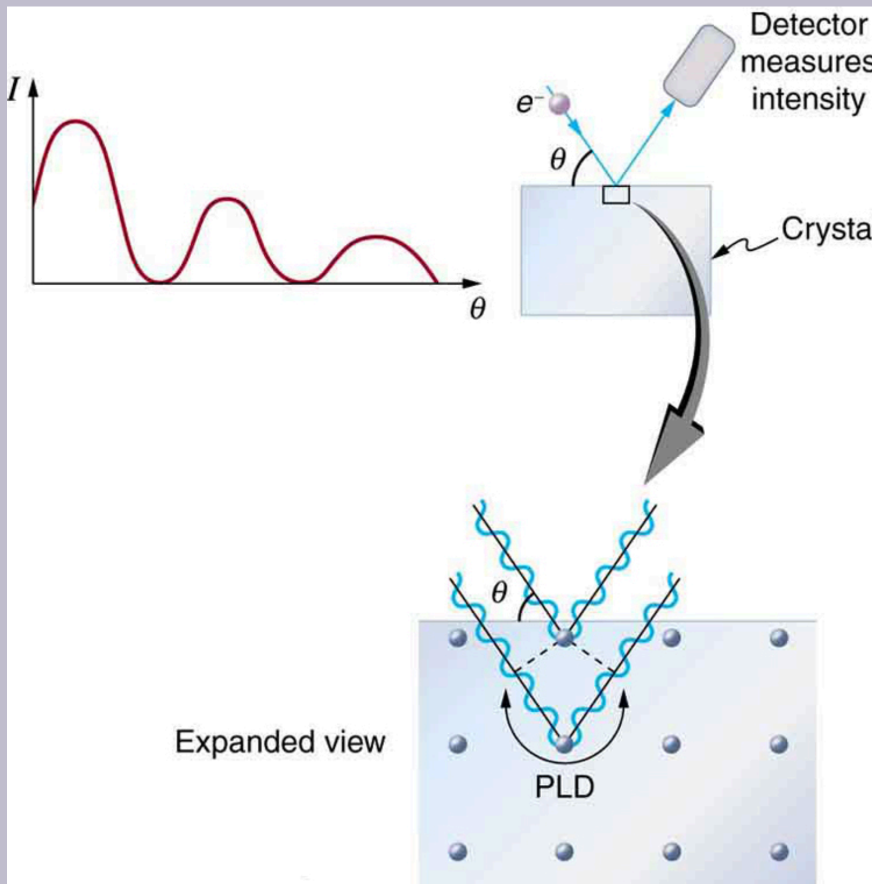


Figure 3: *The diffraction pattern at top left is produced by scattering electrons from a crystal and is graphed as a function of incident angle relative to the regular array of atoms in a crystal, as shown at bottom. Electrons scattering from the second layer of atoms travel farther than those scattered from the top layer. If the path length difference (PLD) is an integral wavelength, there is constructive interference.*

Section Summary

- Particles of matter also have a wavelength, called the de Broglie wavelength, given by pλ=hpλ=hp size 12{λ = { {h} over {p} } } {}">=h/λ, where pp size 12{p} {}">p is momentum.
- Matter is found to have the same *interference characteristics* as any other wave.

Homework Problems

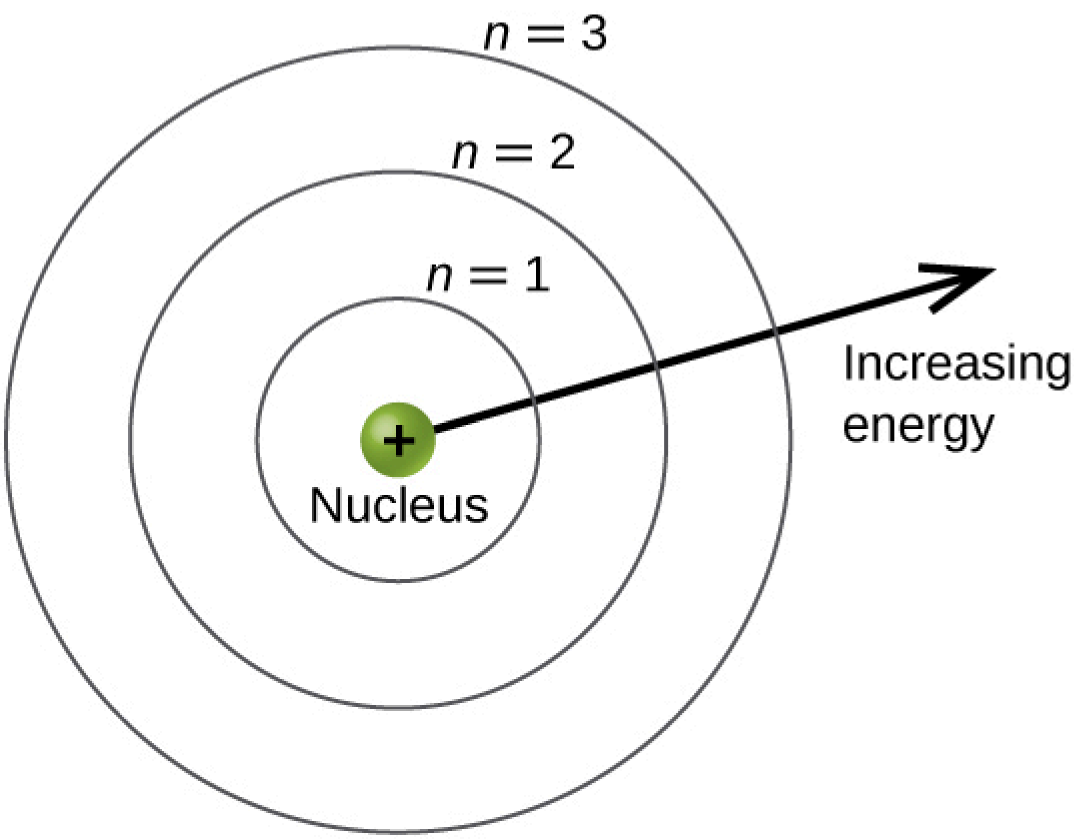
Problem 26: Find the wavelength of a golf ball.

Problem 27: Given an electron’s wavelength, what is its speed?

10. Fundamentals of "Particle in a Box"

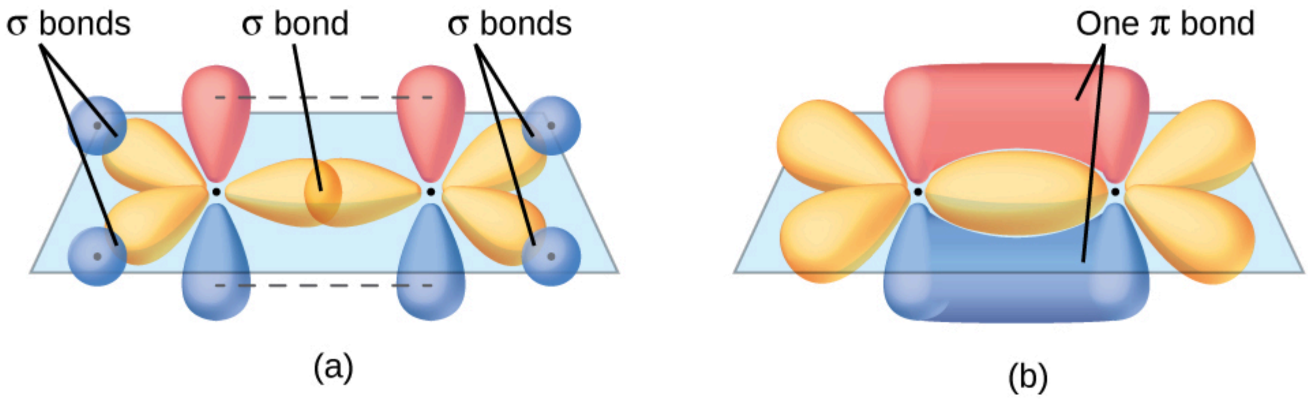
Boxes and Electrons in Atoms: The Essential Questions

Our goal for this unit is to understand why electrons in atoms exist in discrete energy levels; you probably know from chemistry that they do. Typically, these discrete energy levels are depicted as in the figure below. The question we want to answer is, "*WHY*?" Why do electrons only have specific energies?



Electrons in atoms exist in discrete energy levels. Different shells are numbered by principal quantum numbers. Figure 6.19 in OpenStax Chemistry 2e.

You have also probably been exposed in chemistry class to molecular orbital shapes like the one below. Again, we want to know "*WHY*?" Why do the orbitals for electrons in such molecules have these shapes? Can we predict what the orbitals will look like for different energy levels?



In the ethene molecule, C_2H_4, there are (a) five **σ** *bonds. One C–C* **σ** *bond results from overlap of sp^2 hybrid orbitals on the carbon atom with one sp^2 hybrid orbital on the other carbon atom. Four C–H bonds result from the overlap between the C atoms' sp^2 orbitals with s orbitals on the hydrogen atoms. (b) The π bond is formed by the side-by-side overlap of the two unhybridized p orbitals in the two carbon atoms. The two lobes of the π bond are above and below the plane of the* **σ** *system. [Figure 8.23 in OpenStax Chemistry 2e](#).*

To explore both of these sets of questions, we will investigate a simpler situation: what happens when you put an electron in a simple box. A box may not seem to have much in common with an atom at first glance, but we will see in class how these two situations are related.

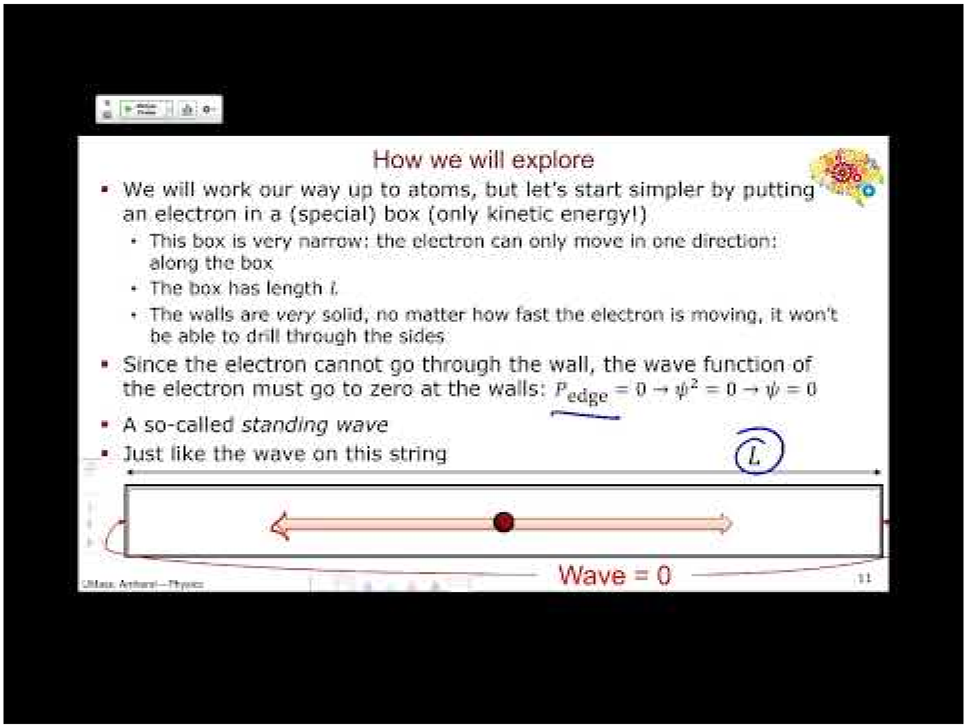
University of Massachusetts Amherst BE REVOLUTIONARY

Instructor's Notes

What I want you to take away from this reading, are the specifics details of the box we will consider and how this impacts the particle-as-wave. These ideas will be the context of your homework problems and your in-class quiz. We will then use this model to understand why electrons have discrete energy levels and how we can predict the shapes of molecular orbitals.

The particulars of our box model

The following text is available as both video and text. The content is the same. Feel free to engage however suits you best.



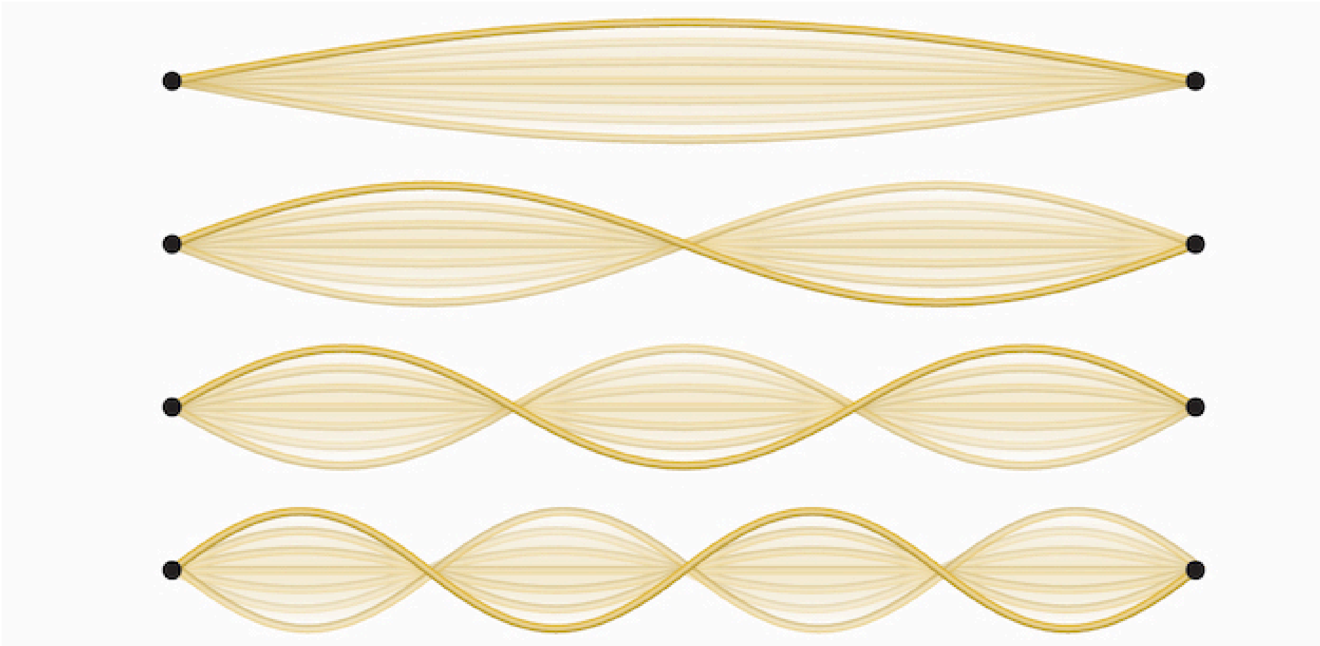
A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=1727

Why do electrons in atoms have energy levels? You know they do from chemistry, right? Electrons exist in discrete energy levels. Why? To answer this we're going to be physicists and do something simple and stupid first: we're going to put a particle in a box. I know you're thinking that sounds dumb, but it works! So, we're going to put a particle in a box. In fact, we're going to put it in a very special box:

1. the box has only one dimension to it – it has a length L. The electron can only move along the box – it it can't move in any other direction.
2. the electron cannot got through the walls of the box – it cannot drill through them.

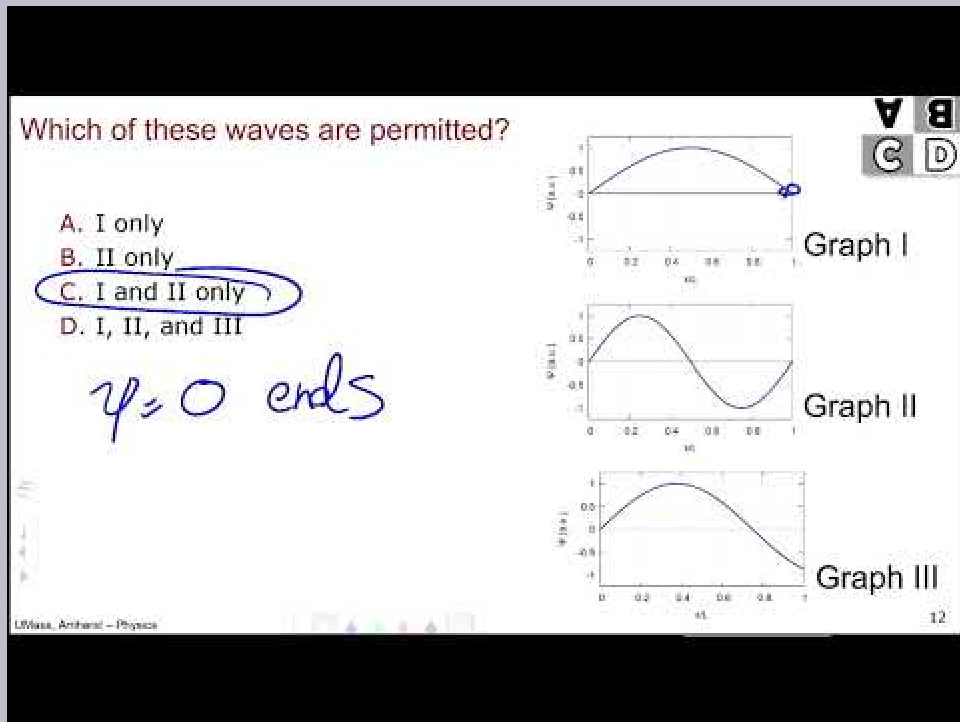
For those of you thinking this is unrealistic and goofy, long carbon chains act like this. Beta-carotene, those electrons act like they're in a box very very well. Similarly, 1-3 butadiene, propene, etc.: for all of these, the electron in a box is a very good model.

The fact that the particle cannot drill through the walls tells us that the amplitude of the wave associated with the particle ψ must go to zero at the ends. As we will see in class, the amplitude ψ is related to the probability of finding the particle at a given spot. Since the electron cannot go through the walls, the amplitude must be zero at the ends. The result is a so-called standing wave just like the one on the string in the figure below: it is waving up and down, but goes to zero on both ends.



For a string fixed on both ends, only certain waves, so-called standing waves, are possible as the wave must be zero on the ends. [Figure 6.7 from OpenStax Chemistry 2e](#).

Examples



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=1727*](http://openbooks.library.umass.edu/toggerson-132/?p=1727)

Key Takeaways

The key elements of our box model that I want you to take away are:

- The box is one-dimensional: it has only a length L and no other dimensions.
  - Consequently, a particle in the box can only move along its length.
- The walls of the box are such that the contained particle cannot "drill through" the walls.
  - Consequently, the amplitude ψ of the wave-aspect of the particle must be zero at both ends.

- The waves on a string are a good way to visualize this.
- While such a box may seem silly, it is a good model for many long organic molecules.

Homework Problems

Problem 28: Which graphs are allowed for a particle in a box.

11. All Homework Problems

Homework

The list below is the list of homework problems in Edfinity. The numbering is the same. You can click on a problem, and it will take you to the relevant section of the book!

1. What is the charge of the Δ?
2. How much energy is produced in the sun from converting 2 protons and two neutrons into a helium nucleus?
3. How much energy is released when an electron and an anti-electron (positron) annihilate during positron emission tomography?
4. Compare the momenta of elephants, humans, and tranquilizer darts!
5. Assuming negligible air resistance, what is the final speed of a rock thrown from a bridge?
6. How many DNA molecules can a single electron from an old-fashioned TV break?
7. How long can you play tennis off a candy bar?
8. How long for a car of fixed power to get up to speed?
9. Which element has an outer electron with the lowest potential energy?
10. Exploring the relationship between momentum and kinetic energy.
11. What is the frequency of a stroboscope?
12. Label the parts of a wave.
13. If the frequency of a wave is changed, which of the other properties must also change assuming the speed of the wave remains fixed?
14. Speed, wavelength, and frequency for sound.
15. How long to collect a certain amount of sunlight?
16. What is the power output of an ultrasound machine for a given intensity and area?
17. Which situations will create electromagnetic radiation?
18. Speed dependencies for electromagnetic waves.
19. What is the frequency of a radio station given the wavelength?
20. Rank the types of waves in the EM spectrum by wavelength.
21. Find the momentum of a microwave photon.
22. From momentum, calculate the wavelength and energy of a photon.
23. An AM radio transmitter radiates some power at a given frequency. How many photons per second does the emitter emit?
24. If the brightness of a beam of light is increased, the ________ of the ____________ will also increase.
25. Determine the wavelength of the third Balmer line (transition from n = 5 to n = 2).
26. Find the wavelength of a golf ball.
27. Given an electron's wavelength, what is its speed?
28. Which graphs are possible for a particle in a box?

PART II

UNIT II

Unit II On-a-Page

Terminology

Instructor's Note

This Unit is very heavy on vocabulary, there is a set on flashcards on [Quizlet](#) to help you.

- An **optical element** is a lens or mirror.
- An **image** is the apparent reproduction of an object, formed by an optical element (or collection of them) reflecting and/or refracting light.
- Images can either be **erect** (same orientation as object) or **inverted** (upside-down with respect to object).
- The optical axis is a line that passes through the optical element perpendicular to it
- The point where the optical axis meets the optical element is called the **vertex**.
- The **center** of a lens is where the lens is thickest (for a converging lens) or thinnest (for a diverging lens).
- The **center** of a mirror is the center of curvature.
- Image i and object o distances are measured along the optical axis.
- Signs are relative to the path of the light.
- If the object is on the same side as the incoming light, then $o > 0$ 0 " title="Rendered by QuickLaTeX.com" height="12" width="38" style="vertical-align: 0px;">, otherwise $olt; 0$.
- If the image is on the same side as the outgoing light, then $i > 0$ 0 " title="Rendered by QuickLaTeX.com" height="12" width="36" style="vertical-align: 0px;">, otherwise $ilt; 0$.
- For a lens, the incoming and outgoing sides are different (light goes through a lens).
- For a mirror, the incoming and outgoing sides are the same (light bounces off a mirror).
- Focal lengths are also signed. If the element tends to **diverge** incoming parallel light (a **diverging lens** or a **convex mirror**), then the focal length is negative. If the element **converges** incoming parallel light (**convex lens** or **concave mirror**), the focal length is positive.

Principles for Unit II

This unit in particular has a lot of terminology. To help you stay focused, the principles (where we will begin analyzing situations) are:

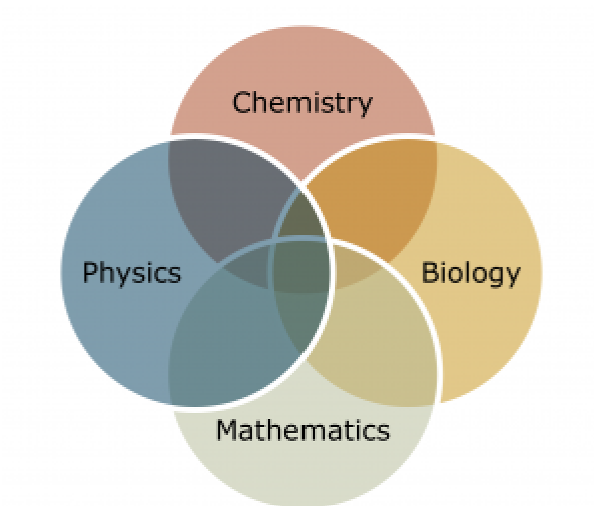
How light interacts with surfaces and materials

- Law of reflection which is best understood in the particle picture.
- Law of refraction which is best understood in the wave picture.
- Light slows when it enters a medium, but the energy of a photon cannot change. So the wavelength must!

Optical elements and ray diagrams

- The position of the image formed by an optical system is located at $\frac{1}{i}+\frac{1}{o}=\frac{1}{f}$ where:
    - o is the object distance, using the sign conventions used above.
    - i is the image distance, using the sign conventions used above.
    - f is the focal length with its correct sign.
- For ray diagrams: one ray in parallel and out through focal point, one ray in through focal point and out parallel, one ray using the center of the system.
- And point where photons seem to emerge can be used as an object. Such a point could be a real source of photons (an object) or an image from another optical element.

12. Motivating Context for Unit II



This unit will focus on how light (and in some cases electrons) travel through both empty space and matter as well as how those interactions can be used to manipulate the paths and make **images**. The most familiar optical system to most of you is the one you are probably using to read these very words: your eyes! We will therefore be looking at the eye quite a bit throughout this unit, both human eyes and simpler eyes in the animal kingdom. Another common biological application of optics is the microscope. To make sure that everyone is on the same page, you will find below some information about the anatomy of the human eye as well as some basic information about microscopes. Please be familiar with this terminology as we will use it in class.

The Human Eye. Derived from 36.5 Vision by OpenStax *Biology*

Vision is the ability to detect light patterns from the outside environment and interpret them into images. Animals are bombarded with sensory information, and the sheer volume of visual information can be problematic. Fortunately, the visual systems of species have evolved to attend to the most-important stimuli. The importance of vision to humans is further substantiated by the fact that about one-third of the human cerebral cortex is dedicated to analyzing and perceiving visual information.

Anatomy of the Eye

The photoreceptive cells of the eye, where the conversion of light to nervous impulses occurs, are located in the **retina** (shown in Figure 2) on the inner surface of the back of the eye. But light does not impinge on the retina unaltered. It passes through other layers that process it so that it can be interpreted by the retina (Figure 2b). The **cornea**, the front transparent layer of the eye, and the crystalline **lens**, a transparent convex structure behind the cornea, both refract (bend) light to focus the image on the retina. The **iris**, which is conspicuous as the colored part of the eye, is a circular muscular ring lying between the lens and cornea that regulates the amount of light entering the eye. In conditions of high ambient light, the iris contracts, reducing the size of the

pupil at its center. In conditions of low light, the iris relaxes and the pupil enlarges.

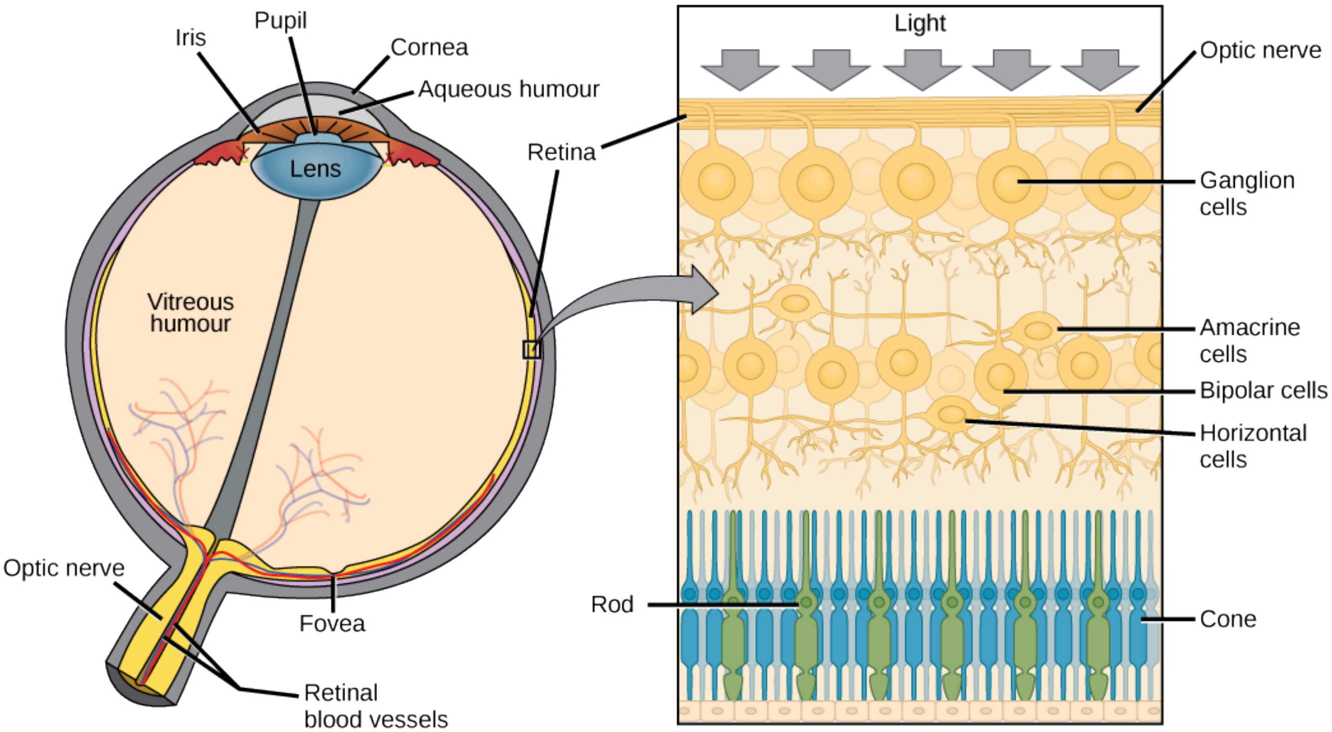


Figure 2. *(a) The human eye is shown in cross section. (b) A blowup shows the layers of the retina.*

Changes in material are crucial to image formation using lenses. Each material has an index of refraction which will be discussed in [The Ray Aspect of Light: The Speed of Light in Materials](). The biggest change in material, and consequently the biggest bending of rays, actually occurs at the cornea rather than the lens. The cornea provides about two-thirds of the power of the eye, owing to the fact that speed of light changes considerably while traveling from air into cornea. The lens provides the remaining power needed to produce an image on the retina. The cornea and lens can be treated as a single thin lens, even though the light rays pass through several layers of material (such as cornea, aqueous humor, several layers in the lens, and vitreous humor), changing direction at each interface. The image formed is much like the one produced by a single convex lens.

The goal of the cornea and lens is to focus light on the retina and, in particular, the fovea centralis. The fovea centralis is a small, central pit composed of closelpacked cones in the eye. The fovea is responsible for sharp central vision, which is necessary in humans for activities for which visual detail is of primary importance, such as reading and driving. The lens is dynamic, focusing and re-focusing light as the eye rests on near and far objects in the visual field. The lens is operated by muscles that stretch it flat or allow it to thicken, changing the focal length of light coming through it to focus it sharply on the retina. With age comes the loss of the flexibility of the lens, and a form of farsightedness called **presbyopia** results. Presbyopia occurs because the image focuses behind the retina. Presbyopia is a deficit similar to a different type of farsightedness called **hyperopia** caused by an eyeball that is too short. For both defects, images in the distance are clear but images nearby are blurry. **Myopia** (nearsightedness) occurs when an eyeball is elongated and the image focus falls in front of the retina. In this case, images in the distance are blurry but images nearby are clear.

There are two types of photoreceptors in the retina: **rods** and **cones**, named for their general appearance as illustrated in . Rods are strongly photosensitive and are located in the outer edges of the retina. They detect dim light and are used primarily for peripheral and nighttime vision. Cones are weakly photosensitive and are located near the center of the retina. They respond to bright light, and their primary role is in daytime, color vision.

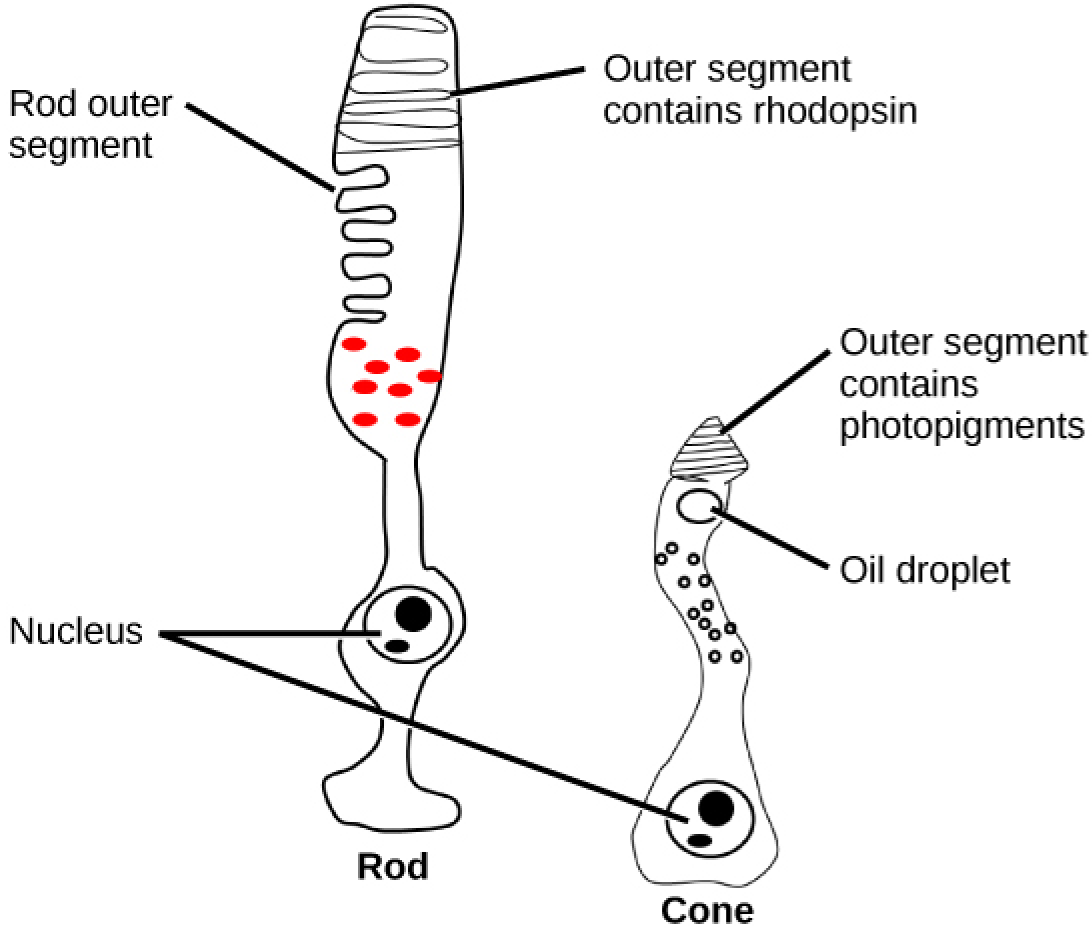


Figure 3. *Rods and cones are photoreceptors in the retina. Rods respond in low light and can detect only shades of gray. Cones respond in intense light and are responsible for color vision. (credit: modification of work by Piotr Sliwa)*

The **fovea** is the region in the center back of the eye that is responsible for acute vision. The fovea has a high density of cones. When you bring your gaze to an object to examine it intently in bright light, the eyes orient so that the object's image falls on the fovea. However, when looking at a star in the night sky or other object in dim light, the object can be better viewed by the peripheral vision because it is the rods at the edges of the retina, rather than the cones at the center, that operate better in low light. In humans, cones far outnumber rods in the fovea.

Homework Problems

Problem 1: Identify the portion of the eye responsible for most of the focusing of light.

Problem 2: What is the purpose of the iris?

Transduction of Light

The rods and cones are the site of transduction of light to a neural signal. Both rods and cones contain photopigments. In vertebrates, the main photopigment, **rhodopsin**, has two main parts (Figure 4) an opsin, which is a membrane protein (in the form of a cluster of α-helices that span the membrane), and retinal—a molecule that absorbs light. When light hits a photoreceptor, it causes a shape change in the retinal, altering its structure from a bent (cis) form of the molecule to its linear (trans) isomer. This isomerization of retinal activates the rhodopsin, starting a cascade of events that ends with the closing of Na^+ channels in the membrane of the photoreceptor. Thus, unlike most other sensory neurons (which become depolarized by exposure to a stimulus) visual receptors become hyperpolarized and thus driven away from threshold (Figure 5).

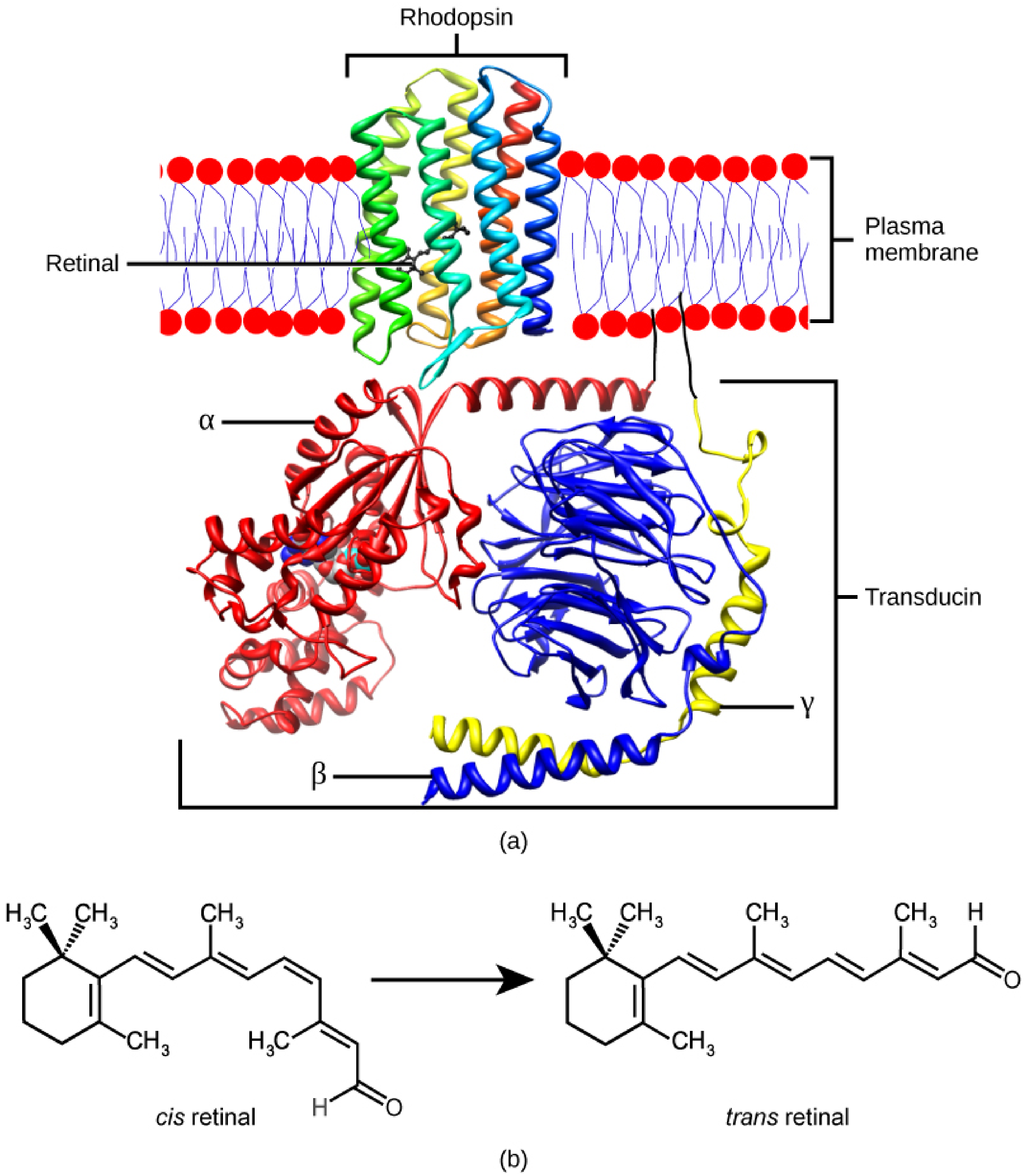


Figure 4. *(a) Rhodopsin, the photoreceptor in vertebrates, has two parts: the trans-membrane protein opsin, and retinal. When light strikes retinal, it changes shape from (b) a cis to a trans form. The signal is passed to a G-protein called transducin, triggering a series of downstream events.*

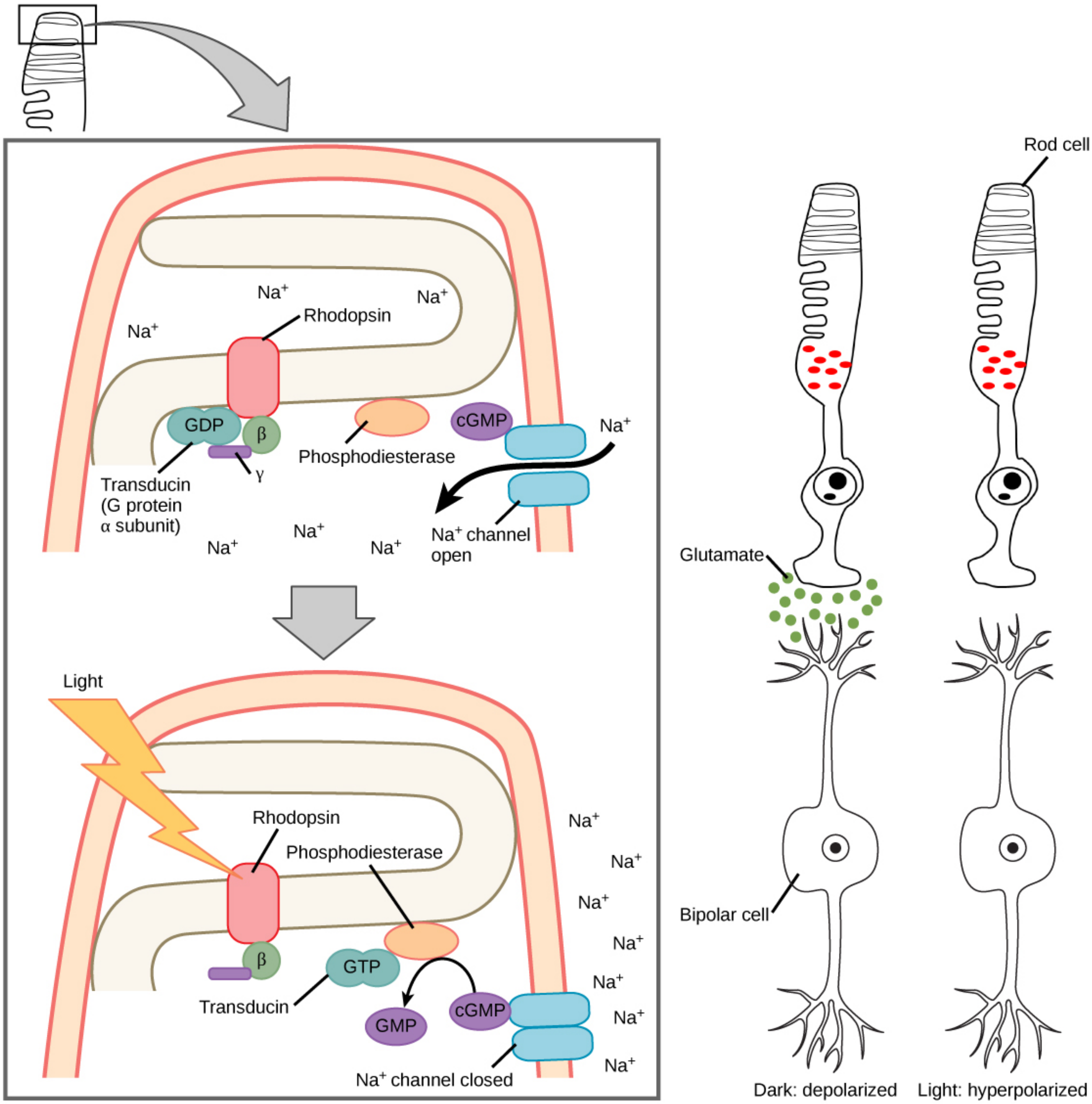


Figure 5. *When light strikes rhodopsin, the G-protein transducin is activated, which in turn activates phosphodiesterase. Phosphodiesterase converts cGMP to GMP, thereby closing sodium channels. As a result, the membrane becomes hyperpolarized. The hyperpolarized membrane does not release glutamate to the bipolar cell.*

Trichromatic Coding

There are three types of cones (with different photopsins), and they differ in the wavelength to which they are most responsive, as shown in Figure 6 . Some cones are maximally responsive to short light waves of 420 nm, so they are called S cones ("S" for "short"); others respond maximally to waves of 530 nm (M cones, for "medium"); a third group responds maximally to light of longer wavelengths, at 560 nm (L, or "long" cones). With only one type of cone, color vision would not be possible, and a two-cone (dichromatic) system has limitations. Primates use a three-cone (trichromatic) system, resulting in full color vision. The color we perceive is a result of the ratio of activity of our three types of cones. The colors of the visual spectrum, running from long-wavelength light to short, are red (700 nm), orange (600 nm), yellow (565 nm), green (497 nm), blue (470 nm), indigo (450 nm), and violet (425 nm). Humans have very sensitive perception of color and can distinguish about 500 levels of brightness, 200 different hues, and 20 steps of saturation, or about 2 million distinct colors.

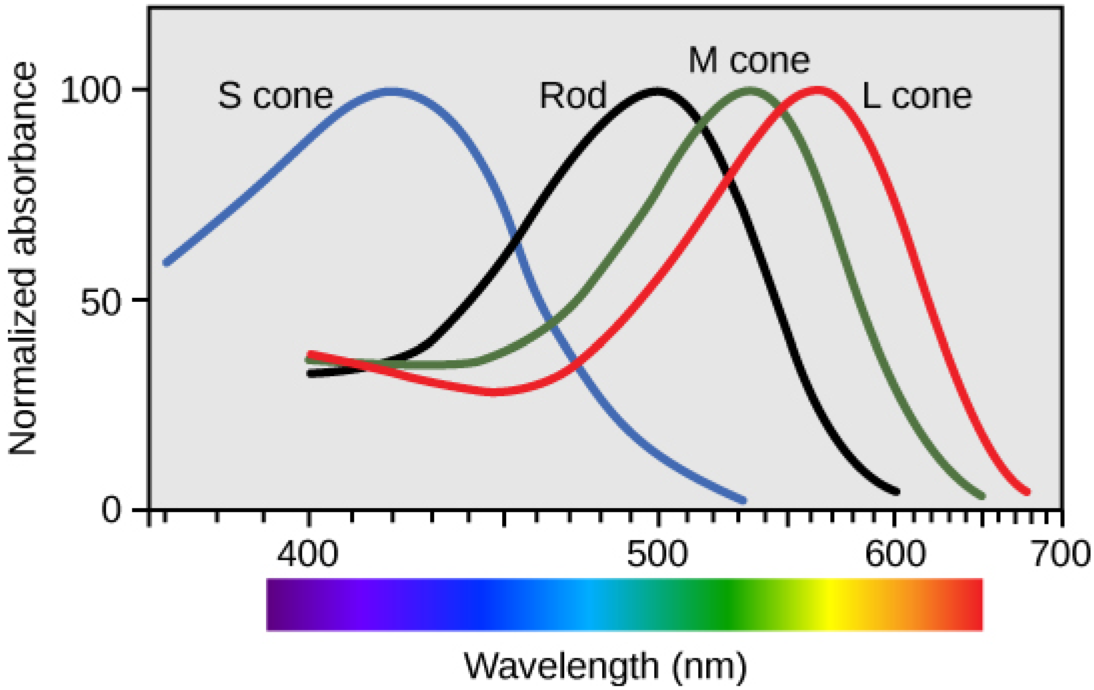


Figure 6. *Human rod cells and the different types of cone cells each have an optimal wavelength. However, there is considerable overlap in the wavelengths of light detected.*

Retinal Processing

Visual signals leave the cones and rods, travel to the bipolar cells, and then to ganglion cells. A large degree of processing of visual information occurs in the retina itself, before visual information is sent to the brain.

Photoreceptors in the retina continuously undergo **tonic activity**. That is, they are always slightly active even when not stimulated by light. In neurons that exhibit tonic activity, the absence of stimuli maintains a firing rate at a baseline; while some stimuli increase firing rate from the baseline, and other stimuli decrease firing rate. In the absence of light, the bipolar neurons that connect rods and cones to ganglion cells are continuously and actively inhibited by the rods and cones. Exposure of the retina to light hyperpolarizes the rods and cones and removes their inhibition of bipolar cells. The now active bipolar cells in turn stimulate the ganglion cells, which send action potentials along their axons (which leave the eye as the optic nerve). Thus, the visual system relies on change in retinal activity, rather than the absence or presence of activity, to encode visual signals for the brain. Sometimes horizontal cells carry signals from one rod or cone to other photoreceptors and to several bipolar cells. When a rod or cone stimulates a horizontal cell, the horizontal cell inhibits more distant photoreceptors and bipolar cells, creating lateral inhibition. This inhibition sharpens edges and enhances contrast in the images by making regions receiving light appear lighter and dark surroundings appear darker. Amacrine cells can distribute information from one bipolar cell to many ganglion cells.

You can demonstrate this using an easy demonstration to “trick” your retina and brain about the colors you are observing in your visual field. Look fixedly at Figure 7 for about 45 seconds. Then quickly shift your gaze to a sheet of blank white paper or a white wall. You should see an afterimage of the Norwegian flag in its correct colors. At this point, close your eyes for a moment, then reopen them, looking again at the white paper or wall; the afterimage of the flag should continue to appear as red, white, and blue. What causes this? According to an explanation called opponent process theory, as you gazed fixedly at the green, black, and yellow flag, your retinal ganglion cells that respond positively to green, black, and yellow increased their firing dramatically. When you shifted your gaze to the neutral white ground, these ganglion cells abruptly decreased their activity and the brain interpreted this abrupt downshift as if the ganglion cells were responding now to their "opponent" colors: red, white, and blue, respectively, in the visual field. Once the ganglion cells return to their baseline activity state, the false perception of color will disappear.

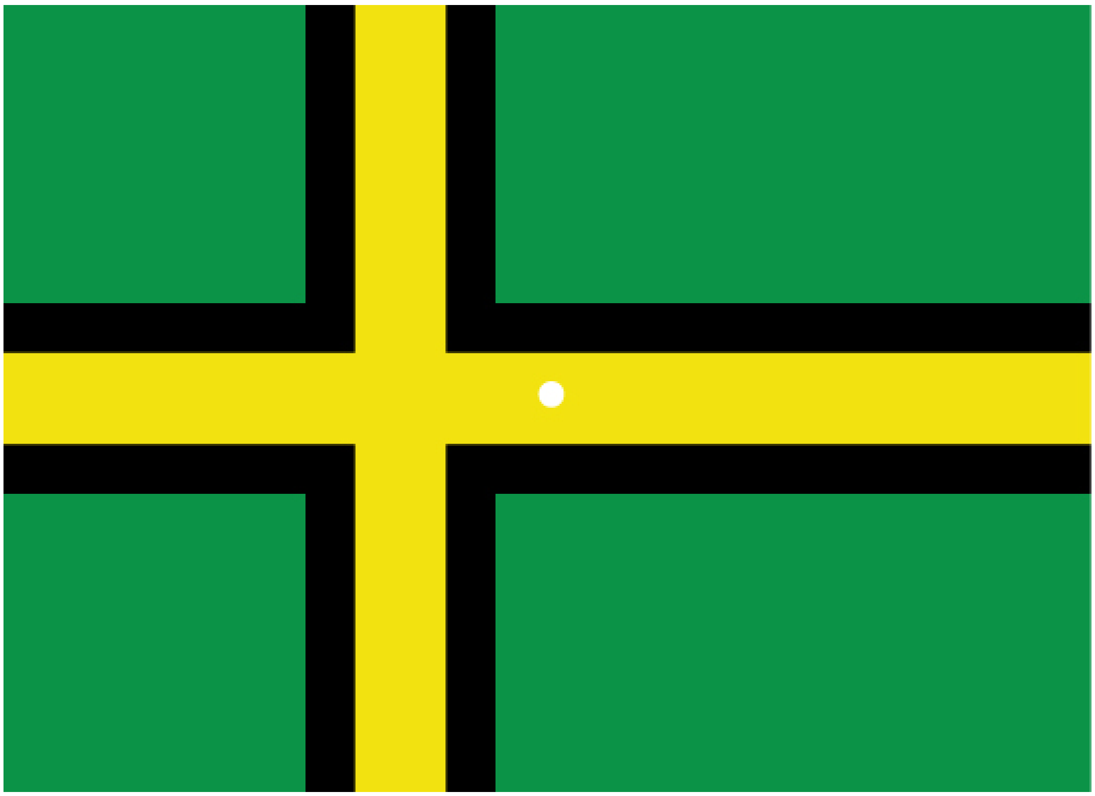


Figure 7. *View this flag to understand how retinal processing works. Stare at the center of the flag (indicated by the white dot) for 45 seconds, and then quickly look at a white background, noticing how colors appear.*

13. Introduction to Geometric Optics

The Ray Aspect of Light

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

In Unit I we talked about how light and electrons have both a wave and particle nature. In this particular unit, we will be using that fact extensively: sometimes thinking about waves and sometimes thinking about particles. the 'ray' aspect of light being described here is most easily pictured as the path of the individual particles, either electrons or photons, traveling in a straight line.

There are three ways in which light can travel from a source to another location. (See Figure 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).

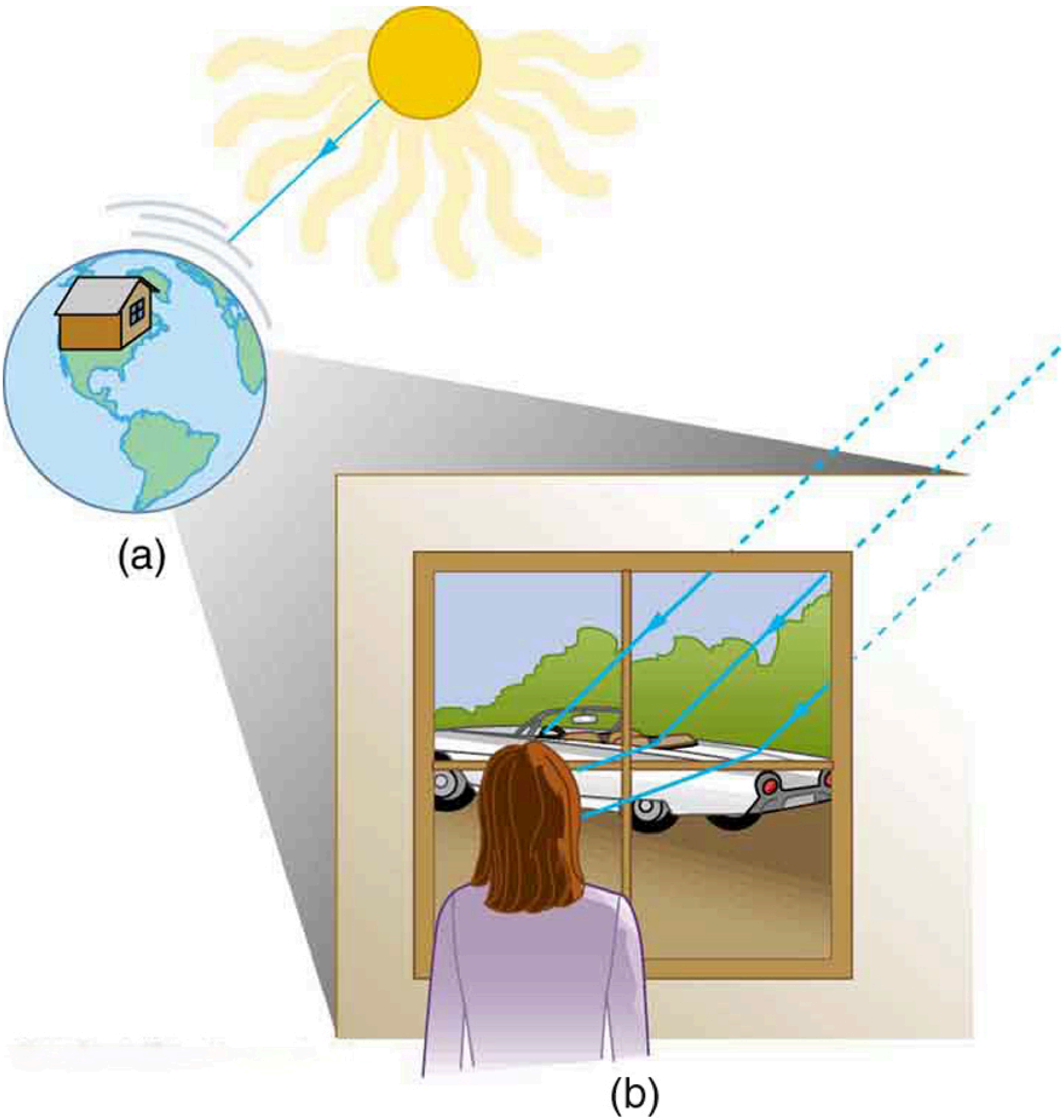


__Figure 1.__ Three methods for light to travel from a source to another location. (a) Light reaches the upper atmosphere of Earth traveling through empty space directly from the source. (b) Light can reach a person in one of two ways. It can travel through media like air and glass. It can also reflect 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 micrometer μm or 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.

Section 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.

Homework Problems

In geometric optics, we will use a lot of, well, geometry. Here are a few problems to help you review the needed geometric concepts. If you need some review, have a look at:

- OpenStax *Pre-algebra* – Section 9.3: Use Properties of Angles, Triangles, and the Pythagorean Theorem.
- OpenStax Pre-algebra – Section 9.4: Properties of Rectangles, Triangles, and Trapezoids.

Problem 3: In geometric optics, we do analyses using similar triangles. This problem is here to help you practice working on these again.

Problem 4: Look at this map and determine the angle.

Problem 5: For this set of intersecting lines, use the following information to find the missing values.

The Law of Reflection

Whenever we look into a **mirror**, or squint at sunlight glinting from a lake, we are seeing a reflection. When you look at the page of a printed book, you are also seeing light reflected from it. Large telescopes use reflection to form an image of stars and other astronomical objects.

The law of reflection is illustrated in Figure 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 1 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. Diffused light is what allows us to see a sheet of paper from any angle, as illustrated in Figure 1. 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 1. When the moon reflects from a lake, as shown in Figure 1, a combination of these effects takes place.

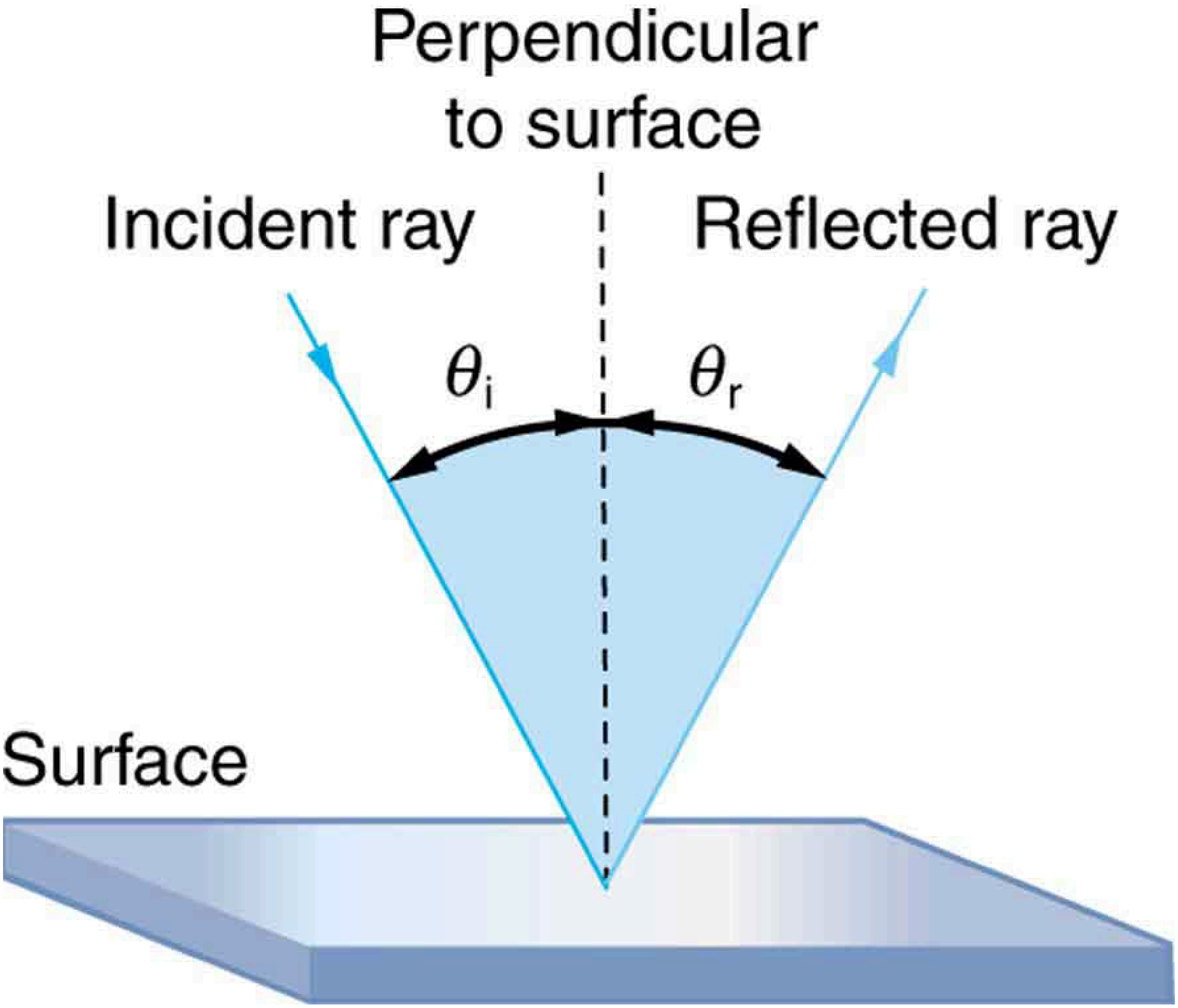


Figure 1. *The law of reflection states that the angle of reflection equals the angle of incidence* $-\theta_r = \theta_i$*. The angles are measured relative to the perpendicular to the surface at the point where the ray strikes the surface.*

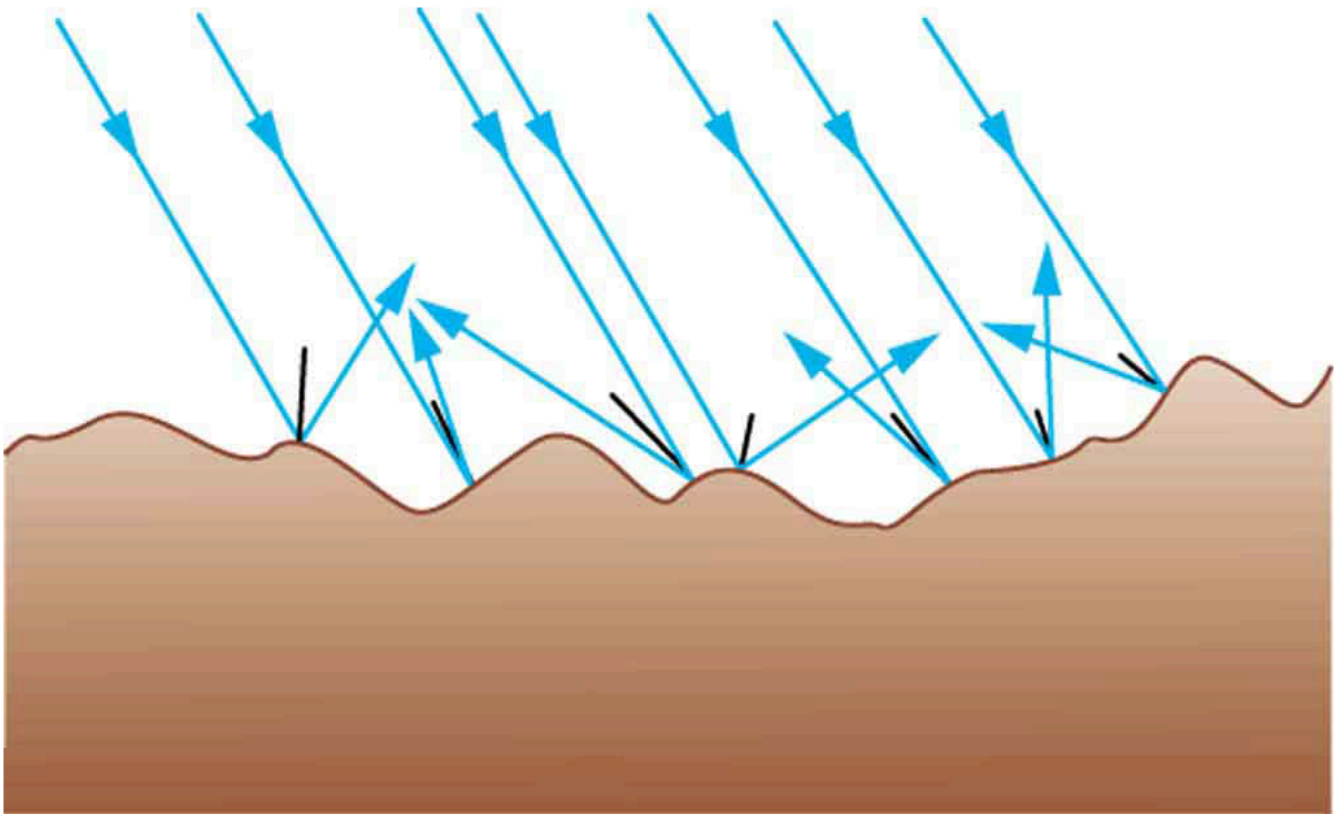


Figure 2. *Light is diffused when it reflects from a rough surface. Here many parallel rays are incident, but they are reflected at many different angles since the surface is rough.*

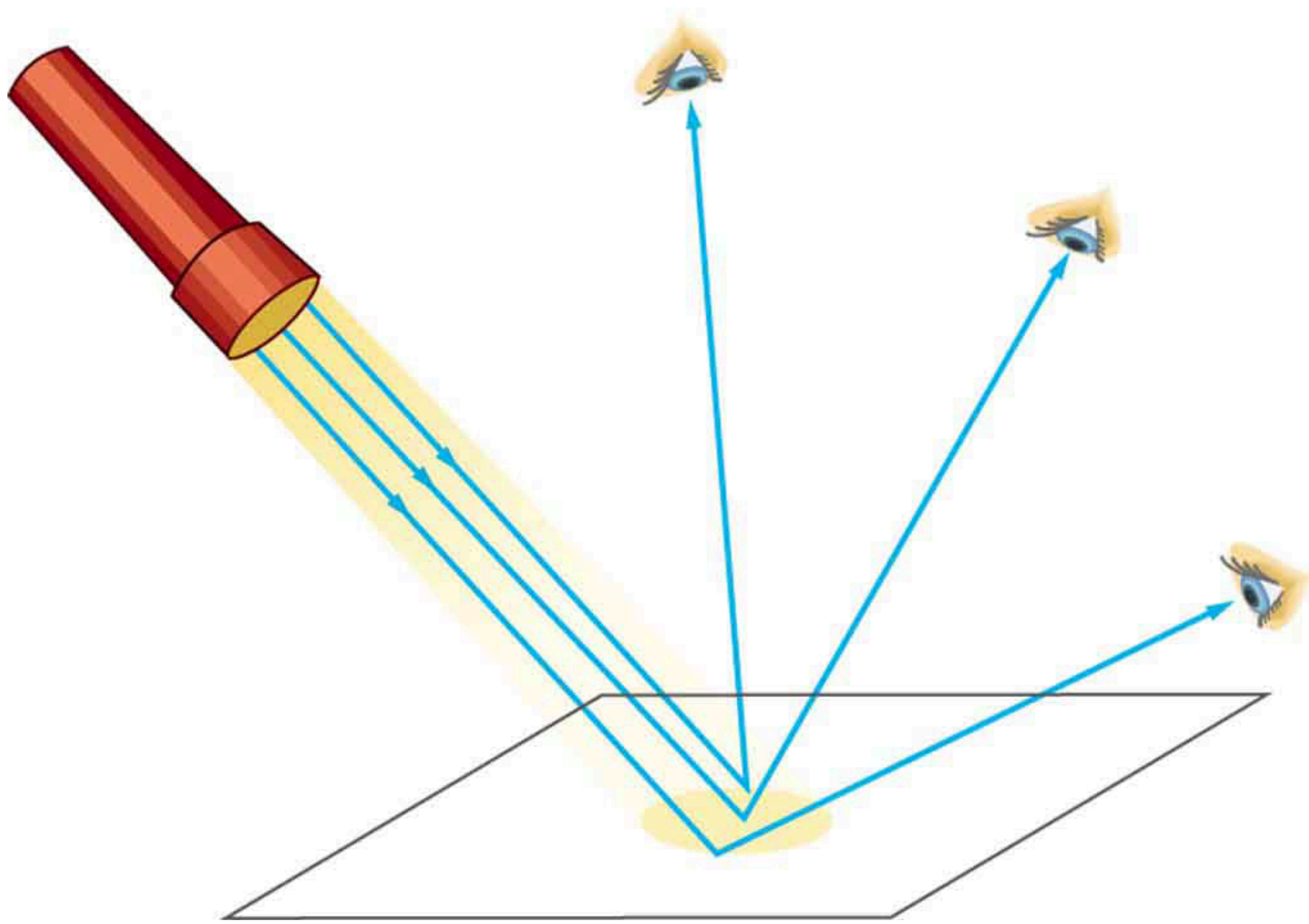


Figure 3. *When a sheet of paper is illuminated with many parallel incident rays, it can be seen at many different angles, because its surface is rough and diffuses the light.*

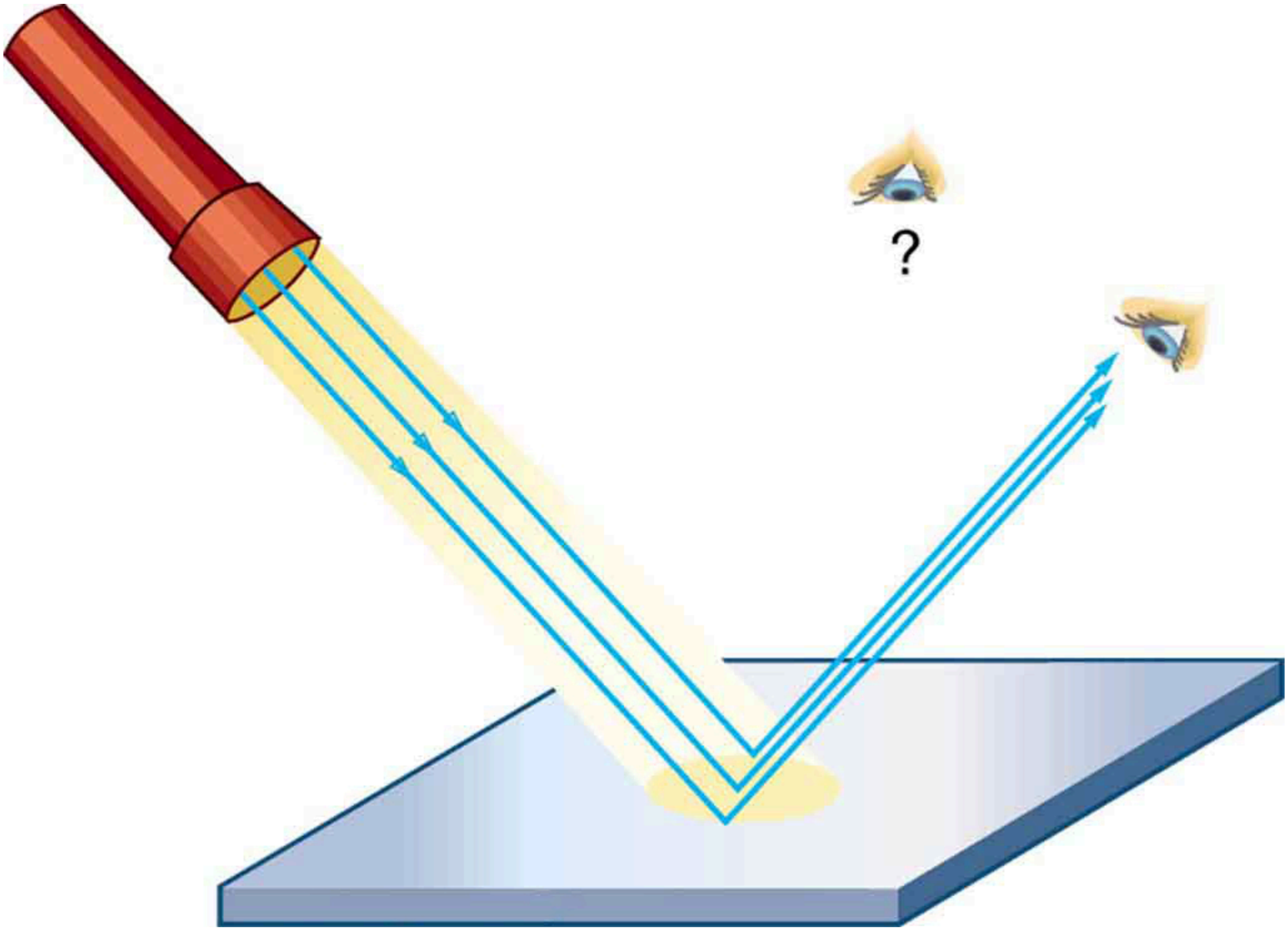


Figure 4. *A mirror illuminated by many parallel rays reflects them in only one direction, since its surface is very smooth. Only the observer at a particular angle will see the reflected light.*

Figure 5. *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.

THE LAW OF REFLECTION

The angle of reflection equals the angle of incidence.

Instructor's Note

This is important! You should ALWAYS measure your angles with respect to a line perpendicular to the surface. This line perpendicular to the surface is called a 'normal' line.

When we see ourselves in a mirror, it appears that our image is actually behind the mirror. This is illustrated in Figure 6. 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 Applications of Geometric Optics: Terminology of Images.

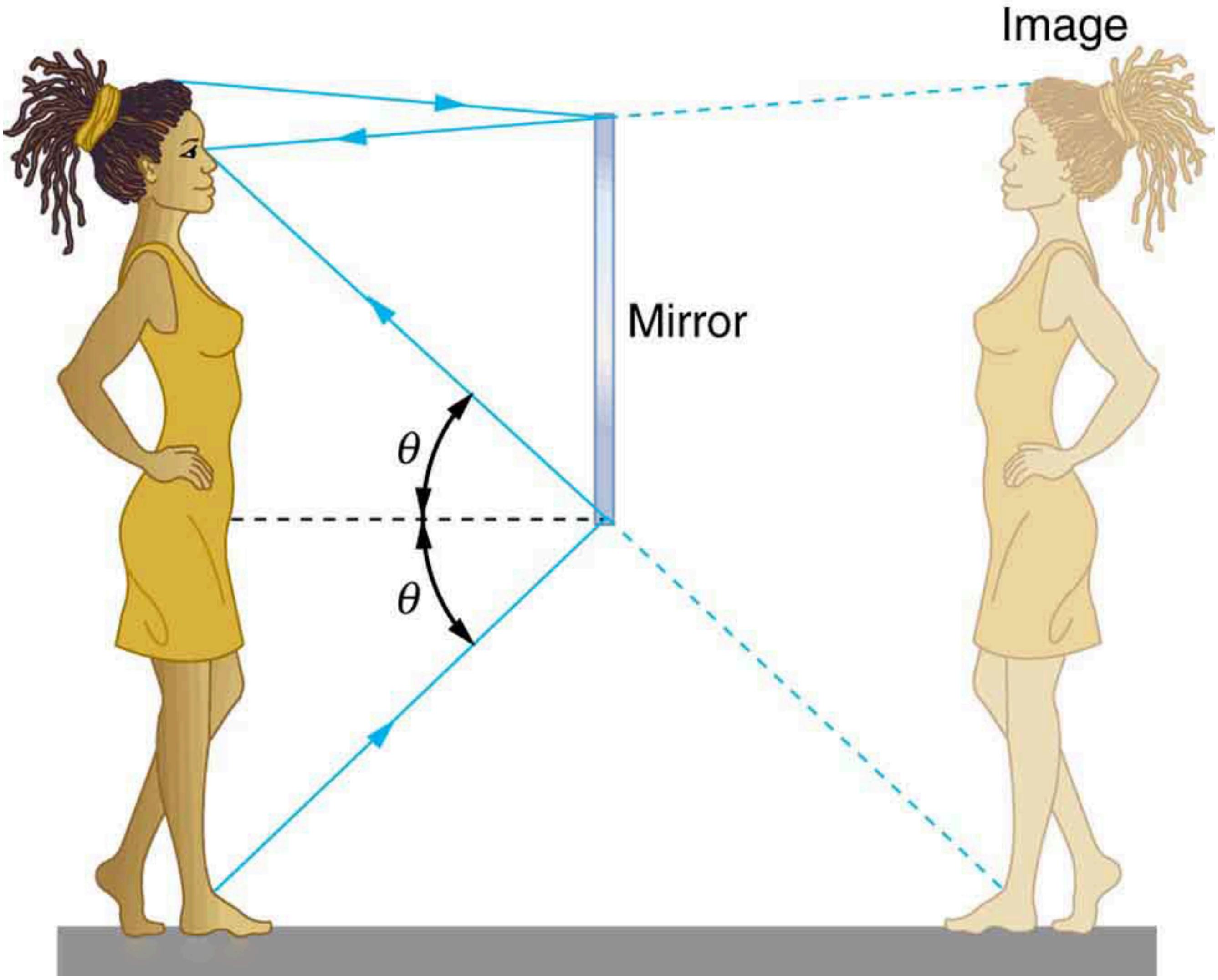


Figure 6. *Our image in a mirror is behind the mirror. The two rays shown are those that strike the mirror at just the correct angles to be reflected into the eyes of the person. The image appears to be in the direction the rays are coming from when they enter the eyes.*

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 3. Now shine the flashlight at a mirror at an angle. Do your observations confirm the predictions in Figure 3 and Figure 4? 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.)

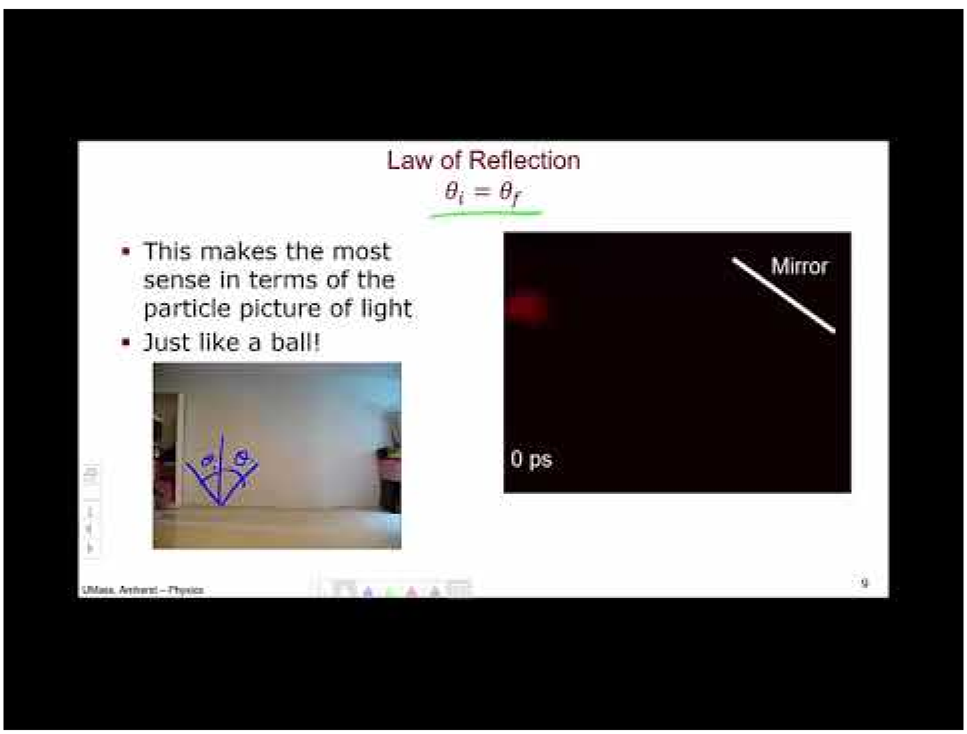
Section Summary

- 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

Homework Problem

Problem 6: Indicate where the outgoing ray from a mirror intersects the dotted line.

Law of Reflection in Terms of the Particle Picture of Light



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=497

We're going to begin with the basics of reflection. The Law of Reflection is simply, $\theta_i = \theta_f$, the incident angle equals the final angle. Reflection makes the most sense in terms of the particle picture of light. It's easiest to see if we just imagine light, a photon, as a ball. When a ball is bounced, the incident angle and the final angle are the same before and after it bounces. You can see that relative to the normal, which is how you should always measure your angles, the incident angle and the final angle are the same.

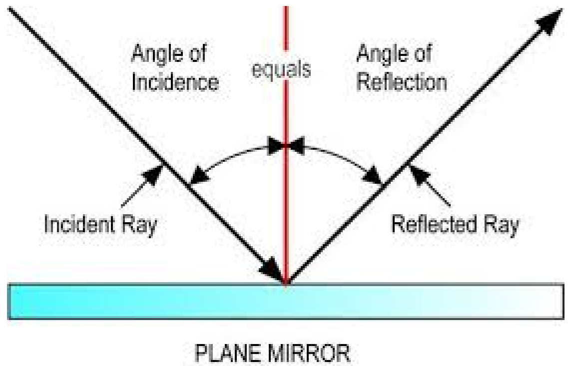


Figure 1. *The Law of Reflection (Credit: Wikipedia, C M Vidyashree)*

Speed of Light in Materials

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. (See Figure 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.

REFRACTION

The changing of a light ray's direction (loosely called bending) when it passes through variations in matter is called refraction.

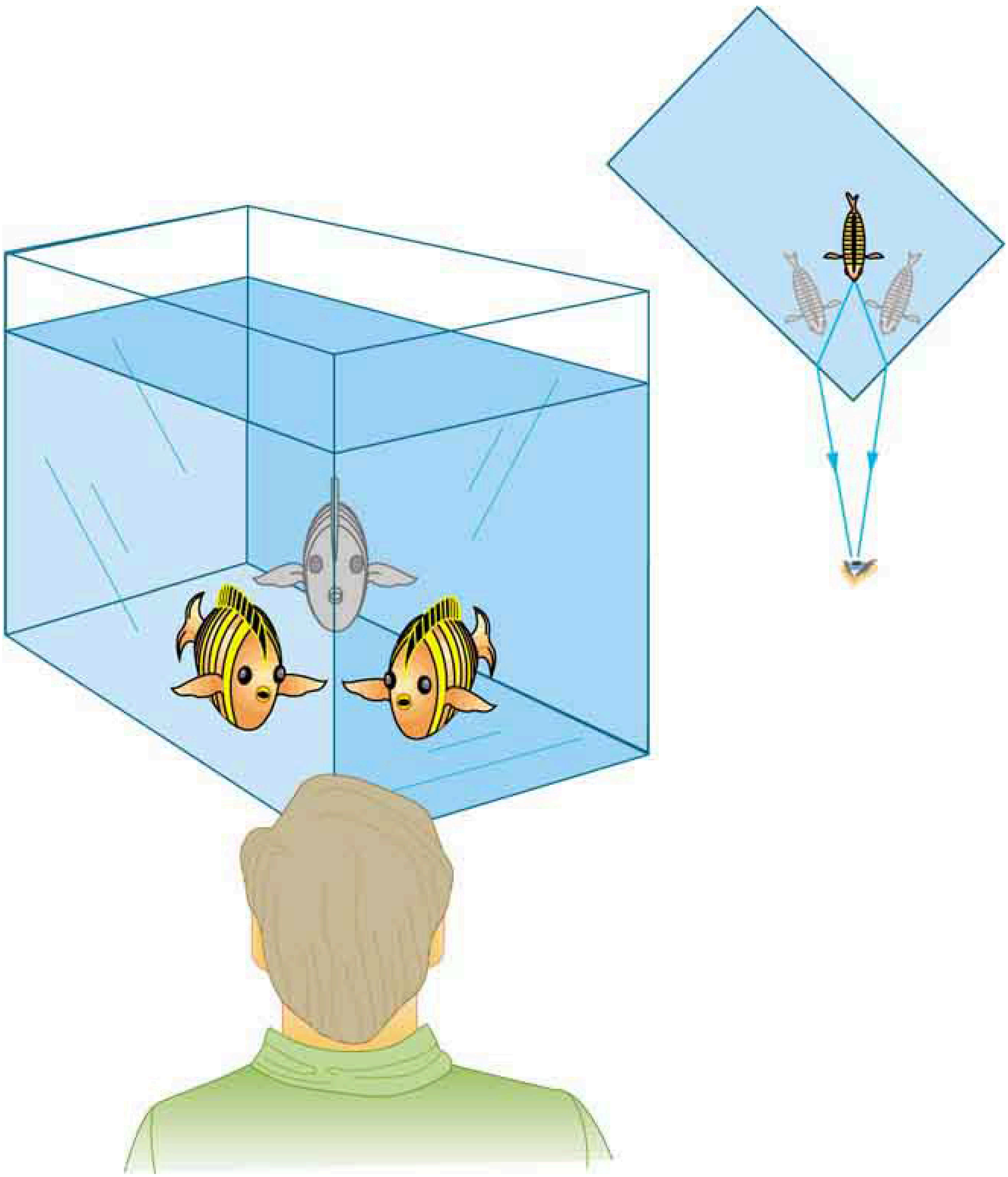


Figure 1. Looking at the fish tank as shown, we can see the same fish in two different locations, because light changes directions when it passes from water to air. In this case, the light can reach the observer by two different paths, and so the fish seems to be in two different places. This bending 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 than 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 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.

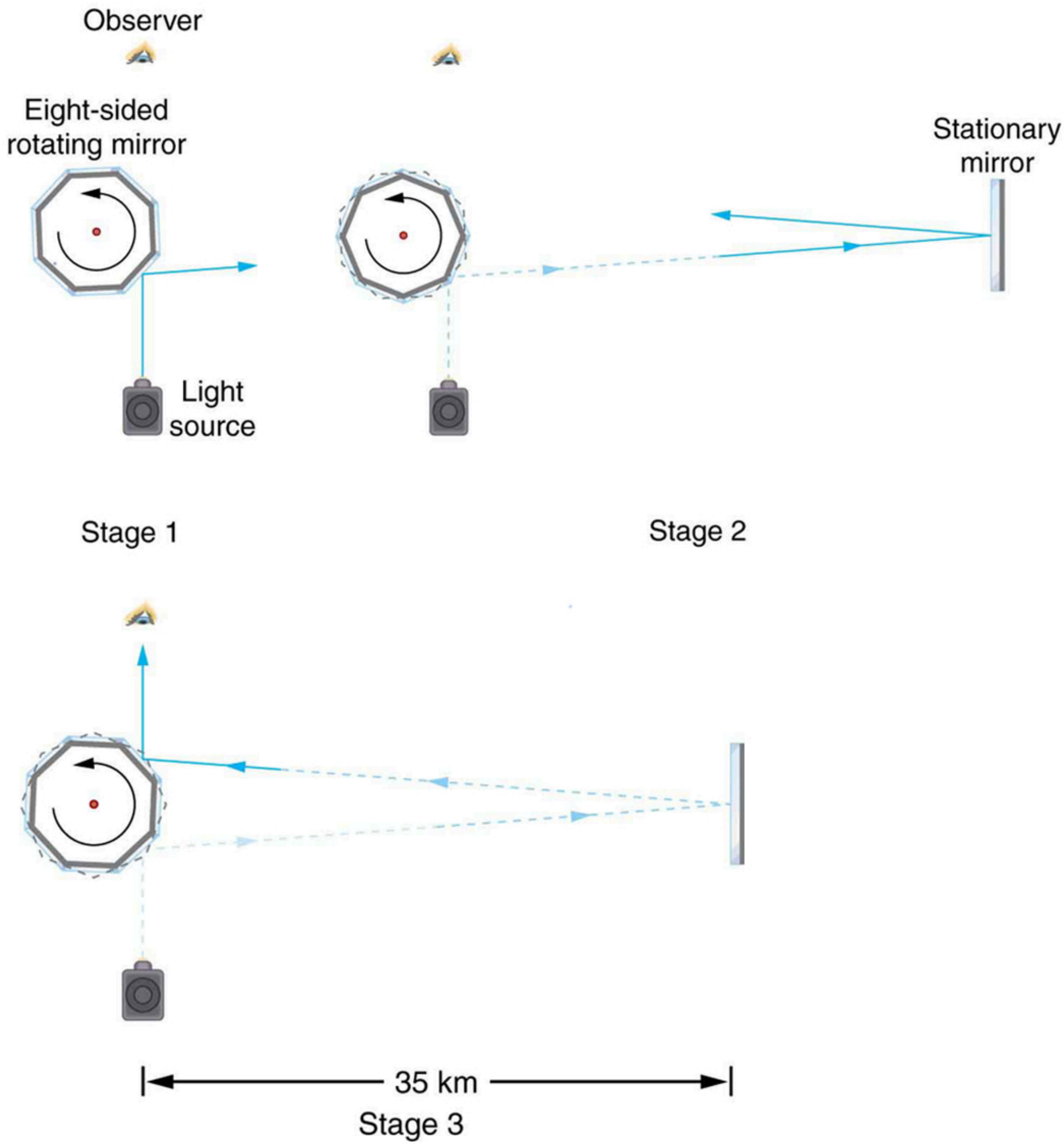


Figure 2. *A schematic of early apparatus used by Michelson and others to determine the speed of light. As the mirrors rotate, the reflected ray is only briefly directed at the stationary mirror. The returning ray will be reflected into the observer's eye only if the next mirror has rotated into the 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.

$$2.99792458 \times 10^8 m/s \approx 3.0 \times 10^8 m/s,$$

where the approximate value of $3.0 \times 10^8 m/s$ 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. We define the **index of refraction** n of a material to be

$$n = \frac{c}{v},$$

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.

VALUE OF THE SPEED OF LIGHT

$$2.99792458 \times 10^8 m/s \approx 3.0 \times 10^8 m/s$$

INDEX OF REFRACTION

$$n = \frac{c}{v}$$

That is, $n \geq 1$ Table 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.

Index of Refraction in Various Media

Index of refraction in Various Media	
Medium	n
Gases at $0°C, 1atm$	
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
Plexiglass	1.51
Quartz, crystalline	1.544
Quartz, fused	1.458
Sodium chloride	1.544
Zircon	1.923

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 $n = \frac{c}{v}$

Solution

The equation for index of refraction states that $n = \frac{c}{v}$, Rearranging this to determine v gives

$$v = \frac{c}{n}$$

The index of refraction for zircon is given as 1.923 in Table 1, and c is given in the equation for speed of light. Entering these values in the last expression gives

$$v = \frac{3.0 \times 10^8 m/s}{1.923}$$
$$= 1.56 \times 10^8 m/s$$

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 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.

Section Summary

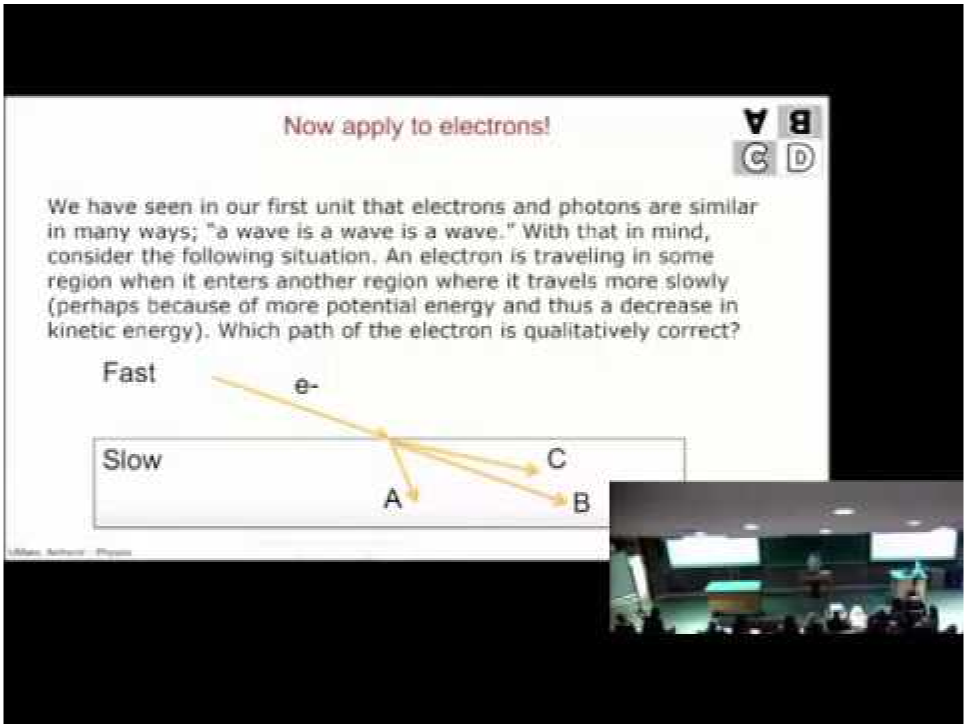
- The changing of a light ray's direction when it passes through variations in matter is called refraction.
- The speed of light in vacuum $2.99792458 \times 10^8 m/s \approx 3.0 \times 10^8 m/s$
- Index of refraction $n = \frac{c}{v}$, where v the speed of light in the material, c is the speed of light in vacuum, and n is the index of refraction.

Homework Problems

Problem 7: What is the speed of light in water? In glycerine? The indices of refraction for water is 1.333 and for glycerine is 1.473.

Problem 8: Calculate the index of refraction for a medium in which the speed of light is 1.416×10^8 m/s.

Why Light Bends



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=497*](http://openbooks.library.umass.edu/toggerson-132/?p=497)

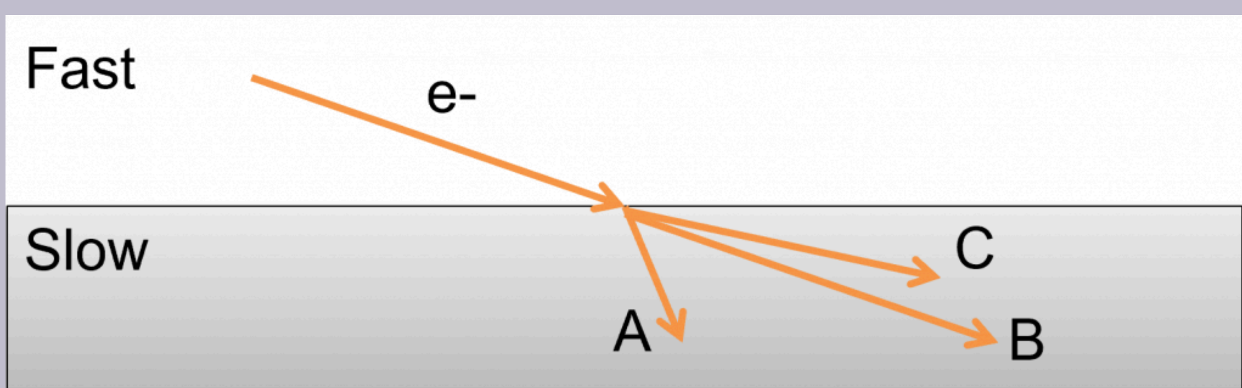
So let's talk about why light bends. So this is what makes starting with wave particle duality so useful: to understand why light bends, a wave picture is very useful, to understand which properties of light change it is useful to think about the particle nature.

Let's watch, now we're thinking about light as a wave. So my wave is coming in from the upper left and you can see that this part of the wave hits the interface first and slows down, while this part of the wave, keeps going. That's what makes it bend, different parts of the wave hit it at different times. The right edge of this wave hits the interface first and slows down, while the left edge keeps traveling at the faster speed and as a consequence the light as a whole bends. You can see in this video they even got the fact that the wavelength shrinks, they even got that right.

A light ray bends because one side hits the interface first and slows while the rest continues at the faster speed. The wavelength in the slower material is smaller.

A wave is a wave is a wave

We have seen in our first unit that electrons and photons are similar in many ways; "a wave is a wave is a wave." With that in mind, consider the following situation. An electron is traveling in some region when it enters another region where it travels more slowly (perhaps because of more potential energy and thus a decrease in kinetic energy). Which path of the electron is qualitatively correct?



Solution:

A is correct.

Just like a light wave, if I can get the electron to slow down I can get it to bend towards the normal just like a light wave.

I can make it bend, I can build a lens.

If I can build a lens, I can build a microscope.

This is why electron microscopes work, because I can bend electrons by speeding them up or slowing them down.

Wave hitting interface perpendicular

If the wave comes in straight perpendicular, does it bend?

Solution:

Remember that the bending came from the fact that different parts of the wave slowed down at different times, if we come straight at it, that doesn't happen and so they all slow down together, and you don't get a bend: the wave goes straight.

Homework Problems

Problem 9: Consider two materials. When light passes through the space between the two materials at $0° < \theta < 90°$, there is no change in the direction of the propagation of the light. What can you infer about the two materials?

Digging More into Wave-Particle Duality and Refraction[1]

Now, let's think about some of the other properties of the light wave, beyond speed, and how they might change as we go from one material to another. Starting with wavelength.

1. A note to more advanced readers - the following derivation of why the wavelength changes and not the frequency is not 100% correct, there are more complex effects at play due to Einstein's Theories of Relativity. However, the essence of the argument depending on energy conservation is correct and so is the result.

Wavelength

We know from Unit I that light is made of photons and that these photons have energy

$$E_\gamma = \frac{hc}{\lambda}.$$

The c in this equation, however, is trying to tell us something. The value $c = 3 \times 10^8\ \mathrm{m}s$ is the speed of light *in vacuum*. We know now that, in a material, light will slow to some $vlt; c$, and our resulting expression will now be

$$E_\gamma^{\text{material}} = \frac{hv}{\lambda}$$

where v is *the speed of the light in the material*. Because of conservation of energy, the energy of the photon cannot change. Thus, according to our equation, if the speed goes down, *the wavelength must also decrease by the same factor.*

Example: Reduction of wavelength in materials

Say we have a light source in a vacuum that emits light with a wavelength of $\lambda_0 = 6\ \mathrm{nm} = 6 \times 10^{-9}\ \mathrm{m}$. The light then enters a material where the speed of light is only $2 \times 10^8\ \mathrm{m/s}$. What is the wavelength λ_n in this new material?

Solution:

Given that

$E_\gamma^{\text{vacuum}} = \frac{hc}{\lambda_0}$

and

$E_\gamma^{\text{material}} = \frac{hv}{\lambda_n}$

and given he energy of the photon cannot change due to conservation of energy:

$E_\gamma^{\text{vacuum}} = E_\gamma^{\text{material}}$

we can set the two expressions equal to another:

$$\frac{hc}{\lambda_0} = \frac{hv}{\lambda_n}$$
$$\frac{c}{\lambda_0} = \frac{v}{\lambda_n}$$
$$\frac{c}{v} = \frac{\lambda_0}{\lambda_n}.$$

We recognize the quantity c/v as the index of refraction $n = c/v$, which in this case is

$n = \frac{c}{v} = \frac{3\times10^8\ \mathrm{m/s}}{2\times10^8\ \mathrm{m/s}} = 1.5$

Thus, we have

$$n = \frac{\lambda_0}{\lambda_n}$$
$$\lambda_n = \frac{\lambda_0}{n}$$

Substituting in our values, we have:

$$\lambda_n = \frac{6\ \mathrm{nm}}{1.5} = 4\ \mathrm{nm}.$$

Discussion:

The speed of light went down by a factor of 2/3 and so did the wavelength!

Frequency

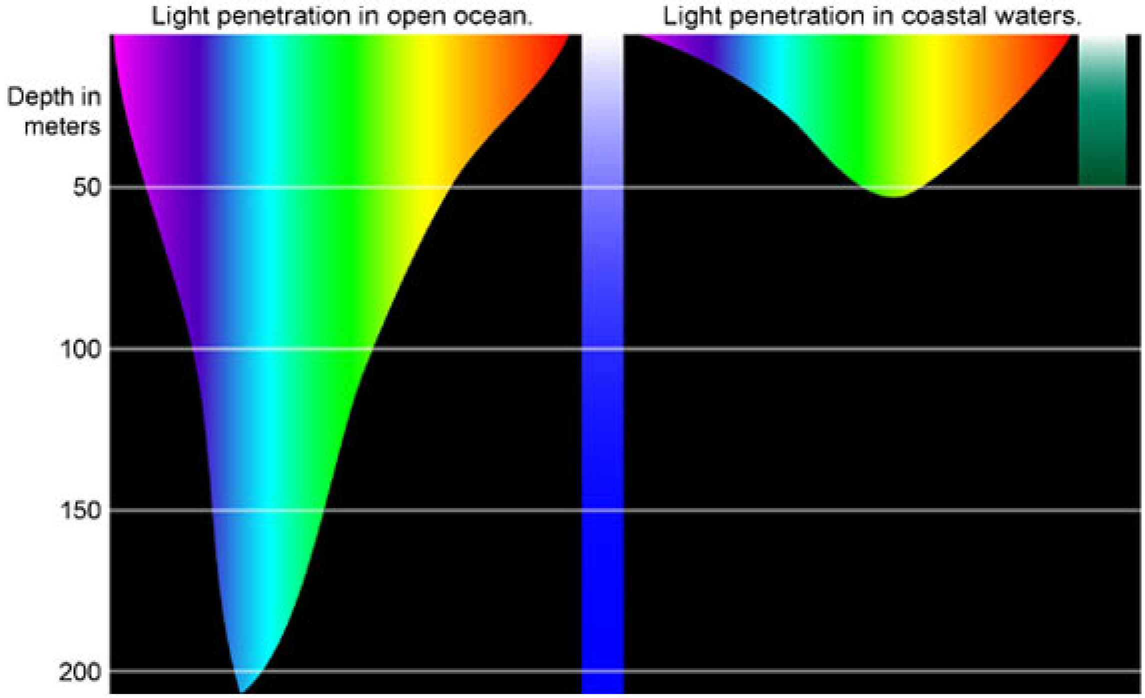
What happens to the frequency of a light wave in matter? Well the fundamental relationship for all waves $v = \lambda\nu$ must still be obeyed. As a light wave goes from a vacuum into a material, the speed changes $c \to v = c/n$, and so does the wavelength $\lambda_0 \to \lambda_n = \lambda_0/n$. What happens to the frequency?

$$v = \lambda_n \nu$$
$$\Downarrow$$
$$\frac{c}{n} = \frac{\lambda_0}{n}\nu$$

cancel the n and we are left with the true statement $c = \lambda_0\nu$. What to make of this? It means that *the frequency of the light wave does not change!*

Amplitude / Number of Photons

As you know from Unit I, the amplitude of the light wave and the number of photons are both related to the light's **intensity**. Thus, these quantities are more about how much light is absorbed by the material than its index of refraction. Glass and air absorb very little light in the visible range, meaning that the amplitude and number of photons is not very much reduced in these materials. Water on the other hand, is very effective at absorbing visible light photons. As shown in the Figure, at a depth of 200m, almost no visible light penetrates resulting in creatures with special adaptations to live in complete darkness.



Light penetration as a function of color and depth: NOAA – National Oceanic and Atmospheric Administration [Public domain]

Homework Problems

Problem 10: Which of the properties of a light ray change as it goes from glass to vacuum?

Problem 11: What are the wavelengths of visible light in crown glass?

The Law of Refraction

Figure 1 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 1, 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 1 (a), the direction of the ray moves closer to the perpendicular when it slows down. Conversely, as shown in Figure 1 (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.

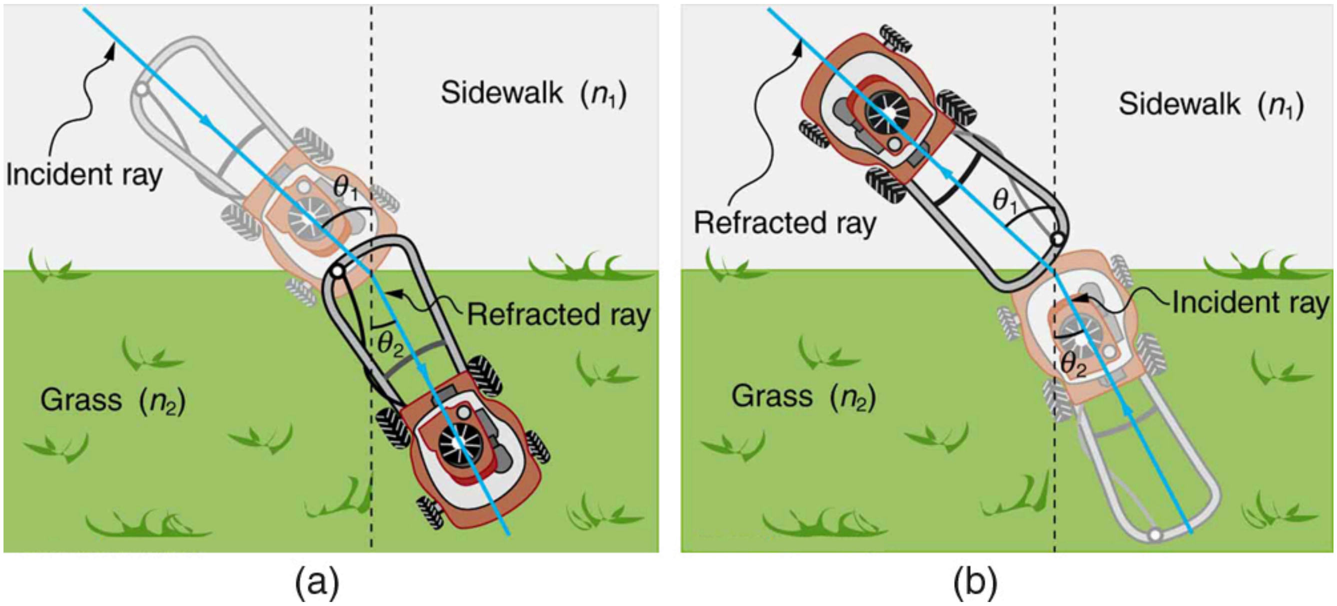


Figure 1. *The change in direction of a light ray depends on how the speed of light changes when it crosses from one medium to another. The speed of light is greater in medium 1 than in medium 2 in the situations shown here. (a) A ray of light moves closer to the perpendicular when it slows down. This 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 reflection, or "Snell's Law," which is stated in equation form

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Here n_1 and n_2 are the indices of refraction for medium 1 and 2, and θ_1 and θ_2 are the angles between the rays and the perpendicular in medium 1 and 2, as shown in Figure 1. 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. Below is a simulation where you can shine light through different materials and see how it bends. I encourage you to play with it to get a feel for refraction. You can even add a protractor and see that the simulation obeys Snell's Law.

An interactive or media element has been excluded from this version of the text. You can view it

online here:
http://openbooks.library.umass.edu/toggerson-132/?p=497

THE LAW OF REFRACTION

$n_1 \sin\theta_1 = n_2 \sin\theta_2$

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.

Determine the Index of Refraction form Refraction Data

Find the index of refraction for medium 2 in Figure 1 (a), assuming medium 1 is air and given the incident angle is 30° and the angle of refraction is 22°

Strategy

The index of reflection for air is taken to be 1 in most cases (and up to four significant figures, it is 1.00). Thus $n_1 = 1.00$. Form the given information, $\theta_1 = 30^\circ$ and $\theta_2 = 22^\circ$. With this information, the only unknown in Snell's law is n_2, so it can be used to find this unknown.

Solution

Snell's law is

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Rearranging to isolate n_2 gives

$$n_2 = n_1 \frac{\sin\theta_1}{\sin\theta_2}$$

Entering known values,

$$n_2 = 1.00\frac{\sin 30.0^\circ}{\sin 22.0^\circ} = \frac{0.500}{0.375}$$
$$n_2 = 1.33$$

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.

A Larger Change in Direction

Find the index of refraction for medium 2 in Figure 1 (a), assuming medium 1 is air and given the incident angle is 30° and the angle of refraction is 22°

Strategy

Again the index of refraction for air is taken to be $n_1 = 1.00$ and we are given $\theta_1 = 30^\circ$. We can look up the index of refraction for a diamond in Table 1 (Speed of Light in Materials), finding $n_2 = 2.419$. The only unknown in Snell's law is θ_2, which we wish to determine.

Solution

Solving Snell's law for $sin\theta_2$ yields

$$sin\theta_2 = \ sin\theta_1 \frac{n_1}{n_2}$$

Entering known values,

$$sin\theta_2 = \ sin 30^\circ \frac{1.00}{2.419} = (0.413)(0.500) = 0.207$$

And angle is thus

$$\theta_2 = \sin^{-1} 0.207 = 11.9^\circ$$

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.

For the same 30° angle of incidence, the angle of refraction in diamond is significantly smaller than in water (11.9° rather than 22°—see the preceding example). This means there is a larger change in direction in diamond. The cause of a large change in direction is a large change in the index of refraction (or speed). In general, the larger the change in speed, the greater the effect on the direction of the ray.

Section Summary

- Snell's law, the law of refraction, is stated in equation form as $n_1\theta_1 = n_2\theta_2$

Homework Problems

Problem 12: Suppose you have an unknown clear substance immersed in water, and you wish to identify it by finding its index of refraction. You arrange to have a beam of light enter it at an angle of 48.6°, and you observe the angle of refraction to be 32.4°. What is the index of refraction of the substance? Water has an index of refraction equal to 1.333.

Problem 13: A beam of white light goes from air into water at an incident angle of 83.0˚. At what angles are the red 660 nm and violet 410 nm parts of the light refracted? Red light in water has an index of refraction equal to 1.331 and that of violet light is 1.342.1.342">

Problem 14: Given that the angle between the ray in the water and the perpendicular to the water is 28.3˚, and using information in the figure above, find the height of the instructor's head above the water. Water has an index of refraction equal to 1.333.

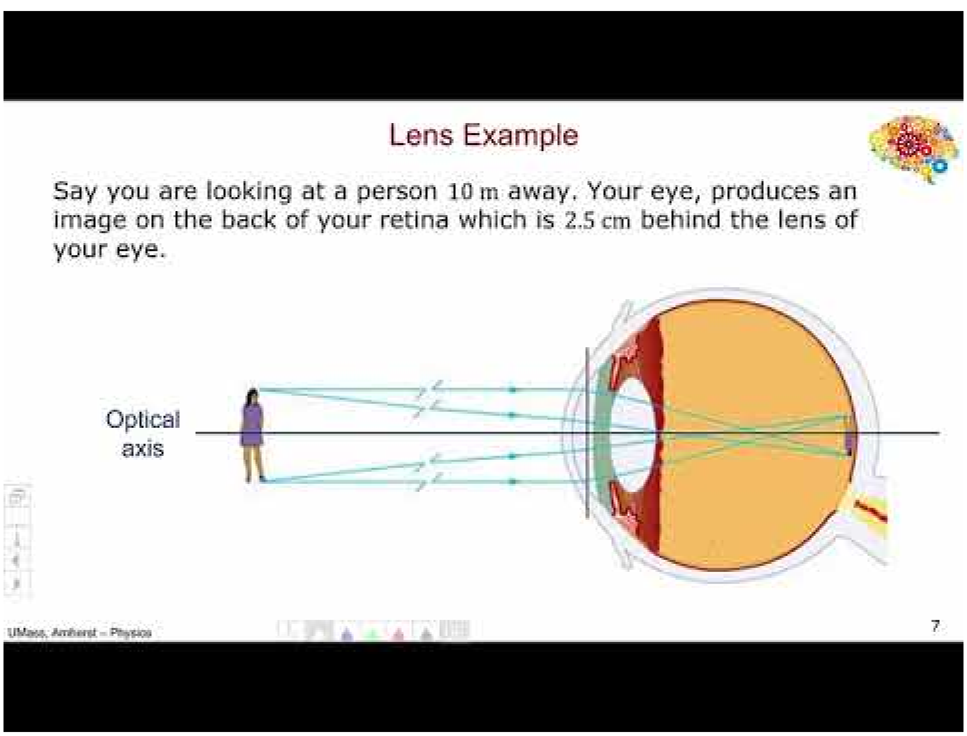
14. Producing Images with Geometric Optics

Terminology of Images and Optical Elements

Instructor's Note

By the end of this section you should:

- Be able to define several terms associated with images and optics.
- Know and apply the sign conventions associated with objects and images in optics.



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=499

An **optical element** is any lens or mirror. A few examples include the lens in your eye that you're using to read this or the aforementioned shaving or makeup mirror that many people have to help them get ready in the morning, it makes your face look a little bigger than it is. The lens or the mirror are examples of optical elements. Anything that's a lens or a mirror is an optical element.

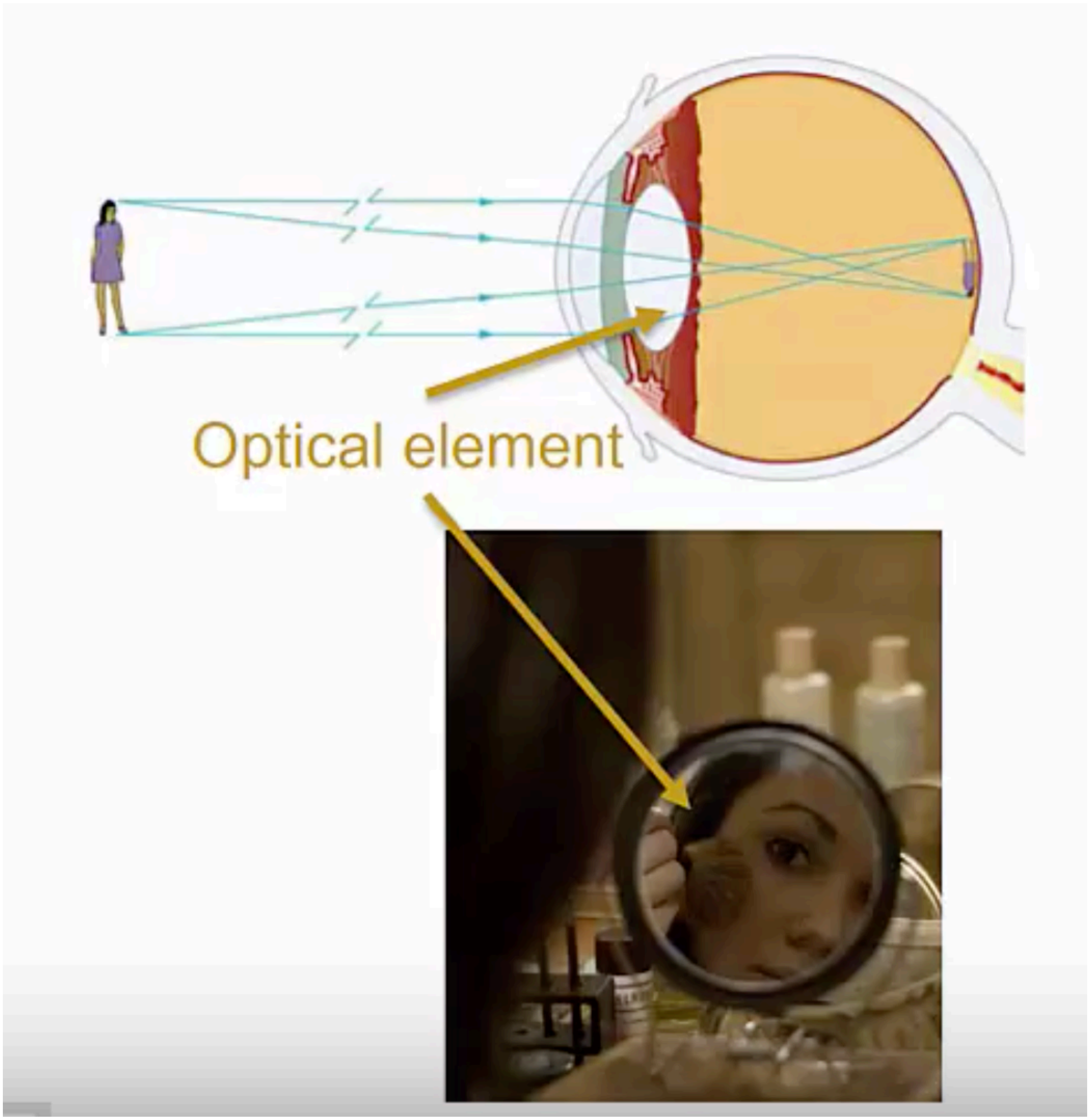


Figure 1.

Optical elements or combinations of them can be used to make **images**. An image is the apparent reproduction of an object formed by an optical element, or collection of them, through the reflection and/or refraction of light. In these two examples, the person is our object and the images are the image on the back of your retina formed by your eyeball, or the image of your reflection in the mirror.

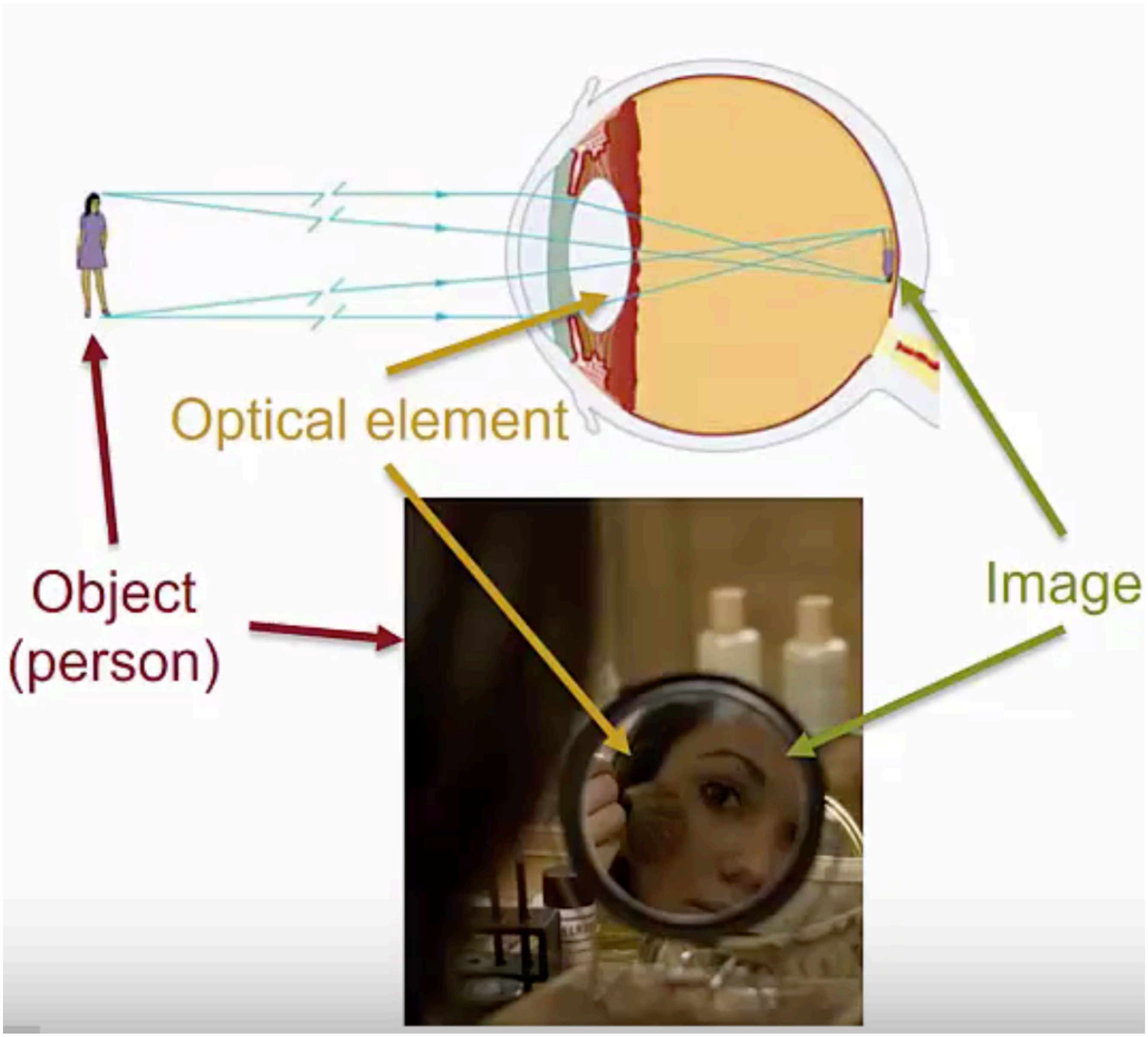


Figure 2.

Images can either be **erect**, with the same orientation as the object, or **inverted**, upside down with respect to the object. In the makeup mirror your face is right-side up, and so the image is erect. The images on your retina are actually upside down your brain corrects them to put them right side up, and this is an example of an inverted image.

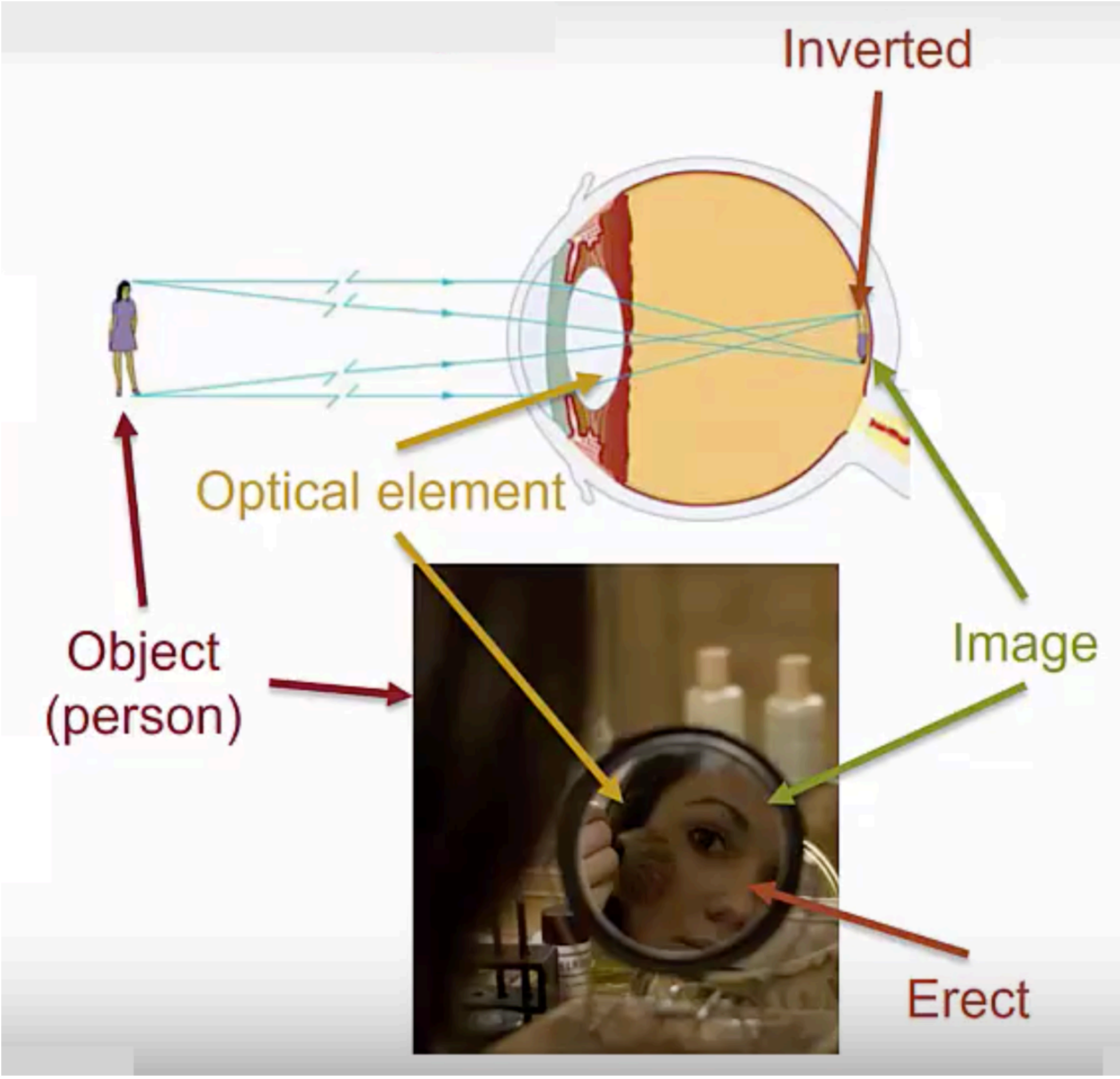


Figure 3.

Before we move on to talk about image and object distances, we need to introduce two more terms: the **optical axis** and the **vertex**.

The **optical axis** is an imaginary line that passes through the optical element in a way that's perpendicular to it. Below, we have a converging lens at the top, with a diverging lens below that, with a converging mirror below that, and a diverging mirror at the bottom. This dashed line that always meets the lens or mirror perpendicular is what we call the optical axis.

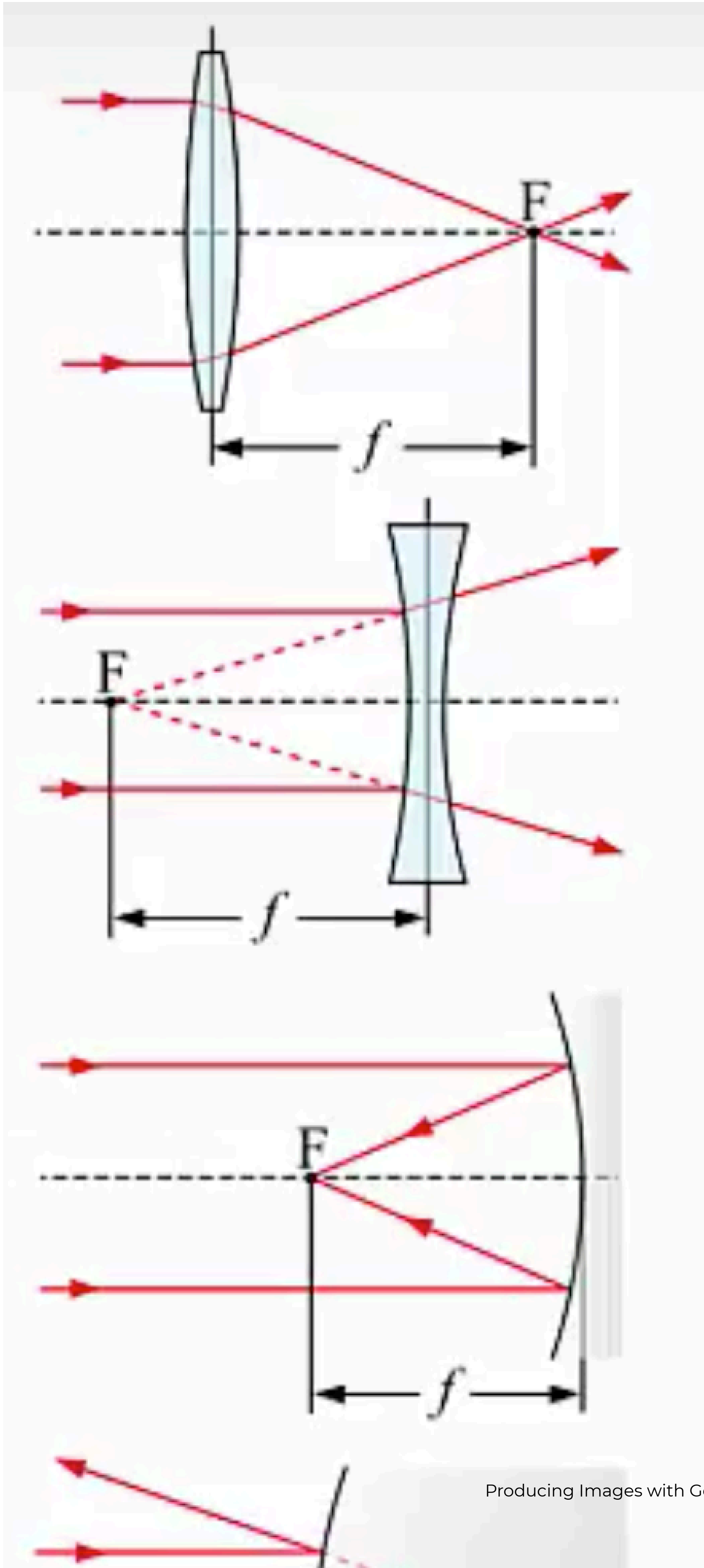

F
f
F
f
F
f

Figure 4.

Looking at a few examples, we can see that the optical axis meets the lens perpendicular here in the middle, same for the diverging lens. Moving on to the mirrors we see the optical axis meets the two mirrors in a way that's perpendicular. The point where the optical axis meets the optical element is called the vertex.

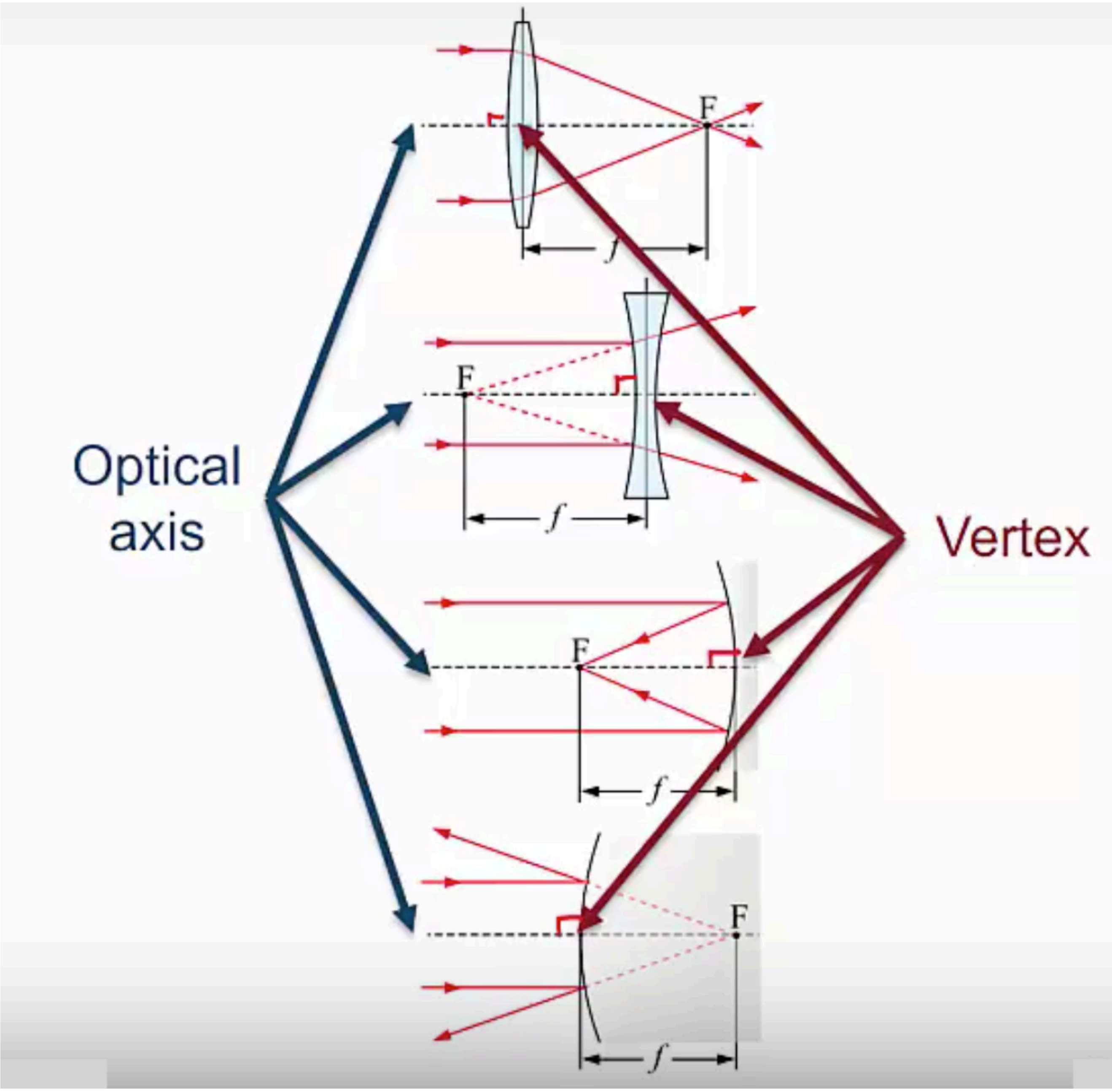


Figure 5.

The reason we needed these terms is because the image distance i and object distance o are measured along the optical axis from the vertex and these distances have signs that can be positive or negative and the sides are relative to the path of the light if the object is on the same side as the incoming light, then the object distance will be positive, otherwise the object distance is negative. If the image is on the same side as the outgoing light, Then the image distance is positive otherwise the image distance is negative. Note that for our lens the incoming and outgoing sides are different, light passes through a lens, so the light comes in one side and goes out the other. For a mirror on the other hand, the incoming and outgoing sides are the same, light bounces off of a mirror. These may seem like a rather convoluted set of rules, but it turns out that this is actually the simplest set of rules that works for all lenses and mirrors, so any other set of rules you might try to come up with will necessarily be more complicated.

Let's employ these sign conventions in the terms of two examples, one with a lens and one with a mirror.

To begin with a lens, let's say you're looking at a person about 10 meters away. Your eye produces an image on the back of your retina, which is about an inch behind the lens of your eye or 2.5 centimeters, what are the image and object distances including the signs?

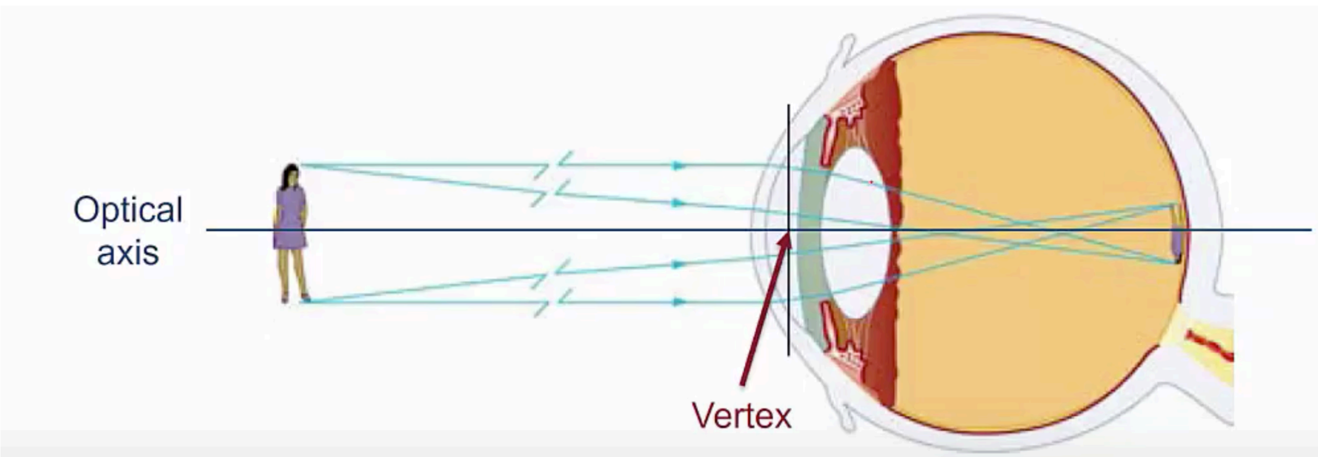


Figure 6.

First, we define our optical axis, passing through the lens you'll notice that the light is bending here, so that's what we're actually defining as our lens. The point where the optical axis meets the lens is the vertex, this is the point from which we are measuring our image and object distances. Now the person of course doesn't shine off by their own light, but from light bouncing off of them, which means the light is coming from the left.

Since the person is on the same side as where the light is coming from the object distance is going to be positive. Now we also know the person is 10 meters away, so we would say the object distance is 10 meters. For the image on the back of your retina, the image is 0.25 centimeters past the vertex on the side where the light is going out, which means the image distance is also positive leading to an image distance of 2.5 centimeters.

Now let's look at an example with a mirror. A can of shaving cream sits 30 centimeters in front of a flat plane mirror, like you have in your bathroom. You see the image of the shaving can apparently 30 centimeters behind the mirror. What are the image distances, i, and the object distance, o?

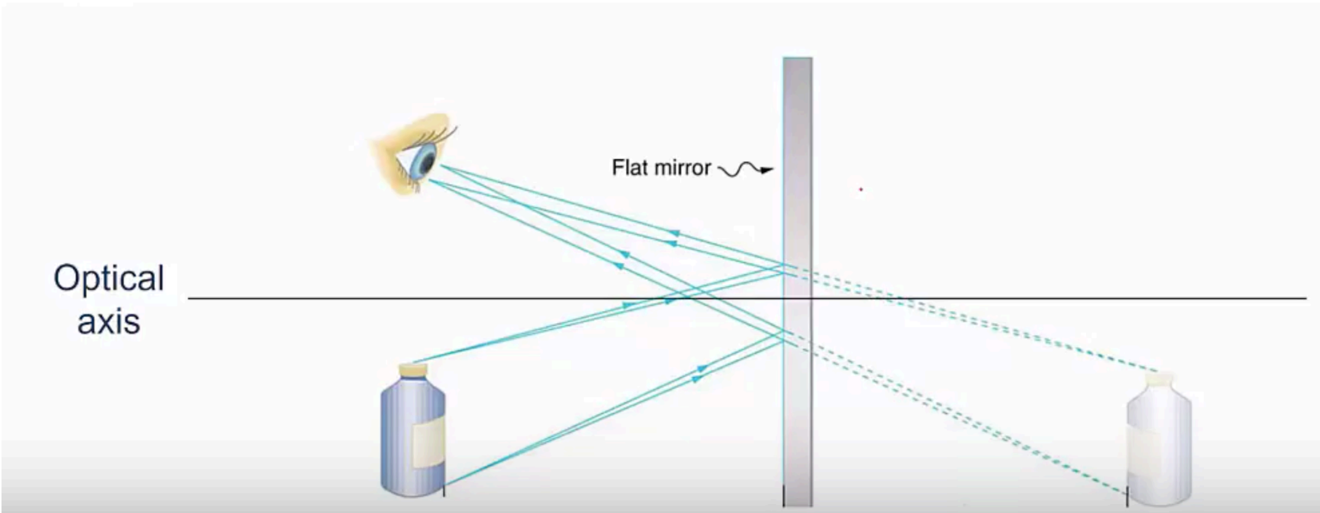


Figure 7.

Once again, we define our optical axis so that it meets the mirror perpendicular. The can does not shine by its own light but from light bouncing off of it, so the light is coming from the left, which means the object is on the same side as the incoming light, which means the object distance is positive and 30 centimeters, so o is 30 centimeters.

The image on the other hand is not on the side of the outgoing light because the light bounces off the mirror and back the way it came from. The image is on the side opposite the outgoing and so the image distance, i, is actually – 30 centimeters.

One more practice – Looking into a Spoon

Go get a metal spoon and look into each side. Pictures from Dr. Toggerson are below, but this really works better if you do it yourself. If you look in the back, as in Figure 8a, your image seems to be behind the spoon. In contrast, if you look carefully in the inside, as in Figure 8b, your image seems to hover in front of the spoon (take a piece of paper or a toothpick and try to poke it and you will see what I mean).

What are the sign conventions for o and i in these cases?

Figure 8: Dr. Toggerson’s reflections in a spoon.

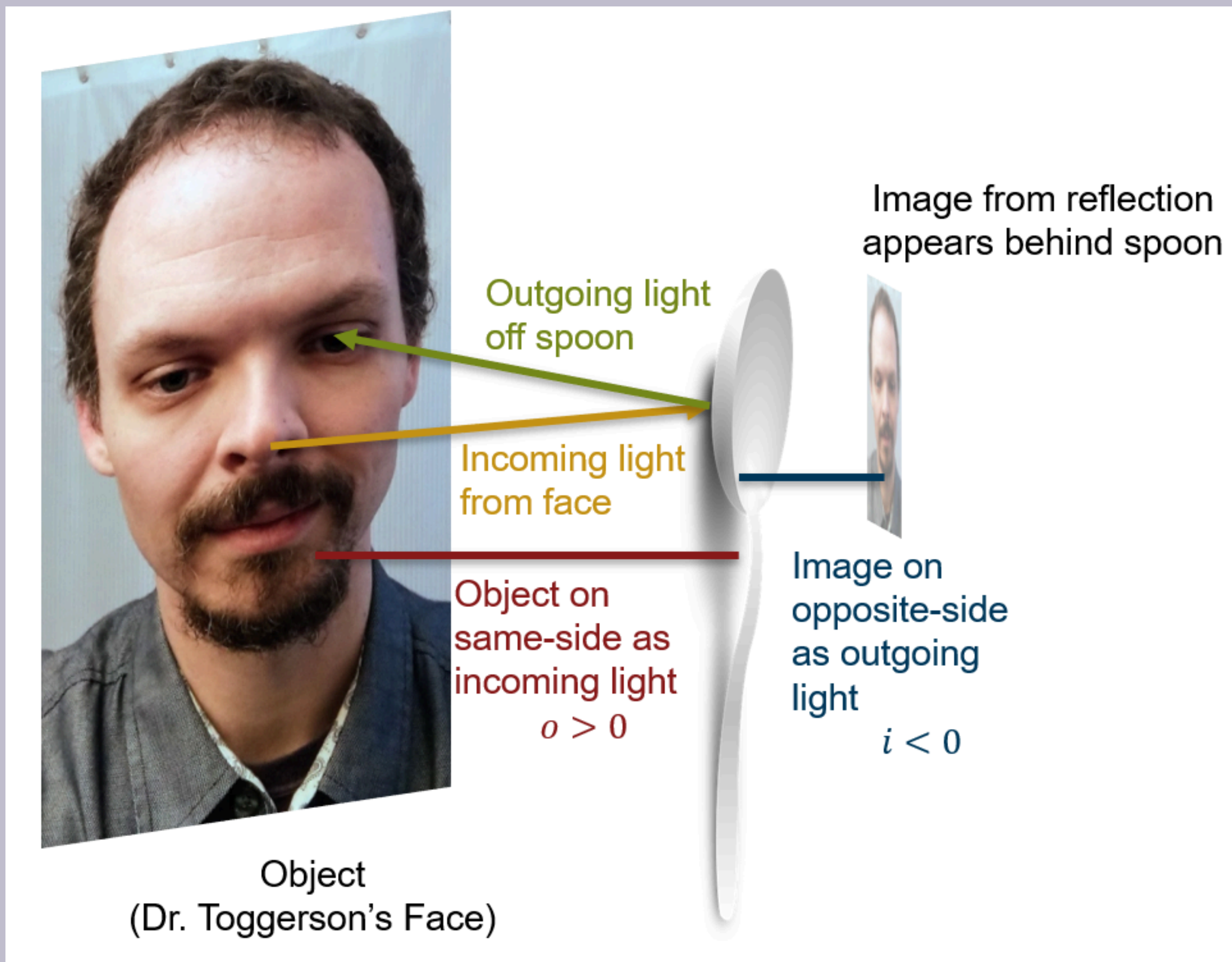
Figure 8a: *Dr. Toggerson in the back of a spoon. His image is behind the spoon and definitely smaller than he is!*

Figure 8b: *Dr. Toggerson’s reflection on the inside of a spoon. Again, the image is smaller than the object. This time, the image is also inverted. Finally, if you look closely, the image appears to hover in front of the spoon!*

Solution – back of spoon:

For the back of the spoon:

- The light is coming off your face and hitting the spoon, thus the object is on the same side as the incoming light: $o > 0$ 0 " title="Rendered by QuickLaTeX.com" height="12" width="38" style="vertical-align: 0px;">.
- The image is on the opposite side as the outgoing light: the light bounces back towards you and does not go behind the spoon. Thus $ilt; 0$

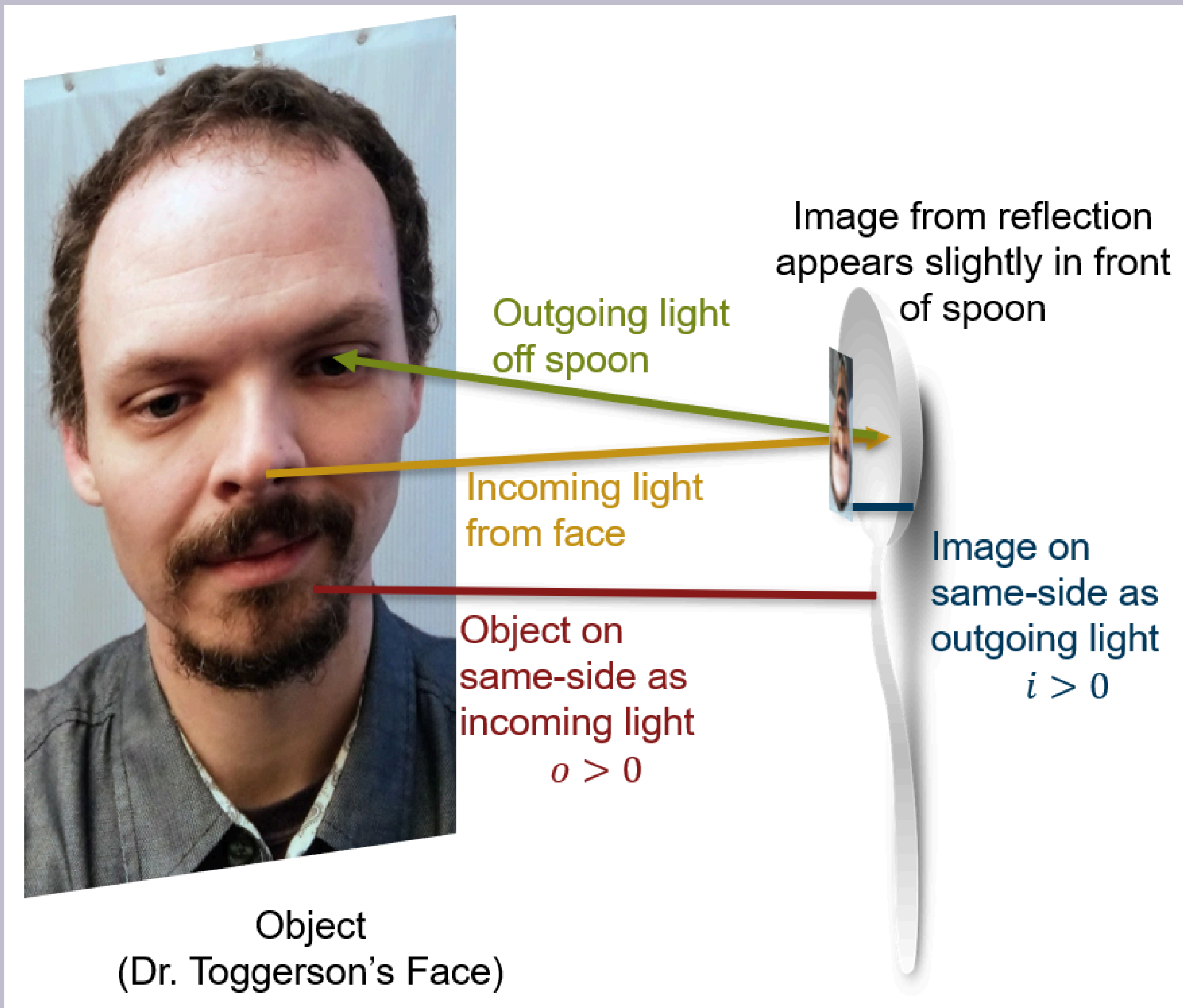


The light comes from the face, bounces off the spoon and into the eye. The object (face) is on the same side as the incoming light so has a positive object distance. The image is on the opposite side as the outgoing light so has a negative image distance.

Solution – Front of spoon

If you are looking in a spoon carefully, you will see that the image appears to hover in front of it. As such:

- The object is still on the same side as the incoming light: $o > 0$ 0 " title="Rendered by QuickLaTeX.com" height="12" width="38" style="vertical-align: 0px;">
- The image is now on the same side as the outgoing light: $i > 0$ 0 " title="Rendered by QuickLaTeX.com" height="12" width="36" style="vertical-align: 0px;">



The light comes from the face, bounces off the spoon and into the eye. The object (face) is still on the same side as the incoming light so has a positive object distance. The image, however, is now on the same side as the outgoing light so also has a positive image distance.

Discussion:

I know that these rules seem odd and obtuse and that it will be tempting to make up your own rules: DON'T. These rules work for all situations. You just need to think about the perspective of where the light is traveling.

Section Summary

- An optical element is a lens or a mirror.
- An image is the apparent reproduction of an object, formed by an optical element or collection of them,

through the refraction or reflection of light.
- Images can either be erect, right side up, or inverted, upside down.
- The optical axis is an imaginary line that passes through the optical element perpendicular to it.
- The point where this optical axis meets the optical element is called the vertex, which is at the center of a lens or the surface of a mirror.
- Image and object distances are measured along the optical axis from the vertex, and the signs of the image distance, i, and the object distance, o, are relative to the path of light.
- If the object is on the same size the incoming light, then the object distance is positive.
- If the image is on the same side is the outgoing light, then the image distance is positive.
- For a lens the incoming and outgoing sides are different because the light goes through it.
- mirror on the other hand, the incoming and outgoing sides are the same because the light bounces.

Homework Problems

Problem 15: Sign conventions for object and image distances: objects and images on opposite sides of the optical element.

Problem 16: Sign conventions for object and image distances: objects and images on same side of the optical element.

Magnification of Images

Instructors Note

Your quiz will cover:

- Given two of magnification, image height, and object height, find the third
- Know that the magnification of an inverted image is negative

When you look in a flat bathroom mirror, your image is the same height as you are. Look at the picture of Dr. Toggerson below in Figure 1. Figure 1a is taken 20cm in front of a mirror (you can see the camera). The image seems to be exactly the same distance, 20cm, behind the mirror as Dr. Toggerson is in front of it. This is shown diagrammatically in Figure 2 where the woman is the same distance in front of the mirror as her reflection is behind it; also she and her reflection are the same height. This is confirmed with Figure 1b, which is taken with the camera at 40cm from Dr. Toggerson's face as the mirror. The height of Dr. Toggerson's face is the same in Figures 1a and 1b. In contrast, look at Figure 3a below which is taken in the back of a spoon, Dr. Toggerson's face is much smaller, about 2.5cm (1 in, it fits on the spoon!). Similarly, in Figure 3b looking at the inside of a spoon, his image, while larger than in the back of the spoon, is again smaller than his face. Moreover, the image on the inside of the spoon is upside down! The flat mirror and the back of the spoon have **erect** images, while the image formed by the inside of the spoon is **inverted**. The different heights are covered by the concept of magnification. Magnification is simply the ratio of image height to object height:

$$m = \frac{h_i}{h_o}$$

where h_i is the height of the image and h_o is the height of the object. As a ratio of heights, magnification is, itself, unit-less and just represents how many times bigger or smaller the image is than the object. **One last note: if the image is inverted, we say that its height is negative. Thus inverted images have negative magnification.** Hopefully, this choice seems reasonable.

Figure 1: Two pictures of Dr. Toggerson. Note that he seems to be the same height in each. The image in a flat mirror is the same height as the object.

Figure 1a: *A selfie of Dr. Toggerson taken 20cm in front of a mirror*

Figure 1b: *A selfie of Dr. Toggerson taken from 40cm away. Twice the face-mirror distance*

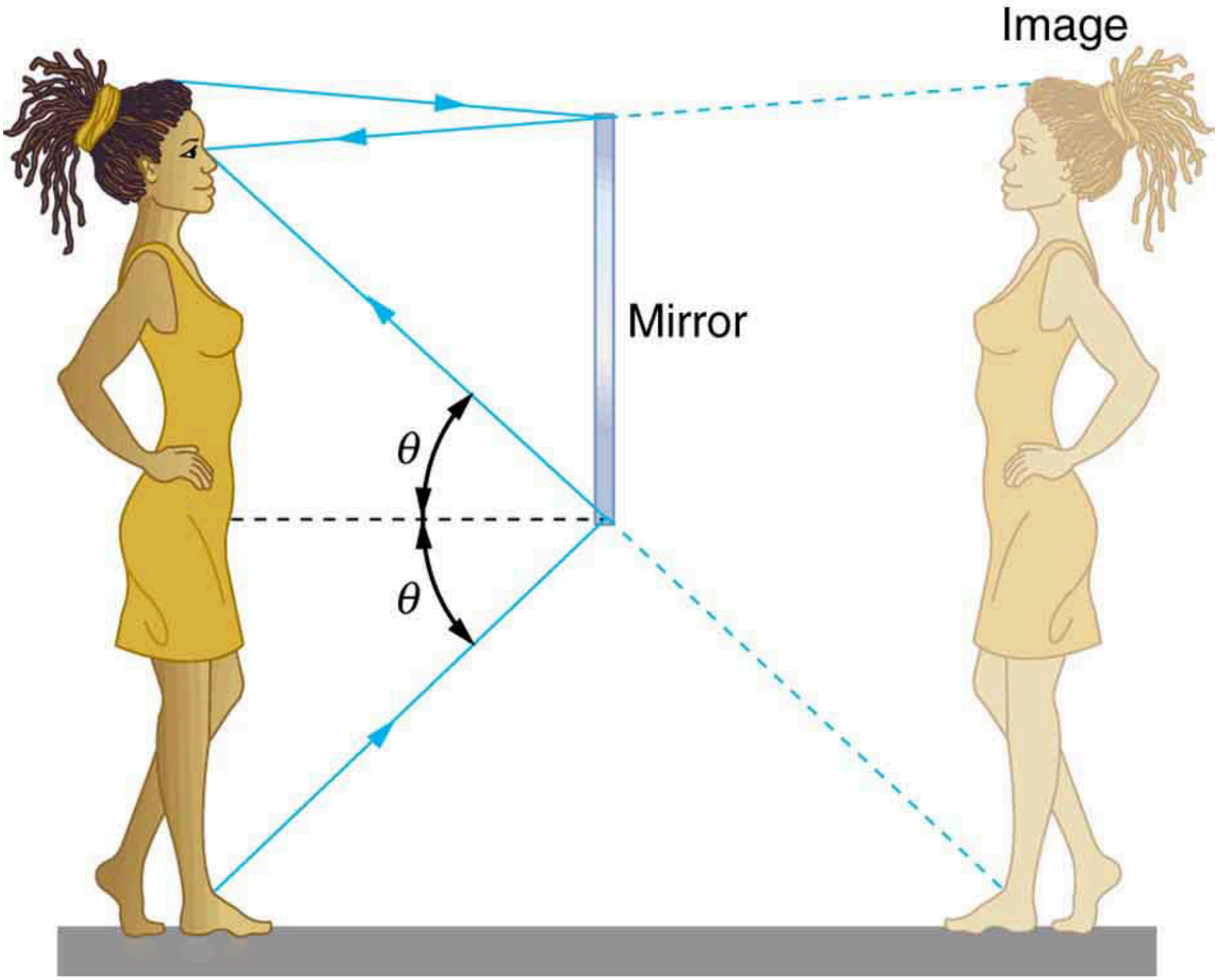


Figure 2: The image in a mirror is the same height as the original object and the same distance behind the mirror as the object is in front of it.

Figure 3: Dr. Toggerson's reflections in a spoon.

Figure 3a: *Dr. Toggerson in the back of a spoon. His image is behind the spoon and definitely smaller than he is!*

Figure 3b: *Dr. Toggerson's reflection on the inside of a spoon. Again, the image is smaller than the object. This time, the image is also inverted. Finally, if you look closely, the image appears to hover in front of the spoon!*

Magnification in a Flat Mirror

As we have seen, in a flat mirror, your image is the same height as you. What is the magnification of a flat mirror?

Solution:

Well, we know the definition of magnification:

$$m = \frac{h_i}{h_o}$$

Let's use Dr. Toggerson as our example. Dr. Toggerson is about 1.7m tall, which means his reflection in a flat mirror is also 1.7m tall. Using these values we see:

$$m = \frac{1.7\,\mathrm{m}}{1.7\,\mathrm{m}} = 1$$

Discussion:

Note, the answer is just 1, no units!

Magnification in a Spoon

Dr. Toggerson is 1.7m tall. His image in the back of a spoon is about 0.5cm. The image in the front of the spoon is about 1cm. What is the magnification of each?

Solution for the back of the spoon:

Again, we know the definition of magnification:

$$m = \frac{h_i}{h_o}$$

All we need to do is substitute the known values. However, for the result to be unit-less we need to express both the image height and the object height in the same units; doesn't matter what units we use, they just have to be the same. Let's use meters:

$$m = \frac{0.005\,m}{1.7\,\mathrm{m}} = 0.0029$$

The image is 0.3% of Dr. Toggerson's height!

Solution for the front of the spoon:

The procedure is essentially the same, the only difference this time is that the image height should be considered to be negative $h_i = 0.5\,\mathrm{cm}$ as the image is inverted. Also, for fun, lets work in cm this time:

$$m = \frac{h_i}{h_0}$$

$$m = \frac{-1.0\,\mathrm{cm}}{170\,\mathrm{cm}} = -0.0059$$

The magnification is negative, again representing the fact that the image is inverted.

Homework Problems

Problem 17: What characterizes an object with a negative magnification?

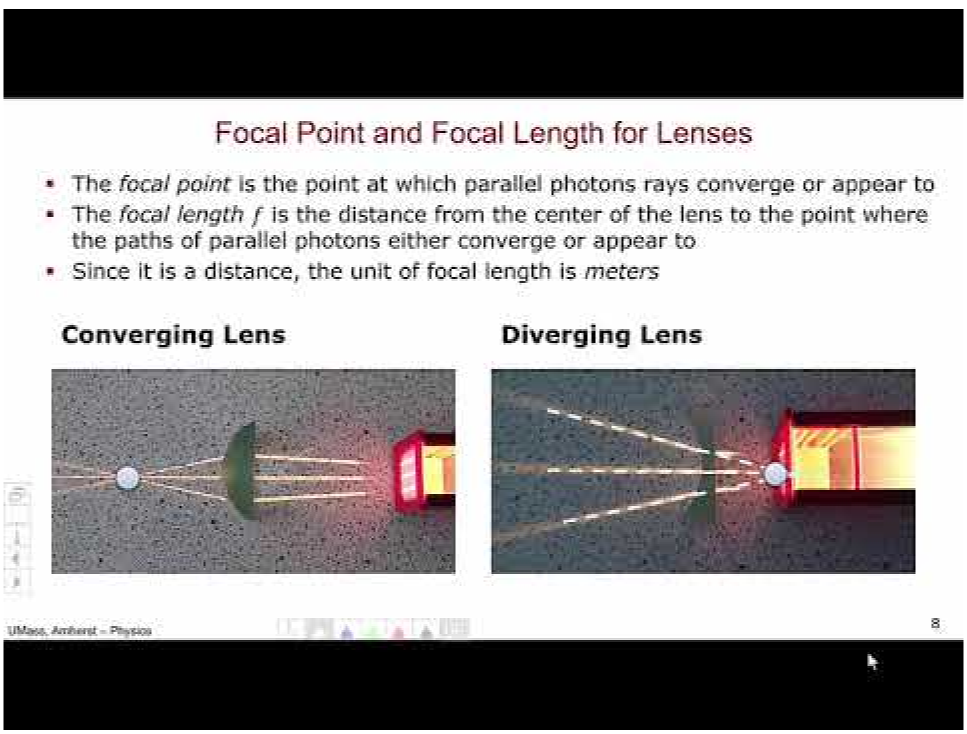
Problem 18: Calculating magnification of a gemstone.

Introduction to Lenses

Instructor's Note

By the end of this section you should:

- Describe what a lens is and describe the two kinds of lenses, converging and diverging.
- Describe what each type does and how these two types of lenses are different.
- Define focal length and focal point for any lens.
- Describe the idea behind the thin lens approximation.



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=499

What is a lens?

A lens is a piece of transparent material designed to take incoming parallel photons and can bend them to a point which we call a converging lens, pictured below, where you have incoming parallel photons from the right passing into the lens and converging to a point on the left.

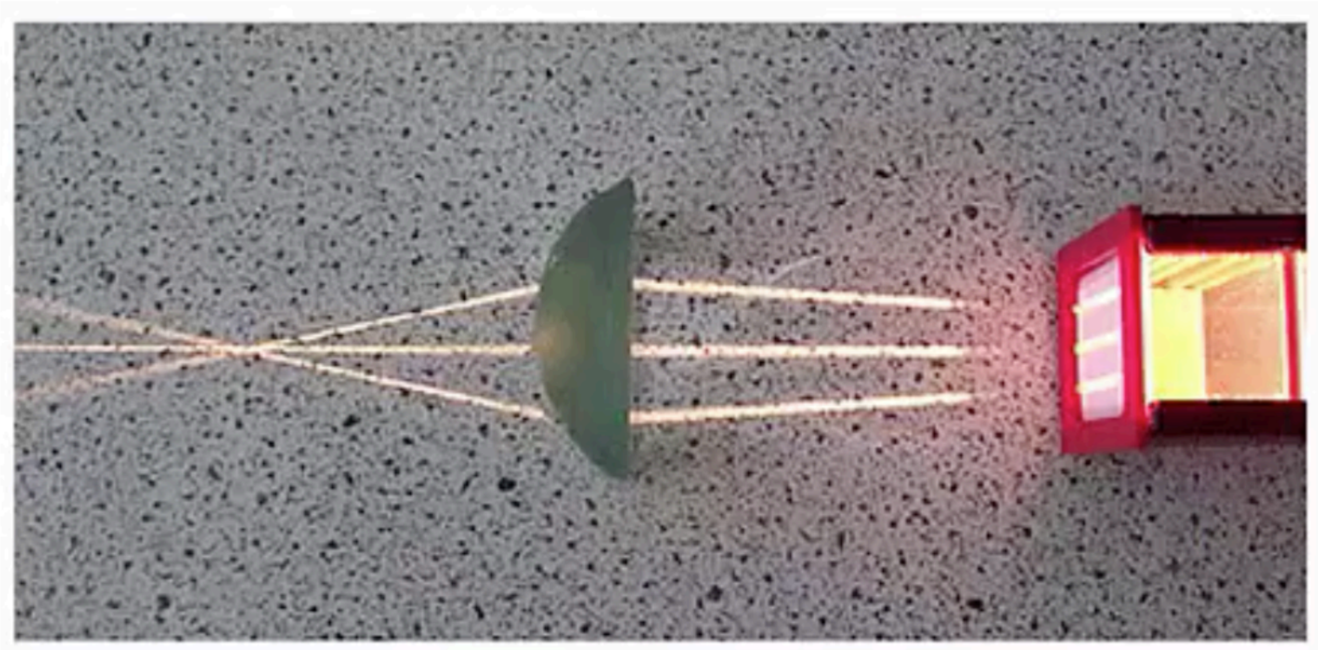


Figure 1.

A common application that you might have experienced with this is taking a regular magnifying glass and focusing the sun's light to a point. The incoming light from the Sun is effectively parallel and the magnifying glass focuses it to a point.

The other thing a lens can do is spread incoming parallel photons as if they came from a point. This is known as a diverging lens. When the light passes through the lens, the light spreads out and from the left side it appears that the light originated at this point.

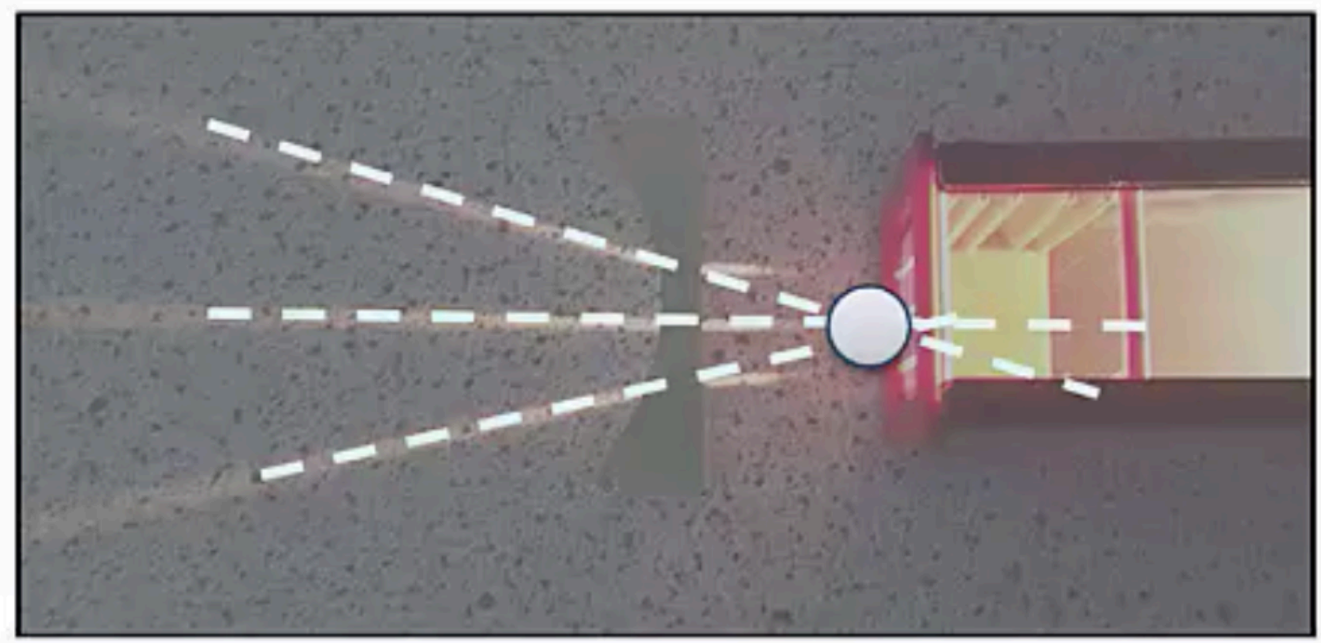


Figure 2.

A common application of diverging lenses is on your face if you happen to be nearsighted, the glasses you're wearing if you're nearsighted are a diverging lens.

Let's talk a little bit about the properties of converging versus diverging lenses. Converging lenses are always thicker at the center than at the edges. On the flip side, diverging lenses are thicker at the edges than they are in the middle.

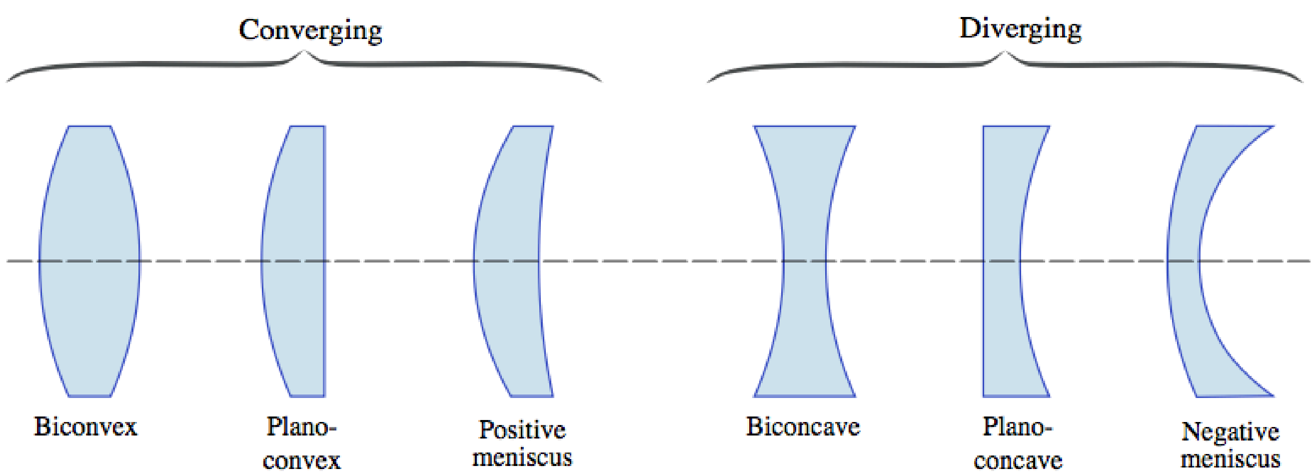


Figure 3.

Instructor's Note

You don't need to know the names of these particular shapes you just need to know that converging lenses are thicker at the middle and diverging lenses are thicker at the edges.

Let's move on to probably the most important property of lens, the focal length and focal points. Here are our two lenses, the converging lens on the left and the diverging lens on the right, in both cases you can see the light coming from the right passing through the lens on its way to the left.

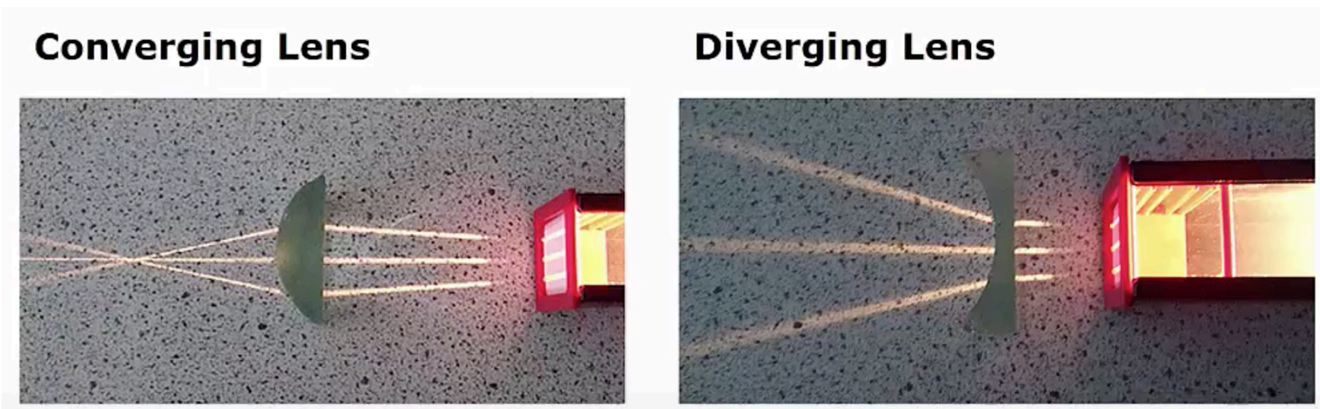


Figure 4.

The focal point is defined as the point at which the parallel photons either converge or appear to. Meanwhile the focal length is the distance from the center of the lens to that point, since the focal length is a distance the unit of focal length is meters.

For the converging lens, the point where the photons appear to intersect is the focal point. Meanwhile for the diverging lens we trace the rays back through the lens and they appear to come from this point behind the lens, this is the focal point of a diverging lens.

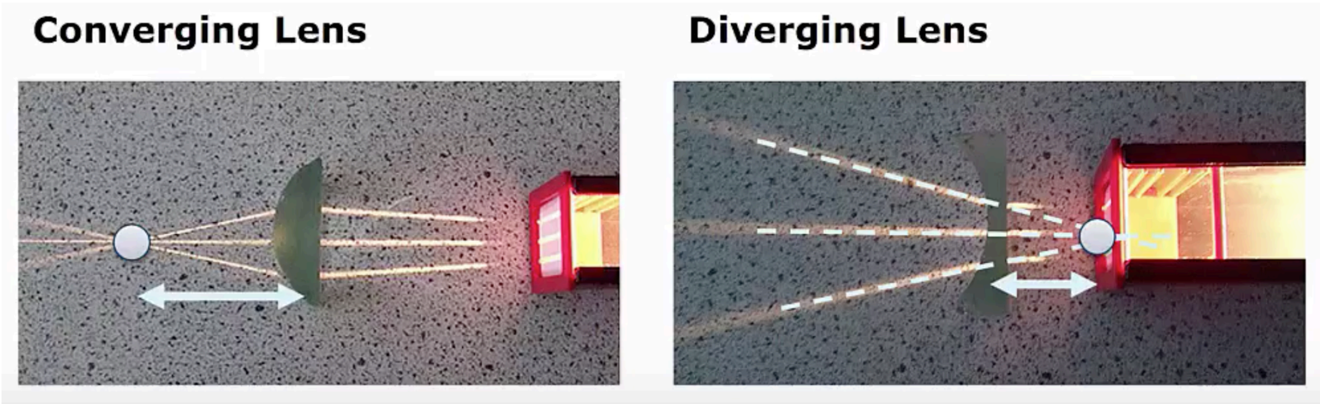


Figure 5.

We always measure the focal length from the center of the lens to the focal point, so you can see it here on the left for the converging lens and here on the right for the diverging lens. Every lens has only one focal length, but two focal points with one on each side.

Below we can see the full set up for converging lens. In this case we've got the light coming in from the left and moving towards the right. We can see the two different focal points, one on each side of the lens, each a focal length from the center of the lens and these two focal lengths are the same.

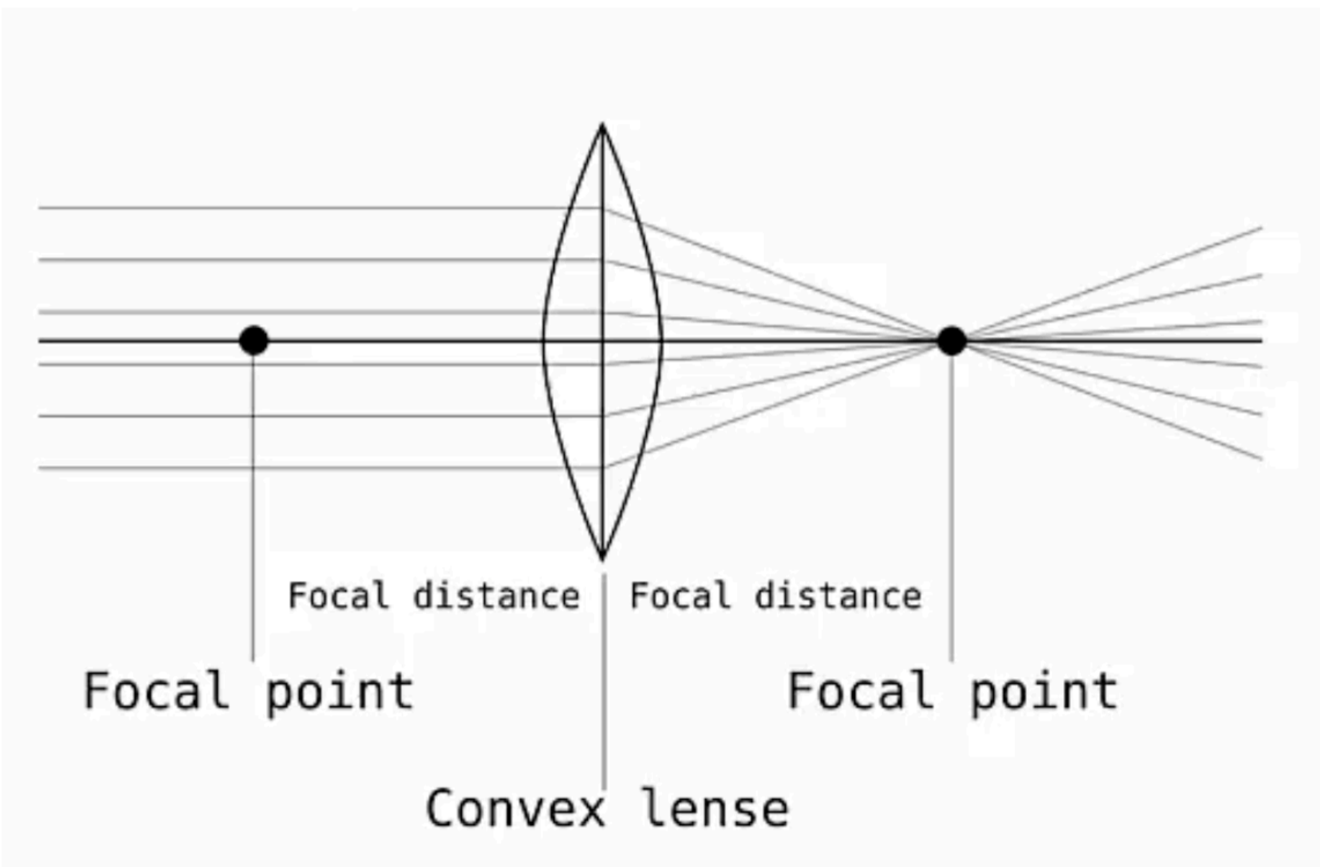


Figure 6.

Similarly, for a diverging lens we have two focal points one on each side of the lens, each the focal length from the center. The convention is that focal lengths for converging lenses are positive while diverging lenses have negative focal lengths.

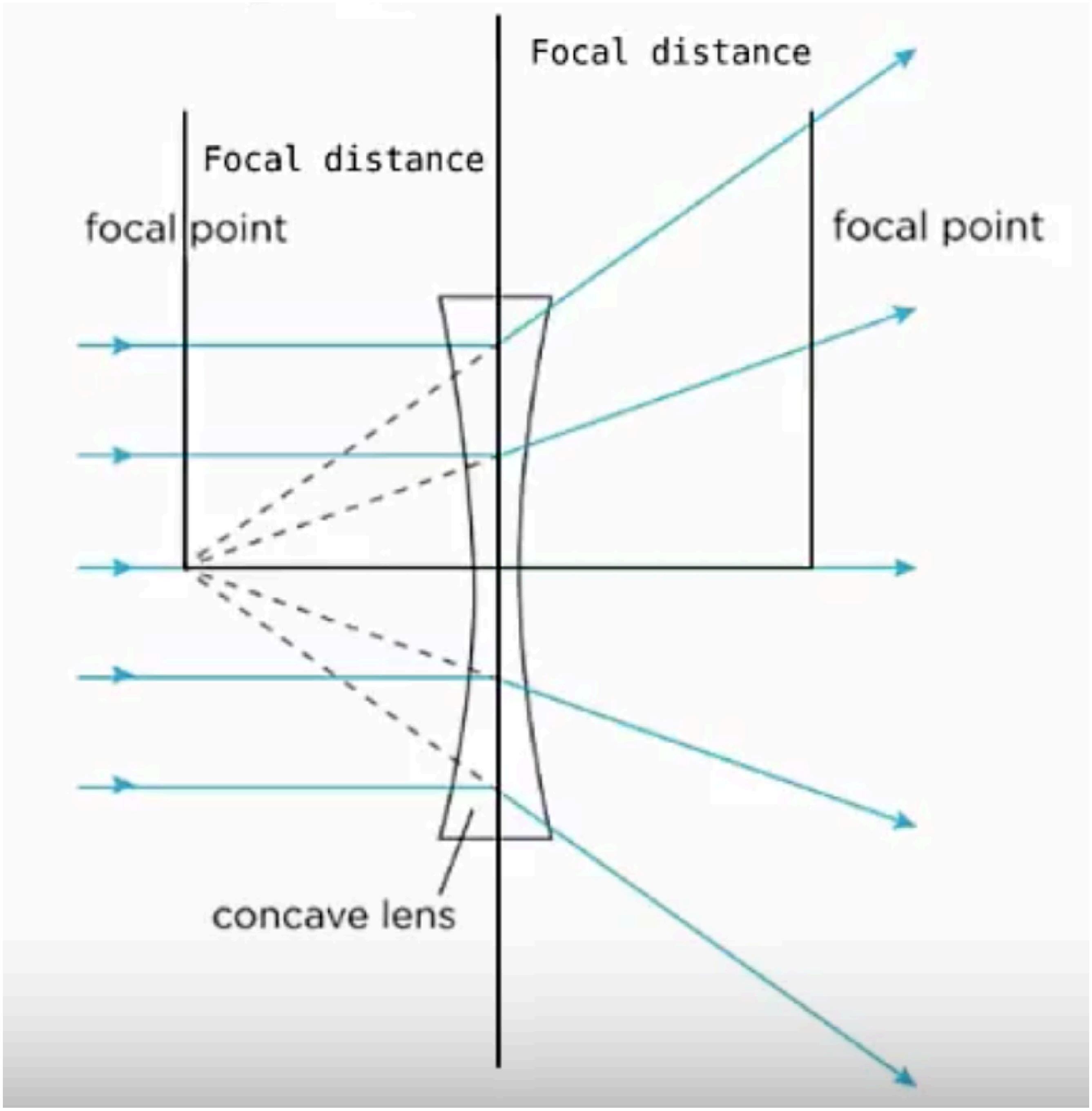


Figure 7.

Instructor's Note

In addition to a lot of vocabulary optics also has a lot of sign conventions that I am expecting you to learn from this prep, there is a set on flashcards on [Quizlet]() to help you.

Finally let's move on to the thin lens approximation. In a real lens, light will bend at each surface, it will bend as it goes from the air to glass, or whatever material the lens is made of, and then it will bend again by Snell's law when the light moves from the glass back into the air.

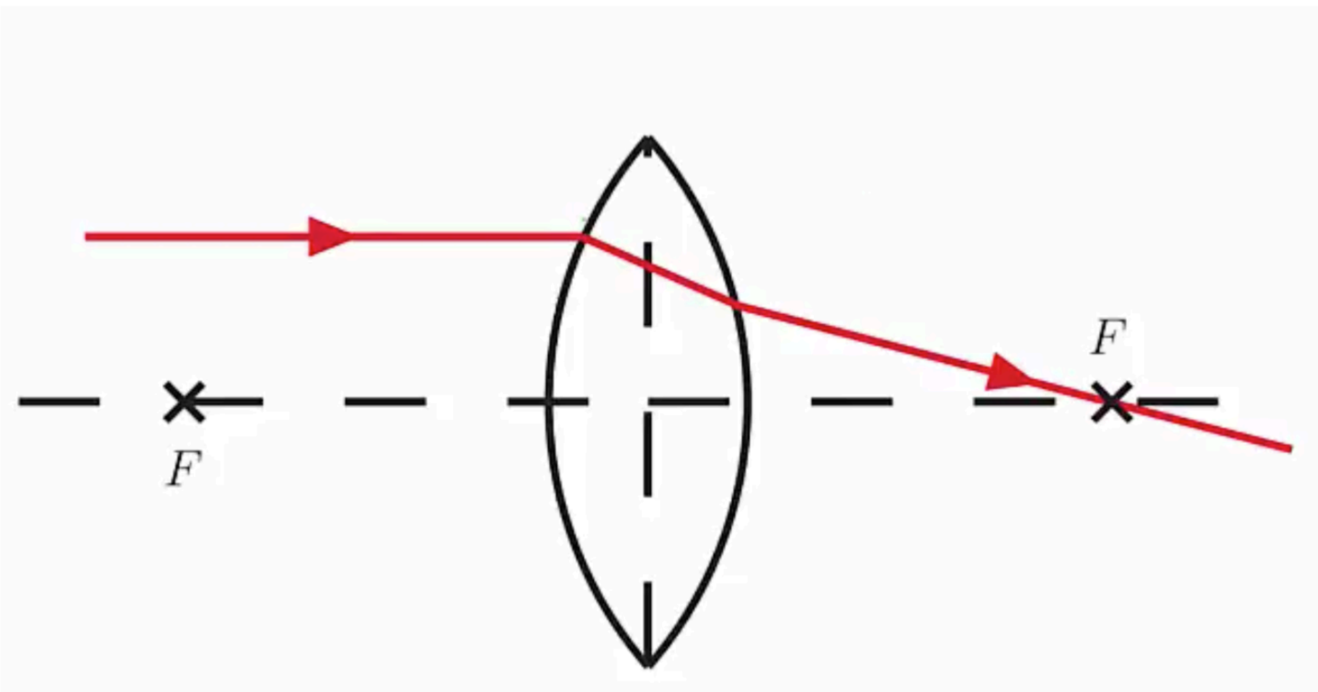


Figure 8.

However, we will assume that the lens is very thin. What do we mean by thin? We mean that the thickness of the lens is very small relative to the focal length. Under this approximation it's essentially as if all the light bending happens at the center. Now this is not actually what happens, remember the light does bend at each interface, air to glass and glass to air, however, if the lens is very thin these two interfaces are so close together and so close to the center of the lens that we can ignore it.

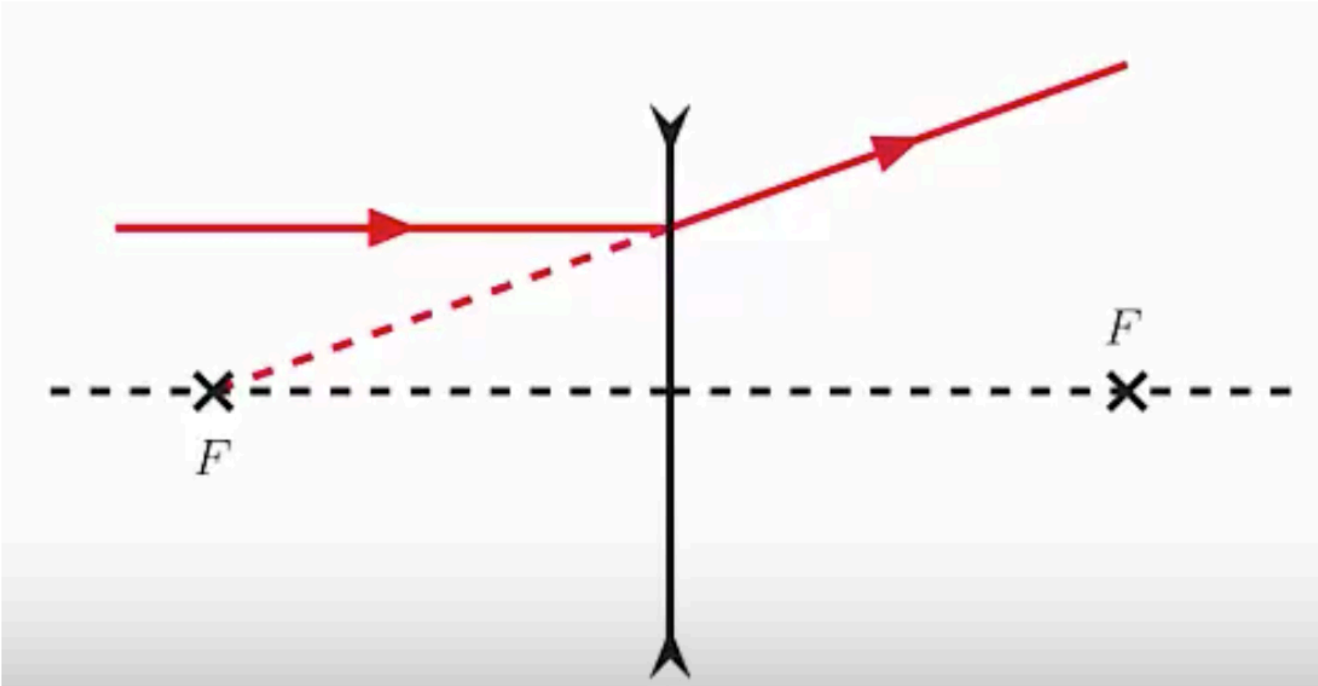


Figure 9.

Section Summary

- A lens is a transparent material that either bends parallel light towards a focal point in the case of a converging lens or bends the light away in such a way that it appears to originate from a point in the case of a diverging lens.
- Each lens has two focal points one on each side of the lens because we can of course send the light through the lens either from the left or from the right, and we've seen examples of light traveling both ways.
- Focal lengths for converging lenses are positive and focal lengths for diverging lenses are by convention negative.
- In reality, light bends at each surface of the lens, however, if the lens is thin relative to the focal length then we can treat all of the bending as if it is happening at the center of the lens, this is known as the thin lens approximation.

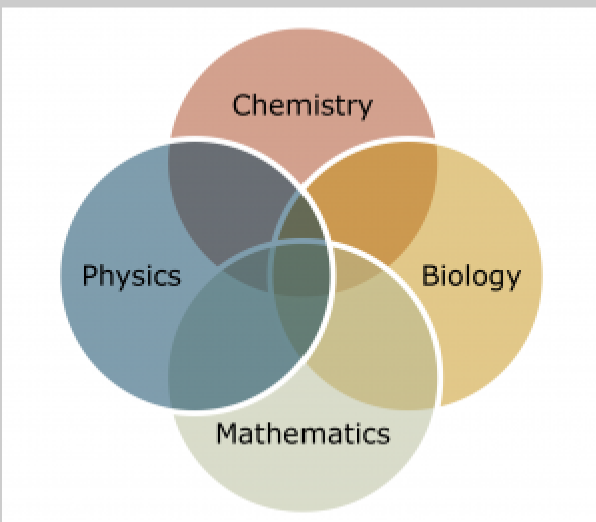
Homework Problems

Problem 19: Characterizing lenses

Lenses Specifically as Applied to the Human Eye

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor’s Note



In this section, you will return to the overview of the human eye that was introduced in the motivation for this unit. Here, you will begin to see how to connect what you may know from biology, to what you have just read about lenses. What we expect you to know from this reading is:

- Most of the focusing of the eye does NOT happen at the lens, but instead happens at the air-to-water interface at the front of the cornea.
- The lens of the eye provides the fine-tuning of the focus to get the image distance to land on

the retina.

- The cornea and lens together act as a single lens system.
- In medicine, people often talk of the **power** of a lens: this is 1/f where f is the **focal length**. Your glasses prescription, if you have one, is listed in diopters.
  - Note, that while this power shares the same name as the power discussed in Unit I, it is a different quantity (Homonyms – they are not just for English!)
  - The unit of optical power is 1/m or a *diopter.*
  - This is consistent as a lens with a *shorter* focal length, bends the light *more strongly* – has a higher *power.*

The eye is remarkable in how it forms images and in the richness of detail and color it can detect. However, our eyes often need some correction to reach what is called “normal” vision. Actually, normal vision should be called “ideal” vision because nearly one-half of the human population requires some sort of eyesight correction, so requiring glasses is by no means “abnormal.” Image formation by our eyes and common vision correction can be analyzed with the optics discussed earlier in this chapter.

Figure 1 shows the basic anatomy of the eye. The corneaand lens form a system that, to a good approximation, acts as a single thin lens. For clear vision, a real image must be projected onto the light-sensitive retina, which lies a fixed distance from the lens. The flexible lens of the eye allows it to adjust the radius of curvature of the lens to produce an image on the retina for objects at different distances. The center of the image falls on the fovea, which has the greatest density of light receptors and the greatest acuity (sharpness) in the visual field. The variable opening (i.e., the pupil) of the eye, along with chemical adaptation, allows the eye to detect light intensities from the lowest observable to 10101010">10^{10} times greater (without damage). This is an incredible range of detection. Processing of visual nerve impulses begins with interconnections in the retina and continues in the brain. The optic nerve conveys the signals received by the eye to the brain.

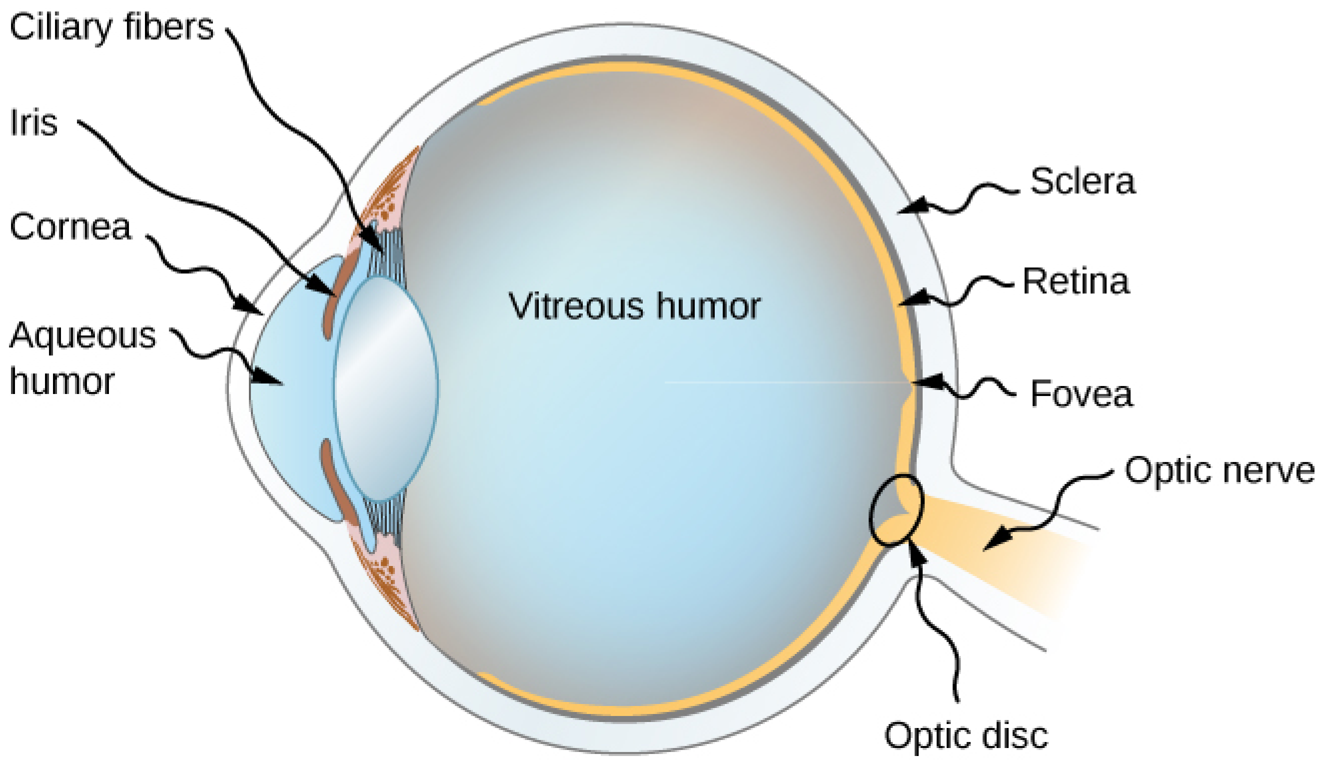


Figure 1: *The cornea and lens of the eye act together to form a real image on the light-sensing retina, which has its densest concentration of receptors in the fovea and a blind spot over the optic nerve. The radius of curvature of the lens of an eye is adjustable to form an image on the retina for different object distances. Layers of tissues with varying indices of refraction in the lens are shown here. However, they have been omitted from other pictures for clarity.*

The indices of refraction in the eye are crucial to its ability to form images. Table 1 lists the indices of refraction relevant to the eye. The biggest change in the index of refraction, which is where the light rays are most bent, occurs at the air-cornea interface rather than at the aqueous humor-lens interface. The ray diagram in Figure 2 shows image formation by the cornea and lens of the eye. The cornea, which is itself a converging lens with a focal length of approximately 2.3 cm, provides most of the focusing power of the eye. The lens, which is a converging lens with a focal length of about 6.4 cm, provides the finer focus needed to produce a clear image on the retina. The cornea and lens can be treated as a single thin lens, even though the light rays pass through several layers of material (such as cornea, aqueous humor, several layers in the lens, and vitreous humor), changing direction at each interface. The image formed is much like the one produced by a single convex lens (i.e., a real, inverted image). Although images formed in the eye are inverted, the brain inverts them once more to make them seem upright.

>Table 2 Refractive Indices Relevant to the Eye *This is an average value. The actual index of refraction varies throughout the lens and is greatest in center of the lens.

Material	Index of Refraction
Water	1.33
Air	1.0
Cornea	1.38
Aqueous humor	1.34
Lens	1.41*
Vitreous humor	1.34

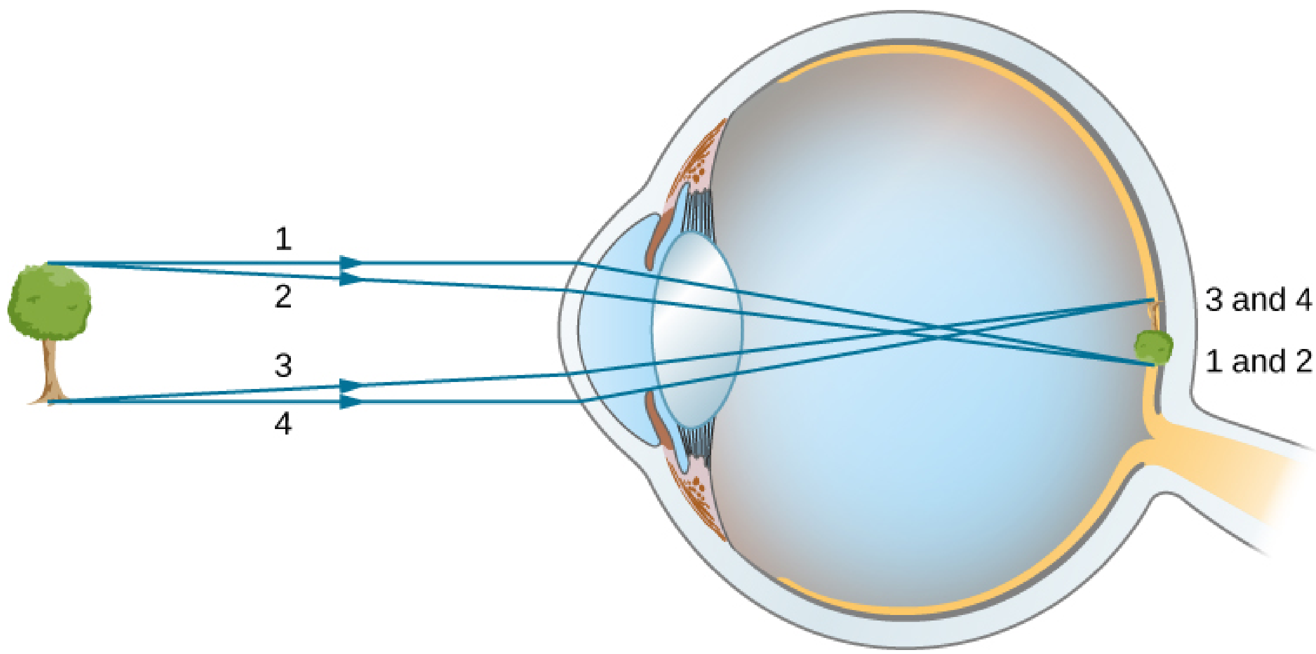


Figure 2: In the human eye, an image forms on the retina. Rays from the top and bottom of the object are traced to show how a real, inverted image is produced on the retina. The distance to the object is not to scale.

As noted, the image must fall precisely on the retina to produce clear vision—that is, the image distance i must equal the lens-to-retina distance. Because the lens-to-retina distance does not change, the image distance i must be the same for objects at all distances. The ciliary muscles adjust the shape of the eye lens for focusing on nearby or far objects. By changing the shape of the eye lens, the eye changes the focal length of the lens. This mechanism of the eye is called accommodation.

The nearest point an object can be placed so that the eye can form a clear image on the retina is called the near point of the eye. Similarly, the far point is the farthest distance at which an object is clearly visible. A person with normal vision can see objects clearly at distances ranging from 25 cm to essentially infinity. The near point increases with age, becoming several meters for some older people. In this text, we consider the near point to be 25 cm.

We define the optical power of a lens as

$$P = \frac{1}{f}$$

with the focal length *f* given in meters. The units of optical power are called "diopters" (D). That is $\mathrm{D} = 1/\mathrm{m} = \mathrm{m}^{-1}$. Optometrists prescribe common eyeglasses and contact lenses in units of diopters.

Working with optical power is convenient because, for two or more lenses close together, the effective optical power of the lens system is approximately the sum of the optical power of the individual lenses:

$$P_{\mathrm{total}} = P_{\mathrm{lens}_1} + P_{\mathrm{lens}_2} + P_{\mathrm{lens}_3} + \ldots$$

Homework Problems

Problem 20: Diopters to focal length.

Problem 21: Focal length to diopters.

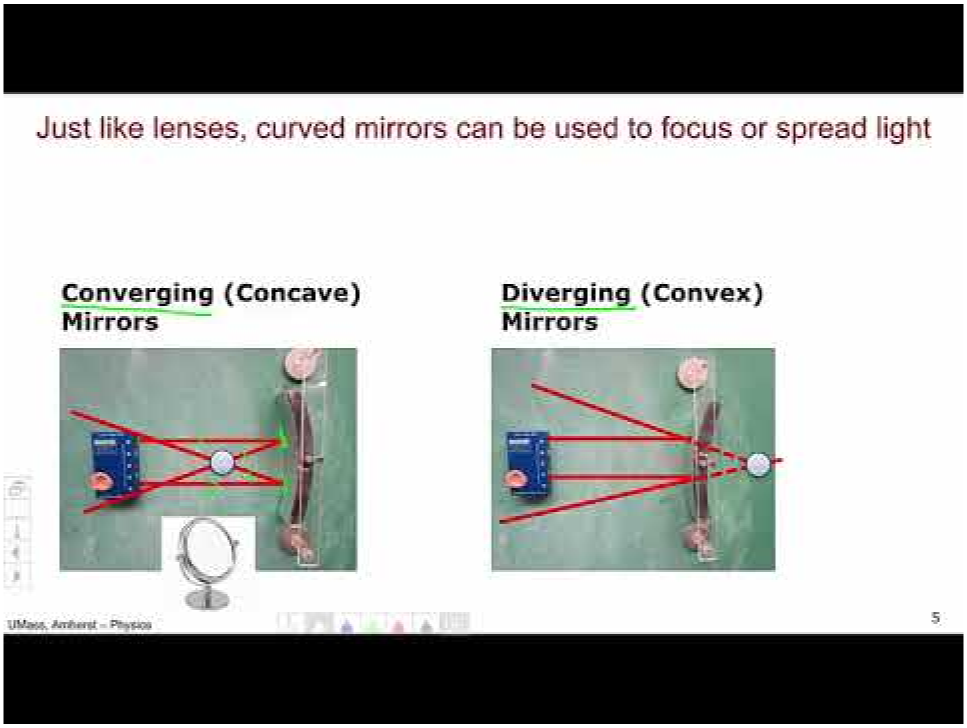
Introduction to Mirrors

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

By the end of this section you should:

- Be able to look at a mirror and say if it's either converging or diverging.
- Know the connection between a mirrors radius and its focal length.



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=499*](http://openbooks.library.umass.edu/toggerson-132/?p=499)

Just like lenses, curved mirrors can be used to focus light to a point or spread it out. Mirrors come in two basic shapes, we have concave mirrors, such as this mirror below on the left, which has a concave shape bending towards the light source, or we can have a convex mirror, pictured below on the right, that curves away from the light source.

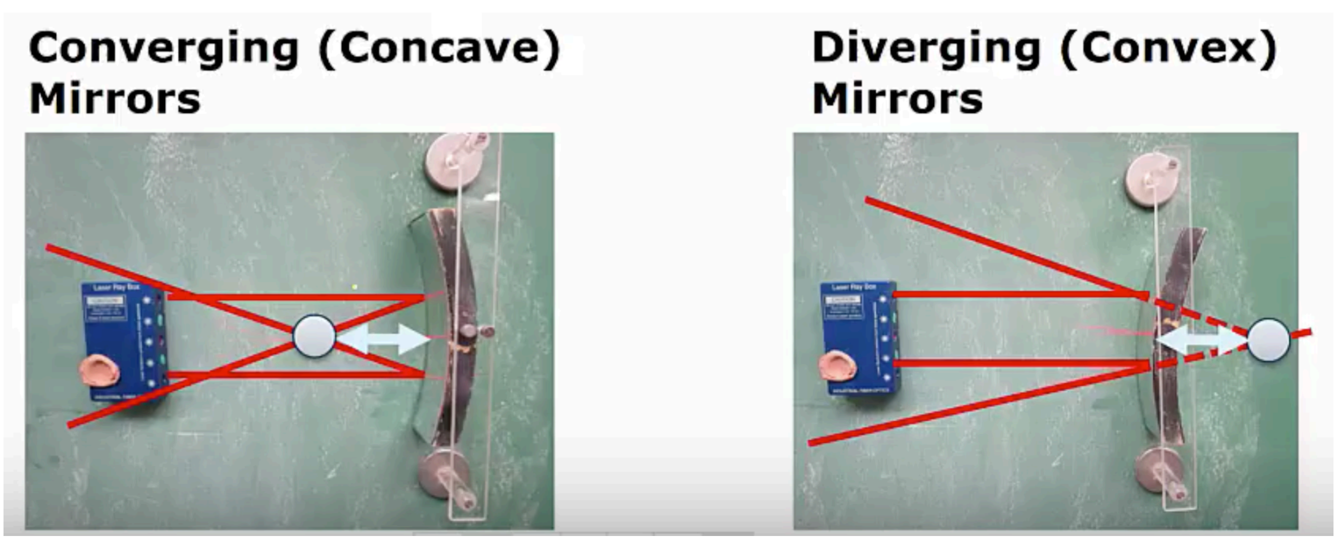


Figure 1.

Concave mirrors are known also as converging mirrors and convex mirrors are also known as diverging mirrors. We can see for the concave mirror that the two light rays have been converged to a point. In this sense, the concave converging mirror is functioning similar to a converging lens, it's taking incoming parallel light rays and converging them to a point. A common example of this particular type of mirror in use is the shaving mirror or makeup mirror that you might have in your bathroom that lets you see your face a little bit larger.

On the flip side, a convex mirror takes the incoming light rays and causes them to spread out. Just as we did with the lens, if we sort of imagine what someone to the left of the light source sees, our eyes kind of assume that light travels in straight lines so we trace the rays back and the light rays appear to originate from a point behind the mirror.

In this case, a convex mirror is providing a very similar function to a diverging lens, taking incoming parallel light and producing outgoing light that's diverging as if it came from some point. You can see convex mirrors quite frequently as security mirrors in buildings.

Instructor's Note

All the mirrors we will talk about in this particular class will be called spherical mirrors, meaning that the mirror is part of a sphere, while most mirrors are not actually spherical, it turns out that studying spherical mirrors is a very good approximation for most real curved mirrors.

Let's talk about focal lengths and focal points for mirrors. Just like lenses, the focal point is the point where the photons either converge to or appear to emanate from. For the concave mirror, which is converging on the left, we've already seen that light comes in and converges to some point, this is known as the focal point. For the diverging mirror, the light comes in and bounces off, light comes in and bounces off, and the light appears to emanate from some point behind the mirror, so this would be the focal point for this diverging mirror.

Unlike lenses, mirrors only have a single focal point, and this stems from the fact that light can go through a lens from either direction, we can put the light on either side of the lens and the light will go through. For a mirror on the other hand, a mirror only has one reflective surface, there's only one side that will act as a mirror, and therefore mirrors only have a single focal point.

As with lenses, the focal length, which we designate f is the distance in meters from the surface of the mirror to the focal point. For the converging lens, below on the left, this blue arrow represents the focal length f and for the diverging mirror we have this focal length below on the right.fff">

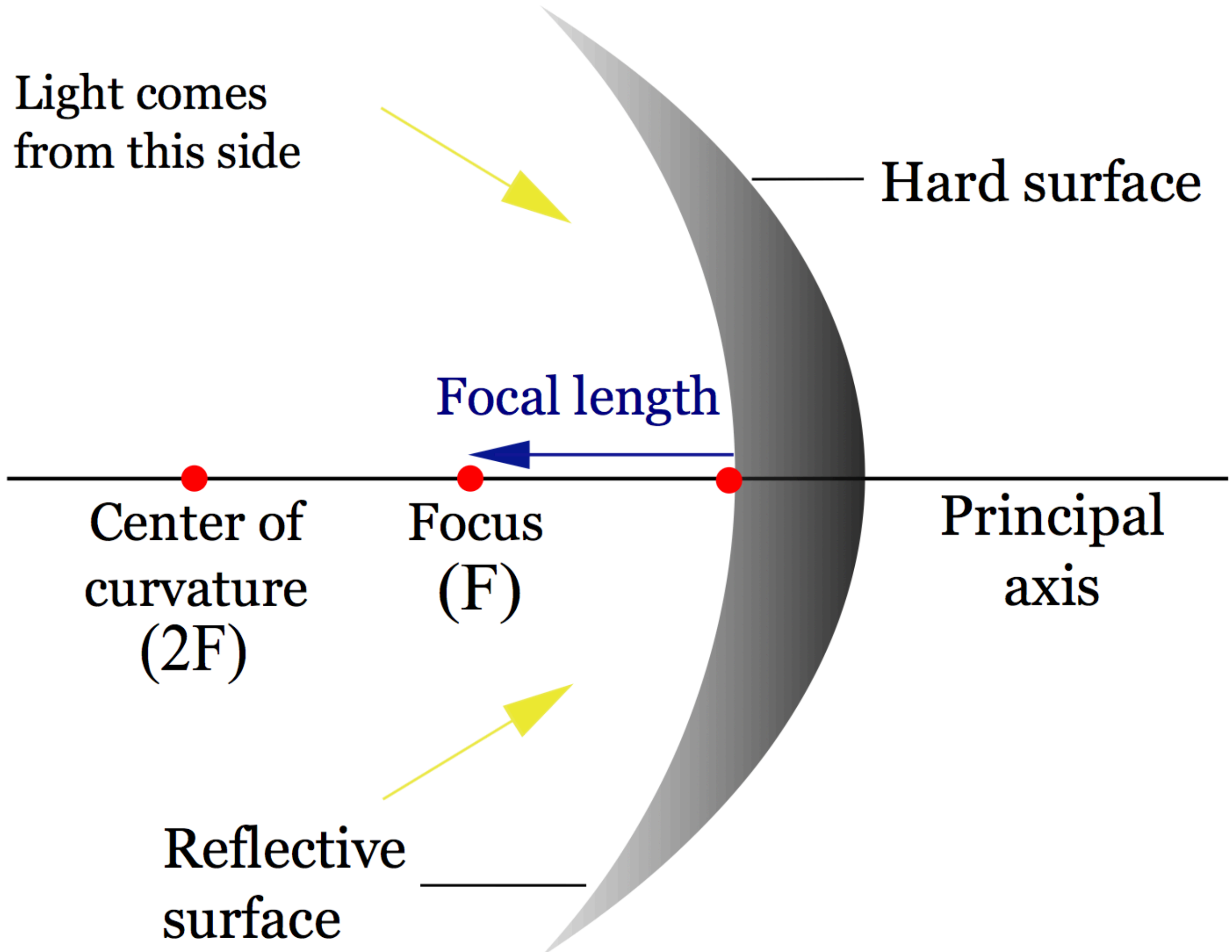


Figure 2. *Concave Mirror (Credit: Cronholm144)*

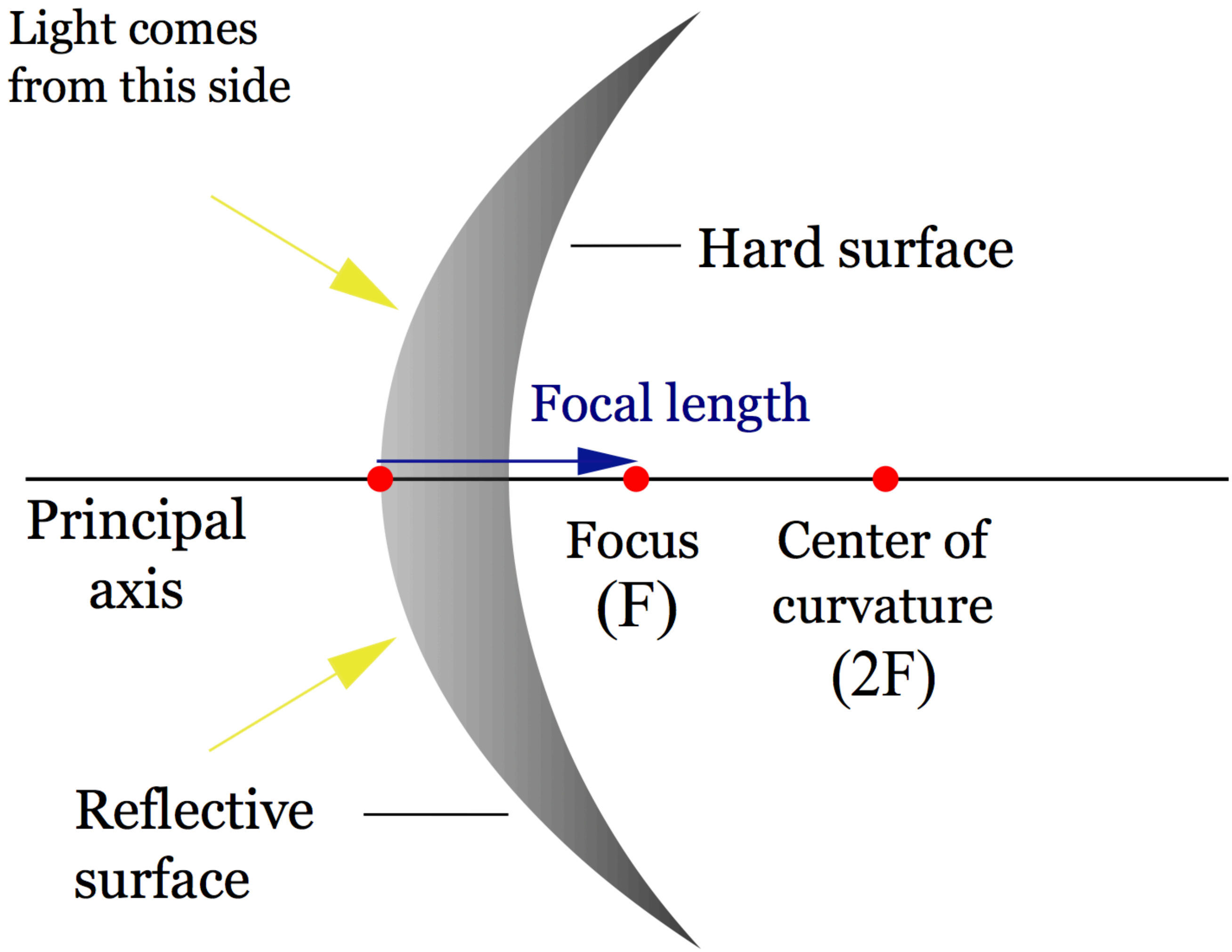


Figure 3. *Convex Mirror (Credit: Cronholm144)*

And just as with lenses, converging mirrors have a positive focal length and diverging mirrors have a negative focal length. You can start to see some similarity between mirrors and lenses. When things are converging you have positive focal lengths, and when things are diverging you have negative focal lengths.

Now let's talk a little bit about focal lengths and the radius of curvature.

Instructor's Note

I want you to remember that all the mirrors we will be talking about here are parts of a sphere.

We need two new terms, the center of curvature, which represents the center of the sphere of which the mirror is a part, and the radius of curvature, which is the radius of which the mirror is a part.

The focal length of any spherical mirror is half the radius, so above we have a concave mirror on the top with positive focal length and a convex mirror on the bottom with negative focal length, the distance R is the radius of the sphere of which this mirror is a part and the focal length is half that radius

$f = \frac{R}{2}$.rrr">

Section Summary

- Mirrors, like lenses, can be used to focus or spread out light.
- Just as with lenses, mirrors have a focal point where the photons either appear to converge to or emanate from.
- However, mirrors only have one focal point as opposed to two for lenses which is basically due to the fact that you can't shine light through a mirror.
- Like lenses, mirrors have a focal length and this focal length is measured from the surface of the mirror to the focal point.
- The signs for focal length are the same for lenses and mirrors, for converging focal lengths are positive and for diverging focal lengths are negative.
- The mirrors that we will deal with in this class are parts of spheres and will have a focal length that is one half of the radius of the associated sphere, which we were write mathematically as $f = \frac{r}{2}$

Homework Problem

Problem 22: Focal lengths of mirrors.

15. Ray Tracing

Ray Tracing

We now know what images are: apparent reproductions of objects. In the last chapter, we characterized images as **erect** and **inverted**, and discussed the idea of **magnification**. Our next goal is to determine where these images will be for a given optical system. To solve this, we will use a new type of problem solving called ray tracing. Ray tracing is a type of problem solving quite different than what you are used to for physics classes; this method of problem solving uses a lot more diagrams and, while equations will still be used, they will play a comparatively smaller role. These multiple problem solving approaches goes back to [Physics Goal #2: Representing physics Ideas in different ways]().

So what is ray tracing? If an object emits light, it emits light in all directions. If the object is visible by reflecting light from another source, your face is visible because it reflects light from the surroundings, that reflected light is diffuse and goes in all directions. Also keep in mind, that typical objects emit huge (10^{30}) numbers of photons. Thus, there are effectively photons going in every conceivable direction as shown in Figure 1.

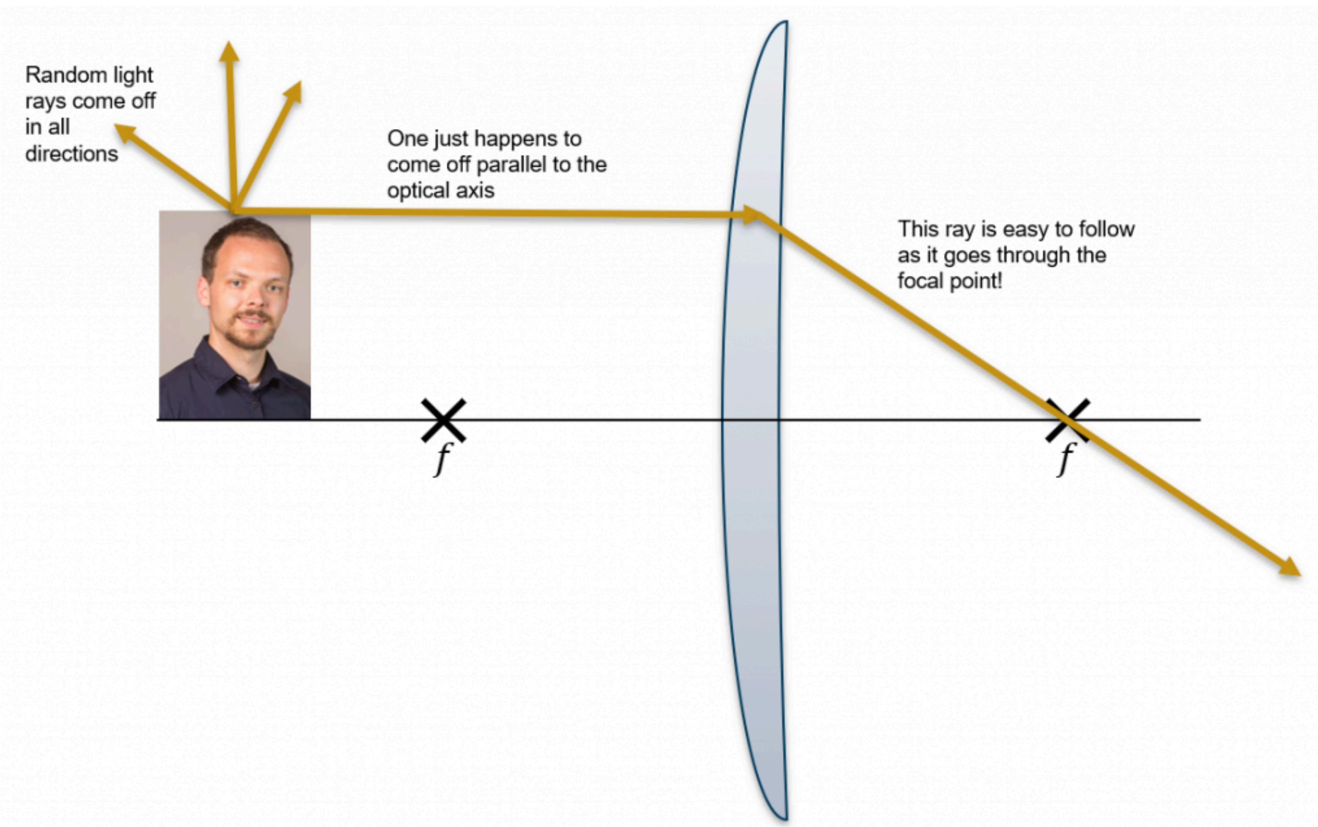


Figure 1: *While rays come off in all directions, we follow the rays which are easy – like one that goes through the lens parallel to the optical axis and then through the focal point.*

Ray tracing allows us to follow very specific photons: photons which will easy to follow because of the paths they take. For example, we know from the last chapter, that a photon that enters a converging lens parallel to the optical axis will go through the focal point as shown in Figure 2 below. Since there are so many photons, leaving your face, one will go parallel to the optical axis and then through the lens and towards the focal point as shown in Figure 1.

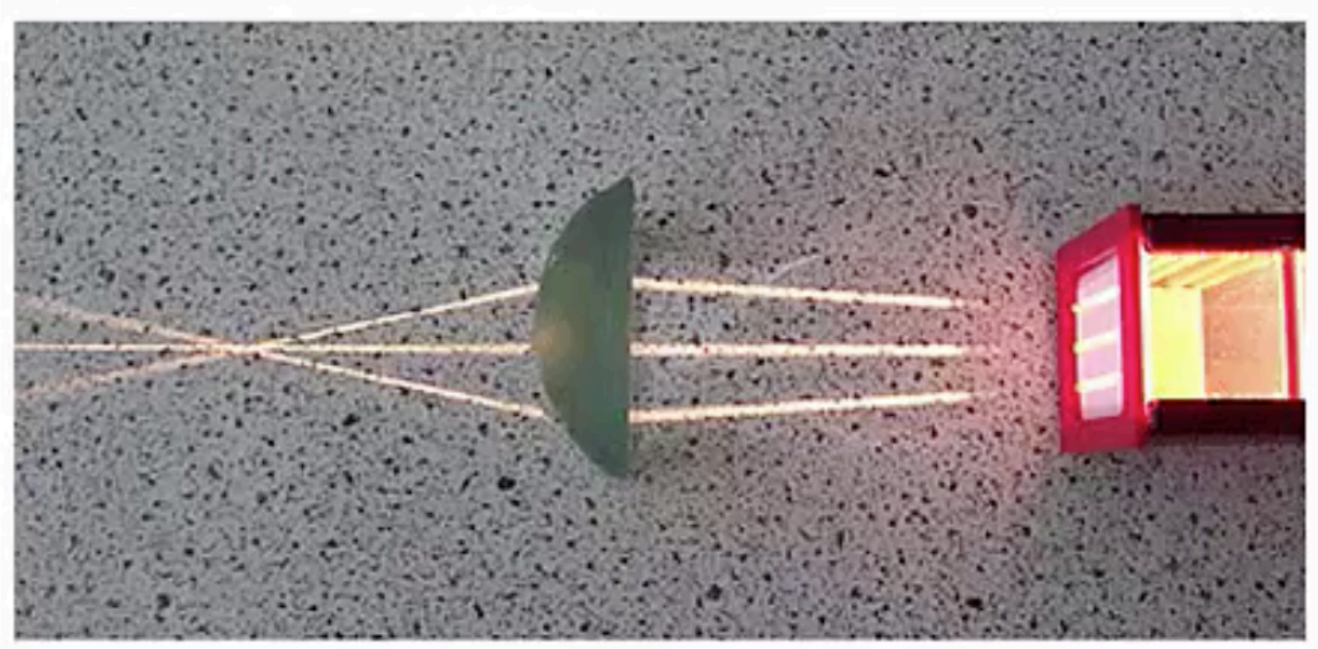


Figure 2: *Incoming parallel rays converge to the focal point of a converging lens.*

Throughout the next few sections, we will go through the particular rays to follow for the different **optical elements**:

- **Convex mirrors**
- **Concave mirrors**
- **Converging lenses**
- **Diverging lenses**

For each optical element, there will be three rays to follow for any object a finite distance away (if the object is infinitely far away, then the rays come in parallel and we saw what happens with parallel rays in the previous chapter). The three rays to follow are:

Our three rays that we will follow

1. A ray that comes in parallel to the optical axis leaves *using* a focal point. If the optical element is **converging** (like a concave mirror or convex lens) then use a focal point that brings the ray towards the optical axis.
2. A ray that aims for the center of a lens or mirror will go straight.

- The center of a lens is the middle: this ray will travel un-deflected as though the lens was not there.
- The center of a mirror is the geometric center: this ray will hit the mirror at a 90° angle and bounce straight back.

3. A ray that comes in *using* a focal point will exit parallel to the optical axis.

Instructor's Note

Your quiz will cover:

- For a given ray, you need to be able to determine where it will go for each of these basic optical elements

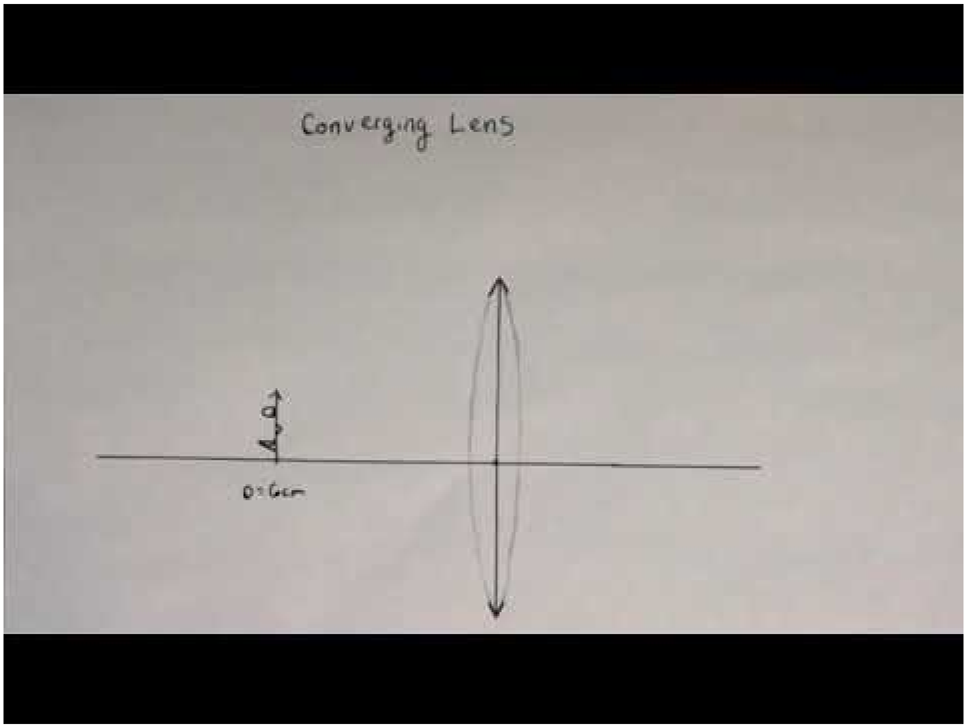
We will NOT expect you to be able to interpret the results, do calculations, or consider multiple elements. We will do that in class.

Ray Tracing for Converging Lenses

University of Massachusetts Amherst BE REVOLUTIONARY™

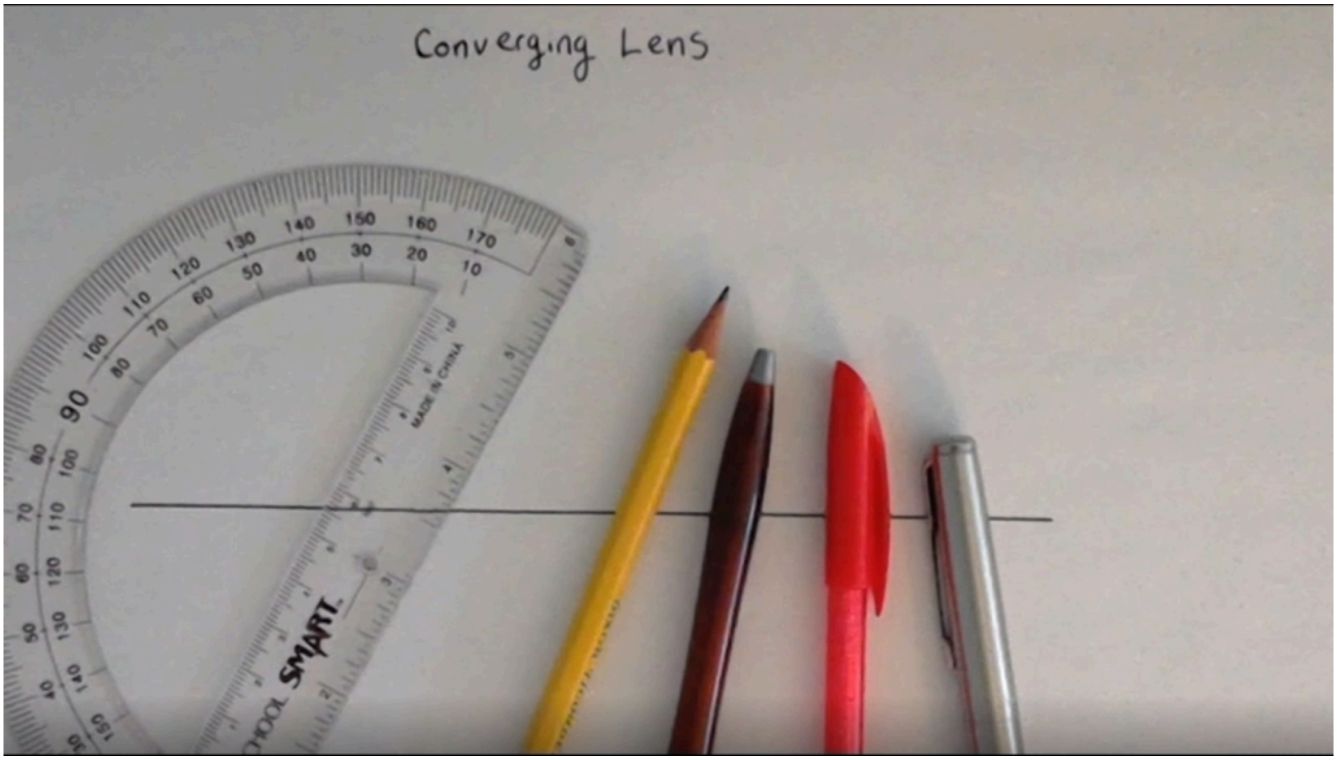
Instructor's Note

All of the following sections are presented both as video and with a text-based transcript. Normally, I feel that the two presentations are equivalent. In this particular case, however, I feel that most students will learn more by watching the videos *particularly if you follow along on your own sheet of paper.*

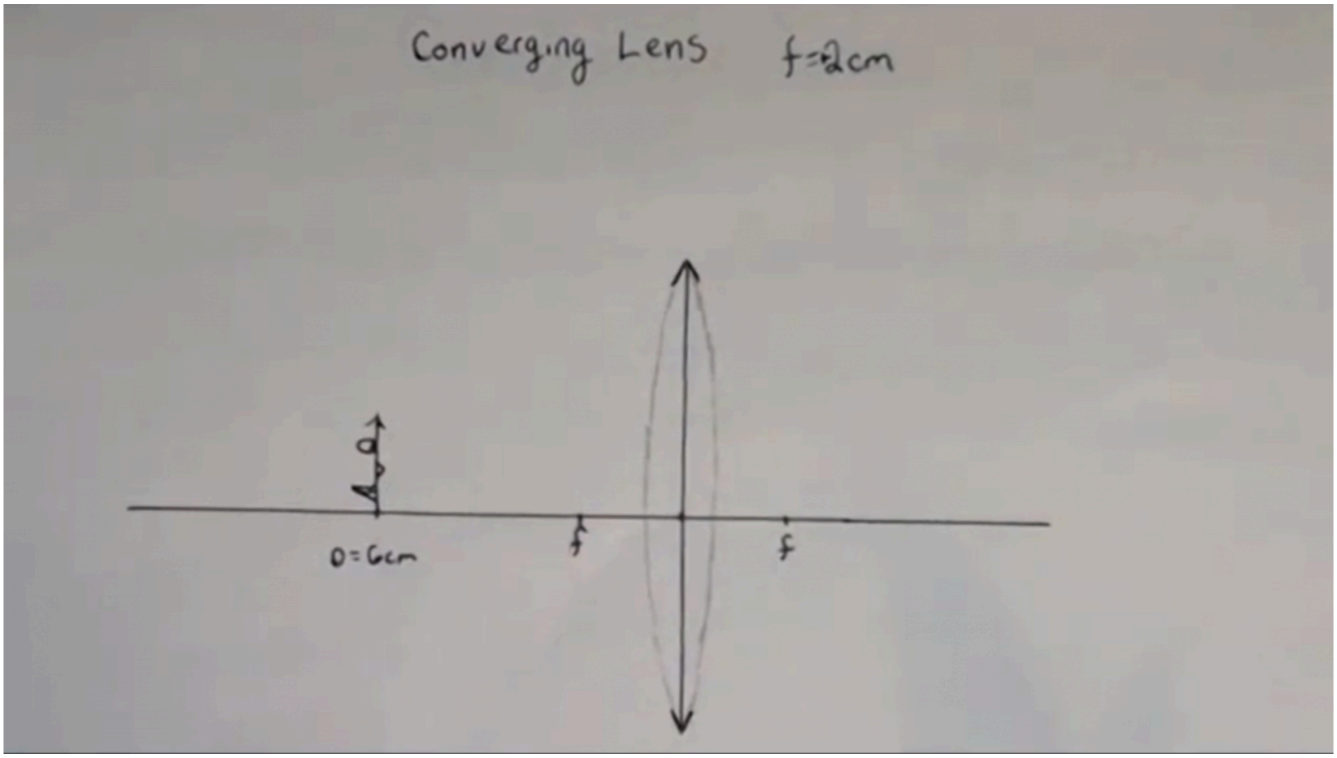


A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=501

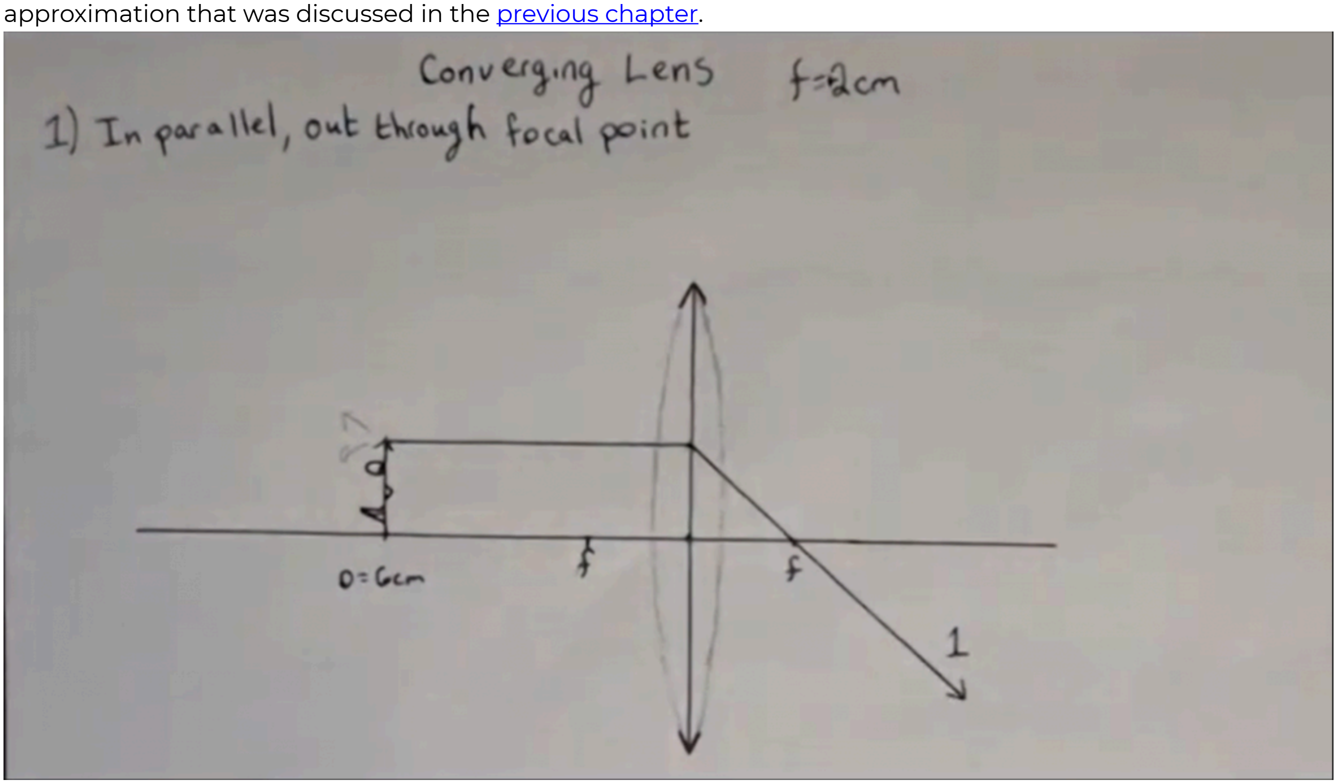
In this section, we are going to explore how to draw ray diagrams for the first of our four types of optical elements: the converging lens. Now, these videos use pictures of paper drawings to really show you the mechanics of how to draw these things out by hand. You will be expected to draw ray diagrams on an exam. To draw ray diagrams, it really helps to have a few things: one it helps to have pens in a couple of different colors: black, red, and blue. I also like to have a pencil which gives me four colors. You also really need a protractor, which was labeled on the syllabus as one of the things you need for this course.



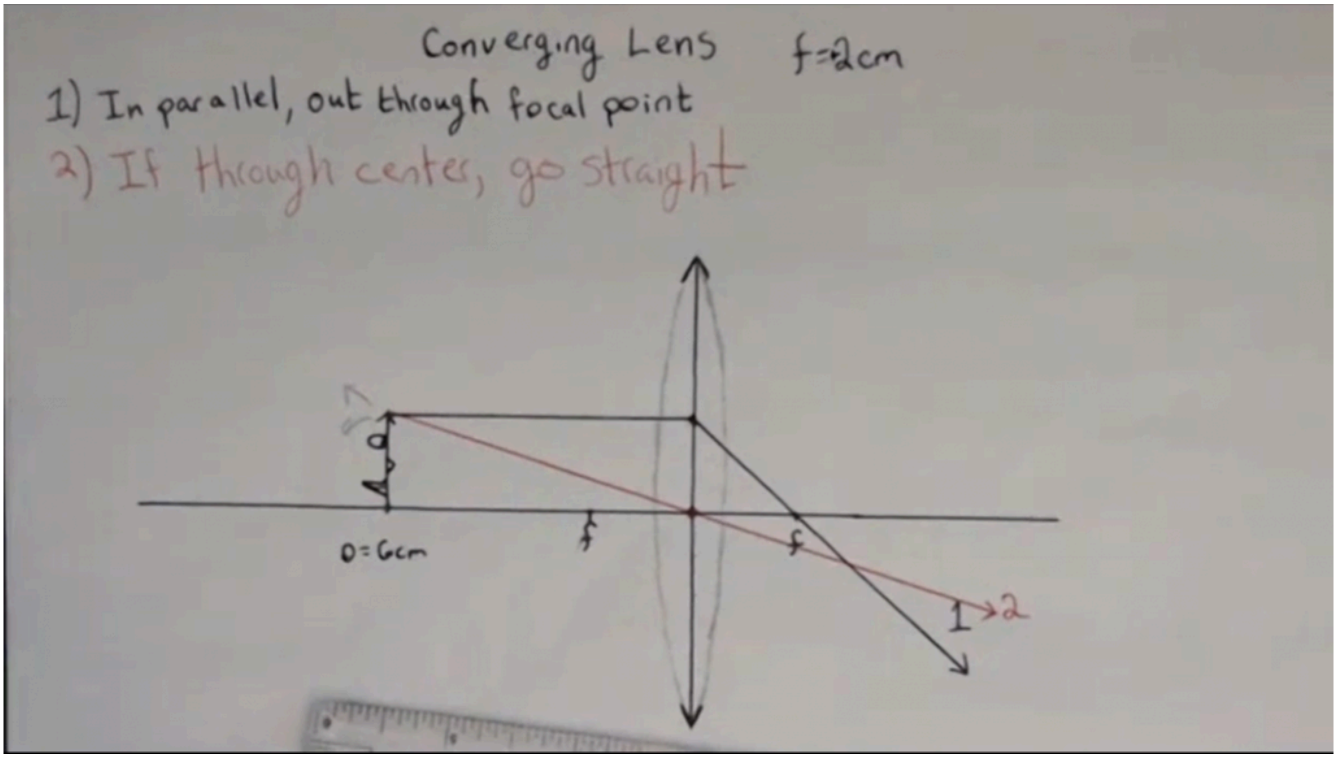
Get started drawing a ray diagram for converging lens, by drawing an optical axis. Now, put your lens on this axis somewhere, kind of in the middle to give yourself some room to work. Draw the center of the lens first, using your protractor to make sure it is perpendicular to your optical axis and that all your lines are straight. Trying to ray diagrams freehand will NOT work! You need your lines to be straight you need your angles to be precise there's no way you'll ever get these diagrams to work properly if you're trying to draw them freehand. Since I'm not an artist, and drawing the lens itself takes time, you'll often see me identify a converging lens just with a symbol that looks like a line with arrows
on both ends. These arrows are meant to indicate that the lens gets thicker towards the middle than it does towards the ends. You'll see when we do the diverging lens, we'll use a different symbol. Now put an object about six centimeters away from my lens. Your object might be a little face or something like that. Again I'm no artist, often you'll see people put a little arrow just to help identify which way is up. We also need to know the focal lengths for our lens: let's put our focal length for this example at two centimeters. It's going to be a positive focal length because our lens is converging. Measure out your two focal points and label them with f. Now, we have a set-up, and at this point we can go through and actually start drawing our rays.



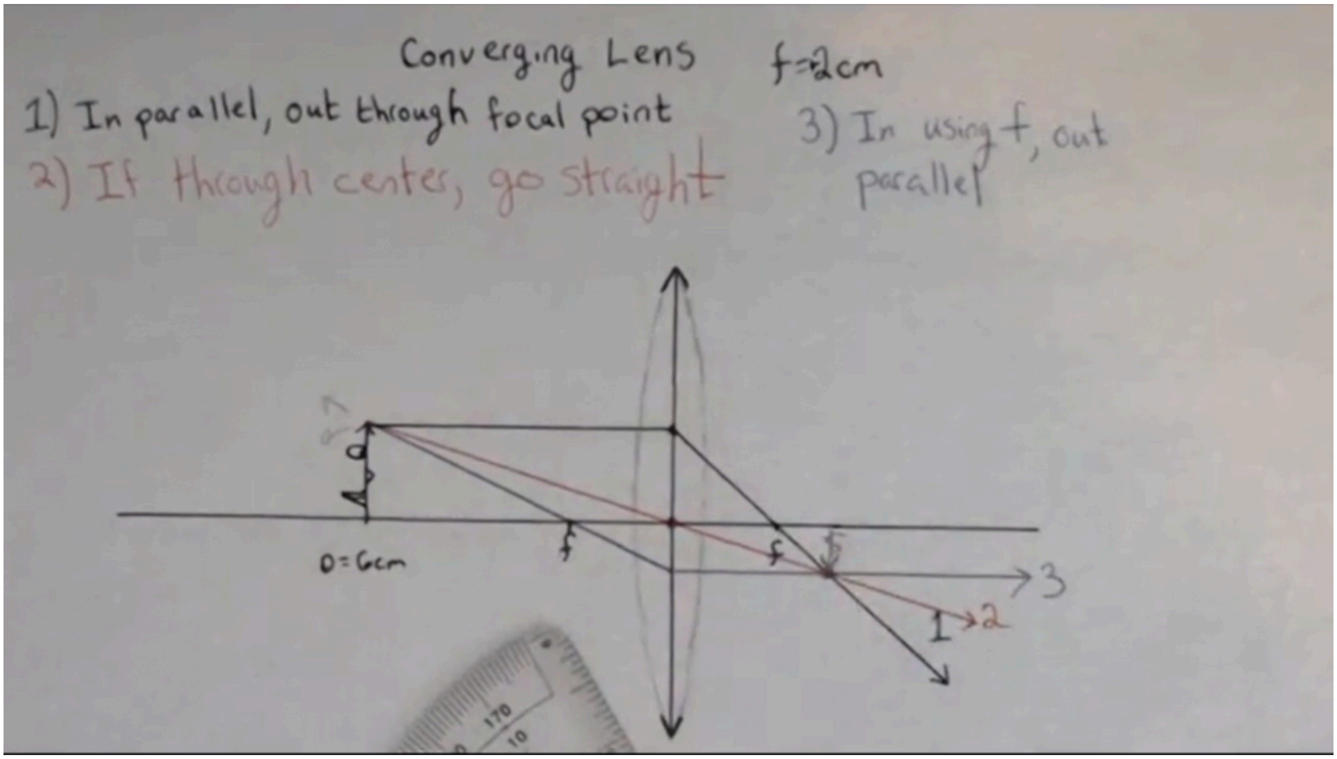
Ray number one is in parallel out through focal point. Recall, light rays are going in all directions off the top of this object. One of these is a ray going off towards the lens parallel to the optical axis. When this ray hits the lens, it's going to go out through the focal point because that's what a focal point is for a converging lens: converging lenses convert incoming light that's parallel and bend them to their focal points. You'll notice we're pretending that all the bending happens here at the center of the lens and that's due to the thin lens approximation that was discussed in the previous chapter.



The second ray is the rule that if a ray through the center goes straight. In the case of a lens, the center is where the lens meets the optical axis: by definition the optical axis goes through the center of the lens (It may not look like it in my drawings because I'm not much of an artist!). So ray number two goes straight through the middle.



Ray number three, remember, is in using focal point out parallel. Now in this case, we've already used the focal point on the far side of the lens from the object. Therefore, this time we're going to come in using the focal point we haven't used yet, the one on the same side as the object. Out of all the infinite numbers of rays, we are going to follow the one which just happens to come in towards the lens going for this near focal point. This ray is going to go out parallel to our optical axis.



You can see that in this particular case, all of our ray's happen to converge on the far side of the lens. As we'll discuss in class, that is going to be the location of our image. We'll talk more about that in class. What you need to know right now is how to draw these three rays:

1. In parallel, out using the focal point.
2. Straight through the center.
3. In through the focal point, and out parallel

Simulation

Below is a flash simulation (you may need to click and allow flash) that shows the paths of rays through a converging lens. A few things to explore:

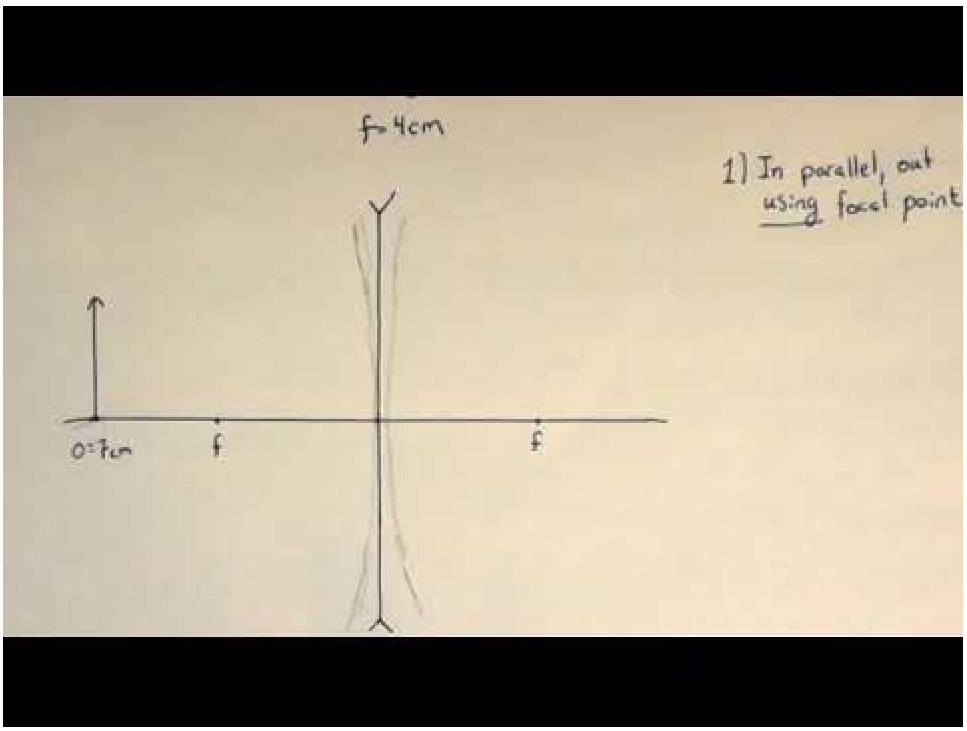
- Click "Principal Rays" to see the rays used in the ray diagrams discussed in the video.
- Click "Many Rays" to see the fact that there are a bunch of rays coming from the object that all converge to the point, not just the ones we saw in the ray diagram. You will see that some even miss the lens entirely and just go straight!

An interactive or media element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=501

Homework Problem

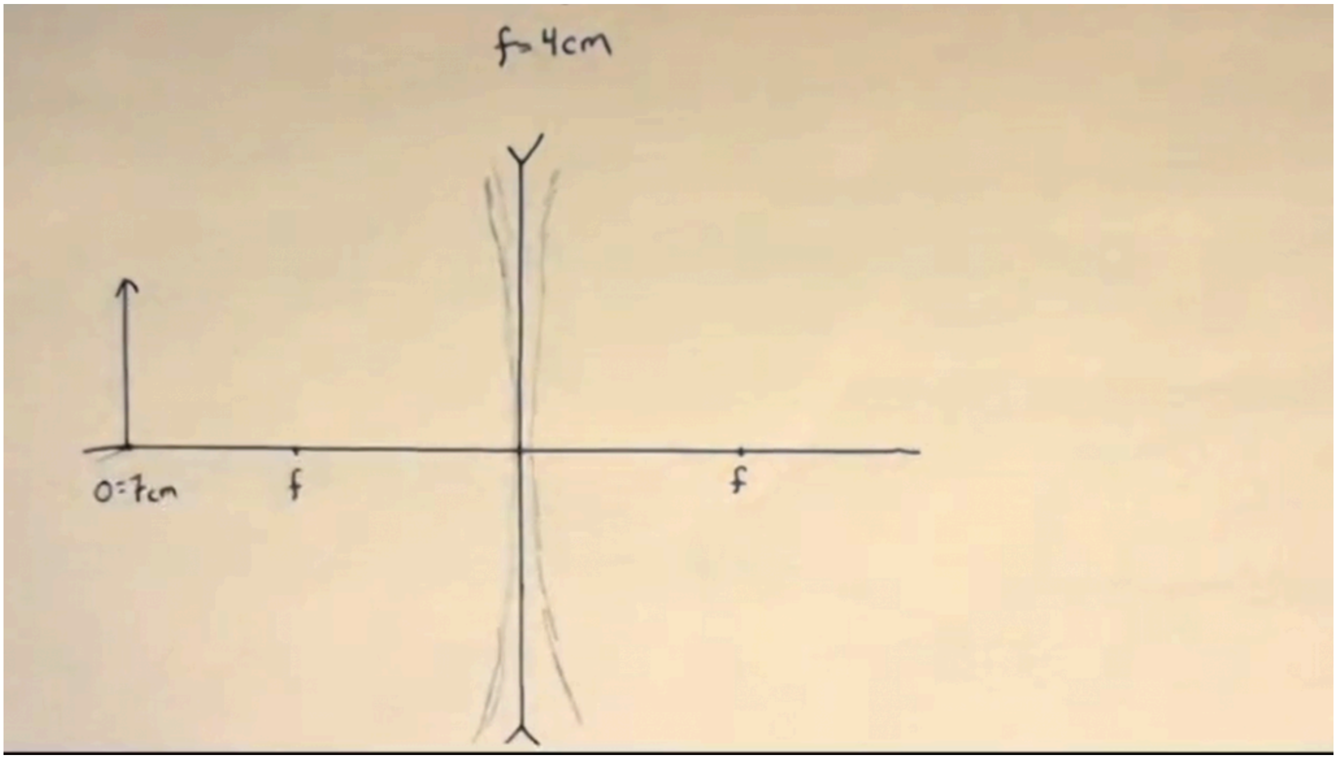
Problem 23: Drawing ray diagrams for converging lenses.

Ray tracing for Diverging Lenses

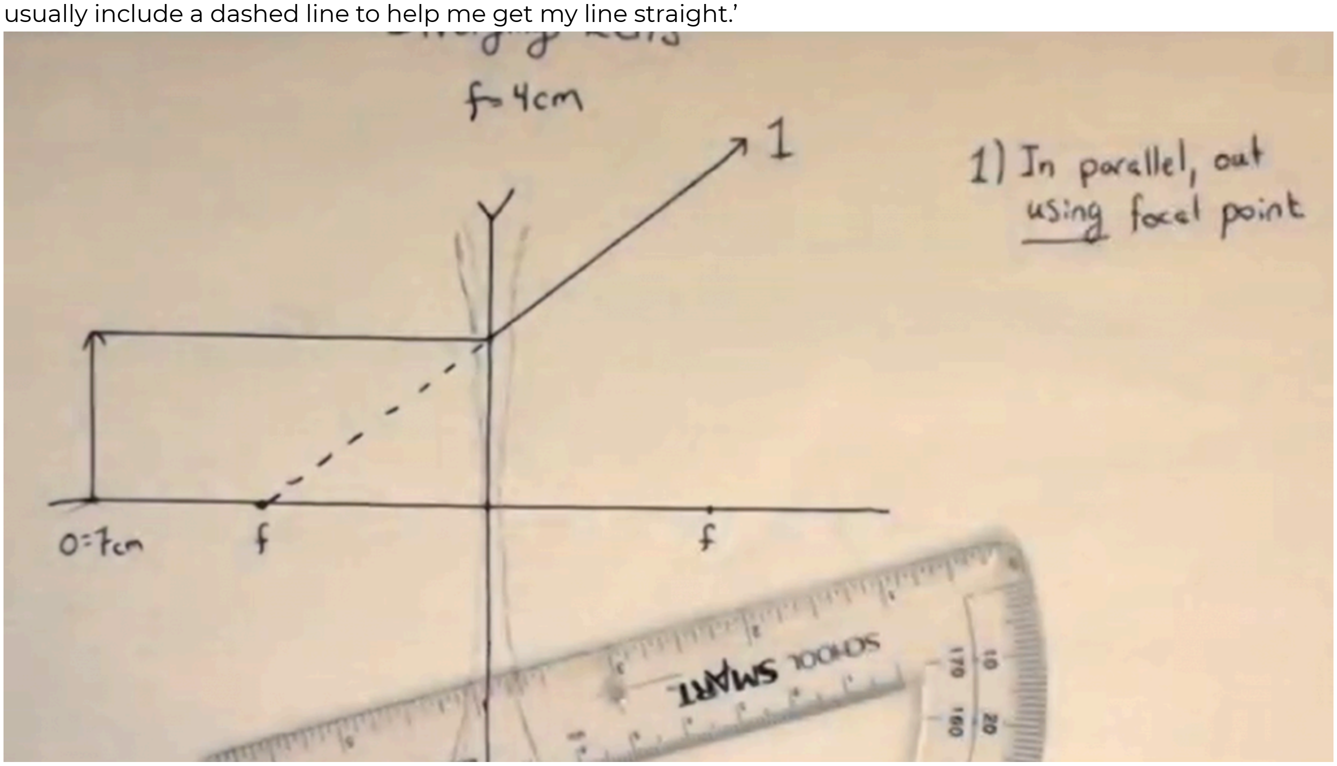


A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=501

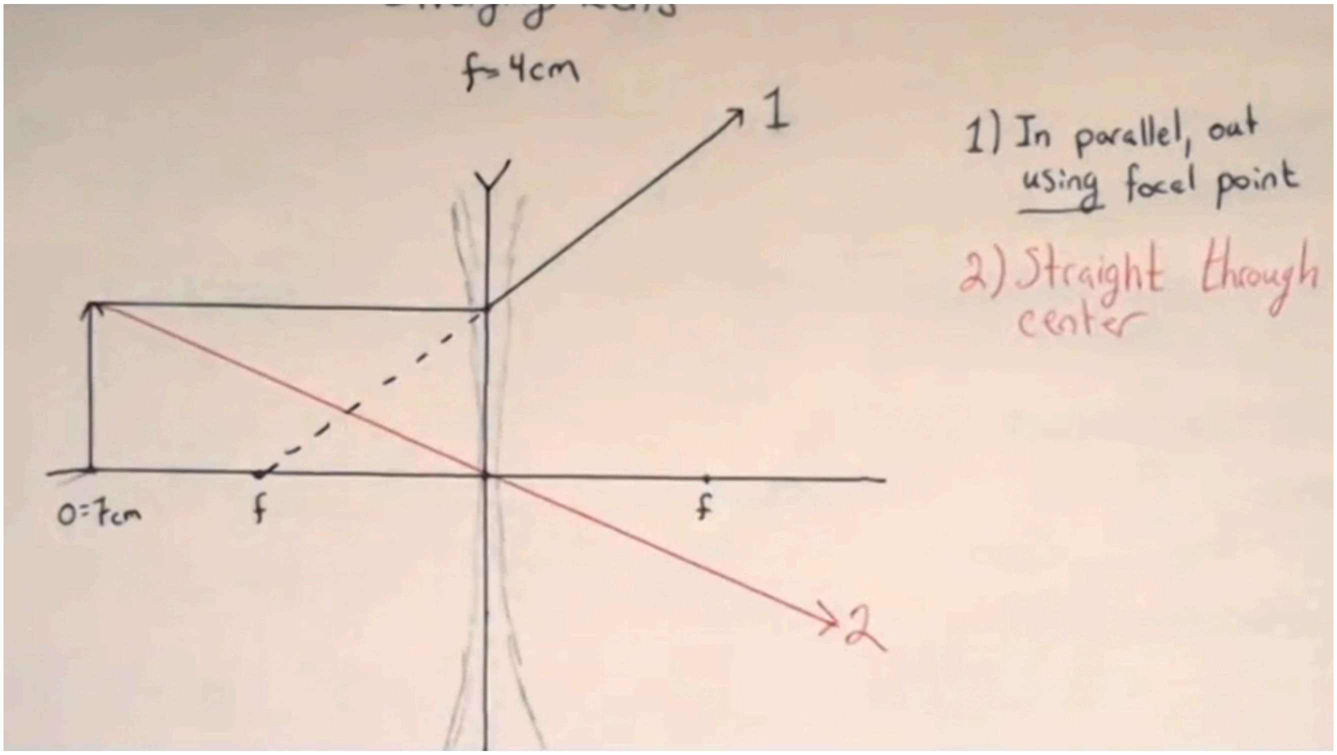
In this section, we will be drawing ray diagrams with diverging lenses. To get setup draw your optical axis and put your lens in the middle: make it pretty big. Always give yourself room to work when you're drawing these things. This time we're doing not a converging lens but a diverging lens. A diverging lens is, as you know, much thinner in the middle than it is on the outside. Because clearly my ability to draw diverging lens is even worse than my ability to draw converging lenses, I'll often use arrow heads pointing towards the center of the lens to indicate that it's thinner in the middle than it is on the outside. For this example, we're going to use a focal length of 4 centimeters: one focal point on each side labelled f. We will put our object 7 centimeters away. Give it a pretty good height: make it like three centimeters tall.



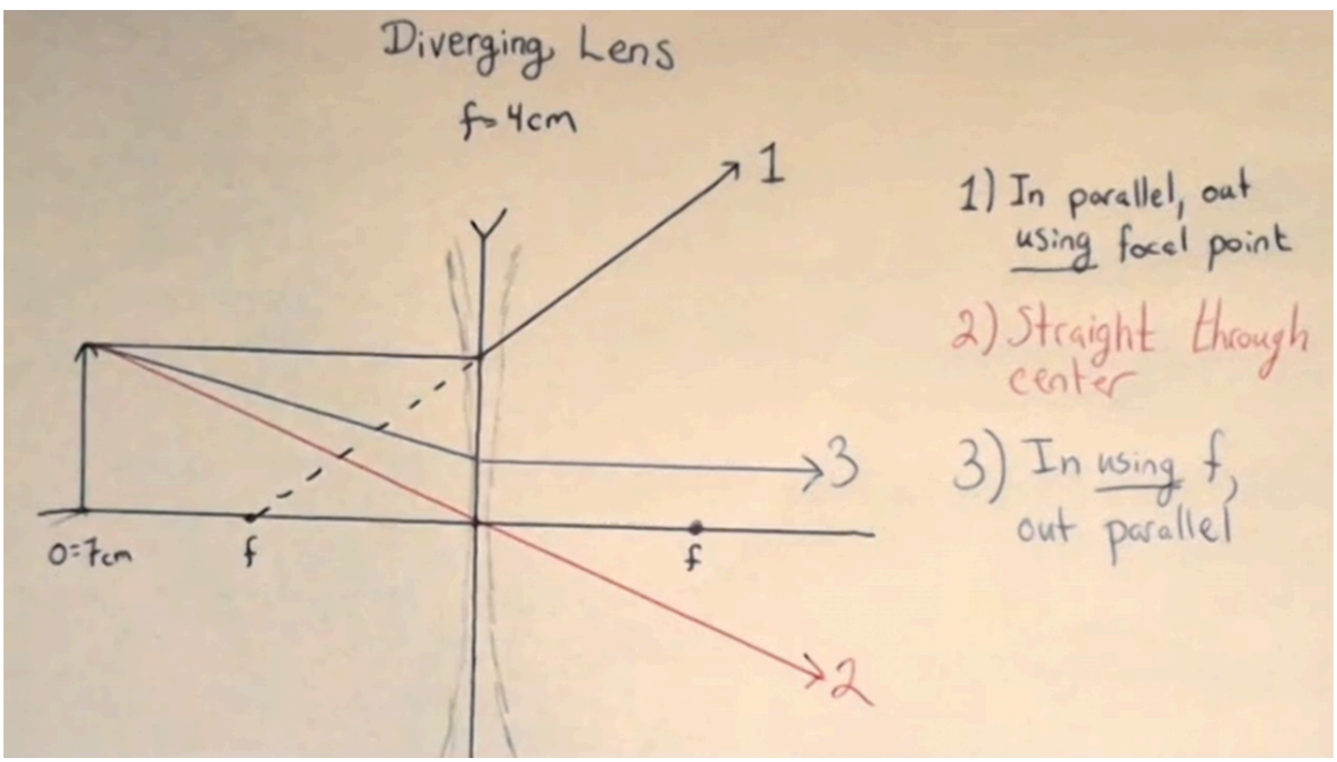
Now we once again go with our same process: ray number one is “in parallel out using a focal point.” You will notice notice I’m using the word *using* not the word *through* because the ray will not actually go through a focal point. The ray comes in parallel, use your protractor to make sure that the ray is parallel to the optical axis. Since this is a
diverging lens, it’s not going to bring the ray towards the optical axis; it’s going to cause the ray to diverge *away* from the optical axis as if the light had come from this focal point on the same side as the object. That is why I say not *through* but *using*. *T*he light is going to diverge as if it came from this focal point on the left. I usually include a dashed line to help me get my line straight.’



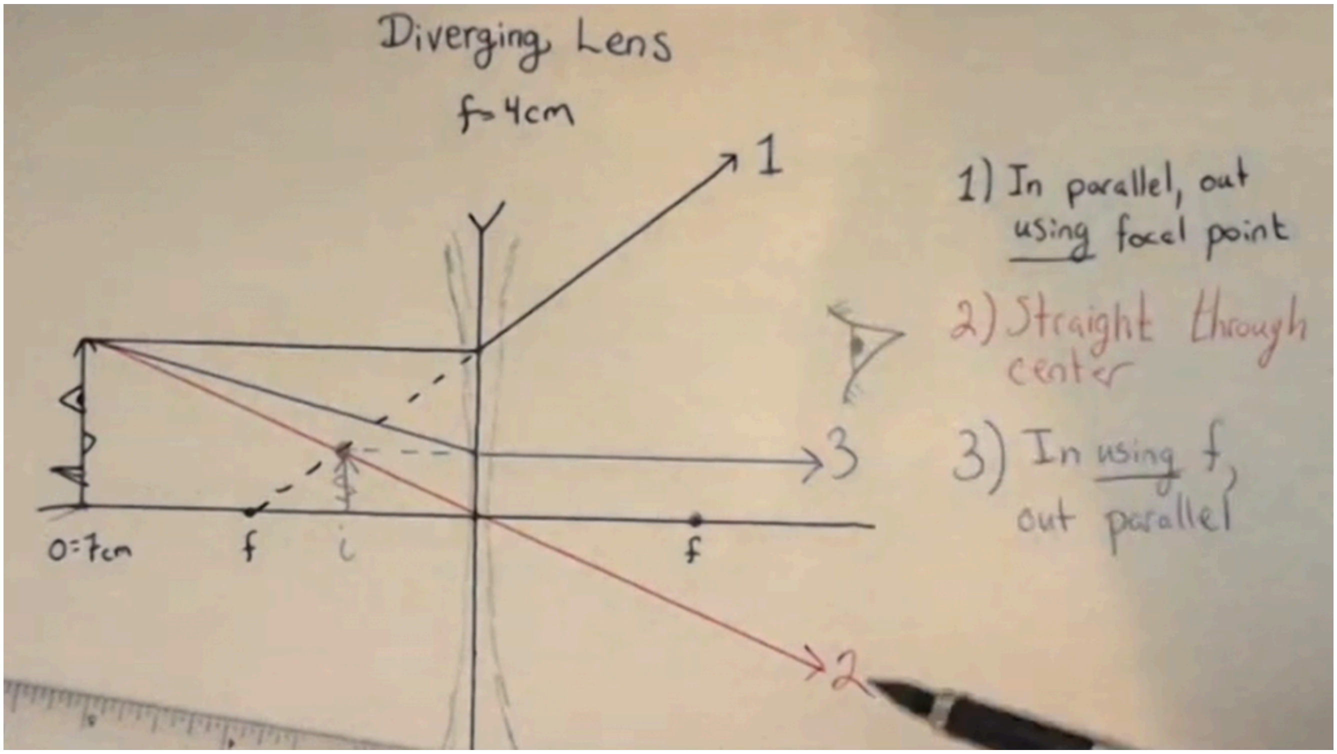
Ray number two, “straight through the middle,” which is pretty straightforward. The middle is where our optical axis meets our optical element.



Then we’ve got ray number three, “in using a focal point, out parallel.” Again, you’ll notice I’m using the word *using*. In this particular case, I’ve already used the left focal point, and I can’t use it again. Therefore, I have to use the other one on the far side from the object. So I’m gonna come in as if I were going for that focal point, but then I hit the lens and, instead of continuing on, I go out parallel to the optical axis. That is ray number 3.



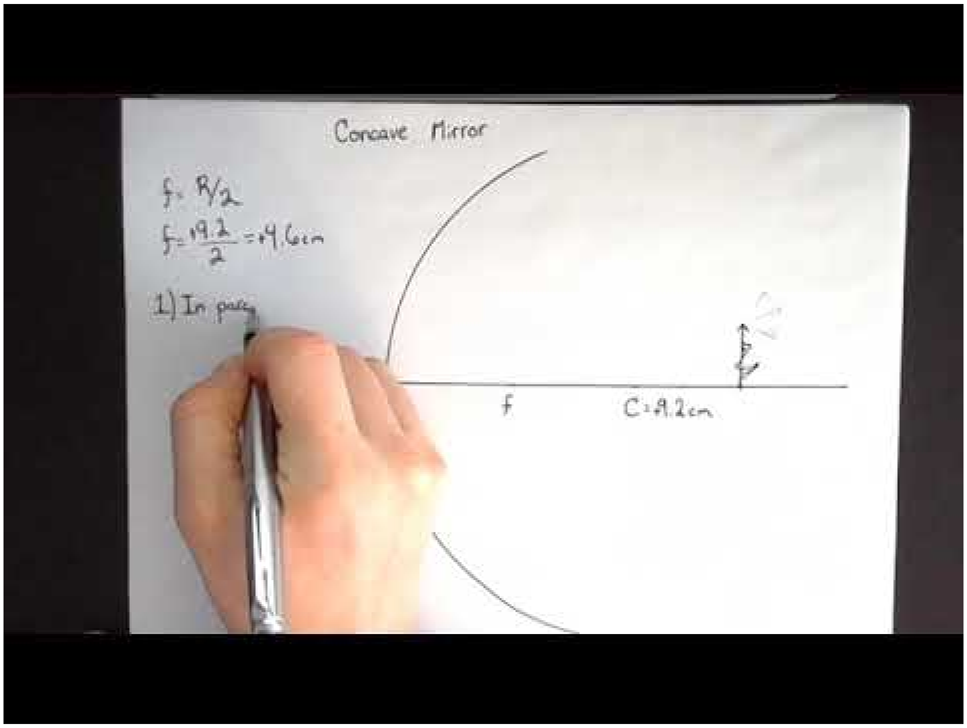
You'll notice these rays are spraying out, which they should. This is a diverging lens after all, the rays should diverge, and they do. Thus, they don't converge to a point anywhere like we did saw in the last example. However, what if your eye were over here on the right looking through the lens at the object, what would you see? Well your brain assumes that light travels in straight lines, because in most of your experience it does. So your brain is going to assume that all these light rays emanated from this point and traveled in a straight lines. Thus, we will have our image over here on the left, the same side as the object! We'll discuss this part more in class, but I thought I would just expose you to it while we're here. What you need to know right now is how to draw these three rays.



Homework Problem

Problem 24: Ray diagrams with diverging lenses.

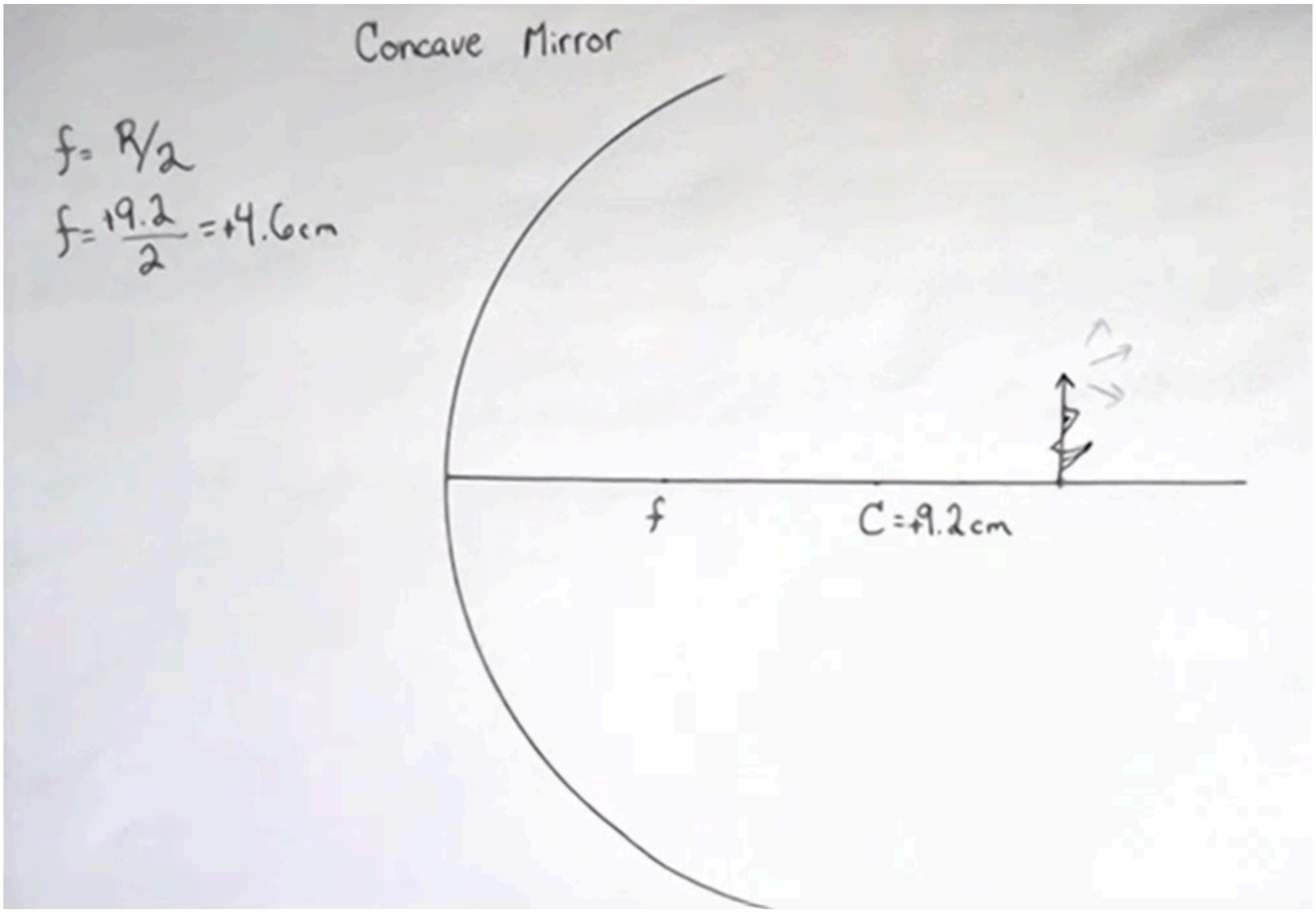
Ray Tracing for Concave Mirrors



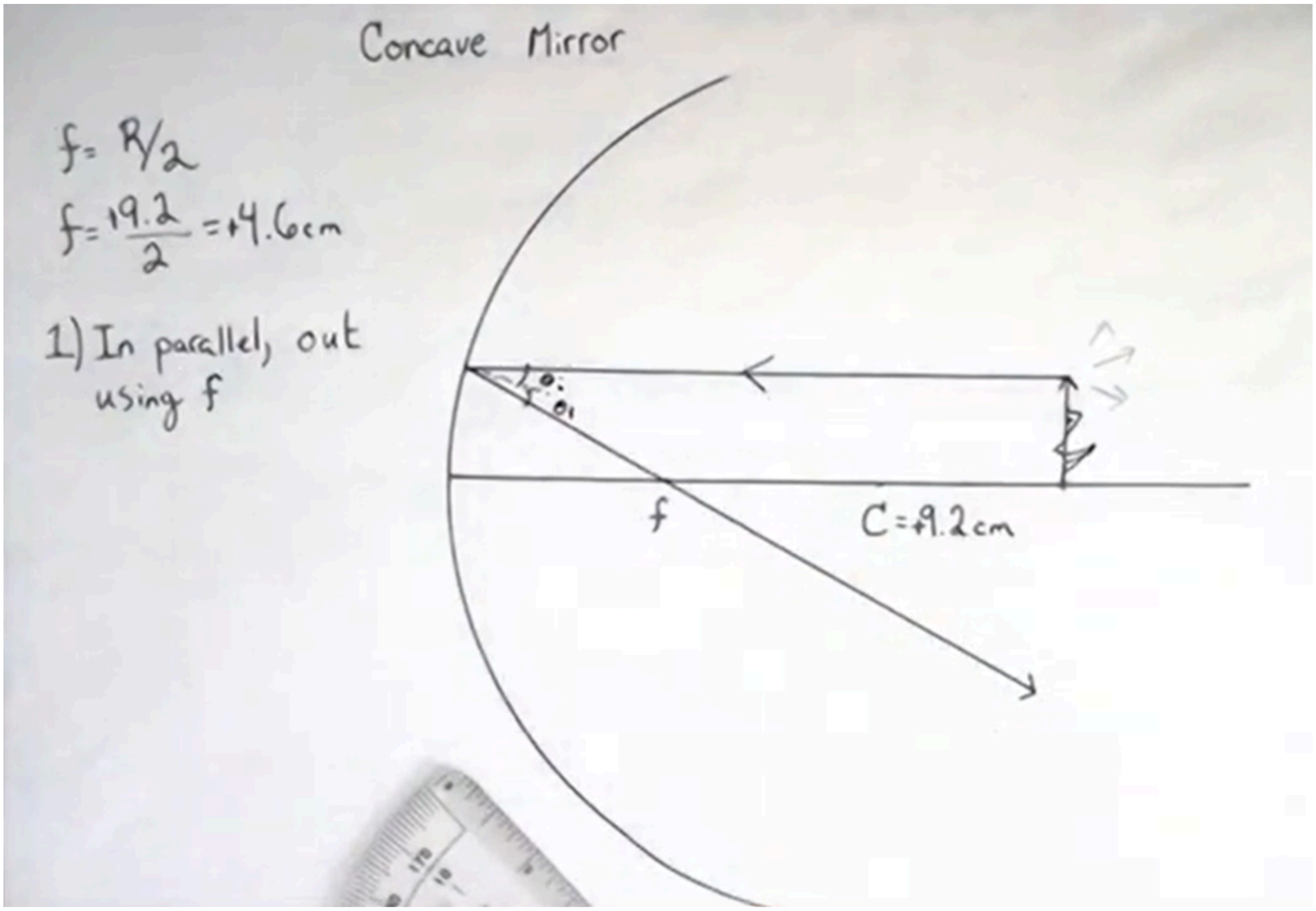
A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=501

In this section, I am going to show you how to draw the ray diagram for a concave mirror now. In this particular class, we're only interested in mirrors that are parts of circles or flat. In class, we will provide you a nice little drafting tool to help you draw mirrors that are circles but for the purposes of this video, I'm going to use a compass. We will begin by drawing a circular mirror using my compass. Note that for mirrors, the **center** is the geometric center of the circle. For lenses, in contrast, the center is the middle part so it's a slight difference in terminology. Now I will add my optical axis which connects the center of the mirror and goes away from it (you're going to really need a protractor to do all of this stuff!). Next, I need to measure how big my circle actually is. I can see from measuring it that my center is 9.2 cm away from the vertex, which by the sign conventions discussed in the [last chapter](), since this is a concave mirror, we're going to think of our radius as being *positive*.

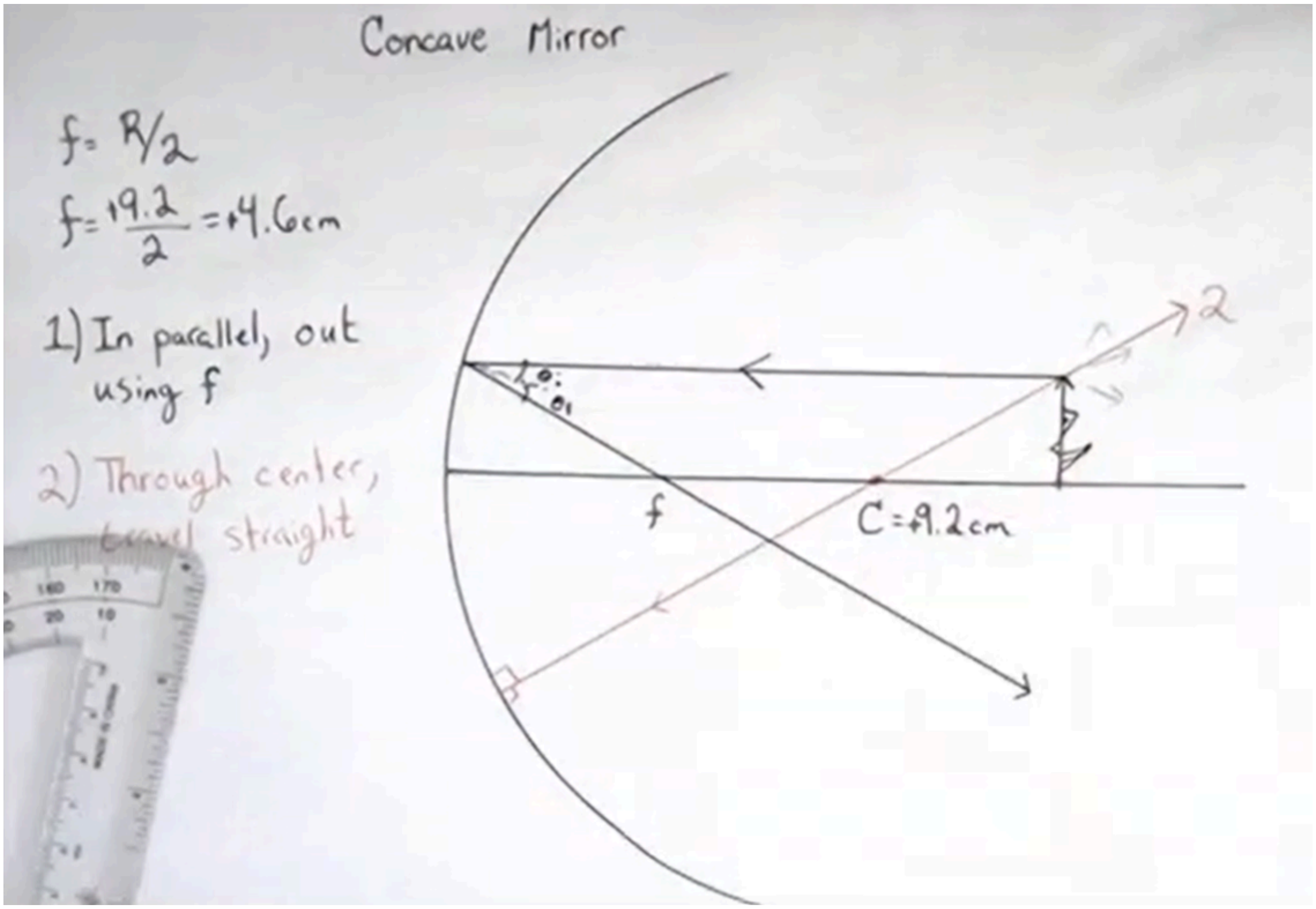
Now, a fundamental property of all spherical mirrors is that the focal length is half of the radius $f = R/2$. That is generally true for all spherical mirrors. Therefore, in this case, the focal length is going to be 4.6cm. I can, therefore, measure 4.6 cm away from the vertex and mark a little point which will my focal point. Now, all we need is an object. In this example, we're going to place our object nice and and far outside our center. It will be outside the center of curvature. Let's make an object that's gonna have some some height to it. I'm actually gonna make it as tall as my protractor ruler for a reason that'll become hopefully apparent in a moment. As usual, we like to put a little arrow on our object so we know which way is up but you can think of it as being a face, or a tree, or a candle, or whatever you want. Arrows are just easy to draw. Recall, there are photons coming off of this object in all directions and we're just going to choose the three photons that are easiest for us to follow.



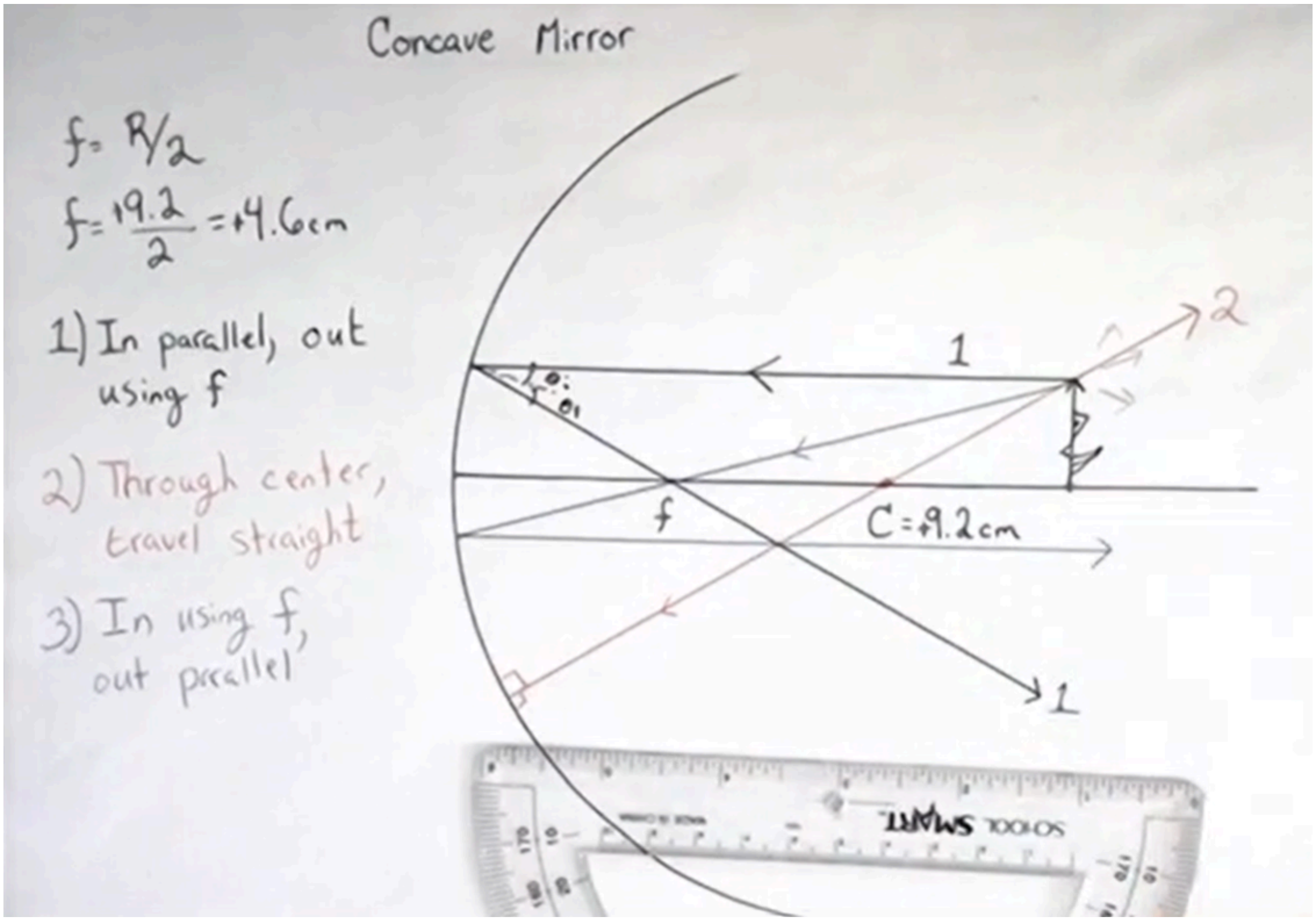
The nice thing about phrasing the ray diagram rules the way I have at the beginning of this chapter is the same three rays that we've been doing for lenses still work: we just have to think about them a little bit differently. The first ray, just as for lenses, is "in parallel, out using the focal point." We are therefore, going to draw our first ray coming in parallel (now you can see why I've made my object the exact same thickness as my protractor – it makes drawing a parallel ray pretty easy) and it's going to then go out through my focal point. You can see why by zooming in over here, our law of reflection: the incoming angle and the outgoing angle are the same.



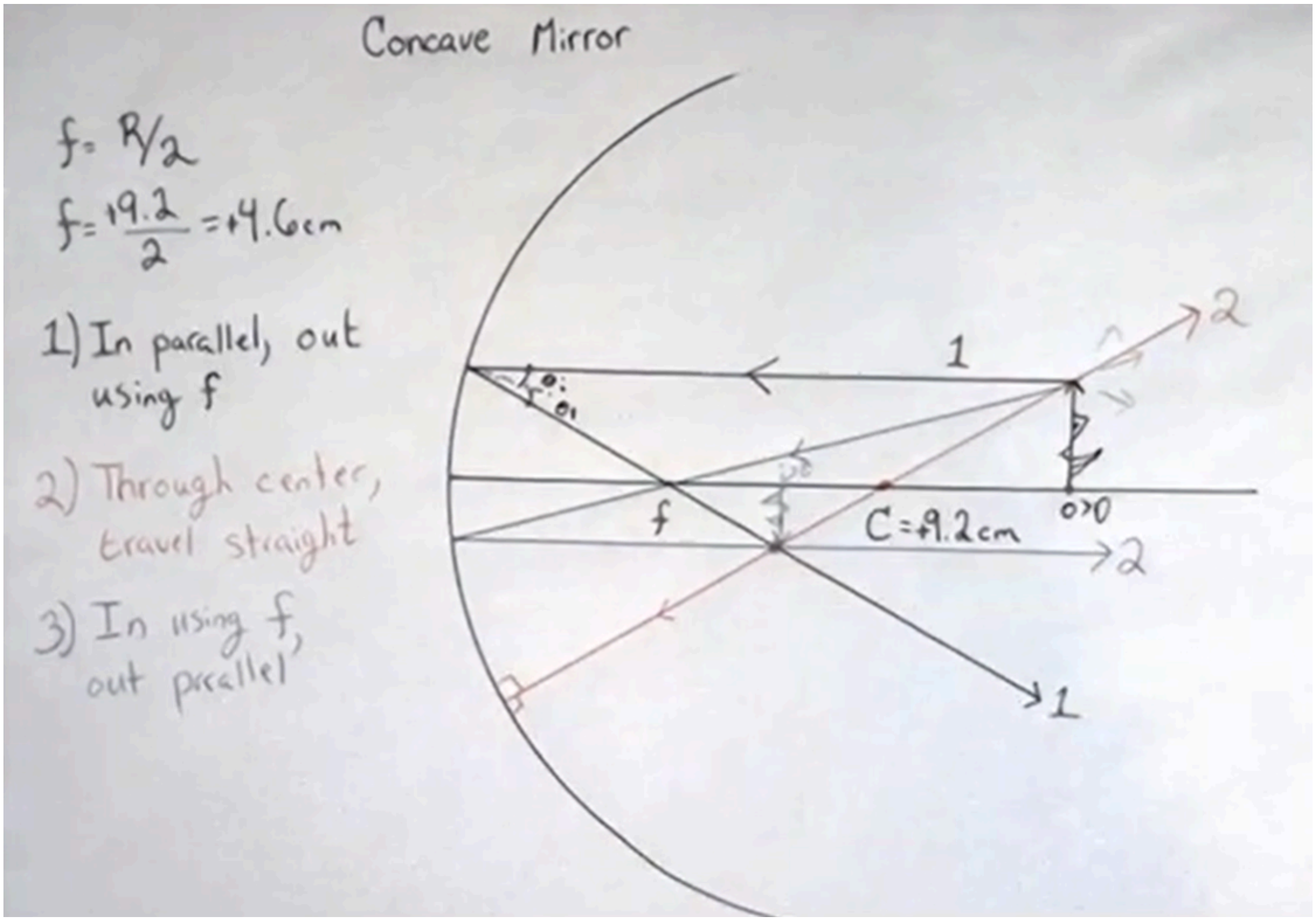
Ray number two: “rays that go through the center travel straight.” Notice again that the phrasing of how we describe our rays is exactly the same for mirrors and lenses. In this case, however, the center is the center of the circle, not physically on the mirror. That point, where the mirror meets the optical axis, is called the vertex. A ray that goes through the center will travel in a straight line: that ray will come in hit the mirror and end up bouncing to travel straight back the way it came. You can see that such a ray meets the mirror at 90 degrees so it bounces straight in-and-out.



Finally, ray number three again follow the same rules as the lens: “in using the focal point out parallel.” Now, lenses have two focal points, one on each side, but mirrors only only have the one. Thus, we don’t have to worry about which focal point to use: mirror only has one, so that makes it maybe a little bit easier. Our ray is going to come in using our focal point and then going to go out parallel to our optical axis.



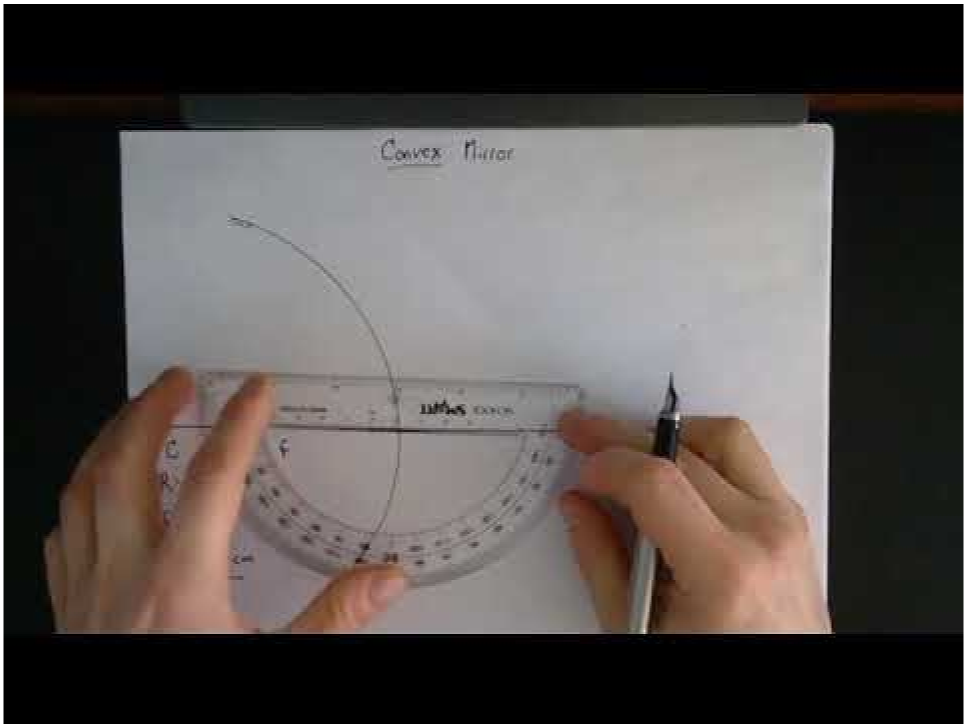
You will see that all of these rays actually converge at this point right in front of the mirror. That point is where our image is going to be. Thus, we are
going to have an image distance that's positive as it is on the side of the outgoing light. The thing with mirrors is that the incoming side and the outgoing side are the same side of the mirror. We are therefore going to have a positive image distance and a positive object distance. Moreover, our image is going to be in the previous chapter when we discussed looking in the inside of a spoon at some length. We mentioned that the image appears to hover in front of the spoon a little bit: here we see exactly where that image is. We will talk more in class about why this is the image, how to characterize that image, and some other aspects of interpreting this diagram. What I really need you to know for right now is how to draw these three rays.



Homework Problem

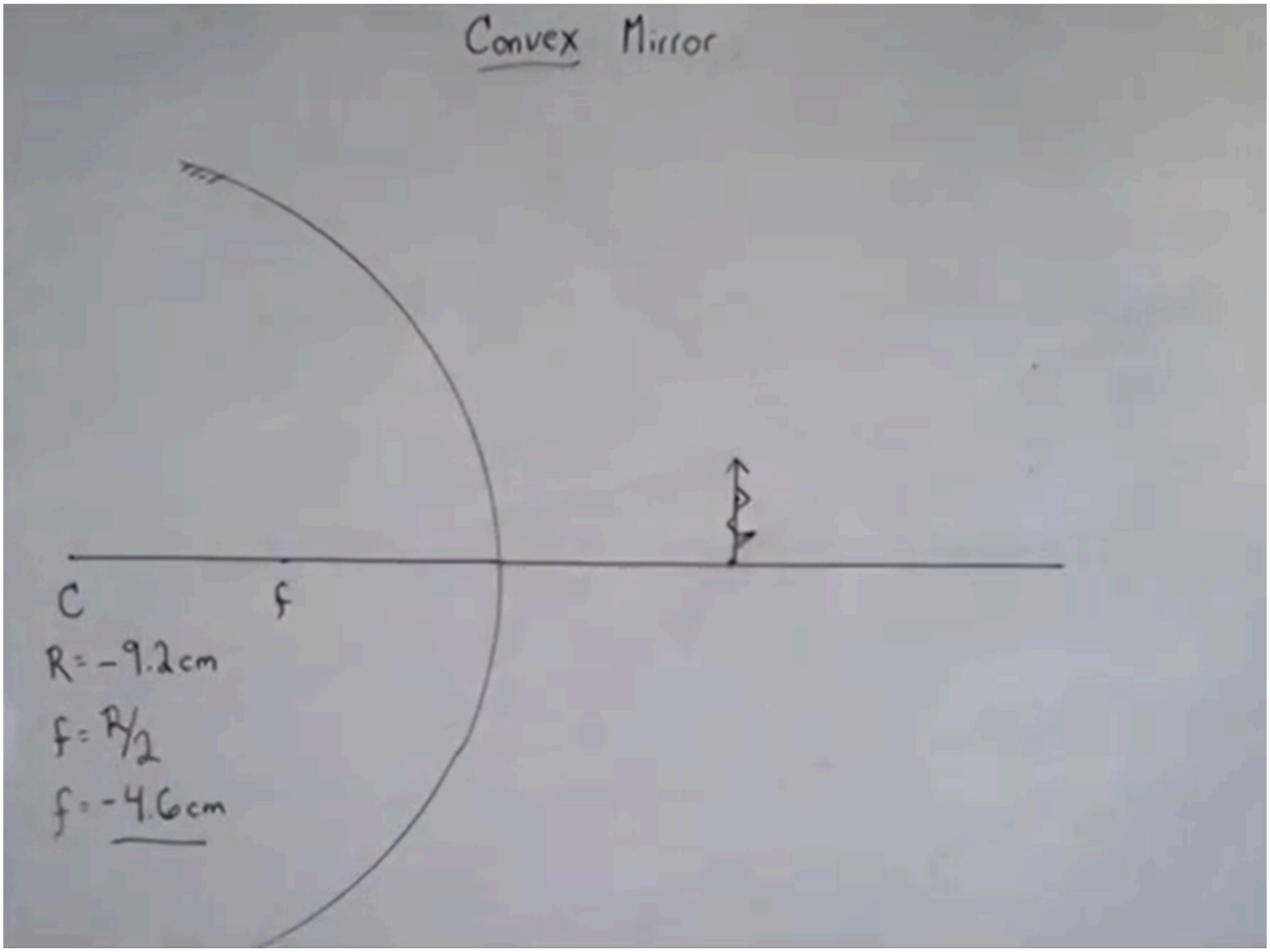
Problem 25: Ray diagrams with concave mirrors.

Ray Tracing for Convex Mirrors

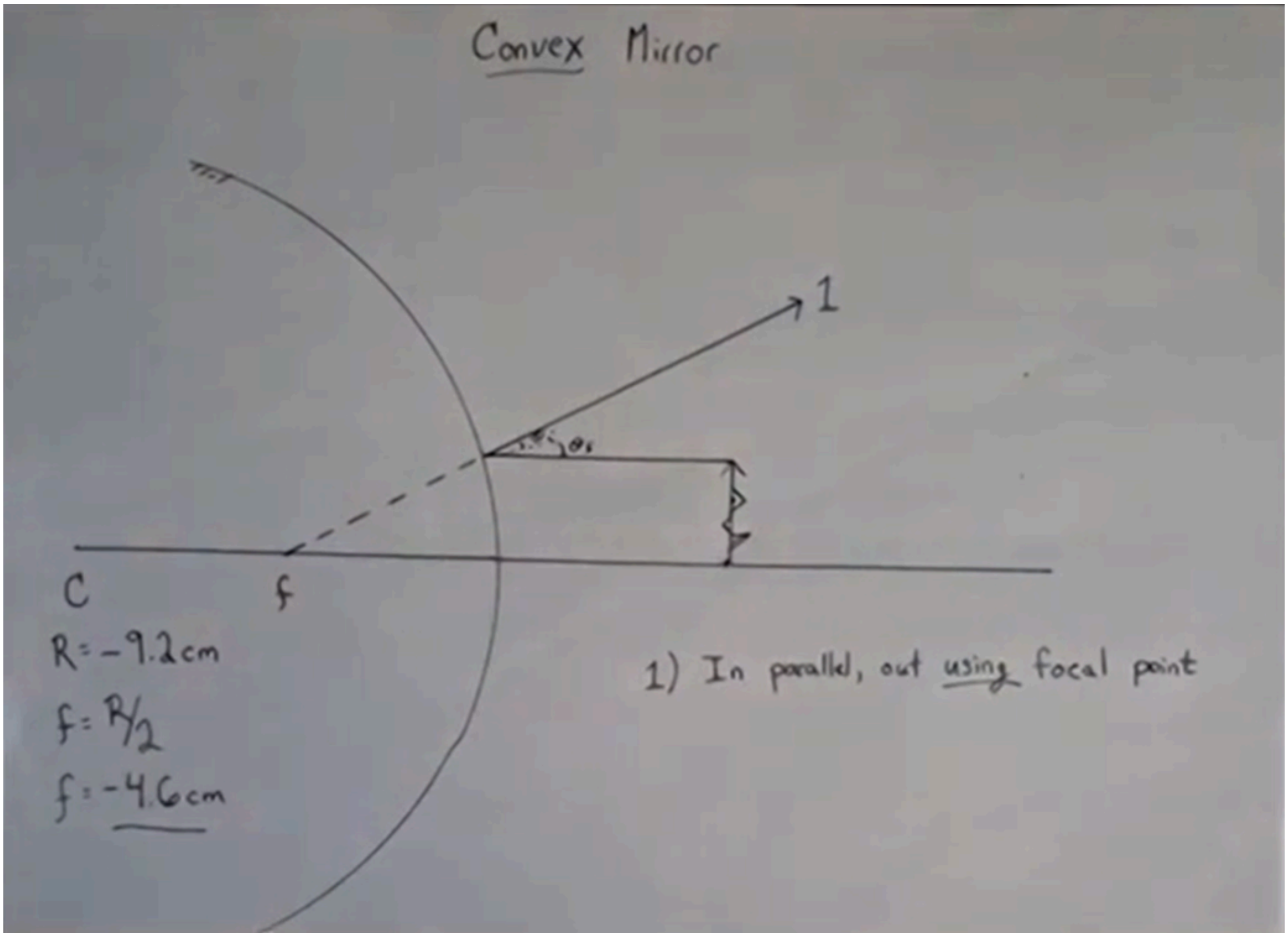


A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=501

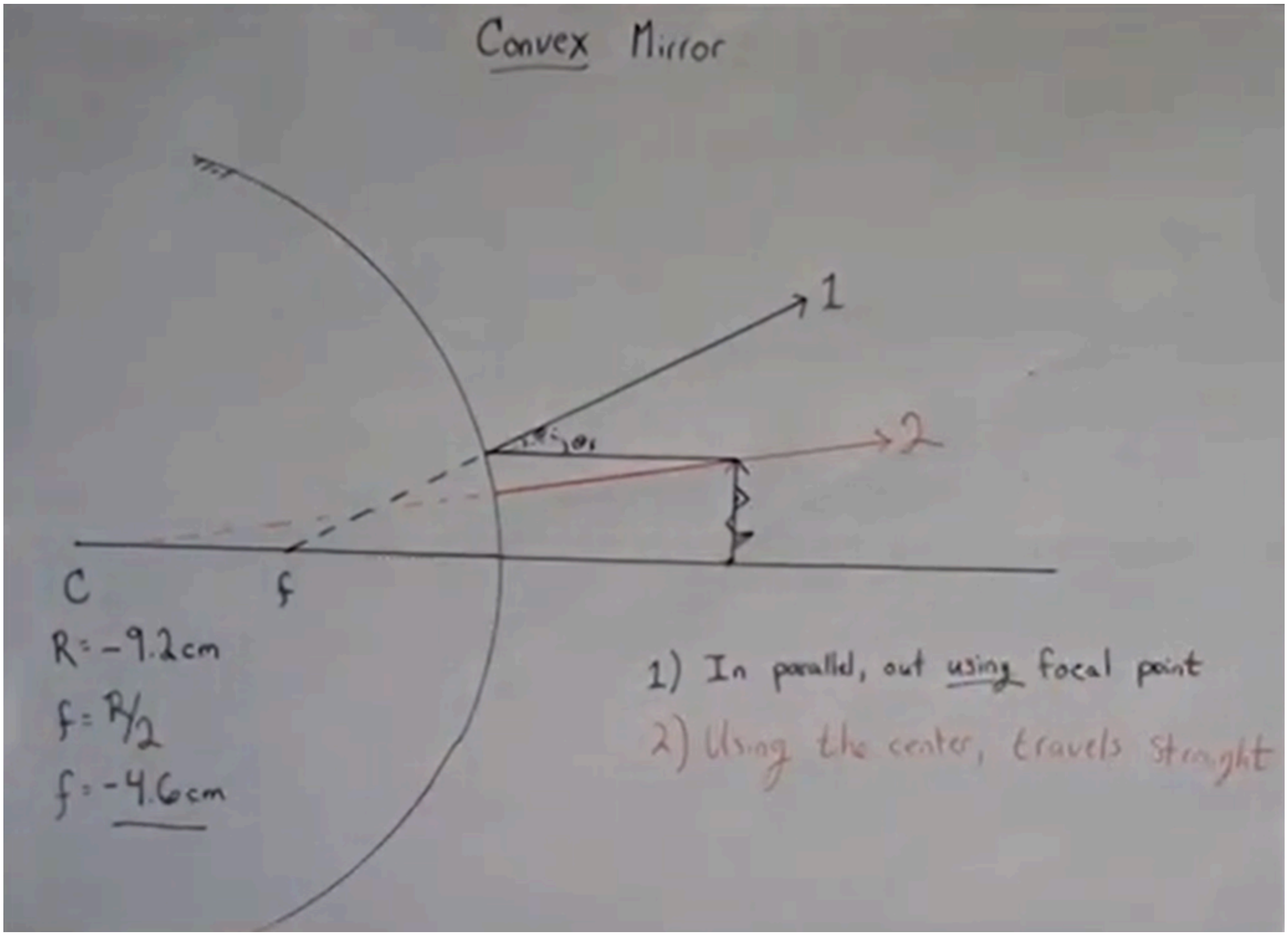
Now for the last optical element, where we draw the ray diagram for a convex mirror. Once again I'm going to use my compass to draw the spherical mirror to begin. In class, I'll give you a tool to help with this but here I'm going to use a compass to help me draw a nice circle. Again, I use the center of the circle to define my optical axis. My circle has a radius of 9.2cm, but since it's a convex mirror our sign conventions say that the radius should be negative: $R = -9.2\,\mathrm{cm}$. Just as with a concave mirror, the focal length is half the radius $f = R/2$. Therefore, the focal length is going to be negative: $f = -4.6\,\mathrm{cm}$ measured from the vertex. Now go and put an object on the other side from the center, because, remember we're looking at a convex mirror – we're looking at the back of a spoon. In this example, let's put it 5 cm outside. Just as in the last section, I am going to make my object the thickness of my protractor ruler. That will make drawing a nice incoming parallel right easy. As always, we draw it as an arrow just so it's easy to tell which way is up but you can think of it as a little face, a tree, a candle, it could be anything.



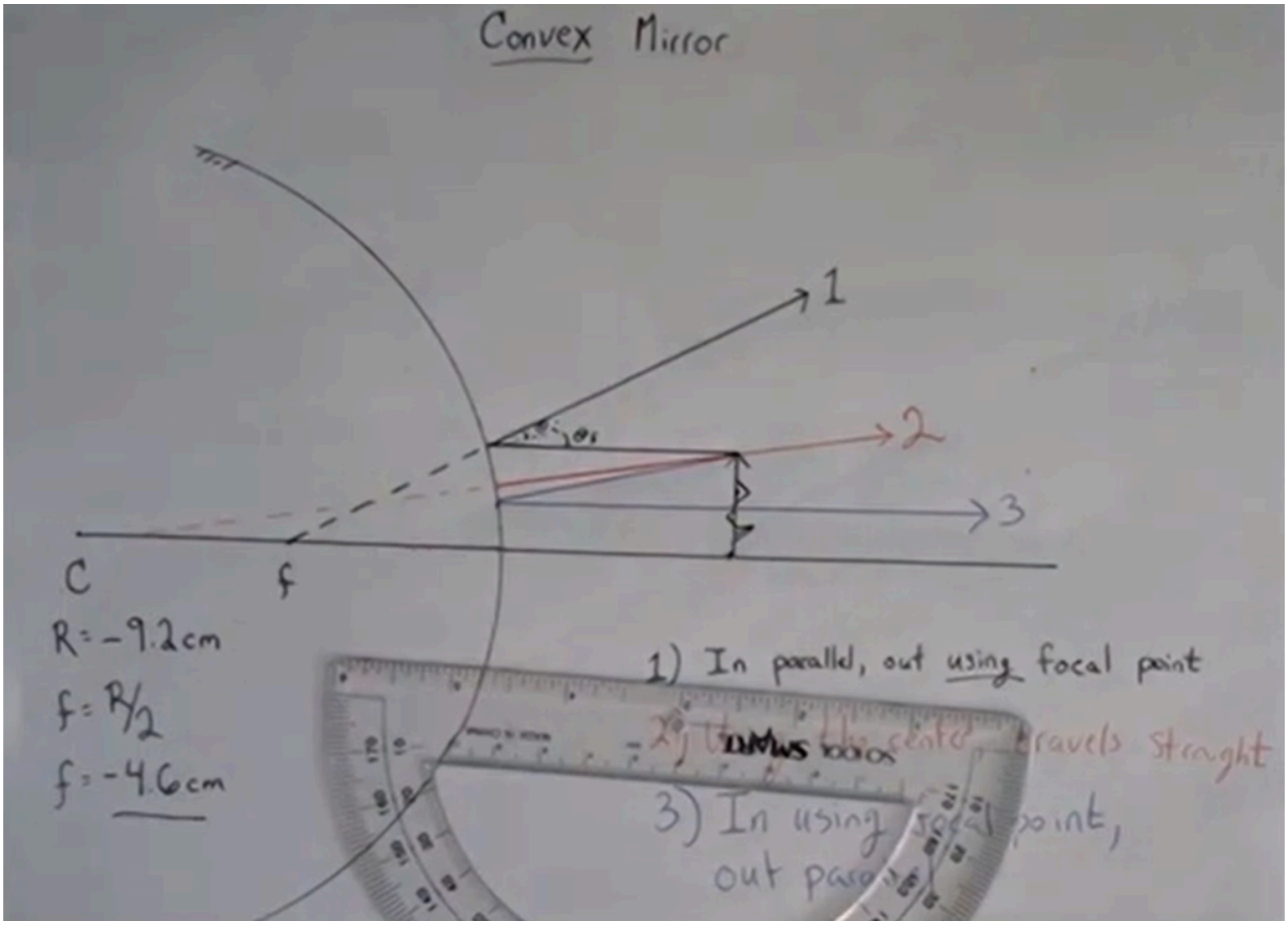
Once again, our rules are the same: the first ray is, as always, “in parallel out using the focal point.” Thus, I draw my incoming ray parallel. now I can’t go through because it’s a mirror, the ray has to bounce instead. You will also notice that again I use the verb *using* instead of *through* to help me remember that I can’t go throughthe focal point that’s behind the mirror. Thus, the ray is going to go out using the focal point: this ray is gonna bounce as if it had come from the focal point. That is what the focal point for a convex mirror means. Thus, we have ray number one and, as always, we can see that $\theta_i = \theta_f$.



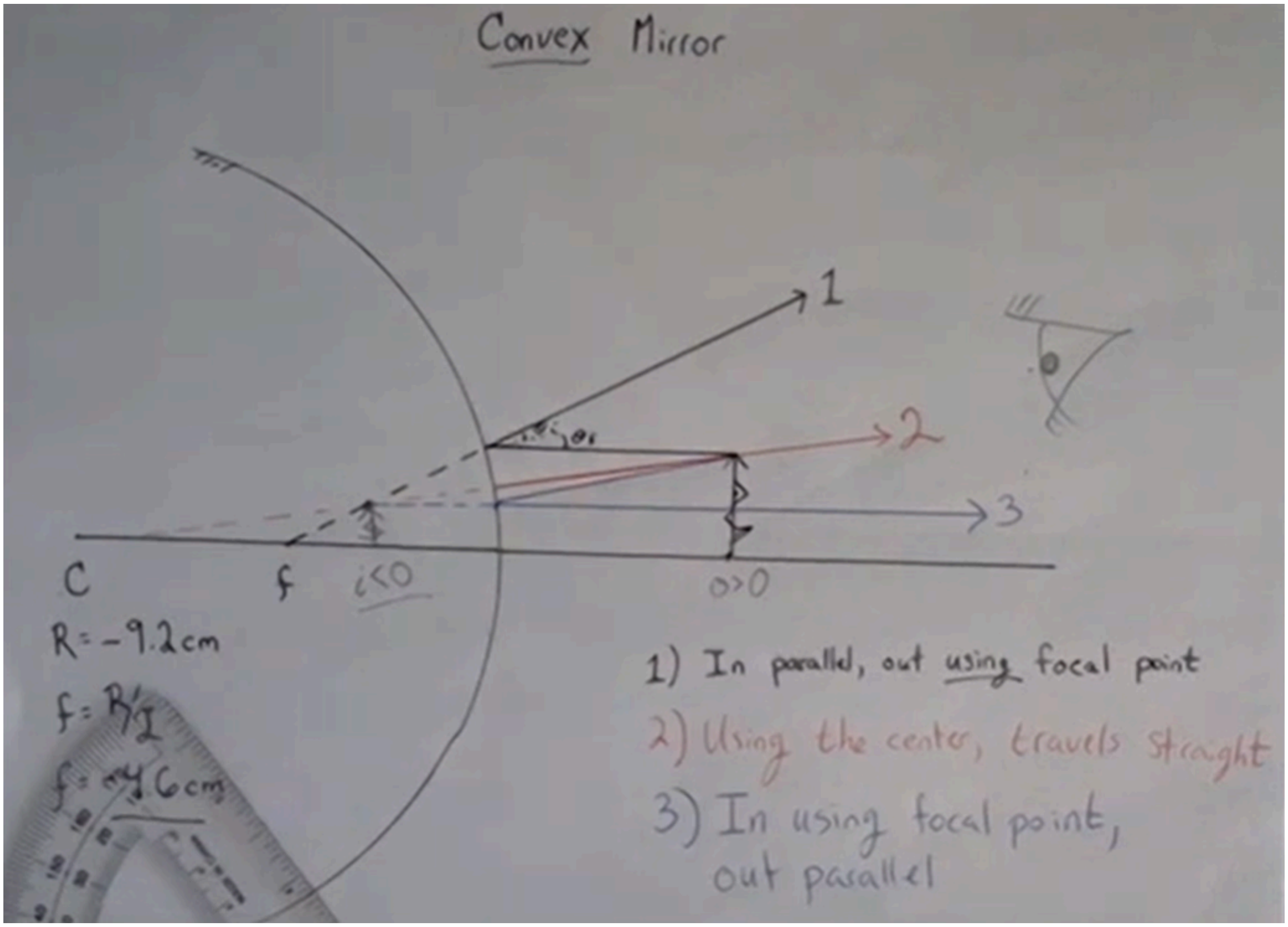
Ray number two: “a ray using the center travels straight.” For this ray, we are going to aim for the center of the mirror’s circle, but it can’t go through (because that’s a mirror) so instead it will bounce and head back off directly the way it came.



Ray number three: “in using the focal point out parallel.” This time we’re going to consider a ray that comes in as if it were going for the focal point behind the mirror and is going to go out parallel to the optical axis.



Now, as with the diverging lens describe above, these three rays don't converge anywhere; they're spraying out as they leave the mirror. This is a diverging optical element so they should spread out. However as with the diverging lens, if your eye is behind the object looking at this whole thing what your eye going to see is these three photons and it's going to assume these photons traveled in straight lines. All the photons appear to have originated from a point behind the mirror. That point is where our image is going to be. In this particular case, we have an object distance to the right that is positive, because the object is on the same side is the incoming light but we have a negative image distance because the light while it comes in on the right, also goes out on the right because of the mirror. The image is on the reverse side from this outgoing light, so the image distance is negative. Ultimately, we see a little upright erect image behind the mirror. If you look in the back of a spoon, as was shown in the last chapter, that's what you see you see: a littlemini version of yourself erect behind the spoon.



Homework Problem

Problem 26: Ray diagrams with convex mirrors.

16. Homework Problems

Homework

The list below is the list of homework problems in Edfinity. The numbering is the same. You can click on a problem, and it will take you to the relevant section of the book!

1. Identify the portion of the eye responsible for most of the focusing of light.
2. What is the purpose of the iris?
3. In geometric optics, we do analyses using similar triangles. This problem is here to help you practice working on these again.
4. Look at this map and determine the angle.
5. For this set of intersecting lines, use the following information to find the missing values.
6. Indicate where the outgoing ray from a mirror intersects the dotted line.
7. What is the speed of light in water? In glycerine? The indices of refraction for water is 1.333 and for glycerine is 1.473.
8. Calculate the index of refraction for a medium in which the speed of light is 1.416×108 m/s.
9. Consider two materials. When light passes through the space between the two materials at 0°<θ<90°, there is no change in the direction of the propagation of the light. What can you infer about the two materials?
10. Which of the properties of a light ray change as it goes from glass to vacuum?
11. What are the wavelengths of visible light in crown glass?
12. Suppose you have an unknown clear substance immersed in water, and you wish to identify it by finding its index of refraction. You arrange to have a beam of light enter it at an angle of 48.6∘, and you observe the angle of refraction to be 32.4∘. What is the index of refraction of the substance? Water has an index of refraction equal to 1.333.
13. A beam of white light goes from air into water at an incident angle of 83.0∘. At what angles are the red 660 nm and violet 410 nm parts of the light refracted? Red light in water has an index of refraction equal to 1.331 and that of violet light is 1.342.
14. Given that the angle between the ray in the water and the perpendicular to the water is 28.3˚, and using information in the figure above, find the height of the instructor's head above the water. Water has an index of refraction equal to 1.333.
15. Sign conventions for object and image distances: objects and images on opposite sides of the optical element.
16. Sign conventions for object and image distances: objects and images on same side of the optical element.
17. What characterizes an object with a negative magnification?
18. Calculating magnification of a gemstone.
19. Characterizing lenses.
20. Diopters to focal length.
21. Focal length to diopters.
22. Focal lengths of mirrors.

23. Ray diagrams: converging lenses.
24. Ray diagrams: diverging lenses.
25. Ray diagrams: concave mirrors.
26. Ray diagrams: convex mirrors.

PART III

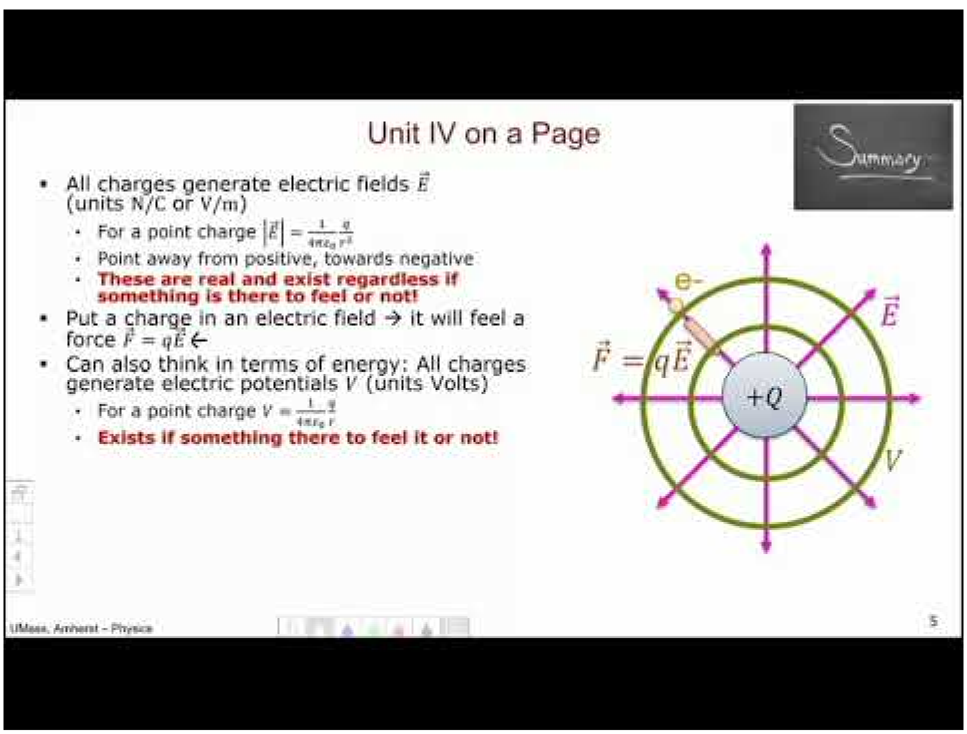
UNIT III

17. Unit III On-a-Page

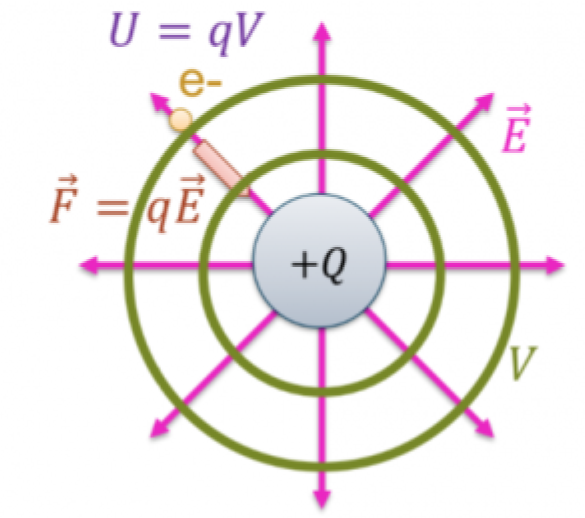
University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

This chapter very strongly follows the ideas of a few distinct **principles** discussed in [Unit I On-a-Page](#).



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=238



This image effectively summarizes the connections between the two new ideas in this chapter, electric fields and electric potentials, as well as showing how they connect to ideas with which you are already familiar: forces and potential energy.

- All charges generate electric fields $\vec{E}$ (units N/C orV/m)
- For a point charge $|\vec{E}| = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}$
- Point away from positive, towards negative
- These are real and exist regardless if something is there to feel or not!
- Put a charge in an electric field ⇒ it will feel a force $\vec{F} = q\vec{E}$
- Can also think in terms of energy: All charges generate electric potentials V (units Volts)
- For a point charge $V = \frac{1}{4\pi\varepsilon_0}\frac{q}{r}$
- Exists if something is there to feel or not!
- Put a charge in a potential it has a potential energy $U = qV$
- Electric potentials and electric fields are two sides of the same coin
- Just like in P131: Can think of forces or energy
- $\vec{E}$ points 'down the potential hill'
- $\vec{E} = -\frac{\Delta V}{\Delta x}$

18. Introduction

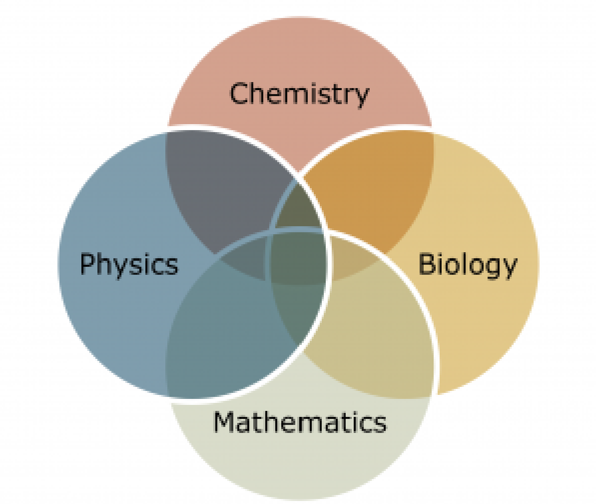
We know that electrons have electric charge. Having charge is one of the key ways in which electrons differ from photons. Up to this point, however, we have been completely neglecting the fact that electrons have charge. How does having charge impact how an electron behaves? Exploring this charged aspect of an electron's identity will be the focus of this chapter. Along the way, we will need to introduce two other concepts: the electric field and the electric potential.

As you may recall from Physics 131, there are four fundamental forces in the Universe: gravity, electricity/magnetism, the weak nuclear force, and the strong nuclear force. With the exception of gravity, all of the other forces in your everyday life are either electrical or magnetic in nature. The normal force that keeps you from falling through your chair is, in reality, electrical in origin: the electrons in the chair are electrically repelling the electrons in your body. The tension forces in ropes are also electrical: they arise from the chemical bonds in the rope which ultimately arise from the electrical attractions between protons and electrons. Since the electrical force is the ultimate source of all electrical bonds, that means that it is also the most relevant force for biology and chemistry!

The readings for this unit are ultimately divided up into four main parts. First we motivate the study of electricity with some applications to biology and chemistry. The second part is review. Our study of electrical forces will be, as all forces are, heavily dependent on being able to use vectors. Since it may have been a while since you studied vectors, we have included the chapter on vectors from the 131 textbook for your reference. If you feel comfortable with vectors, feel free to skip this section: it is just there for your review, but there are some homework problems to make sure you are fresh. Afterwards there are some problems reviewing charge conservation. Following these refreshers, we get to the meat of this unit: we will introduce the idea of electric field $\vec{E}$ which is the ultimate source of electrical forces. Finally, just as you can think about a falling ball in terms of the gravitational force or in terms of the gravitational potential energy, the same is true for electrical forces. We will, therefore, then consider the interactions between charges from an energy perspective and introduce the idea of electric potential V Note, electric potential and electric potential energy are two different, but related, ideas – be careful with your vocabulary here!

You may have seen a Coulomb's Law for electrical forces in other courses, but I really want you to try to think of electrical forces as arising from fields and potentials. Try to visualize them. Fields and potentials are just as real as the electrons and photons we have been studying. Being able to think of the interactions between charged particles in terms of electric fields and electric potentials will be key to being successful in the rest of this course.

19. Motivating Context for Unit III



This unit will focus on the electric field $\vec{E}$ and the electric potential V. As described in the introduction, the electric force is the underpinning of chemistry, so I would like you to refresh some ideas about chemical bonds that you probably saw in your chemistry classes that we will use in this unit.

Another case that we will explore in some detail in this unit is gel electrophoresis: a process you have probably discussed in your biology class. This laboratory technique is fundamentally based upon the ideas of electric field and potential we will discuss in this unit. Thus, a review of the procedure based upon OpenStax Microbiology chapter 12.2 is included below for your review.

Molecular Bond Basics

The following material is from *OpenStax Chemistry: Atoms First 2e – Chapter 5.1 Valence Bond Theory*

Valence bond theory describes a covalent bond as the overlap of half-filled atomic orbitals (each containing a single electron) that yield a pair of electrons shared between the two bonded atoms. We say that orbitals on two different atoms overlap when a portion of one orbital and a portion of a second orbital occupy the same region of space. According to valence bond theory, a covalent bond results when two conditions are met: (1) an orbital on one atom overlaps an orbital on a second atom and (2) the single electrons in each orbital combine to form an electron pair. The mutual attraction between this negatively charged electron pair and the two atoms' positively charged nuclei serves to physically link the two atoms through a force we define as a covalent bond. The strength of a covalent bond depends on the extent of overlap of the orbitals involved. Orbitals that overlap extensively form bonds that are stronger than those that have less overlap. **We calculated the shapes of these orbitals for simple molecules such as 1,3-butadine in-class during Unit I by modeling the electrons as a particle in a box!**

The energy of the system depends on how much the orbitals overlap. Figure 1 illustrates how the sum of the energies of two hydrogen atoms (the colored curve) changes as they approach each other. When the atoms are far apart there is no overlap, and by the convention described in Unit I – Chapter 5 Some Energy-Related Ideas that Might be New or Are Particularly Important: The Potential Energy of Atoms and Molecules the potential energy of each atom is zero. As the atoms move together, the electron waves (orbitals) begin to overlap. Each electron begins to feel the attraction of the nucleus in the other atom. In addition, the electrons begin to repel each other, as do the nuclei. The result is the electron waves change shape in response.

While the atoms are still widely separated, the attraction is slightly stronger than the repulsion, and the energy of the system decreases. (A bond begins to form.) As the atoms move closer together, the overlap increases, so the attraction of the nuclei for the electrons continues to increase (as do the repulsions among electrons and between the nuclei). At some specific distance between the atoms, which varies depending on the atoms involved, the energy reaches its lowest (most stable) value. This optimum distance between the two bonded nuclei is the bond distance between the two atoms. The bond is stable because at this point, the attractive and repulsive forces combine to create the lowest possible energy configuration. If the distance between the nuclei were to decrease further, the repulsions between nuclei and the repulsions as electrons are confined in closer proximity to each other would become stronger than the attractive forces. The energy of the system would then rise (making the system destabilized), as shown at the far left of Figure 1.

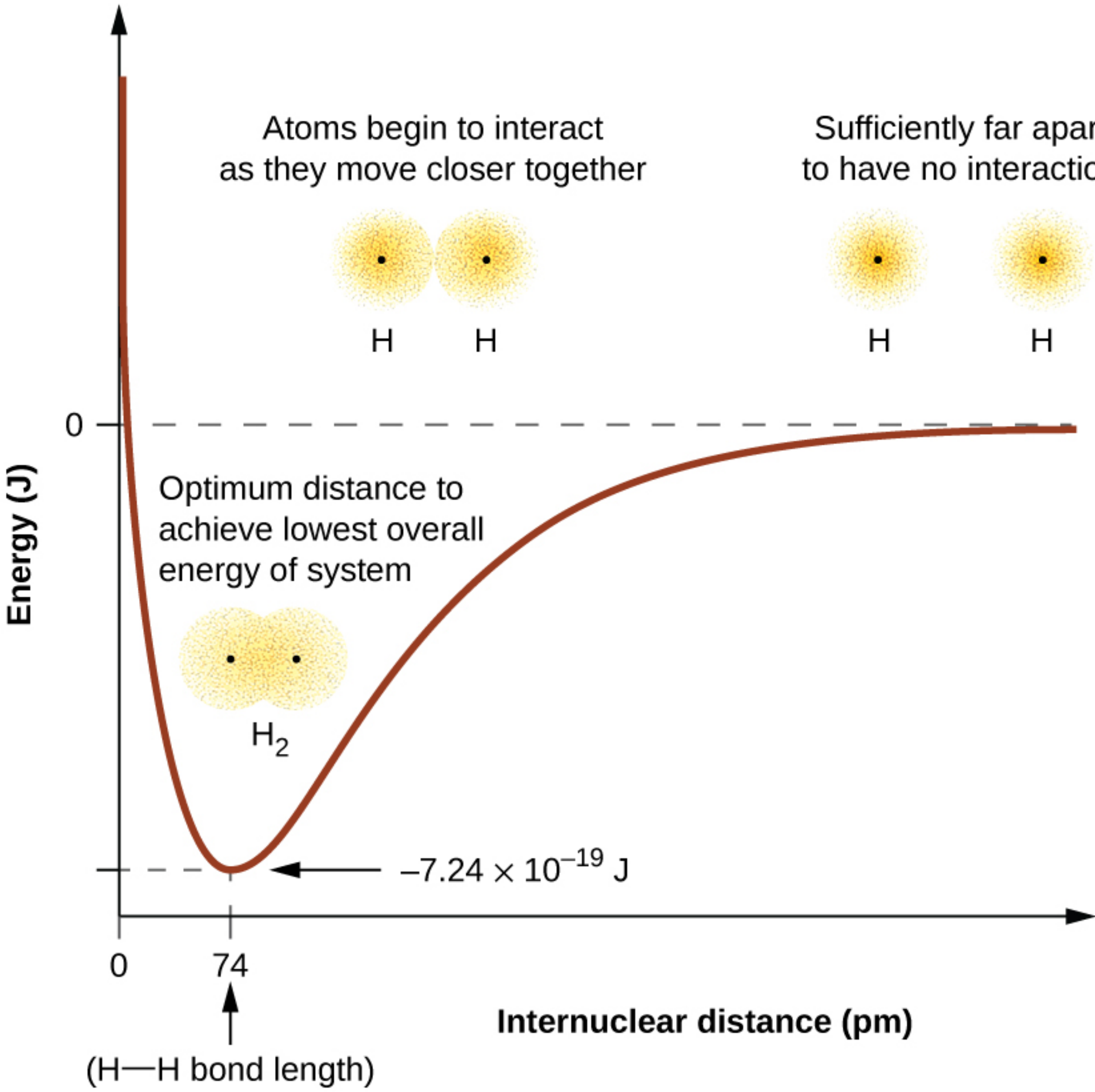


Figure 1 (a) The interaction of two hydrogen atoms changes as a function of distance. (b) The energy of the system changes as the atoms interact. The lowest (most stable) energy occurs at a distance of 74 pm, which is the bond length observed for the H2 molecule.

Play with the energies of atoms and bonds

Below is a simulation where you can see the energies of atoms and their bonds. Play around with it. In this simulation, one atom is "pinned down" and the other is free to move.

- Drag the free atom wherever you wish and let it go. It will move according to its potential energy.

- You can see the overlap in the electron waves (electron clouds) at the bottom.
- You can change the types of atoms using the menu in the upper right.
- You can turn on the forces as well to see how the forces are responding to the potential energy.

You will notice that the atoms do not stay at a fixed distance; they bounce as skateboarders on a hill! You can even get neon atoms to "bond" although there is not a lot of room to do so! This is true; at small enough temperatures you can get neon to bond. At 24.56K it will actually become a solid. You can see O=O, on the other hand, has a much deeper well and is therefore a much stronger bond.

An interactive or media element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=1470

Homework

Problem 1: Comparing the energies of bound atoms vs. free atoms.

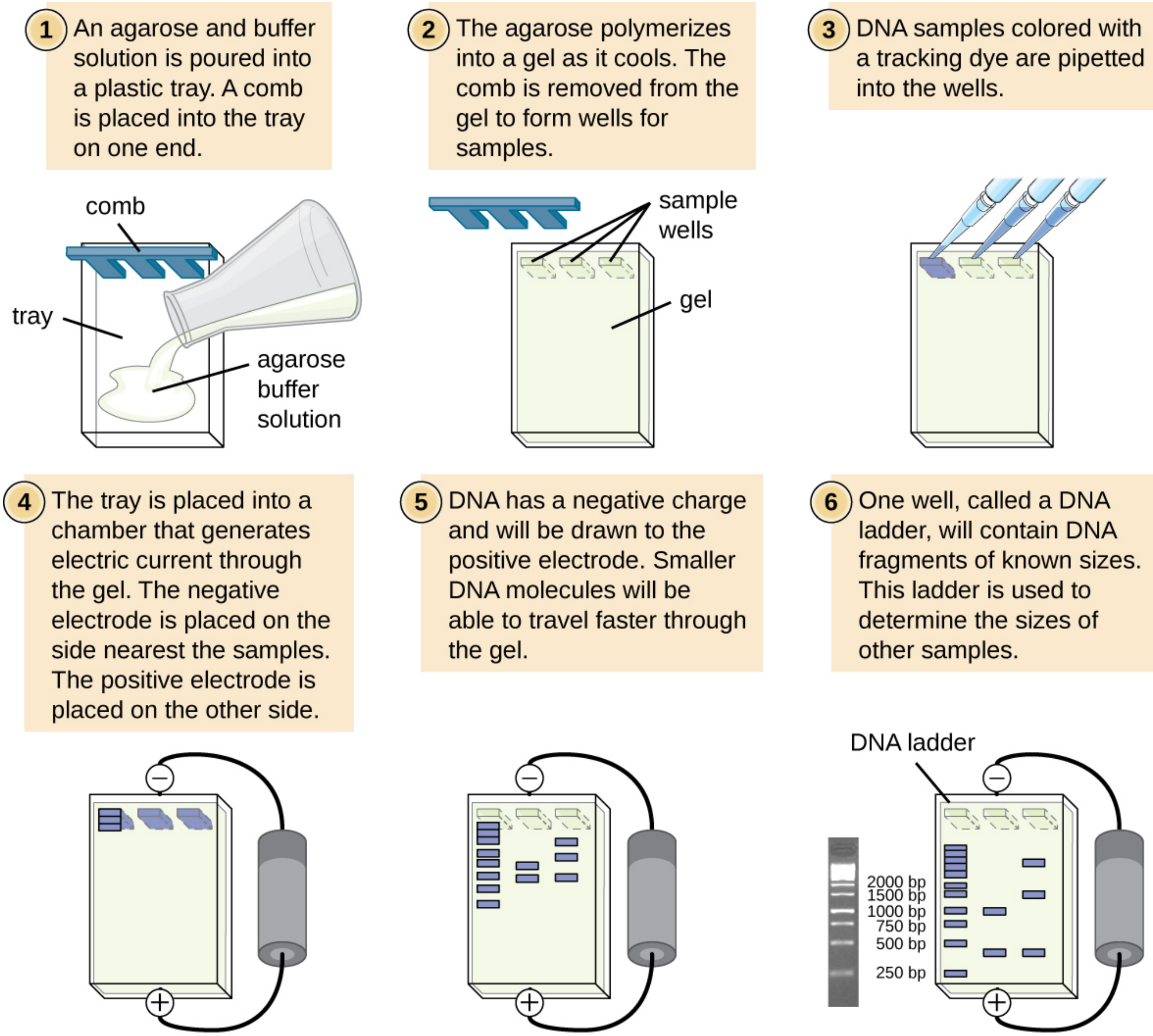
Basic Description of Gel Electrophoresis

We will explore this device in class, so it is probably beneficial if you have already familiar with how it works. This material is from OpenStax Microbiology – Chapter 12.2 Visualizing and Characterizing DNA, RNA, and Protein.

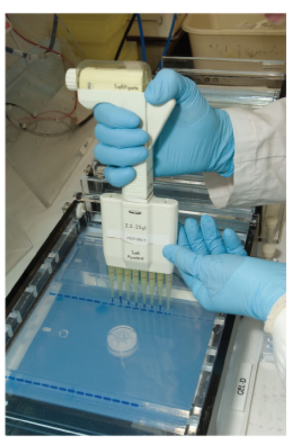
There are a number of situations in which a researcher might want to physically separate a collection of DNA fragments of different sizes. A researcher may also digest a DNA sample with a restriction enzyme to form fragments. The resulting size and fragment distribution pattern can often yield useful information about the sequence of DNA bases that can be used, much like a bar-code scan, to identify the individual or species to which the DNA belongs.

Gel electrophoresis is a technique commonly used to separate biological molecules based on size and biochemical characteristics, such as charge and polarity. Agarose gel electrophoresis is widely used to separate DNA (or RNA) of varying sizes that may be generated by restriction enzyme digestion or by other means, such as the PCR (Figure 2).

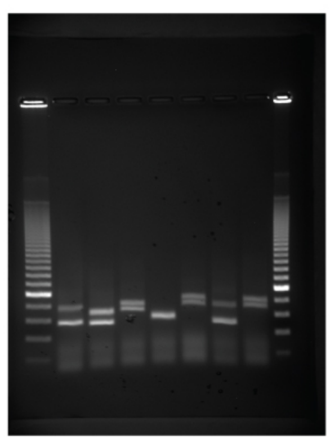
Due to its negatively charged backbone, DNA is strongly attracted to a positive electrode. In agarose gel electrophoresis, the gel is oriented horizontally in a buffer solution. Samples are loaded into sample wells on the side of the gel closest to the negative electrode, then drawn through the molecular sieve of the agarose matrix toward the positive electrode. The agarose matrix impedes the movement of larger molecules through the gel, whereas smaller molecules pass through more readily. Thus, the distance of migration is inversely correlated to the size of the DNA fragment, with smaller fragments traveling a longer distance through the gel. Sizes of DNA fragments within a sample can be estimated by comparison to fragments of known size in a DNA ladder also run on the same gel. To separate very large DNA fragments, such as chromosomes or viral genomes, agarose gel electrophoresis can be modified by periodically alternating the orientation of the electric field during pulsed-field gel electrophoresis (PFGE). In PFGE, smaller fragments can reorient themselves and migrate slightly faster than larger fragments and this technique can thus serve to separate very large fragments that would otherwise travel together during standard agarose gel electrophoresis. In any of these electrophoresis techniques, the locations of the DNA or RNA fragments in the gel can be detected by various methods. One common method is adding ethidium bromide, a stain that inserts into the nucleic acids at non-specific locations and can be visualized when exposed to ultraviolet light. Other stains that are safer than ethidium bromide, a potential carcinogen, are now available.


1 An agarose and buffer solution is poured into a plastic tray. A comb is placed into the tray on one end.
2 The agarose polymerizes into a gel as it cools. The comb is removed from the gel to form wells for samples.
3 DNA samples colored with a tracking dye are pipetted into the wells.
comb
tray
agarose buffer solution
sample wells
gel
4 The tray is placed into a chamber that generates electric current through the gel. The negative electrode is placed on the side nearest the samples. The positive electrode is placed on the other side.
5 DNA has a negative charge and will be drawn to the positive electrode. Smaller DNA molecules will be able to travel faster through the gel.
6 One well, called a DNA ladder, will contain DNA fragments of known sizes. This ladder is used to determine the sizes of other samples.
DNA ladder
2000 bp
1500 bp
1000 bp
750 bp
500 bp
250 bp

(a)



(b)



(c)

Figure 2 (a) The process of agarose gel electrophoresis. (b) A researcher loading samples into a gel. (c) This photograph shows a completed electrophoresis run on an agarose gel. The DNA ladder is located in lanes 1 and 9. Seven samples are located in lanes 2 through 8. The gel was stained with ethidium bromide and photographed under ultraviolet light. (credit a: modification of work by Magnus Manske; credit b: modification of work by U.S. Department of Agriculture; credit c: modification of work by James Jacob)

Homework

Problem 2: Reviewing gel electrophoresis.

20. Basics of Charge

Static Electricity and Charge

Editors' note: This section is derived from Derived from 18.2 Static Electricity and Charge: Conservation of Charge by OpenStax, Bobby Bailey

Figure 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)*

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 (see Figure 1). The very word *electric* derives from the Greek word for amber (*electron*).

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 aluminum foil on the bottoms to avoid creating sparks which may ignite the oxygen being used.

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 2 shows how these simple materials can be used to explore the nature of the force between charges.

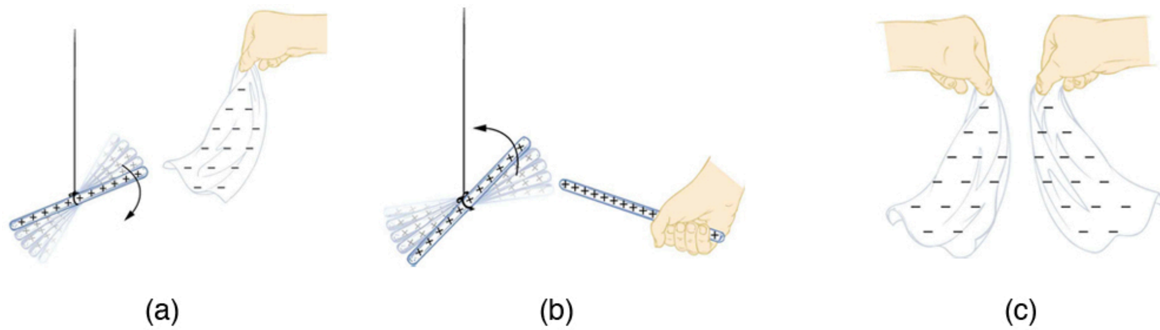


Figure 2: *A glass rod becomes positively charged when rubbed with silk, while the silk becomes negatively charged. (a) The glass rod is attracted to the silk because their charges are opposite. (b) Two similarly charged glass rods repel. (c) Two similarly charged silk cloths repel.*

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 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.

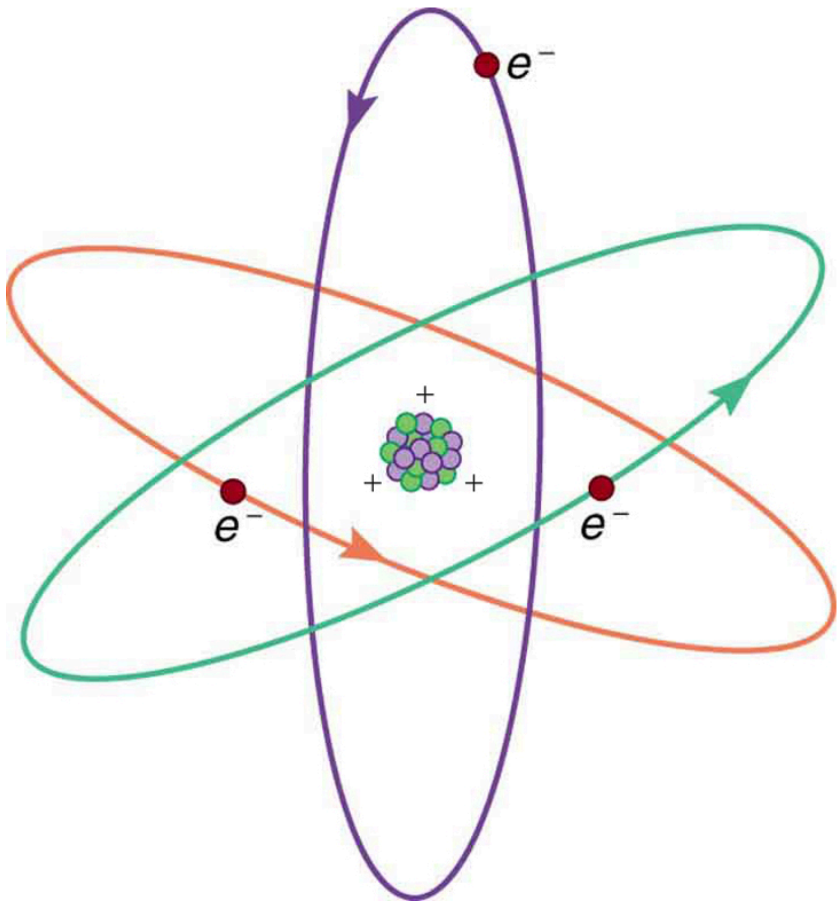


Figure 3: *This simplified (and not to scale) view of an atom is called the planetary model of the atom. Negative electrons orbit a much heavier positive nucleus, as the planets orbit the much heavier sun. There the similarity ends, because forces in the atom are electromagnetic, whereas those in the 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

$$|q_e| = 1.602 \times 10^{-19}\mathrm{C}$$

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

$$1.00\,\mathrm{C} \times \frac{1\,proton}{1.602\times10^{-19}\,\mathrm{C}} = 6.25 \times 10^{18}\,\mathrm{protons}$$

Similarly, 6.25×10^{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 $|q_e|$, and all observed charges are integral multiples of $|q_e|$.

Figure 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.

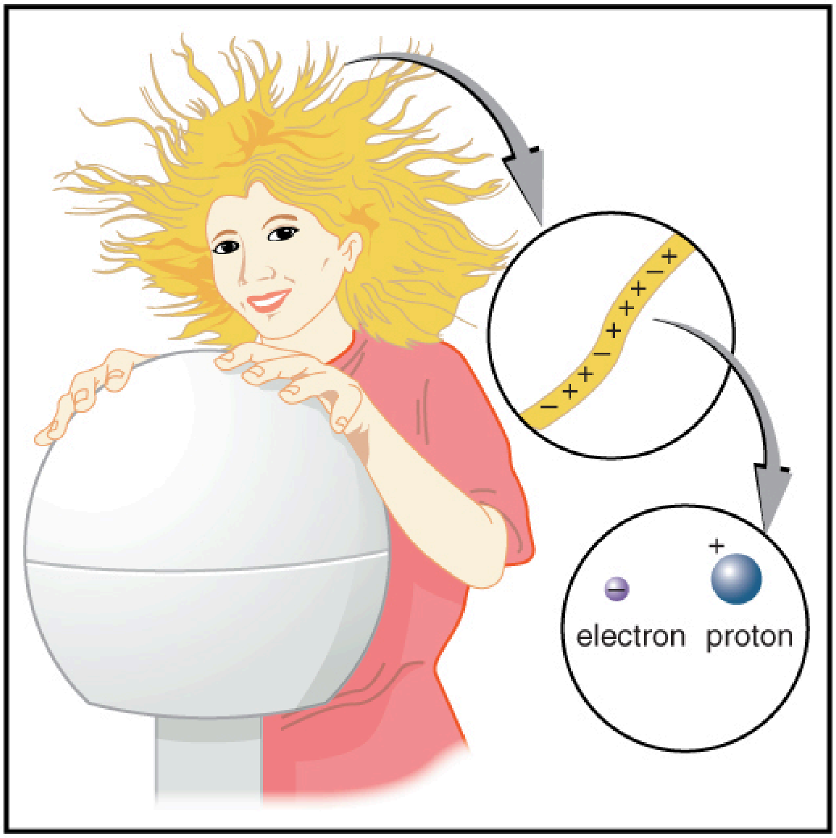


Figure 4: *When this person touches a Van de Graaff generator, she receives an excess of positive charge, causing her hair to stand on end. The charges in one hair are shown. An artist's conception of an electron and a proton illustrate the particles carrying the negative and positive charges. We cannot 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.*

Key Takeaways: 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. (See Figure 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.

Instructor's Note

This fact that there are no observed free particles with less than $|q_e|$ of charge is important and will be used in some of your homework problems.

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. (See Figure 5.) 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.

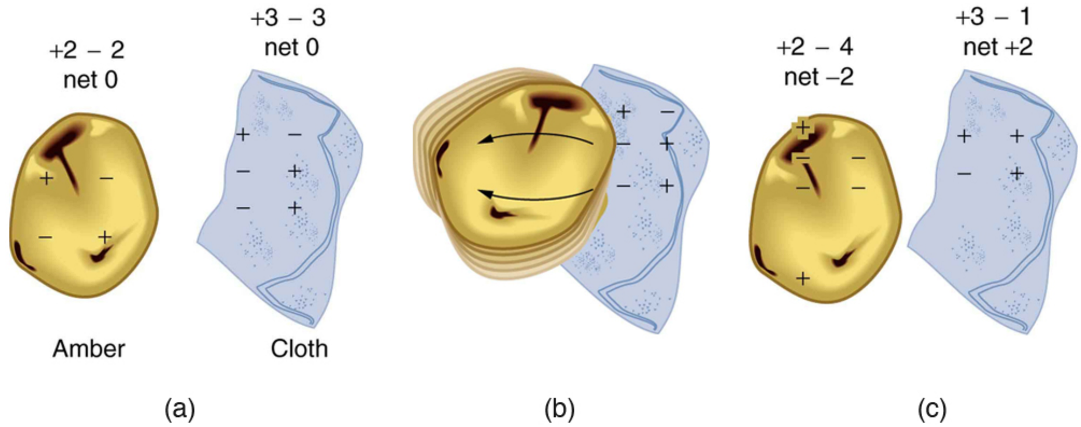


Figure 5: *When materials are rubbed together, charges can be separated, particularly if one material has a greater affinity for electrons than another. (a) Both the amber and cloth are originally neutral, with equal positive and negative charges. Only a tiny fraction of the charges are involved, and 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.

Play with the Simulation

Below is a simulation of a balloon and a sweater. As you probably know, if you rub a balloon on a sweater, it will stick to a wall.

A few things to note:

- The total number of charges is conserved – electrons move from the sweater to the balloon.
- If you have two balloons with negative charge, they will repel, just like in real life (check it for real if you don't believe us!)
- When you bring the balloon near the wall, what happens to the electrons in the wall?

An interactive or media element has been excluded from this version of the text. You can view it online here:
http://openbooks.library.umass.edu/toggerson-132/?p=184

Law of Conservation of Charge

The total charge is constant in *any* process.

Section 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 $|q_e| = 1.602 \times 10^{-19}\,\mathrm{C}$
- 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.

Homework Problems

Problem 3: Converting from Coulombs to numbers of particles.

21. Vector Review

Instructor's Note

You should already be familiar with vectors, but we will use them in this unit so this chapter from Physics 131 is a review. If this material is familiar, feel free to go to the homework problems at the end (there are none embedded in the sections as it *is* review)

Your Quiz would Cover

- A vector is a quantity with a magnitude and direction
- Converting between magnitude/direction and the component form for any vector. This ties into the Pythagorean Theorem

Kinematics in Two Dimensions: an Introduction

Figure 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.

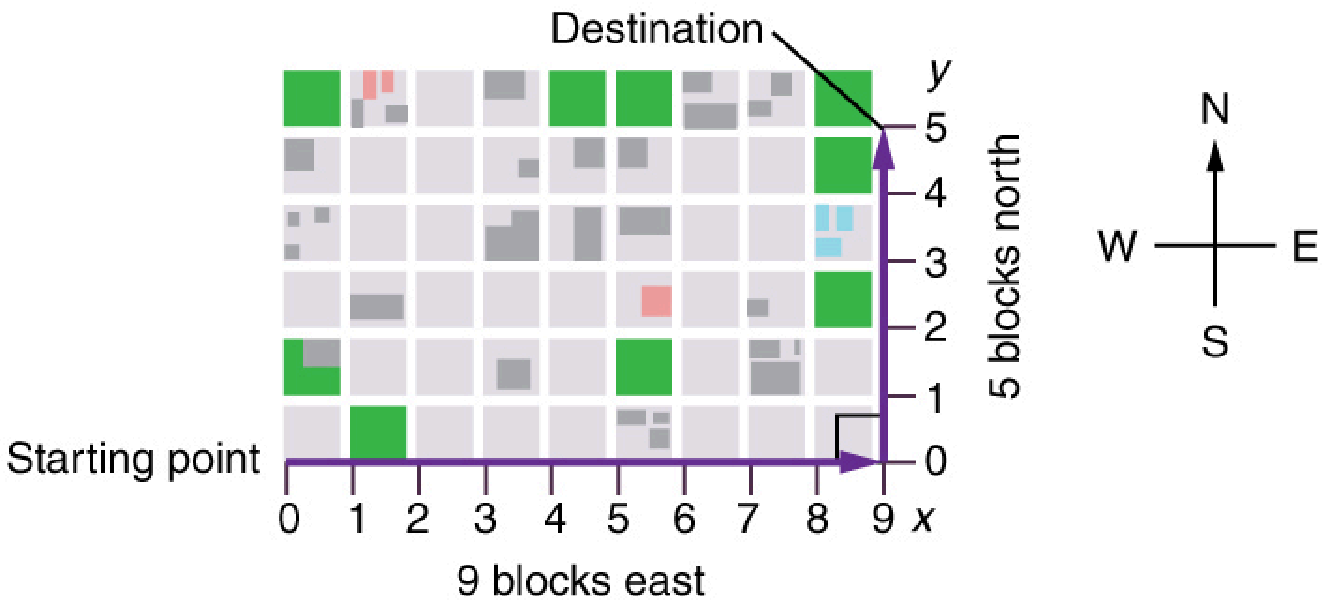


Figure 2*. A pedestrian walks a two-dimensional path between two points in a city. In this scene, all blocks are square and are the same size.*

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, $a^2 + b^2 = c^2$, can be used to find the straight-line distance.

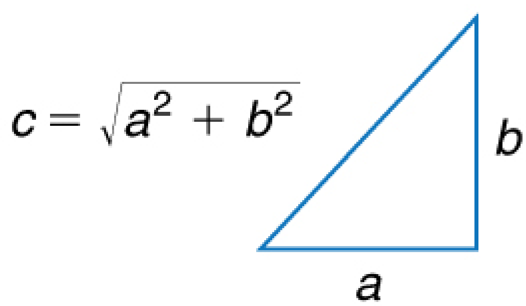


Figure 3. *The Pythagorean theorem relates the length of the legs of a right triangle, labeled a and b, with the hypotenuse, labeled c. The relationship is given by:* $a^2 + b^2 = c^2$*. This can be rewritten, solving for c :* $c = \sqrt{a^2 + b^2}$

University *of* Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

We will be using the Pythagorean Theorem all throughout two-dimensional kinematics, as well as throughout this entire course. If you are uncomfortable or unfamiliar with the Pythagorean Theorem, or even if it's just been a long time since you've used it, please come see your instructor as soon as possible and they will get you up to speed.

The hypotenuse of the triangle is the straight-line path, and so in this case its length in units of city blocks is $\sqrt{(9\text{ blocks})^2+(5\text{ blocks})^2}=10.3\text{ blocks}$, 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.)

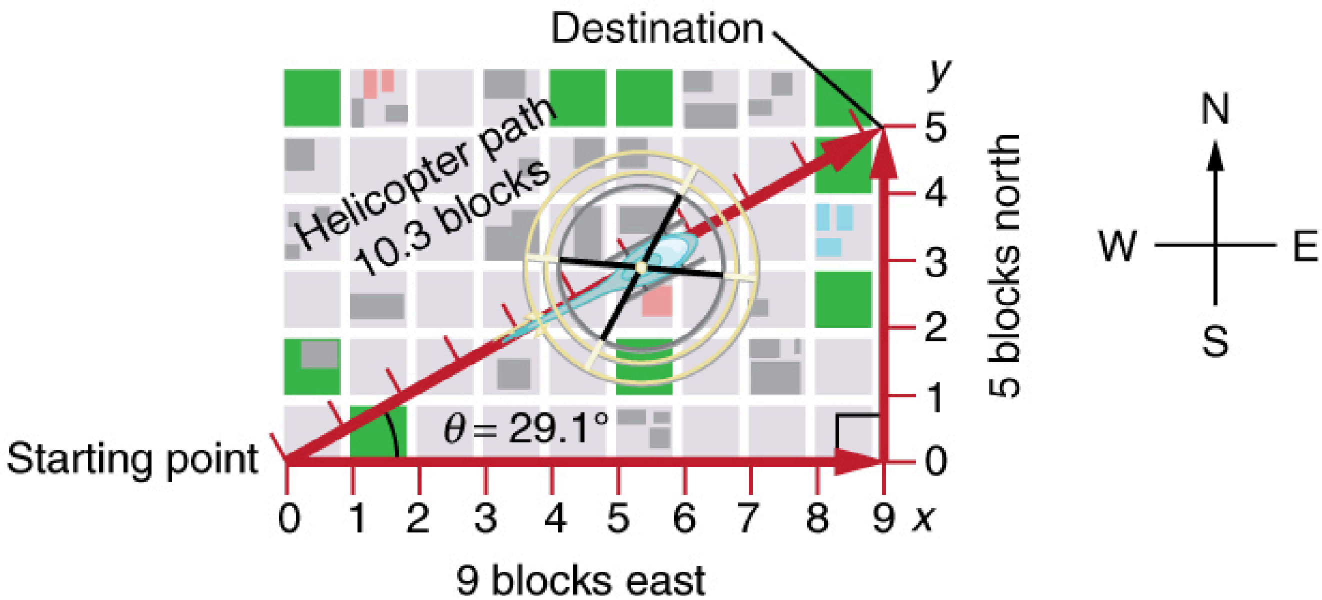


***Figure 4**. The straight-line path followed by a helicopter between the two points is shorter than the 14 blocks walked by the pedestrian. All blocks are square and the same size.*

The fact that the straight-line distance (10.3 blocks) in Figure 4 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 2 and Figure 4. 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 4. 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

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

The idea of the independence of perpendicular motion is a fundamental one that you should take some time to think about, and there are some questions about this on the homework.

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.

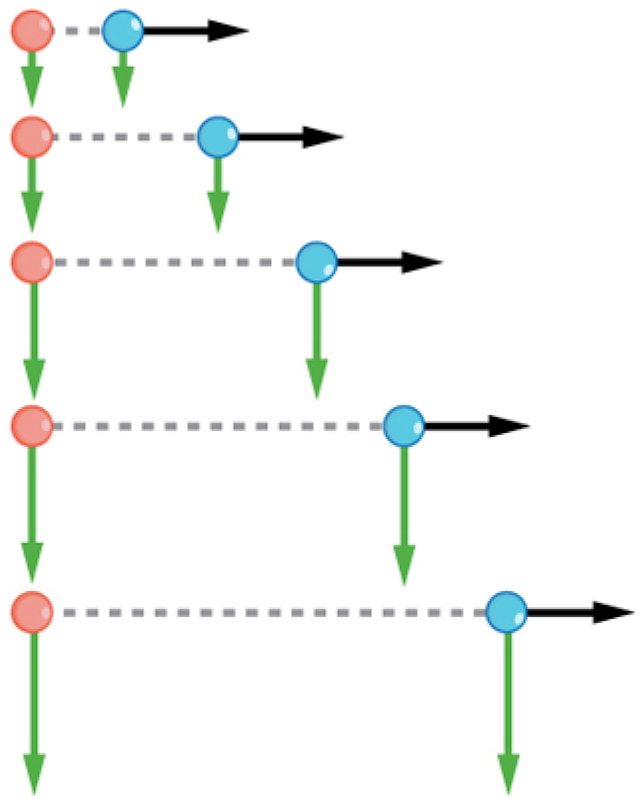


Figure 5. *This shows the motions of two identical balls—one falls from rest, the other has an initial horizontal velocity. Each subsequent position is an equal time interval. Arrows represent horizontal and vertical velocities at each position. The ball on the right has an initial horizontal velocity, 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.*

Instructor's Note

This graphic displays this concept quite nicely; notice how both balls fall downward at the same speed at each point, even though one of the balls has a horizontal velocity. Basically, the velocity of the ball in the x-direction has no effect on the velocity in the y-direction, and vice-versa. This will be an important idea, especially when working with vectors.

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.

Section 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.

Vector Addition and Subtraction: Graphical Methods

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Your Quiz would Cover

- Given two graphical representations of vectors, be able to draw the sum or difference. There are some simple procedures to follow. Solidify your understanding of these procedures and we can work on why this makes sense in class
- Describe both visually and mathematically what happens when a scalar is multiplied by a vector. If I give you a vector and a number, you should be able to turn the crank and multiply them mathematically. I am NOT expecting you to be able to do this graphically and will not ask you what it means. Just focus on the mechanics of how to do it.
- Convert between magnitude/direction and component form for any vector

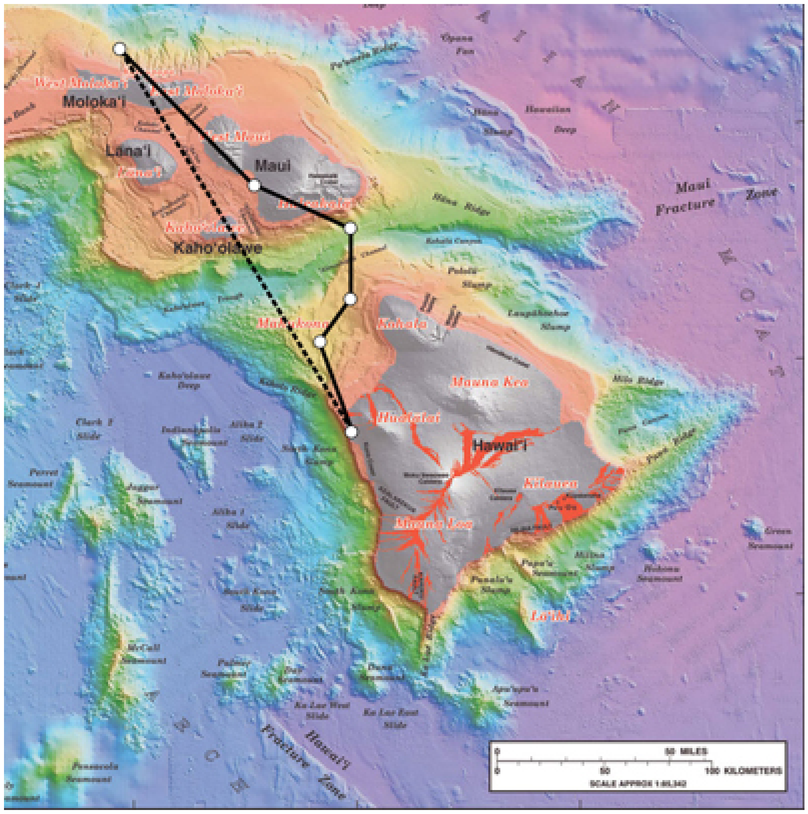


Figure 1. *Displacement can be determined graphically using a scale map, such as this one of the Hawaiian Islands. A journey from Hawai'i to Moloka'i has a number of legs, or journey segments. These segments can be added graphically with a ruler to determine the total two-dimensional displacement of the journey. (credit: US Geological Survey)*

Vectors in Two Dimensions

A vector is a quantity that has magnitude and direction. Displacement, velocity, acceleration, and force, for example, are all vectors. In one-dimensional, or straight-line, motion, the direction of a vector can be given simply by a plus or minus sign. In two dimensions (2-d), however, we specify the direction of a vector relative to some reference frame (i.e., coordinate system), using an arrow having length proportional to the vector's magnitude and pointing in the direction of the vector.

Figure 2 shows such a *graphical representation of a vector*, using as an example the total displacement for the person walking in a city considered in Kinematics in Two Dimensions: An Introduction. We shall use the notation that a boldface symbol, such as $\mathbf{D}$, stands for a vector. Its magnitude is represented by the symbol in italics, D, and its direction by θ

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

There's some notation in the following note that would be useful to pay attention to.

Vectors in this Text

In this text, we will represent a vector with a boldface variable. For example, we will represent the quantity force with the vector $\mathbf{F}$, which has both magnitude and direction. The magnitude of the vector will be represented by a variable in italics, such as F, and the direction of the variable will be given by an angle θ.

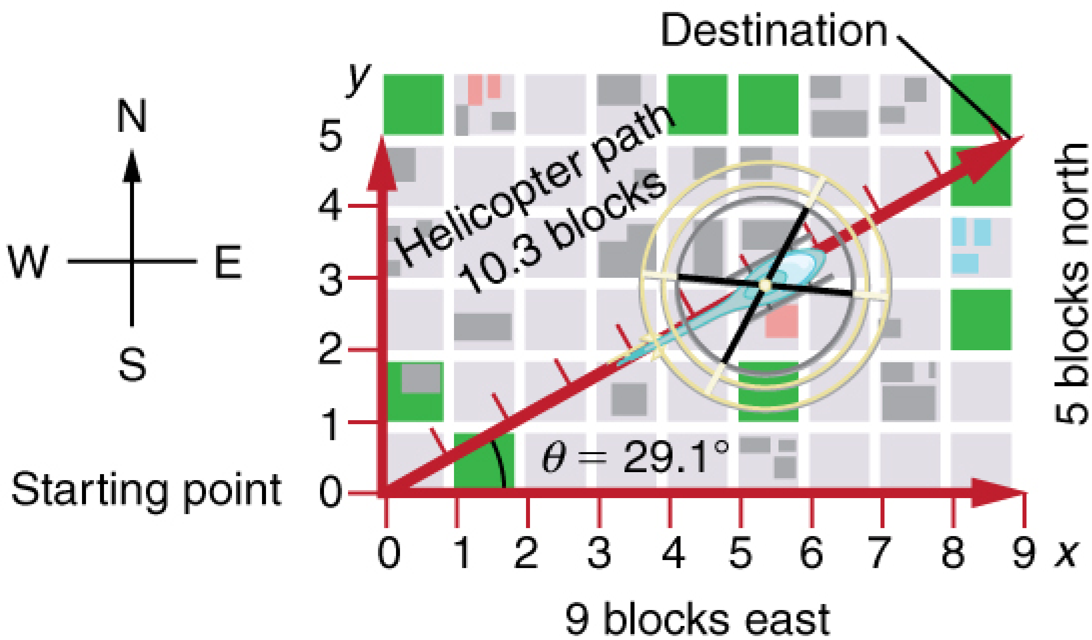


Figure 2. A person walks 9 blocks east and 5 blocks north. The displacement is 10.3 blocks at an angle $29.1°$ north of east.

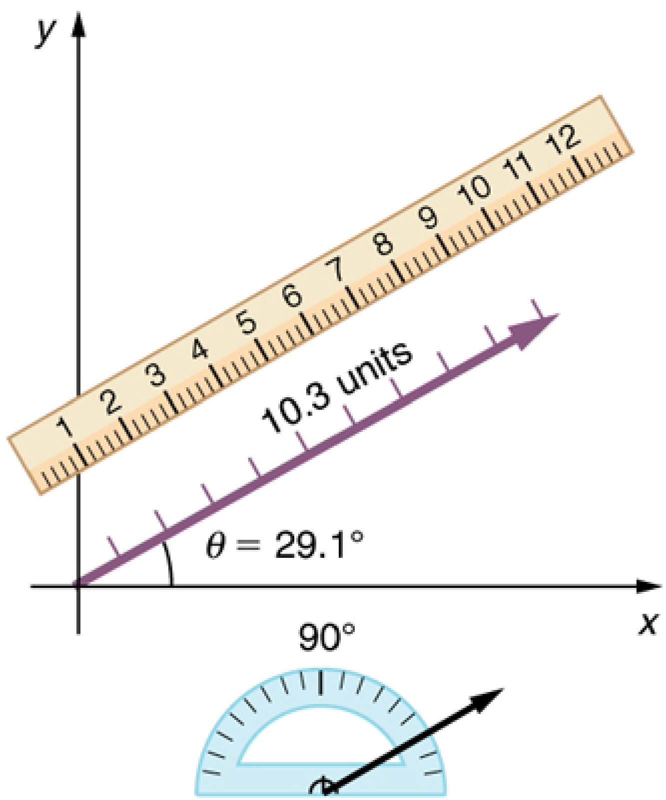


Figure 3. *To describe the resultant vector for the person walking in a city considered in Figure graphically, draw an arrow to represent the total displacement vector* $\mathbf{D}$*. Using a protractor, draw a line at an angle* θ *relative to the east-west axis. The length* D *of the arrow is proportional to the vector's magnitude and is measured along the line with a ruler. In this example, the magnitude* D *of the vector is 10.3 units, and the direction* θ *is* $29.1°$ *north of east.*

Instructor's Note

Taking some time to understand and practice the head-to-tail method is recommended, you'll notice that there's a series of algorithmic steps, so you just need to learn the process, and it will work for any two vectors.

Vector Addition: Head-to-Tail Method

The **head-to-tail method** is a graphical way to add vectors, described in Figure 4 below and in the steps following. The **tail[/pb_glossary]** of the vector is the starting point of the vector, and the **head (or tip)** of a vector is the final, pointed end of the arrow.

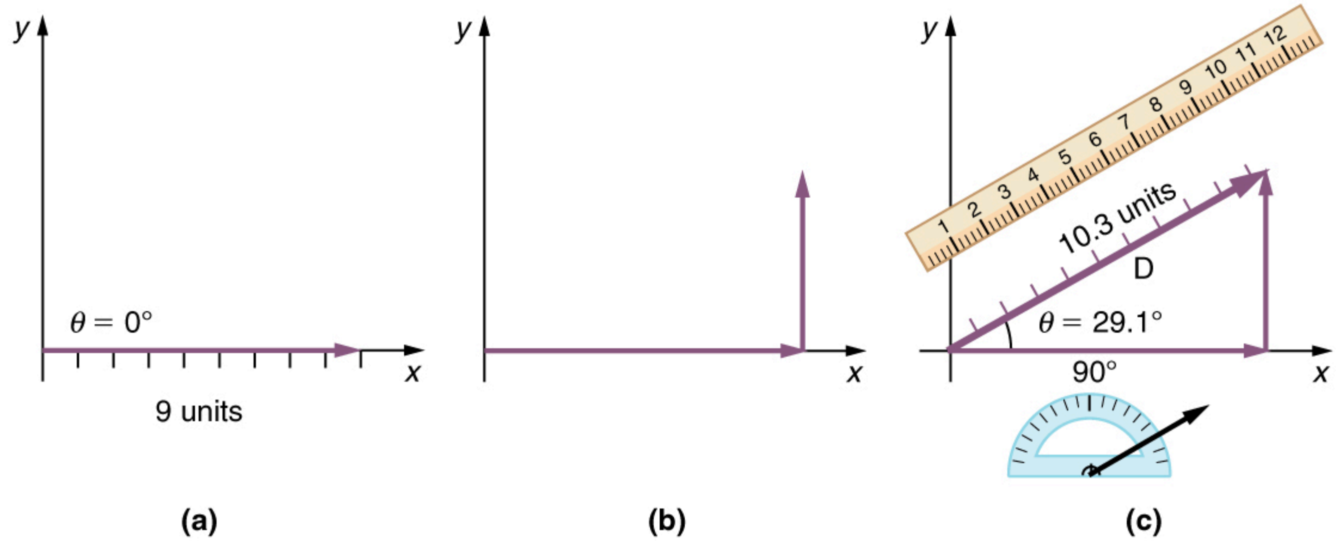


Figure 4. Head-to-Tail Method: *The head-to-tail method of graphically adding vectors is illustrated for the two displacements of the person walking in a city considered in Figure. (a) Draw a vector representing the displacement to the east. (b) Draw a vector representing the displacement to the north. The tail of this vector should originate from the head of the first, east-pointing vector. (c) Draw a line from the tail of the east-pointing vector to the head of the north-pointing vector to form the sum or **resultant vector** $\mathbf{D}$. The length of the arrow D is proportional to the vector's magnitude and is measured to be 10.3 units . Its direction, described as the angle with respect to the east (or horizontal axis) θ is measured with a protractor to be 29.1°.*

Step 1. *Draw an arrow to represent the first vector (9 blocks to the east) using a ruler and protractor.*

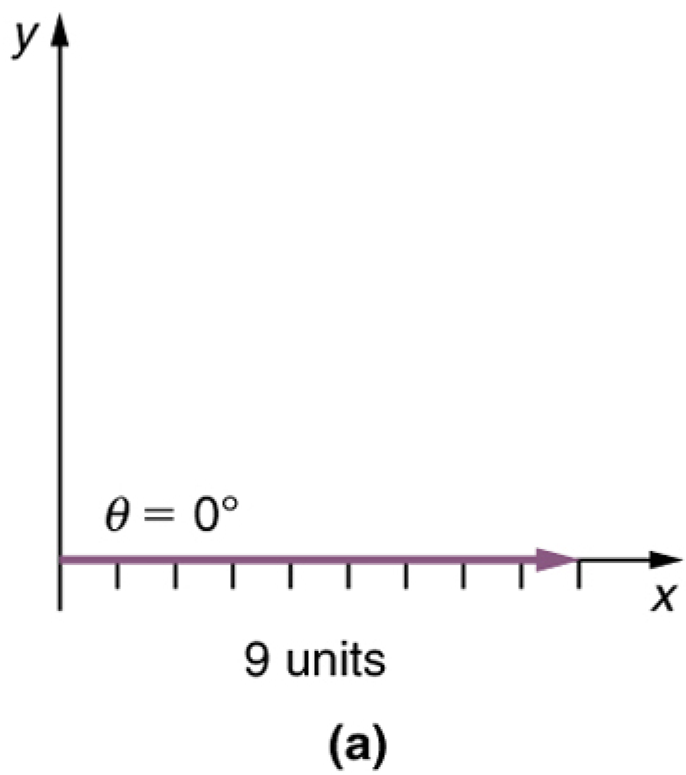


Figure 5.

Step 2. Now draw an arrow to represent the second vector (5 blocks to the north). *Place the tail of the second vector at the head of the first vector*.

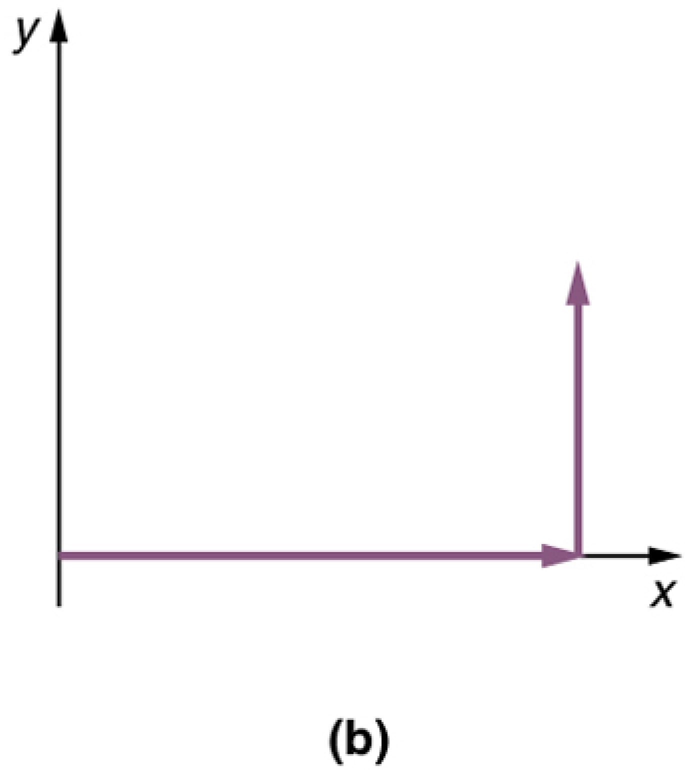


Figure 6.

Step 3. *If there are more than two vectors, continue this process for each vector to be added. Note that in our example, we have only two vectors, so we have finished placing arrows tip to tail*.

Step 4. *Draw an arrow from the tail of the first vector to the head of the last vector*. This is the **resultant**, or the sum, of the other vectors.

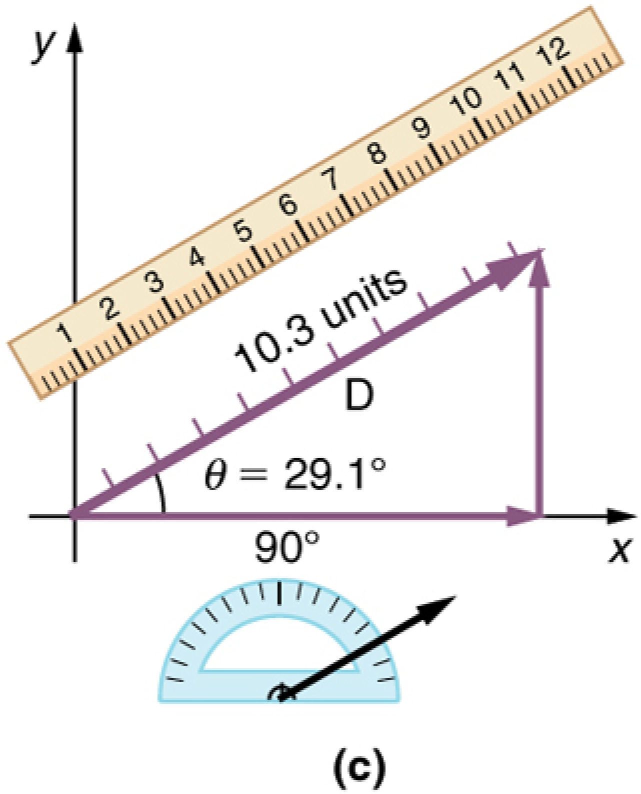


Figure 7.

Step 5. To get the **magnitude** of the resultant, *measure its length with a ruler. (Note that in most calculations, we will use the Pythagorean theorem to determine this length.)*

Step 6. To get the **direction** of the resultant, *measure the angle it makes with the reference frame using a protractor. (Note that in most calculations, we will use trigonometric relationships to determine this angle.)*

The graphical addition of vectors is limited in accuracy only by the precision with which the drawings can be made and the precision of the measuring tools. It is valid for any number of vectors.

Adding Vectors Graphically using the Head-to-Tail Method: A Woman Takes a Walk

Use the graphical technique for adding vectors to find the total displacement of a person who walks the following three paths (displacements) on a flat field. First, she walks 25.0 m in a direction $49.0°$ north of east. Then, she walks 23.0 m heading $15.0°$ north of east. Finally, she turns and walks 32.0 m in a direction $68.0°$ south of east.

Strategy

Represent each displacement vector graphically with an arrow, labeling the first $\mathbf{A}$, the second $\mathbf{B}$, and the third $\mathbf{C}$, making the lengths proportional to the distance and the directions as specified relative to an east-west line. The head-to-tail method outlined above will give a way to determine the magnitude and direction of the resultant displacement, denoted $\mathbf{R}$.

Solution

(1) Draw the three displacement vectors.

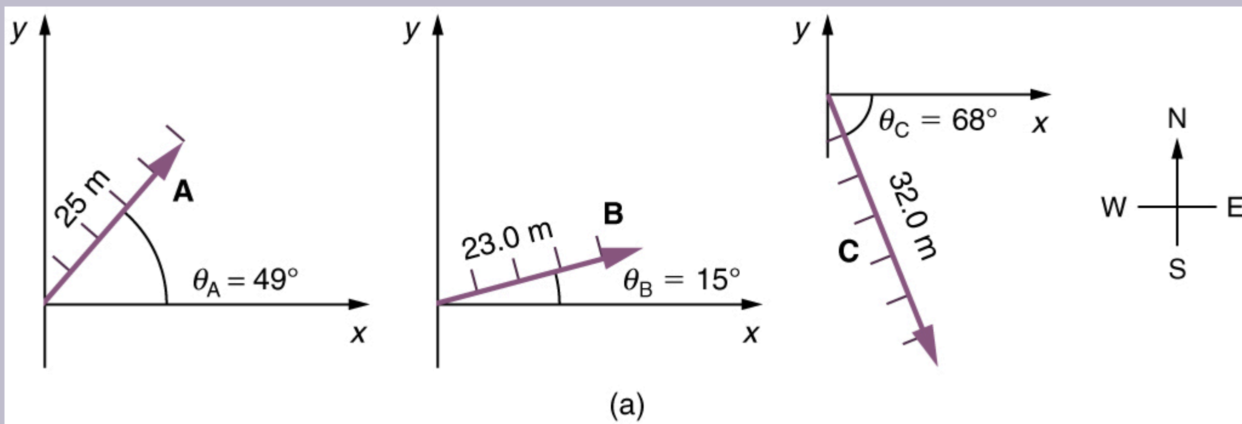


Figure 8.

(2) Place the vectors head to tail retaining both their initial magnitude and direction.

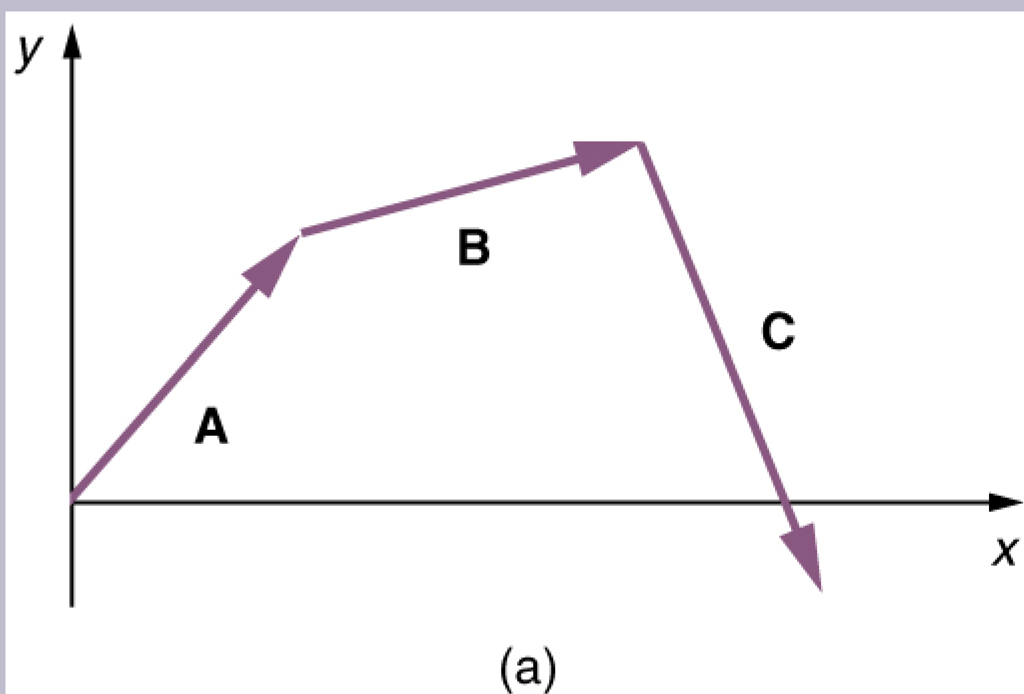


Figure 9.

(3) Draw the resultant vector, $\mathbf{R}$.

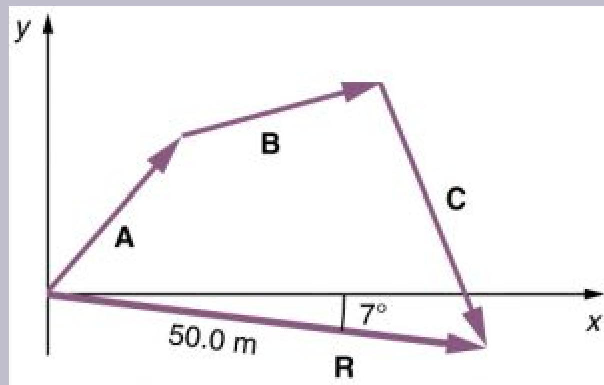


Figure 10.

(4) Use a ruler to measure the magnitude of $\mathbf{R}$, and a protractor to measure the direction of $\mathbf{R}$. While the direction of the vector can be specified in many ways, the easiest way is to measure the angle between the vector and the nearest horizontal or vertical axis. Since the resultant vector is south of the eastward pointing axis, we flip the protractor upside down and measure the angle between the eastward axis and the vector.

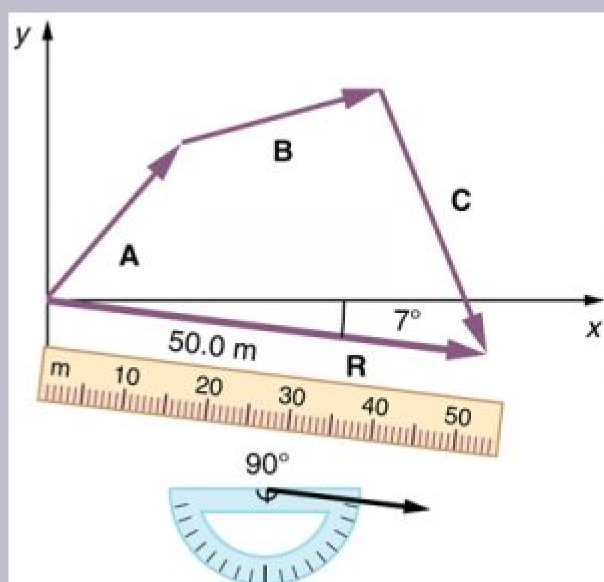


Figure 11.

In this case, the total displacement $\mathbf{R}$ is seen to have a magnitude of 50.0 m and to lie in a direction $7.0°$ south of east. By using its magnitude and direction, this vector can be expressed as $R = 50.0$ meters and $\theta = 7.0°$ south of east.

Discussion

The head-to-tail graphical method of vector addition works for any number of vectors. It is also important to note that the resultant is independent of the order in which the vectors are added. Therefore, we could add the vectors in any order as illustrated in Figure 12 and we will still get the same solution.

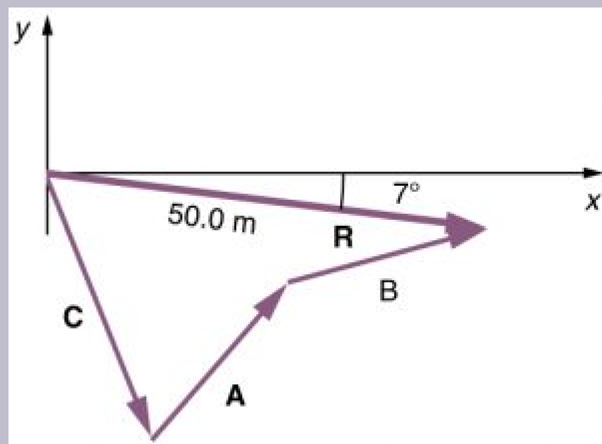


Figure 12.

Here, we see that when the same vectors are added in a different order, the result is the same. This characteristic is true in every case and is an important characteristic of vectors. Vector addition is **commutative**. Vectors can be added in any order.

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$

(This is true for the addition of ordinary numbers as well—you get the same result whether you add $2+3$ or $3+2$ for example).

Play with Vectors

In the simulation below, choose "Explore 2-D" you can then

- Move vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ on to the grid
- Change their magnitude and direction by clicking on the tip and dragging it around.
- Manipulate the actual size using the values at the top.
- Turn on using the menus on the right:
    - The components
    - Angles
    - Values
- Have the simulation draw the sum for you checking the tip-to-tail method

An interactive or media element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=241

Vector Subtraction

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Understanding vector subtraction is necessary to understand other physics ideas. For example, acceleration is $\frac{\Delta v}{\Delta t}$, and Δv is $v_f - v_i$. Velocity is a vector, so you're looking at a vector subtraction whenever you're working with acceleration.

Vector subtraction is a straightforward extension of vector addition. To define subtraction (say we want to subtract $\mathbf{B}$ from $\mathbf{A}$, written $\mathbf{A} - \mathbf{B}$, we must first define what we mean by subtraction. The *negative* of a vector $\mathbf{B}$ is defined to be $-\mathbf{B}$; that is, graphically *the negative of any vector has the same magnitude but the opposite direction*, as shown in Figure 13. In other words, $\mathbf{B}$ has the same length as $-\mathbf{B}$, but points in the opposite direction. Essentially, we just flip the vector so it points in the opposite direction.

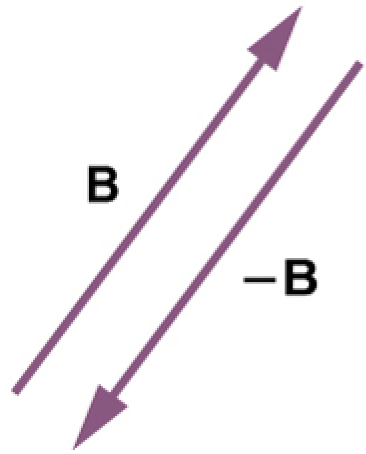


Figure 13. *The negative of a vector is just another vector of the same magnitude but pointing in the opposite direction. So $\mathbf{B}$ is the negative of $\mathbf{B}$; it has the same length but opposite direction.*

The *subtraction* of vector $\mathbf{B}$ from vector $\mathbf{A}$ is then simply defined to be the addition of $-\mathbf{B}$ to $\mathbf{A}$. Note that vector subtraction is the addition of a negative vector. The order of subtraction does not affect the results.

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$$

This is analogous to the subtraction of scalars (where, for example, $5 - 2 = 5 + (-2)$. Again, the result is independent of the order in which the subtraction is made. When vectors are subtracted graphically, the techniques outlined above are used, as the following example illustrates.

Subtracting Vectors Graphically: A Woman Sailing a Boat

A woman sailing a boat at night is following directions to a dock. The instructions read to first sail 27.5 m in a direction 66.0° north of east from her current location, and then travel 30.0 m in a direction 112° north of east (or 22.0° west of north). If the woman makes a mistake and travels in the *opposite* direction for the second leg of the trip, where will she end up? Compare this location with the location of the dock.

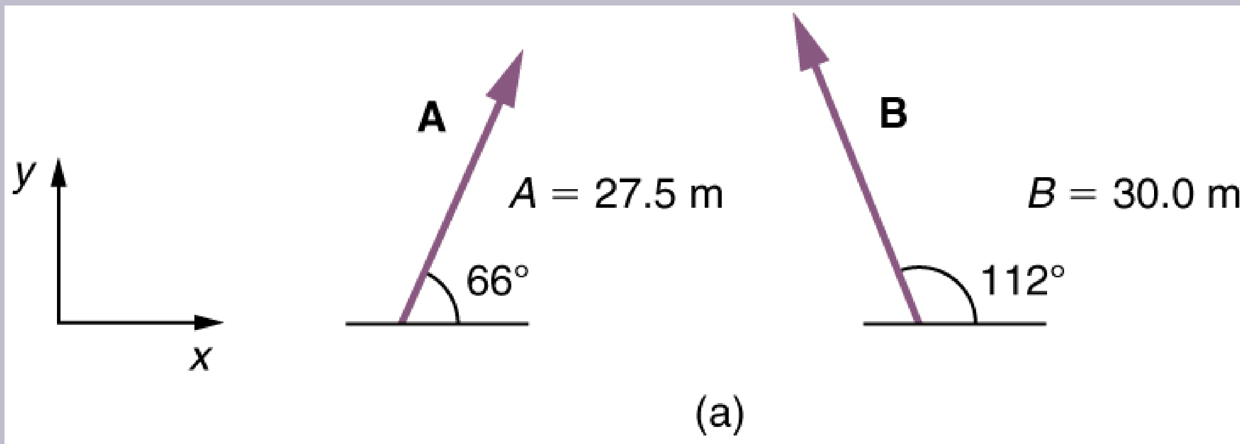


Figure 14.

Strategy

We can represent the first leg of the trip with a vector $\mathbf{A}$, and the second leg of the trip with a vector $\mathbf{A}$. The dock is located at a location $\mathbf{A} + \mathbf{B}$. If the woman mistakenly travels in the *opposite* direction for the second leg of the journey, she will travel a distance $\mathbf{A}$ (30.0 m) in the direction $180° - 112° = 68°$ south of east. We represent this as $-\mathbf{B}$, as shown below. The vector $-\mathbf{B}$ has the same magnitude as $\mathbf{B}$ but is in the opposite direction. Thus, she will end up at a location $\mathbf{A} + (-\mathbf{B})$ or $\mathbf{A} - \mathbf{B}$.

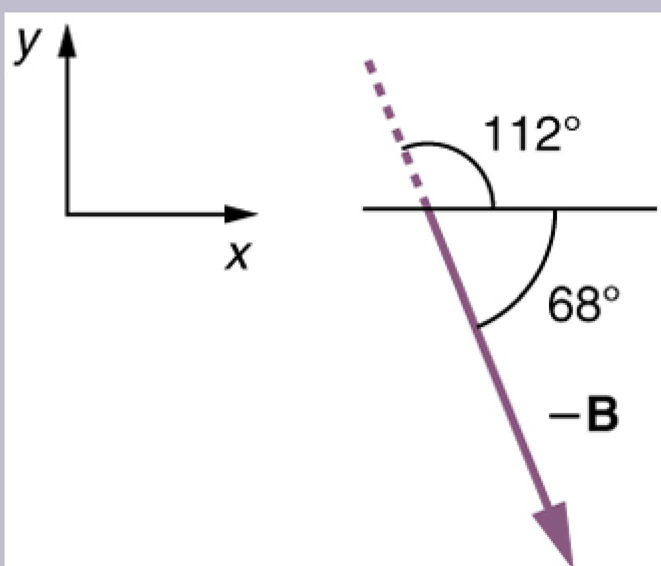


Figure 15.

We will perform vector addition to compare the location of the dock, $\mathbf{A} + \mathbf{B}$, with the location at which the woman mistakenly arrives, $\mathbf{A} + (-\mathbf{B})$.

Solution

(1) To determine the location at which the woman arrives by accident, draw vectors $\mathbf{A}$ and $-\mathbf{B}$.

(2) Place the vectors head to tail.

(3) Draw the resultant vector $\mathbf{R}$.

(4) Use a ruler and protractor to measure the magnitude and direction of $\mathbf{R}$.

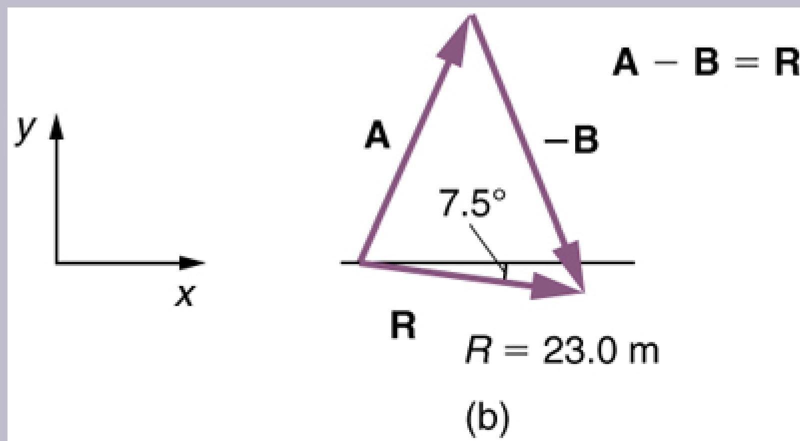


Figure 16.

In this case, $R = 23.0\mathrm{m}$ and $\theta < spanstyle = "font - size : 1em; text - align : initial" >= 7.5^{\circ}$ = 7.5^{\circ} " title="Rendered by QuickLaTeX.com" height="16" width="488" style="vertical-align: -3px;"> south of east.

(5) To determine the location of the dock, we repeat this method to add vectors $\mathbf{A}$ and $\mathbf{B}$. We obtain the resultant vector $\mathbf{R}'$:

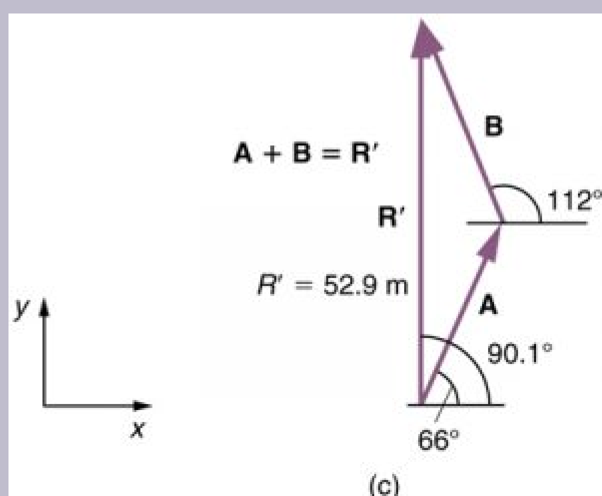


Figure 17.

In this case $R = 52.9\mathrm{m}$ and $\theta < spanstyle = "font - size : 1em; text - align : initial" >= 90.1^{\circ}$ = 90.1^{\circ} " title="Rendered by QuickLaTeX.com" height="15" width="496" style="vertical-align: -3px;"> north of east.

We can see that the woman will end up a significant distance from the dock if she travels in the opposite direction for the second leg of the trip.

Discussion

Because subtraction of a vector is the same as addition of a vector with the opposite direction, the graphical method of subtracting vectors works the same as for addition.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

If you've taken another physics course, you've probably seen $\mathbf{F} = m\mathbf{a}$. This equation will play a significant role in this class, and you'll notice that mass is a scalar, and acceleration is a vector, so understanding how scalars and vectors multiply will be important.

Multiplication of Vectors and Scalars

If we decided to walk three times as far on the first leg of the trip considered in the preceding example, then we would walk $3 \times 27.5\text{m}$ or 82.5 m, in a direction $66°$ north of east. This is an example of multiplying a vector by a positive **scalar**. Notice that the magnitude changes, but the direction stays the same.

If the scalar is negative, then multiplying a vector by it changes the vector's magnitude and gives the new vector the *opposite* direction. For example, if you multiply by –2, the magnitude doubles but the direction changes. We can summarize these rules in the following way: When vector $\mathbf{A}$ is multiplied by a scalar c,

- The magnitude of the vector becomes the absolute value of $c \cdot \mathbf{A}$
- If c is positive, the direction of the vector does not change
- If c is negative, the direction is reversed.

In our case, $c = 3$ and $A = 27.5\text{m}$. Vectors are multiplied by scalars in many situations. Note that division is the inverse of multiplication. For example, dividing by 2 is the same as multiplying by the value $\frac{1}{2}$. The rules for multiplication of vectors by scalars are the same for division; simply treat the divisor as a scalar between 0 and 1.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

When dealing with vectors using analytic methods (which is covered in the next section), you need to break down vectors into essentially x-components and y-components. This next part covers this idea, so try to familiarize yourself with breaking down vectors as you read.

Resolving a Vector into Components

In the examples above, we have been adding vectors to determine the resultant vector. In many cases, however, we will need to do the opposite. We will need to take a single vector and find what other vectors added together produce it. In most cases, this involves determining the perpendicular **components** of a single vector, for example the *x– and y*-components, or the north-south and east-west components.

For example, we may know that the total displacement of a person walking in a city is 10.3 blocks in a direction 29.0° north of east and want to find out how many blocks east and north had to be walked. This method is called *finding the components (or parts)* of the displacement in the east and north directions, and it is the inverse of the process followed to find the total displacement. It is one example of finding the components of a vector. There are many applications in physics where this is a useful thing to do. We will see this soon in Projectile Motion, and much more when we cover **forces** in Dynamics: Newton's Laws of Motion. Most of these involve finding components along perpendicular axes (such as north and east), so that right triangles are involved. The analytical techniques presented in Vector Addition and Subtraction: Analytical Methods are ideal for finding vector components.

Section Summary

- The **graphical method of adding vectors** $\mathbf{A}$ and $\mathbf{B}$ involves drawing vectors on a graph and adding them using the head-to-tail method. The resultant vector $\mathbf{R}$is defined such that $\mathbf{A}+\mathbf{B}=\mathbf{R}$. The magnitude and direction of $\mathbf{R}$ are then determined with a ruler and protractor, respectively.
- The **graphical method of subtracting vector** $\mathbf{B}$ from $\mathbf{A}$ involves adding the opposite of vector $\mathbf{B}$, which is defined as $-\mathbf{B}$. In this case, A–B=A+(–B)=RA–B=A+(–B)=R">A–B=A+(–B)=RA–B=A+(–B)=R">A–B=A+(–B)=R. Then, the head-to-tail method of addition is followed in the usual way to obtain the resultant vector $\mathbf{R}$.
- Addition of vectors is **commutative** such that $\mathrm{A}+\mathrm{B}=\mathrm{A}+\mathrm{B}$
- The head-to-tail method of adding vectors involves drawing the first vector on a graph and then placing the tail of each subsequent vector at the head of the previous vector. The resultant vector is then drawn from the tail of the first vector to the head of the final vector.
- If a vector $\mathbf{A}$ is multiplied by a scalar quantity c, the magnitude of the product is given by $|c \cdot \mathbf{A}|$. If c is positive, the direction of the product points in the same direction as $\mathbf{A}$; if c is negative, the direction of the product points in the opposite direction as $\mathbf{A}$

Vector Addition and Subtraction: Analytical Methods

Instructor's Note

Your Quiz would Cover

- Adding vectors by components. Don't focus too much on what it means to add vectors. Just learn the mechanics of how to do it. We will talk about the meaning in class.

Analytical methods of vector addition and subtraction employ geometry and simple trigonometry rather than the ruler and protractor of graphical methods. Part of the graphical technique is retained, because vectors are still represented by arrows for easy visualization. However, analytical methods are more concise, accurate, and precise than graphical methods, which are limited by the accuracy with which a drawing can be made. Analytical methods are limited only by the accuracy and precision with which physical quantities are known.

Resolving a Vector into Perpendicular Components

Analytical techniques and right triangles go hand-in-hand in physics because (among other things) motions along perpendicular directions are independent. We very often need to separate a vector into perpendicular components. For example, given a vector like $\mathbf{A}$in Figure 1, we may wish to find which two perpendicular vectors, $\mathbf{A}_x$ and $\mathbf{A}_y$, add to produce it.

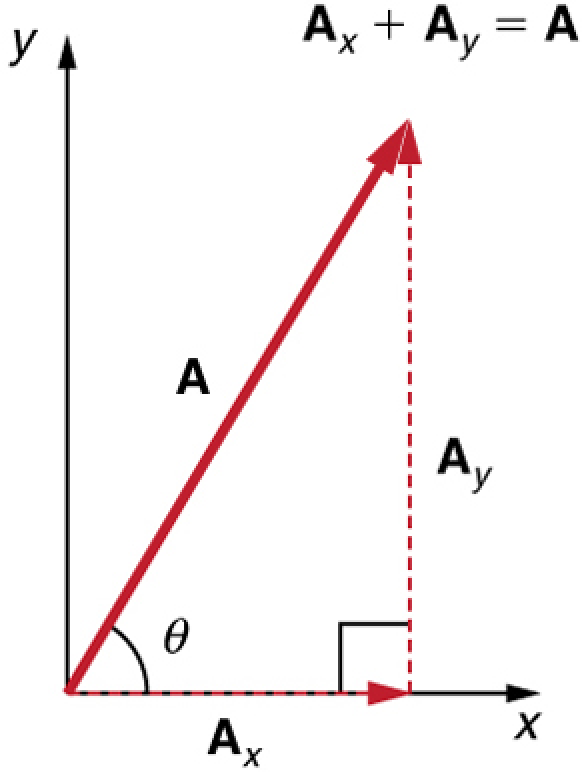


Figure 1. *The vector A, with its tail at the origin of an x, y-coordinate system, is shown together with its x- and y-components, Ax and Ay. These vectors form a right triangle. The analytical relationships among these vectors are summarized below.*

$\mathbf{A}_x$ and $\mathbf{A}_y$, are defined to be the components of $\mathbf{A}$ along the *x*– and *y*-axes. The three vectors $\mathbf{A}$, $\mathbf{A}_x$ and $\mathbf{A}_y$, form a right triangle:

$$\mathbf{A}_x + \mathbf{A}_y = \mathbf{A}$$

Note that this relationship between vector components and the resultant vector holds only for vector quantities (which include both magnitude and direction). The relationship does not apply for the magnitudes alone. For example, if $\mathbf{A}_x = 3\mathrm{m}$ east, $\mathbf{A}_y = 4\mathrm{m}$ north, and $\mathbf{A} = 5\mathrm{m}$ north-east, then it is true that the vectors $\mathbf{A}_x + \mathbf{A}_y = \mathbf{A}$. However, it is *not* true that the sum of the magnitudes of the vectors is also equal. That is,

$$3\mathrm{m} + 4\mathrm{m} \neq 5\mathrm{m}$$

Thus,

$$A_x + A_y \neq A$$

If the vector $\mathbf{A}$ is known, then its magnitude A(its length) and its angle θ (its direction) are known. To find A_x and A_y, its *x*– and *y*-components, we use the following relationships for a right triangle.

$$A_x = A\cos\theta$$

and

$$A_y = A\sin\theta.$$

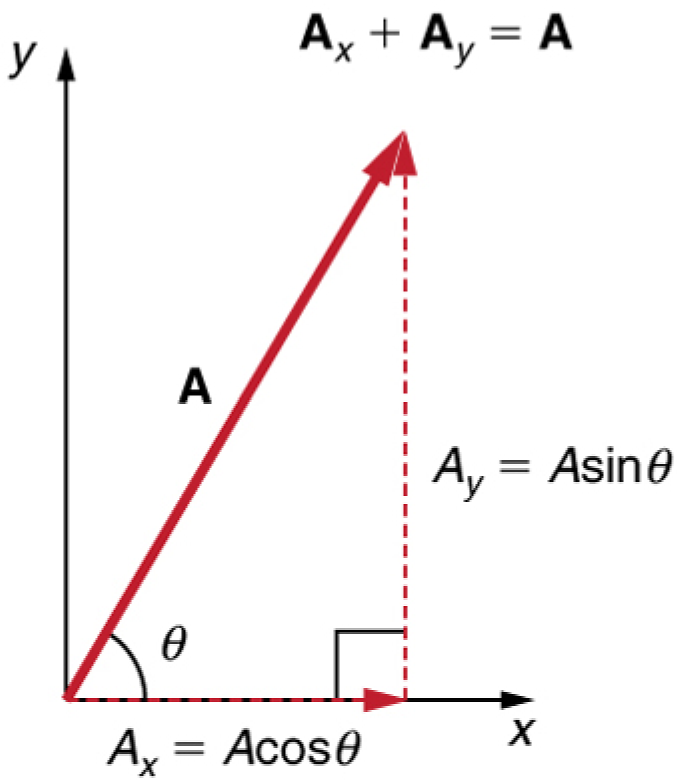


Figure 2. *The magnitudes of the vector components A_x and A_y can be related to the resultant vector $\mathbf{A}$ and the angle θ with trigonometric identities. Here we see that $A_x = A\cos\theta$ and $A_y = A\sin\theta$.*

Suppose, for example, that $\mathbf{A}$ is the vector representing the total displacement of the person walking in a city considered in Kinematics in Two Dimensions: An Introduction and Vector Addition and Subtraction: Graphical Methods.

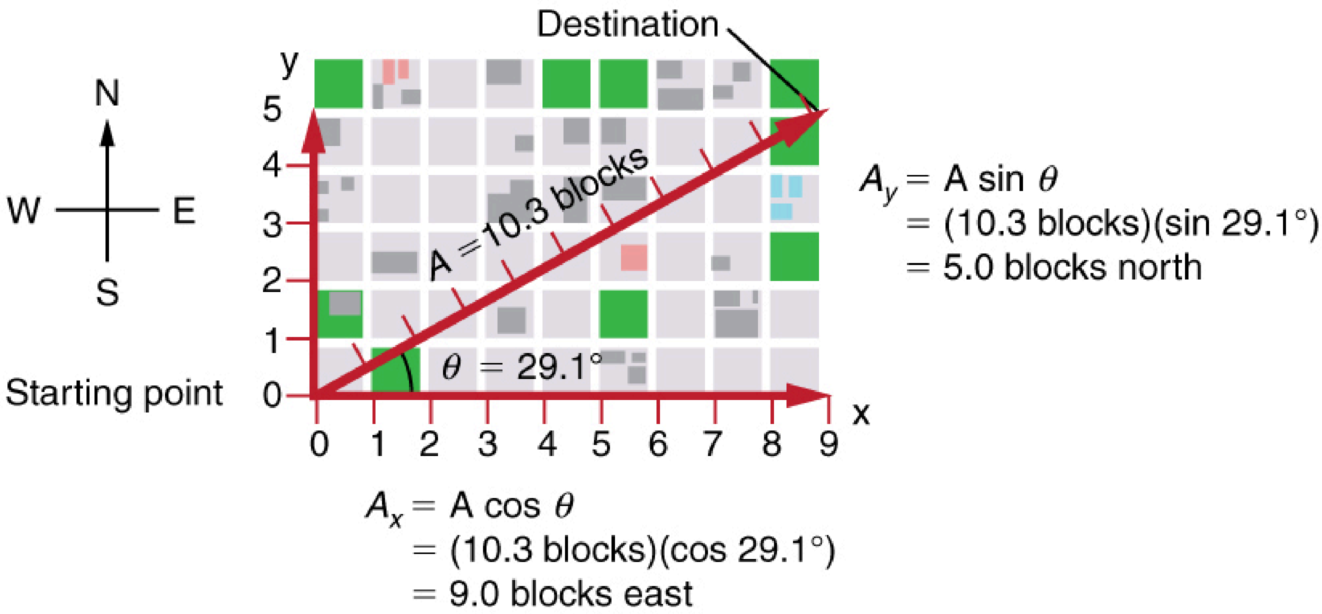


Figure 3. *We can use the relationships $A_x = A\cos\theta$ and $A_y = A\sin\theta$ to determine the magnitude of the horizontal and vertical component vectors in this example.*

Then $A = 10.3$blocks and $\theta = 29.6°$, so that

$$A_x = A\cos\theta = (10.3\text{ blocks})(\cos 29.1°) = 9.0\text{ blocks}$$
$$A_y = A\sin\theta = (10.3\text{ blocks})(\cos 29.1°) = 5.0\text{ blocks}$$

Calculating a Resultant Vector

If the perpendicular components $\mathbf{A}_x$ and $\mathbf{A}_y$AyAy size 12{A rSub { size 8{y} } } {}"> of a vector are known, then $\mathbf{A}$AA size 12{A} {}"> AA size 12{A} {}"> can also be found analytically. To find the magnitude AAA size 12{A} {}"> and direction θ of a vector from its perpendicular components $\mathbf{A}_x$ and $\mathbf{A}_y$, we use the following relationships:θ=tan−1(Ay/Ax).θ=tan−1(Ay/Ax). size 12{θ="tan" rSup { size 8{ - 1} } \(A rSub { size 8{y} } /A rSub { size 8{x} } \) } {}">

$$A = \sqrt{A_x^2 + A_y^2}$$
$$\theta = \tan^{-1}\frac{A_y}{A_x}.$$

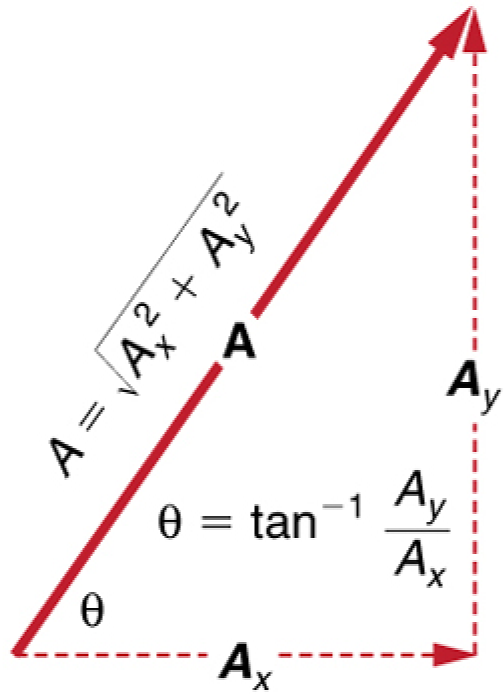


Figure 4. *The magnitude and direction of the resultant vector can be determined once the horizontal and vertical components* $\mathbf{A}_x$ *and* $\mathbf{A}_y$AyAy size 12{A rSub { size 8{y} } } {}"> *have been determined.*

Note that the equation $A = \sqrt{A_x^2 + A_y^2}$ is just the Pythagorean theorem relating the legs of a right triangle to the length of the hypotenuse. For example, if AxAx size 12{A rSub { size 8{x} } } {}"> A_x and A_y are 9 and 5 blocks, respectively, then $A = \sqrt{9^2 + 5^2} = 10.3\text{blocks}$, again consistent with the example of the person walking in a city. Finally, the direction is $\theta = \tan^{-1} \frac{5}{9} = 29.1°$, as before.

Determining Vectors and Vector Components with Analytical Methods

Equations $A_x = A \cos\theta$ and $A_y = A \sin\theta$ are used to find the perpendicular components of a vector—that is, to go from $\mathbf{A}$ and θ to $\mathbf{A}_x$ and $\mathbf{A}_y$AyAy size 12{A rSub { size 8{y} } } {}"> . Equations $A = \sqrt{A_x^2 + A_y^2}$ and $\theta = \tan^{-1} \frac{A_y}{A_x}$ are used to find a vector from its perpendicular components—that is, to go from $\mathbf{A}_x$ and $\mathbf{A}_y$ to $\mathbf{A}$ and θ. Both processes are crucial to analytical methods of vector addition and subtraction.

Instructor's Note

Now that you know how to break down vectors into components, here's a procedure to adding vectors analytically. There's some trigonometry involved, so, again, if you're not familiar or comfortable with trigonometry, come see your instructor. You should be familiar with both methods. you should be able to add two vectors given their x and y components, and you should be able to draw the resulting vector of two added vectors. Also, we will go over how to use these to solve problems, so focus primarily on the methods of adding vectors.

Adding Vectors Using Analytical Methods

To see how to add vectors using perpendicular components, consider Figure 5, in which the vectors $\mathbf{A}$ and $\mathbf{B}$ are added to produce the resultant $\mathbf{R}$.

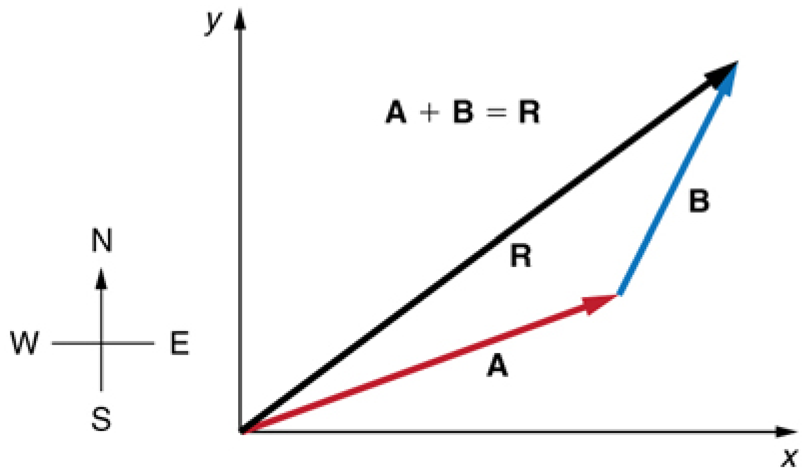


Figure 5. *Vectors $\mathbf{A}$ and $\mathbf{B}$ are two legs of a walk, and $\mathbf{R}$ is the resultant or total displacement. You can use analytical methods to determine the magnitude and direction of $\mathbf{R}$.*

If $\mathbf{A}$ and $\mathbf{B}$ represent two legs of a walk (two displacements), then $\mathbf{R}$ is the total displacement. The person taking the walk ends up at the tip of $\mathbf{R}$. There are many ways to arrive at the same point. In particular, the person could have walked first in the *x*-direction and then in the *y*-direction. Those paths are the *x*– and *y*-components of the resultant, $\mathbf{R}_x$ and $\mathbf{R}_y$. If we know $\mathbf{R}_x$ and $\mathbf{R}_y$, we can find $\mathbf{R}$ and θ using the equations $A = \sqrt{A_x^2 + A_y^2}$ and $\theta = \tan^{-1}\frac{A_y}{A_x}$. When you use the analytical method of vector addition, you can determine the components or the magnitude and direction of a vector.

Step 1. *Identify the x- and y-axes that will be used in the problem. Then, find the components of each vector to be added along the chosen perpendicular axes. Use the equations $A_x = A\cos\theta$ and $A_y = A\sin\theta$ to find the components. In Figure 6, these components are $\mathbf{A}_x$, $\mathbf{A}_y$, $\mathbf{B}_x$ and $\mathbf{B}_y$. The angles that vectors $\mathbf{A}$ and $\mathbf{B}$ make with the x-axis are θ_A and θ_B, respectively.*

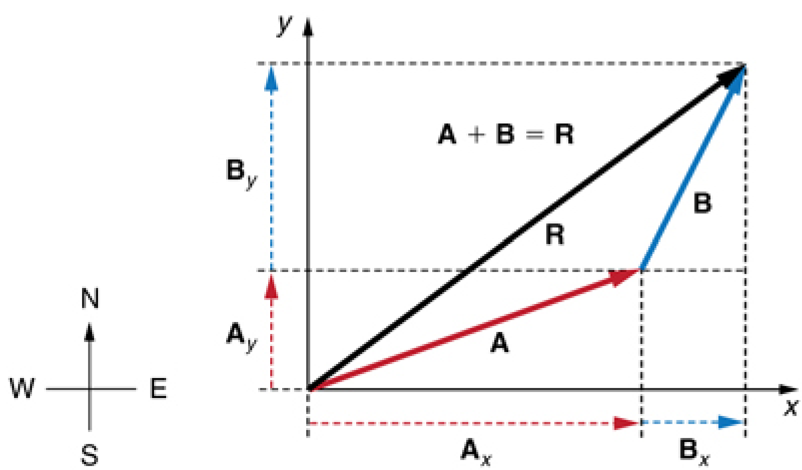


Figure 6. *To add vectors $\mathbf{A}$ and $\mathbf{B}$, first determine the horizontal and vertical components of each vector. These are the dotted vectors $\mathbf{A}_x$, $\mathbf{A}_y$, $\mathbf{B}_x$ and $\mathbf{B}_y$ shown in the image.*

Step 2. *Find the components of the resultant along each axis by adding the components of the individual vectors along that axis. That is, as shown in Figure 7,Ry=Ay+By.Ry=Ay+By. size 12{R rSub { size 8{y} } =A rSub { size 8{y} } +B rSub { size 8{y} } } {}">*

$$R_x = A_x + B_x$$
$$R_y = A_y + B_y.$$

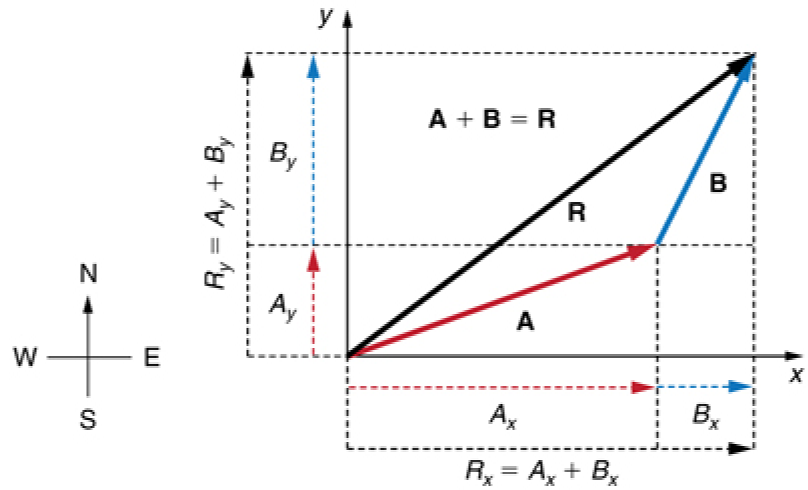


Figure 7. *The magnitude of the vectors $\mathbf{A}_x$ and $\mathbf{B}_x$ add to give the magnitude R_x of the resultant vector in the horizontal direction. Similarly, the magnitudes of the vectors $\mathbf{A}_y$ and $\mathbf{B}_y$add to give the magnitude R_y of the resultant vector in the vertical direction.*

Components along the same axis, say the x-axis, are vectors along the same line and, thus, can be added to one another like ordinary numbers. The same is true for components along the y-axis. (For example, a 9-block eastward walk could be taken in two legs, the first 3 blocks east and the second 6 blocks east, for a total of 9, because they are along the same direction.) So resolving vectors into components along common axes makes it easier to add them. Now that the components of $\mathbf{R}$ are known, its magnitude and direction can be found.

Step 3. *To get the magnitude R of the resultant, use the Pythagorean theorem:*

$$A = \sqrt{A_x^2 + A_y^2}$$

Step 4. *To get the direction of the resultant:*

$$\theta = \tan^{-1} \frac{A_y}{A_x}.$$

The following example illustrates this technique for adding vectors using perpendicular components.

Adding Vectors Using Analytical Methods

Add the vector $\mathbf{A}$ to the vector $\mathbf{B}$ shown in Figure 8, using perpendicular components along the *x*– and *y*-axes. The *x*– and *y*-axes are along the east–west and north–south directions, respectively. Vector

$\mathbf{A}$ represents the first leg of a walk in which a person walks $53.0 \mathrm{m}$ in a direction 20.0° north of east. Vector $\mathbf{B}$ represents the second leg, a displacement of $34.0 \mathrm{m}$ in a direction 63.0° north of east.

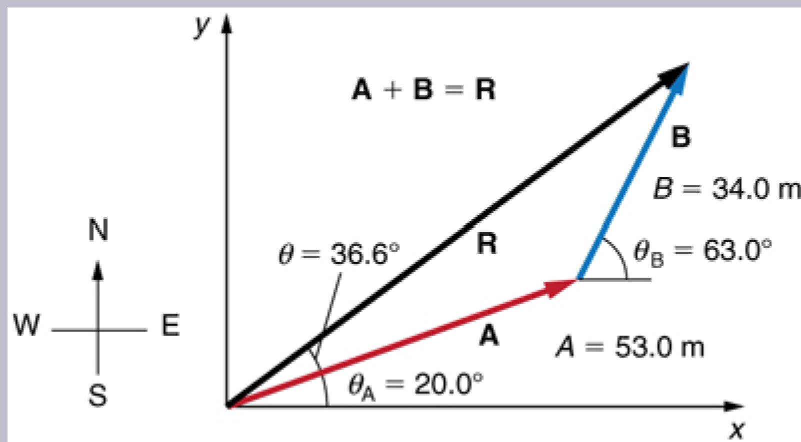


Figure 8. *Vector $\mathbf{A}$ has magnitude $53.0 \mathrm{m}$ and direction 20.0° north of the x-axis. Vector $\mathbf{B}$ has magnitude $34.0 \mathrm{m}$ and direction 63.0° north of the x-axis. You can use analytical methods to determine the magnitude and direction of $\mathbf{R}$.*

Strategy

The components of $\mathbf{A}$ and $\mathbf{B}$ along the *x*– and *y*-axes represent walking due east and due north to get to the same ending point. Once found, they are combined to produce the resultant.

Solution

Following the method outlined above, we first find the components of $\mathbf{A}$ and $\mathbf{B}$ along the *x*– and *y*-axes. Note that $A = 53.0 \mathrm{m}$, $\theta_A = 20.0^{\circ}$A=53.0 mA=53.0 m size 12{"A" "=" "53.0 m"} {}">, $B = 34.0 \mathrm{m}$, and $\theta_B = 63.0^{\circ}$. We find the *x*-components by using $A_x = A \cos \theta_A$, which gives:

$$A_x = A \cos \theta_A = (53.0 \text{ m})(\cos 20.0^{\circ})$$
$$A_x = (53.0 \text{ m})(0.940) = 49.8 \text{ m}$$

and

$$B_x = B \cos \theta_B = (34.0 \text{ m})(\cos 63.0^{\circ})$$
$$B_x = (34.0 \text{ m})(0.545) = 15.4 \text{ m}$$

Similarly, the *y*-components are found using $A_y = A \sin \theta_A$:

$$A_y = A \sin \theta_A = (53.0 \text{ m})(\sin 20.0^{\circ})$$
$$A_y = (53.0 \text{ m})(0.342) = 18.1 \text{ m}$$

and

$$B_y = B \sin \theta_B = (34.0 \text{ m})(\sin 63.0^{\circ})$$
$$B_y = (34.0 \text{ m})(0.891) = 30.3 \text{ m}$$

-components of the resultant are thus

$$R_x = A_x + B_x = 49.8 \text{ m} + 15.4 \text{ m} = 65.2 \text{ m}$$

and

$$R_y = A_y + B_y = 18.1\text{m} + 30.3\text{m} = 48.4\text{m}$$

Now we can find the magnitude of the resultant by using the Pythagorean theorem:

$$R = \sqrt{R_x^2 + R_y^2} = \sqrt{(65.2)^2 + (48.4)^2}$$

so that

$$R = 81.2\text{m}$$

Finally, we find the direction of the resultant:

$$\theta = \tan^{-1}(R_y/R_x) = +\tan^{-1}(48.4/65.2)$$

Thus,

$$\theta = \tan^{-1}(0.742) = 36.6°$$

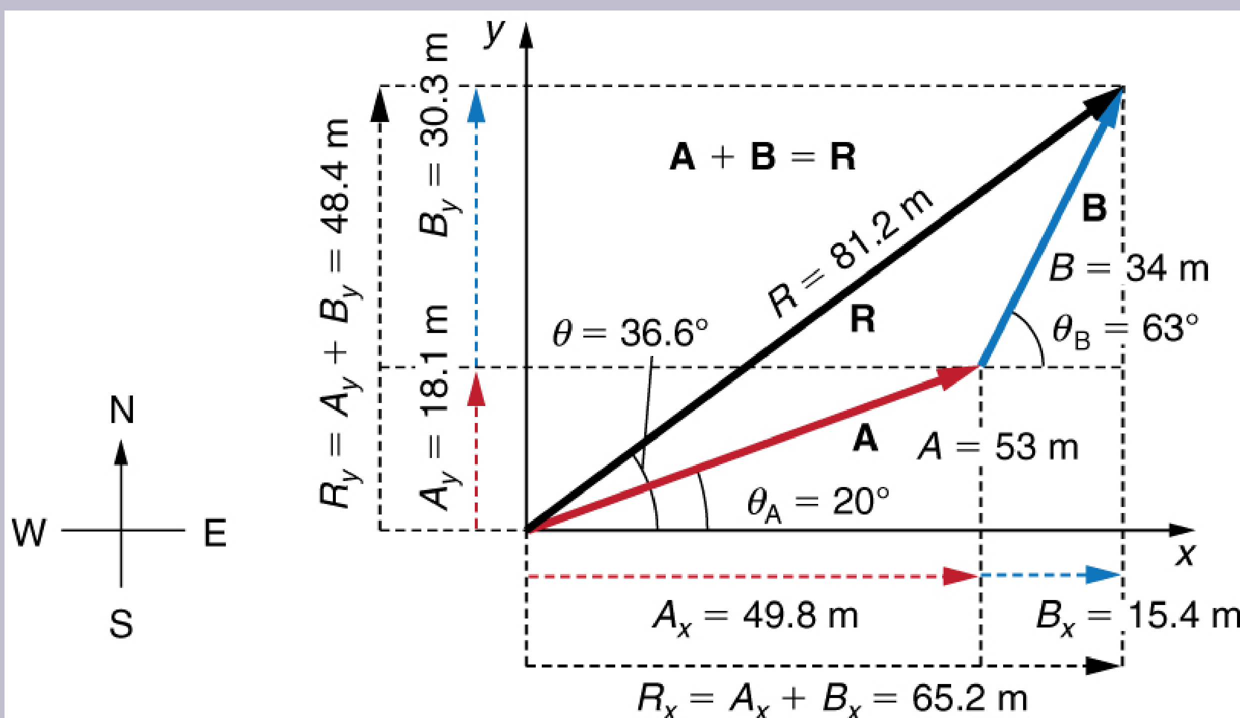


Figure 9. *Using analytical methods, we see that the magnitude of R is* 81.2 *m and its direction is* $36.6°$ *north of east.*

Discussion

This example illustrates the addition of vectors using perpendicular components. Vector subtraction using perpendicular components is very similar—it is just the addition of a negative vector.

Subtraction of vectors is accomplished by the addition of a negative vector. That is, $\mathbf{A} - \mathbf{B} \equiv \mathbf{A} + (-\mathbf{B})$. Thus, *the method for the subtraction of vectors using perpendicular* *components is identical to that for addition*. The components of $-\mathbf{B}$ are the negatives of the components of $\mathbf{B}$. The *x*– and *y*-components of the resultant $\mathbf{A}-\mathbf{B}=\mathbf{R}$ are thus

$$R_x = A_x + (-B_x)$$

and

$$R_y = A_y + (-B_y)$$

and the rest of the method outlined above is identical to that for addition. (See Figure 10.)

Analyzing vectors using perpendicular components is very useful in many areas of physics, because perpendicular quantities are often independent of one another. The next module, Projectile Motion, is one of many in which using perpendicular components helps make the picture clear and simplifies the physics.

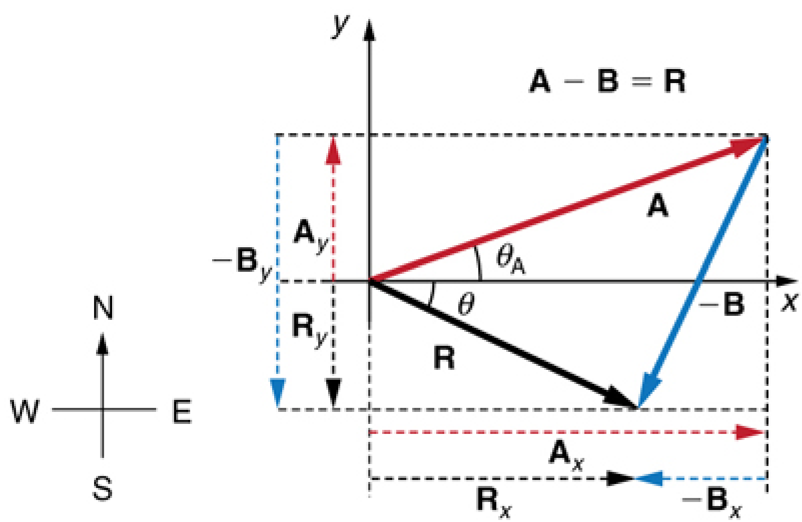


Figure 10. *The subtraction of the two vectors shown in Figure. The components of* $-\mathbf{B}$ *are the negatives of the components of* $\mathbf{B}$. *The method of subtraction is the same as that for addition.*

More PhET Explorations: Vector Addition

This is the same simulation as above. However, now, I would recommend you look at the "Equations" option!

—

An interactive or media element has been excluded from this version of the text. You can view it online here:
http://openbooks.library.umass.edu/toggerson-132/?p=241

Section Summary

- The analytical method of vector addition and subtraction involves using the Pythagorean theorem and trigonometric identities to determine the magnitude and direction of a resultant vector.
- The steps to add vectors $\mathbf{A}$ and $\mathbf{B}$ using the analytical method are as follows:

1. Determine the coordinate system for the vectors. Then, determine the horizontal and vertical components of each vector using the equations:

   $A_x = A\cos\theta$ and $B_x = B\cos\theta$
   $A_y = A\sin\theta$ and $B_y = B\sin\theta$.

2. Add the horizontal and vertical components of each vector to determine the components R_x and R_y of the resultant vector, $\mathbf{R}$:

   $R_x = A_x + B_x$
   $R_y = A_y + B_y$.

3. Use the Pythagorean theorem to determine the magnitude, R, of the resultant vector $\mathbf{R}$:
   $R = \sqrt{R_x^2 + R_y^2}$.

4. Use a trigonometric identity to determine the direction, θ, of $\mathbf{R}$:
   $\theta = \tan^{-1}\frac{R_y}{R_x}$

Homework Problems

Homework Problems

Problem 4: Find the magnitudes of the forces F1 and F2 and that add to give the total force Ftotal shown above. This may be done either graphically or by using trigonometry.

Problem 5: Suppose you walk straight west and then straight south. How far are you from your starting point, and what is the compass direction of a line connecting your starting point to your final position?

22. Electric Fields

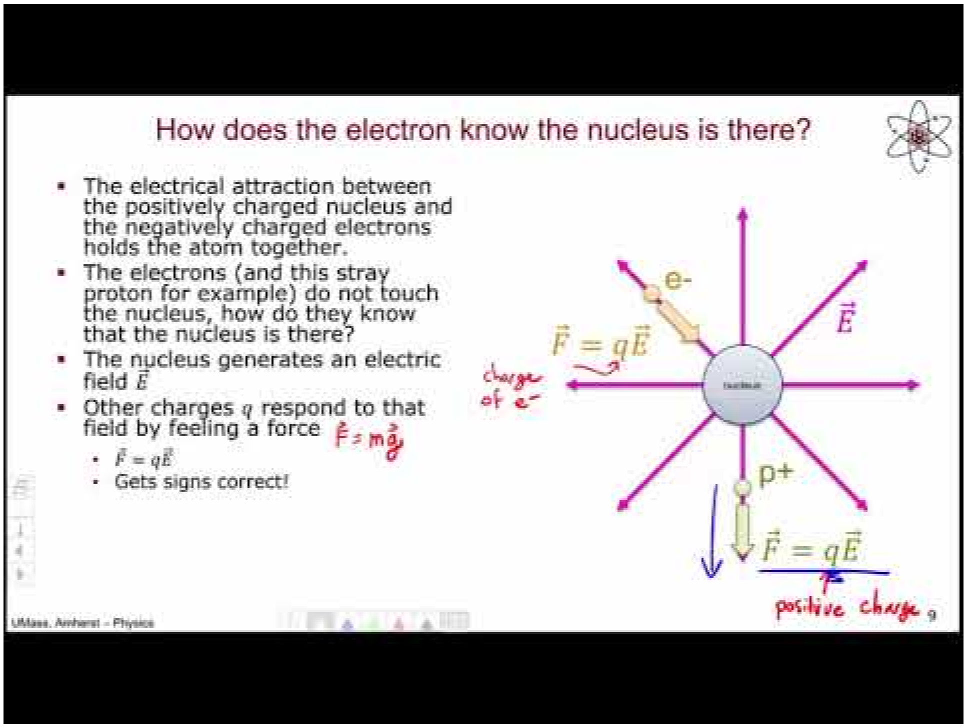
Introduction to Electric Field

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

By the end of this section you should be able to:

- Explain what a field is
- Explain how forces arise from objects interacting with field
- Justify the units of the electric field



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=243*](http://openbooks.library.umass.edu/toggerson-132/?p=243)

For the past few units we have been really focusing on the ideas of light, now, everything we've been talking about in our units on geometric optics and physical optics applies to matter waves as well. We've just really been focusing on the applications to light. Now here in Unit 3, we will return two electrons, and these will be the focus of our attention for the next two units. Up to now, we've been ignoring the fact that electrons have charge, and we know that they do, so it's time to fix that. By thinking about the charge attribute of electrons we will begin our study of the electric force.

To begin, I want to remind you from 131 that there are fundamentally only four forces: The strong nuclear force, which is responsible for holding all of the positively charged protons in the nucleus together. The electrical and magnetic forces, which we will see through this course are really two deeply connected sides of the same coin, and these are the ideas of opposite charges attract like charges repel and magnets.

Instructor's Note

While these may seem like wildly different phenomena at first glance. We'll see through this course that electricity and magnetism are deeply connected.

Next in strength we have the weak nuclear force responsible for radioactive decay, and then finally the gravitational force, which holds you to the Earth and holds the Earth in orbit around the Sun.

Gravity was talked about in some depth in Physics 131, whereas electricity and magnetism will consume our attention for the rest of 132. It's worth pointing out that the idea of opposite charges attract and like charges repel is the fundamental origin of all of the other forces taught in these courses. For example, in 131, you discuss the normal force, when a physics book sits on a table you have the weight of the book, but you also have a normal force from the table on the book, which keeps it from falling through the table. This force is perpendicular, or normal, to the surface between the table in the book and really arises because of the electrons in the table repelling the electrons in the book.

Another force discussed in 131 is the idea of tension. For example, when a block is just hung by a string, you have the weight of the block being countered by the tension in the rope, but where does this tension ultimately come from? This tension comes from the atomic bonds which are fundamentally the electrical attraction between one molecule of rope and the next. Similarly, the spring force and Hooke's law are a result of atomic bonds, which are electrical at the microscopic scale. Finally, the frictional forces arise from Van der Waals interactions between the molecules and different surfaces. Once again, you're talking about the electrical interactions between atoms.

Before we begin with the electric field we're going to explore everyone's favorite number from Physics 131, this

9.8

that was used extensively in that course, and we're going to deconstruct exactly what this number is. Gravity is a bit more familiar to us, as we experience it on a daily basis. And we will build up the idea of the electric field in parallel, using gravity as a crutch.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor’s Note

It is important however to keep in mind that gravity and electricity are fundamentally different forces.

Absolutely everything in the universe experiences gravity. We saw in class, through our gravitational redshift of light, that even light experiences gravity. However only those objects with charges will experience electrical forces. Can you perhaps think of something that experiences gravity but not electricity? You would need something that has mass but not an electric charge.

Let’s begin thinking about gravity. We know that the force of gravity is what holds the moon in orbit around the Earth, but how does the moon know that the Earth is there? I mean the moon is 3×10^8 meters away from the Earth. That’s a long way. How does the moon know that the Earth is even there? To simplify and bring it a little more down to Earth, no pun intended, we’re going to think about a ball. We all know that when I release a ball it falls due to the Earth’s gravity. We could ask the same question; how does the ball know that the Earth is there? They’re not in contact. Yes, the distances are a lot smaller, but they don’t touch each other. How does the ball know that the Earth is there?

The way that this is explained is that the Earth generates a gravitational field and this gravitational field is what the letter g from Physics 131 represents, and the strength of that gravitational field is 9.8 Newton’s per kilogram. Now you might be used to thinking about it being meters per second squared but if you look a little bit at Newton’s second law, $F = ma$, you will see that meters per second squared and Newton’s per kilogram are equivalent units. We’re going to think in terms of Newton’s per kilogram, because it’s a more useful unit for our purposes right now.

The Earth generates this gravitational field and it points straight down and it has a magnitude of 9.8 Newton’s per kilogram. Now you will notice that the ball does actually touch the Earth’s gravitational field, it doesn’t touch the Earth, but it does touch this gravitational field. And through this contact the ball reacts to that field by feeling a gravitational force, mg. The force is the mass of the ball, m, times the strength of the gravitational field, $\vec{g}$. That’s what gives rise to the force.

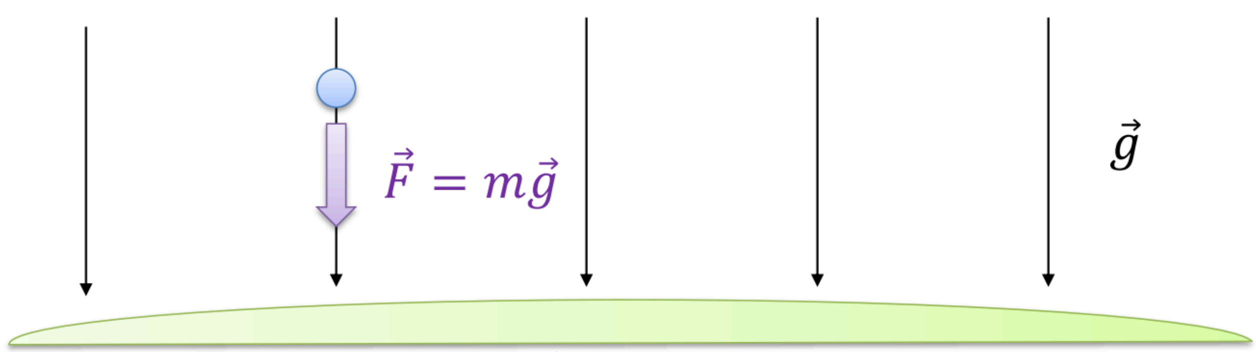


Figure 1. A ball in a gravitational field.

Now let's return to these units for $\vec{g}$, these Newton's per kilogram that you might not be used to thinking about. For every one kilogram of ball, the ball feels 9.8. Newton's of force. That's what $m\vec{g}$ means, a two-kilogram ball will feel 2 times 9.8 Newton's of force. That's why these units are somewhat useful for our purposes right now, is it tells us how many Newton's of force we get for each kilogram of ball, that's what 9.8 represents.

Why did we introduce this middle step of a gravitational field that's invisible and extends from the Earth? Why did we invent this? Why can't we just say that something causes the ball to fall and the magnitude of the force is $m\vec{g}$? That's what we did in 131.

Well the reason we invent this is because this field, while we cannot see, it is a fundamentally real object. It's just as real as a ball. Yes, it's invisible, but so are atoms, and you believe atoms are real so the field is a similar thing. It is fundamentally a real quantity, it contains energy, in fact we can think about the energy when we let the ball go. When the ball falls it gains kinetic energy, where does that energy come from? That energy comes from the gravitational field. This is a fundamentally different way of thinking a little more deeply about what gravity actually is and where this 9.8 Newton's per kilogram number comes from.

Now let's return to the idea of electricity, the main focus of our course. How does the electron know the nucleus is there? We know that it's the positively charged nucleus attracting the negatively charged electron that holds the atom together. But again, the nucleus and the electron do not touch, so how does the nucleus know that the electron is there? We're going to add a proton to this situation for purposes of illustration, but the same argument can be made. How does this stray proton know that the nucleus is there?

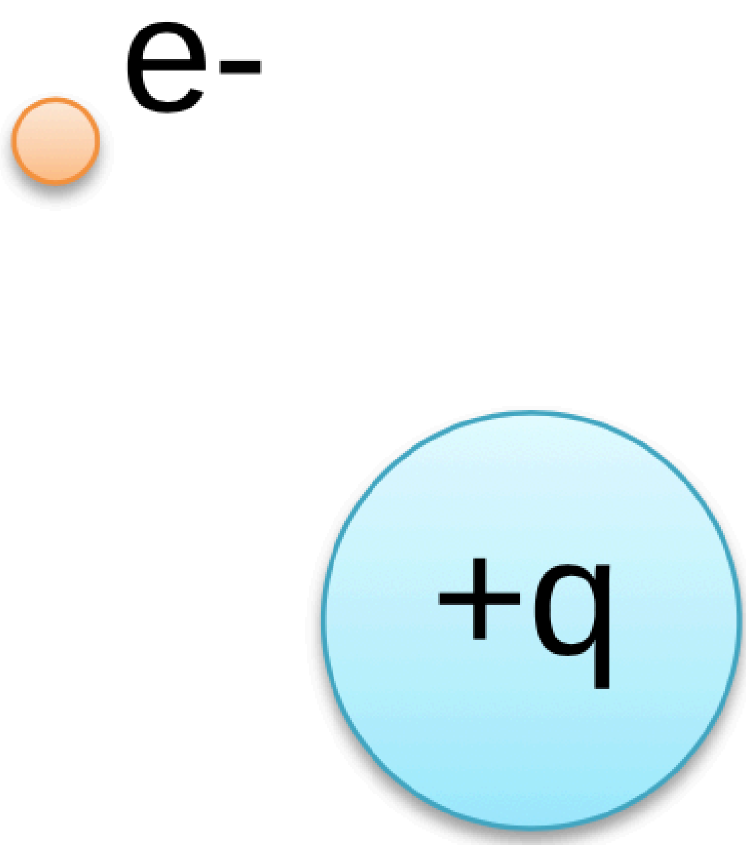


Figure 2. An electron and a proton. Thinking in terms of particles, how does the electron know the proton is there?

Well, we use the same sort of idea, the nucleus generates an electric field that we call $\vec{E}$ and then the other charges, this electron and striped proton, respond to that field by feeling a force.

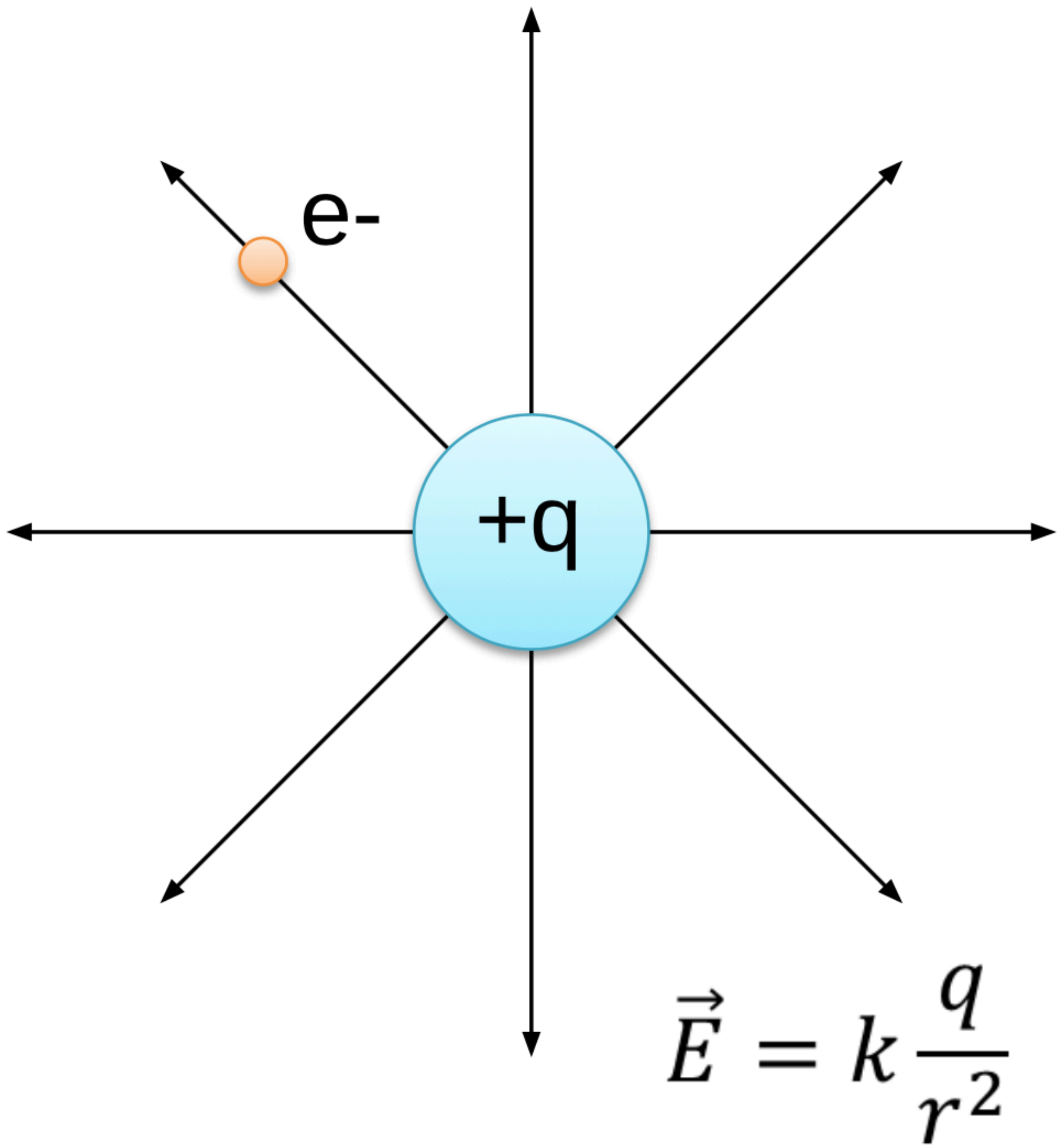


Figure 3. The electric field surrounding the proton.

The force felt by the electron is $\vec{F} = q\vec{E}$ and, with analogy to gravity, where the force was the mass times the gravitational field, the force felt by these other charges is going to be the charge times the electric field. The stray proton feels a force $q\vec{E}$, where q is the positive charge of proton, and similarly the electron also feels a force $q\vec{E}$, where now this qqq">q is the charge of the electron.

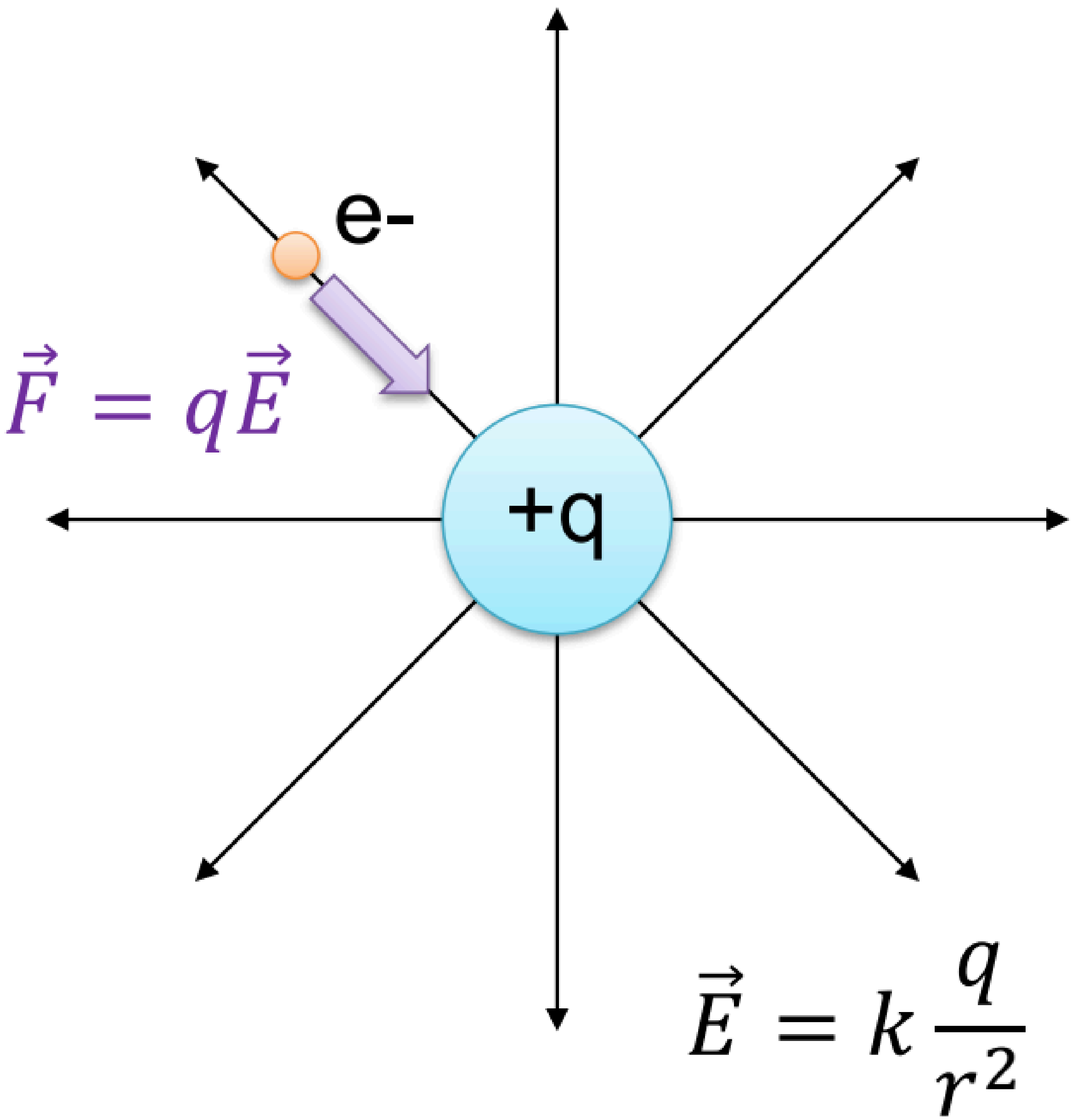


Figure 4. The forces felt by the electron and the proton.

Now it's worth noting that this expression gets the sign correct in the case of the proton. The charge is positive, and the force and the electric field are in the same direction. For the electron, the charge is negative and so the force and the electric field are in opposite directions, the force is inward, and the electric field is outwards. So, it's important to keep in mind that this expression actually does get the signs correct, in fact so does our expression for gravity, mass is always positive and so the force of gravity is always in the same direction as the gravitational field. Once we know the forces we can then go on to calculate accelerations using Newton's second law.

Let's do some examples, we'll do an example with gravity first, and then do one with electricity.

Gravitational Example

Let's say we have a $5\,\mathrm{kg}$ object with a charge of $2\,\mathrm{C}$ and it's sitting above the surface of the Moon where $\vec{g} = 1.6\,\mathrm{N}/\mathrm{kg}$. Different planetary body different gravitational field. What is the acceleration of this five-kilogram object?

We begin by thinking about what force is acting, the weight force. We know that the weight force is the mass of the object times the gravitational field, in this case: $5\,\mathrm{kg}$ times the $1.6\,\mathrm{N}/\mathrm{kg}$, giving us a weight force of $8\,\mathrm{N}$.

Then we can move on to calculating the acceleration using Newton's second law,

$$F = ma.$$

The only force here is weight, the $8\,\mathrm{N}$. We also know the mass of the ball $5\,\mathrm{kg}$, and we are left with the acceleration of $1.6\mathrm{m}/\mathrm{s}^2$.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Just as you might expect from 131 the acceleration and

g

are the same number. This results from the fact that both in the definition of the weight force and the definition of Newton's second law are both dependent upon the mass of the object.

Electrical Example

Let’s go to electricity to see a case where it’s not always dependent upon mass.

Let’s have the same $5\,\mathrm{kg}$ object with $2\,\mathrm{C}$ of charge, only this time, it’s sitting in an electric field of $20\,\mathrm{N/C}$ *instead of* a gravitational field of $1.6\,\mathrm{N/kg}$. What is the acceleration in this case?

Well the force at a play is now the electrical force which is going to be the electric charge times the electric field $\vec{F} = q\vec{E}$.

We know the charge is $2\,\mathrm{C}$. We know the electric field is $20\,\mathrm{N/C}$, giving us an electric force of

$\vec{F} = q\vec{E} = (2\,\mathrm{C})(20\,\mathrm{N/C}) = 40\,\mathrm{N}$.

Now we move into calculating the acceleration, $F = ma$. Our force is $40\,\mathrm{N}$, the electric force we just calculated. Our mass is $5\,\mathrm{kg}$ and so we get an acceleration of

$\sum \vec{F} = m\vec{a}$[]
[latex] $(40 \quad \mathrm{N}) = (5 \quad \mathrm{kg})\ \mathrm{a}$
$a = 8\,\mathrm{m/s^2}$

You’ll notice that in this case q and m are different: we have a $5\,\mathrm{kg}$ object with $2\,\mathrm{C}$ of charge. The charge is what’s relevant for calculating the electric force, the mass is what’s relevant for calculating the acceleration, and since these are different the acceleration and the electric field are not the same number.

This idea where the acceleration and the gravitational field are the same number is unique to gravity because both the force and Newton’s second law depend upon m. This uniqueness, and thinking about it deeply, is what actually led Einstein to developing the general theory of relativity.

Instructor’s Note

I want you to work on visualizing these fields and thinking of interactions in terms of them.

The people who will be successful for the rest of this course will begin to really think about the fields as being real objects. These fields, gravitational fields, electric fields, magnetic fields, are truly present and they are everywhere around us.

This is how you “touch” things: the atoms in your finger never actually contact the atoms in the object you are touching. The charge in your finger generates an electric field, the charge in the wall interacts with that field and is repelled. That’s the origin of your sensation of pushing against the wall is through this intermediary of the electric field.

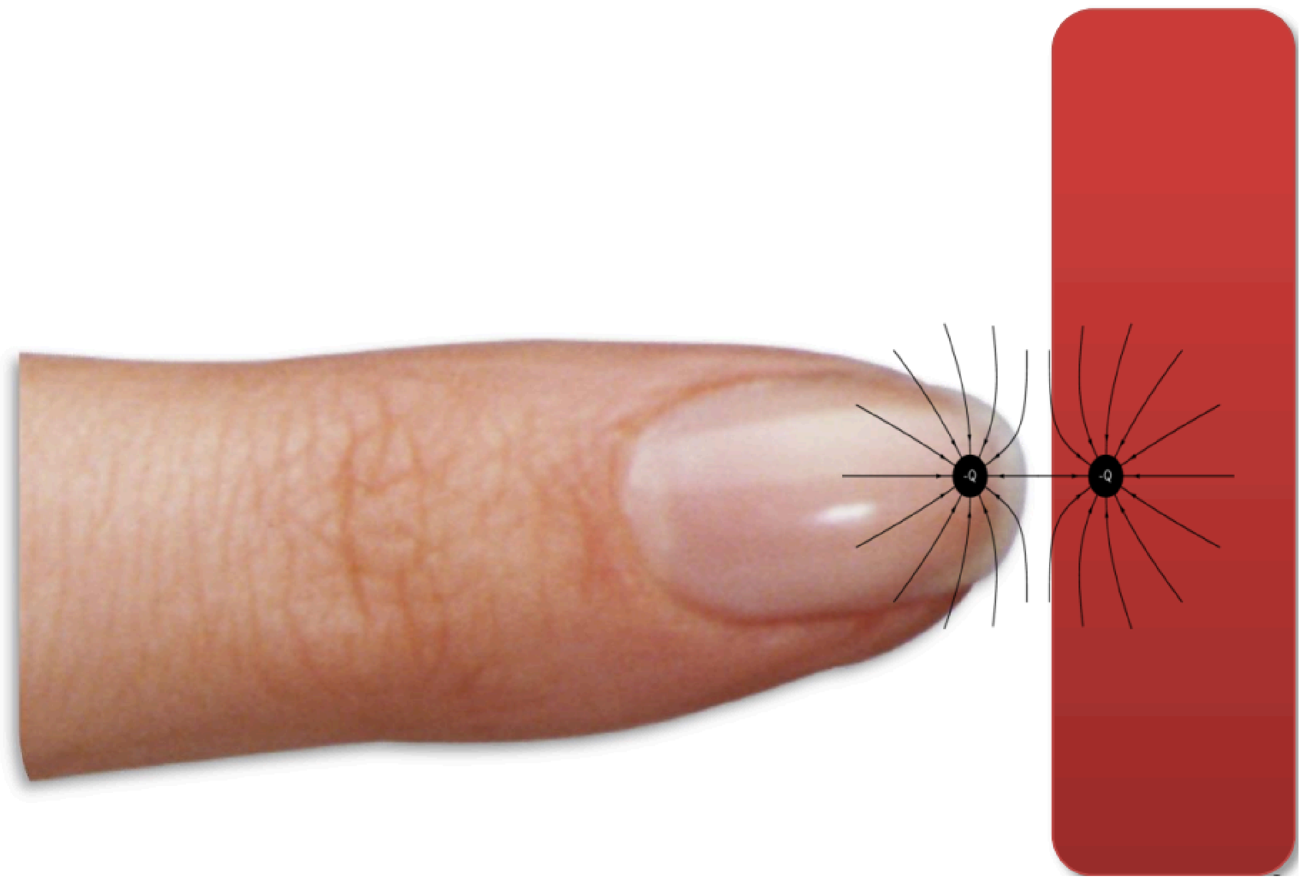


Figure 5. Charges interacting between your finger and the wall.

Instructor's Note

In summary, both gravity and electricity are fundamental forces that can act without contact between the objects. So how has this force actually transmitted if there's no contact? The answer is through fields, all massive objects generate gravitational fields, $\vec{g}$, and all electrically charged objects generate electric fields, $\vec{E}$.

Gravitational fields

- A distant object can interact with these fields and feel a force, an object with mass, m, in a gravitational field, $\vec{g}$, feels a force,$\vec{F}_g = m\vec{g}$: mass times the field.
- For gravity, the force $\vec{F}$ is always in the same direction as $\vec{g}$ because mass is always positive.
- The units of g: we're now going to think of them as Newton's per kilogram.

Electric fields

- An object with charge q in electric field $\vec{E}$ feels an electric force, charge times electric field, $\vec{F}_E = qE$.
- These parallels are what caused people to sometimes call mass gravitational charge, because mass is playing the same role for gravity as charge does for electricity.
- Unlike gravity, the force can be opposite the field because charge can be positive or negative, and just like the units of gravitational field are Newton's per kilogram, the units of electric field are Newton's per Coulomb.

Homework Problems

Problem 6: Electrically neutral objects can exert a gravitational force on each other, but they cannot exert an electrical force on each other.

Problem 7: What is the magnitude of the force exerted on a 4.49 C charge by a 310.24 N/C electric field that points due east?

Problem 8: Find the direction and magnitude of an electric field that exerts a 9×10^{-17} N westward force on an electron.

Problem 9: Calculate the initial (from rest) acceleration of a proton in a 3.09×10^{6} N/C electric field (such as created by a research Van de Graaff).

Problem 10: Suppose there is a single electron a set distance from a point charge Q, which quantities does the electric field experienced by the electron depend on?

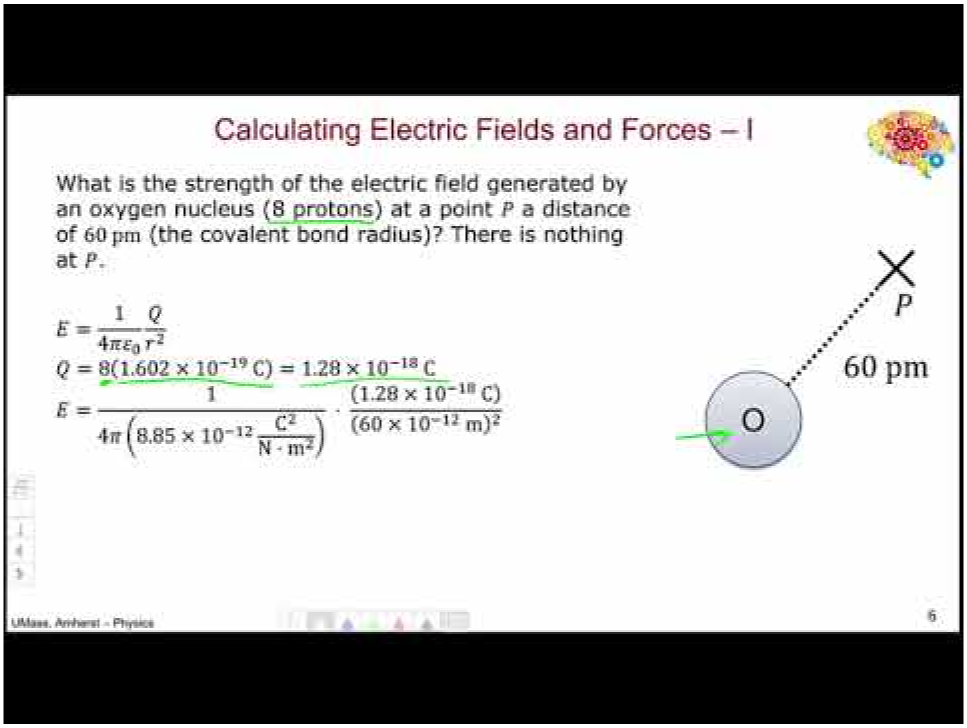
Calculating an Electric Field from a Point Charge

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

By the end of this section you should:

- Know that the quantity $g = 9.8\,\mathrm{N/kg}$ is a calculable quantity
- Be able to calculate the electric field from a given point charge.



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=243

In the last section, we discussed what electric fields are and how we visualize electrical forces in terms of electrical fields. In this section, we're going to actually calculate them. Just as with the last section, we will be making use of gravity as an analogy.

Let's begin by calculating fields and thinking in terms of gravity using our experience of Physics 131. Essentially the question we're looking to answer is where does this $g = 9.8\,\mathrm{N/kg}$ come from?

The value $g = 9.8\,\mathrm{N/kg}$ is the strength of the gravitational field from the Earth at the surface of the Earth. If you are at the space station, which is just outside of the atmosphere, the value of g is a little bit smaller, it's about $9.7\,\mathrm{N/kg}$. From this we can conclude that the strength of the gravitational field depends upon distance, the further you are from the planet, the weaker the gravitational field. On the moon on the other hand, the dominant gravitational field is not the gravitational field generated by the Earth, but the gravitational field generated by the moon and has a value of $1.6\,\mathrm{N/kg}$. In comparison, near the surface of the Sun, the dominant gravitational field is from the Sun and has a value of about $275\,\mathrm{N/kg}$. From this we can conclude that the strength of the gravitational field depends upon the mass of the object which is creating it. The moon is less massive than the Earth, so it generates a weaker gravitational field, the Sun is more massive than the Earth and generates a corresponding stronger gravitational field.

It turns out that the strength of the gravitational field is actually calculable using the equation below.

$$g = \frac{Gm}{r^2}$$

Where m is the mass of the object, which as we've already seen is one of the things that g depends on, a more massive object will result in a larger gravitational field. It also depends upon the distance from the center, r, because as we've seen with a Space Station the further away we get, the smaller the strength of the gravitational field. And then this G quantity is a constant of the universe, like the speed of light is a constant of the universe, and the value of this G constant is 6.67×10^{-11}.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

You don't actually need to know this expression for the gravitational field, I'm just using it to sort of build up an analogy because we're more familiar with gravity.

But now let's look at that analogy and think about calculating electric fields. Just as with gravitational fields, the strength of the electric field depends upon three things. The first being the amount of source, in this case the amount of charge that's generating the electric field, not the amount of mass, and we indicate the amount of charge with the letter Q. Just as with gravitational fields, electric fields also decrease with distance from the center of the charge, and finally there is also a constant, this ε_0 constant $\varepsilon_0 = 8.85 \times 10^{-12} \frac{\mathrm{C}^2}{\mathrm{N{\cdot}m}^2}$. When you go to write the formula for the electric field, it's very similar to the formula for the strength of the gravitational field. Gravitational field was some constant mass over distance squared, electrical field is some constant charge over distance squared.

$$\left|\vec{E}\right| = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}$$

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

I will ask you to use this particular equation.

You'll notice that we have the electric field vector inside of absolute value bars, and what that's indicating to you is that this formula only tells you the strength of the magnetic field, it doesn't tell you the direction. As for the direction, electric fields point away from positive charges and towards negative charges.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

While the formula for gravitational fields and electric fields appear very similar, again I want to stress that you should keep in mind that electric fields and gravitational fields are completely different things, electric fields transmit electrical forces which act on anything with charge, gravitational fields act on anything with mass. An object with a mass and a charge will generate both a gravitational field and an electrical field.

Let's take a moment to talk about the constant. The constant out front in the electric field expression is $\frac{1}{4\pi\varepsilon_0}$ and some people will write this constant as k and if you put in the value of $\varepsilon_0 = [latex]8.85 \times 10^{-12}\,\frac{\mathrm{C}^2}{\mathrm{N}\cdot\mathrm{m}^2}$ and calculate out this value, you will see that $\frac{1}{4\pi\varepsilon_0} = 9 \times 10^9\,\frac{\mathrm{N}\cdot\mathrm{m}^2}{\mathrm{C}^2}$

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

In class and in all of our examples I will use

$$\frac{1}{4\pi\varepsilon_0}$$

and the reason for this is it will be easier in the long run when we talk about materials and light waves thinking in terms of

$$\varepsilon_0$$

will be much more straightforward.

However, both ε_0 and k are on your equation sheet, in fact we have already seen the constant ε_0 in one of our equations, the equation relating the amplitude of light to the intensity of light $\mathcal{I} = \frac{1}{2}c\varepsilon_0 E^2$, that's the same ε_0. And in fact, looking ahead, this E in this equation stands for electric field. This is already starting to tell us something, that electric fields and light are going to be deeply connected in some interesting way that we'll talk about in our last unit.

Let's move on and actually try to use this expression for calculating the electric fields and forces.

What is the strength of an electric field generated by an oxygen nucleus, which has 8 protons in it, at a point P a distance of $60\ \mathrm{pm}$?

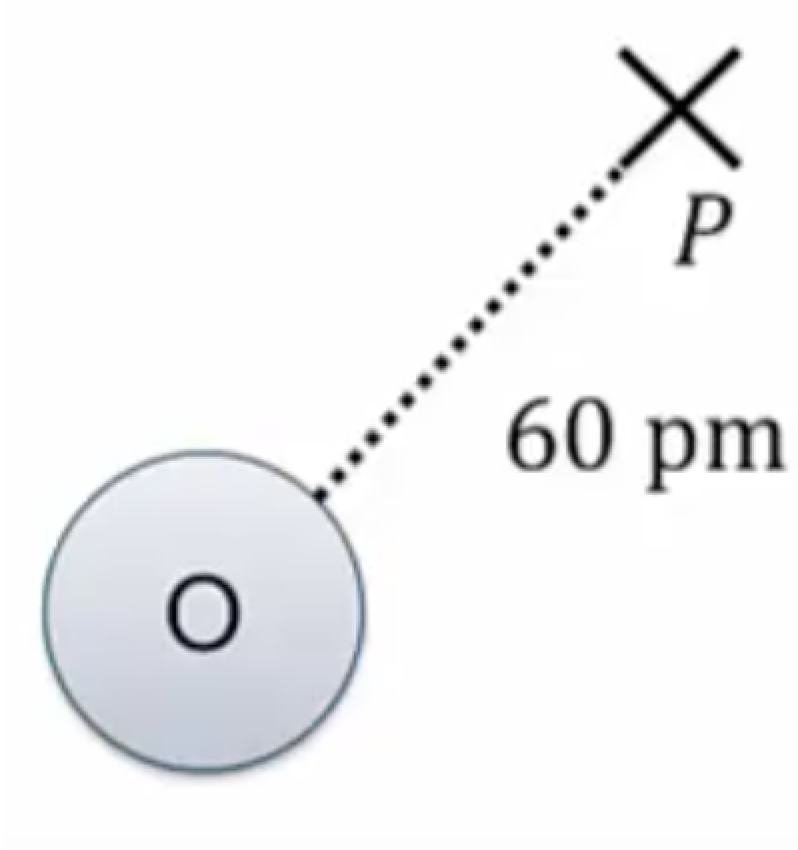


Figure 6: A random point P 60pm from an oxygen nucleus.

There is absolutely nothing currently at the point P. We're going to ignore the effects of all the surrounding electrons and just think about the nucleus for the moment. We begin with our definition for the strength of the electric field. It is the charges within the nucleus that are generating this field, and inside the nucleus we have 8 protons, so we have 8 times the charge on each proton, giving us a total charge of $Q = 8q_e = 1.28 \times 10^{-18}\,\mathrm{C}$. We also know that $r = 60\,\mathrm{pm} = 60 \times 10^{-12}\,\mathrm{m}$ from the problem. When we plug in all of our numbers into the equation above we get an electric field of $3.2 \times 10^{12}\,\mathrm{N/C}$.

Think about this for a second, the strength of the gravitational field generated by the entire Earth is essentially $10\,\mathrm{N/kg}$! The electric fields inside of an atom are much larger. Since we know that the nucleus is positively charged we know that the electric field will point away from the source.

Instructor's Note

It's worth pointing out that this field is a real thing and exists regardless if there is something to experience it at P or not. The oxygen nucleus will generate a field that surrounds it which at point P has this value, respective of if there is actually an object at point P.

Now let's see what happens when we put an object at point P

We're going to go and add an electron at the point we've been talking about.

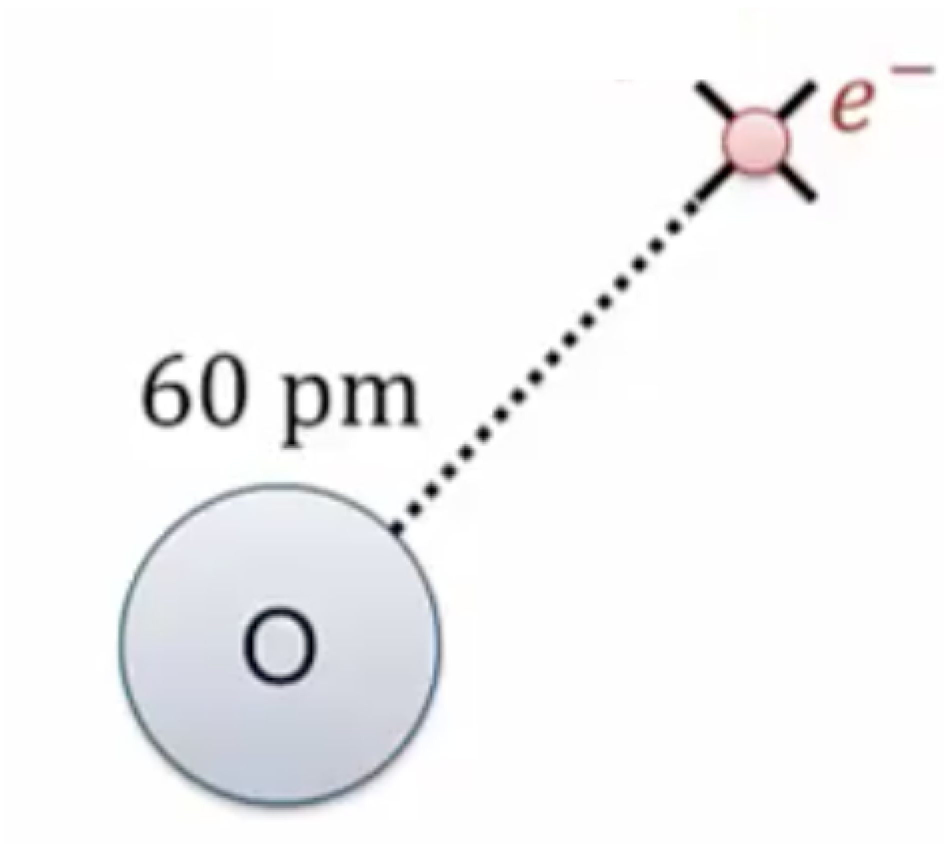


Figure 7: Add an electron to the point P

We know the strength of the electric field $3.2 \times 10^{12}\,\mathrm{N/C}$, calculated in the last part, we know that the electron will feel a force

$$\vec{F} = q\vec{E}$$

as discussed in a previous section, and we know the charge of the electron is $q = 1.602 \times 10^{-19}\,\mathrm{C}$. Substituting that charge and our value for electric field we get a force of $F = -5.13 \times 10^{-7}\,\mathrm{N}$ where the negative sign means that the force is opposite the electric field.

Then we can finally move on to calculating acceleration. We begin with Newton's second law

$$\sum \vec{F} = m\vec{a}$$

The only force present on this particular electron is the electrical force, $5.13 \times 10^{-7}\,\mathrm{N}$. We know the mass of the electron from our equation sheet 9.11×10^{-31} kilograms, and then we solve for the acceleration and we get an enormous value, $a = 5.63 \times 10^{23}\,\mathrm{m/s^2}$. Remember for comparison, accelerations in your everyday life are 10 meters per second squared, and at an acceleration of about 80 meters per second squared you're blacking out. This electron is accelerating with an astronomically large acceleration.

Instructor's Note

In Summary:

- The electric field from a point charge can be calculated by this $\left|\vec{E}\right| = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}$, which depends upon the charge, making the field make sense and the distance from the charge to the point of interest, r.
- It also depends upon this constant $\varepsilon_0 = 8.85 \times 10^{-12}\frac{\mathrm{C}^2}{\mathrm{N{\cdot}m}^2}$.
- This field exists regardless if there is something there to feel it or not and the forces and fields in electricity we've already seen in the few examples are much bigger than gravitational forces.
- As sort of a reference point say you have two electrons, the electrical repulsion between the two electrons is 10^{40} times larger than their gravitational attraction

Homework Problems

Problem 11: Calculate the magnitude of the electric field 8.41">8.41 m from a point charge of 8.92">8.92 mC (such as found on the terminal of a Van de Graaff).

Problem 12: What magnitude point charge creates a 15947.46">15947.46 N/C electric field at a distance of 0.513">0.513 m? How large is the field at 18.85">18.85 m?

Visualizing Electric Fields

Drawings using lines to represent electric fields around charged objects are very useful in visualizing field strength and direction. Since the electric field has both magnitude and direction, it is a vector. Like all vectors, the electric field can be represented by an arrow that has length proportional to its magnitude and that points in the correct direction. (We have used arrows extensively to represent force vectors, for example.)

Figure 6 shows two pictorial representations of the same electric field created by a positive point charge QQ size 12{Q} {}">Q. Figure 6b shows the standard representation using continuous lines. Figure 6a shows numerous individual arrows with each arrow representing the force on a test charge qq size 12{q} {}">q. Field lines are essentially a map of infinitesimal force vectors.

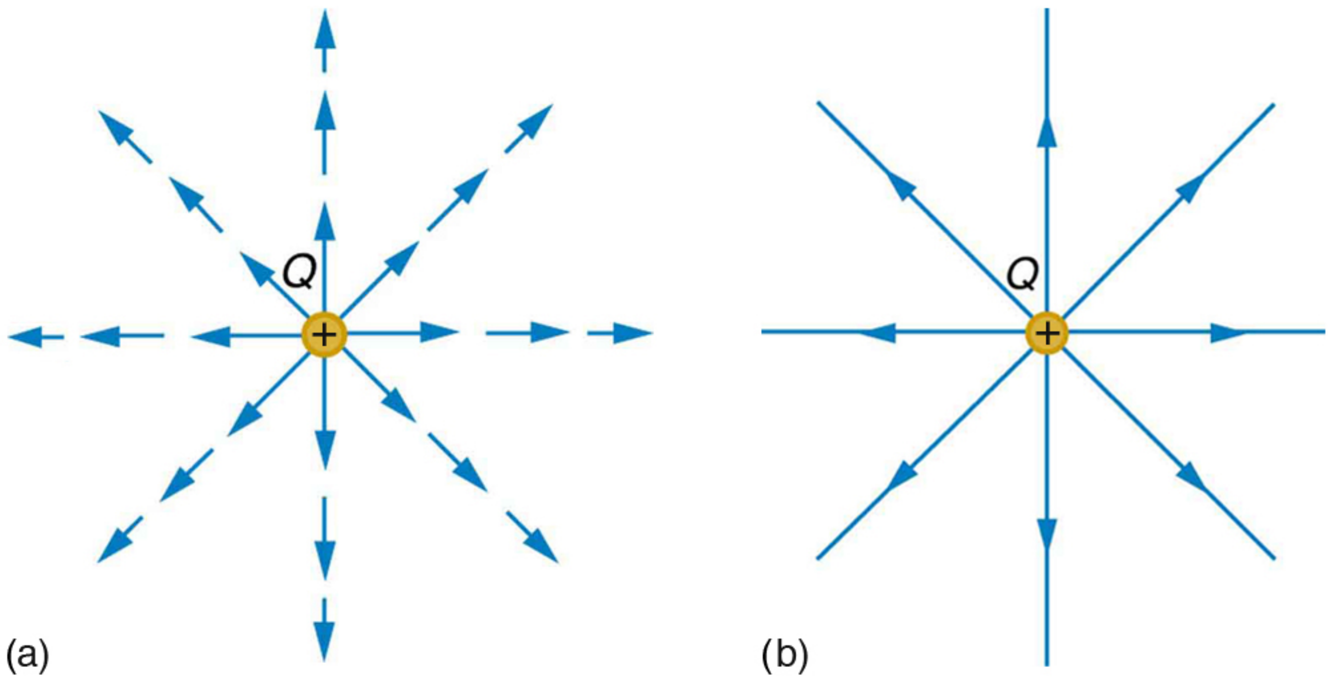


Figure 8 Two equivalent representations of the electric field due to a positive charge Q. (a) Arrows representing the electric field's magnitude and direction. (b) In the standard representation, the arrows are replaced by continuous field lines having the same direction at any point as the electric field. The closeness of the lines is directly related to the strength of the electric field. A test charge placed anywhere will feel a force in the direction of the field line; this force will have a strength proportional to the density of the lines (being greater near the charge, for example).

Note that the electric field is defined for a positive test charge qq size 12{q} {}">q, so that the field lines point away from a positive charge and toward a negative charge. (See Figure 7.) The electric field strength is exactly proportional to the number of field lines per unit area, since the magnitude of the electric field for a point charge is $E\ frac14\pi\epsilon\frac{q}{r^2}$E=k|Q|/r2E=k|Q|/r2 size 12{E= { ital "kQ"} slash {r rSup { size 8{2} } } } {}"> and area is proportional to r2r2 size 12{r rSup { size 8{2} } } {}">r^2. This pictorial representation, in which field lines represent the direction and their closeness (that is, their areal density or the number of lines crossing a unit area) represents strength, is used for all fields: electrostatic, gravitational, magnetic, and others.

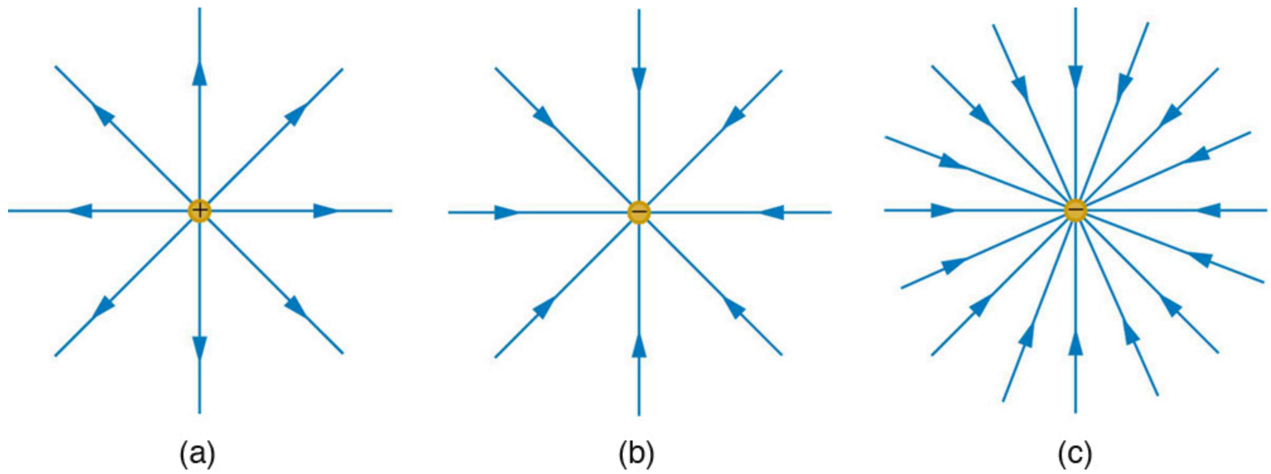


Figure 9 The electric field surrounding three different point charges. (a) A positive charge. (b) A negative charge of equal magnitude. (c) A larger negative charge.

In many situations, there are multiple charges. The total electric field created by multiple charges is the vector sum of the individual fields created by each charge. Figure 10 shows how the electric field from two point charges can be drawn by finding the total field at representative points and drawing electric field lines consistent with those points. While the electric fields from multiple charges are more complex than those of single charges, some simple features are easily noticed. For example, the field is weaker between like charges, as shown by the lines being farther apart in that region. (This is because the fields from each charge point in opposite directions.) (See Figure 10 and Figure 11 (a).) Furthermore, at a great distance from two like charges, the field becomes identical to the field from a single, larger charge.

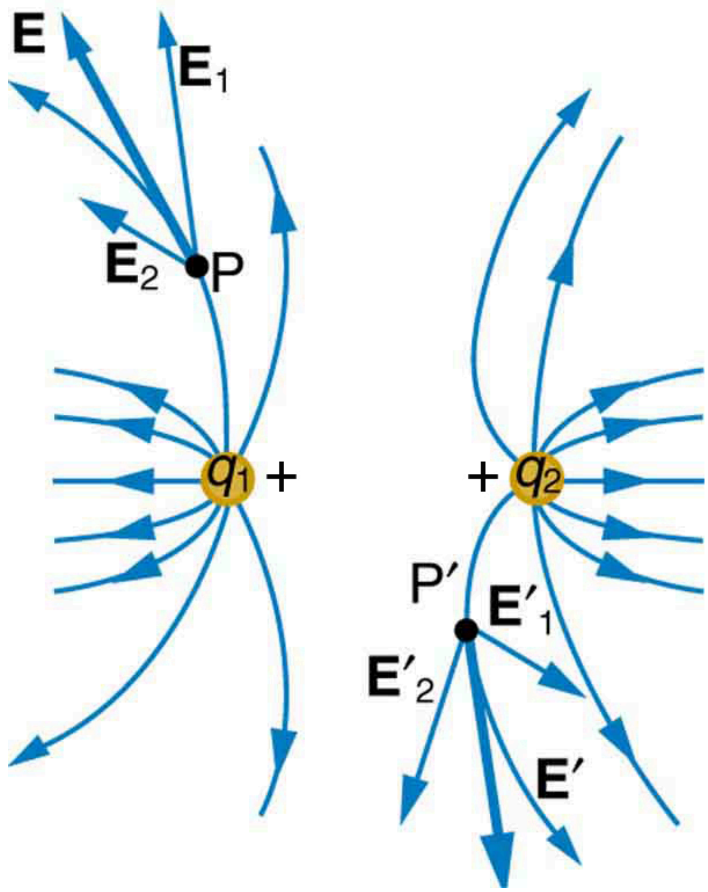


Figure 10: Two positive point charges q1 and q2 produce the resultant electric field shown. The field is calculated at representative points and then smooth field lines drawn following the rules outlined in the text.

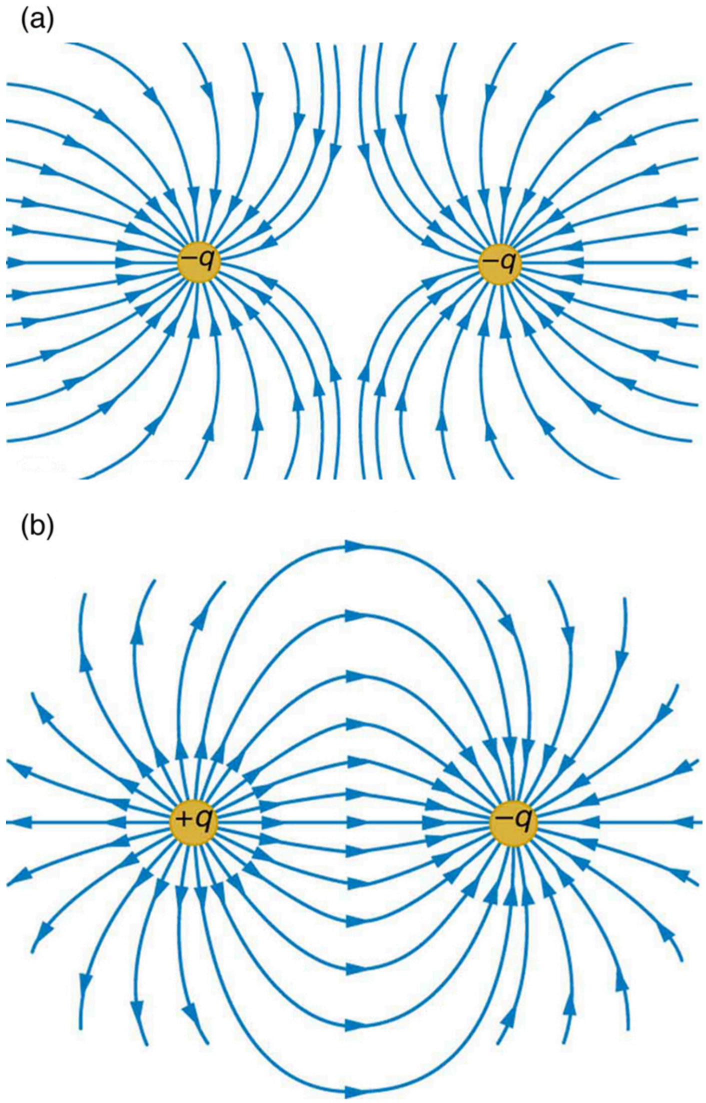


Figure 11 (a) Two negative charges produce the fields shown. It is very similar to the field produced by two positive charges, except that the directions are reversed. The field is clearly weaker between the charges. The individual forces on a test charge in that region are in opposite directions. (b) Two opposite charges produce the field shown, which is stronger in the region between the charges.

Figure 11(b) shows the electric field of two unlike charges. The field is stronger between the charges. In that region, the fields from each charge are in the same direction, and so their strengths add. The field of two unlike charges is weak at large distances, because the fields of the individual charges are in opposite directions and so their strengths subtract. At very large distances, the field of two unlike charges looks like that of a smaller single charge.

We use electric field lines to visualize and analyze electric fields (the lines are a pictorial tool, not a physical entity in themselves). The properties of electric field lines for any charge distribution can be summarized as follows:

1. Field lines must begin on positive charges and terminate on negative charges, or at infinity in the

hypothetical case of isolated charges.

2. The number of field lines leaving a positive charge or entering a negative charge is proportional to the magnitude of the charge.
3. The strength of the field is proportional to the closeness of the field lines—more precisely, it is proportional to the number of lines per unit area perpendicular to the lines.
4. The direction of the electric field is tangent to the field line at any point in space.
5. Field lines can never cross.

The last property means that the field is unique at any point. The field line represents the direction of the field; so if they crossed, the field would have two directions at that location (an impossibility if the field is unique).

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Notes

I will expect you to know these five rules.

Homework Problems

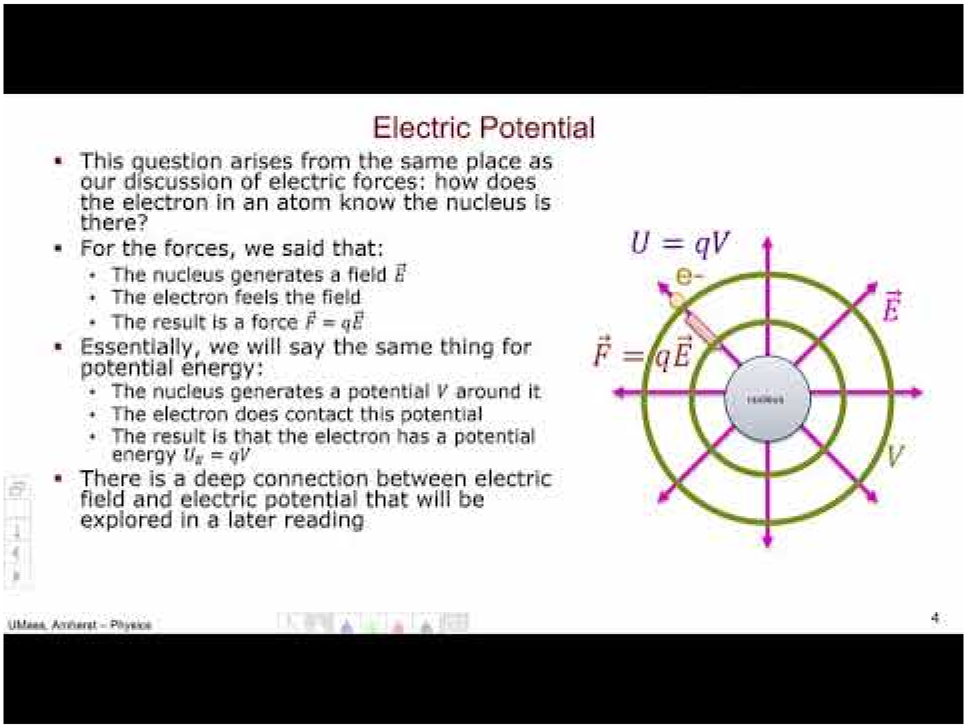
Problem 13: Below you see an unknown charge generating an electric field. Also indicated are two empty regions of space "A" and "B." Which of the following statements are true from this picture of field lines?

23. Electric Potential

Introduction to Potential

Instructor's Note

By the end of this section you should know that the electric force is conservative (i.e. there is a potential energy associated with the electric force), you should be able to define what a potential is and be able to calculate potential energy from a potential.



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=245*](http://openbooks.library.umass.edu/toggerson-132/?p=245)

We will begin by stating that the electric force is a conservative force, which means that an electric potential energy must exist. In one of your problems, you will explore the idea of work done by a charge moving in a uniform electric field. You will see that as the charge moves around the work done is independent of the path of the charged takes, and that the work done around a closed loop path is in fact zero. This fact that the work done around a closed loop path is indicative of the fact that the electric field must be a conservative force, this should be familiar to you from previous sections.

Since the electric force is conservative, we know that we can write down a potential energy, U_E for the electrical force.

Instructor's Note

As usual, some places will use PE for potential energy but we in class will use the letter capital U

We've actually been using the idea of electric potential energy already, throughout both this class and Physics 131. The chemical energies discussed in Physics 131 are actually electric potential energies. Similarly, the potential energies of the electrons that we discussed in Units 1 and 2, unless we stated explicitly that they were gravitational potential energies, were electric potential energies.

What is the electric potential? In the figure below, we have an electron surrounding a nucleus.

Figure 1: An electron around a nucleus.

The question arises from the same place as our discussion of electrical forces, how does the electron know that the nucleus is there?

In the case of forces, we said that the nucleus generates an electric field, $\vec{E}$. The electron is in contact with this field, and as a result feels a force, $\vec{F} = q\vec{E}$.

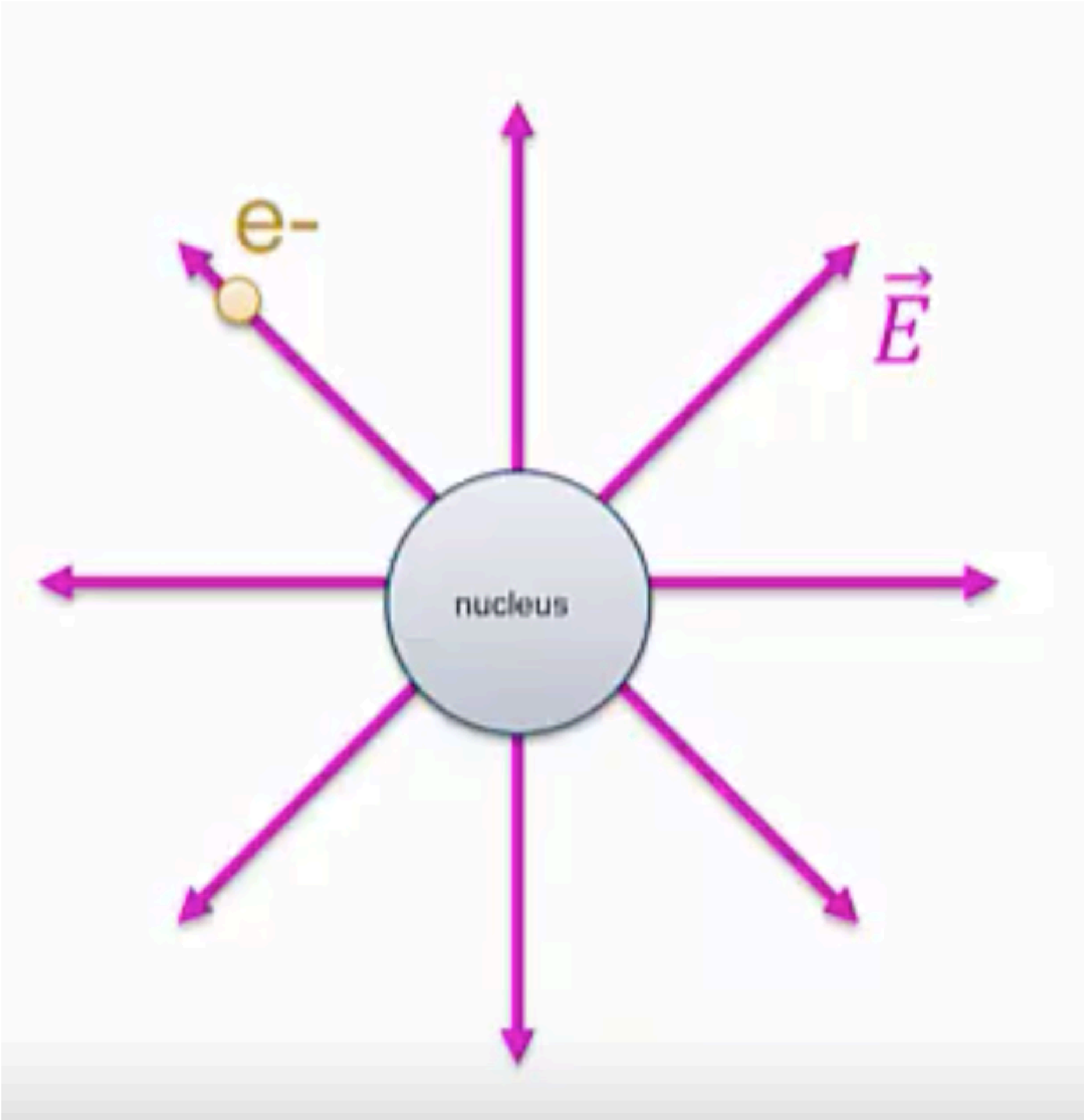


Figure 2: The electron knows of the nucleus's existence because the nucleus generates an electric field and the electron, which is in contact with that field, experiences a force in response.

Essentially, we're going to say the same thing for potential energy, the nucleus is going to generate an electric potential, V, around it, you'll learn how to calculate these potentials from point charges in the next section. The electron does contact the potential and as a result feels a potential energy, $U = qV$

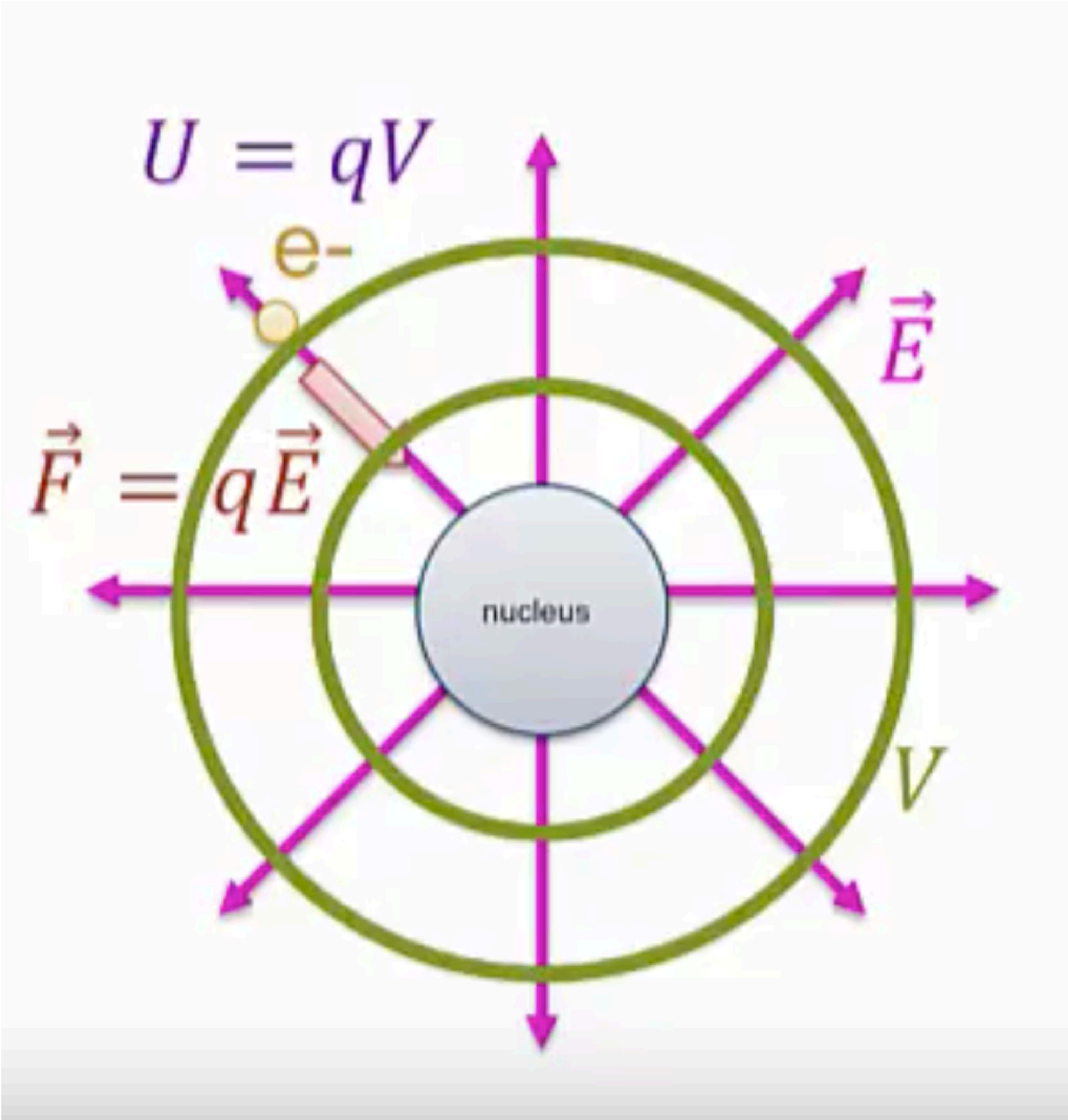


Figure 3: From an energy perspective, the nucleus generates a potential V around it and the electron responds to that potential be feeling a potential energy U.

There's a deep connection between electric field and electric potential that will be explored in a later section. Just as with electric field the potential exists even if there is something to feel there or not, so even if we were to remove the electron the potential would still be present.

Now let's talk about the units of potential. The unit of potential is the Volt, V. Yes, it has the same symbol as the quantity that we're using for potential, but we have to deal with it. One Volt is one Joule per Coulomb, $1\,\mathrm{V} = \frac{1\,\mathrm{J}}{1\,\mathrm{C}}$. This definition is visible from the equation connecting potential and potential energy. If we rewrite

$$U = qV$$

into

$$V = \frac{U}{q},$$

we know that U has units of Joules, q has units of Coulombs, so$V$, potential, is going to have units of Joules per Coulomb which we call Volts. This second way of writing potential as

$$\frac{U}{q},$$

is why some people call potential, potential energy per unit charge. On the other hand, I want you to think of it as an invisible field around charges, that gives rise to potential energy when other charged particles interact with it.

Now, let's do an example

An electron has 20eV of kinetic energy in a region where the potential is 10V. The electron then travels to a region with a lower potential of $5\,\text{V}$.

What are the initial and final potential energies? What is the change in potential energy?

Solution:

Let's begin by looking at the initial potential energy

$$U = qV.$$

We know the charge of the electron $q = 1.60210^{-19}\,\text{C}$ and our initial potential is $10\,\text{V}$. So, the initial potential energy will be

$$U = qV \rightarrow U = (1.602 \times 10^{-19}\,\text{C})(10\,\text{V}) = 16.0210^{-19}\,\text{J}$$

Joules or converting that to electron volts, we get $10\,\text{eV}$.

Now let's do the final potential energy, we know the charge of the electron. Again, our final potential is $5\,\text{V}$. Multiplying this together we get a final potential of

$$U = qV \rightarrow U = (1.602 \times 10^{-19}\,\text{C})(5\,\text{V}) = 8.0110^{-19}\,\text{J}$$

or

$$U = 5\,\text{eV}.$$

Now let's think about the change in potential energy, $\Delta U = U_f U_i$. We solved for $U_f = 5\,\text{eV}$. Moreover, we saw our initial potential energy was $U_i 10\,\text{eV}$. So the result is a change of

$$\Delta U = U_f - U_i = (-5\,\text{eV}) - (-10\,\text{eV}) = +5\,\text{eV}.$$

Even though the potential *dropped* from 10V to 5V, the potential energy actually *increased*. This is due to the fact that the electron has a negative charge.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Throughout our calculations, we've been multiplying the potential V by a negative charge q.

If we had instead considered a proton, then q would be positive, and a positive drop in potential would result in a drop in potential energy.

Once we have changes in potential energy, we can then move on to solve problems using conservation of energy as we've been doing throughout this course.

One last point to discuss is the connection between the volt and the electron volt. You may have already started to see this connection in the last problem. Throughout this course and in Physics 131 we've been using the electron volt as a unit of energy, and we've just been using it as a straight conversion factor,

$$1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}$$

Now however, you have enough information to understand where this unit of energy comes from: 1eV is the increase in energy of an electron as it goes across a 1-volt potential drop. To solve it out, we know

$$U = qV$$

so the change in U is the charge times the change in potential. The charge of the electron is $-1.602 \times 10^{-19}\,\mathrm{C}$. A unit potential drop would be a change in potential of $1\,\mathrm{V}$ and so multiplying it all out we see that an electron going across a 1-volt potential drop has an increase in potential energy of $1.602 \times 10^{-19}\,\mathrm{J}$ which we recognize as 1eV.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

In Summary:

- Potential is to potential energy as electric field is to electric force
- Forces result in charged objects interacting, forces result from charged particles interacting with the fields generated by other charged objects through

$$F = qE$$

- Potential energies result from charged objects interacting with the potentials generated by other charged objects, mathematically written as

$$U = qV$$

- Fields and potentials have the same sort of relationship as forces and potential energies
- We can solve many problems by looking at it either in terms of fields and potentials, just like we can solve many problems by looking at it in terms of forces or potential energies
- The unit of the potential is the volt, where one volt is equal to one Joule per Coulomb and the electron volt as a unit of energy arises from the amount of energy gained by an electron going across a one-volt potential difference.

Some Common Misconceptions About Potential

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

$$\Delta U = q\Delta V.$$

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 $\Delta U = q\Delta V$. The car battery can move more charge than the motorcycle battery, although both are 12 V batteries.

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 $\Delta U = q\Delta V$.

So to find the energy output, we multiply the charge moved by the potential difference.

Solution

For the motorcycle battery, $q = 5000\,\mathrm{C}$ and $\Delta V = 12.0\,\mathrm{V}$. The total energy delivered by the motorcycle battery is

$$\begin{aligned}\Delta U_{\text{cycle}} &= (5000\,\mathrm{C})(12.0\,\mathrm{V})\\ &= (5000\,\mathrm{C})(12.0\,\mathrm{J/C})\\ &= 6.00 \times 10^4\,\mathrm{J}.\end{aligned}$$

Similarly, for the car battery, $q = 60,000\,\mathrm{C}$ and

$$\begin{aligned}\Delta U_{\text{car}} &= (60,000\,\mathrm{C})(12.0\,\mathrm{V})\\ &= 7.20 \times 10^5\,\mathrm{J}.\end{aligned}$$

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. The change in potential is $\Delta V = V_B - V_A = +12\ \mathrm{V}$ and the charge q is negative, so that $\Delta U = q\Delta V$ is negative, meaning the potential energy of the battery has decreased when q has moved from A to B.

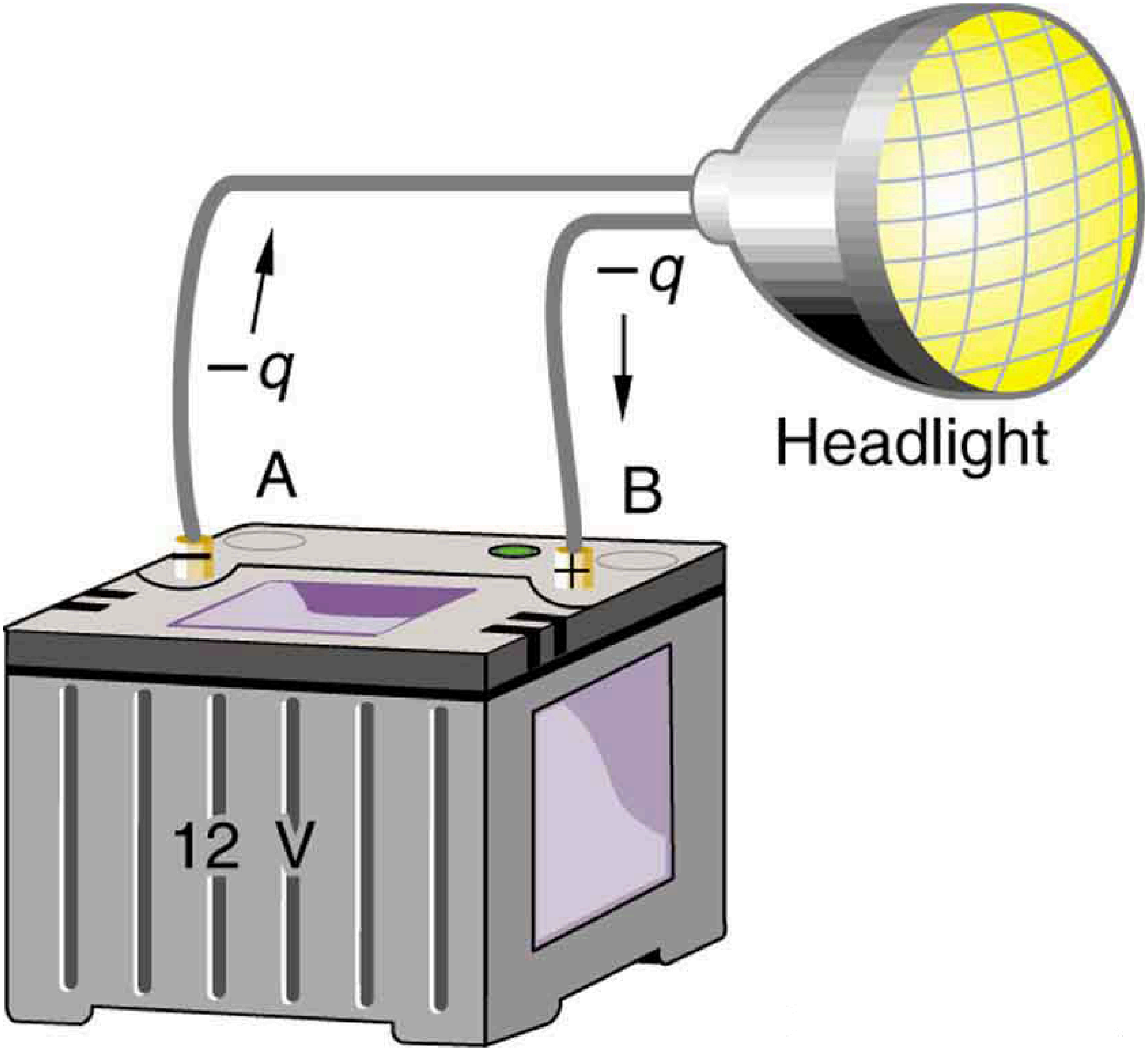


Figure 4: A battery moves negative charge from its negative terminal through a headlight to its positive terminal. Appropriate combinations of chemicals in the battery separate charges so that the negative terminal has an excess of negative charge, which is repelled by it and attracted to the excess 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. We will discuss batteries in more detail in our next unit.

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 $\Delta U = q\Delta V$. A 30.0 W lamp uses 30.0 joules per second. Since the battery loses energy, we have $\Delta U = -30.0 \text{ J}$ and, since the electrons are going from the negative terminal to the positive, we see that $\Delta V = +12.0 \text{ V}$.

Solution

To find the charge $q$moved, we solve the equation $\Delta U = q\Delta V$:

$$q = \frac{\Delta U}{\Delta V}.$$

Entering the values for ΔU and ΔV, we get

$$q = \frac{-30.0 \text{ J}}{+12.0 \text{ V}} = \frac{-30.0 \text{ J}}{+12.0 \text{ J/C}} = -2.50 \text{ C}.$$

The number of electrons n_e is the total charge divided by the charge per electron. That is,

$$\mathrm{n}_e = \frac{-2.50 \text{ C}}{-1.60\times10^{-19} \text{ C/e}^-} = 1.56 \times 10^{19} \text{ electrons}.$$

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.

Homework Problems

Problem 14: A lightning bolt strikes a tree, moving charge through a potential difference. What energy was dissipated?

Problem 15: An evacuated tube uses an accelerating voltage to accelerate electrons to hit a copper plate and produce x rays. What would be the final speed of such an electron?

Electrical Potential Due to a Point Charge

Point charges, such as electrons, are among the fundamental building blocks of matter. Furthermore, spherical charge distributions (like on a metal sphere) create external electric fields exactly like a point charge. The electric potential due to a point charge is, thus, a case we need to consider. Using calculus to find the work done by a non-conservative force to move a small charge q from a large distance away, against the electric field, to a distance of r from a point charge Q, it can be shown that the *electric potential* V *of a point charge* is

$$V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}\text{(Point Charge)},$$

where $\frac{1}{4\pi\epsilon_0} = 9.0 \times 10^9 \text{ N } \cdot \text{m}^2/\text{C}^2$ as usual.

As with potential energy, the potential at infinity is chosen to be zero. Thus V for a positive point charge decreases with distance.

Recall that the electric potential V is a scalar and has no direction, whereas the electric field $\mathbf{E}$ is a vector. To find the voltage due to a combination of point charges, you add the individual voltages as numbers. To find the total electric field, you must add the individual fields as *vectors*, taking magnitude and direction into account. This is consistent with the fact that V is closely associated with energy, a scalar, whereas $\mathbf{E}$ is closely associated with force, a vector.

What Voltage Is Produced by a Small Charge on a Metal Sphere?

Charges in static electricity are typically in the nanocoulomb (nC) to microcoulomb (μC) range. What is the voltage 5.00 cm away from the center of a 1-cm diameter metal sphere that has a -3.00 nC static charge?

Strategy

As we have discussed in Electric Charge and Electric Field, charge on a metal sphere spreads out uniformly and produces a field like that of a point charge located at its center. Thus we can find the voltage using the equation $V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}$.

Solution

Entering known values into the expression for the potential of a point charge, we obtain

$$V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}$$
$$= (8.99 \times 10^9 \text{ N } \cdot \text{ m}^2/\text{ C}^2)(\frac{-3.00{\times}10^9 \text{ C}}{5.0010^2 \text{ m}})$$
$$= -539\text{V}.$$

Discussion

The negative value for voltage means a positive charge would be attracted from a larger distance, since the potential is lower (more negative) than at larger distances. Conversely, a negative charge would be repelled, as expected.

What Is the Excess Charge on a Van de Graaff Generator

A demonstration Van de Graaff generator has a 25.0 cm diameter metal sphere that produces a voltage of 100 kV near its surface. (See Figure 1.) What excess charge resides on the sphere? (Assume that each numerical value here is shown with three significant figures.)

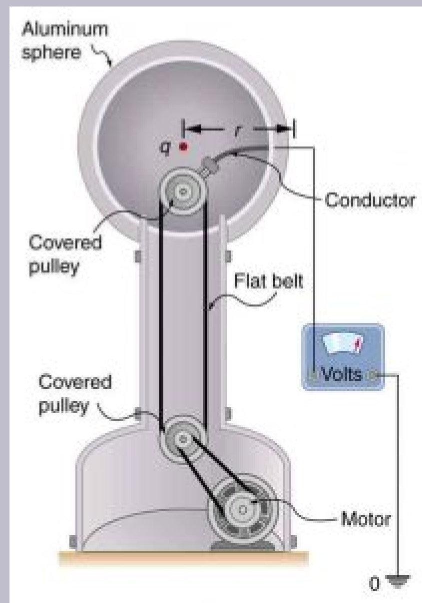


Figure 5: The voltage of this demonstration Van de Graaff generator is measured between the charged sphere and ground. Earth's potential is taken to be zero as a reference. The potential of the charged conducting sphere is the same as that of an equal point charge at its center.

Strategy

The potential on the surface will be the same as that of a point charge at the center of the sphere, 12.5 cm away. (The radius of the sphere is 12.5 cm.) We can thus determine the excess charge using the equation

$$V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}.$$

Solution

Solving for Q and entering known values gives

$$Q = 4\pi\varepsilon_0 rV$$

$$= 4\pi(8.85\times10^{-12}\ \frac{\mathrm{C}^2}{\mathrm{N{\cdot}m}})(0.125\ \mathrm{m})(100\times10^3\ \mathrm{V})$$

$$= 1.39\times10^{-6}\ \mathrm{C} = 1.39\ \mu C.$$

Discussion

This is a relatively small charge, but it produces a rather large voltage. We have another indication here that it is difficult to store isolated charges.

The voltages in both of these examples could be measured with a meter that compares the measured potential with ground potential. Ground potential is often taken to be zero (instead of taking the potential at infinity to be zero). It is the potential difference between two points that is of importance, and very often there is a tacit assumption that some reference point, such as Earth or a very distant point, is at zero potential similar to the process described in Unit I – Chapter 5 Some Energy Ideas that Might Be New or Are Particularly Important. Choosing the Earth or some other reference point to be $V=0$ is analogous to taking sea level as $h=0$when considering gravitational potential energy, $U_g=mgh$.

Homework Problems

Problem 16: What is the potential 52.92 pm from a proton (the average distance between the proton and electron in a hydrogen atom)?

Problem 17: A research Van de Graaff generator has a metal sphere with a charge on it. What is the potential near its surface?

Equipotential Lines

We can represent electric potentials (voltages) pictorially, just as we drew pictures to illustrate electric fields. Of course, the two are related. Consider Figure 1, which shows an isolated positive point charge and its electric field lines. Electric field lines radiate out from a positive charge and terminate on negative charges. While we use blue arrows to represent the magnitude and direction of the electric field, we use green lines to represent places where the electric potential is constant. These are called **equipotential lines** in two dimensions, or *equipotential surfaces* in three dimensions. The term *equipotential* is also used as a noun, referring to an equipotential line or surface. The potential for a point charge is the same anywhere on an imaginary sphere of radius r surrounding the charge. This is true since the potential for a point charge is given by $V=\frac{1}{4\pi\varepsilon_0}\frac{q}{r}$ and, thus, has the same value at any point that is a given distance r from the charge. An equipotential sphere is a circle in the two-dimensional view of Figure 1. Since the electric field lines point radially away from the charge, they are perpendicular to the equipotential lines.

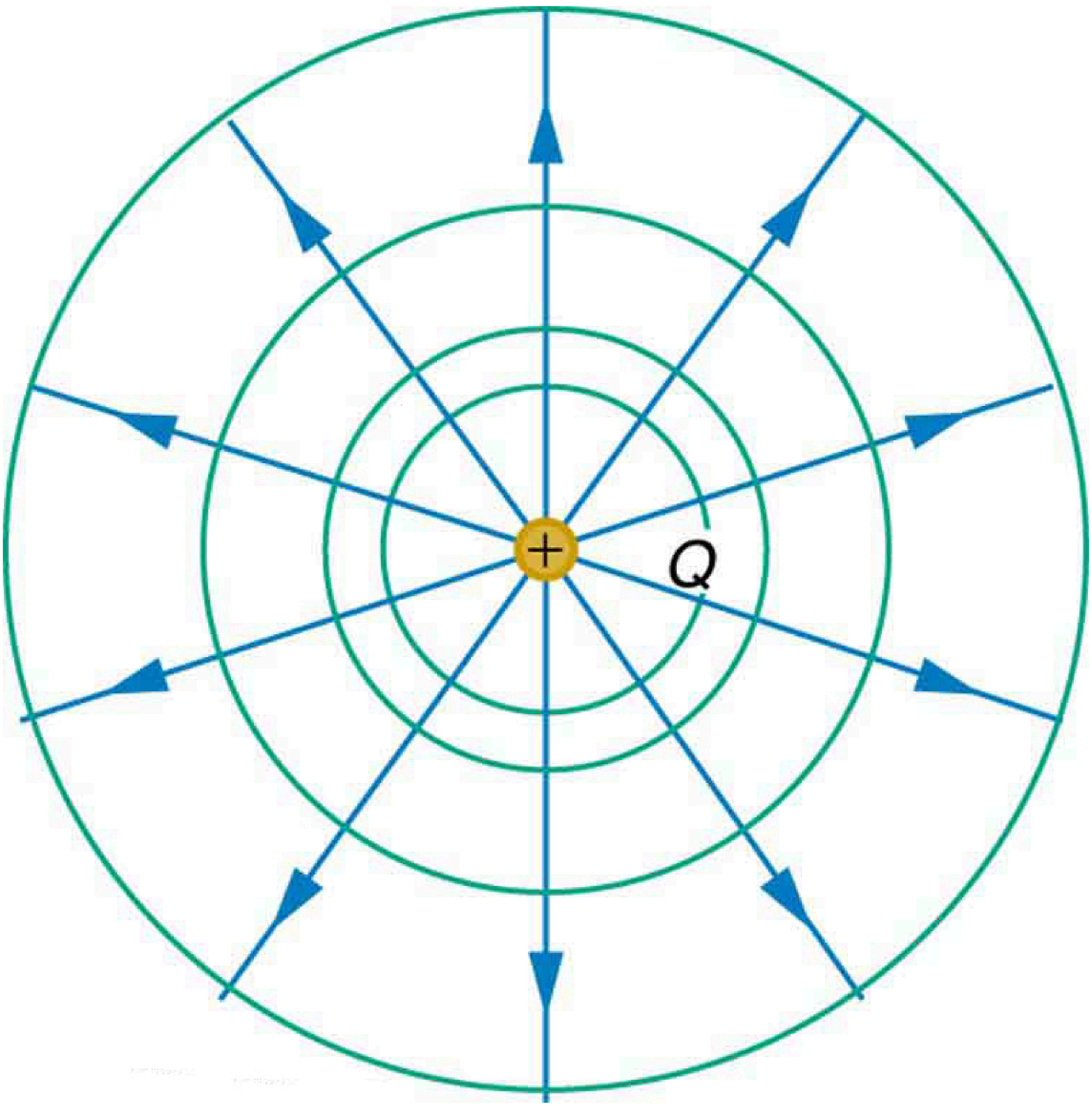


Figure 6: An isolated point charge Q with its electric field lines in blue and equipotential lines in green. The potential is the same along each equipotential line, meaning that no work is required to move a charge anywhere along one of those lines. Work is needed to move a charge from one equipotential line to another. Equipotential lines are perpendicular to electric field lines in every case.

It is important to note that *equipotential lines are always perpendicular to electric field lines*. No work is required to move a charge along an equipotential, since $\Delta V = 0$. Work is zero if force is perpendicular to motion. Force is in the same direction as E, so that motion along an equipotential must be perpendicular to E. More precisely, work is related to the electric field by

$$W = Fd\cos\theta = qEd\cos\theta = 0.$$

Note that in the above equation, E and F symbolize the magnitudes of the electric field strength and force, respectively. Neither q nor E nor d is zero, and so $\cos\theta$ must be 0, meaning θ must be 90°90°">90°. In other words, motion along an equipotential is perpendicular to E.

One of the rules for static electric fields and conductors is that the electric field must be perpendicular to the surface of any conductor. This implies that a *conductor is an equipotential surface in static situations*. There can be no voltage difference across the surface of a conductor, or charges will flow.

Because a conductor is an equipotential, it can replace any equipotential surface. For example, in Figure 1 a charged spherical conductor can replace the point charge, and the electric field and potential surfaces outside of it will be unchanged, confirming the contention that a spherical charge distribution is equivalent to a point charge at its center.

Figure 2 shows the electric field and equipotential lines for two equal and opposite charges. Given the electric field lines, the equipotential lines can be drawn simply by making them perpendicular to the electric field lines. Conversely, given the equipotential lines, as in Figure 3(a), the electric field lines can be drawn by making them perpendicular to the equipotentials, as in Figure 3(b).

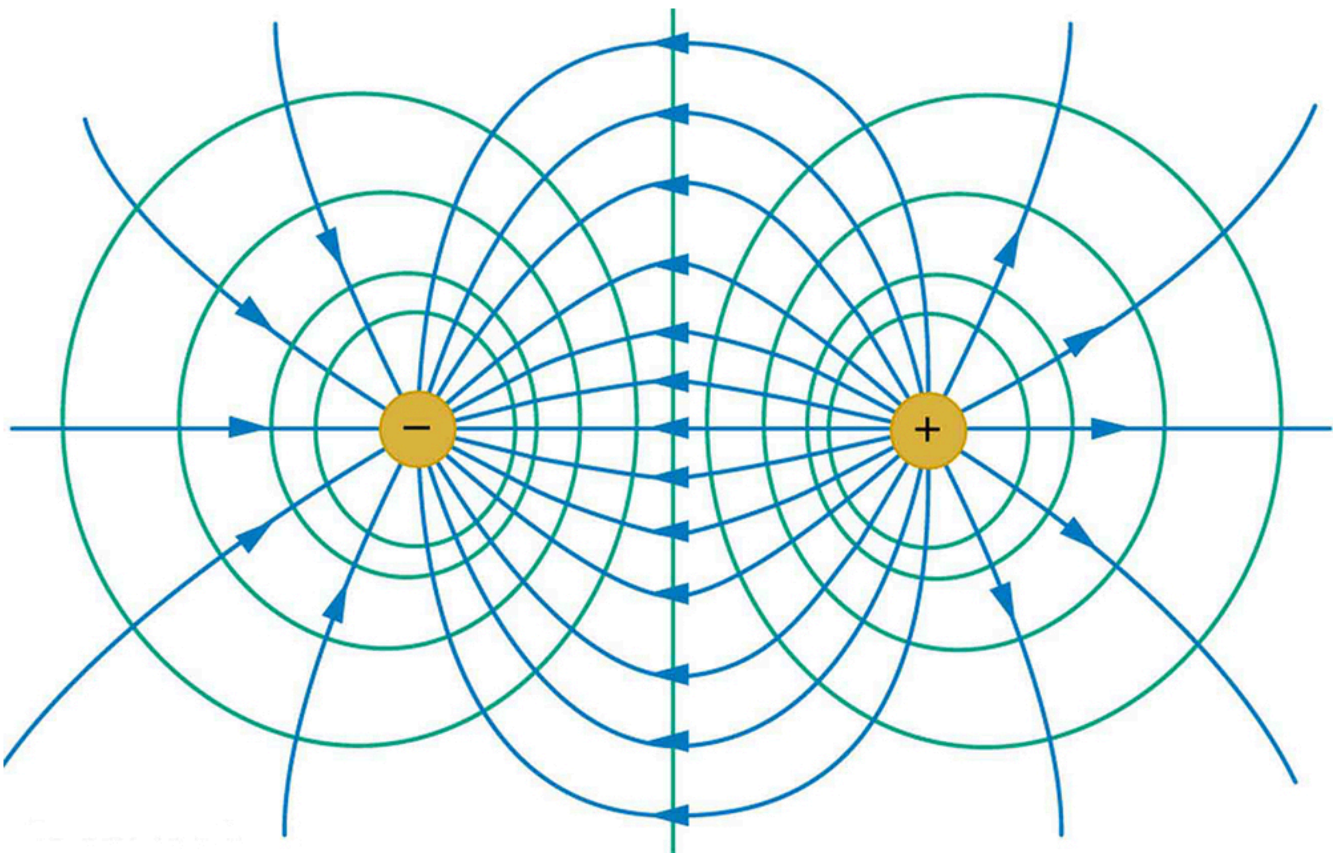


Figure 7: The electric field lines and equipotential lines for two equal but opposite charges. The equipotential lines can be drawn by making them perpendicular to the electric field lines, if those are known. Note that the potential is greatest (most positive) near the positive charge and least (most negative) near the negative charge.

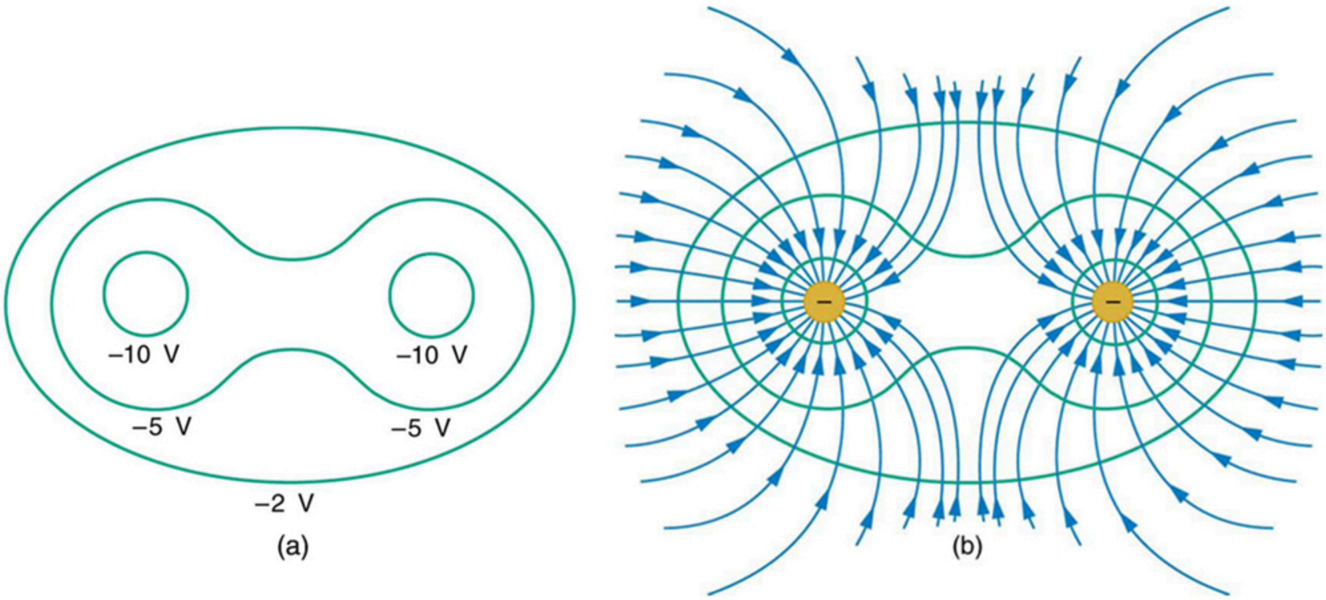


Figure 8: (a) These equipotential lines might be measured with a voltmeter in a laboratory experiment. (b) The corresponding electric field lines are found by drawing them perpendicular to the equipotentials. Note that these fields are consistent with two equal negative charges.

Section Summary

- An equipotential line is a line along which the electric potential is constant.
- An equipotential surface is a three-dimensional version of equipotential lines.
- Equipotential lines are always perpendicular to electric field lines.

Homework Problems

Problem 18: Electric field lines are always___________.

Problem 19: Electric field lines ____________.

The Relationship Between Electric Potential and Electric Field

We now want to explore the relationship between electric field $\vec{E}$ and electric potential V. These ideas are simply two different ways of looking at the same phenomenon: two like-charges repel and opposite charges attract. One approach uses forces and the other uses energy. An analogous situation from 131 would be the falling of a ball due to gravity: I can either think about the force of gravity $\vec{F}_g = m\vec{g}$ causing the ball to accelerate down and increase speed, or I can think about the ball exchanging gravitational potential energy $U_g = mgh$ for kinetic energy $K = \frac{1}{2}mv^2$. The final answer for the ball's speed when it hits the ground is the same regardless of approach. In this section, we want to explore how to convert from one of these pictures to the other, and, along the way, discover a different (but equivalent) unit for electric field.

In Figure 6 we see the two approaches applied to a nucleus attracting an electron. In one picture, the nucleus generates an electric field

$$\left|\vec{E}\right| = \frac{1}{4\pi\varepsilon_0}\frac{q_{\text{nucleus}}}{r^2}$$

The electric field points away from the nucleus. The electron then interacts with that field and feels a force $\vec{F} = q\vec{E}$. In the second picture, the nucleus generates an electric potential

$$V = \frac{1}{4\pi\varepsilon_0}\frac{q_{\text{nucleus}}}{r}$$

which decreases from the nucleus outwards. The electron then interacts with this potential by feeling a potential energy $U = qV$.

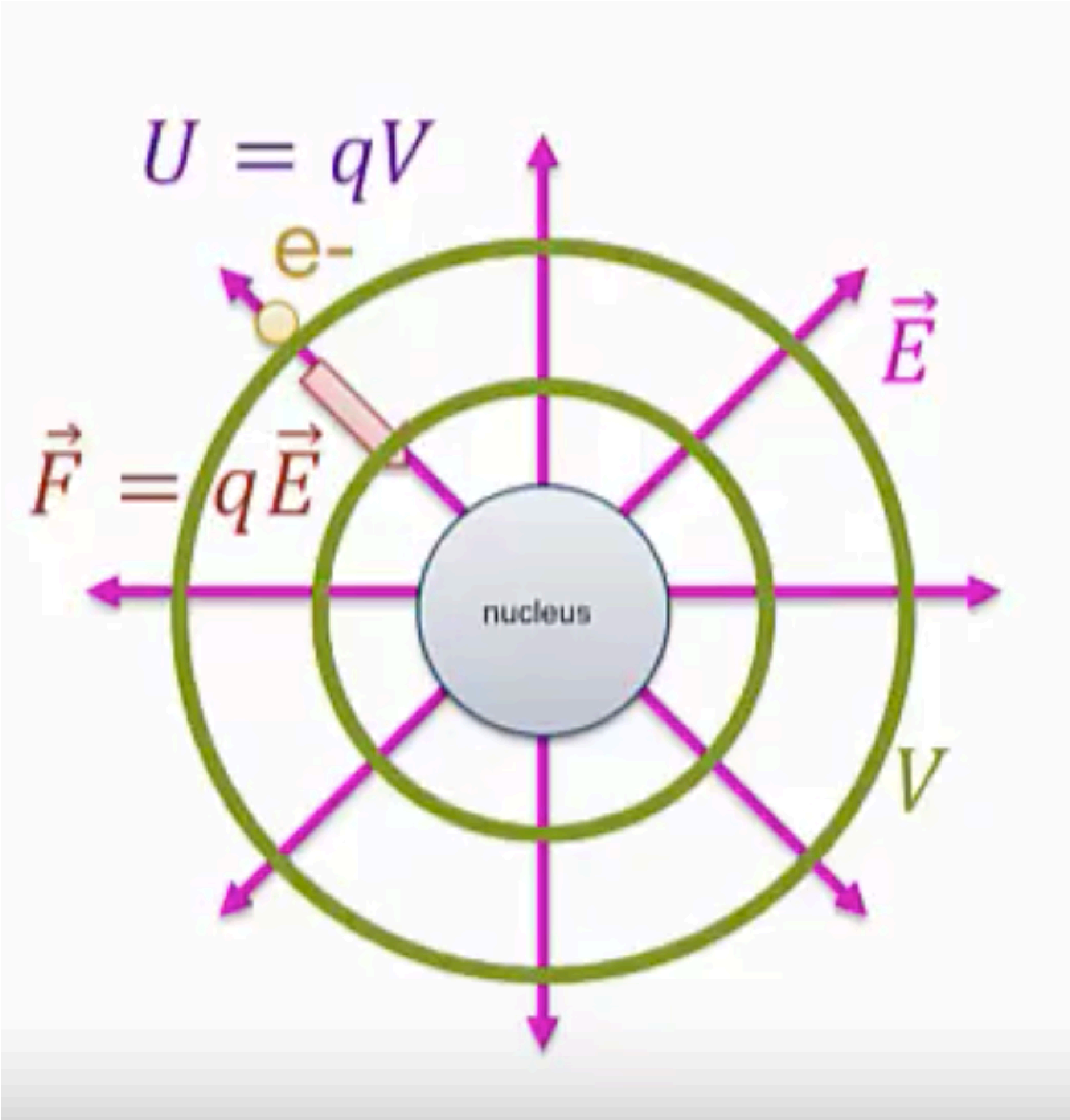


Figure 6: The two ways of picturing how an electron and a nucleus interact: electric fields causing electric forces and electric potentials causing electric potential energies.

How are these two pictures, electric fields $\vec{E}$, and electric potentials V related? Looking at the two formulas

$$\left|\vec{E}\right| = \frac{1}{4\pi\varepsilon_0}\frac{q_{\text{nucleus}}}{r^2} V = \frac{1}{4\pi\varepsilon_0}\frac{q_{\text{nucleus}}}{r}$$

we can see a relationship: the electric field is simply the potential divided by r! While there are formally some holes in this mathematical reasoning, the fundamental result is correct:

$$|E| = \frac{\Delta V}{\Delta s}$$

The magnitude of the electric field is the change in potential between the two points divided by the distance between those two points. This is a slope (i.e. derivative) thing like velocity or acceleration: to get the electric field at a point, you need to look at the change in potential immediately on either side and divide by the tiny distance between them. *Only for uniform fields will this equation give exact results, otherwise it gives an average electric field* value.

One thing to note, is that the equation $|E| = \frac{\Delta V}{\Delta s}$ indicates that the units for electric field will be Volts/meter $\mathrm{V/m}$. However, we already know from $\vec{F} = q\vec{E}$ that the units for electric field are Newtons/Coulomb $\mathrm{N/C}$. We are therefore, forced to conclude that these two units are the same:

$$1\frac{\mathrm{Newtons}}{\mathrm{Coulomb}} = 1\frac{\mathrm{Volt}}{\mathrm{meter}}$$

If we "break up" a Newton $1\mathrm{N} = \frac{\mathrm{kg \cdot m}}{\mathrm{s^2}}$, a Volt $1\mathrm{V} = 1\frac{\mathrm{J}}{\mathrm{C}}$, and a Joule $1\mathrm{J} = \frac{\mathrm{kg \cdot m^2}}{\mathrm{s^2}} = \mathrm{N \cdot m}$, we can see that they are the same:

$$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{V}}{\mathrm{m}}$$
$$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{J}}{\mathrm{C}}\frac{1}{\mathrm{m}}$$
$$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{N \cdot m}}{\mathrm{C}}\frac{1}{\mathrm{m}}$$
$$\frac{\mathrm{N}}{\mathrm{C}} = \frac{\mathrm{N}}{\mathrm{C}}$$

(The fact that we end up with an obviously true statement of N/C = N/C means that our starting assertion that N/C = V/m was true). Since Volts are much easier to measure and control in the lab, the units of V/m are probably more commonly used than N/C.

There is one final issue we need to address: the electric field is a *vector* having magnitude and direction, while the potential is a *scalar*, having only a magnitude. How can we determine the direction? Again, we turn to gravity for an analogy. We notice that the force of gravity points towards lower potential energy (down the hill). This is true for electricity as well: the electric field points down the "potential hill." In the example in Figure 6 above, the electric potential from the nucleus decreases with distance and the electric field points away from the nucleus. In general, the electric field points down the steepest slope in electric potential. We write this mathematically as

$$\vec{E} = -\frac{\Delta V}{\Delta s}$$

where the negative sign tells us that the electric field points down hill: from one equipotential to the next lower, always perpendicular to the equipotential lines as described in the previous section. We will practice this idea more in class.

What is the Highest Voltage Possible between Two Plates?

Dry air will support a maximum electric field strength of about $3.0 \times 10^6\ V/m$. Above that value, the field creates enough ionization in the air to make the air a conductor. This allows a discharge or spark that reduces the field. What, then, is the maximum voltage between two parallel conducting plates separated by 2.5 cm of dry air (as we will see in class, two parallel plates generate a uniform electric field)?

Strategy

We are given the maximum electric field E between the plates and the distance d between them. The equation $E = \frac{\Delta V}{\Delta s}$ with $\Delta s = d$ and $\Delta V = V_{AB}$ can thus be used to calculate the maximum voltage.

Solution

The potential difference or voltage between the plates is

$$V_{AB} = Ed.$$

Entering the given values for E and d gives

$$V_{AB} = (3.010^6 \text{ V/m})(0.025 \text{ m}) = 7.510^4 \ V$$

or

$$V_{AB} = 75 \ kV.$$

(The answer is quoted to only two digits, since the maximum field strength is approximate.)

Discussion

One of the implications of this result is that it takes about 75 kV to make a spark jump across a 2.5 cm (1 in.) gap, or 150 kV for a 5 cm spark. This limits the voltages that can exist between conductors, perhaps on a power transmission line. A smaller voltage will cause a spark if there are points on the surface, since points create greater fields than smooth surfaces. Humid air breaks down at a lower field strength, meaning that a smaller voltage will make a spark jump through humid air. The largest voltages can be built up, say with static electricity, on dry days.


scintillation counter
high voltage switch
scintillation counter

Figure 7. A spark chamber is used to trace the paths of high-energy particles. Ionization created by the particles as they pass through the gas between the plates allows a spark to jump. The sparks are perpendicular to the plates, following electric field lines between them. The potential difference between adjacent plates is not high enough to cause sparks without the ionization produced by particles from accelerator experiments (or cosmic rays). (credit: Daderot, Wikimedia Commons)

Field and Force inside an Electron Gun

(a) An electron gun has parallel plates separated by 4.00 cm and gives electrons 25.0 keV of energy. What is the electric field strength between the plates? (b) What force would this field exert on a piece of plastic with a $0.500\ \mu C$ charge that gets between the plates?

Strategy

Since the voltage and plate separation are given, the electric field strength can be calculated directly from the expression $E=\frac{\Delta V}{\Delta s}$. Once the electric field strength is known, the force on a charge is found using $\mathbf{F}=q\mathbf{E}$. Since the electric field is in only one direction, we can write this equation in terms of the magnitudes, $F=qE$.

Solution for (a)

The expression for the magnitude of the electric field between two uniform metal plates is

$$E=\frac{\Delta V}{\Delta s}.$$

Since the electron is a single charge and is given 25.0 keV of energy, the potential difference must be 25.0 kV. Entering this value for V_{AB} and the plate separation of 0.0400 m, we obtain

$$E=\frac{25.0\ \mathrm{kV}}{0.0400\ \mathrm{m}}=6.2510^{5}\ \mathrm{V/m}.$$

Solution for (b)

The magnitude of the force on a charge in an electric field is obtained from the equation

$$F=qE$$

Substituting known values gives

$$F=(0.50010^{-6}\ \mathrm{C})(6.2510^{5}\ \mathrm{V/m})=0.313\ \mathrm{N}.$$

Discussion

Note that the units are newtons, since $1\mathrm{V/m}=1\ \mathrm{N/C}$. The force on the charge is the same no matter where the charge is located between the plates. This is because the electric field is uniform between the plates.

Homework Problems

Problem 20: Which of the following are units of electric field?

Problem 21: Membrane walls of living cells have surprisingly large electric fields across them due to separation of ions. What is the voltage across a membrane given an electric field strength across it?

Problem 22: What is the potential difference between the plates given the electric field and separation? The plate with the lowest potential is taken to be at zero volts. What is the potential 1.0">1.0 cm">cm from that plate?

Problem 23: Find the maximum potential difference between two parallel conducting plates separated by some amount of air, given the maximum sustainable electric field strength in air to be 3.00 MV/m.

A PhET to Explore These Ideas

An interactive or media element has been excluded from this version of the text. You can view it online here:
http://openbooks.library.umass.edu/toggerson-132/?p=245

A few things to play around with in the simulation above:

1. Add a positive and negative charge with about 5 cm of space between them. Describe what the electric field looks like.
2. Use the device to plot equipotential lines (locations where the electric potential is the same). Describe what the equipotentials look like.
3. Is there a relationship between the electric field and the equipotentials?
4. What happens if you add more charges?

24. Homework Problems

Homework

The list below is the list of homework problems in Edfinity. The numbering is the same. You can click on a problem, and it will take you to the relevant section of the book!

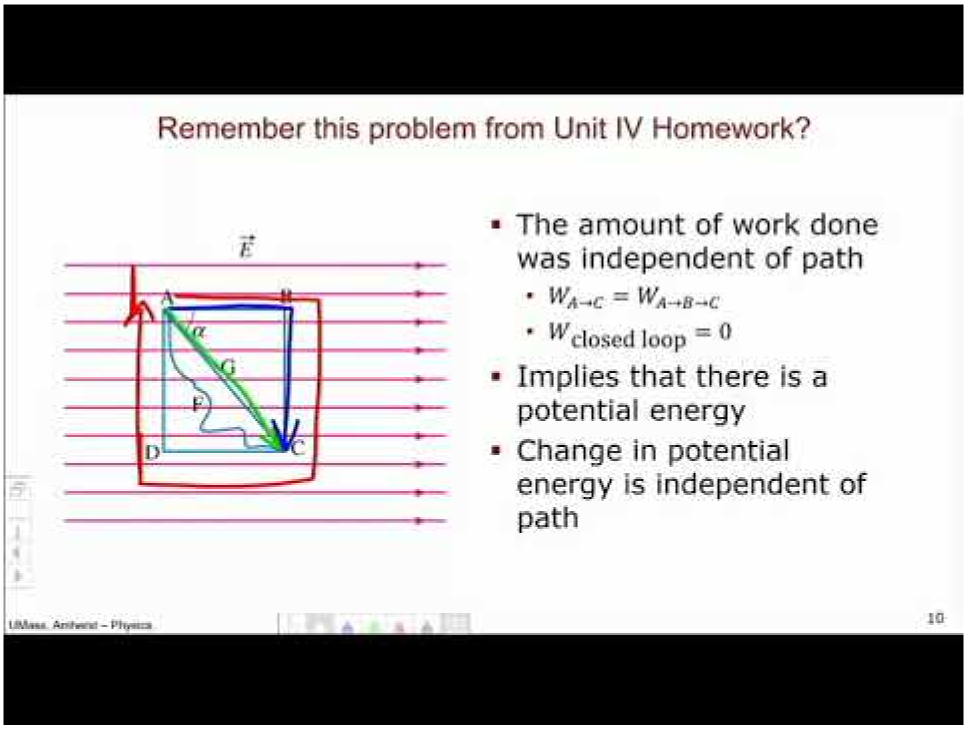
1. Comparing the energies of bound atoms vs. free atoms.
2. Reviewing gel electrophoresis.
3. Converting from charge to number of particles.
4. Find the magnitudes of the forces F_1 and F_2 and that add to give the total force F_{total} shown above. This may be done either graphically or by using trigonometry.
5. Suppose you walk 17.0m straight west and then 23.0m straight south. How far are you from your starting point, and what is the compass direction of a line connecting your starting point to your final position?
6. To start a car engine, the car battery moves 3.76×10^{21} electrons through the starter motor. How many coulombs of charge were moved?
7. A certain lightning bolt moves 35.8 C of charge. How many fundamental units of charge $|q_e|$ is this?
8. Electrically neutral objects can exert a gravitational force on each other, but they cannot exert an electrical force on each other.
9. What is the magnitude of the force exerted on a 4.49 C charge by a 310.24 N/C electric field that points due east?
10. Find the direction and magnitude of an electric field that exerts a 9×10^{-17} N westward force on an electron.
11. Calculate the initial (from rest) acceleration of a proton in a 3.09×10^{6} N/C electric field (such as created by a research Van de Graaff).
12. Suppose there is a single electron a set distance from a point charge Q, which quantities does the electric field experienced by the electron depend on?
13. Calculate the magnitude of the electric field 8.41">8.41 m from a point charge of 8.92">8.92 mC (such as found on the terminal of a Van de Graaff).
14. What magnitude point charge creates a 15947.46">15947.46 N/C electric field at a distance of 0.513">0.513 m? How large is the field at 18.85">18.85 m?
15. Below you see an unknown charge generating an electric field. Also indicated are two empty regions of space “A” and “B.” Which of the following statements are true from this picture of field lines?
16. A lightning bolt strikes a tree, moving 27 C of charge through a potential difference of 1.10×10^{2} MV. What energy was dissipated?
17. An evacuated tube uses an accelerating voltage of 41">41 kV">kV to accelerate electrons to hit a copper plate and produce x rays. What would be the final speed of such an electron?
18. What is the potential 52.92 pm from a proton (the average distance between the proton and electron in a hydrogen atom)?
19. A research Van de Graaff generator has a 2.01-m-diameter metal sphere with a charge of 5.25 mC on it. What is the potential near its surface?
20. Electric field lines are always___________.

21. Electric field lines ____________.
22. Which of the following are units of electric field?
23. Membrane walls of living cells have surprisingly large electric fields across them due to separation of ions. What is the voltage across an 8.00 nmthick membrane if the electric field strength across it is 5.25 MV/m? You may assume a uniform electric field.
24. The electric field strength between two parallel conducting plates separated by 6.4 cm is 7.0×10^4 V/m. What is the potential difference between the plates? The plate with the lowest potential is taken to be at zero volts. What is the potential 1.0 cm from that plate?
25. Find the maximum potential difference between two parallel conducting plates separated by 0.55 cm of air, given the maximum sustainable electric field strength in air to be 3.00 MV/m.

PART IV

UNIT IV

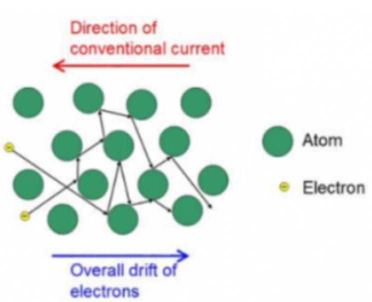
Unit IV On-a-Page



A YouTube element has been excluded from this version of the text. You can view it online here: [*http://openbooks.library.umass.edu/toggerson-132/?p=429*](http://openbooks.library.umass.edu/toggerson-132/?p=429)

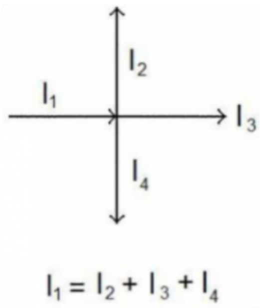
Current

- Current and charge must be conserved item.
- The current is the amount of charge per second $I = \frac{\Delta Q}{\Delta t}$
- units: C/s = A
- Direction of current is direction of positive charges' motion



Kirchhoff's Rules (Or "How to analyze a circuit")

- The amount of current going into a junction must equal the amount going out (otherwise it would pile up)

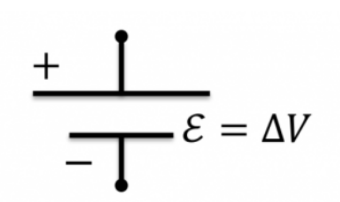


- The changes in potential around any closed loop is zero

Circuit Elements

Batteries

- The potential drop across a battery is fixed



Resistors

- Resistors dissipate electrical potential energy into something else (heat, motion, light, etc.)
- The potential drop across a resistor is related to the current and resistance

$$\Delta V = IR$$

- The resistance is fixed by the material, units:

$$\frac{V}{A} = \Omega$$

- The potential drop across a wire is zero ($R = 0$)

Capacitors

- A capacitor is two pieces of metal that don't touch
- Capacitors store charges separated. Therefore, they store energy
- The capacitance is a property of the geometry of the plates
- The potential drop across a capacitor is related to the charge and the capacitance:

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

- For parallel plates

$$C_{\|} = \varepsilon \frac{A}{d}$$

- Units

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

Power

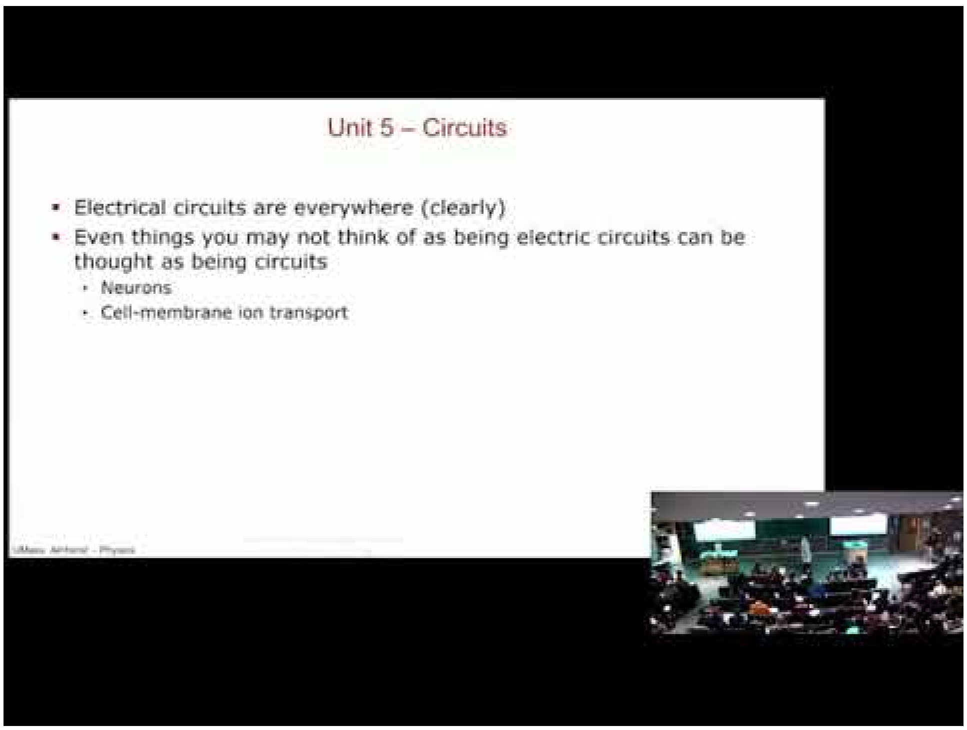
- The power provided/stored/dissipated by a circuit element is

$$P = IV$$

- Must do element-by-element

25. Introduction and Motivating Biological Context for Unit IV

Introduction



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=432

How many electrical devices do you have on you at the moment?

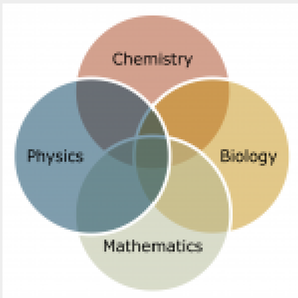
Electrical circuits are everywhere, I have my tablet, tablet pen, the remote, the microphone, my phone, and my watch. I've got six different electrical circuits on me right now.

There are a lot of things that you might not think of as being an electric circuit, but we can analyze using the properties of circuits, which we are going to discuss over the course of this unit.

By the end of this unit, we will be able to draw a rough circuit and think about a neuron as a circuit. A cell membrane is essentially a capacitor, so we can think about cell membrane ion transport, so we'll be able to talk a little bit about cell membranes in this context.

We will introduce what a current is but other than that there are not a lot of new fundamental physics in this unit, we're mostly going to be dealing with the idea that if I put a charge and a potential, I get a potential energy. That is going to be the vast majority of what we're going to be working with. It's really a lot of application of those ideas and seeing how things work.

Motivating Biological Context for Unit IV – The Neuron



In this unit, we will be looking at the neuron throughout. This section, from [OpenStax Biology – How Neurons Communicate] is to refresh your biology knowledge on these cells.

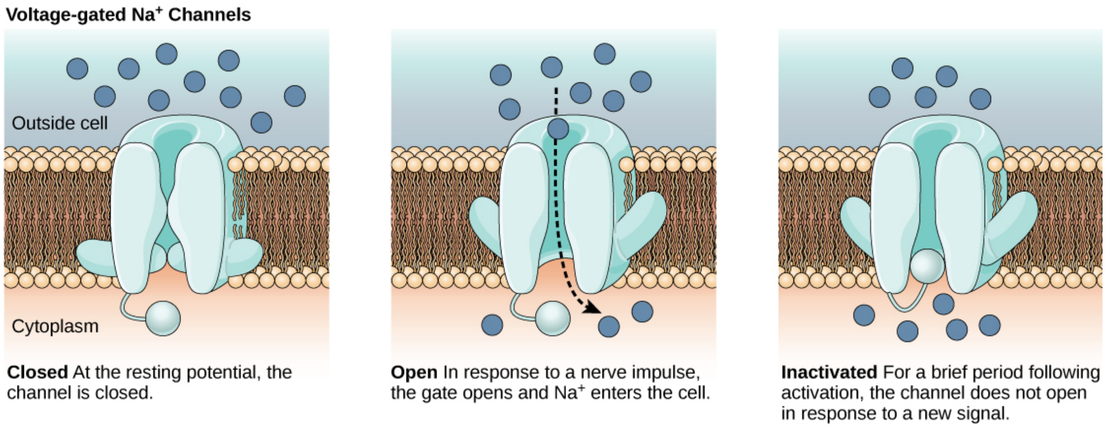
All functions performed by the nervous system—from a simple motor reflex to more advanced functions like making a memory or a decision—require neurons to communicate with one another. While humans use words and body language to communicate, neurons use electrical and chemical signals. Just like a person in a committee, one neuron usually receives and synthesizes messages from multiple other neurons before "making the decision" to send the message on to other neurons.

Nerve Impulse Transmission within a Neuron

For the nervous system to function, neurons must be able to send and receive signals. These signals are possible because each neuron has a charged cellular membrane (a voltage difference between the inside and the outside), and the charge of this membrane can change in response to neurotransmitter molecules released from other neurons and environmental stimuli. To understand how neurons communicate, one must first understand the basis of the baseline or 'resting' membrane charge.

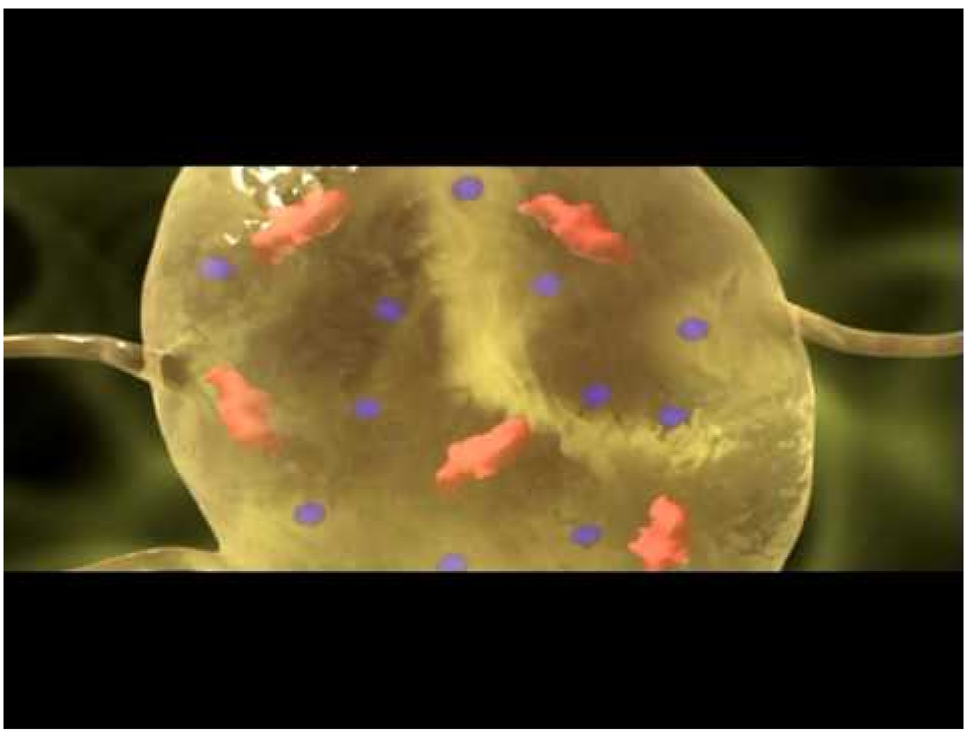
Neuronal Charged Membranes

The lipid bilayer membrane that surrounds a neuron is impermeable to charged molecules or ions. To enter or exit the neuron, ions must pass through special proteins called ion channels that span the membrane. Ion channels have different configurations: open, closed, and inactive, as illustrated in the figure below. Some ion channels need to be activated in order to open and allow ions to pass into or out of the cell. These ion channels are sensitive to the environment and can change their shape accordingly. Ion channels that change their structure in response to voltage changes are called voltage-gated ion channels. Voltage-gated ion channels regulate the relative concentrations of different ions inside and outside the cell. The difference in total charge between the inside and outside of the cell is called the membrane potential.



Voltage-gated ion channels open in response to changes in membrane voltage. After activation, they become inactivated for a brief period and will no longer open in response to a signal.

This video discusses the basis of the resting membrane potential.



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=432

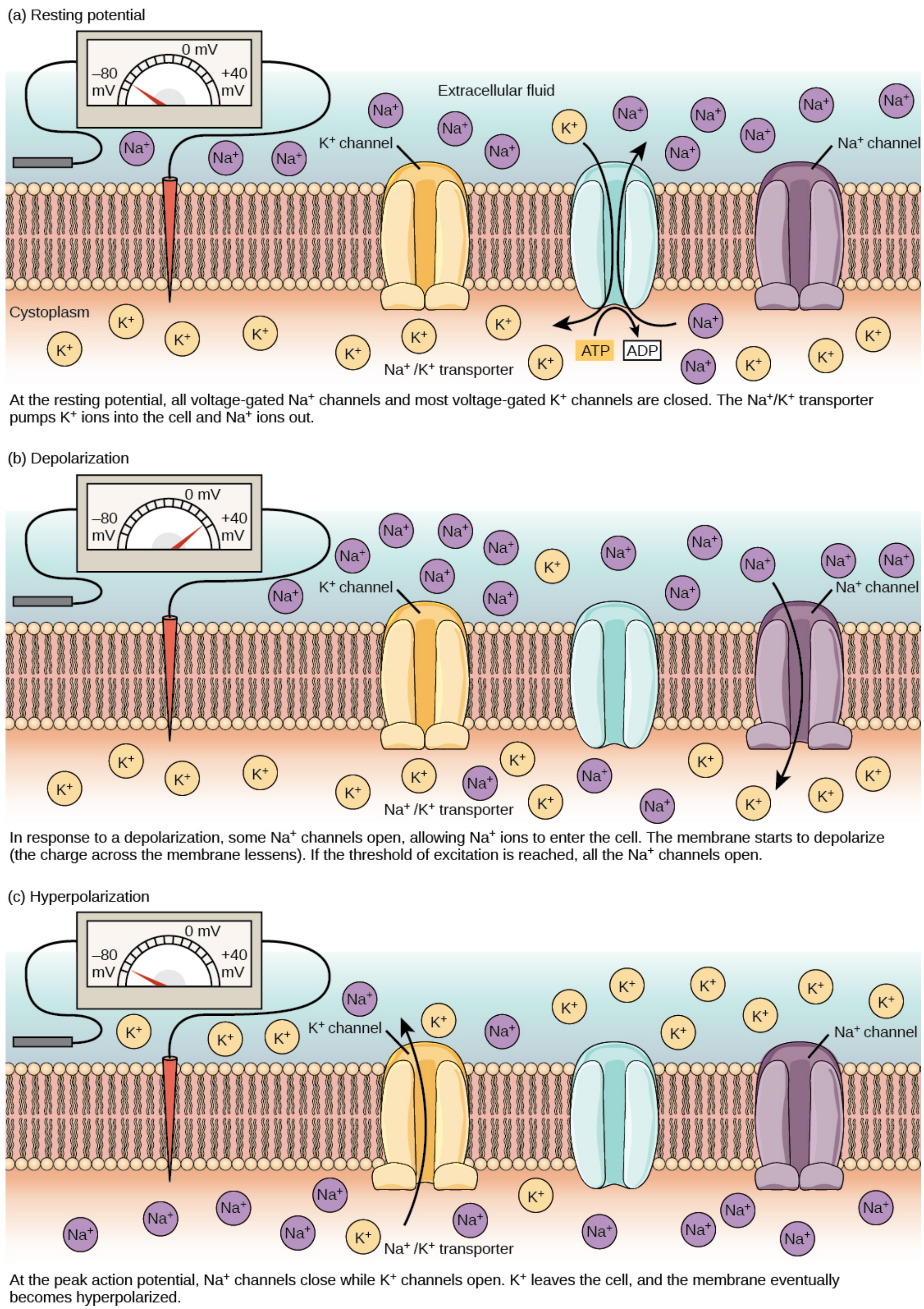
Resting Membrane Potential

A neuron at rest is negatively charged: the inside of a cell is approximately 70 millivolts more negative than the outside (−70 mV, note that this number varies by neuron type and by species). This voltage is called the resting membrane potential; it is caused by differences in the concentrations of ions inside and outside the cell. If the membrane were equally permeable to all ions, each type of ion would flow across the membrane and the system would reach equilibrium. Because ions cannot simply cross the membrane at will, there are different concentrations of several ions inside and outside the cell, as shown in the table below. The difference in the number of positively charged potassium ions (K^+) inside and outside the cell dominates the resting membrane potential (figure below table). When the membrane is at rest, K^+ ions accumulate inside the cell due to a net movement with the concentration gradient. The negative resting membrane potential is created and maintained by increasing the concentration of cations outside the cell (in the extracellular fluid) relative to inside the cell (in the cytoplasm). The negative charge within the cell is created by the cell membrane being more permeable to potassium ion movement than sodium ion movement. In neurons, potassium ions are maintained at high concentrations within the cell while sodium ions are maintained at high concentrations outside of the cell. The cell possesses potassium and sodium leakage channels that allow the two cations to diffuse down their concentration gradient. However, the neurons have far more potassium leakage channels than sodium leakage channels. Therefore, potassium diffuses out of the cell at a much faster rate than sodium leaks in. Because more cations are leaving the cell than are entering, this causes the interior of the cell to be negatively charged relative to the outside of the cell. The actions of the sodium potassium pump help to maintain the resting potential, once established. Recall that sodium potassium pumps brings two K^+ ions into the cell while removing three Na^+ ions per ATP consumed. As more cations are expelled from the cell than taken in, the inside of the cell remains negatively charged relative to the extracellular fluid. It should be noted that chloride ions (Cl^-) tend to accumulate outside of the cell because they are repelled by negatively-charged proteins within the cytoplasm.

Ion Concentration Inside and Outside Neurons

Ion	Extracellular concentration (mM)	Intracellular concentration (mM)	Ratio outside/inside
Na^+	145	12	12
K+	4	155	0.026
Cl^-	120	4	30
Organic anions (A−)	—	100	

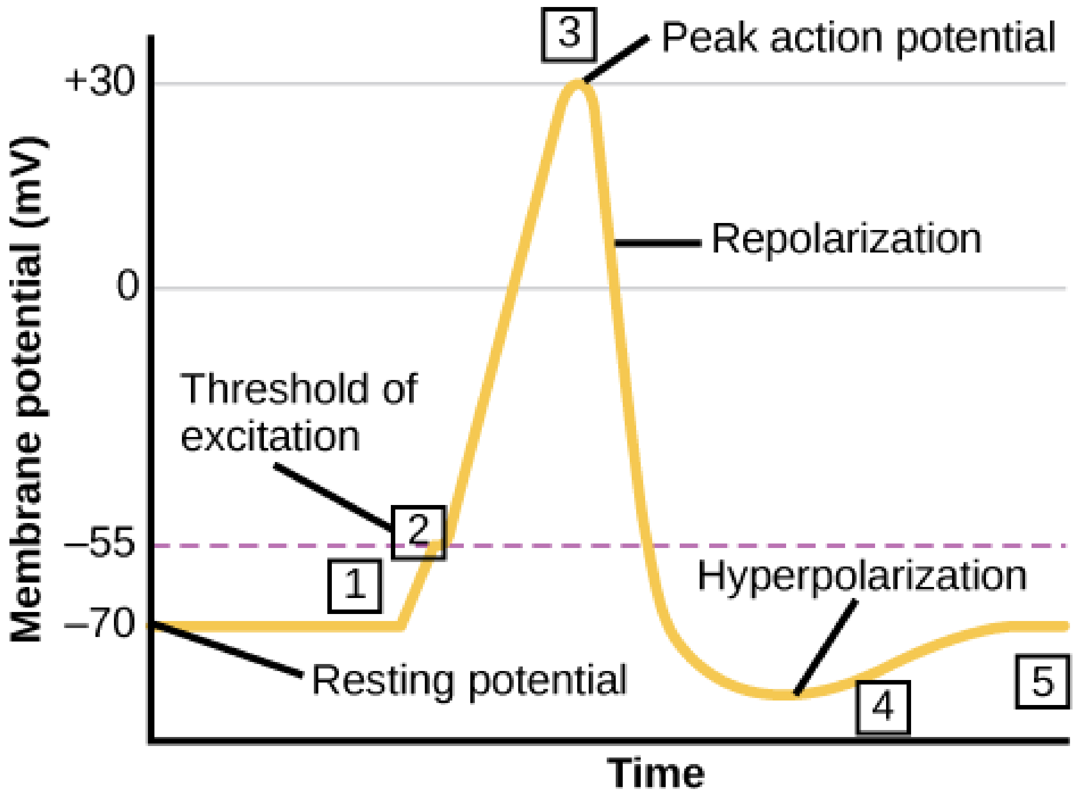
The resting membrane potential is a result of different concentrations inside and outside the cell.



The (a) resting membrane potential is a result of different concentrations of Na+ and K+ ions inside and outside the cell. A nerve impulse causes Na+ to enter the cell, resulting in (b) depolarization. At the peak action potential, K+ channels open and the cell becomes (c) hyperpolarized.

Action Potential

A neuron can receive input from other neurons and, if this input is strong enough, send the signal to downstream neurons. Transmission of a signal between neurons is generally carried by a chemical called a neurotransmitter. Transmission of a signal within a neuron (from dendrite to axon terminal) is carried by a brief reversal of the resting membrane potential called an action potential. When neurotransmitter molecules bind to receptors located on a neuron's dendrites, ion channels open. At excitatory synapses, this opening allows positive ions to enter the neuron and results in depolarization of the membrane—a decrease in the difference in voltage between the inside and outside of the neuron. A stimulus from a sensory cell or another neuron depolarizes the target neuron to its threshold potential (-55 mV). Na^+ channels in the axon hillock open, allowing positive ions to enter the cell (figure above and graph below). Once the sodium channels open, the neuron completely depolarizes to a membrane potential of about +40 mV. Action potentials are considered an "all-or nothing" event, in that, once the threshold potential is reached, the neuron always completely depolarizes. Once depolarization is complete, the cell must now "reset" its membrane voltage back to the resting potential. To accomplish this, the Na^+ channels close and cannot be opened. This begins the neuron's refractory period, in which it cannot produce another action potential because its sodium channels will not open. At the same time, voltage-gated K^+ channels open, allowing K^+ to leave the cell. As K^+ ions leave the cell, the membrane potential once again becomes negative. The diffusion of K^+ out of the cell actually hyperpolarizes the cell, in that the membrane potential becomes more negative than the cell's normal resting potential. At this point, the sodium channels will return to their resting state, meaning they are ready to open again if the membrane potential again exceeds the threshold potential. Eventually the extra K^+ ions diffuse out of the cell through the potassium leakage channels, bringing the cell from its hyperpolarized state, back to its resting membrane potential.



The formation of an action potential can be divided into five steps: (1) A stimulus from a sensory cell or another neuron causes the target cell to depolarize toward the threshold potential. (2) If the threshold of excitation is reached, all Na+ channels open and the membrane depolarizes. (3) At the peak action potential, K+ channels open and K+ begins to leave the cell. At the same time, Na+ channels close. (4) The membrane becomes hyperpolarized as K+ ions continue to leave the cell. The hyperpolarized membrane is in a refractory period and cannot fire. (5) The K+ channels close and the Na+/K+ transporter restores the resting potential.

Potassium channel blockers, such as amiodarone and procainamide, which are used to treat abnormal electrical activity in the heart, called cardiac dysrhythmia, impede the movement of K^+ through voltage-gated K^+ channels. Which part of the action potential would you expect potassium channels to affect?

Homework

1. Which ions are important in understanding neuron function?

26. Current

Electric Current

Editors' note: This section is derived from [Current](#) by OpenStax

Instructor's Notes

By the end of this section you should know:

- Electric current I is the rate at which charge flows, given by

$$I = \frac{\Delta Q}{\Delta t},$$

where ΔQ is the amount of charge passing through an area in time Δt.

- 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 $1\ \text{A} = 1\ \text{C/s}$.
- Current is the flow of free charges, such as electrons and ions.
- Drift velocity 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 $I = nqAv_\text{d}$. Here, I is the current through a wire of cross-sectional area A. The wire's material has 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.

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

$$I = \frac{\Delta Q}{\Delta t},$$

where ΔQ

is the amount of charge passing through a given area in time Δt. (As in previous chapters, initial time is often taken to be zero, in which case $\Delta t = t$.) (See Figure 1.) The SI unit for current is the **ampere** (A), named for the French physicist André-Marie Ampère (1775–1836). Since $I = \Delta Q / \Delta t$, we see that an ampere is one coulomb per second:

$$1 \text{ A} = 1 \text{ C/s}$$

Not only are fuses and circuit breakers rated in amperes (or amps), so are many electrical appliances.

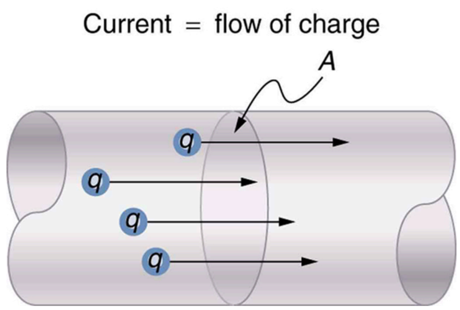


Figure 1. *The rate of flow of charge is current. An ampere is the flow of one coulomb through an area in one second.*

Calculating Currents: Current in a Truck Battery and 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 for (a)

Entering the given values for charge and time into the definition of current gives

$$I = \frac{\Delta Q}{\Delta t} = \frac{720 \text{ C}}{4.00 \text{ s}} = 180 \text{ C/s}$$
$$= 180 \text{ A}$$

Discussion for (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 for (b)

Solving the relationship $I = \Delta Q/\Delta t$ for time Δt, and entering the known values for charge and current gives

$$\Delta t = \frac{\Delta Q}{I} = \frac{1.00\ \text{C}}{0.300 \times 10^{-3}\ \text{C/s}}$$

$$= 3.33 \times 10^{3}\ \text{s}$$

Discussion for (b)

This time is slight less than an hour. The small current used by the handheld 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 a 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 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 2(b), 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.

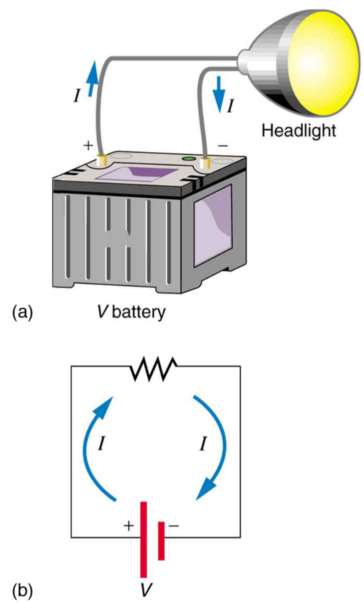


Figure 2. *(a) A simple electric circuit. A closed path for current to flow through is supplied by conducting wires connecting a load to the terminals of a battery. (b) In this schematic, the battery is represented by the two parallel red lines, conducting wires are shown as straight lines, and the zigzag represents the load. The schematic represents a wide variety of similar circuits.*

Note that the direction of current flow in Figure 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 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.

It is important to realize that there is an electric field in conductors responsible for producing the current, as illustrated in Figure 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.

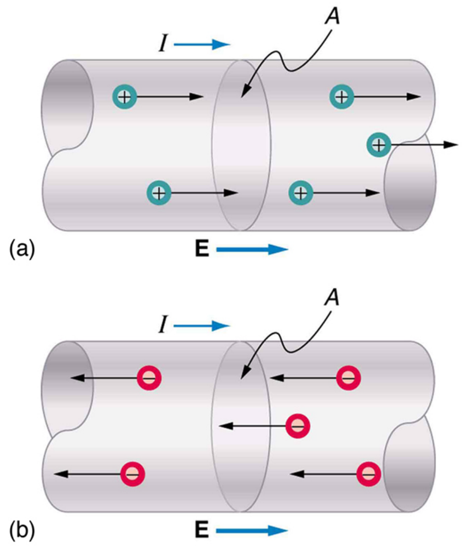


Figure 3. *Current I is the rate at which charge moves through an area A, such as the cross-section of a wire. Conventional current is defined to move in the direction of the electric field. (a) Positive charges move in the direction of the electric field and the same direction as conventional current. (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.*

Calculating the Number of Electrons that Move through a Calculator

If the 0.300-mA current through the calculator mentioned in the Example above 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, $I_{electrons} = -0.300 \times 10^{-3}\ \mathrm{C/s}$. Since each electron ($e^{-}$) has a charge of $-1.60 \times 10^{-19}\ \mathrm{C}$, we can convert the current in coulombs per second to electrons per second.

Solution

Starting with the definition of current, we have

$$I_{electrons} = \frac{\Delta Q_{electrons}}{\Delta t} = \frac{-0.300 \times 10^{-3}\ \mathrm{C}}{s}.$$

We divide this by the charge per electron, so that

$$\frac{e}{s} = \frac{-0.300 \times 10^{3}\ \mathrm{C}}{s} \times \frac{1\ \mathrm{e}}{-1.60 \times 10^{-19}\ \mathrm{C}}$$

$$= 1.88 \times 10^{15} \frac{e}{s}$$

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.

Homework

2. Direction of charge flow and current
3. How much charge in a defibrillator?
4. How long is a lightning bolt?

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 $10^{8}\ \mathrm{m/s}$, 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 $10^{-4}\ \mathrm{m/s}$. 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 Figure 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.

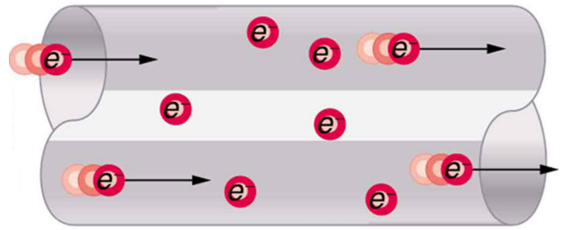


Figure 4. *When charged particles are forced into this volume of a conductor, an equal number are quickly forced to leave. The repulsion between like charges makes it difficult to increase the number of charges in a volume. Thus, as one charge enters, another leaves almost immediately, carrying the signal rapidly forward.*

Good conductors have large numbers of free charges in them. In metals, the free charges are free electrons. Figure 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} is 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.

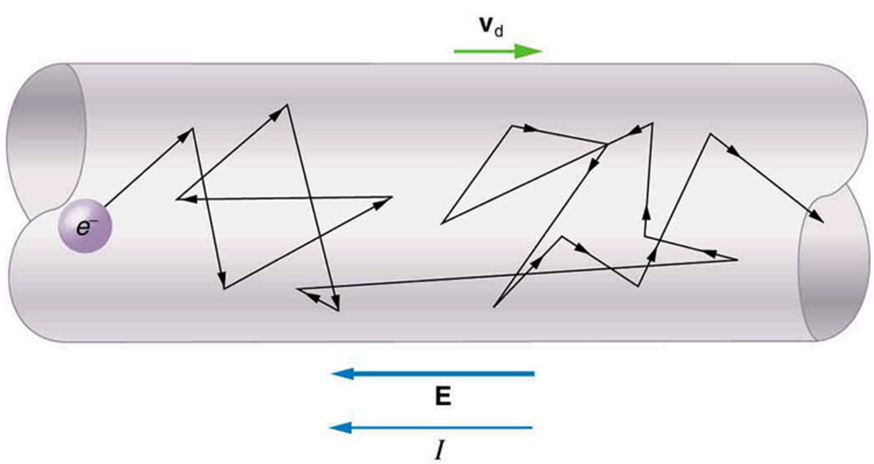


Figure 5. *Free electrons moving in a conductor make many collisions with other electrons and atoms. The path of one electron is shown. The average velocity of the free charges is called the drift velocity, and it is in the direction opposite to the electric field for electrons. The collisions normally 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 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 ΔQ in this segment is thus $qnAx$ where q is the amount of charge on each carrier. (Recall that for electrons, q is $-1.60 \times 10^{-19}\ \text{C}$. Current is charge moved per unit time; thus, if all the original charges move out of this segment in time Δt, the current is

$$I = \frac{\Delta Q}{\Delta t} = \frac{qnAx}{\Delta t}.$$

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 Δt. Rearranging terms gives

$$I = nqAv_{\text{d}},$$

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}.

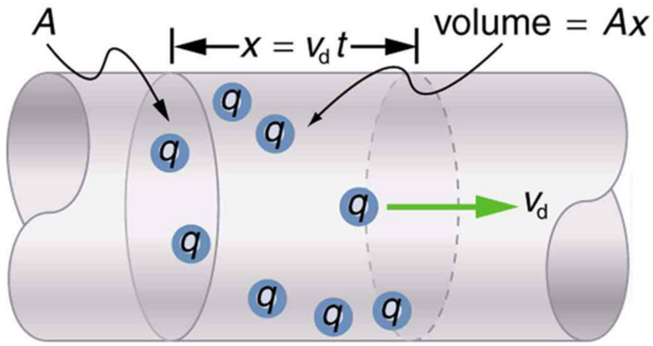


***Figure 6**. All the charges in the shaded volume of this wire move out in a time t, having a drift velocity of magnitude $v_\mathrm{d} = x/t$. See text for further discussion.*

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.

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 $8.80 \times 10^3\ \mathrm{kg/m}^3$.

Strategy

We can calculate the drift velocity using the equation $I = nqAv_\mathrm{d}$. The current $I = 20.0\ \mathrm{A}$ is given, and $q = -1.60 \times 10^{-19}\ \mathrm{C}$ 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, $8.80 \times 10^3\ \mathrm{kg/m}^3$, and the periodic table shows that the atomic mass of copper is 63.54 g/mol. We can use these two quantities with Avogadro's number, $6.02 \times 10^{23}\ \mathrm{atoms/mol}$, 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:

$$n = \frac{1\ e^-}{atom} \times \frac{6.02 \times 10^{23}\ \text{atoms}}{mol} \times \frac{1\ \text{mol}}{63.54\ \text{g}} \times \frac{1000\ \text{g}}{kg} \times \frac{8.80 \times 10^3\ \text{kg}}{1\ \text{m}^3}$$
$$= 8.342 \times 10^{23}\ e^-/\text{m}^3.$$

The cross-sectional area of the wire is

$$A = \pi r^2$$
$$= \pi \left(\frac{2.053 \times 10^{-3}\ \text{m}}{2}\right)^2$$
$$= 3.310 \times 10^{-6}\ \text{m}^2.$$

Rearranging $I = nqAv_\text{d}$ to isolate drift velocity gives

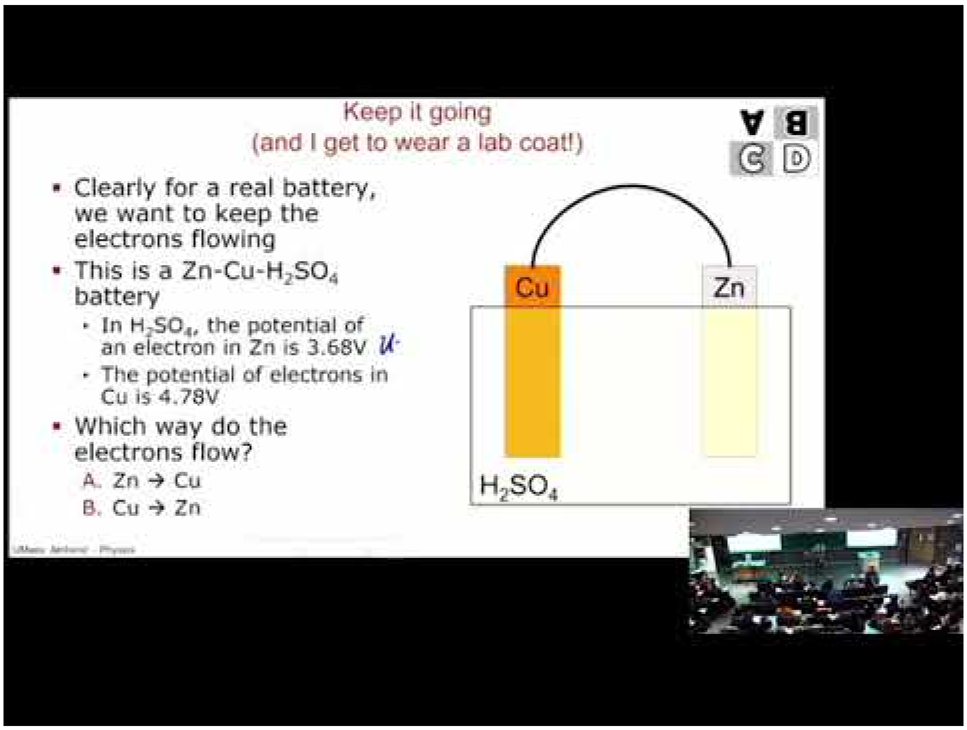
$$v_\text{d} = \frac{I}{nqA}$$
$$= \frac{20.0\ \text{A}}{(8.342 \times 10^{23}/\text{m}^3)(-1.60 \times 10^{19}\ \text{C})(3.310 \times 10^6\ \text{m}^2)}$$
$$= -4.53 \times 10^{-4}\ \text{m/s}$$

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 $10^{-4}\ \text{m/s}$) confirms that the signal moves on the order of 10^{12} times faster (about $10^8\ \text{m/s}$) than the charges that carry it.

27. Circuit Elements

Ideal Batteries



A YouTube element has been excluded from this version of the text. You can view it online here: *http://openbooks.library.umass.edu/toggerson-132/?p=451*

This section is about the power coming out of the wall or a battery, which is probably the first thing you think of when you start thinking about electricity and electric circuits.

One of the important things is what is the electron volt as a unit, to go through the math, we're starting with the definition of potential, if we put a charge in a potential we get a potential energy, ΔV, charge, Q is $1.602 \times 10^{-19}\,\mathrm{C}$ which gives me $1.602 \times 10^{-19}\,\mathrm{J}$ or one electron volt (1 eV). And in fact, this is what an electron volt is, the potential energy gained by a single electron going over a volt. We've mentioned this number before, the energy you need to ionize hydrogen gas is $13.6\,\mathrm{eV}$ electron volts, meaning that you have to give that electron 13.6 electron volts to rip it away from its nucleus. Based upon this you can say that the potential difference, not the potential energy, between the ground state of the electron and very far away is 13.6 volts.

Instructor’s Note

Keeping this distinction between electron volts as energy and volts as potential is going to be really important for this section.

Our question for this section is: what is a battery?

Essentially, it’s two pieces of metal in contact. If you think back to Unit 1 we talked about this photoelectric effect, and we explained how much energy you need to remove electrons from a metal. That was described by the work function, which we used to find how much energy we need to remove an electron from the surface of a metal.

Instructor’s Note

We connected this with photons and other things but the key thing you need right now is that there is this material dependent number that essentially tells you how much energy you need to rip an electron off the surface and it's called that work function.

If we take gold and platinum and touch them together, we can make a simple battery. Gold has a work function of about $5\ \mathrm{eV}$ and platinum have a work function of about $6\frac{1}{3}\ \mathrm{eV}$. This means that an electron in the surface of the gold has a potential energy of $-5.1\ \mathrm{eV}$ because the charge of an electron is negative, so the potential energy might take this positive and we get a negative potential.

What's going to happen when we touch these two things together?

Well the electron in the gold is going to slide down because now that's a lower energy state in gold it has a potential energy of $-5.1\ \mathrm{eV}$ in platinum it has a potential energy of $-6.35\ \mathrm{eV}$, it can lower its potential energy by $1\frac{1}{4}\ \mathrm{eV}$ by sliding down from the gold to the platinum.

The energy sliding the electrons sliding over it is losing $1.25\ \mathrm{eV}$, if the potential energy difference is $1.25\ \mathrm{eV}$ then the potential difference is $1.25\ \mathrm{V}$, so we have a potential difference and electrons spontaneously moving from two metals in contact.

So, what is a battery?

A potential difference that, if I connect the two metals across the terminals the electrons spontaneously move. This is essentially the simplest battery I can think of. Now this isn't a very good battery because the electrons from the gold are going to kind of run across which means we're going to end up with a negatively charged piece of platinum and a positively charged piece of gold. And eventually this repulsion is going to stop the flow, you're going to build up a negatively charged thing and electrons aren't going to flow anymore.

So that's what makes this kind of a crummy battery, the electrons flow once and then they stop pretty quickly, you get kind of a spark, you don't get a constant flow. But the potential difference between the two metals states that you keep that potential difference, the electrons just stop flowing because the repulsion.

The battery we just talked about, the gold and platinum in contact, well that's not particularly helpful because the electrons flow once and then you're done. We want the battery to keep running, so we're actually going to repeat a famous experiment by a guy named Volta, and so for his setup, shown below, we have a piece of copper and a piece of zinc, and they're not touching each other, and we have no potential difference between them. Then we bring them into contact.

Then we've got 1.5 moles per liter sulfuric acid and add this in until there's enough in there to bring the two metals into contact, through the sulfuric acid. So now we're getting a potential difference between these two pieces of metal, because now we've dropped them into contact using the sulfuric acid.

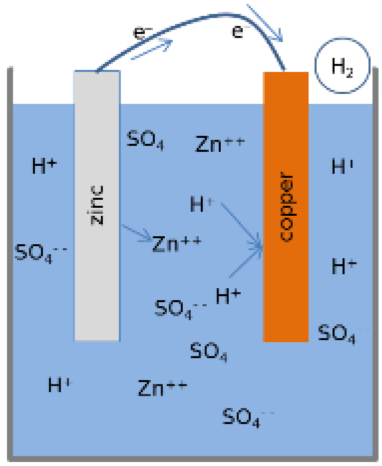


Figure 1. *A voltaic cell. (credit: Wikipedia)*

This is the first battery voltage, and you can see this potential difference is just sitting there, and these electrons keep flowing. In sulfuric acid, the potential of an electron in zinc is $3.68\ \mathrm{V}$, the fact that there is sulfuric acid changes these work function numbers. The charge of an electron is negative, so zinc has an electron and zinc has a potential energy of $-6.38\ \mathrm{eV}$, an electron in the copper has a potential energy of $-4.78\ \mathrm{eV}$.

This number is bigger, and so electrons will flow to the lower energy state, which is towards the copper, because electrons go against the potential, from low potential to high potential.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Be careful here we're talking about electrons moving and usually when you're dealing with electronics that's what's going on, but not always in a cell. For example, you can and often do have positive charges moving potassium ions, sodium ions, calcium ions, so those charges would move from high potential to low potential, because they're positive, electrons go low to high, so you got to think about what's actually doing the moving.

Now let's go through and actually talk through how this thing keeps running. So, the gold-platinum, when we put them in contact the electrons run across and they stopped, because it's a buildup of charge and the repulsion stops the thing from going. Why does the zinc-copper keep going? Well zinc dissolves in sulfuric acid, and you spit off a Zinc 2^+ ion, which will eventually precipitate out a zinc sulfate. What does that leave behind in the metal? Electrons. Zinc 2^+ popped off, now I've got two electrons, and now my zinc has a slight negative charge to it. Well we just said that copper is the lower energy state, so the electrons would prefer to be in copper because it's a lower energy state.

They go over and run through the wire because it's less resistance to run through the wire than to run through the solution. We would have the same problem that we did with our platinum-gold we'd eventually build up charges on this side and the whole thing would stop, but the solution comes in again. Because it's sulfuric acid, there's these positive hydrogen ions sitting in there and these electrons are attracted to these positive hydrogen ions and jump off. Which then form a nice diatomic hydrogen, which bubbles out and now we're back where we started.

We have neutral-zinc, neutral-copper and the whole thing can just keep going, and so we can continue moving these electrons across and the reaction doesn't stop because we come back to a neutral state each time and we maintain a fixed potential difference.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

The reason this reaction keeps going is probably the number one thing that people miss in this unit.

Batteries have a fixed potential difference between the terminals, they don't have a fixed current, and they don't have a fixed power output. This is the number one thing people miss in this unit right here.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

The big takeaway is the battery has a fixed potential difference between its terminals, they do not maintain constant current, they do not maintain constant power output fixed potential difference fixed potential difference.

Homework

5. What are the properties of an ideal battery?

Capacitors and Dielectrics

Instructor's Note

By the end of this section you should know:

- A capacitor is a device used to store charge.
- The amount of charge Q a capacitor can store depends on two major factors—the voltage applied and the capacitor's physical characteristics, such as its size.
- The capacitance C is the amount of charge stored per volt, or

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

- The capacitance of a parallel plate capacitor is $C = \varepsilon_0 \frac{A}{d}$, when the plates are separated by air or free space. ε_0 is called the permittivity of free space.
- A parallel plate capacitor with a dielectric between its plates has a capacitance given by

$$C = \epsilon \frac{A}{d},$$

where ϵ is the value for the material.

- The maximum electric field strength above which an insulating material begins to break down and conduct is called dielectric strength.

A **capacitor** is a device used to store electric charge. Capacitors have applications ranging from filtering static out of radio reception to energy storage in heart defibrillators. Typically, commercial capacitors have two conducting parts close to one another, but not touching, such as those in Figure 1. (Most of the time an insulator is used between the two plates to provide separation—see the discussion on dielectrics below.) When battery terminals are connected to an initially uncharged capacitor, equal amounts of positive and negative charge, $+Q$ and $-Q$, are separated into its two plates. The capacitor remains neutral overall, but we refer to it as storing a charge Q in this circumstance.

Capacitor

A capacitor is a device used to store electric charge.

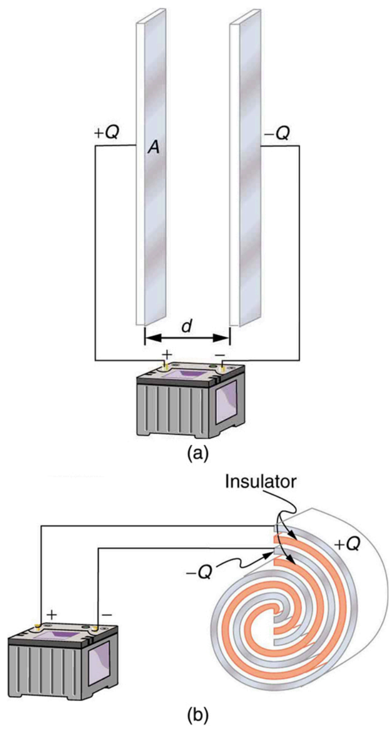


Figure 1. *Both capacitors shown here were initially uncharged before being connected to a battery. They now have separated charges of $+Q$ and $-Q$ on their two halves. (a) A parallel plate capacitor. (b) A rolled capacitor with an insulating material between its two conducting sheets.*

The amount of charge Q a *capacitor* can store depends on two major factors—the voltage applied and the capacitor's physical characteristics, such as its size.

The Amount of Charge Q A Capacitor Can Store

The amount of charge Q a *capacitor* can store depends on two major factors—the voltage applied and the capacitor's physical characteristics, such as its size.

A system composed of two identical, parallel conducting plates separated by a distance, as in Figure 2, is called a **parallel plate capacitor.** It is easy to see the relationship between the voltage and the stored charge for a parallel plate capacitor, as shown in Figure 2. Each electric field line starts on an individual positive charge and ends on a negative one, so that there will be more field lines if there is more charge. (Drawing a single field line per charge is a convenience, only. We can draw many field lines for each charge, but the total number is proportional to the number of charges.) The electric field strength is, thus, directly proportional to Q.QQ size 12{Q} {}">

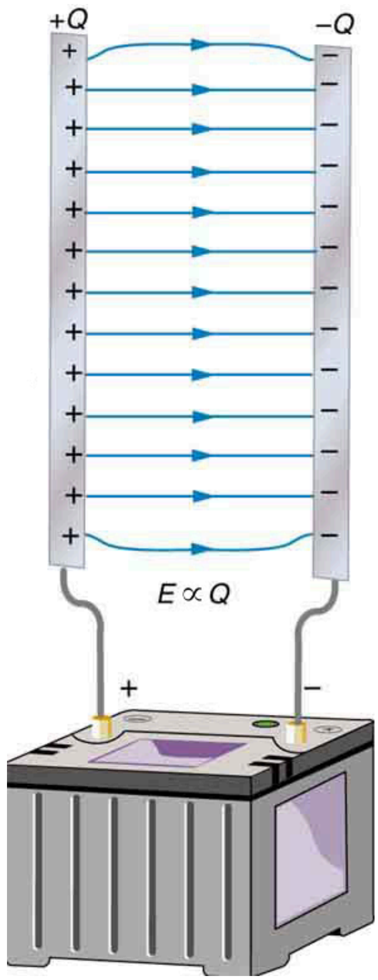


Figure 2*. Electric field lines in this parallel plate capacitor, as always, start on positive charges and end on negative charges. Since the electric field strength is proportional to the density of field lines, it is also proportional to the amount of charge on the capacitor.*

The field is proportional to the charge:

$$E \alpha Q,$$

where the symbol α means "proportional to." From the discussion in The Relationship between Electric and Potential and Electric Field, we know that the voltage across parallel plates is $\Delta V = E\Delta x$. Thus,

$$\Delta V \propto E.$$

It follows, then, that $V \propto Q$, and conversely,

$$Q \propto V.$$

This is true in general: The greater the voltage applied to any capacitor, the greater the charge stored in it.

Different capacitors will store different amounts of charge for the same applied voltage, depending on their physical characteristics. We define their **capacitance** C to be such that the charge Q stored in a capacitor is proportional to C. The charge stored in a capacitor is given by

$$Q = CV.$$

This equation expresses the two major factors affecting the amount of charge stored. Those factors are the physical characteristics of the capacitor. C, and the voltage, V. Rearranging the equation, we see that *capacitance* C is the amount of charge stored per volt, or

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

Capacitance

Capacitance C CC size 12{C} {}">is the amount of charge stored per volt, orC=QV.C=QV. size 12{C=Q/V} {}">

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

The unit of capacitance is the farad (F), named for Michael Faraday (1791–1867), an English scientist who contributed to the fields of electromagnetism and electrochemistry. Since capacitance is charge per unit voltage, we see that a farad is a coulomb per volt, or

$$1 \text{ F} = \frac{1 \text{ C}}{1 \text{ V}}$$

A 1-farad capacitor would be able to store 1 coulomb (a very large amount of charge) with the application of only 1 volt. One farad is, thus, a very large capacitance. Typical capacitors range from fractions of a picofarad $(1 \text{ pF} = 10^{-12} \text{ F})$ to millifarads $(1 \text{ mF} = 10^{-3} \text{ F})$.

Figure 3 shows some common capacitors. Capacitors are primarily made of ceramic, glass, or plastic, depending upon purpose and size. Insulating materials, called dielectrics, are commonly used in their construction, as discussed below.

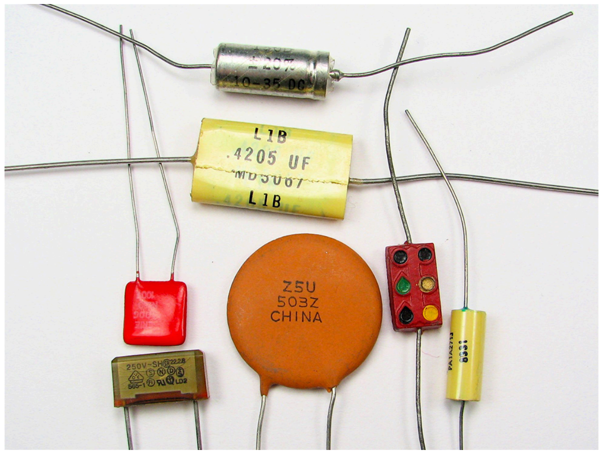


Figure 3. *Some typical capacitors. Size and value of capacitance are not necessarily related. (credit: Windell Oskay)*

Homework

6. Charge stored on a capacitor

Parallel Plate Capacitor

The parallel plate capacitor shown in Figure 4 has two identical conducting plates, each having a surface area A, separated by a distance d dd size 12{d} {}">(with no material between the plates). When a voltage V is applied to the capacitor, it stores a charge Q, as shown. We can see how its capacitance depends on A and d by considering the characteristics of the Coulomb force. We know that like charges repel, unlike charges attract, and the force between charges decreases with distance. So it seems quite reasonable that the bigger the plates are, the more charge they can store—because the charges can spread out more. Thus C should be greater for larger A. Similarly, the closer the plates are together, the greater the attraction of the opposite charges on them. So C should be greater for smaller d.

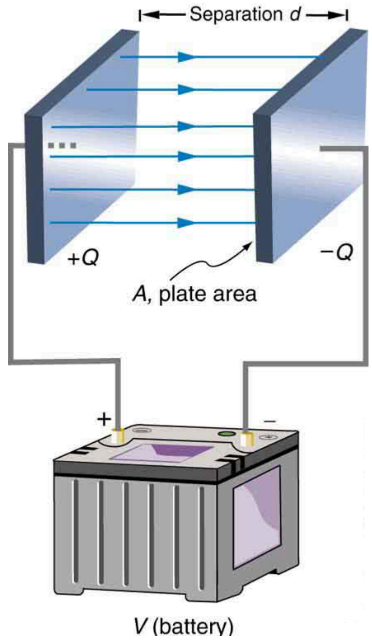


Figure 4. Parallel plate capacitor with plates separated by a distance d. Each plate has an area A.

It can be shown that for a parallel plate capacitor, with vacuum (or the very similar air) between the plates, there are only two factors (A and d) that affect its capacitance C. The capacitance of a parallel plate capacitor in equation form is given byC=ε0Ad.C=ε0Ad. size 12{C=e rSub { size 8{0} } A/d} {}">

$$C = \epsilon_0 \frac{A}{d}.$$

The situation with something other than air is discussed below.

Capacitance of a Parallel Plate Capacitor with Vacuum or Air Between the Plates

$$C = \epsilon_0 \frac{A}{d}$$

A is the area of one plate in square meters, and d is the distance between the plates in meters. The constant ϵ_0 is the $\epsilon_0 = 8.85 \times 10^{-12} \frac{\mathrm{C}^2}{\mathrm{N{\cdot}m}^2}$ we have seen before. Now, we can write it in a new way as Farads/meters: $\mathrm{F/m} = \frac{\mathrm{C}^2}{\mathrm{N{\cdot}m}^2}$. The small numerical value of ϵ_0 is related to the large size of the farad. A parallel plate capacitor must have a large area to have a capacitance approaching a farad. (Note that the above equation is valid when the parallel plates are separated by air or free space. When another material is placed between the plates, the equation is modified, as discussed below.)

Capacitance and Charge Stored in a Parallel Plate Capacitor

a) What is the capacitance of a parallel plate capacitor with metal plates, each of area $1.00\ \text{m}^2$ 1.00m21.00m2 size 12{m rSup { size 8{2} } } {}">, separated by 1.00 mm? (b) What charge is stored in this capacitor if a voltage of $3.00 \times 10^3\ \text{V}$ is applied to it?

Strategy

Finding the capacitance C is a straightforward application of the equation $C = \varepsilon_0 A/d$. Once C is found, the charge stored can be found using the equation $Q = CV$.

Solution for (a)

Entering the given values into the equation for the capacitance of a parallel plate capacitor yields

$$C = \varepsilon_0 \frac{A}{d} = (8.85 \times 10^{-12} \frac{F}{m}) \frac{1.00\ \text{m}^2}{1.00{\times}10^{-3}\ \text{m}}$$
$$= 8.85 \times 10^{-9}\ \text{F} = 8.85\ \text{nF}.$$

Discussion for (a)

This small value for the capacitance indicates how difficult it is to make a device with a large capacitance. Special techniques help, such as using very large area thin foils placed close together.

Solution for (b)

The charge stored in any capacitor is given by the equation $Q = CV$. Entering the known values into this equation gives

$$Q = CV = (8.85 \times 10^{-9}\ \text{F})(3.00 \times 10^3\ \text{V})$$
$$26.6\mu C.$$

Discussion for (b)

This charge is only slightly greater than those found in typical static electricity. Since air breaks down at about $3.00 \times 10^6\ \text{V/m}$, more charge cannot be stored on this capacitor by increasing the voltage.

Another interesting biological example dealing with electric potential is found in the cell's plasma membrane. The membrane sets a cell off from its surroundings and also allows ions to selectively pass in and out of the cell. There is a potential difference across the membrane of about $-70\ \text{mV}$. This is due to the mainly negatively charged ions in the cell and the predominance of positively charged sodium (Na^+) ions outside. Things change when a nerve cell is stimulated. Na^+ ions are allowed to pass through the membrane into the cell, producing a positive membrane potential—the nerve signal. The cell membrane is about 7 to 10 nm thick. An approximate value of the electric field across it is given by

$$|E| = \frac{V}{\Delta x} = \frac{-70{\times}10^{-3}\ \text{V}}{8{\times}10^{-9}\ \text{m}} = 9 \times 10^6\ \text{V/m}$$

This electric field is enough to cause a breakdown in air.

Homework

7. Bookshelf capacitor

Dielectric

The previous example highlights the difficulty of storing a large amount of charge in capacitors. If d is made smaller to produce a larger capacitance, then the maximum voltage must be reduced proportionally to avoid breakdown (since $E = V/d$). An important solution to this difficulty is to put an insulating material, called a dielectric, between the plates of a capacitor and allow d to be as small as possible. Not only does the smaller d make the capacitance greater, but many insulators can withstand greater electric fields than air before breaking down.

There is another benefit to using a dielectric in a capacitor. As discussed in class during Unit III, the electric field in a material is smaller than in vacuum due to the polarization of the material. In those analyses, we swapped $\epsilon_0 \rightarrow \epsilon$. We will do the same here. Thus, for a parallel plate capacitor filled with material, $C = \epsilon \frac{A}{d}$. Since essentially all materials have $\epsilon > \epsilon_0$ \epsilon_0 " title="Rendered by QuickLaTeX.com" height="12" width="42" style="vertical-align: -3px;"> (everything is more polarizeable than vacuum), the capacitance of a capacitor filled with material will be larger than one filled with air or vacuum.

Also as discussed in class during Unit III, tables of materials typically do not list ϵ but instead ϵ/ϵ_0 a factor called the *dielectric constant* κ. Values of the dielectric constant κ for various materials are given in the Table below. If Teflon is placed between the plates of the capacitor, as in the example after the table, then the capacitance is greater by the factor κ, which for Teflon is 2.1.

Take-Home Experiment: Building A Capacitor

How large a capacitor can you make using a chewing gum wrapper? The plates will be the aluminum foil, and the separation (dielectric) in between will be the paper.

Dielectric Constants and Dielectric Strengths for Various Materials at 20ºC

Material	Dielectric constant κ	Dielectric strength (V/m)
Vacuum	1.00000	—
Air	1.00059	3×10^6
Bakelite	4.9	24×10^6
Fused quartz	3.78	8×10^6
Neoprene rubber	6.7	12×10^6
Nylon	3.4	14×10^6
Paper	3.7	16×10^6
Polystyrene	2.56	24×10^6
Pyrex glass	5.6	14×10^6
Silicon oil	2.5	15×10^6
Strontium titanate	233	8×10^6
Teflon	2.1	60×10^6
Water	80	—

Note also that the dielectric constant for air is very close to 1, so that air-filled capacitors act much like those with vacuum between their plates *except* that the air can become conductive if the electric field strength becomes too great. (Recall that $E = \Delta V/\Delta x$.) Also shown in Table are maximum electric field strengths in V/m, called dielectric strengths, for several materials. These are the fields above which the material begins to break down and conduct. The dielectric strength imposes a limit on the voltage that can be applied for a given plate separation. For instance, in Example, the separation is 1.00 mm, and so the voltage limit for air is

$$\begin{aligned} V &= E \cdot \Delta x \\ &= (3 \times 10^6 \mathrm{V/m})(1.00 \times 10^{-3}\ \mathrm{m}) \\ &= 3000\ \mathrm{V}. \end{aligned}$$

However, the limit for a 1.00 mm separation filled with Teflon is 60,000 V, since the dielectric strength of Teflon is $60 \times 10^6 \mathrm{V/m}$. So the same capacitor filled with Teflon has a greater capacitance and can be subjected to a much greater voltage. Using the capacitance we calculated in the above example for the air-filled parallel plate capacitor, we find that the Teflon-filled capacitor can store a maximum charge of

$$\begin{aligned} Q &= CV \\ &= \kappa C_{\text{air}} V \\ &= (2.1)(8.85\ \mathrm{nF})(6.0 \times 10^4\ \mathrm{V}) \\ &= 1.1\ \mathrm{mC} \end{aligned}$$

This is 42 times the charge of the same air-filled capacitor.

Dielectric Strength

The maximum electric field strength above which an insulating material begins to break down and conduct is called its dielectric strength.

PhET Explorations: Capacitor Lab

Explore how a capacitor works! Change the size of the plates and add a dielectric to see the effect on capacitance. Change the voltage and see charges built up on the plates. Observe the electric field in the capacitor. Measure the voltage and the electric field.

An interactive or media element has been excluded from this version of the text. You can view it online here:
http://openbooks.library.umass.edu/toggerson-132/?p=451

Homework

8. Capacitor with neoprene.
9. What does capacitance depend upon?

Ohm's Law: Resistance and Simple Circuits

Derived from Ohm's Law: Resistance and Simple Circuits by OpenStax

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 VV size 12{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*:

$$I \propto V$$

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

$$I \propto \frac{1}{R}$$

Thus, for example, current is cut in half if resistance doubles. Combining the relationships of current to voltage and current to resistance gives

$$I = \frac{V}{R}.$$

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 the 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 Ω (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:

$$1\ \Omega = 1\frac{V}{A}$$

Figure 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.RR size 12{R} {}">

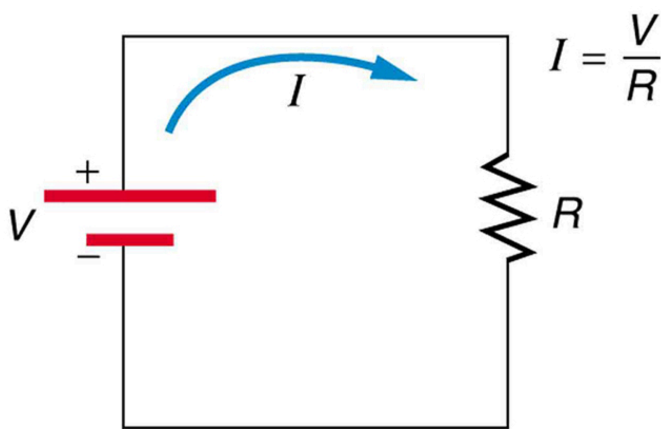


***Figure 1**. A simple electric circuit in which a closed path for current to flow is supplied by conductors (usually metal wires) connecting a load to the terminals of a battery, represented by the red parallel lines. The zigzag symbol represents the single resistor and includes any resistance in the connections to the voltage source.*

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 $I = V/R$ and substituting known values gives

$$R = \frac{V}{I} = \frac{12.0\ \text{V}}{2.50\ \text{A}} = 4.80\ \Omega$$

Discussion

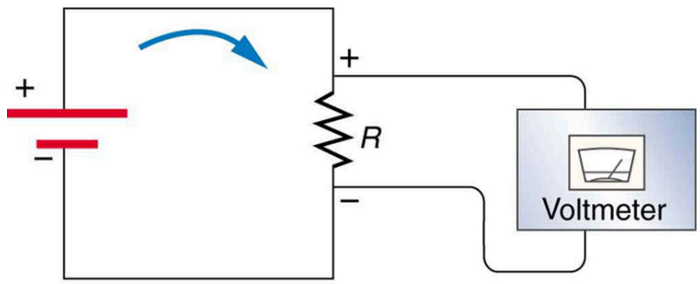
This is a relatively small resistance, but it is larger than the cold resistance of the headlight. As we shall see in *Resistance and Resistivity* below, resistance usually increases with temperature, and so the bulb has a lower resistance when it is first switched on and will draw considerably more 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* below.

Additional insight is gained by solving $I = V/R$ for V, yielding

$$V = IR$$

This expression for V can be interpreted as the *voltage drop across a resistor produced by the flow of current* I. The phrase IR *drop* is often used for this voltage. For instance, the headlight in the automobile headlight example above 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 $U = q\Delta V$, and the same q flows through each. Thus the energy supplied by the voltage source and the energy converted by the resistor are equal. (See Figure 2.)



$$V = IR = 18 \text{ V}$$

***Figure 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.

PhET Explorations: Ohm's Law

See how the equation form of Ohm's law relates to a simple circuit. Adjust the voltage and resistance, and see the current change according to Ohm's law. The sizes of the symbols in the equation change to match the circuit diagram.

An interactive or media element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=451

Homework

10. Voltage in defibrillator.
11. Resistance of defibrillator.
12. Simple circuit.

Section 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 a simple circuit to be $I = \frac{V}{R}$.
- Resistance has units of ohms (Ω), related to volts and amperes by $1\ \Omega = 1\ \text{V}/\text{A}$.
- There is a voltage or IR drop across a resistor, caused by the current flowing through it, given by $V = IR$.

Resistance and Resistivity

Derived from Resistance and Resistivity by OpenStax

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

For this section I am NOT expecting you to remember the formula exactly. What I want you to know is that resistance increases with length (more atoms to run into), decreases with area, and is dependent upon the resistivity of the material.

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.

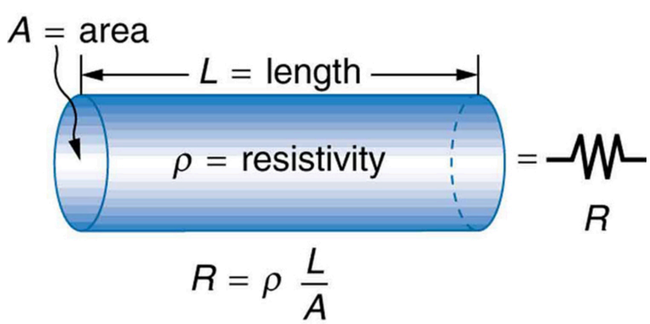


Figure 1. *A uniform cylinder of length L and cross-sectional area A. Its resistance to the flow of current is similar to the resistance posed by a pipe to fluid flow. The longer the cylinder, the greater its resistance. The larger its cross-sectional area A, the smaller its resistance.*

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 ρ. Resistivity ρ 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 ρ, is

$$R = \frac{\rho L}{A}$$

The table below 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.

Material	**Resistivity** ρ $(\Omega \cdot \mathrm{m})$
Conductors	
Silver	1.59×10^{-8}
Copper	1.72×10^{-8}
Gold	2.44×10^{-8}
Aluminum	2.65×10^{-8}
Tungsten	5.6×10^{-8}
Iron	9.71×10^{-8}
Platinum	10.6×10^{-8}
Steel	20×10^{-8}
Lead	22×10^{-8}
Manganin (Cu, Mn, Ni alloy)	44×10^{-8}
Constantan (Cu, Ni alloy)	49×10^{-8}
Mercury	96×10^{-8}
Nichrome (Ni, Fe, Cr alloy)	100×10^{-8}
Semiconductors	
Carbon (pure)	3.5×10^{-5}
Carbon	$(3.5 - 60) \times 10^{-5}$
Germanium (pure)	600×10^{-3}
Germanium	$(1 - 600) \times 10^{-3}$
Silicon (pure)	2300
Silicon	$(0.1 - 2300)$
Insulators	
Amber	5×10^{14}
Glass	$10^{9} - 10^{14}$

Lucite	$> 10^{13}$ 10^{13}" title="Rendered by QuickLaTeX.com" height="16" width="45" style="vertical-align: -1px;">
Mica	$10^{11} - 10^{15}$
Quartz (fused)	75×10^{16}
Rubber (hard)	$10^{13} - 10^{16}$
Sulfur	10^{15}
Teflon	$> 10^{13}$ 10^{13}" title="Rendered by QuickLaTeX.com" height="16" width="45" style="vertical-align: -1px;">
Wood	$10^{8} - 10^{11}$

Resistivity ρ of Various materials at 20°C

Calculating Resistor Diameter: A Headlight Filament

A car headlight filament is made of tungsten and has a cold resistance of $0.350\ \Omega$. 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 the table, yields

$$\begin{aligned}A&=\frac{(5.6\times10^{-8}\ \Omega\cdot\mathrm{m})(4.00\times10^{-2}\ \mathrm{m})}{0.350\ \Omega}\\&=6.40\times10^{-9}\ \mathrm{m}^2.\end{aligned}$$

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

$$\begin{aligned}D&=2(\frac{A}{P})^{\frac{1}{2}}=2(\frac{6.40\times10^{9}\ \mathrm{m}^2}{3.14})^{\frac{1}{2}}\\&=9.0\times10^{-5}\ \mathrm{m}\end{aligned}$$

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. Semiconductors, on the other hand, have resistivities which decrease 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 ρ 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 ρ. 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 ρ. (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 ρ.) 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 3. *These familiar thermometers are based on the automated measurement of a thermistor's temperature-dependent resistance. (credit: Biol, Wikimedia Commons)*

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.

An interactive or media element has been excluded from this version of the text. You can view it online here:
http://openbooks.library.umass.edu/toggerson-132/?p=451

Section Summary

- The resistance R increases with length and decreases with cross sectional area.
- The resistance is also dependent on the resistivity ρ of the material.
- Values of ρ fall into three groups—*conductors, semiconductors, and insulators*.
- For a metal, resistivity, and therefore resistance, increases with temperature as the nuclei are jiggling

around more resulting in more collisions as the electrons try to travel.

Homework

13. Resistance of extension cords.

28. Circuits

Electric Power and Energy

Derived from Electric Power and Energy by OpenStax. You may find it useful to review the section on Power from Some Energy-Related Ideas that Might be New or are Particularly Important.

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 light bulbs in terms of their power ratings in watts. Let us compare a 25-W bulb with a 60-W bulb. (See Figure 1(a).) 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 (think about Ohm's Law!). 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?

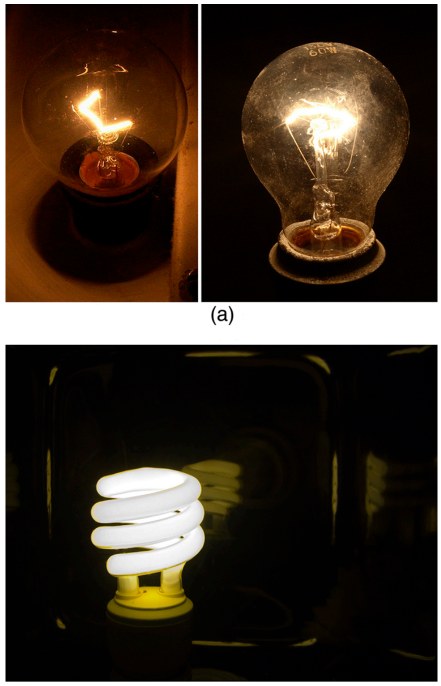


Figure 1. *(a) Which of these lightbulbs, the 25-W bulb (upper left) or the 60-W bulb (upper right), has the higher resistance? Which draws more current? Which uses the most energy? Can you tell from the color that the 25-W filament is cooler? Is the brighter bulb a different color and if so why? (credits: 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 $U = qV$, 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

$$P = \frac{U}{t} = \frac{qV}{t}.$$

Recognizing that current is $I = q/t$ (note that $\Delta t = t$ here), the expression for power becomes

$$P = IV.$$

Electric power (P) is simply the product of current times voltage. Power has familiar units of watts. Thus, $1\ \mathrm{A} \cdot\ \mathrm{V} = 1\ \mathrm{W}$. 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 $P = IV = (20\ \mathrm{A})(12\ \mathrm{V}) = 240\ \mathrm{W}$. In some applications, electric power may be expressed as volt-amperes or even kilovolt-amperes ($1\ \mathrm{kA} \cdot\ \mathrm{V} = 1\ \mathrm{kW}$).

To see the relationship of power to resistance, we combine Ohm's law with $P = IV$. Substituting $I = V/R$ gives $P = (V/R)V = V^2/R$. Similarly, substituting $V = IR$ gives $P = I(IR) = I^2R$. Three expressions for electric power are listed together here for convenience:

$$P = IV$$

$$P = \frac{V^2}{R}$$
$$P = I^2R.$$

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.

Calculating Power Dissipation and Current: Hot and Cold Power

(a) Consider the examples given in the last chapter on Ohm's Law. Then find the power dissipated by the car headlight in this example.

Strategy for (a)

For the hot headlight, we know voltage and current, so we can use $P = IV$ to find the power.

Solution for (a)

Entering the known values of current and voltage for the hot headlight, we obtain

$$P = IV = (2.50 \text{ A})(12.0 \text{ V}) = 30.0 \text{ W}.$$

Discussion for (a)

The 30 W dissipated by the hot headlight is typical.

Homework

14. Truck starter motor.

15. Current draws from different appliances.

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

$$E = Pt$$

is the energy used by a device using power P PP size 12{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 kilowatt-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 $1\ \text{kW} \cdot\ \text{h} = 3.6 \times 10^{6}\ \text{J}$.

The electrical 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 1(b).) 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.

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.

Calculating the Cost Effectiveness of Compact Fluorescent Lights (CFL)

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? (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?

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 for (a)

The energy used in kilowatt-hours is found by entering the power and time into the expression for energy:

$$E = Pt = (60 \text{ W})(1000 \text{ h}) = 60,000 \text{ W} \cdot \text{ h}.$$

In kilowatt-hours, this is

$$E = 60.0 \text{ kW} \cdot \text{ h}.$$

Now the electricity cost is

$$\text{cost} = (60.0 \text{ kW} \cdot \text{ h})(0.12/ \text{ kW} \cdot \text{h}) = 7.20.$$

The total cost will be $7.20 for 1000 hours (about one-half year at 5 hours per day)

Solution for (b)

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

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?

Section Summary

- Electric power is the rate (in watts) that energy is supplied by a source or dissipated by a device.
- Three expressions for electrical power are

$$P = IV,$$
$$P = \frac{V^2}{R},$$

and

$$P = I^2R.$$

- The energy used by a device with a power P over time t is $E = Pt$.

Homework

16. Battery life-time.

Kirchhoff's Principles

This section is adapted from The University of Maryland BERG group.

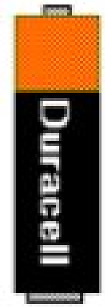
The basic ideas that we have developed about how electric charges move in matter serve as a basis for analyzing a wide variety of electric circuits and devices and for modeling the electrical behavior of biological systems. But these circuits, devices, and models can quickly become quite complex. It becomes useful to establish a set of foothold ideas — principles that we can hold on to and refer back to in order to organize our thinking in a complex situations — to provide a "stake in the ground" that we can trust and use to support our safety net of coherent and linked ideas.

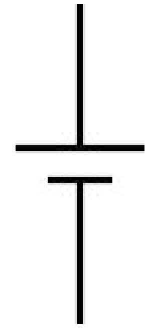
The foothold principles for understanding electric currents were developed by the 19th century German physicist, Gustav Kirchhoff (yes, two "h"s) are called Kirchhoffs laws (or principles). (He also formulated laws of spectroscopy and thermochemistry.)

The (idealized) context for Kirchhoff's principles

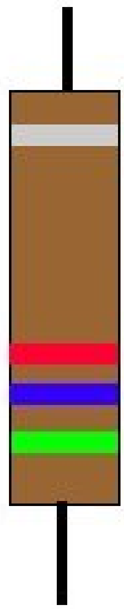
Kirchhoff's principles are restrictions of more general electromagnetic laws (Maxwell's equations, conservation of charge) to standard situations in electrical circuits. We'll talk about them and use them in the context of analyzing connected networks of electrical devices — batteries, resistors, capacitors, and wires. Here's how we will represent and idealize them:

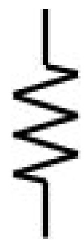
Batteries — devices that maintain a constant electrical pressure difference (voltage) across their terminals (like a water pump that raises water to a certain height). We use the symbol drawn at the right with the longer line corresponding to the end of the battery with the higher potential.



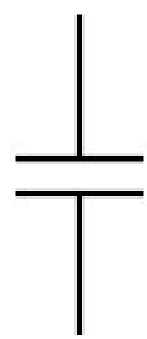


Resistors — devices that have significant drag and oppose current. Pressure will drop across them when current is flowing through them. The resistor shown in the figure is of the kind used in hand-built electrical circuits. The stripes color-code the size of the resistor and its precision. We indicate them with the zig-zag symbol shown on the right.





Capacitors — devices that can maintain a charge separation in response to a pressure differential (voltage) applied across its plates. The symbol for a capacitor is the pair of parallel lines shown on the right. Be sure to distinguish between the symbol for a capacitor (parallel lines of equal length) and a battery (parallel lines of different length)!



Wires — conductors that connect the other devices. We treat them as perfect conductors — having zero resistance. This approximation only works when there are other resistors in the circuit having significant resistance, compared to which the resistance of the wires are negligible. The symbol for a wire is just a line and it could be straight or curved — makes no difference.

Kirchhoff's 1st (Flow) Principle

The first principle is basically a combination of two ideas:

- conservation of charge (the total amount of positive charge minus the total amount of negative charge is a constant)
- in electrical circuits, due to the strong repulsive forces between like charges, electrical elements remain neutral — there is no build-up of charge anywhere.

The principle is often called "the flow rule" and is stated as follows:

The total amount of current flowing into any volume in an electrical network equals the amount flowing out.

From our analysis of how a capacitor and a resistor both work, we know that this idea doesn't hold when things are just getting started.

For example, when we charge a capacitor, charge is flowing into one side of the capacitor and out of the other: charge (of opposite sign) is building up on each plate of the capacitor in violation of the flow rule. But if we put a box around the capacitor and don't look inside, the rule works. It also works when the system is in the steady state and things have stabilized.

A similar thing holds for a resistor. When a current just starts to build up through the resistor, a build-up of equal and opposite charges at the two ends of the resistor are what is responsible for establishing the electric field in the resistor (creating a potential drop across the resistor) that keeps the charges moving through the drag of the resistor at a constant velocity, consistent with Newton's laws of motion.

In both these cases, when we hook up these devices to a circuit, the first principle violations that take place in the interior of the device happen fast — in nanoseconds or less. And if we consider the whole device instead of just a part of it, the principle still works even on that time scale.

Kirchhoff's 2nd (Resistance) principle

The second principle tells what happens when there is a current in a resistor — there is a potential drop in the direction of the current which is proportional to the current times a property of the resistor. This is just Ohm's law and it hold for any device in which the drag resisting the flow is proportional to the velocity. (See Resistive electric flow: Ohm's law.) We can even stretch its validity by letting the resistance be a function of the current. Mostly, we wont need to do this.

$$\Delta V = IR$$

For capacitors we have the analogous result:

$$Q = C\Delta V$$

Kirchhoff’s 3rd (Loop) principle

Where Kirchhoff’s first principle controls the current in an electrical network, the second deals with the voltage drops in the network. We can understand it by using the water analogy. The electric potential is analogous in the water model to the height that the water has been raised. One of the things we know about heights is that if you make a loop and come back to the same point, you will be at the same height from which you started. Whatever drops (descents) you made had to be cancelled by and equal sum of rises (climbs) in order to get back to your starting point.

The same thing is true of electric potential (voltage). As we travel through a circuit, we may have rises, say if we go through a battery from its low end to its high end, and we may have drops, say if we go through a resistance in the direction of the current flow. Kirchhoff’s third principle states:

Following around any loop in an electrical network the potential has to come back to the same value (sum of drops = sum of rises).

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor’s Note

This can be a bit tricky to apply! Just as if you go up a hill you are rising, but if you walk down that same hill you are descending, whether you have a rise or a drop in electric potential as you go through a device depends on which way you are following your loop If you go through a battery from the positive end to the negative end, it gives you a drop! If you go through a resistor in the direction opposite the direction a current is flowing you get a rise!

Useful heuristics

Applying Kirchhoff's principles to a complex circuit is sometimes complicated. There are two variables to be solved for — the voltage (electrical pressure) and the current. These are independent variables. They affect each other, but your intuitions as to what is happening sometimes refers to one, sometimes to the other — but it's easy to get confused!

A useful way to think of the voltage throughout the circuit is as analogous to pressure (in the air flow model) or height (in the water flow model). Moving throughout the circuit there are different values of this variable — the voltage (electric pressure) — but it doesn't move or change. it is the difference between voltages (say at opposite ends of a resistor) that drives current through a resistor. One of the best ways to start analyzing an electrical network is by figuring out what you know about the voltage. And here's a corollary to Ohm's law that helps a great deal:

A conductor in a circuit that can be treated as having 0 resistance, e.g., a wire, is an equipotential (has the same value of the potential everywhere along it) even if there is current flowing through it, since for that wire, the voltage drop across the wire is given by Ohm's law: $\Delta V = IR = 0$, even if I is NOT zero.

The best advice in handling the current in an electrical circuit problem is to choose some directions for the directions you think the currents are flowing in and take those as positive. Then just apply Kirchhoff's principles to generate relationships (equations) among the various variables. If you have chosen wrong the signs will come out negative. No problem! It just tells you that your initial assumption was wrong and that the current is flowing in the opposite direction from the one you expected.

Homework

17. Adding current.

29. Review of Solving Systems of Equations

Instructor's Notes

One mathematical skill you will need in this unit is the ability to solve systems of linear equations. In class, and on exams, we will stick to two-equations with two-unkowns (what I call 2×2). However, in some of the additional practice problems, you will need to go to three-equations with three-unknowns (3×3). To help everyone refresh their memories on how to do this, I am assigning a few problems. If you are familiar, just go ahead and try them. If you need some review, I include Section 7.1 – Systems of Linear Equations: Two Variables from the OpenStax College Algebra 2e textbook below.

Homework Problems

18. Solve a system with two equations and two unknowns.
19. Solve a system with three equations and three unknowns.

Systems of Linear Equations: Two Variables

Learning Objectives

In this section, you will:

- Solve systems of equations by graphing.
- Solve systems of equations by substitution.
- Solve systems of equations by addition.
- Identify inconsistent systems of equations containing two variables.
- Express the solution of a system of dependent equations containing two variables

Figure 1. *(credit: Thomas Sørenes)*

A skateboard manufacturer introduces a new line of boards. The manufacturer tracks its costs, which is the amount it spends to produce the boards, and its revenue, which is the amount it earns through sales of its boards. How can the company determine if it is making a profit with its new line? How many skateboards must be produced and sold before a profit is possible? In this section, we will consider linear equations with two variables to answer these and similar questions.

Introduction to Systems of Equations

In order to investigate situations such as that of the skateboard manufacturer, we need to recognize that we are dealing with more than one variable and likely more than one equation. A system of linear equations consists of two or more linear equations made up of two or more variables such that all equations in the system are considered simultaneously. To find the unique solution to a system of linear equations, we must find a numerical value for each variable in the system that will satisfy all equations in the system at the same time. Some linear systems may not have a solution and others may have an infinite number of solutions. In order for a linear system to have a unique solution, there must be at least as many equations as there are variables. Even so, this does not guarantee a unique solution.

In this section, we will look at systems of linear equations in two variables, which consist of two equations that contain two different variables. For example, consider the following system of linear equations in two variables.

$$\begin{aligned} 2x+y &= 15 \\ 3x-y &= 5 \end{aligned}$$

The *solution* to a system of linear equations in two variables is any ordered pair that satisfies each equation independently. In this example, the ordered pair (4, 7) is the solution to the system of linear equations. We can verify the solution by substituting the values into each equation to see if the ordered pair satisfies both equations. Shortly we will investigate methods of finding such a solution if it exists.

$$\begin{aligned} 2\left(4\right)+\left(7\right) &= 15 \text{True} \\ 3\left(4\right)-\left(7\right) &= 5 \quad \text{True} \end{aligned}$$

In addition to considering the number of equations and variables, we can categorize systems of linear equations by the number of solutions. A consistent system of equations has at least one solution. A consistent system is considered to be an **independent system** if it has a single solution, such as the example we just explored. The two lines have different slopes and intersect at one point in the plane. A consistent system is considered to be a **dependent system** if the equations have the same slope and the same *y*-intercepts. In other words, the lines coincide so the equations represent the same line. Every point on the line represents a coordinate pair that satisfies the system. Thus, there are an infinite number of solutions.

Another type of system of linear equations is an **inconsistent system**, which is one in which the equations represent two parallel lines. The lines have the same slope and different *y*-intercepts. There are no points common to both lines; hence, there is no solution to the system.

Types of Linear Systems

There are three types of systems of linear equations in two variables, and three types of solutions.

- An independent system has exactly one solution pair (x,y). The point where the two lines intersect is the only solution.
- An inconsistent system has no solution. Notice that the two lines are parallel and will never intersect.
- A dependent system has infinitely many solutions. The lines are coincident. They are the same line, so every coordinate pair on the line is a solution to both equations.

(Figure) compares graphical representations of each type of system.

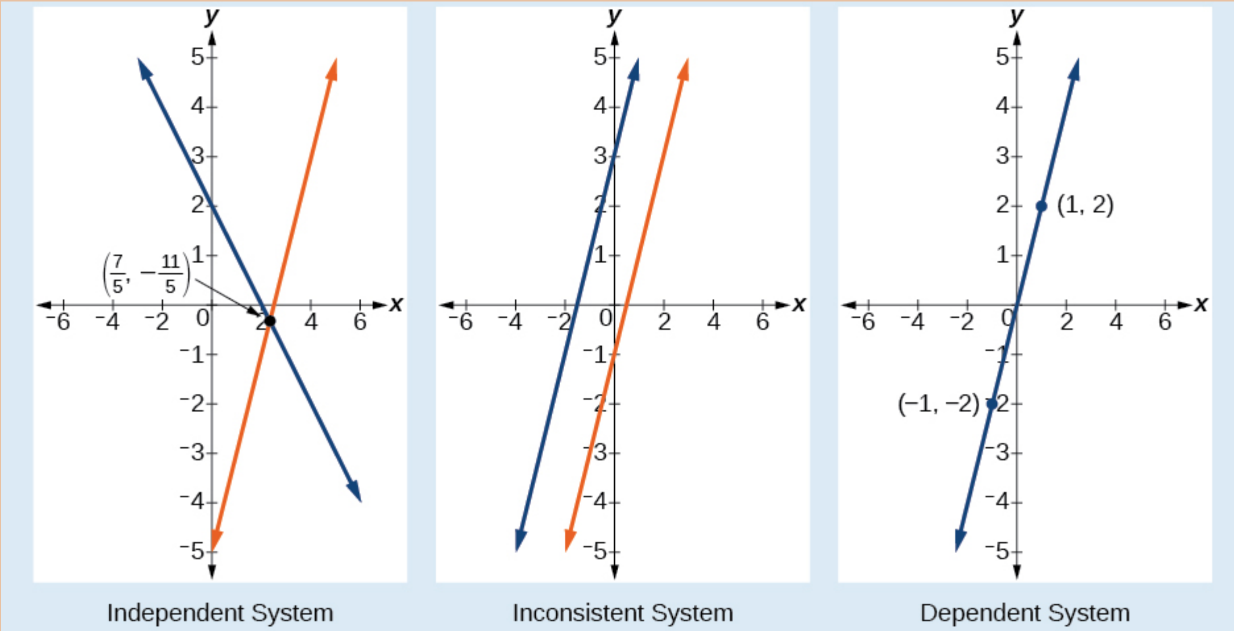


Figure 2.

How To

Given a system of linear equations and an ordered pair, determine whether the ordered pair is a solution.

1. Substitute the ordered pair into each equation in the system.
2. Determine whether true statements result from the substitution in both equations; if so, the ordered pair is a solution.

Determining Whether an Ordered Pair Is a Solution to a System of Equations

Determine whether the ordered pair $(5, 1)$ is a solution to the given system of equations.

$$x + 3y = 8$$
$$2x - 9 = y$$

Substitute the ordered pair $(5,1)$ into both equations.

$$\begin{array}{ll} (5)+3(1)=8 & \\ 8=8 & \text{True} \\ 2(5)-9=(1) & \\ 1=1 & \text{True} \end{array}$$

The ordered pair $(5,1)$ satisfies both equations, so it is the solution to the system.

Analysis

We can see the solution clearly by plotting the graph of each equation. Since the solution is an ordered pair that satisfies both equations, it is a point on both of the lines and thus the point of intersection of the two lines. See (Figure).

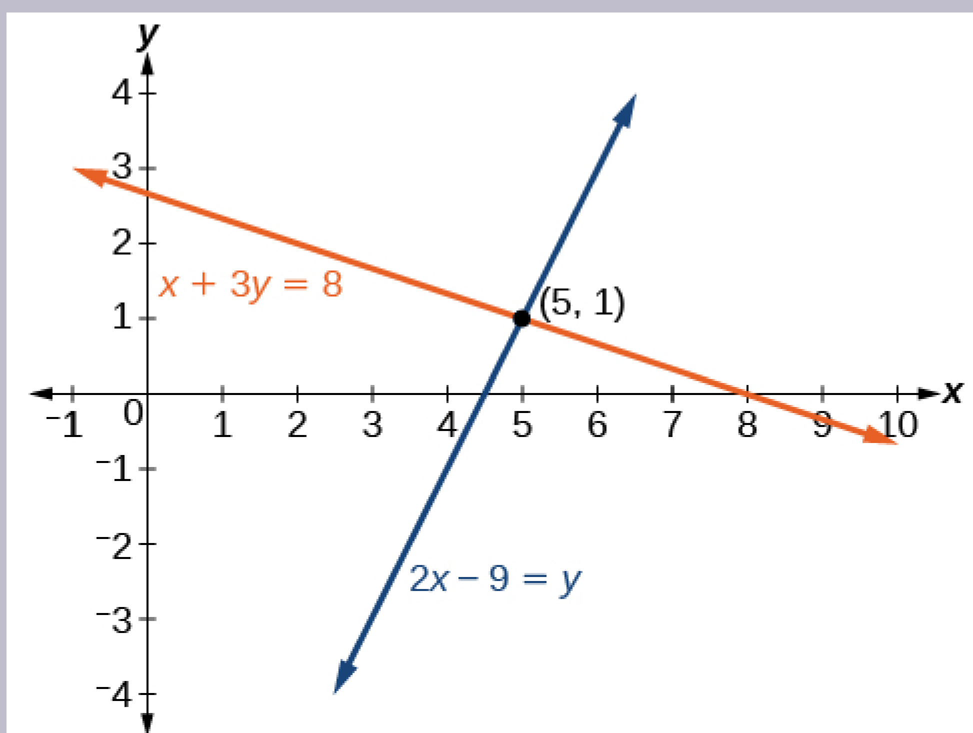


Figure 3.

Try It

Determine whether the ordered pair $(8, 5)$ is a solution to the following system.

$$\begin{aligned} 5x - 4y &= 20 \\ 2x + 1 &= 3y \end{aligned}$$

Not a solution.

Solving Systems of Equations by Graphing

There are multiple methods of solving systems of linear equations. For a system of linear equations in two variables, we can determine both the type of system and the solution by graphing the system of equations on the same set of axes.

Solving a System of Equations in Two Variables by Graphing

Solve the following system of equations by graphing. Identify the type of system.

$$\begin{aligned} 2x + y &= -8 \\ x - y &= -1 \end{aligned}$$

Solve the first equation for y.

$$\begin{aligned} 2x + y &= -8 \\ y &= -2x - 8 \end{aligned}$$

Solve the second equation for y.

$$\begin{aligned} x - y &= -1 \\ y &= x + 1 \end{aligned}$$

Graph both equations on the same set of axes as in (Figure).

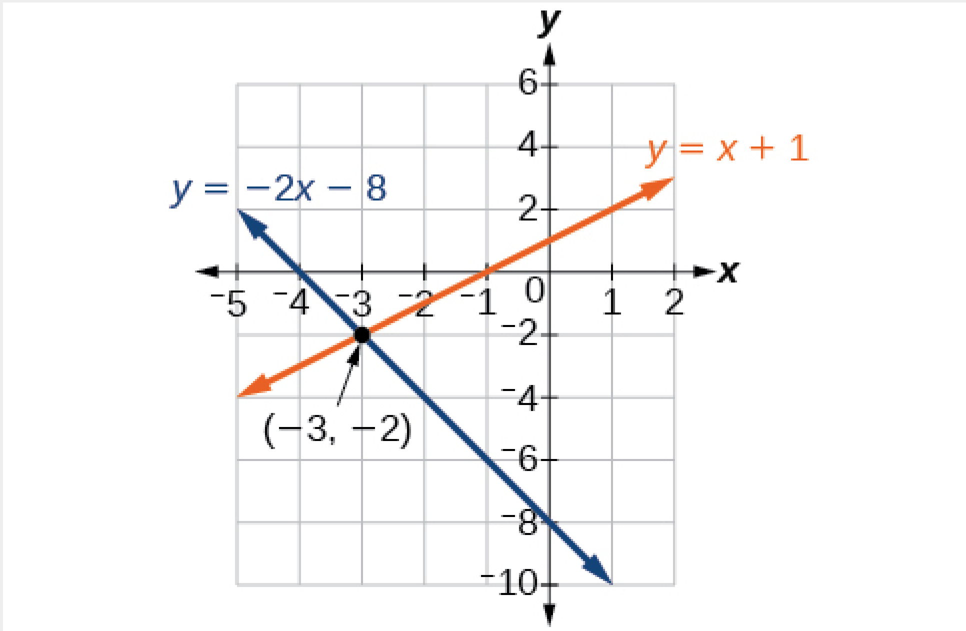


Figure 4.

The lines appear to intersect at the point $(-3,-2)$. We can check to make sure that this is the solution to the system by substituting the ordered pair into both equations.

$$\begin{array}{ll} 2\left(-3\right)+\left(-2\right)=-8 & \\ \quad -8=-8 & \text{True} \\ \left(-3\right)-\left(-2\right)=-1 & \\ \quad -1=-1 & \text{True} \end{array}$$

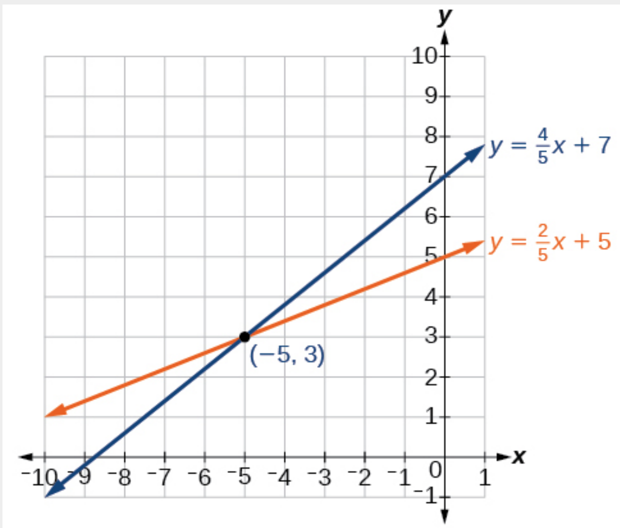
The solution to the system is the ordered pair $(-3,-2)$, so the system is independent.

Try It

Solve the following system of equations by graphing.

$$\begin{aligned} 2x - 5y &= -25 \\ -4x + 5y &= 35 \end{aligned}$$

The solution to the system is the ordered pair $(-5, 3)$.



Can graphing be used if the system is inconsistent or dependent?

Yes, in both cases we can still graph the system to determine the type of system and solution. If the two lines are parallel, the system has no solution and is inconsistent. If the two lines are identical, the system has infinite solutions and is a dependent system.

Solving Systems of Equations by Substitution

Solving a linear system in two variables by graphing works well when the solution consists of integer values, but if our solution contains decimals or fractions, it is not the most precise method. We will consider two more methods of solving a system of linear equations that are more precise than graphing. One such method is solving a system of equations by the substitution method, in which we solve one of the equations for one variable and then substitute the result into the second equation to solve for the second variable. Recall that we can solve for only one variable at a time, which is the reason the substitution method is both valuable and practical.

How To

Given a system of two equations in two variables, solve using the substitution method.

1. Solve one of the two equations for one of the variables in terms of the other.
2. Substitute the expression for this variable into the second equation, then solve for the remaining variable.
3. Substitute that solution into either of the original equations to find the value of the first variable. If possible, write the solution as an ordered pair.
4. Check the solution in both equations.

Solving a System of Equations in Two Variables by Substitution

Solve the following system of equations by substitution.

$$-x+y=-5$$
$$2x-5y=1$$

First, we will solve the first equation for y.

$$-x+y=-5$$
$$y=x-5$$

Now we can substitute the expression $x-5$ for y in the second equation.

$$2x-5y=1$$
$$2x-5\left(x-5\right)=1$$
$$2x-5x+25=1$$
$$-3x=-24$$
$$x=8$$

Now, we substitute $x=8$ into the first equation and solve for y.

$$-(8) + y = -5$$
$$y = 3$$

Our solution is $(8, 3)$.

Check the solution by substituting $(8, 3)$ into both equations.

$$-x + y = -5$$
$$-(8) + (3) = -5 \qquad \text{True}$$
$$2x - 5y = 1$$
$$2(8) - 5(3) = 1 \qquad \text{True}$$

Try It

Solve the following system of equations by substitution.

$$x = y + 3$$
$$4 = 3x - 2y$$

[reveal-answer q=”fs-id1165135516681"]Show Solution[/reveal-answer]
[hidden-answer a=”fs-id1165135516681"]
$(-2, -5)$

[/hidden-answer]

Can the substitution method be used to solve any linear system in two variables?

Yes, but the method works best if one of the equations contains a coefficient of 1 or –1 so that we do not have to deal with fractions.

Solving Systems of Equations in Two Variables by the Addition Method

A third method of solving systems of linear equations is the addition method. In this method, we add two terms with the same variable, but opposite coefficients, so that the sum is zero. Of course, not all systems are set up with the two terms of one variable having opposite coefficients. Often we must adjust one or both of the equations by multiplication so that one variable will be eliminated by addition.

How To

Given a system of equations, solve using the addition method.

1. Write both equations with *x*– and *y*-variables on the left side of the equal sign and constants on the right.
2. Write one equation above the other, lining up corresponding variables. If one of the variables in the top equation has the opposite coefficient of the same variable in the bottom equation, add the equations together, eliminating one variable. If not, use multiplication by a nonzero number so that one of the variables in the top equation has the opposite coefficient of the same variable in the bottom equation, then add the equations to eliminate the variable.
3. Solve the resulting equation for the remaining variable.
4. Substitute that value into one of the original equations and solve for the second variable.
5. Check the solution by substituting the values into the other equation.

Solving a System by the Addition Method

Solve the given system of equations by addition.

$$x + 2y = -1$$
$$-x + y = 3$$

Both equations are already set equal to a constant. Notice that the coefficient of x in the second equation, –1, is the opposite of the coefficient of x in the first equation, 1. We can add the two equations to eliminate x without needing to multiply by a constant.

$$\begin{array}{r} x + 2y = -1 \\ -x + y = 3 \\ \hline 3y = 2 \end{array}$$

Now that we have eliminated $x,$ we can solve the resulting equation for $y.$

$$\begin{aligned} 3y &= 2 \\ y &= \frac{2}{3} \end{aligned}$$

Then, we substitute this value for y into one of the original equations and solve for $x.$

$$\begin{aligned} -x + y &= 3 \\ -x + \frac{2}{3} &= 3 \\ -x &= 3 - \frac{2}{3} \\ -x &= \frac{7}{3} \\ x &= -\frac{7}{3} \end{aligned}$$

The solution to this system is $\left(-\frac{7}{3}, \frac{2}{3}\right).$

Check the solution in the first equation.

$$\begin{aligned} x + 2y &= -1 \\ \left(-\frac{7}{3}\right) + 2\left(\frac{2}{3}\right) &= \\ -\frac{7}{3} + \frac{4}{3} &= \\ -\frac{3}{3} &= \\ -1 &= -1 \qquad \text{True} \end{aligned}$$

Analysis

We gain an important perspective on systems of equations by looking at the graphical representation. See (Figure) to find that the equations intersect at the solution. We do not need to ask whether there may be a second solution because observing the graph confirms that the system has exactly one solution.

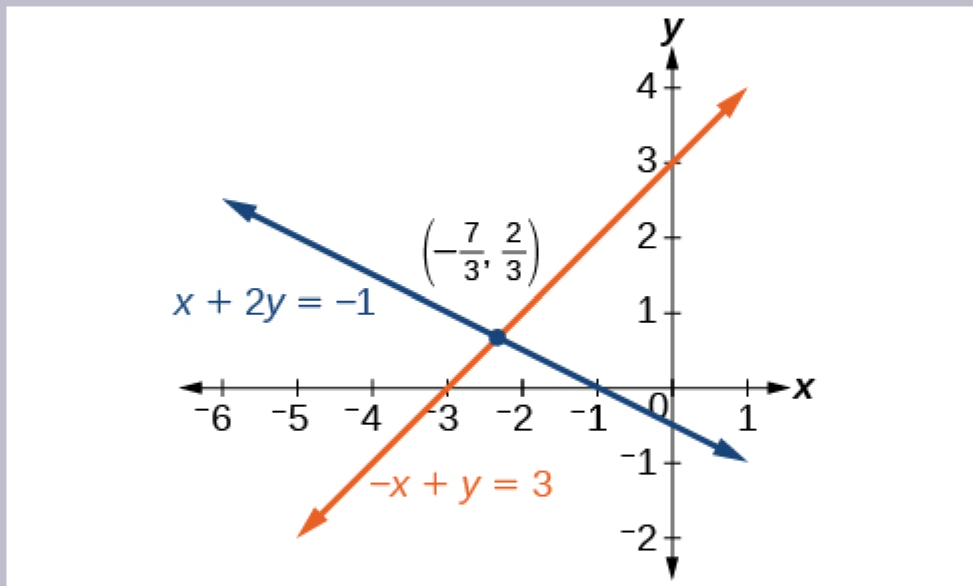


Figure 5.

Using the Addition Method When Multiplication of One Equation Is Required

Solve the given system of equations by the addition method.

$$3x + 5y = -11$$

$$x - 2y = 11$$

Adding these equations as presented will not eliminate a variable. However, we see that the first equation has $3x$ in it and the second equation has x. So if we multiply the second equation by $-3,$ the *x*-terms will add to zero.

$$\begin{array}{ll} x - 2y = 11 & \\ -3\left(x - 2y\right) = -3\left(11\right) & \text{Multiply both sides by } -3. \\ -3x + 6y = -33 & \text{Use the distributive property.} \end{array}$$

Now, let's add them.

$$
\begin{aligned}
3x+5y&=-11\\
-3x+6y&=-33
\end{aligned}
$$

$$
\begin{aligned}
11y&=-44\\
y&=-4
\end{aligned}
$$

For the last step, we substitute $y=-4$ into one of the original equations and solve for x.

$$
\begin{aligned}
3x+5y&=-11\\
3x+5(-4)&=-11\\
3x-20&=-11\\
3x&=9\\
x&=3
\end{aligned}
$$

Our solution is the ordered pair $(3,-4)$. See (Figure). Check the solution in the original second equation.

$$
\begin{aligned}
x-2y&=11\\
(3)-2(-4)&=3+8\\
11&=11 \qquad \text{True}
\end{aligned}
$$

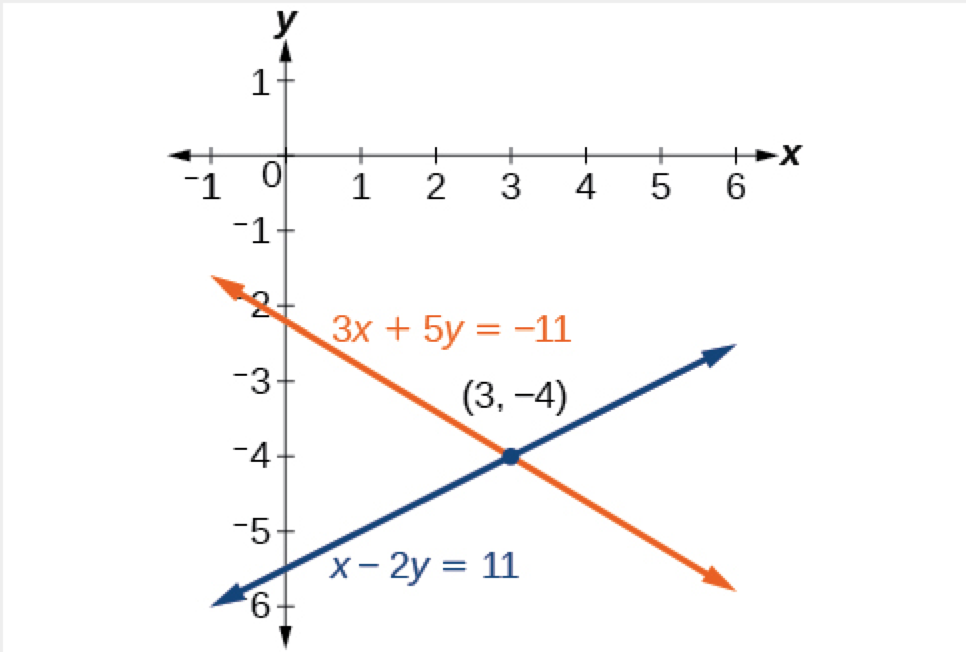


Figure 6.

Try It

Solve the system of equations by addition.

$$2x - 7y = 2$$
$$3x + y = -20$$

$(-6, -2)$

Using the Addition Method When Multiplication of Both Equations Is Required

Solve the given system of equations in two variables by addition.

$$2x+3y=-16$$
$$5x-10y=30$$

One equation has $2x$ and the other has $5x.$ The least common multiple is $10x$ so we will have to multiply both equations by a constant in order to eliminate one variable. Let's eliminate x by multiplying the first equation by -5 and the second equation by $2.$

$$\begin{aligned}-5\left(2x+3y\right)&=-5\left(-16\right)\\ -10x-15y&=80\\ 2\left(5x-10y\right)&=2\left(30\right)\\ 10x-20y&=60\end{aligned}$$

Then, we add the two equations together.

$$\begin{aligned}-10x-15y&=80\\ 10x-20y&=60\\ \hline -35y&=140\\ y&=-4\end{aligned}$$

Substitute $y=-4$ into the original first equation.

$$\begin{aligned}2x+3\left(-4\right)&=-16\\ 2x-12&=-16\\ 2x&=-4\\ x&=-2\end{aligned}$$

The solution is $\left(-2,-4\right).$ Check it in the other equation.

$$\begin{aligned}5x-10y&=30\\ 5\left(-2\right)-10\left(-4\right)&=30\\ -10+40&=30\\ 30&=30\end{aligned}$$

See (Figure).

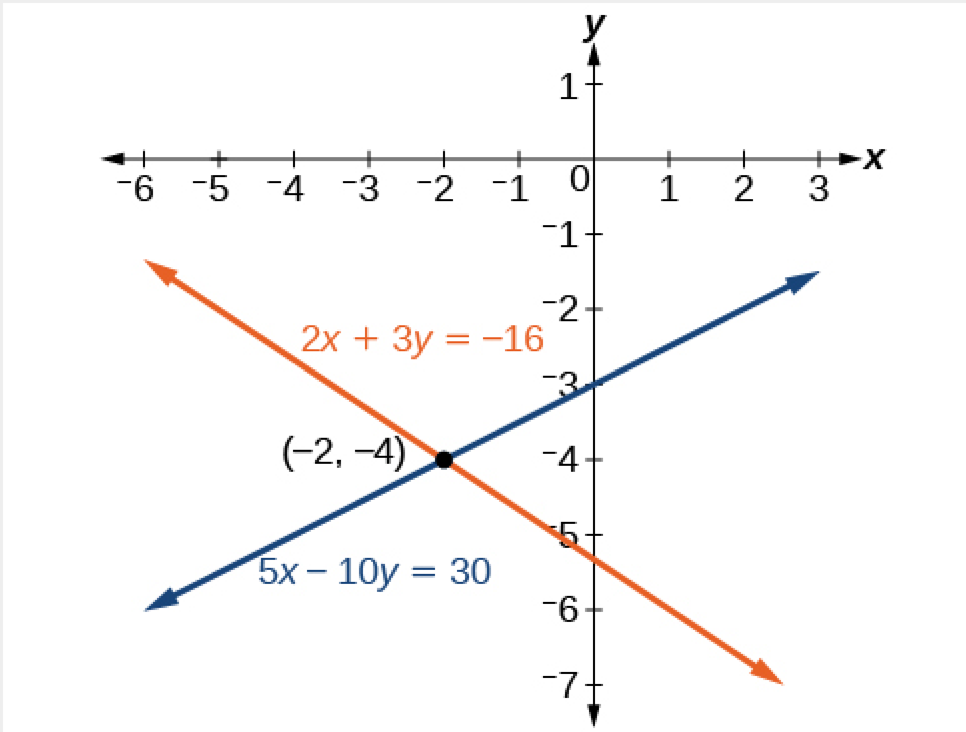


Figure 7.

Using the Addition Method in Systems of Equations Containing Fractions

Solve the given system of equations in two variables by addition.

$$\frac{x}{3}+\frac{y}{6}=3$$
$$\frac{x}{2}-\frac{y}{4}=\ 1$$

First clear each equation of fractions by multiplying both sides of the equation by the least common denominator.

$$6\left(\frac{x}{3}+\frac{y}{6}\right)=6\left(3\right)$$
$$2x+y=18$$
$$4\left(\frac{x}{2}-\frac{y}{4}\right)=4\left(1\right)$$
$$2x-y=4$$

Now multiply the second equation by -1 so that we can eliminate the *x*-variable.

$$-1\left(2x-y\right)=-1\left(4\right)$$
$$-2x+y=-4$$

Add the two equations to eliminate the *x*-variable and solve the resulting equation.

$$\begin{aligned} 2x+y&=18\\ -2x+y&=-4\\ \hline 2y&=14\\ y&=7 \end{aligned}$$

Substitute $y=7$ into the first equation.

$$2x+\left(7\right)=18$$
$$2x=11$$
$$x=\frac{11}{2}$$
$$=5.5$$

The solution is $\left(\frac{11}{2},7\right)$. Check it in the other equation.

$$\frac{x}{2}-\frac{y}{4}=1$$
$$\frac{\frac{11}{2}}{2}-\frac{7}{4}=1$$
$$\frac{11}{4}-\frac{7}{4}=1$$
$$\frac{4}{4}=1$$

Try It

Solve the system of equations by addition.

$$2x+3y=8$$
$$3x+5y=10$$

$(10, -4)$

Identifying Inconsistent Systems of Equations Containing Two Variables

Now that we have several methods for solving systems of equations, we can use the methods to identify inconsistent systems. Recall that an inconsistent system consists of parallel lines that have the same slope but different y-intercepts. They will never intersect. When searching for a solution to an inconsistent system, we will come up with a false statement, such as $12 = 0$.

Solving an Inconsistent System of Equations

Solve the following system of equations.

$$\begin{aligned} x &= 9 - 2y \\ x + 2y &= 13 \end{aligned}$$

We can approach this problem in two ways. Because one equation is already solved for x, the most obvious step is to use substitution.

$$\begin{aligned} x + 2y &= 13 \\ (9 - 2y) + 2y &= 13 \\ 9 + 0y &= 13 \\ 9 &= 13 \end{aligned}$$

Clearly, this statement is a contradiction because $9 \neq 13$. Therefore, the system has no solution.

The second approach would be to first manipulate the equations so that they are both in slope-intercept form. We manipulate the first equation as follows.

$$\begin{aligned} x &= 9 - 2y \\ 2y &= -x + 9 \\ y &= -\tfrac{1}{2}x + \tfrac{9}{2} \end{aligned}$$

We then convert the second equation expressed to slope-intercept form.

$$\begin{aligned} x+2y &= 13 \\ 2y &= -x+13 \\ y &= -\tfrac{1}{2}x+\tfrac{13}{2} \end{aligned}$$

Comparing the equations, we see that they have the same slope but different *y*-intercepts. Therefore, the lines are parallel and do not intersect.

$$\begin{aligned} y &= -\tfrac{1}{2}x+\tfrac{9}{2} \\ y &= -\tfrac{1}{2}x+\tfrac{13}{2} \end{aligned}$$

Analysis

Writing the equations in slope-intercept form confirms that the system is inconsistent because all lines will intersect eventually unless they are parallel. Parallel lines will never intersect; thus, the two lines have no points in common. The graphs of the equations in this example are shown in (Figure).

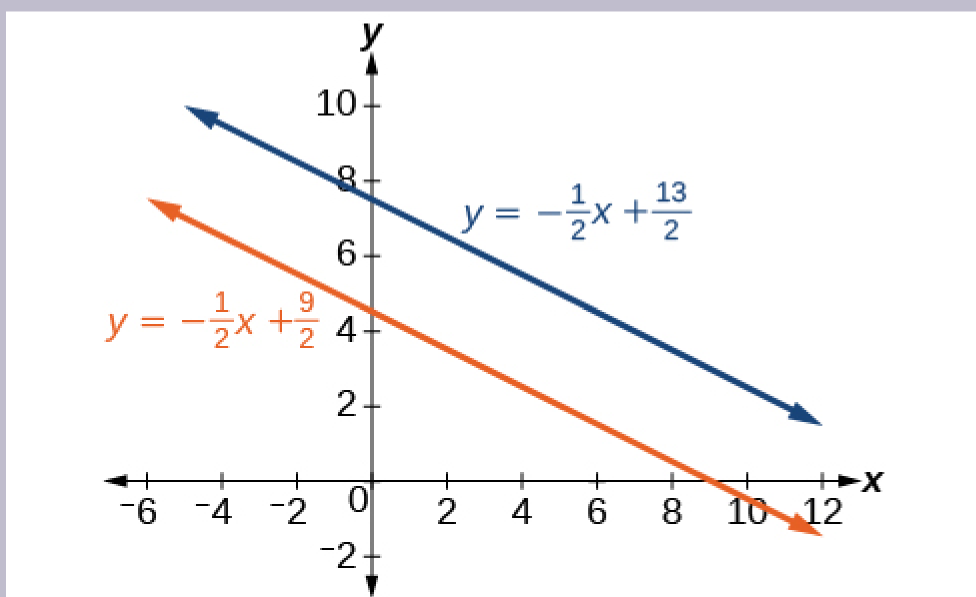


Figure 8.

Try It

Solve the following system of equations in two variables.

$$2y - 2x = 2$$
$$2y - 2x = 6$$

No solution. It is an inconsistent system.

Expressing the Solution of a System of Dependent Equations Containing Two Variables

Recall that a dependent system of equations in two variables is a system in which the two equations represent the same line. Dependent systems have an infinite number of solutions because all of the points on one line are also on the other line. After using substitution or addition, the resulting equation will be an identity, such as $0 = 0.$

Finding a Solution to a Dependent System of Linear Equations

Find a solution to the system of equations using the addition method.

$$x + 3y = 2$$
$$3x + 9y = 6$$

With the addition method, we want to eliminate one of the variables by adding the equations. In this case, let's focus on eliminating x. If we multiply both sides of the first equation by $-3,$ then we will be able to eliminate the x-variable.

$$x + 3y = 2$$
$$(-3)\,(x + 3y) = (-3)\,(2)$$
$$-3x - 9y = -6$$

Now add the equations.

$$\begin{array}{rl} -3x - 9y & = -6 \\ + \quad 3x + 9y & = 6 \\ \hline 0 & = 0 \end{array}$$

We can see that there will be an infinite number of solutions that satisfy both equations.

Analysis

If we rewrote both equations in the slope-intercept form, we might know what the solution would look like before adding. Let's look at what happens when we convert the system to slope-intercept form.

$$\begin{aligned} x + 3y &= 2 \\ 3y &= -x + 2 \\ y &= -\frac{1}{3}x + \frac{2}{3} \end{aligned}$$

$$\begin{aligned} 3x + 9y &= 6 \\ 9y &= -3x + 6 \\ y &= -\frac{3}{9}x + \frac{6}{9} \\ y &= -\frac{1}{3}x + \frac{2}{3} \end{aligned}$$

See (Figure). Notice the results are the same. The general solution to the system is $\left(x, -\frac{1}{3}x + \frac{2}{3}\right)$.

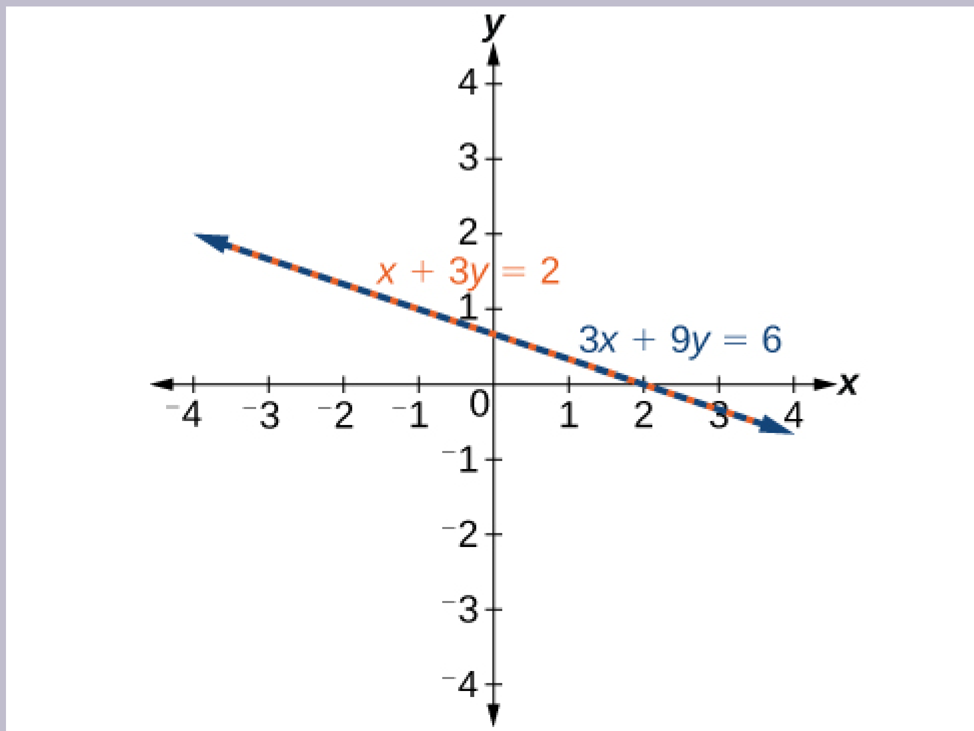


Figure 9.

Try It

Solve the following system of equations in two variables.

$$\begin{aligned} y - 2x &= 5 \\ -3y + 6x &= -15 \end{aligned}$$

The system is dependent so there are infinite solutions of the form $(x, 2x + 5)$.

Using Systems of Equations to Investigate Profits

Using what we have learned about systems of equations, we can return to the skateboard manufacturing problem at the beginning of the section. The skateboard manufacturer's revenue function is the function used to calculate the amount of money that comes into the business. It can be represented by the equation $R = xp,$ where $x =$ quantity and $p =$ price. The revenue function is shown in orange in (Figure).

The cost function is the function used to calculate the costs of doing business. It includes fixed costs, such as rent and salaries, and variable costs, such as utilities. The cost function is shown in blue in (Figure). The x-axis represents quantity in hundreds of units. The *y*-axis represents either cost or revenue in hundreds of dollars.

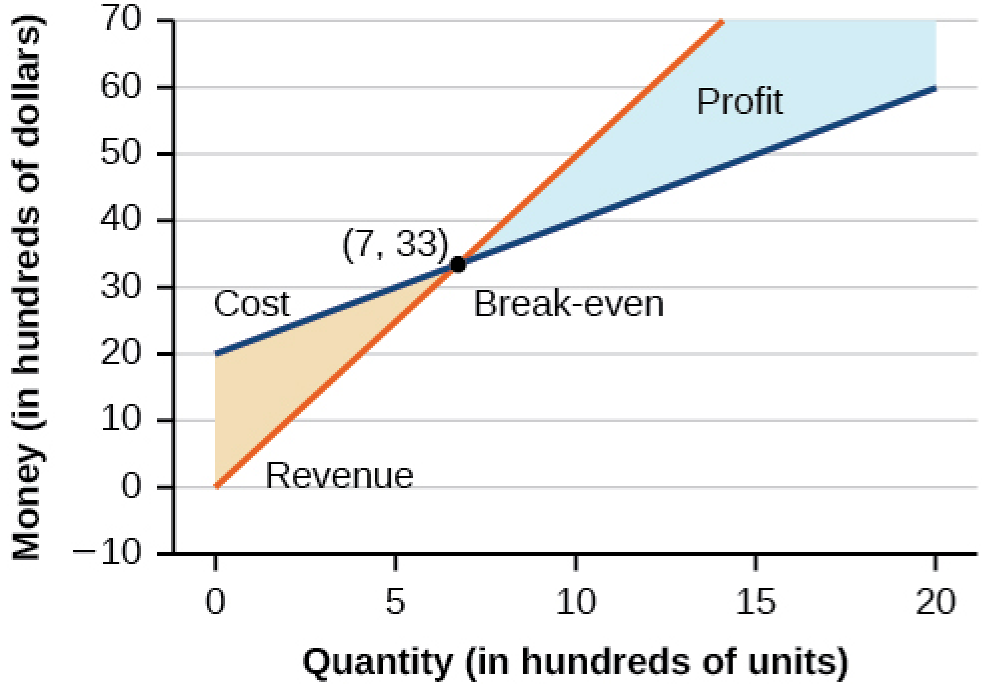


Figure 10.

The point at which the two lines intersect is called the break-even point. We can see from the graph that if 700 units are produced, the cost is $3,300 and the revenue is also $3,300. In other words, the company breaks even if they produce and sell 700 units. They neither make money nor lose money.

The shaded region to the right of the break-even point represents quantities for which the company makes a profit. The shaded region to the left represents quantities for which the company suffers a loss. The profit function is the revenue function minus the cost function, written as $P\left(x\right) = R\left(x\right) - C\left(x\right)$. Clearly, knowing the quantity for which the cost equals the revenue is of great importance to businesses.

Finding the Break-Even Point and the Profit Function Using Substitution

Given the cost function $C\left(x\right) = 0.85x + 35,000$ and the revenue function $R\left(x\right) = 1.55x,$ find the break-even point and the profit function.

Write the system of equations using y to replace function notation.

$$y = 0.85x + 35,000$$
$$y = 1.55x$$

Substitute the expression $0.85x + 35,000$ from the first equation into the second equation and solve for x.

$$\begin{aligned} 0.85x + 35,000 &= 1.55x \\ 35,000 &= 0.7x \\ 50,000 &= x \end{aligned}$$

Then, we substitute $x = 50,000$ into either the cost function or the revenue function.

$$1.55\left(50,000\right) = 77,500$$

The break-even point is $\left(50,000, 77,500\right)$.

The profit function is found using the formula $P\left(x\right) = R\left(x\right) - C\left(x\right)$.

$$\begin{aligned} P\left(x\right) &= 1.55x - \left(0.85x + 35,000\right) \\ &= 0.7x - 35,000 \end{aligned}$$

The profit function is $P\left(x\right) = 0.7x - 35,000.$

Analysis

The cost to produce 50,000 units is $77,500, and the revenue from the sales of 50,000 units is also $77,500. To make a profit, the business must produce and sell more than 50,000 units. See (Figure).

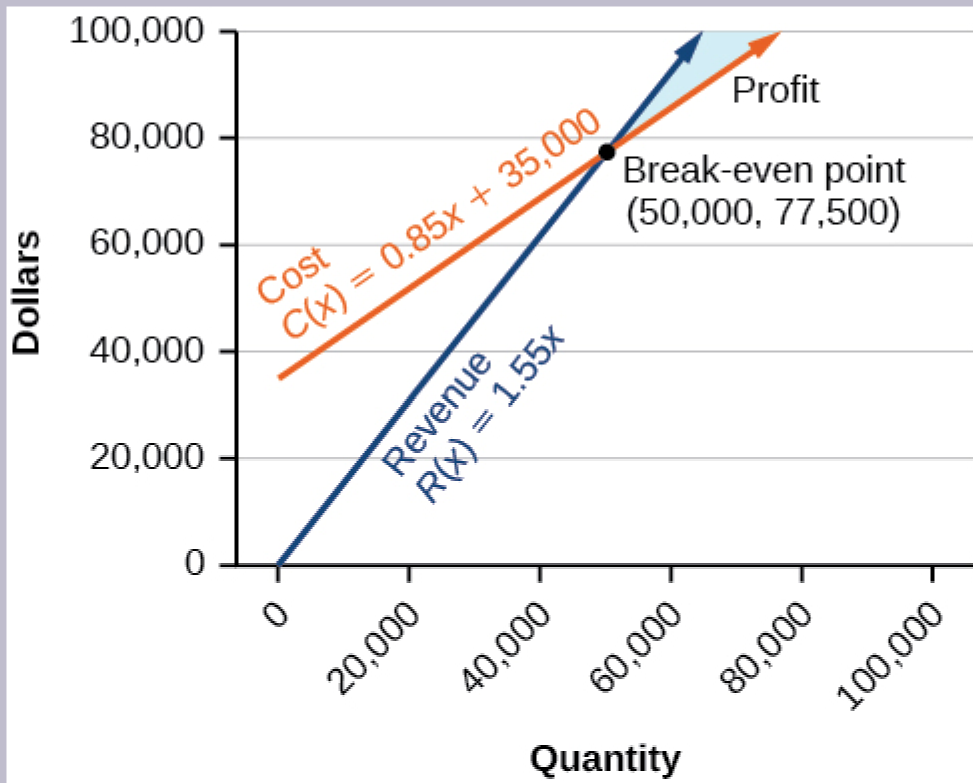


Figure 12.

We see from the graph in (Figure) that the profit function has a negative value until $x = 50,000,$ when the graph crosses the *x*-axis. Then, the graph emerges into positive *y*-values and continues on this path as the profit function is a straight line. This illustrates that the break-even point for businesses occurs when the profit function is 0. The area to the left of the break-even point represents operating at a loss.

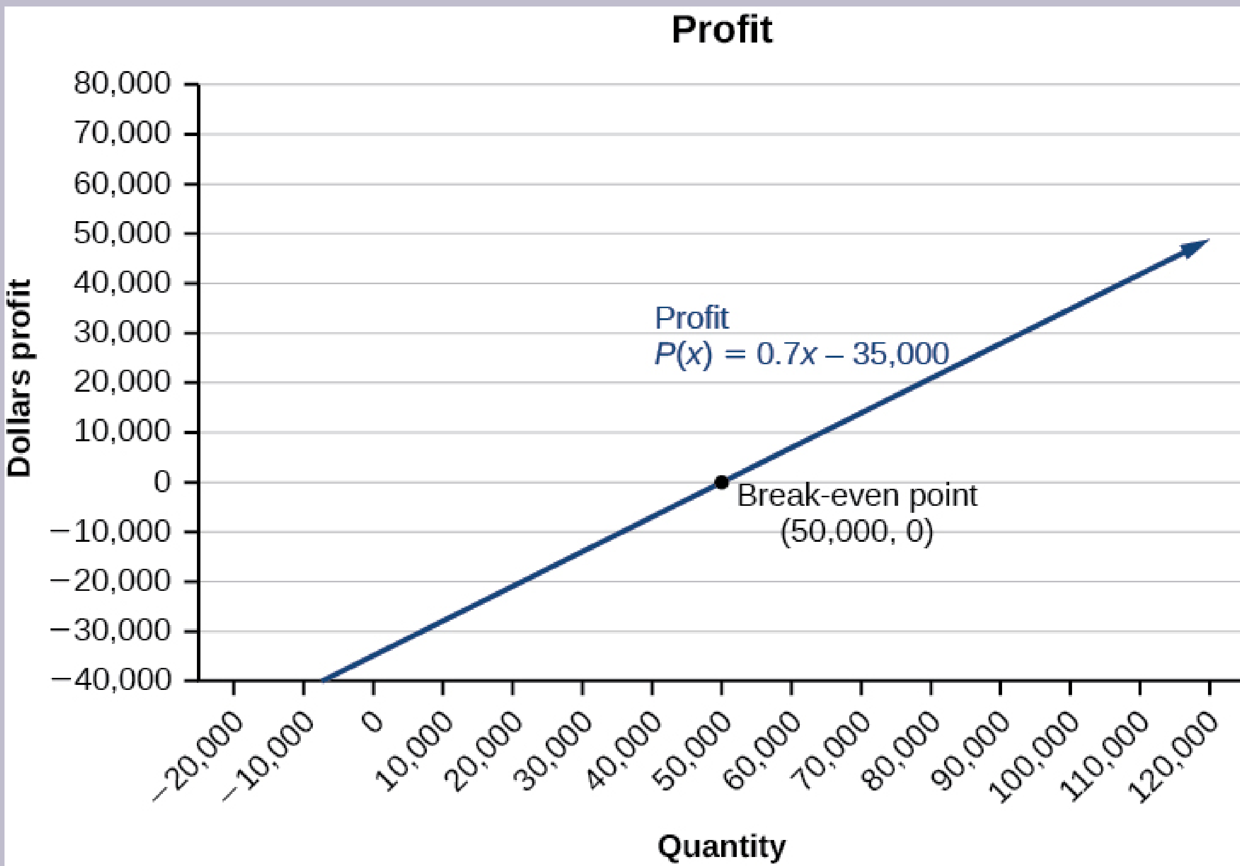


Figure 13.

Writing and Solving a System of Equations in Two Variables

The cost of a ticket to the circus is $\$25.00$ for children and $\$50.00$ for adults. On a certain day, attendance at the circus is $2,000$ and the total gate revenue is $\$70,000.$ How many children and how many adults bought tickets?

Let *c* = the number of children and *a* = the number of adults in attendance.

The total number of people is $2,000.$ We can use this to write an equation for the number of people at the circus that day.

$$c + a = 2,000$$

The revenue from all children can be found by multiplying $\$25.00$ by the number of children, $25c$. The revenue from all adults can be found by multiplying $\$50.00$ by the number of adults, $50a$. The total revenue is $\$70,000.$ We can use this to write an equation for the revenue.

$$25c + 50a = 70,000$$

We now have a system of linear equations in two variables.

$$\begin{aligned} c + a &= 2,000 \\ 25c + 50a &= 70,000 \end{aligned}$$

In the first equation, the coefficient of both variables is 1. We can quickly solve the first equation for either c or a. We will solve for a.

$$\begin{aligned} c + a &= 2,000 \\ a &= 2,000 - c \end{aligned}$$

Substitute the expression $2,000 - c$ in the second equation for a and solve for c.

$$\begin{aligned} 25c + 50\left(2,000 - c\right) &= 70,000 \\ 25c + 100,000 - 50c &= 70,000 \\ -25c &= -30,000 \\ c &= 1,200 \end{aligned}$$

Substitute $c = 1,200$ into the first equation to solve for a.

$$\begin{aligned} 1,200 + a &= 2,000 \\ a &= 800 \end{aligned}$$

We find that $1,200$ children and 800 adults bought tickets to the circus that day.

Try It

Meal tickets at the circus cost $\$4.00$ for children and $\$12.00$ for adults. If $1,650$ meal tickets were bought for a total of $\$14,200,$ how many children and how many adults bought meal tickets?

700 children, 950 adults

Access these online resources for additional instruction and practice with systems of linear equations.

- Solving Systems of Equations Using Substitution
- Solving Systems of Equations Using Elimination
- Applications of Systems of Equations

Key Concepts

- A system of linear equations consists of two or more equations made up of two or more variables such that all equations in the system are considered simultaneously.
- The solution to a system of linear equations in two variables is any ordered pair that satisfies each equation independently. See (Figure).
- Systems of equations are classified as independent with one solution, dependent with an infinite number of solutions, or inconsistent with no solution.
- One method of solving a system of linear equations in two variables is by graphing. In this method, we graph the equations on the same set of axes. See (Figure).
- Another method of solving a system of linear equations is by substitution. In this method, we solve for one variable in one equation and substitute the result into the second equation. See (Figure).
- A third method of solving a system of linear equations is by addition, in which we can eliminate a variable by adding opposite coefficients of corresponding variables. See (Figure).
- It is often necessary to multiply one or both equations by a constant to facilitate elimination of a variable when adding the two equations together. See (Figure), (Figure), and (Figure).
- Either method of solving a system of equations results in a false statement for inconsistent systems because they are made up of parallel lines that never intersect. See (Figure).
- The solution to a system of dependent equations will always be true because both equations describe the same line. See (Figure).
- Systems of equations can be used to solve real-world problems that involve more than one variable, such as those relating to revenue, cost, and profit. See (Figure) and (Figure).

addition method

an algebraic technique used to solve systems of linear equations in which the equations are added in a way that eliminates one variable, allowing the resulting equation to be solved for the remaining variable; substitution is then used to solve for the first variable

break-even point

the point at which a cost function intersects a revenue function; where profit is zero

consistent system

a system for which there is a single solution to all equations in the system and it is an independent system, or if there are an infinite number of solutions and it is a dependent system

cost function

the function used to calculate the costs of doing business; it usually has two parts, fixed costs and variable costs

dependent system

a system of linear equations in which the two equations represent the same line; there are an infinite number of solutions to a dependent system

inconsistent system

a system of linear equations with no common solution because they represent parallel lines, which have no point or line in common

independent system

a system of linear equations with exactly one solution pair (x, y)

profit function

the profit function is written as $P\left(x\right)=R\left(x\right)-C\left(x\right)$, revenue minus cost

revenue function

the function that is used to calculate revenue, simply written as $R=xp,$ where $x=$ quantity and $p=$ price

substitution method

an algebraic technique used to solve systems of linear equations in which one of the two equations is solved for one variable and then substituted into the second equation to solve for the second variable

system of linear equations

a set of two or more equations in two or more variables that must be considered simultaneously.

30. Homework Problems

Homework

The list below is the list of homework problems in Edfinity. The numbering is the same. You can click on a problem, and it will take you to the relevant section of the book!

1. Which ions are important in understanding neuron function
2. Direction of charge flow and current.
3. How much charge in a defibrillator?
4. How long is a lightning bolt?
5. What are the properties of an ideal battery?
6. Charge stored on a capacitor.
7. Bookshelf capacitor.
8. Capacitor with neoprene.
9. What does capacitance depend upon?
10. Voltage in defibrillator.
11. Resistance in defibrillator.
12. Simple circuit.
13. Resistance in extension cords.
14. Truck starter motor.
15. Current from electric appliances.
16. Battery life-time.
17. Adding current.
18. Two equations with two unknowns.
19. Three equations with three unknowns.

PART V

UNIT V

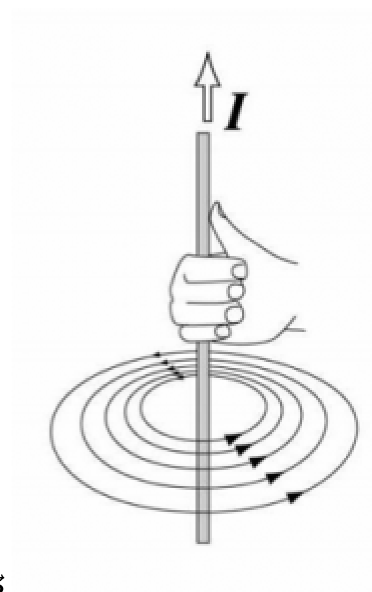
Unit V On-a-Page

Homework

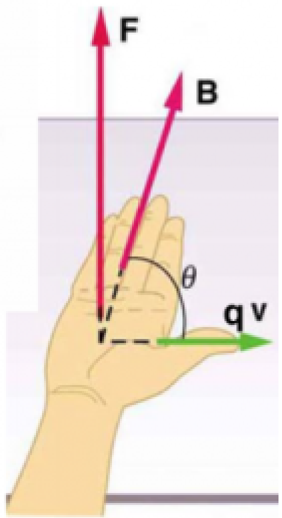
[All of the homework for this unit can be found at this link.]

Given that this unit is short, so is the homework!

- Only moving charges (i.e. currents) generate and feel magnetic fields
- The field is "real" and the key object (just like for electricity)
- Moving charges make fields which can then exert forces on other moving charges
- Use your right hand: thumb in current ($\vec{I}$ or $q\vec{v}$), fingers in $\vec{B}$



- For the direction of a magnetic field, fingers curl in direction of $\vec{B}$
- $|\vec{B}| = \frac{I\mu_0}{2\pi r}$
- μ_0 is another universal constant like h, c, ε_0, and $\mu_0 = 4\pi \times 10^{-7}\,\frac{N}{A^2}$
- For the force, palm "pushes" in the direction of $\vec{F}$

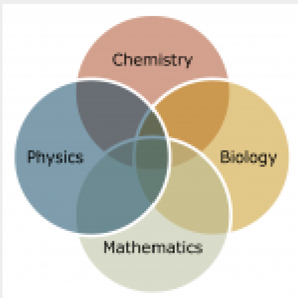


$F = qvB \sin\theta$

- $F = qVBsin\theta$ or $F = ILBsin\theta$
- Result is particles in uniform $\vec{B}$ move in circles!!

31. Introduction

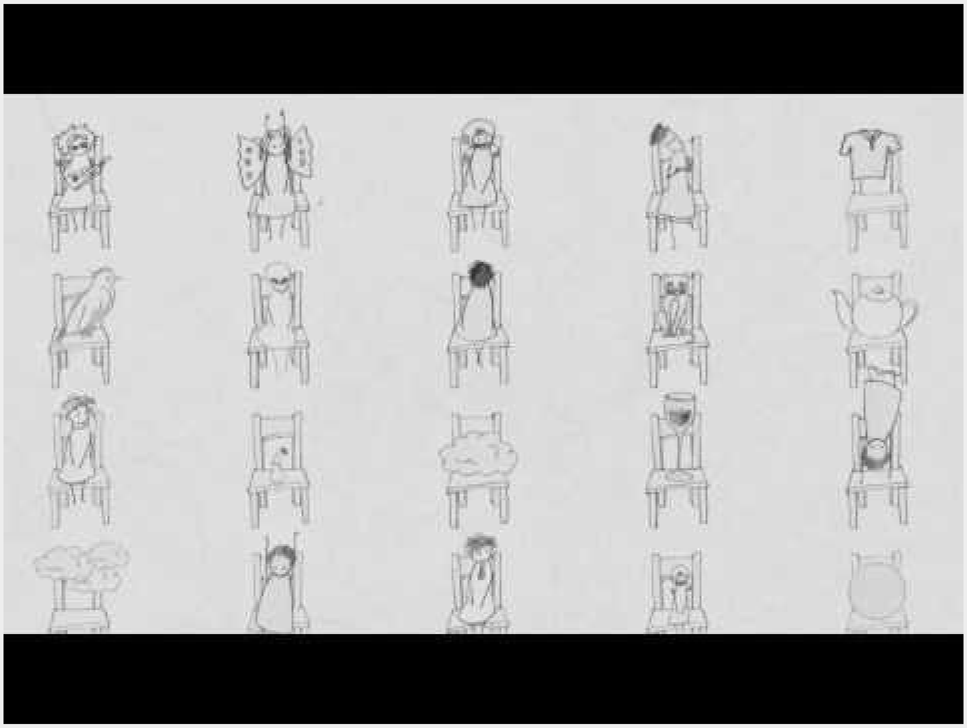
Organizing Principle for this Unit



Physics sees beauty in simplicity. Meanwhile biology sees beauty in complexity.

– I heard this somewhere, but I cannot find the source

I think this quote sums things up nicely. In biology, we look with wonder at the huge diversity of life on Earth and all the solutions evolution has developed over the eons. In physics, as I hope this course has demonstrated, we like to try to explain as many different phenomena with the smallest number of ideas. I feel that this is summarized nicely in the video below:



A YouTube element has been excluded from this version of the text. You can view it online here: http://openbooks.library.umass.edu/toggerson-132/?p=467

Unlike prior units, which had an explicit connection to your other courses we were exploring, this unit is about really all about physics' idea of beauty in simplicity. In this unit, we will bring together all of the different ideas that we have talked about over the duration of this course: light, electrons, charge, wave-particle duality, electric field, and potential into a beautiful whole showing that everything is connected to everything else!

Introduction to Magnetism

***Figure 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)*

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.

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 2. *Engineering of technology like iPods would not be possible without a deep understanding magnetism. (credit: Jesse! S?, Flickr)*

Magnets

Derived from [Magnets](#) by OpenStax

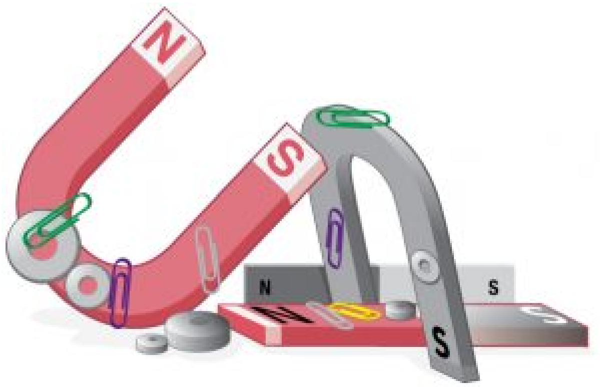


Figure 1. *Magnets come in various shapes, sizes, and strengths. All have both a north pole and a south pole. There is never an isolated pole (a monopole).*

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 Magnetic 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.

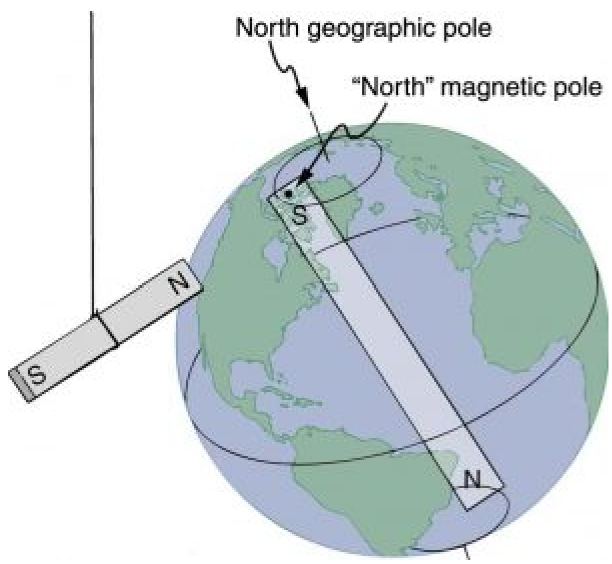


Figure 2. *One end of a bar magnet is suspended from a thread that points toward north. The magnet's two poles are labeled N and S for north-seeking and south-seeking poles, respectively.*

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.

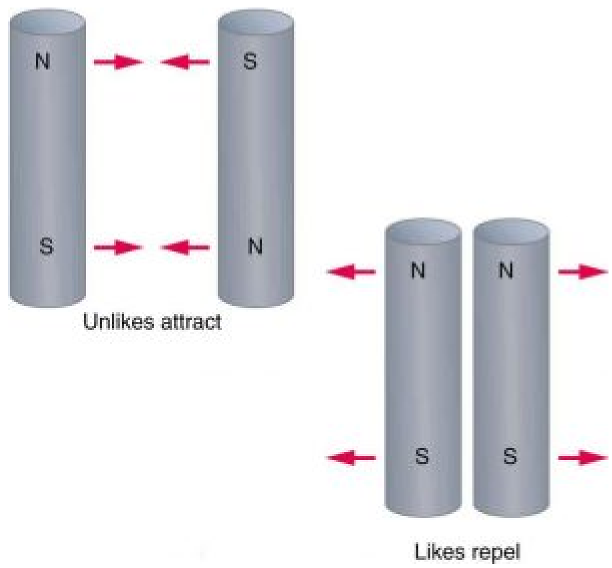


Figure 3. *Unlike poles attract, whereas like poles repel.*

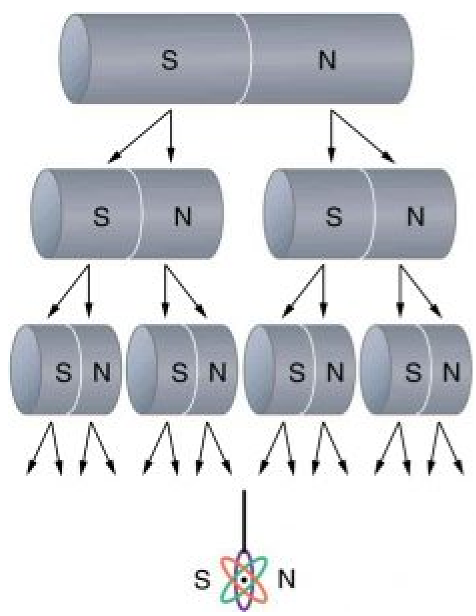


Figure 4*. North and south poles always occur in pairs. Attempts to separate them result in more pairs of poles. If we continue to split the magnet, we will eventually get down to an iron atom with a north pole and a south pole—these, too, cannot be separated.*

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?

Section 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.

Sources of Magnetism

Derived from Ferromagnets and Electromagnets by OpenStax

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.

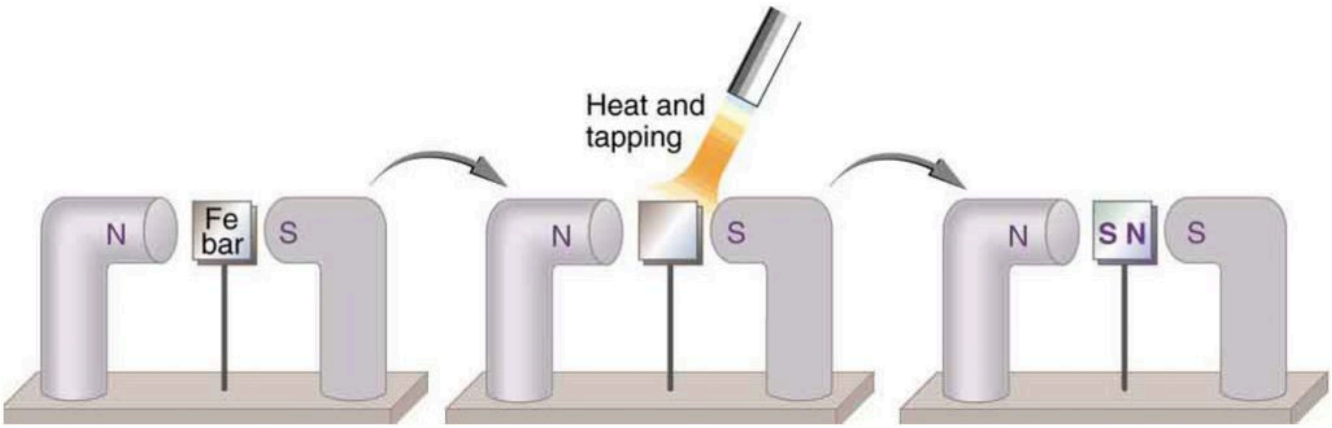


Figure 1. *An unmagnetized piece of iron is placed between two magnets, heated, and then cooled, or simply tapped when cold. The iron becomes a permanent magnet with the poles aligned as shown: its south pole is adjacent to the north pole of the original magnet, and its north pole is adjacent to the 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 1. (This results in the attraction of the previously unmagnetized material to the magnet.) What happens on a microscopic scale is illustrated in Figure 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 2(b). This induced magnetization can be made permanent if the material is heated and then cooled, or simply tapped in the presence of other magnets.

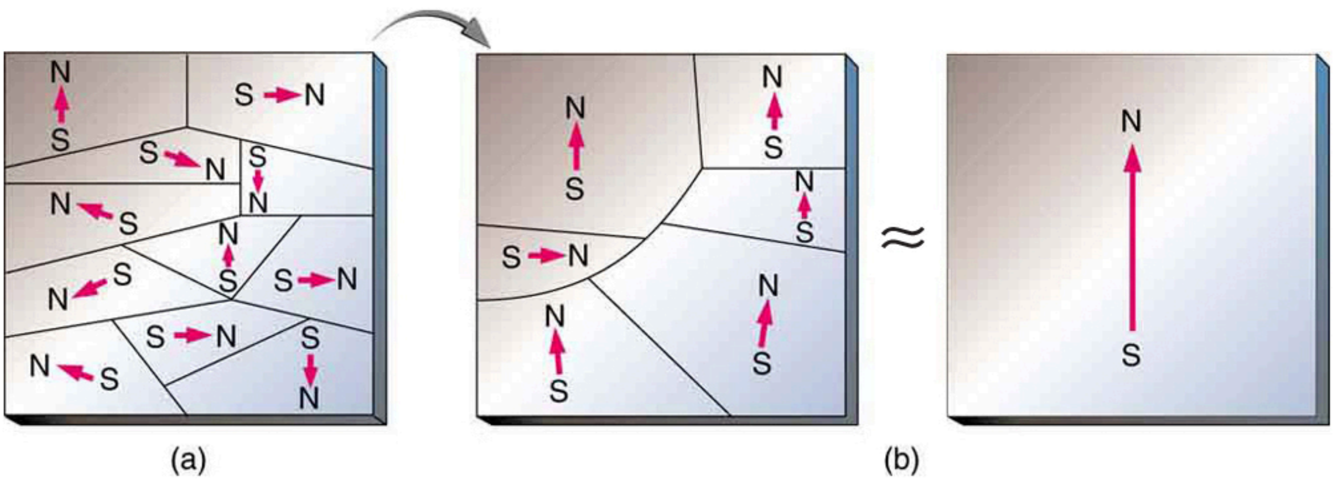


Figure 2. *(a) An unmagnetized piece of iron (or other ferromagnetic material) has randomly oriented domains. (b) When magnetized by an external field, the domains show greater alignment, and some grow at the expense of others. Individual atoms are aligned within domains; each atom acts like a tiny bar 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 $770C$ 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 (See Figure 3).

Figure 4. *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 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.

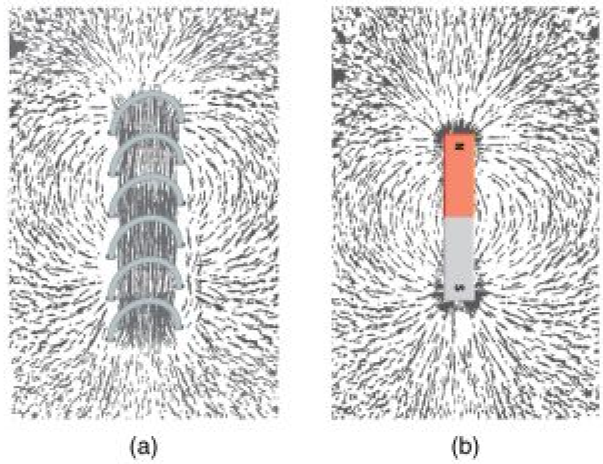


Figure 4. Iron filings near (a) a current-carrying coil and (b) a magnet act like tiny compass needles, showing the shape of their fields. Their response to a current-carrying coil and a permanent magnet is seen to be very similar, especially near the ends of the coil and the magnet.

Combining a ferromagnet with an electromagnet can produce particularly strong magnetic effects. (See Figure.) 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 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 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.

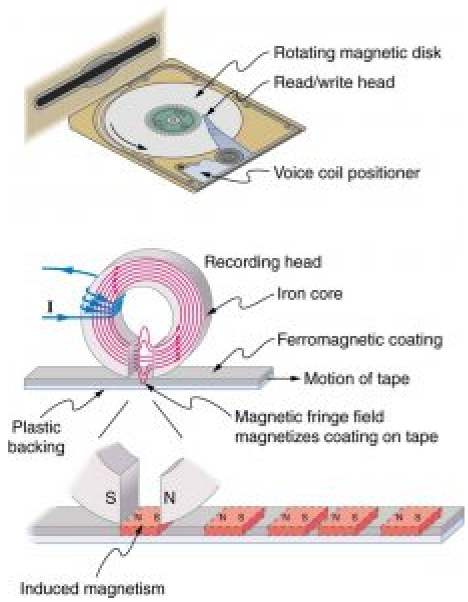


Figure 6. *An electromagnet induces regions of permanent magnetism on a floppy disk coated with a ferromagnetic material. The information stored here is digital (a region is either magnetic or not); in other applications, it can be analog (with a varying strength), such as on audiotapes.*

Current: The Source of All Magnetism

Instructor's Note

The fact that current is the source of all magnetism is the IMPORTANT POINT- all magnetic fields are ultimately created by moving charges.

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 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 sub-nuclear 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.

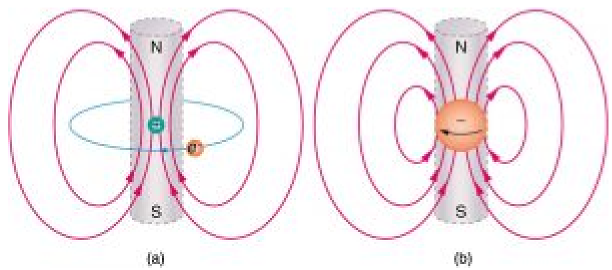


Figure 7. *(a) In the planetary model of the atom, an electron orbits a nucleus, forming a closed-current loop and producing a magnetic field with a north pole and a south pole. (b) Electrons have spin and can be crudely pictured as rotating charge, forming a current that produces a magnetic field with 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.*

Homework

1. The source of magnetic fields

Section 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.

32. Magnet Fields and What They Do

Magnetic Fields and Magnetic Field Lines

Derived from Magnetic Fields and Magnetic Field Lines by OpenStax

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 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.**

Instructor's Note

Why do we use B for magnetic field? I have no idea.

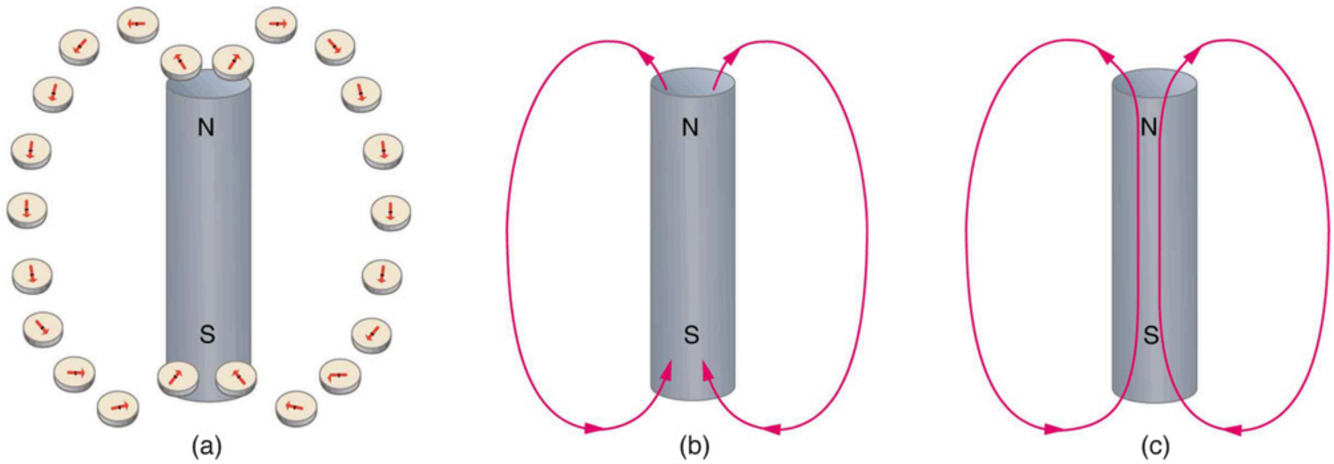


Figure 1. *Magnetic field lines are defined to have the direction that a small compass points when placed at a location. (a) If small compasses are used to map the magnetic field around a bar magnet, they will point in the directions shown: away from the north pole of the magnet, toward the south pole 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 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.

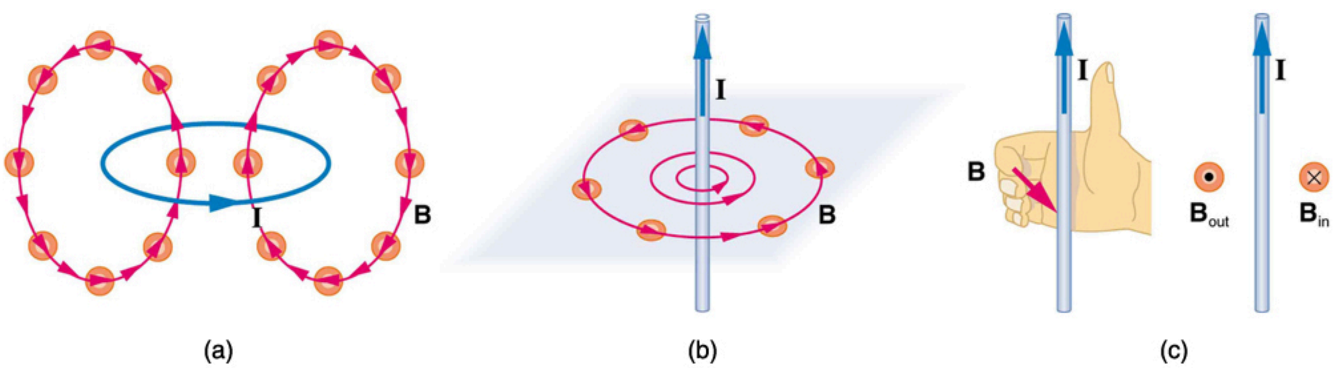


Figure 2. *Small compasses could be used to map the fields shown here. (a) The magnetic field of a circular current loop is similar to that of a bar magnet. (b) A long and straight wire creates a field with magnetic field lines forming circular loops. (c) When the wire is in the plane of the paper, the 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 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.

Homework

2. Compasses and magnetic field lines

Section 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.

Magnetic Field Strength: Force on a Moving Charge in a Magnetic Field

Derived from [Magnetic Field Strength: Force on a Moving Charge in a Magnetic Field] by OpenStax

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

In this section I am looking for you to understand:

- Only moving charges (i.e. currents) experience forces due to magnetic fields
- The letters in the expression $F = qvB$. As far as prep is concerned, the angle theta will always be 90-degrees, so sine will always be 1.

What is the mechanism by which one magnet exerts a force on another? The answer is related to the fact that all magnetism is caused by current, the flow of charge. *Magnetic fields exert forces on moving charges*, and so they exert forces on other magnets, all of which have moving charges.

Right Hand Rule 1

The magnetic force on a moving charge is one of the most fundamental known. Magnetic force is as important as the electrostatic or Coulomb force. Yet the magnetic force is more complex, in both the number of factors that affects it and in its direction, than the relatively simple Coulomb force. The magnitude of the **magnetic force** F on a charge q moving at a speed v in a magnetic field of strength B is given by

$$F = qvB\sin\theta$$

Instructor's Note

The only thing about the angle you need to know is that if the velocity if PARALLEL to the field, then there is no force!

where θ is the angle between the directions of v and B. This force is often called the **Lorentz force**. In fact, this is how we define the magnetic field strength B—in terms of the force on a charged particle moving in a magnetic field. The SI unit for magnetic field strength B is called the **tesla** (T) after the eccentric but brilliant inventor Nikola Tesla (1856–1943). To determine how the tesla relates to other SI units, we solve $F = qvB\sin\theta$ for B.

$$B = \frac{F}{qv\sin\theta}$$

Instructor's Note

We will spend time in class thinking about these complex directions!

Because $\sin\theta$ is unitless, the tesla is

$$1\ \text{T} = \frac{1\ \text{N}}{\text{C}\cdot\text{m/s}} = \frac{1\ \text{N}}{\text{A}\cdot\text{m}}$$

(note that C/s = A).

Another smaller unit, called the **gauss** (G), where $1\ \text{G} = 10^{-4}\ \text{T}$, is sometimes used. The strongest permanent magnets have fields near 2 T; superconducting electromagnets may attain 10 T or more. The Earth's magnetic field on its surface is only about $5\ \times\ 10^{-5}\ \text{T}$, or 0.5 G.

The *direction* of the magnetic force F is perpendicular to the plane formed by v and B, as determined by the **right hand rule 1** (or RHR-1), which is illustrated in Figure 1. RHR-1 states that, to determine the direction of the magnetic force on a positive moving charge, you point the thumb of the right hand in the direction of v, the fingers in the direction of B, and a perpendicular to the palm points in the direction of F. One way to remember this is that there is one velocity, and so the thumb represents it. There are many field lines, and so the fingers represent them. The force is in the direction you would push with your palm. The force on a negative charge is in exactly the opposite direction to that on a positive charge.

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

Again, we will spend more time in class going over these complex directions!

Magnetism is WAY more interesting then just bar magnets!

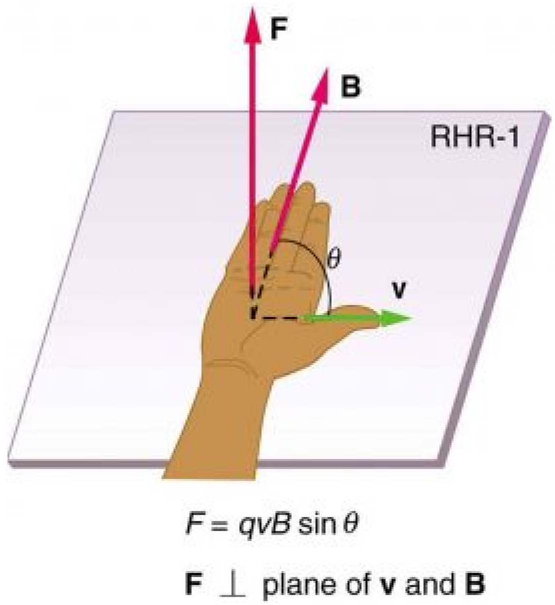


Figure 1. *Magnetic fields exert forces on moving charges. This force is one of the most basic known. The direction of the magnetic force on a moving charge is perpendicular to the plane formed by* v *and* B *and follows right hand rule–1 (RHR-1) as shown. The magnitude of the force is proportional to* q, v, B*, and the sine of the angle between* v *and* B*.*

Making Connections: Charges and Magnets

There is no magnetic force on static charges. However, there is a magnetic force on moving charges. When charges are stationary, their electric fields do not affect magnets. But, when charges move, they produce magnetic fields that exert forces on other magnets. When there is relative motion, a connection between electric and magnetic fields emerges—each affects the other.

Calculating Magnetic forces

Instructor's Note

This is a good example as I will expect you to solve for the numbers in problems such as this on your homework and on your quiz. I will NOT ask you about the directions.

With the exception of compasses, you seldom see or personally experience forces due to the Earth's small magnetic field. To illustrate this, suppose that in a physics lab you rub a glass rod with silk, placing a 20-nC positive charge on it. Calculate the force on the rod due to the Earth's magnetic field $5.0 \times 10^{-5}\ \mathrm{T}$, if you throw it with a horizontal velocity of 10 m/s due west in a place where the Earth's field is due north parallel to the ground. (The direction of the force is determined with right hand rule 1 as shown in Figure 2.)

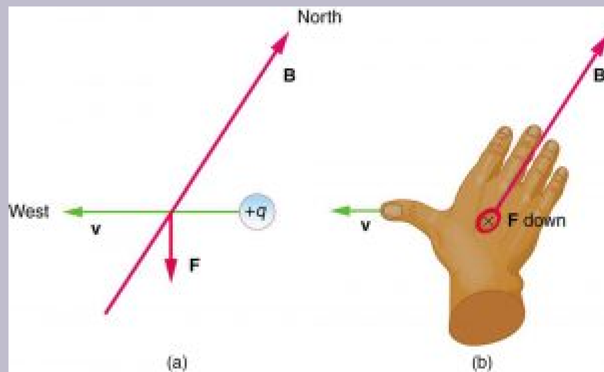


Figure 2. *A positively charged object moving due west in a region where the Earth's magnetic field is due north experiences a force that is straight down as shown. A negative charge moving in the same direction would feel a force straight up.*

Strategy

We are given the charge, its velocity, and the magnetic field strength and direction. We can thus use the equation $F = qvB\sin\theta$ to find the force.

Solution

The magnetic force is

$$F = qvB\sin\theta$$

We see that $\sin\theta = 1$, since the angle between the velocity and the direction of the field is 90°90° size 12{"90" rSup { size 8{ circ } } } {}">90°. Entering the other given quantities yields

$$F = (20 \times 10^{-9}\ \mathrm{C})(10\ \mathrm{m/s})(5 \times 10^{-5}\ \mathrm{T})$$
$$= 1 \times 10^{-11}(\mathrm{C}\ \cdot\ \mathrm{m/s})(\frac{\mathrm{N}}{\mathrm{C{\cdot}m/s}}) = 1 \times 10^{-11}\ \mathrm{N}.$$

Discussion

This force is completely negligible on any macroscopic object, consistent with experience. (It is calculated to only one digit, since the Earth's field varies with location and is given to only one digit.) The Earth's magnetic field, however, does produce very important effects, particularly on submicroscopic particles. Some of these are explored in Force on a Moving Charge in a Magnetic Field: Examples and Applications.

Section Summary

- Magnetic fields exert a force on a moving charge *q*, the magnitude of which is

$$F = qvB\sin\theta$$

where θ is the angle between the directions of v and B.

- The SI unit for magnetic field strength B is the tesla (T), which is related to other units by

$$1\ \mathrm{T} = \frac{1\ \mathrm{N}}{\mathrm{C} \cdot \mathrm{m/s}} = \frac{1\ \mathrm{N}}{\mathrm{A} \cdot \mathrm{m}}$$

- The *direction* of the force on a moving charge is given by right hand rule 1 (RHR-1): Point the thumb of the right hand in the direction of v, the fingers in the direction of B, and a perpendicular to the palm points in the direction of F.
- The force is perpendicular to the plane formed by $\mathbf{v}$ and $\mathbf{B}$. Since the force is zero if $\mathbf{v}$ is parallel to $\mathbf{B}$, charged particles often follow magnetic field lines rather than cross them.

Homework

3. Units of magnetic field.
4. What will feel a magnetic force.
5. What will feel a magnetic force – macroscopic objects.
6. Thinking about the magnetic forces on charged particles.
7. Magnetic force on an airplane.
8. Magnetic force on a baseball.
9. Magnetic force on an electron at an angle to $\vec{B}$.

Magnetic Force on a Current-Carrying Conductor

Derived from [Magnetic Force on a Current-Carrying Conductor]() by OpenStax

Instructor's Note

As with the section on magnetic forces on moving charged particles, I am NOT expecting you to master all of the material in this section. I am hoping that you will, by the end of this section:

- Know that only currents perpendicular to magnetic fields experience magnetic forces
- Be able to calculate the magnetic force on a section of wire of length L carrying a current I perpendicular to a magnetic field B using F=ILB

We will, as with charged particles, deal with the directions and the case of currents neither parallel or perpendicular to magnetic fields in class.

Because charges ordinarily cannot escape a conductor, the magnetic force on charges moving in a conductor is transmitted to the conductor itself.

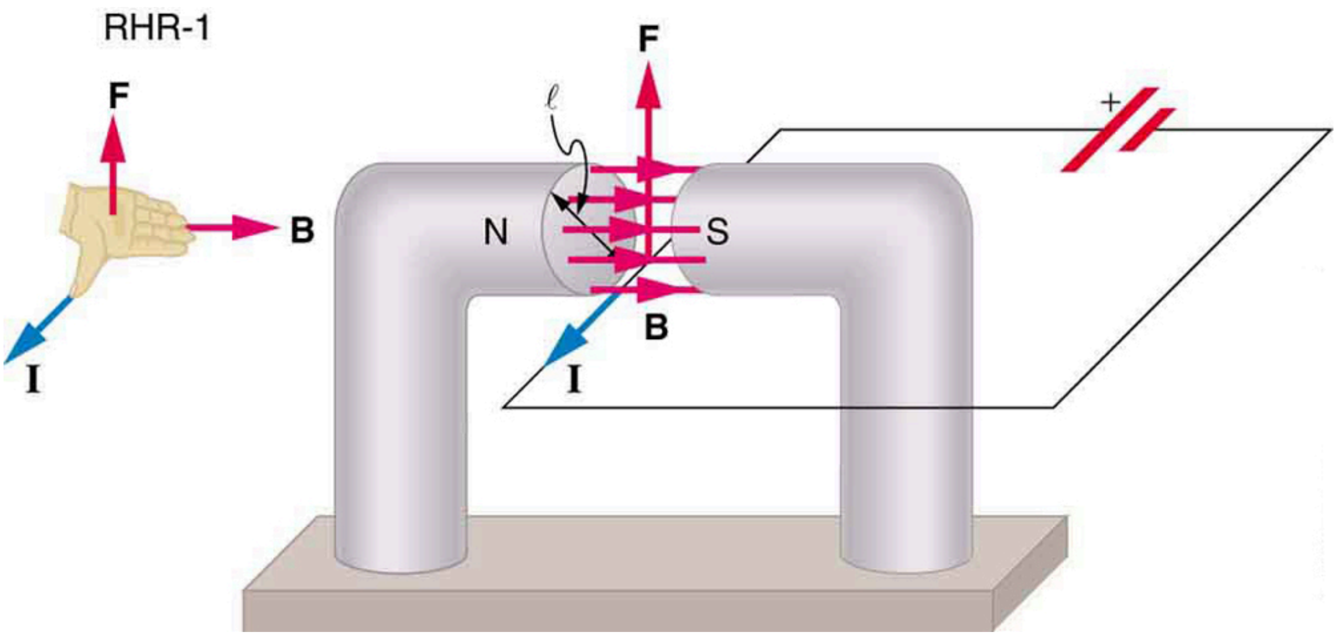


Figure 1. *The magnetic field exerts a force on a current-carrying wire in a direction given by the right hand rule 1 (the same direction as that on the individual moving charges). This force can easily be large enough to move the wire, since typical currents consist of very large numbers of moving charges.*

We can derive an expression for the magnetic force on a current by taking a sum of the magnetic forces on individual charges. (The forces add because they are in the same direction.) The force on an individual charge moving at the drift velocity v_d is given by $F = qv_dB\sin\theta$. Taking B to be uniform over a length of wire l and zero elsewhere, the total magnetic force on the wire is then $F = (qv_dB\sin\theta)(N)$, where N is the number of charge carriers in the section of wire of length l. Now, $N = nV$, where n is the number of charge carriers per unit volume and V is the volume of wire in the field. Noting that $V = Al$, where A is the cross-sectional area of the wire, then the force on the wire is $F = (qv_dB\sin\theta)(nAl)$. Gathering terms,

$$F = (nqAv_d)lB\sin\theta$$

Because $nqAv_d = I$ (see Current),

$$F = IlB\sin\theta$$

is the equation for *magnetic force on a length l of wire carrying a current I in a uniform magnetic field B*, as shown in Figure 2. If we divide both sides of this expression by l, we find that the magnetic force per unit length of wire in a uniform field is $\frac{F}{l} = IB\sin\theta$. The direction of this force is given by RHR-1, with the thumb in the direction of the current I. Then, with the fingers in the direction of B, a perpendicular to the palm points in the direction of F, as in Figure 2.

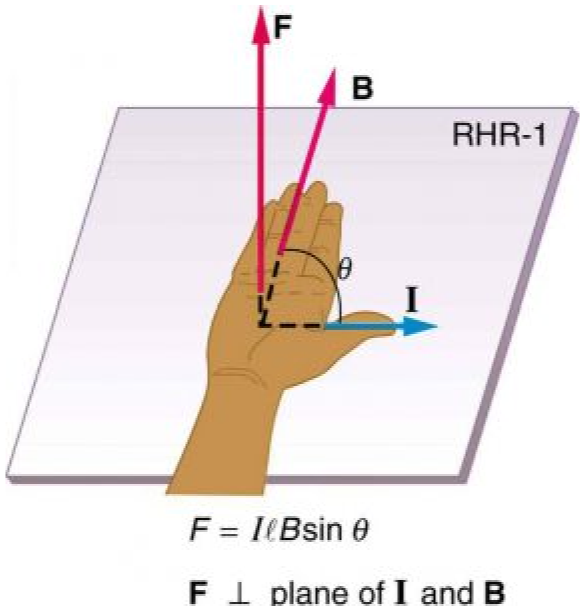


***Figure 2.** The force on a current-carrying wire in a magnetic field is $F = IlB\sin\theta$. Its direction is given by RHR-1.*

Calculating Magnetic Force on a Current-Carrying Wire: A Strong Magnetic Field

Calculate the force on the wire shown in Figure 1, given $B = 1.50\ \mathrm{T}, l = 5.00\ \mathrm{cm}$ and $I = 20.0\ \mathrm{A}$.

Strategy

The force can be found with the given information by using $F = IlB\sin\theta$ and noting that the angle θ between I and B is 90°, so that $\sin\theta$.

Solution

Entering the given values into $F = IlB\sin\theta$ yields

$$F = IlB\sin\theta = (20.0\ \text{A})(0.0500\ \text{m})(1.50\ \text{T})(1).$$

The units for tesla are $1\ \text{T} = \frac{N}{\text{A}\cdot\text{m}}$; thus,

$$F = 1.50\ \text{N}.$$

Discussion

This large magnetic field creates a significant force on a small length of wire.

Magnetic force on current-carrying conductors is used to convert electric energy to work. (Motors are a prime example—they employ loops of wire and are considered in the next section.) Magnetohydrodynamics (MHD) is the technical name given to a clever application where magnetic force pumps fluids without moving mechanical parts. (See Figure 3.)

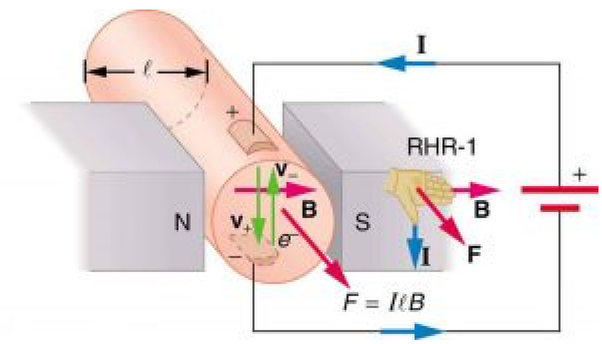


Figure 3. *Magnetohydrodynamics. The magnetic force on the current passed through this fluid can be used as a nonmechanical pump.*

A strong magnetic field is applied across a tube and a current is passed through the fluid at right angles to the field, resulting in a force on the fluid parallel to the tube axis as shown. The absence of moving parts makes this attractive for moving a hot, chemically active substance, such as the liquid sodium employed in some nuclear reactors. Experimental artificial hearts are testing with this technique for pumping blood, perhaps circumventing the adverse effects of mechanical pumps. (Cell membranes, however, are affected by the large fields needed in MHD, delaying its practical application in humans.) MHD propulsion for nuclear submarines has been proposed, because it could be considerably quieter than conventional propeller drives. The deterrent value of nuclear submarines is based on their ability to hide and survive a first or second nuclear strike. As we slowly disassemble our nuclear weapons arsenals, the submarine branch will be the last to be decommissioned because of this ability (See Figure 4.) Existing MHD drives are heavy and inefficient—much development work is needed.

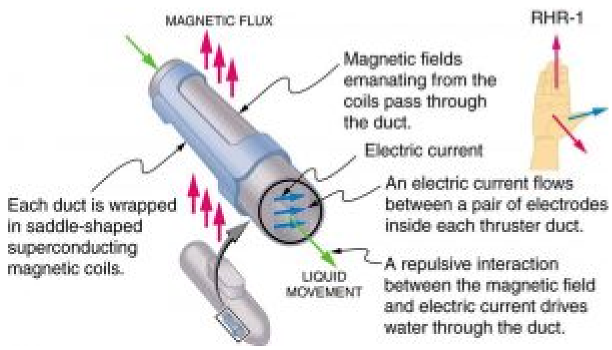


Figure 4. *An MHD propulsion system in a nuclear submarine could produce significantly less turbulence than propellers and allow it to run more silently. The development of a silent drive submarine was dramatized in the book and the film The Hunt for Red October.*

Section Summary

- The magnetic force on current-carrying conductors is given by

$$F = IlB\sin\theta,$$

where I is the current, l is the length of a straight conductor in a uniform magnetic field B, and θ is the angle between I and B. The force follows RHR-1 with the thumb in the direction of I.

Homework

10. When is the force on a current carrying wire the strongest?
11. Determine the field from the force on a wire.
12. Determine the angle between a wire and the field.

33. Sources of Magnetic Fields

Magnetic Fields Produced by Currents: Ampere's Law

Derived from [Magnetic Fields Produced by Currents: Ampere's Law]() by OpenStax

Instructor's Note

As before, I am NOT expecting you to fully understand everything in this section. I hope you:

- Are reminded that moving charges (i.e. currents) are ultimately the source of all magnetic fields
- Are reminded that magnetic field lines travel in closed loops
- How to calculate the magnitude of the magnetic field some distance from a straight wire

How much current is needed to produce a significant magnetic field, perhaps as strong as the Earth's field? Surveyors will tell you that overhead electric power lines create magnetic fields that interfere with their compass readings. Indeed, when Oersted discovered in 1820 that a current in a wire affected a compass needle, he was not dealing with extremely large currents. How does the shape of wires carrying current affect the shape of the magnetic field created? We noted earlier that a current loop created a magnetic field similar to that of a bar magnet, but what about a straight wire or a toroid (doughnut)? How is the direction of a current-created field related to the direction of the current? Answers to these questions are explored in this section, together with a brief discussion of the law governing the fields created by currents.

Magnetic Field Created by a Long Straight Current-Carrying Wire: Right Hand Rule 2

Magnetic fields have both direction and magnitude. As noted before, one way to explore the direction of a magnetic field is with compasses, as shown for a long straight current-carrying wire in Figure 1. Hall probes can determine the magnitude of the field. The field around a long straight wire is found to be in circular loops. The **right hand rule 2** (RHR-2) emerges from this exploration and is valid for any current segment—*point the thumb in the direction of the current, and the fingers curl in the direction of the magnetic field loops* created by it.

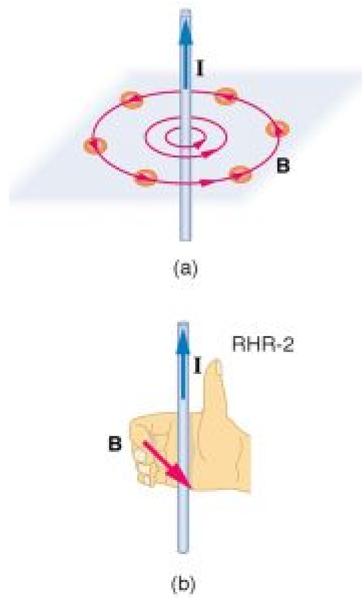


Figure 1. *(a) Compasses placed near a long straight current-carrying wire indicate that field lines form circular loops centered on the wire. (b) Right hand rule 2 states that, if the right hand thumb points in the direction of the current, the fingers curl in the direction of the field. This rule is consistent with the field mapped for the long straight wire and is valid for any current segment.*

The **magnetic field strength (magnitude) produced by a long straight current-carrying wire** is found by experiment to be

$$B = \frac{\mu_0 I}{2\pi r}(\text{long straight wire})$$

University of Massachusetts Amherst BE REVOLUTIONARY™

Instructor's Note

I would like you to be able to use the equation above to get a numerical value.

where I is the current, r is the shortest distance to the wire, and the constant $\mu_0 = 4\pi \times 10^{-7}\ \mathrm{T \cdot\ m/A}$ is the **permeability of free space.** (μ_0 one of the basic constants in nature. We will see later that μ_0 is related to the speed of light.) Since the wire is very long, the magnitude of the field depends only on distance from the wire r, not on position along the wire.

Calculating Current that Produces a Magnetic Field

Find the current in a long straight wire that would produce a magnetic field twice the strength of the Earth's at a distance of 5.0 cm from the wire.

Strategy

The Earth's field is about $5.0 \times 10^{-5}\ \mathrm{T}$ and so here B due to the wire is taken to be $1.0 \times 10^{-4}\ \mathrm{T}$. The equation $B = \frac{\mu_0 I}{2\pi r}$ can be used to find I, since all other quantities are known.

Solution

Solving for I and entering known values gives

$$I = \frac{2\pi r B}{/mu_0} = \frac{2\pi(5.0\times10^{-2}\ \mathrm{m})(1.0\times10^{-4}\ \mathrm{T})}{4\pi\times10^{-7}\ \mathrm{T{\cdot}m/A}}$$

$$25\ \mathrm{A}$$

Discussion

So a moderately large current produces a significant magnetic field at a distance of 5.0 cm from a long straight wire. Note that the answer is stated to only two digits, since the Earth's field is specified to only two digits in this example.

Ampere's Law and Others

The magnetic field of a long straight wire has more implications than you might at first suspect. *Each segment of current produces a magnetic field like that of a long straight wire, and the total field of any shape current is the vector sum of the fields due to each segment.* The formal statement of the direction and magnitude of the field due to each segment is called the **Biot-Savart law**. Integral calculus is needed to sum the field for an arbitrary shape current. This results in a more complete law, called **Ampere's law**, which relates magnetic field and current in a general way. Ampere's law in turn is a part of **Maxwell's equations**, which give a complete theory of all electromagnetic phenomena. Considerations of how Maxwell's equations appear to different observers led to the modern theory of relativity, and the realization that electric and magnetic fields are different manifestations of the same thing. Most of this is beyond the scope of this text in both mathematical level, requiring calculus, and in the amount of space that can be devoted to it. But for the interested student, and particularly for those who continue in physics, engineering, or similar pursuits, delving into these matters further will reveal descriptions of nature that are elegant as well as profound. In this text, we shall keep the general features in mind, such as RHR-2 and the rules for magnetic field lines listed in Magnetic Fields and Magnetic Field Lines, while concentrating on the fields created in certain important situations.

Making Connections: Relativity

Hearing all we do about Einstein, we sometimes get the impression that he invented relativity out of nothing. On the contrary, one of Einstein's motivations was to solve difficulties in knowing how different observers see magnetic and electric fields.

Section Summary

- The strength of the magnetic field created by current in a long straight wire is given by

$$B = \frac{\mu_0 I}{2\pi r} (\text{long straight wire}),$$

where I is the current, r is the shortest distance to the wire, and the constant $\mu_0 = 4\pi \times 10^{-7}\ \text{T} \cdot \text{m/A}$ is the permeability of free space.

- The direction of the magnetic field created by a long straight wire is given by right hand rule 2 (RHR-2): *Point the thumb of the right hand in the direction of current, and the fingers curl in the direction of the magnetic field loops* created by it.
- The magnetic field created by current following any path is the sum (or integral) of the fields due to

segments along the path (magnitude and direction as for a straight wire), resulting in a general relationship between current and field known as Ampere's law.

Homework

13. Magnetic fields from power lines.
14. Magnetic fields in your car.

34. Homework Problems

Homework

The list below is the list of homework problems in Edfinity. The numbering is the same. You can click on a problem, and it will take you to the relevant section of the book!

1. Sources of magnetism
2. Compasses and magnetic field lines.
3. Units of magnetic field.
4. What will feel a magnetic force.
5. What will feel a magnetic force – macroscopic objects.
6. Thinking about the magnetic forces on charged particles.
7. Magnetic force on an airplane.
8. Magnetic force on a baseball.
9. Magnetic force on an electron at an angle to $\vec{B}$.
10. When is the force on a current carrying wire the strongest?
11. Determine the field from the force on a wire.
12. Determine the angle between a wire and the field.
13. Magnetic fields from power lines.
14. Magnetic fields in your car.

Glossary

accommodation

the ability of the eye to adjust its focal length is known as accommodation

amplitude

The size of the wave. For a physical wave like a water wave, this will be the actual height in meters. For a sound wave (a pressure wave in the air) this will be in units of pressure Pa.

analytical methods

the method of determining the magnitude and direction of a resultant vector using the Pythagorean theorem and trigonometric identities

antimatter

For each type of particle in the universe, there is an antimatter counterpart. These antimatter counterparts have the same mass as the usual particles, but the sign of the electric charge is reversed: anti-electrons are positively charged, while anti-protons are negatively charged. When matter and antimatter are brought into contact, the result is their mutual destruction (or annihilation) into pure energy.

capacitor

a device that stores electric charge

center (optics)

For a lens, where the lens is thickest.
For a mirror, the center is the center of curvature.

circadian

describes a time cycle about one day in length

commutative

refers to the interchangeability of order in a function; vector addition is commutative because the order in which vectors are added together does not affect the final sum

components

a piece of a vector that points in either the vertical or the horizontal direction; every 2-d vector can be expressed as a sum of two vertical and horizontal vector components

concave mirror

A mirror that bends towards the incoming light)

cones

A set of weakly photosensitive, cone-shaped neurons in the fovea of the retina that detects bright light and is used in daytime color vision. There are cones responsible for red light, green light, and blue light.

converges (optics)

Bring light rays to a point.

converging (or convex) lens

a convex lens in which light rays that enter it parallel to its axis converge at a single point on the opposite side

convex mirror

A mirror that bends away from the light source (

cornea

The transparent layer over the front of the eye that helps focus light waves. Most of the focusing of the eye actually happens at the cornea, not in the lens.

definitions

An equation representing a common quantity. This equation does not elucidate a fundamental truth of the Universe, it just defines an idea. For example, velocity = distance / time. There are no fundamental truths here, that is just the definition of velocity.

direction

the orientation of a vector in space

diverge (optics)

Spread the light rays apart.

diverging lens

a concave lens in which light rays that enter it parallel to its axis bend away (diverge) from its axis

energy

From physics 131: The ability of an object to do work. I.e. its capability to exert a force for a distance. Whether that ability is realized is not relevant.

Energy comes, ultimately, in only two types: kinetic and potential. Kinetic energy is the capability to do work due to motion; thermal energy due to temperature, is at a fundamental level, kinetic energy due to molecular motion. Potential energy is the energy due to the relative positions of two objects: gravitational potential energy arises from the relative positions of an object and the Earth. "What is the gravitational potential energy of the ball?" is, technically, a meaningless question. The question only has relevance when considered in conjunction with the fact that the Earth exists.

erect

When the image is the same orientation as the object

far point

The furthest an object can be from an eye and still be seen clearly.

focal length

distance from the center of a lens or curved mirror to its focal point

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

fovea centralis

region in the center of the retina with a high density of photoreceptors and which is responsible for acute vision

frequency

The number of wave crests passing a point per second. The unit is 1/s or, equivalently, Hertz Hz.
The frequency will be 1 divided by the period T.

geometric optics

part of optics dealing with the ray aspect of light

head (or tip)

the end point of a vector; the location of the tip of the vector's arrowhead; also referred to as the "tip"

head-to-tail method

a method of adding vectors in which the tail of each vector is placed at the head of the previous vector

heat

The transfer of energy through microscopic collisions: fast moving (high-temperature) atoms colliding with slow moving (low-temperature) atoms results in the movement of energy from hot-to-cold.
Relevant to this course, the collisions could also be with photons.

hyperopia (also, farsightedness)

Visual defect in which the image focus falls behind the retina, thereby making images in the distance clear, but close-up images blurry.

image

The apparent reproduction of an object, formed by an optical element (or collection of them) reflecting and/or refracting light.

image distance

the distance of the image from the center of a lens

incident ray

Incoming ray

index of refraction

for a material, the ratio of the speed of light in vacuum to that in the material [latex] n = c/v [/latex]. Always greater than 1.

intensity

Power per area:
I = P/A
or, using P = E/t,
I = E/(At)

inverted

When the image is upside-down with respect to the object.

ionizing radiation

radiation that ionizes materials that absorb it

iris

The pigmented, circular muscle at the front of the eye that regulates the amount of light entering the eye.

Law of Reflection

The angle of reflection equals the angle of incidence.

Law of Refraction (Snell's Law)

n_1 sin (theta_1) = n_2 sin (theta_2)

lens (eye)

The transparent, convex structure behind the cornea that helps focus light waves on the retina. The lens is for the fine-tuning.

magnification

ratio of image height to object height

magnitude

the length or size of a vector; magnitude is a scalar quantity

metaphysics

Metaphysics is the branch of philosophy that examines the fundamental nature of reality, including the relationship[1] between mind and matter, between substance and attribute, and between potentiality and actuality.[2] The word "metaphysics" comes from two Greek words that, together, literally mean "after or behind or among [the study of] the natural" -Wikipedia

mirror

A smooth surface that reflects light at specific angles, forming an image of the person or object in front of it

myopia (also, nearsightedness)

Visual defect in which the image focus falls in front of the retina, thereby making images in the distance blurry, but close-up images clear

near point

The closest an object can be to the eye and still be seen clearly.

object distance

the distance of an object from the center of a lens

ontology

Ontology is the philosophical study of being. More broadly, it studies concepts that directly relate to being, in particular becoming, existence, reality, as well as the basic categories of being and their relations. Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology often deals with questions concerning what entities exist or may be said to exist and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences. -Wikipedia

optical axis

an imaginary line that passes through the optical element in a way that's perpendicular to it

optical element

any lens or mirror

particle

The simplest conception is a ball. Particles have a fixed position and speed. Particles are characterized, by their energy, momentum, and how many of them there are.

photons

Particles of light. For a given frequency [latex] \nu [/latex], the smallest amount of energy that you can have is one photon's worth: [latex] E_\gamma = h \nu [/latex].

pi bond

a type of covalent bond that results from the side-by-side overlap of two p orbitals

power

The energy or work per time. The unit is the J/s or the Watt (W)

power (optics)

inverse of focal length

presbyopia

a condition in which the lens of the eye becomes progressively unable to focus on objects close to the viewer

principles

A fundamental relationship that describes how the Universe works. These are the fundamental truths of Nature. When writing a principle as an equation, the "=" is translated as "causes." For example, Newton's 2nd Law, F = ma, a force F causes an object m to accelerate (change its speed or direction) a. These principles are where we begin our analyses.

pupil

The small opening at the front of the eye though which light enters. Appears black (or red in flash photographs!). The size is controlled by the iris.

ray

straight line that originates at some point

real image

image that can be projected

refracted ray

A ray that has been bent by a refraction, such as in a lens.

refraction

changing of a light ray’s direction when it passes through variations in matter

resultant

the sum of two or more vectors

resultant vector

the vector sum of two or more vectors

retina

layer of photoreceptive and supporting cells on the inner surface of the back of the eye

rhodopsin

main photopigment in vertebrates

rods

Strongly photosensitive, achromatic, cylindrical neuron in the outer edges of the retina that detects dim light and is used in peripheral and nighttime vision. Can only see in black and white.

scalar

a quantity with magnitude but no direction

sigma bond

a covalent bond in which the electron density is concentrated in the region along the internuclear axis; that is, a line between the nuclei would pass through the center of the overlap region. Single bonds in Lewis structures are described as σ bonds in valence bond theory.

superior colliculus

paired structure in the top of the midbrain, which manages eye movements and auditory integration

suprachiasmatic nucleus

cluster of cells in the hypothalamus that plays a role in the circadian cycle

tail

the start point of a vector; opposite to the head or tip of the arrow

thin lens

a lens whose thickness allows rays to refract but does not allow properties such as dispersion and aberrations.

tonic activity

in a neuron, slight continuous activity while at rest

valence bond theory

Describes a covalent bond as the overlap of half-filled atomic orbitals (each containing a single electron) that yield a pair of electrons shared between the two bonded atoms.

vectors

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

vertex

The point where the optical axis meets the optical element.

vertex (optics)

The point where the optical axis meets the optical element

virtual image

image that cannot be projected

vision

sense of sight

wave-particle duality

A description for the fundamentally new nature of very small objects like electrons and photons: sometimes they behave like waves and sometimes they behave like particles. Neither picture is 100% correct: electrons are neither waves nor particles, but have properties of both.

In fact, all objects exhibit wave particle duality. You have a wavelength! However, your wavelength is too small to notice (check with de Broglie if you want). The effect is really only noticeable for small objects.

wavelength

The distance from one point in a wave to the same point on the next wave: for example, crest-to-crest. This is a distance measured in meters.

work

The exchange of energy through the application of a force through some distance.