Essential MODERN PHYSICS

Study Guide Workbook

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

$$E^2 = p^2c^2 + m_0^2c^4$$

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

$$\mathcal{E}_n = nhf \qquad p = \frac{h}{\lambda}$$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

with Full Solutions!

Chris McMullen, Ph.D.

ESSENTIAL MODERN PHYSICS Study Guide Workbook

(with Full Solutions)

Chris McMullen, Ph.D.

Essential

MODERN PHYSICS

Study Guide Workbook

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

$$E^2 = p^2c^2 + m_0^2c^4$$

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

$$\mathcal{E}_n = nhf \qquad p = \frac{h}{\lambda}$$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

with Full Solutions!

Chris McMullen, Ph.D.

Essential Modern Physics
Study Guide Workbook
Chris McMullen, Ph.D.

monkeyphysicsblog.wordpress.com
improveyourmathfluency.com
chrismcmullen.com

Zishka Publishing
ISBN: 978-1-941691-28-1
Textbooks > Science > Physics
Study Guides > Workbooks> Science

CONTENTS

INTRODUCTION

The goal of this study guide workbook is to provide practice and help carrying out essential problem-solving strategies that are standard in modern physics. The aim here is not to overwhelm the student with comprehensive coverage of every type of problem, but to focus on the main strategies and techniques with which most physics students struggle.

This workbook is not intended to serve as a substitute for lectures or for a textbook, but is rather intended to serve as a valuable supplement. Each chapter includes a concise review of the essential information, a handy outline of the problem-solving strategies, and examples which show step-by-step how to carry out the procedure. This is not intended to teach the material, but is designed to serve as a time-saving review for students who have already been exposed to the material in class or in a textbook. Students who would like more examples or a more thorough introduction to the material should review their lecture notes or read their textbooks.

Every problem in this study guide workbook applies the same strategy which is solved step-by-step in at least one example within the chapter. Study the examples and then follow them closely in order to complete the exercises. Many of the exercises are broken down into parts to help guide the student through the exercises. Each exercise tabulates the corresponding answers on the same page. Students can find full solutions at the end of each chapter.

The prerequisites for using this workbook include first-year physics (including energy, waves, and electricity and magnetism) and calculus (including derivatives and integrals). Although Schrödinger's equation in quantum mechanics is a differential equation, students do not need previous exposure to differential equations: This workbook provides a concise introduction to basic differential equations in Chapter 12, and shows how to apply these methods in the remaining chapters.

1 SPECIAL RELATIVITY CONCEPTS

Relevant Terminology

Galilean relativity – our experience with the relative motion of objects traveling at speeds much slower than the speed of light.

Special relativity – the physics of the relative motion of objects where at least one object is traveling at a very high speed (compared to the speed of light).

Ether – a hypothetical substance once believed to fill space; it was believed to serve as a medium for the transmission of light waves.

Photon – a single particle in a beam of light.

Interferometer – a device involving the interference of two beams of light, which was used by Michelson and Morley to measure the speed of light relative to the earth.

Time dilation – the phenomenon whereby time appears to travel more slowly for objects moving fast (close to light speed) relative to other observers.

Length contraction – the phenomenon whereby objects moving fast (close to light speed) appear shorter relative to other observers.

Simultaneity – when two events occur at the exact same moment relative to an observer, the events are said to be simultaneous for that observer.

Momentum – mass times velocity.

Inertia – the natural tendency of any object to maintain constant momentum.

Mass – a measure of inertia.

Vacuum – a region of space completely devoid of matter (it doesn’t even contain air).

Inertial reference frame – a frame that travels with constant velocity.

Galilean Relativity

In our everyday experience with objects that travel much slower than the speed of light, relative velocities appear to obey the formula for **vector addition**. Suppose that one observer (designated R) is at rest while a second observer (designated M) is moving with speed v relative to the first observer. Suppose also that each observer sets up a coordinate system with the x-axis along the relative velocity

$$\vec{v}$$

If each observer measures the velocity of an object, the x-components of the velocities that they measure (u_R and u_M) are related by the following vector addition formula, provided that neither observer nor the object are traveling close to light speed.

$$u_M = u_R - v$$

This equation is actually pretty simple: It's just subtraction. The challenge is to remember what the notation means (M stands for moving, while R stands for rest) so that you can apply the equation correctly. We will explore this equation further in the examples that follow.

The Mysterious Ether

Physicists once believed in a hypothetical substance called the **ether**, which was believed to permeate all of space. At the time, all other waves besides light were known to travel in a medium. You can see ripples travel along the surface of water. Sound waves create alternating regions of compression (high pressure) and rarefaction (low pressure) in a medium such as air, water, wood, or metal. Light was also known to be a wave, yet sunlight can travel through space (a near-perfect vacuum). Since all other waves required a medium in which to propagate, the concept of the ether could explain the transmission of light through space.

It turns out that the ether hypothesis is **incorrect**, as demonstrated by the Michelson-Morley experiment. Light can travel through a perfect **vacuum** (without an ether).

The Earth, Light, and the Hypothetical Ether

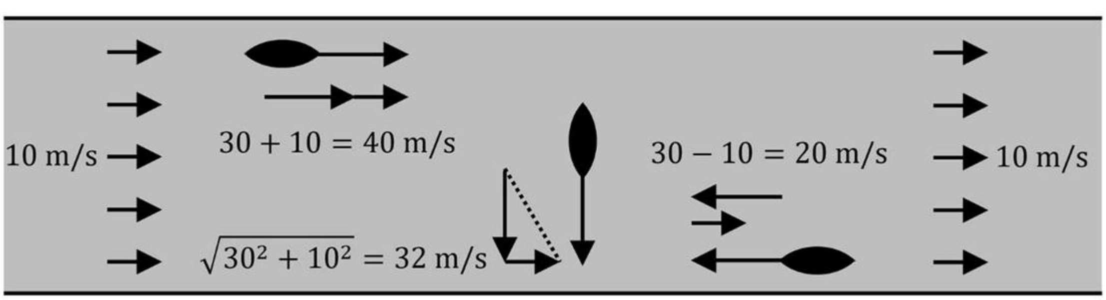
The result of the Michelson-Morley experiment—that **the ether doesn't exist**—came as a big shock to the physics community at the time. To understand why, we must explore the ether as it had been believed to exist. The ether was believed to permeate all of space. The reference frame of the ether was believed to serve as an absolute reference frame. That is, the speed of light was only believed to travel $c = 2.9979 \times 10^8$ m/s in a reference frame that was at rest relative to the ether. Furthermore, it was believed that the vector addition equation of Galilean relativity applied to objects moving any speed, including light itself. (Like the ether, this also proved to be **incorrect**.)

The earth orbits the sun and therefore must be moving relative to the hypothetical ether (as the direction of earth's velocity is constantly changing). From the point of view of the ether, the earth is moving relative to the ether with instantaneous speed v. From the point of view of earthlings, the earth seems to be stationary and we would instead interpret the ether to be moving with speed v. (You should have experience with this. If you are sitting in a bus that is moving, objects outside of the bus appear to be moving relative to you.) From the perspective of "stationary" earthlings, the ether is seen as an "**ether wind**" (when you run through air that is originally still, the air seems to pass by you like a sort of wind).

Imagine that you get in a motorboat and travel along the surface of a river. The boat would travel with a speed of 30 m/s on a still pond, but there is a river current of 10 m/s. When the boat is headed downstream, it would be traveling 30 + 10 = 40 m/s relative to the land. When the boat is headed upstream, it would be traveling 30 – 10 = 20 m/s relative to the land. When the boat is headed cross-stream, apply the Pythagorean theorem to determine that the boat travels

$$\sqrt{30^2 + 10^2} = \sqrt{1000} \approx 32 \text{ m/s}$$

relative to the land.



The same principle as the motorboat and river was believed to apply to the earth traveling through the hypothetical ether. Imagine that you shine a beam of light from earth and wish to measure the speed of light. According to the ether hypothesis, the light would travel with speed $c = 2.9979 \times 10^8$ m/s relative to the ether. The earth travels with speed v = 30,000 m/s relative to the sun. As the earth orbits the sun, the direction of its velocity constantly changes. The speed of light would equal $c + v$ when the light is heading "downwind" (when the earth happens to be traveling along the ether wind), the speed of light would equal $c - v$ when the light is heading "upwind," and the speed of light would equal

$$\sqrt{c^2 + v^2}$$

when the light is heading “across wind.” However, when Michelson and Morley investigated this, **no such changes in the speed of light were detected**. The laws of Galilean relativity do **not** apply to light or to objects moving fast (close to light speed). **The ether does not exist.** The speed of light turned out to be a universal **constant** (independent of the motion of the observer or source).

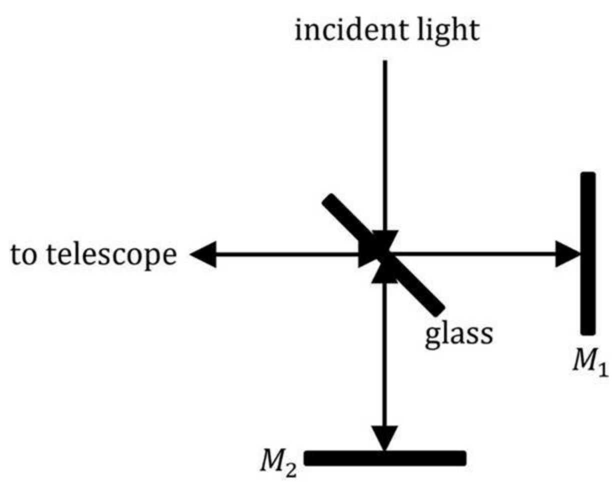
Why $c + v$, $c - v$, and $\sqrt{c^2 + v^2}$ Seemed to Make Sense

Imagine a father and son playing catch on a boat that is at rest. They throw the ball with a speed of 20 m/s east and west. They continue to play catch the same way even when the boat starts traveling 50 m/s to the east. When the ball travels 20 m/s east relative to the boat, it is traveling 50 + 20 = 70 m/s to the east relative to the land. When the ball travels 20 m/s to the west relative to the boat, it is traveling 50 – 20 = 30 m/s to the east relative to the land. This is the way relative velocities work at speeds that are small compared to the speed of light. It agrees with everyday experience.

Now imagine a spaceship traveling $0.7c$ (or 70% the speed of light) relative to earth. The spaceship suddenly turns on its headlights. How fast is the beam of light traveling relative to earth? Based on our experience with low speeds (Galilean relativity), it's intuitive to expect the beam of light to be traveling $0.7c + c = 1.7c$ relative to the earth. But that's not what happens. It turns out that the beam of light travels $c = 2.9979 \times 10^8$ m/s relative to the ship and that the beam of light also travels $c = 2.9979 \times 10^8$ m/s relative to earth, with no contradiction! Although $c + v$ seemed to make sense, **it turned out to contradict experiments**.

The Michelson-Morley Experiment

Michelson and Morley used an **interferometer** that was designed to detect small changes in the speed of light. When an incident beam of light reached a glass slab, it split into two: One beam reflected from the surface of the glass towards mirror M_1, and a second beam refracted through the glass towards mirror M_2. After reflecting off mirrors M_1 and M_2, the beams met back up at the glass, and light from each beam reached a telescope, where an interference pattern was viewed. The entire apparatus could be rotated.



The two beams would travel in different directions relative to the ether, and thus the two beams would have different speeds relative to the earth if the ether hypothesis were true. For example, if the beam heading toward M_1 were traveling **upwind** with speed $c - v$ (in which case it would then travel **downwind** with speed $c + v$ after reflection), then the beam heading toward M_1 would be traveling **crosswind** with speed

$$\sqrt{c^2 + v^2}$$

Thus, the two beams would return to the glass at different times, creating an interference pattern when viewed through the telescope (since the beams would be slightly out of phase due to the time lag).

However, the Michelson-Morley experiment **failed to detect any change in the speed of light** as the apparatus was rotated. It turns out that light (which is an electromagnetic wave) does not require a medium (such as the ether) in order to propagate. Light can travel through a perfect **vacuum**. There is **no** preferred or absolute reference frame for measuring the speed of light. It turns out that the speed of light is a universal **constant**, independent of the motion of the source or the observer.

The Problem

The Michelson-Morley experiment contradicted the expectations of the ether hypothesis. How could the speed of light be the same in each beam of the interferometer, regardless of the orientation of the apparatus?

If a boat is traveling north, it can launch a cannonball farther to the north (relative to the land) than it could if the boat were at rest. This is the principle of vector addition applied in Galilean relativity, which agrees with human experience (with objects traveling much slower than the speed of light).

Imagine a spaceship traveling one-half the speed of light ($0.5c$) relative to earth. Also imagine a beam of light traveling parallel to the spaceship. An observer on earth measures the speed of the beam of light to be $c = 2.9979 \times 10^8$ m/s. What will an observer on the spaceship measure the speed of the beam of light to be? According to the Michelson-Morley experiment, the answer is the **same**: $c = 2.9979 \times 10^8$ m/s. (Note that the answer isn't $0.5c$.)

Your experience with relative motion at low speeds is much different. If you're sitting in a bus traveling 40 m/s and a car passes you traveling 50 m/s, each second the car gets 10 m further ahead of the bus, so the car seems to be traveling 10 m/s relative to you. However, if you're in a spaceship and proceed to measure the speed of light, the Michelson-Morley experiment shows that you will get $c = 2.9979 \times 10^8$ m/s regardless of how fast the ship is traveling. Even if the spaceship travels $0.99c$ relative to earth, an observer on the ship still measures the speed of light to be $c = 2.9979 \times 10^8$ m/s (and not $0.01c$).

Albert Einstein introduced his theory of special relativity to resolve this seeming paradox, but it came with some interesting consequences: **time dilation** and **length contraction**. The underlying issue is that time, space, and light behave much differently than our everyday experience with low-speed motion suggests. The two different observers (in the spaceship and on earth) actually measure distance and time differently due to time dilation and length contraction. If there is a meterstick on the spaceship and an observer on earth proceeds to measure the length of that meterstick as the spaceship travels very fast (close to light speed), the observer on earth measures the meterstick to be significantly less than one meter long. Similarly, if there is a pendulum on the spaceship that oscillates with a period of exactly one second relative to an observer on the spaceship, an observer on earth measures the period of that same pendulum to last significantly longer than one second. When the relative speed between two observers is in the neighborhood of the speed of light, the observers significantly disagree on such basic notions as what meters and seconds are! As bizarre as this may seem, Einstein's theory of special relativity not only explains the Michelson-Morley experiment, it also agrees with countless other scientific tests.

Einstein's Theory of Special Relativity

The theory of special relativity applies to objects that move with **constant velocity** (meaning that they move with constant speed and also travel in a straight line). Einstein developed his theory of special relativity from two fundamental postulates:

1. The laws of physics are the same in any **inertial reference frame**. (An inertial reference frame is any coordinate system that has constant velocity. Recall that constant velocity means both constant speed and traveling in a straight line.) This means that any physics experiment will yield the same results in any laboratory that travels with constant velocity, whether the laboratory is at rest on earth or moving in a spaceship with a velocity of $0.8c$ relative to the earth.

2. *Any* observer in an inertial reference frame would measure the **speed of light** to be $c = 2.9979 \times 10^8$ m/s, regardless of the velocity of the observer and also regardless of the velocity of the light source.

One consequence of special relativity is that there isn't any preferred (or absolute) reference frame (such as an ether). Any inertial reference frame is equally as good as any other.

What does it mean to be at **rest**? You can be at rest relative to earth, but earth is revolving around the sun. Even if an object is at rest relative to the sun, the sun is traveling through space relative to other stars. Since the laws of physics are the same in all inertial reference frames, it would be impossible to find a particular reference frame that you could say is truly at "rest." As long as you're moving with constant velocity, you're entitled to consider yourself to be at "rest" and to consider everything else moving relative to you. Even if you're in a bus that is traveling 30 m/s west past a station, you may consider yourself to be at rest and may consider the station as traveling 30 m/s to the east (opposite to you). A woman at the station may consider herself to be at rest and consider you to be traveling 30 m/s to the west. You are both entitled to be correct when working out physics with your own **inertial reference frames**. According to Einstein, **it's all relative**.

As we will explore mathematically in Chapters 2-5, the theory of special relativity comes with a few seemingly strange consequences. In particular, even seemingly fundamental concepts like length and time are relative, and different observers in different inertial reference frames measure length and time differently:

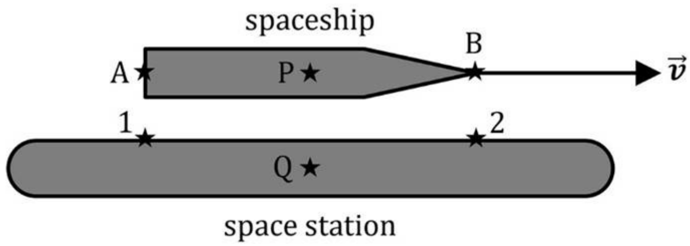
- Two events that appear to occur **simultaneously** in one inertial reference frame may appear to occur at different times in another inertial reference frame.

- **Length contraction**: When an object is moving relative to an inertial reference frame, the object appears shorter (along the direction of motion) than it does relative to an inertial reference frame that is at rest relative to the object.

- **Time dilation**: Time passes more slowly on a clock in a moving inertial reference frame than it does for an inertial reference frame that is at rest relative to the clock.

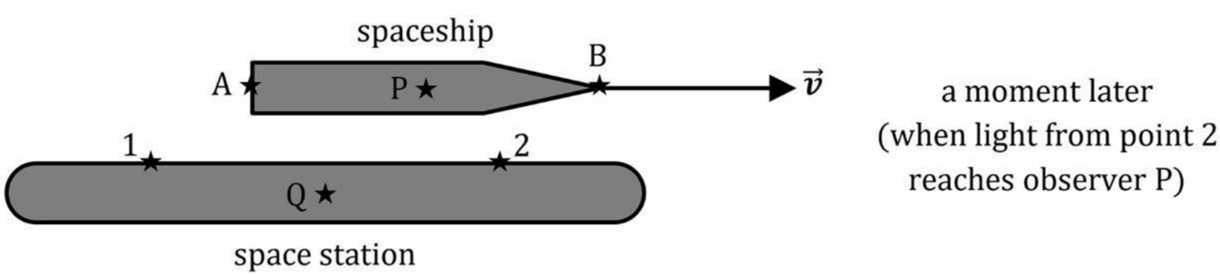
Simultaneity Is Relative

One consequence of the postulates of special relativity is that whether or not two events appear to occur simultaneously (at the same time) depends on the inertial reference frame from which the events are observed. If two events are observed to occur simultaneously in one inertial reference frame, they may not be observed to occur simultaneously in another inertial reference frame.



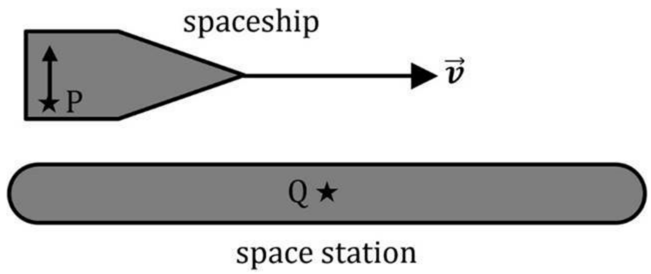
For example, consider the spaceship traveling close to the speed of light in the illustration above. The spaceship passes very close to a space station that is orbiting the earth. The space station is traveling so much slower than the speed of light that it is practically at rest relative to the very fast spaceship. One astronaut is stationed at point Q on the space station, while another astronaut is stationed at point P on the spaceship. There are lights on the space station at the points labeled 1 and 2. These two lights remain off almost all of the time. The lights are programmed to flash quickly at the exact instant that the two ends of the spaceship happen to be (momentarily) positioned directly above the lights. At this exact instant, point A is directly above point 1, point B is directly above point 2, and points P and Q are exactly midway between points 1 and 2 (and are thus midway between points A and B).

Observer Q is at rest relative to the space station and is exactly midway between points 1 and 2, such that the light emitted by each point during the flash travels the same distance to reach observer Q. Thus, observer Q on the space station sees the two lights flash simultaneously. In contrast, observer P is traveling close to the speed of light relative to points 1 and 2. Light from point 2 reaches observer P before light from point 1. Observer P doesn't see the two lights flash simultaneously. Which observer is correct? According to special relativity, **both are correct**. Whether you are practically at rest (like the space station) or traveling close to the speed of light relative to the earth (like the spaceship), the laws of physics are the same. There isn't a preferred inertial reference frame that makes observations more "correct."



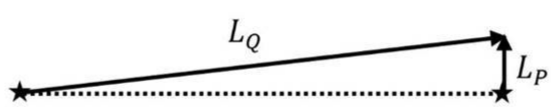
Time Dilation

Time passes more slowly on a clock in a moving inertial reference frame than it does for an inertial reference frame that is at rest relative to the clock. This is known as **time dilation**.



You can see how time dilation is a direct consequence of the postulates of special relativity by considering the illustration above. The spaceship is traveling to the right with a speed that is close to the speed of light. The spaceship passes by a space station that is orbiting the earth. The space station is traveling so much slower than the speed of light that it is practically at rest relative to the very fast spaceship. As the spaceship is passing the space station, an astronaut inside of the spaceship turns on a flashlight, shining a beam of light straight upward in the diagram (perpendicular to the direction that the spaceship is traveling). The spaceship has transparent walls such that an observer inside of the space station can see the beam of the flashlight shining inside of the spaceship as it passes the space station.

Relative to the astronaut inside of the spaceship (observer P), the flashlight beam appears to travel straight upward, as shown in the left diagram above. Relative to an observer inside of the space station (observer Q), the flashlight beam appears to travel diagonally up and to the right, as shown in the right diagram above. (Of course, the photons in the beam of light have **inertia**, which is the natural tendency of all objects to travel with constant momentum. If you ride in an airplane traveling 500 mph and throw a ball straight upwards, you will catch the ball in your hand because of inertia. The ball certainly won't smack the back of the airplane mid-flight. If you've forgotten about inertia, it may help to review an introductory physics textbook.)



Observer P sees the beam of light take a shorter path (distance L_P, which is straight upward), whereas observer Q sees the beam of light take a longer path (distance L_Q, which is diagonal). According to the second postulate of special relativity, both observers must measure the **speed of light** to be the same value. Either observer takes the distance (L) traveled and divides by the corresponding time (t) measured to determine the speed of light. The subscripts P and Q indicate which observer makes the measurement.

$$c = \frac{L_P}{t_P} \quad , \quad c = \frac{L_Q}{t_Q}$$

Since $L_Q > L_P$, in order for both observers to measure the same value for the speed of light, it follows that $t_Q > t_P$. Since observer P is at rest relative to the flashlight, while observer Q is moving relative to the flashlight, the inequality $t_Q > t_P$ means that time travels more slowly for clocks (and observers and objects) that are moving relative to an event. This is called time dilation. We will explore the mathematics of time dilation further in Chapter 2.

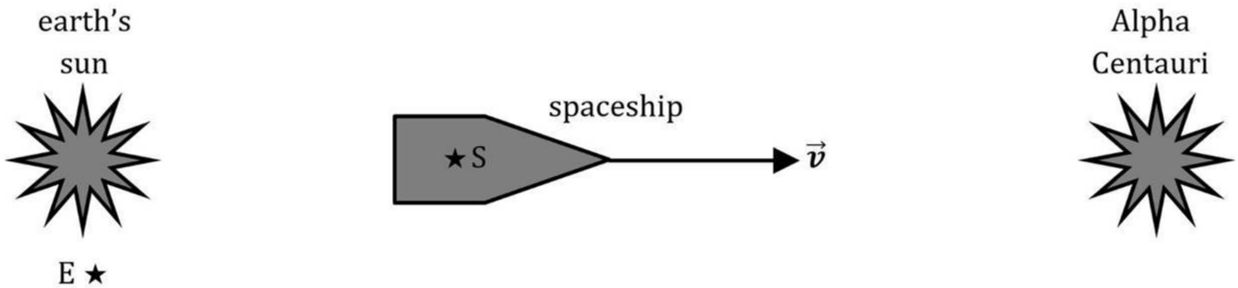
The effect is mutual. Note that observer Q is moving relative to P, but also that observer P is moving relative to Q. Therefore, time dilation depends on your perspective:

- Relative to observer P on the spaceship, time passes more slowly for observer Q who is on the space station. This is what we found in our example since we let observer P shine the flashlight inside of the spaceship.

- Relative to observer Q on the space station, time passes more slowly for observer P who is on the spaceship. If we had let observer Q shine the flashlight inside of the space station (instead of letting observer P shine the flashlight inside of the spaceship), we would have found that the time measured by observer P would have been longer.

Length Contraction

When an object is moving relative to an inertial reference frame, the object appears shorter (along the direction of motion) than it does relative to an inertial reference frame that is at rest relative to the object. This is known as **length contraction**.



You can see how length contraction comes about by considering the illustration above. The spaceship is traveling from earth's sun to Alpha Centauri with a speed that is close to the speed of light. Observer S is an astronaut aboard the spaceship, while observer E is stationed on earth. Each observer proceeds to measure the distance from the sun to Alpha Centauri (L) and the time of the trip (t).

Observer E on earth measures the distance to be L_E and the time to be t_E, such that the speed of the spaceship is $v = L_E/t_E$. Since the earth is "moving" relative to the spaceship, time is **dilated** for observer E, meaning that $t_E > t_S$. Observer S is at "rest" relative to the spaceship, and sees Alpha Centauri getting closer to the spaceship with the same speed v. (The perspective is different—whether earth is at rest and the spaceship is moving, or whether the spaceship is at rest and the stars are moving—but the speed is the same.) Observer S measures the distance to be L_S and the time to be t_S such that $v = L_S/t_S$.

Since observer S is at "rest" relative to the spaceship, time passes "normally" for observer S, such that $t_S < t_E$. If $t_S < t_E$, how can $v = L_S/t_S$ and $v = L_E/t_E$ both result in the same speed? The answer is that $L_S < L_E$. This means that the distance between the stars appears shorter for observer S than it does for observer E. Note that observer S is moving relative to the distance between the stars. Therefore, distance is shorter (along the direction of motion) relative to an inertial reference frame that it is moving relative to the distance than it is for a reference frame that is at rest relative to the distance. This is known as length contraction. We will explore the mathematics of length contraction further in Chapter 2.

The effect is mutual. Note that observer S is moving relative to E, but also that observer E is moving relative to S. Therefore, length contraction depends on your perspective:

- Relative to observer S on the spaceship, distances measured along the direction of the spaceship's motion (such as the distance between the sun and Alpha Centauri) appear shorter than they do for observer E on earth. This is what we found in our example since observer S is moving relative to the distance between the two stars.

• Relative to observer E on earth, the length of the spaceship is shorter than it is for observer S on the spaceship. If we had considered measurements of the length of the spaceship (instead of the distance between the stars) in our example, this is what we would have found.

• Relative to observer S on the spaceship traveling from the sun to Alpha Centauri, a second spaceship parked on earth would appear shorter (along the line connecting the two stars) than it is for observer E on earth.

Note that the effects of time dilation and length contraction that we discussed in this example involve two different perspectives:

• For observer E on earth, $v = L_E/t_E$, we noted that time was **dilated** ($t_E > t_S$) for observer E because the earth is "moving" relative to the spaceship.

• For observer S on the spaceship, $v = L_S/t_S$, we noted that length was **contracted** ($L_S < L_E$) for observer S because the spaceship is "moving" relative to the two stars.

Relativistic Mass

As the relative speed between two inertial reference frames gets closer to the speed of light, the effects of special relativity—including time dilation and length contraction—become more pronounced. In the limit that the relative speed approaches the speed of light, time slows down to a complete stop and length contracts to zero. However, you **can't** actually reach this limit. An object that has mass can be accelerated to nearly light speed (like 0.99*c* or 0.999*c*), but can never reach the speed of light exactly.

In Galilean relativity, it would be very easy to accelerate an object faster than the speed of light. According to **Newton's second law** of motion, the net force acting on an object equals the object's mass times its acceleration:

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

(This equation applies to objects that have constant mass, but that's not a problem for Galilean relativity, where ordinarily an object's mass isn't expected to change while it accelerates.) The mass of an electron is 9.1×10^{-31} kg. If you applied a force of just 1 N (one Newton) to an electron, according to Galilean relativity, the electron would experience an acceleration of

$$a = \frac{F}{m} = \frac{1\text{ N}}{9.1\times10^{-31}\text{kg}} = 1.1 \times 10^{30}\text{m/s}^2.$$

What does this **acceleration** mean? It means that starting from rest, after a just 1 s (one second), an electron would have a speed of $v = 1.1 \times 10^{30}$ m/s (since acceleration describes the rate at which velocity increases). That's way, way faster than the speed of light in vacuum, which is $c = 2.9979 \times 10^8$ m/s.

In the laboratory, it doesn't happen that way. Although it is very easy to accelerate electrons to very high speeds, once the speed of an electron reaches the neighborhood of the speed of light in vacuum, it becomes increasingly harder to accelerate the electron. We can accelerate electrons up to 0.9*c* or even 0.99*c*, but trying to reach 0.99999*c* is extremely difficult. Why? According to Einstein's theory of relativity, mass isn't constant: The faster an electron travels, the more mass the electron has (relative to the laboratory). We call this **relativistic mass**.

Recall that mass is a measure of **inertia** in the following sense: The more mass an object has, the more difficult it is to overcome the object's inertia in order to accelerate the object. The relativistic mass of an object describes the object's inertia. As the object travels closer to the speed of light (relative to an inertial reference frame), the greater its relativistic mass (and relativistic inertia): It becomes harder and harder to accelerate the object.

There are two types of mass:

- An object’s **rest mass** tells you how difficult it is to accelerate the object relative to an observer for which the observer and object are both at rest.

- An object’s **relativistic mass** tells you how difficult it is to accelerate the object relative to an observer for which the object is traveling close to the speed of light.

Proper Time, Proper Length, and Rest Mass

When you compare measurements of time, distance, or mass made by observers in different inertial reference frames, it's important to be able to determine which of the measurements will be larger and which will be smaller. If you can identify the proper time, the proper length, and the rest mass properly, this will help with your comparisons.

- The **proper time** corresponds to a time interval measured by a clock that is at rest relative to the events. Observers who are moving relative to the events measure a greater time interval due to time dilation.

- The **proper length** corresponds a distance measured by a tape measure (or other device used for measuring distance) that is at rest. Observers who are moving relative to the distance measure a shorter distance due to length contraction.

- The **rest mass** corresponds to the mass of an object measured by an observer who is at rest relative to the object. Observers who are moving relative to the object measure a larger mass called relativistic mass.

Note: In some relativity questions, the proper time and proper length are **<u>not</u>** measured by the *same* observer: Proper time and proper length may come from *different* perspectives (that is, two different inertial reference frames).

Symbols and SI Units

Symbol	Name	SI Units
v	relative speed between two observers	m/s
c	speed of light in vacuum	m/s
u_R	the x-component of the velocity of an object as measured by an observer at rest (called R)	m/s
u_M	the x-component of the velocity of an object as measured by an observer (called M) that is moving relative to the observer that is at rest	m/s
L_A	distance measured by observer A	m
t_A	time measured by observer A	s

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s

Strategy for Solving Galilean Relativity Problems

To solve a problem involving Galilean relativity (which applies to problems with relative speeds that are small compared to the speed of light in vacuum), follow these steps:

• Setup a coordinate system with the x-axis along the direction of the relative velocity,

$$\vec{v}$$

• Galilean relativity involves **vector subtraction**. For one-dimensional problems, simply subtract according to the following equation:

$$u_M = u_R - v$$

•• u_R is the velocity of an object as measured by an observer at **rest** (called R).

•• u_M is the velocity of an object as measured by an observer (called M) that is **moving** relative to the observer that is at rest.

•• v is the **relative speed** between the two observers.

Strategy for Solving Conceptual Special Relativity Problems

To solve a conceptual problem involving time dilation, length contraction, or relativistic mass, follow these steps. (For mathematical problems, see Chapters 2-5.)

- It may help to draw a diagram and label both objects and observers.

- The laws of special relativity apply to **inertial observers**—observers who travel with (approximately) **constant velocity** (which means that speed and direction are both constant). Any inertial observer is free to claim to be at rest, and can make equally valid arguments based on this claim.

- When applying time dilation, identify the **proper time**: The proper time is measured by an observer whose clock is at rest relative to the events. Any observer who is moving relative to the events will measure a longer time due to time dilation.

- When applying length contraction, identify the **proper length**: The proper length is measured by an observer who is at rest relative to the distance being measured. Any observer who is moving relative to the distance will measure a shorter distance due to length contraction.

- When working with relativistic mass, identify the **rest mass**: The rest mass is measured by an observer who is at rest relative to the object. Any observer who is moving relative to the object will measure a greater mass called relativistic mass.

Example: A monkey is riding on a boat that is traveling 12 m/s to the east along a river. As the boat passes a boy who is standing on the bank of the river, the monkey throws a banana 18 m/s to the east relative to the monkey. How fast is the banana moving relative to the boy?

Setup a coordinate system with $+x$ directed to the east. Identify the given information:

- The relative speed between the monkey and the boy is $v = 12$ m/s.
- The monkey is the moving observer. The velocity of the banana relative to the monkey is $u_M =$ 18 m/s to the east.
- The boy is at rest. The velocity of the banana relative to the boy is u_R.

Since these speeds are small compared to the speed of light, we may apply the equation for Galilean relativity:

$$u_M = u_R - v$$
$$18 = u_R - 12$$
$$u_R = 18 + 12 = \boxed{30\text{ m/s}}$$

The banana is moving 30 m/s relative to the boy.

Note: It is standard in physics to neglect air resistance unless stated otherwise in a problem.

Example: A monkey is riding on a boat that is traveling 12 m/s to the east along a river. As the boat passes a boy who is standing on the bank of the river, the monkey throws a banana 18 m/s to the west relative to the monkey. How fast is the banana moving relative to the boy?

Compare these two examples carefully. What is different? This time the monkey throws the banana to the west, which is opposite to the boat's motion. Setup a coordinate system with $+x$ directed to the east. Identify the given information:

- The relative speed between the monkey and the boy is $v = 12$ m/s.
- The monkey is the moving observer. The velocity of the banana relative to the monkey is $u_M = -18$ m/s. Velocity includes direction: For one-dimensional problems, minus signs distinguish between forward and backward.
- The boy is at rest. The velocity of the banana relative to the boy is u_R.

Since these speeds are small compared to the speed of light, we may apply the equation for Galilean relativity:

$$u_M = u_R - v$$
$$-18 = u_R - 12$$
$$u_R = -18 + 12 = \boxed{-6 \text{ m/s}}$$

The banana is moving 6 m/s relative to the boy. Since u_R is negative, the banana is heading west (along $-x$) relative to the boy.

Example: A monkey drives a blue car 30 m/s to the north. On the same street, another monkey drives a red car 20 m/s to the south. What is the velocity of the red car relative to the monkey in the blue car?

Don't overthink it. This problem is simpler, since there isn't a banana (or other object) moving relative to both observers. We're just trying to determine the relative velocity

$$\vec{v}$$

between the two monkeys. You can reason this out as follows:

- In one second, the blue car will travel 30 m north relative to the ground.
- In one second, the red car will travel 20 m south relative to the ground.
- Thus, in one second, the red car will be 20 + 30 = 50 m south of the blue car.
- Since the red car will be 50 m further south of the blue car each second, the velocity of the red car is –50 m/s relative to the blue car, meaning 50 m/s to the south.

Example: One astronaut is in a space station that is orbiting the earth. Another astronaut is in a spaceship. The space station is practically at rest relative to the very fast spaceship, and the spaceship is traveling close to the speed of light relative to the earth. The spaceship and space station both have transparent walls such that either astronaut can make observations (with the aid of a telescope) of what's going inside of the other space craft.

(A) How does the astronaut inside of the spaceship appear to age relative to the astronaut inside of the space station?

Identify the proper time. The astronaut on the spaceship measures the proper time because this astronaut is at rest relative to himself (or herself). An observer who is moving relative to the spaceship will measure a longer time due to time dilation. Therefore, relative to the astronaut on the space station, the astronaut on the spaceship appears to age more slowly than the astronaut on the space station. (Let's not worry about make-up, genetic differences in aging, etc. Let's treat the astronauts equally, as if they are identical twins.)

(B) Both astronauts measure the length of the spaceship. Compare their measurements.

Identify the proper length. The astronaut on the spaceship is at "rest" relative to the distance being measured. An observer who is moving relative to the spaceship will measure a shorter distance due to length contraction. Therefore, the spaceship appears to be shorter relative to the astronaut inside of the space station than it does relative to the astronaut inside of the spaceship.

(C) Both astronauts measure the mass of the spaceship. Compare their measurements.

The astronaut in the spaceship measures the spaceship's rest mass, whereas the astronaut in the space station measures the spaceship's relativistic mass, which is larger.

Chapter 1 Problems

1. A monkey is standing on the top of a train that is traveling 36 m/s to the south. A girl is standing on the ground beside the railroad tracks. As the train passes the girl, the monkey throws an apple with a speed of 12 m/s relative to the train.

(A) If the monkey throws the apple to the south, what is the speed of the apple relative to the girl?

(B) If the monkey throws the apple to the north, what is the speed of the apple relative to the girl?

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 48 m/s (B) 24 m/s

2. A monkey is riding inside of a train car that is traveling to the west with constant velocity. A girl is standing on the ground beside the railroad tracks. The monkey paints a red X on the floor of the train car (that is, the X is on the bottom of the train car, not on the ground). As the train car is passing the girl, the monkey holds a banana directly over the X and releases the banana. The train car has large windows (with no curtains), such that the girl is able to watch the banana fall and see where it lands.

(A) Where does the banana land? Explain your answer.

(B) What path does the banana take relative to the monkey?

(C) What path does the banana take relative to the girl?

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) on the X

(B) straight line (C) parabola

3. On a still pond, a monkey rides a boat 32 m/s to the east. This monkey throws an orange 20 m/s to the east relative to himself. The monkey's father rides a boat 24 m/s to the east, the monkey's mother is standing still on a pier, and the monkey's uncle rides a boat 10 m/s to the west.

(A) Find the velocity of the monkey relative to his father.

(B) Find the velocity of the monkey relative to his mother.

(C) Find the velocity of the monkey relative to his uncle.

(D) Find the velocity of the monkey's father relative to the monkey.

(E) Find the velocity of the monkey's mother relative to the monkey.

(F) Find the velocity of the monkey's uncle relative to the monkey.

(G) Find the velocity of the monkey's father relative to the monkey's mother.

(H) Find the velocity of the monkey's mother relative to the monkey's father.

(I) Find the velocity of the monkey's father relative to the monkey's uncle.

(J) Find the velocity of the monkey's uncle relative to the monkey's father.

(K) Find the velocity of the monkey's mother relative to the monkey's uncle.

(L) Find the velocity of the monkey's uncle relative to the monkey's mother.

(M) Find the velocity of the orange relative to the monkey.

(N) Find the velocity of the orange relative to the monkey's father.

(O) Find the velocity of the orange relative to the monkey's mother.

(P) Find the velocity of the orange relative to the monkey's uncle.

Want help? Check the solution at the end of the chapter.

Answers: 3. (A) 8 m/s E (B) 32 m/s E

(C) 42 m/s E (D) 8 m/s W (E) 32 m/s W

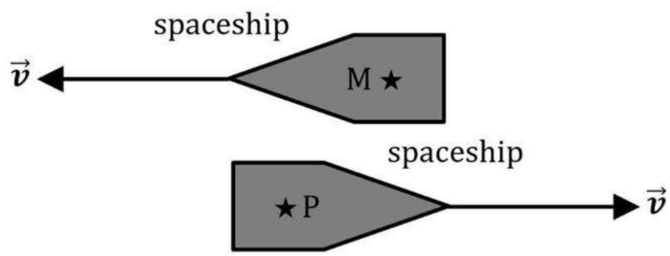
(F) 42 m/s W (G) 24 m/s E (H) 24 m/s W

(I) 34 m/s E (J) 34 m/s W (K) 10 m/s E

(L) 10 m/s W (M) 20 m/s E (N) 28 m/s E

(O) 52 m/s E (P) 62 m/s E

4. Two identical spaceships travel with a relative speed close to the speed of light. As shown below, the spaceships travel in opposite directions. At the exact moment that the spaceships pass one another, scientists aboard each ship create identical clones of a chimpanzee. In this way, the two chimpanzees are effectively identical twins. Their names are Marco and Polo.



(A) Scientists aboard Marco's ship have one photograph of Marco on his 30th birthday, and compare it to a photograph taken with a telescope of exactly how Polo looked after 30 years according to calendars kept by Marco's scientists (after accounting for the time it took for light to reach Marco's ship from Polo's ship). How do the photographs compare?

(B) Scientists aboard Polo's ship have one photograph of Polo on his 30th birthday, and compare it to a photograph taken with a telescope of exactly how Marco looked after 30 years according to calendars kept by Polo's scientists (after accounting for the time it took for light to reach Polo's ship from Marco's ship). How do the photographs compare?

(C) Which team of scientists is correct? Explain your answers to parts (A) and (B).

(D) Which ship appears longer according to measurements made by Marco's scientists?

(E) Which ship appears longer according to measurements made by Polo's scientists?

(F) Which ship has more mass according to measurements made by Marco's scientists?

(G) Which ship has more mass according to measurements made by Polo's scientists?

(H) Which ship appears taller according to measurements made by Marco's scientists?

Want help? Check the solution at the end of the chapter.

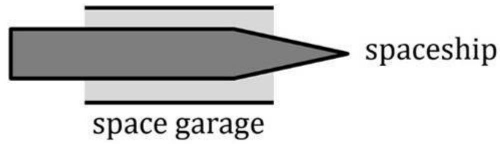
Answers: 4. (A) Polo looks younger than 30

(B) Marco looks younger than 30

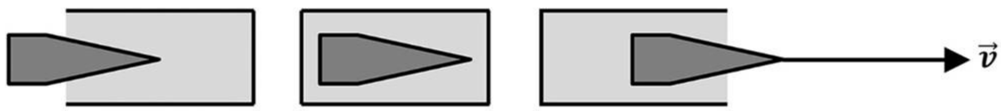
(C) both! (D) Marco's (E) Polo's

(F) Polo's (G) Marco's (H) same

5. As illustrated below, when a spaceship is parked at rest inside of a space garage, the ship is too long to fit inside of the garage.



As illustrated below, when the same spaceship is traveling close to the speed of light, it can momentarily fit inside of the same space garage according to observers stationed in the space garage. (This space garage is special: There are front and back doors which sense the presence of the spaceship, and which open or close almost instantly. Of course, the spaceship will only be inside of the space garage with both doors shut for a tiny fraction of a second relative to observers stationed in the space garage.)



(A) Explain how the spaceship is able to fit inside of the space garage with both doors closed simultaneously (for a tiny fraction of a second) from the point of view of observers stationed in the space garage.

(B) From the point of view of astronauts inside of the spaceship, is the spaceship able to fit inside of the space garage with both doors closed simultaneously? Explain.

Want help? Check the solution at the end of the chapter.

Answers: 5. (A) length contraction

(B) no; the doors do not appear to be closed simultaneously relative to observers in the spaceship

6. When cosmic rays interact with atoms high in earth's atmosphere, particles called muons can be produced. Muons are particles similar to electrons, except that they have about 200 times as much mass and are unstable. Muons decay very quickly: A muon produced at rest only lasts a couple of microseconds before decaying into other particles. Based on the short average lifetime of a muon (as measured in the muon's rest frame), almost none of the muons produced in the upper atmosphere should reach earth's surface, yet a very large number of these muons are detected at earth's surface.

(A) From the point of view of scientists stationed on the ground, explain how a large number of muons are able to reach earth's surface even though their average lifetime (as measured in the muon's rest frame) is too short for them to survive that long.

(B) Now explain this from the reference frame of the muons (instead of the reference frame of scientists stationed on the ground).

Want help? Check the solution at the end of the chapter.

Answers: 6. (A) time dilation

(B) length contraction

7. For each question below, state whether it is theoretically possible or impossible according to special relativity. Explain your answers.

(A) A chimpanzee could travel to a star that is 10,000 light-years away in her natural lifetime. Note: A light-year is the distance that light travels in one year.

(B) A chimpanzee could go on a space trip and appear younger than her own daughter when she returns to earth.

(C) A chimpanzee could go on a space trip and appear younger than she was when she left the earth.

Want help? Check the solution at the end of the chapter.

Answers: 7. (A) possible

(B) possible (C) impossible

Solutions to Chapter 1

1. Setup a coordinate system with $+x$ directed to the south. Identify the given information:

- The relative speed between the monkey and the girl is $v = 36$ m/s.

- The monkey is the moving observer. The velocity of the apple relative to the monkey is u_M. This value will be different in parts (A) and (B) since the apple is thrown in a different direction in each part.

- The girl is at rest. The velocity of the apple relative to the girl is u_R.

Since these speeds are small compared to the speed of light, we may apply the equation for Galilean relativity. As usual, we neglect air resistance unless stated otherwise.

$$u_M = u_R - v$$

(A) Since the apple is thrown south and we chose $+x$ to point south, $u_M = 12$ m/s.

$$12 = u_R - 36$$
$$u_R = 12 + 36 = \boxed{48 \text{ m/s}}$$

The apple is moving 48 m/s relative to the girl.

(B) Since the apple is thrown north and we chose $+x$ to point south, $u_M = -12$ m/s.

$$-12 = u_R - 36$$
$$u_R = -12 + 36 = \boxed{24 \text{ m/s}}$$

The apple is moving 24 m/s relative to the girl.

2. (A) The banana lands directly on the X. Why? Because the banana has **inertia**. Recall from Newton's laws of motion (which are taught in first-year physics) that inertia is the natural tendency of all objects to maintain constant velocity. According to Newton's second law of motion, a net external force is needed to accelerate an object (and thus change the object's velocity), since

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

When the monkey releases the banana, a net gravitational force acts downward, causing the banana to accelerate downward. However, there are **no** forces acting horizontally, so the banana doesn't accelerate horizontally. The banana maintains a constant horizontal component of velocity (v_x), while the vertical component of velocity (v_y) changes. That's why, in projectile motion, horizontally we have

$$\Delta x = v_x t$$

and vertically we have

$$\Delta y = v_{yo} t + \frac{1}{2} a_y t^2$$

Horizontally, v_x is constant (because $a_x = 0$), whereas vertically there is uniform acceleration ($a_y = -g = -9.81$ m/s^2 is constant).

You can verify this by riding in an airplane. If you throw an eraser straight upward relative to you while sitting inside of the airplane, you will catch the eraser because it has inertia. The eraser surely won't land behind you, even if the airplane is traveling 300 m/s and if the eraser is in the air for half a second (in which case the airplane travels 150 m horizontally).

(B) Relative to the monkey, the banana appears to fall in a straight line downward, no different than if the train had been parked when the monkey released the banana. The laws of physics are the same in any inertial reference frame, meaning that the result of dropping a banana from rest will be the same whether the train is at rest or moving with constant velocity.

(C) Relative to the girl, the banana follows the arc of a parabola, beginning with a horizontal tangent. The same path would result if the girl had the banana and threw it horizontally. The banana follows the path of a projectile, which is parabolic.

From the monkey's point of view, the monkey claims that the train is at "rest" (but that the girl and ground are moving). According to the monkey, $v_x = 0$ and the banana falls straight downward.

From the girl's point of view, the girl claims that she is at "rest" (but that the train is moving). According to the girl, v_x isn't 0. Relative to the girl,

$$\Delta x = v_x t$$

and

$$\Delta y = v_{yo} t + \frac{1}{2} a_y t^2$$

which can be combined to make

$$\Delta y = \Delta x \tan\theta_0 - \frac{g}{2v_0^2 \cos^2\theta_0} \Delta x^2$$

which is the equation of a parabola. To derive this equation, you need to combine the following equations:

$$\Delta x = v_x t$$

$$\Delta y = v_{yo} t + \frac{1}{2} a_y t^2$$

$$v_x = v_{x0}$$

$$v_{x0} = v_0 \cos\theta_0$$

$$v_{y0} = v_0 \sin\theta_0$$

$$a_y = -g$$

3. Setup a coordinate system with $+x$ directed to the east.

$$p = \text{pier}, k = \text{monkey}, f = \text{father}$$

$$m = \text{mother}, u = \text{uncle}, o = \text{orange}$$

$$v_{ab} = \text{the } x\text{-component of the velocity of } a \text{ relative to } b$$

$$v_{kp} = 32 \text{ m/s} \quad , \quad v_{fp} = 24 \text{ m/s} \quad , \quad v_{mp} = 0$$

$$v_{up} = -10 \text{ m/s} \quad , \quad v_{om} = 20 \text{ m/s}$$

$$v_{ab} = v_{ac} - v_{bc}$$

(A) $v_{kf} = v_{kp} - v_{fp} = 32 - 24 = 8 \text{ m/s}$ (east)

(B) $v_{km} = v_{kp} - v_{mp} = 32 - 0 = 32 \text{ m/s}$ (east)

(C) $v_{ku} = v_{kp} - v_{up} = 32 - (-10) = 32 + 10 = 42 \text{ m/s}$ (east)

(D) $v_{fk} = v_{fp} - v_{kp} = 24 - 32 = -8 \text{ m/s}$ (west)

(E) $v_{mk} = v_{mp} - v_{kp} = 0 - 32 = -32 \text{ m/s}$ (west)

(F) $v_{uk} = v_{up} - v_{kp} = -10 - 32 = -42 \text{ m/s}$ (west)

(G) $v_{fm} = v_{fp} - v_{mp} = 24 - 0 = 24 \text{ m/s}$ (east)

(H) $v_{mf} = v_{mp} - v_{fp} = 0 - 24 = -24 \text{ m/s}$ (west)

(I) $v_{fu} = v_{fp} - v_{up} = 24 - (-10) = 24 + 10 = 34 \text{ m/s}$ (east)

(J) $v_{uf} = v_{up} - v_{fp} = -10 - 24 = -34 \text{ m/s}$ (west)

(K) $v_{mu} = v_{mp} - v_{up} = 0 - (-10) = 0 + 10 = 10 \text{ m/s}$ (east)

(L) $v_{um} = v_{up} - v_{mp} = -10 - 0 = -10 \text{ m/s}$ (west)

(M) $v_{ok} = 20 \text{ m/s}$ (east); this was stated in the problem

(N) In the formula $v_{ab} = v_{ac} - v_{bc}$ (which you can verify is consistent with the formulas used in parts A thru M), let $a = o$ (orange), $b = f$ (father), and $c = k$ (monkey). Recall from part D that $v_{fk} = -8$ m/s:

$v_{of} = v_{ok} - v_{fk} = 20 - (-8) = 20 + 8 = 28$ m/s (east)

Alternatively, you can reason this out conceptually: The orange is traveling 20 m/s faster than the monkey, who is traveling 8 m/s faster than the father: $20 + 8 = 28$ m/s.

(O) $v_{om} = v_{ok} - v_{mk} = 20 - (-32) = 20 + 32 = 52$ m/s (east)

(P) Recall from part F that $v_{uk} = -42$ m/s.

$v_{ou} = v_{ok} - v_{uk} = 20 - (-42) = 20 + 42 = 62$ m/s (east)

Alternatively, you can reason this out conceptually: The orange is traveling 20 m/s faster than the monkey, who is traveling 42 m/s faster than the uncle (since 32 m/s to the east and 10 m/s to the west have a relative speed of 42 m/s): $20 + 42 = 62$ m/s.

4. (A) Marco appears to age normally relative to the scientists on Marco's ship. Because the proper time for Polo's aging process is measured by Polo's team, Marco's team will measure Polo's aging process to occur more slowly due to time dilation. Therefore, when Marco's team compares their photographs, Polo will appear younger.

(B) Polo appears to age normally relative to the scientists on Polo's ship. Because the proper time for Marco's aging process is measured by Marco's team, Polo's team will measure Marco's aging process to occur more slowly due to time dilation. Therefore, when Polo's team compares their photographs, Marco will appear younger.

(C) Both teams are correct. Marco's team believes that Polo appears younger when Marco celebrates his 30^{th} birthday, and Polo's team believes that Marco appears younger when Polo celebrates his 30^{th} birthday, and both teams are correct because both teams are inertial observers (since both spaceships travel with constant velocity). According to special relativity, there is no preferred reference frame; and the laws of physics are the same for all inertial observers. (For the chimpanzees to actually meet up, note that one would have to accelerate.)

(D) Marco's ship appears normal relative to the scientists on Marco's ship. Because the proper length for Polo's ship is measured by Polo's team, Marco's team will measure the length of Polo's ship to be shorter than normal due to length contraction; Marco's appears longer.

(E) Polo's ship appears normal relative to the scientists on Polo's ship. Because the proper length for Marco's ship is measured by Marco's team, Polo's team will measure the length of Marco's ship to be shorter than normal due to length contraction; Polo's appears longer.

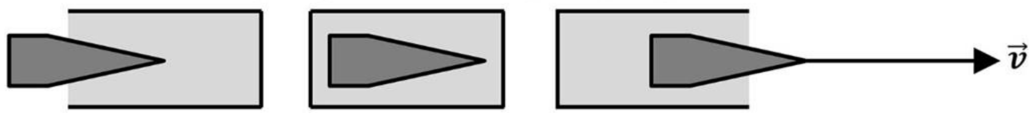
(F) Marco's team measures Marco's ship to have its rest mass. When Marco's team measures the mass of Polo's ship, Polo's ship's relativistic mass appears greater than its rest mass. (Let's assume that both teams of scientists have the same combined rest mass.)

(G) Polo's team measures Polo's ship to have its rest mass. When Polo's team measures the mass of Marco's ship, Marco's ship's relativistic mass appears greater than its rest mass.

(H) The height is the same for each. (Length contraction occurs only along the direction of the velocity. The height, which is perpendicular to the velocity, is unaffected.)

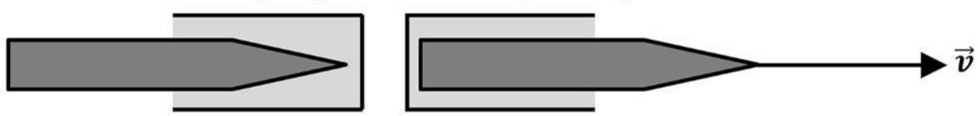
5. (A) Since the proper length of the spaceship is measured by astronauts inside of the ship, astronauts stationed in the space garage measure the length of the spaceship to be shorter than normal due to length contraction. The concept of length contraction explains how the ship is short enough to fit in the garage while traveling close to the speed of light.

from the perspective of the space garage



(B) The spaceship has its usual length relative to astronauts inside of the spaceship. You can't explain this in terms of length contraction because length contraction would make the space garage appear shorter than usual relative to astronauts inside of the spaceship. The answer has to do with the fact that different inertial observers often disagree on whether or not two events occur simultaneously; this is one of those times. Whereas astronauts stationed in the garage momentarily see the spaceship fit inside of the garage with both doors temporarily closed, astronauts aboard the spaceship don't see the two doors closed at the same instant. Rather, astronauts aboard the spaceship see the door on the right side of our diagram close first (while the back of the spaceship sticks out on the left side of our diagram). They then see the front door open. The spaceship continues a short ways until the back of the spaceship is safely inside of the garage. At this point, the door on the left side of our diagram quickly closes (while the front of the spaceship sticks out on the right side of our diagram).

from the perspective of the spaceship



6. (A) Since the proper time would be measured by a reference frame traveling with the muons themselves, scientists on earth's surface measure a longer lifetime due to time dilation. That is, because the muons travel close to the speed of light relative to the earth, their average lifetime is much longer than it would be if they were at rest, which allows them to travel a longer distance before they decay.

(B) From the reference frame of the muons, their lifetime is normal (they measure the proper time). Instead, the muons "see" length contraction: Since scientists on earth's surface measure the proper length for the muons' trip, a reference frame attached to the muons would measure a shorter distance. That is, because the muons travel close to the speed of light (although from their perspective, the muons are at rest and the earth is traveling close to the speed of light towards the muons), the distance between earth's surface and where they are produced in earth's atmosphere is much shorter than would be if they were at rest, which allows them to reach earth's surface in less time. It's instructive to compare how these different perspectives lead to equivalent conclusions through two quite different effects.

7. (A) Theoretically, it is possible, provided that the chimpanzee travels in a spaceship with a speed that is close enough to the speed of light (that's the hard part). Relative to observers on earth, the chimpanzees would age much more slowly than normal through time dilation. (However, the initial part of the trip requires acceleration, and the end of the trip involves deceleration, both of which involve general relativity, which is a step beyond special relativity.)

(B) Theoretically, it is possible, for the same reason as part A, provided that the chimpanzee travels in a spaceship that is close enough to the speed of light. If the chimpanzee ages slowly enough due to time dilation, she could appear younger than her daughter when she returns to earth. However, whatever the age difference is between the chimpanzee and her daughter, at least that number of years must pass on earth for this to be possible. For example, if the chimpanzee is 30 years old and her daughter is 12 years old, at least 18 years must pass on earth during the trip (plus additional years depending on how fast the ship travels; we'll explore the mathematics involved in time dilation in Chapter 2).

(C) This is theoretically impossible (without plastic surgery or age defying medicine). While time can slow down due to time dilation, in special relativity time can't go backwards.

2 TIME DILATION AND LENGTH CONTRACTION

Relevant Terminology

Time dilation – the phenomenon whereby time appears to travel more slowly for objects moving fast (close to light speed) relative to other observers.

Length contraction – the phenomenon whereby objects moving fast (close to light speed) appear shorter relative to other observers.

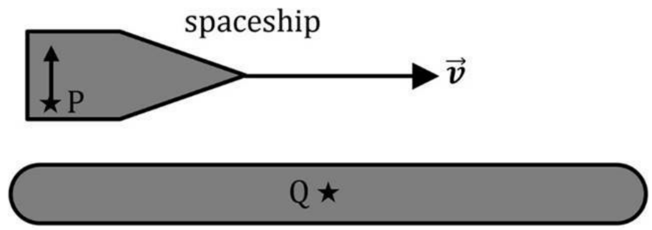
Proper time – a time interval measured by a clock that is at rest relative to the events. An observer who is moving relative to the events measures a greater time interval due to time dilation.

Proper length – a distance measured by an observer who is at rest relative to the distance. An observer who is moving relative to the distance measures a shorter distance due to length contraction.

Inertial reference frame – a frame that travels with constant velocity.

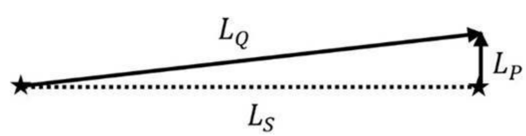
Time Dilation

Time passes more slowly on a clock in a moving inertial reference frame than it does for an inertial reference frame that is at rest relative to the clock. This is known as **time dilation**.



In the diagram above, the spaceship is traveling to the right with a speed that is close to the speed of light. The spaceship passes by a space station that is practically at rest relative to the very fast spaceship. As the spaceship is passing the space station, an astronaut inside of the spaceship turns on a flashlight, shining a beam of light straight upward in the diagram (perpendicular to the direction that the spaceship is traveling).

Relative to the astronaut inside of the spaceship (observer P), the flashlight beam appears to travel straight upward (left diagram). Relative to an observer inside of the space station (observer Q), the flashlight beam appears to travel diagonally (right diagram).



In the illustration above:

- L_P is the distance that the light travels according to observer P.
- L_Q is the distance that the light travels according to observer Q.
- L_S is the distance that the spaceship travels horizontally during this time.

These three distances are related by the Pythagorean theorem:

$$L_Q^2 = L_S^2 + L_P^2$$

According to the second postulate of special relativity, both observers must measure the speed of light to be the **same** value. Either observer takes the distance (L) traveled and divides by the corresponding time (t) measured to determine the speed of light (c). The subscripts P and Q indicate which observer makes the measurement.

$$c = \frac{L_P}{t_P} \quad , \quad c = \frac{L_Q}{t_Q}$$

Observer Q can also measure the speed (v) of the spaceship by dividing the horizontal distance (L_S) traveled by the corresponding time (t_Q).

$$v = \frac{L_S}{t_Q}$$

Multiply each equation by the corresponding time.

$$L_P = c\, t_P$$

$$L_Q = c\, t_Q$$

$$L_S = v\, t_Q$$

Substitute these expressions into the equation from the Pythagorean theorem.

$$L_Q^2 = L_S^2 + L_P^2$$
$$c^2 t_Q^2 = v^2 t_Q^2 + c^2 t_P^2$$

Solve for t_Q.

$$c^2 t_Q^2 - v^2 t_Q^2 = c^2 t_P^2$$
$$t_Q^2 (c^2 - v^2) = c^2 t_P^2$$
$$t_Q^2 = \frac{c^2 t_P^2}{c^2 - v^2}$$

Divide the numerator and denominator each by c^2.

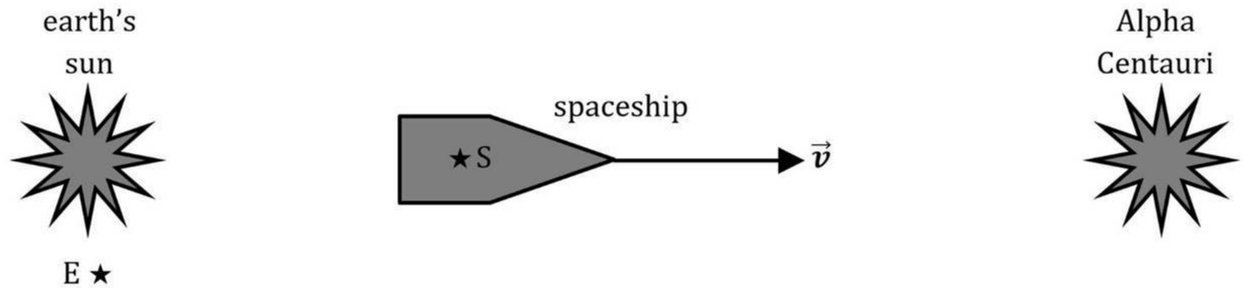
$$t_Q^2 = \frac{t_P^2}{1 - \left(\frac{v}{c}\right)^2}$$

$$t_Q = \frac{t_P}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

Since $v < c$, it follows that $1 - \left(\frac{v}{c}\right)^2 < 1$, such that $t_Q > t_P$: Time is **dilated** for observer Q relative to observer P. Note that t_P is called the **proper time**: the clock is at 'rest' relative to P.

Length Contraction

When an object is moving relative to an inertial reference frame, the object appears shorter (along the direction of motion) than it does relative to an inertial reference frame that is at rest relative to the object. This is known as **length contraction**.



In the diagram above, a spaceship is traveling from earth's sun to Alpha Centauri with a speed that is close to the speed of light. Observer S is an astronaut aboard the spaceship, while observer E is stationed on earth.

Observer E on earth measures the distance to be L_E and the time to be t_E, such that the speed of the spaceship is

$$v = L_E / t_E$$

Observer S on the spaceship measures the distance to be L_S and the time to be t_S such that

$$v = L_S / t_S$$

According to observer E, the ship is "moving" while the stars are at rest, whereas according to observer S, the earth and stars are "moving" while the spaceship is at "rest," but either way the speed of the spaceship is the **same**:

$$\frac{L_E}{t_E} = \frac{L_S}{t_S}$$

Solve for L_S.

$$L_S = \frac{t_S}{t_E} L_E$$

Since the spaceship's clock is at "rest" relative to the journey, observer S measures the **proper time** and the passage of time aboard the spaceship appears **dilated** relative to observer E. Use the time dilation equation with t_S as the proper time:

$$t_E = \frac{t_S}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

Solve for the ratio $\frac{t_S}{t_E}$. Multiply both sides of the equation by $\sqrt{1 - \left(\frac{v}{c}\right)^2}$ and divide by t_E.

$$\frac{t_S}{t_E} = \sqrt{1 - \left(\frac{v}{c}\right)^2}$$

Substitute this expression into the equation $L_S = \frac{t_S}{t_E} L_E$.

$$L_S = L_E \sqrt{1 - \left(\frac{v}{c}\right)^2}$$

Since $v < c$, it follows that $1 - \left(\frac{v}{c}\right)^2 < 1$, such that $L_S < L_E$, showing that length is **contracted** for observer S relative to observer E. Note that L_E is called the **proper length** in this example since observer E is at rest relative to the stars (and since in this example we are measuring the distance between the stars).

Note: When applying the above equation to other problems, the **proper length** will often **not** be measured by an observer on earth. That just happened to be the case in this example. In a given problem, you must apply the concept of proper length to determine which observer measures the proper length.

Note that the effects of time dilation and length contraction that we discussed in this example involve two different perspectives:

- For observer E on earth, $v = L_E / t_E$, we noted that time was **dilated** ($t_E > t_S$) for observer E because the earth is "moving" relative to the spaceship.

- For observer S on the spaceship, $v = L_S / t_S$, we noted that length was **contracted** ($L_S < L_E$) for observer S because the spaceship is "moving" relative to the two stars.

Time Dilation and Length Contraction Equations

In the equations below, t_0 represents the **proper time** (measured by an observer who is at rest relative to the events) and L_0 represents the **proper length** (measured by an observer who is at rest relative to the distance). Note that t_0 and L_0 are not necessarily measured by the same observer in a problem (in fact, these were measured by different observers when we derived the equation for length contraction in the previous section).

$$t_d = \gamma t_0 = \frac{t_0}{\sqrt{1-\beta^2}} = \frac{t_0}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$$

$$L_c = \frac{L_0}{\gamma} = L_0\sqrt{1-\beta^2} = L_0\sqrt{1-\left(\frac{v}{c}\right)^2}$$

In the equations above, the symbols β and γ are defined as follows:

$$\beta = \frac{v}{c} \quad , \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$$

Symbols and SI Units

Symbol	Name	SI Units
v	relative speed between two observers	m/s
c	speed of light in vacuum	m/s
L_A	distance measured by observer A	m
t_A	time measured by observer A	s
L_0	proper length	m
t_0	proper time	s
β	fraction of the speed of light	unitless
γ	time dilation factor	unitless

Note: The symbols β and γ are the lowercase Greek letters beta and gamma.

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s

Strategy for Solving Time Dilation and Length Contraction Problems

To solve a problem involving time dilation or length contraction, follow these steps:

- It may help to draw a diagram and label both objects and observers.

- When applying time dilation, identify the **proper time** (t_0), which is measured by an observer whose clock is at rest relative to the events.

$$t_d = \gamma t_0 = \frac{t_0}{\sqrt{1-\beta^2}} = \frac{t_0}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$$

- When applying length contraction, identify the **proper length** (L_0), which is measured by an observer who is at rest relative to the distance being measured.

$$L_c = \frac{L_0}{\gamma} = L_0\sqrt{1-\beta^2} = L_0\sqrt{1-\left(\frac{v}{c}\right)^2}$$

Note that t_0 and L_0 are not necessarily measured by the same observer in a problem.

Example: A chimpanzee astronaut sleeps for 8 hours according to a spaceship's clock while traveling at 0.5c (half the speed of light) relative to the earth. For how much time does the chimpanzee appear to be sleeping relative to an observer on earth?

Which observer measures the proper time? The spaceship's clock is at rest relative to the chimpanzee. Therefore, the chimpanzee measures the proper time: $t_0 = 8$ hr. The observer on earth measures a greater time, t_d, due to time dilation. The relative speed is $v = 0.5c$. Use the time dilation equation.

$$t_d = \frac{t_0}{\sqrt{1-\left(\frac{v}{c}\right)^2}} = \frac{8}{\sqrt{1-\left(\frac{0.5c}{c}\right)^2}} = \frac{8}{\sqrt{1-(0.5)^2}}$$

$$= \frac{8}{\sqrt{1-\left(\frac{1}{2}\right)^2}} = \frac{8}{\sqrt{1-\frac{1}{4}}} = \frac{8}{\sqrt{\frac{3}{4}}}$$

$$t_d = \frac{8}{\frac{\sqrt{3}}{2}} = 8 \div \frac{\sqrt{3}}{2} = 8 \times \frac{2}{\sqrt{3}}$$

$$= \frac{16}{\sqrt{3}} = \boxed{\frac{16\sqrt{3}}{3}\text{ hr}} = \boxed{9.2\text{ hr}}$$

Recall that the way to divide by a fraction is to multiply by its reciprocal. Note that $\frac{16}{\sqrt{3}} = \frac{16}{\sqrt{3}}\frac{\sqrt{3}}{\sqrt{3}} = \frac{16\sqrt{3}}{3}$ since $\sqrt{3}\sqrt{3} = 3$. We multiplied $\frac{16}{\sqrt{3}}$ by $\frac{\sqrt{3}}{\sqrt{3}}$ in order to rationalize the denominator. You can verify with a calculator that $\frac{16}{\sqrt{3}}$ and $\frac{16\sqrt{3}}{3}$ both equal 9.2 (after rounding to two significant figures).

Example: A spaceship has a length of 20 m when it is parked near the surface of the earth. When a chimpanzee astronaut in the spaceship travels at $0.8c$ relative to the earth, what length does the spaceship appear to have relative to an observer on earth?

Which observer measures the proper length? The chimpanzee is at rest relative to the length of the spaceship. Therefore, the chimpanzee measures the proper length: $L_0 = 20$ m. The observer on earth measures a shorter distance, L_c, due to length contraction. The relative speed is $v = 0.8c$. Use the length contraction equation.

$$L_c = L_0\sqrt{1 - \left(\frac{v}{c}\right)^2} = 20\sqrt{1 - \left(\frac{0.8c}{c}\right)^2}$$

$$= 20\sqrt{1 - (0.8)^2} = 20\sqrt{1 - 0.64} = 20\sqrt{0.36}$$

$$L_c = 20(0.6) = \boxed{12 \text{ m}}$$

Chapter 2 Problems

1. A chimpanzee astronaut travels in a spaceship at $0.6c$ relative to the earth. According to the chimpanzee, the spaceship is 30 m long and the trip takes 12 years.

(A) How long is the spaceship relative to an observer on earth?

(B) How long does the trip take relative to an observer on earth?

(C) How far does the spaceship travel according to an observer on earth?

Note: One light-year (ly) is the distance that light travels in one year.

(D) How far does the spaceship travel according to the chimpanzee?

(E) Explain your answers to parts C and D.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 24 m (B) 15 years

(C) 9.0 ly (D) 7.2 ly

(E) time dilation vs. length contraction

2. How fast must a spaceship travel relative to another observer in order to appear half as long as it really is?

Want help? Check the solution at the end of the chapter.

Answer:

$$\frac{\sqrt{3}}{2}c = 0.87c = 2.6 \times 10^8 \text{ m/s}$$

3. Muons that are produced at rest have an average lifetime of

$$2.2\ \mu\text{s (where } \mu = 10^{-6} \text{ is the prefix micro).}$$

A beam of muons is produced that travels 0.99*c* relative to the earth.

(A) How far does classical physics expect the muons to travel on average? (For this question, pretend that the muons don't follow the laws of special relativity.)

(B) Relative to observers on earth, how far will the muons actually travel on average?

(C) Relative to a reference frame attached to the muons, how far do the muons travel on average?

(D) Explain your answers to parts B and C.

Want help? Check the solution at the end of the chapter.

Answers: 3. (A) 0.65 km

(B) 4.6 km (C) 0.65 km

(D) time dilation vs. length contraction

4. A chimpanzee astronaut travels in a spaceship at $(12/13)c$ relative to the earth. According to observers stationed on earth, the trip takes 26 years.

(A) How far does the spaceship travel relative to the earth?

(B) How far does the spaceship travel relative to the chimpanzee?

(C) How long does the trip take relative to the chimpanzee?

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) 24 ly (B) (120/13) ly = 9.2 ly (C) 10 yr

Solutions to Chapter 2

1. The relative speed is $v = 0.6c$.

Use the equations for β and γ:

$$\beta = \frac{v}{c} = \frac{0.6c}{c} = 0.6$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-0.6^2}} = \frac{1}{\sqrt{1-0.36}}$$

$$= \frac{1}{\sqrt{0.64}} = \frac{1}{0.8} = \frac{1}{4/5} = \frac{5}{4}$$

(A) Regarding the length of the spaceship, the chimpanzee measures the proper length since the chimpanzee is at rest relative to the spaceship: $L_0 = 30$ m. Use the equation for length contraction to determine what an observer on earth measures.

$$L_E = \frac{L_0}{\gamma} = \frac{30}{\frac{5}{4}} = 30 \div \frac{5}{4}$$

$$= 30 \times \frac{4}{5} = \frac{120}{5} = \boxed{24 \text{ m}}$$

Recall that the way to divide by a fraction is to multiply by its reciprocal.

(B) Regarding the time of the trip, the chimpanzee measures the proper time since the ship's clock is at "rest" relative to the journey: $t_0 = 12$ yr. Use the equation for time dilation to determine what an observer on earth measures.

$$t_E = \gamma t_0 = \frac{5}{4}(12) = \frac{60}{4} = \boxed{15 \text{ yr}}$$

(C) Multiply the speed of the ship by the time of the trip as measured by an earth observer.

$$d_E = v\, t_E = (0.6c)(15 \text{ yr}) = \boxed{9.0 \text{ ly}}$$

Note that one lightyear (ly) equals c times 1 yr. That is, the speed of light times one year equals the distance that light travels in one year. (If you prefer meters, use $c = 2.9979 \times 10^8$ m/s and convert 1 yr to 31,536,000 seconds to get 8.5×10^{16} m.) Unlike part A, the observer on earth measures the proper length for this distance (which is at rest relative to earth).

(D) Multiply the speed of the ship by the time of the trip as measured by the chimpanzee.

$$d_C = v\, t_C = v\, t_0 = (0.6c)(12 \text{ yr}) = \boxed{7.2 \text{ ly}}$$

Since the observer on earth measures the proper length of the trip (since the starting and ending points of the journey—which are likely the sun and a nearby star—aren't moving relative to the earth), the chimpanzee measures a shorter length due to length contraction. We could have obtained the same answer from the length contraction equation (using d for the distance of the trip, so as not to confuse it with the L that we used for the length of the ship in part A). It's instructive to compare part D with part A, since in part A the chimpanzee measured the proper length, whereas in part D the observer on earth measures the proper length.

That is, $L_0 = L_C$ and $L_E = \frac{L_0}{\gamma}$ in part A, whereas $d_0 = d_E$ and $d_C = \frac{d_0}{\gamma}$ in part D (where the subscript E stands for earth and the subscript C stands for chimpanzee).

$$d_C = \frac{d_0}{\gamma} = \frac{d_E}{\gamma} = \frac{9}{5/4} = 9 \div \frac{5}{4} = 9 \times \frac{4}{5} = \frac{36}{5} = 7.2 \text{ ly}$$

(E) In part C, the speed of the ship is $\frac{d_E}{t_E} = \frac{9}{15}c = 0.6c$. Compare with part D: $\frac{d_C}{t_C} = \frac{7.2}{12}c = 0.6c$. Either way, the speed of the ship is the $\boxed{\text{same}}$. What's different are the two perspectives. On earth, the distance is the proper length, $d_E = 9.2$ ly, while the time is dilated, $t_E = 15$ yr. To the chimpanzee, the length is contracted, $d_C = 7.2$ ly, while the time is proper, $t_0 = 12$ yr. These two different phenomena (time dilation and length contraction) lead to exactly the same effect in the two different perspectives.

2. According to the problem, $L_c = \frac{L_0}{2}$. Plug this into the equation $L_c = \frac{L_0}{\gamma}$ to get $\frac{L_0}{2} = \frac{L_0}{\gamma}$. The two denominators must be equal for this equation to be true: $\gamma = 2$. Set $\gamma = 2$ in $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ to get $2 = \frac{1}{\sqrt{1-\beta^2}}$. Square both sides to get $4 = \frac{1}{1-\beta^2}$. Multiply both sides by $1 - \beta^2$ and divide by 4 to get $1 - \beta^2 = \frac{1}{4}$. Add β^2 to both sides and subtract $\frac{1}{4}$ from both sides: $1 - \frac{1}{4} = \beta^2$. This simplifies to $\frac{3}{4} = \beta^2$. Squareroot both sides to get $\frac{\sqrt{3}}{2} = \beta$. Plug this into the equation $\beta = \frac{v}{c}$ to get $\frac{\sqrt{3}}{2} = \frac{v}{c}$. Multiply both sides by c to get $v = \boxed{\frac{\sqrt{3}}{2}c} = \boxed{0.87c}$ (to two significant figures). If you prefer meters per second, plug in $c = 2.9979 \times 10^8$ m/s to get $v = \boxed{2.6 \times 10^8 \text{ m/s}}$.

3. Use $c = 2.9979 \times 10^8$ m/s to get $v = 0.99c = 2.9679 \times 10^8$ m/s. Use the equations for β and γ:

$$\beta = \frac{v}{c} = \frac{0.99c}{c} = 0.99$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-0.99^2}} = \frac{1}{\sqrt{1-0.9801}} = \frac{1}{\sqrt{0.0.0199}} = \frac{1}{0.14107} = 7.0888$$

(A) Multiply the speed of the muons by the average lifetime.

$$d_{\text{classical}} = vt_0 = (2.9679 \times 10^8)(2.2 \times 10^{-6})$$
$$= 653 \text{ m} \approx \boxed{0.65 \text{ km}}$$

(B) The average lifetime will be dilated relative to observers on earth.

$$t_E = \gamma t_0 = (7.0888)(2.2 \times 10^{-6})$$
$$= 15.595 \times 10^{-6} \text{ s} \approx 16 \times 10^{-6} \text{ s} = 16\ \mu\text{s}$$
$$d_E = v\, t_E = (2.9679 \times 10^8)(15.595 \times 10^{-6})$$
$$= 4628 \text{ m} \approx \boxed{4.6 \text{ km}}$$

(C) In the muons' reference frame, the lifetime is normal, but the distance that they travel (on average) is contracted.

$$d_M = \frac{d_0}{\gamma} = \frac{4628}{7.0888} = 653 \text{ m} \approx \boxed{0.65 \text{ km}}$$

(D) In part B, the speed of the muons is $\frac{d_E}{t_E} = \frac{4628}{15.595\times10^{-6}} = 0.99c$. Compare with part C: $\frac{d_M}{t_M} = \frac{653}{2.2\times10^{-6}} = 0.99c$. Either way, the speed of the muons is the $\boxed{\text{same}}$. (It's exactly the same, though if you plug in our rounded numbers, they will differ slightly due to round-off error.) What's different are the two perspectives. On earth, the distance is the proper length, $d_E = 4.6$ km, while the time is dilated, $t_E = 16$ μs. For the muons' frame, the length is contracted, $d_M = 0.65$ km, while the time is proper, $t_0 = 2.2$ μs. These two different phenomena (time dilation and length contraction) lead to exactly the same effect in the two different perspectives.

4. The relative speed is $v = \frac{12}{13}c$. Use the equations for β and γ:

$$\beta = \frac{v}{c} = \frac{\frac{12}{13}c}{c} = \frac{12}{13}$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{12}{13}\right)^2}} = \frac{1}{\sqrt{1-\frac{144}{169}}}$$

$$= \frac{1}{\sqrt{\frac{25}{169}}} = \frac{1}{5/13} = \frac{13}{5}$$

(A) Multiply the speed of the ship by the time of the trip as measured by an earth observer.

$$d_E = v\,t_E = \left(\frac{12}{13}c\right)(26\text{ yr}) = \boxed{24\text{ ly}}$$

Note that one lightyear (ly) equals c times 1 yr. That is, the speed of light times one year equals the distance that light travels in one year. (If you prefer meters, use $c = 2.9979 \times 10^8$ m/s and convert 1 yr to 31,536,000 seconds to get 2.3×10^{17} m.)

(B) Since the observer on earth measures the proper length of the trip (since the starting and ending points of the journey—which are likely the sun and a nearby star—aren't moving relative to the earth), the chimpanzee measures a shorter length due to length contraction.

$$d_C = \frac{d_0}{\gamma} = \frac{d_E}{\gamma} = \frac{24}{\frac{13}{5}} = 24 \div \frac{13}{5}$$

$$= 24 \times \frac{5}{13} = \frac{120}{13} \text{ ly} \approx \boxed{9.2 \text{ ly}}$$

(C) Divide the distance by the speed.

$$t_C = \frac{d_C}{v} = \frac{120/13}{12/13} = \boxed{10 \text{ yr}}$$

Alternatively, use the time dilation formula:

$$t_E = \gamma t_0 = \gamma t_C \quad \rightarrow \quad t_C = \frac{t_E}{\gamma} = \frac{26}{13/5} = 10 \text{ yr}$$

3 THE LORENTZ TRANSFORMATION

Relevant Terminology

Time dilation – the phenomenon whereby time appears to travel more slowly for objects moving fast (close to light speed) relative to other observers.

Length contraction – the phenomenon whereby objects moving fast (close to light speed) appear shorter relative to other observers.

Inertial reference frame – a frame that travels with constant velocity.

The Galilean Transformation

The Galilean transformation involves Galilean relativity (Chapter 1), which only applies when the objects and observers are traveling with speeds that are small compared to the speed of light. Suppose that there are two different observers, O and O', where O' moves with constant velocity

$$\vec{v}$$

relative to O. The Galilean transformation relates the coordinates (t, x, y, z) and (t', x', y', z') of the two observers. If we setup our coordinate systems such that the $+x$- and $+x'$-axes are oriented along the relative velocity,

$$\vec{v}$$

and if we start our clocks such that O and O' coincide at $t = t' = 0$, the Galilean transformation is:

$$t' = t$$

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

Note that vt equals the distance between the two coordinate systems at time t. To see that this Galilean transformation agrees with the Galilean relativity equation that we learned in Chapter 1, take a derivative with respect to time (noting that $dt' = dt$ since $t' = t$):

$$\frac{dx'}{dt} = \frac{dx}{dt} - v$$

Since a derivative of x with respect to time equals the x-component of velocity, the above equation is identical to the Galilean relativity equation $u_M = u_R - v$ from Chapter 1.

The Lorentz Transformation

The Galilean transformation **doesn't** apply when the relative velocity is significant compared to the speed of light. The Lorentz transformation accounts for the effects of special relativity, and applies at all speeds (low or high):

$$t' = \gamma\left(t - \frac{v}{c^2}x\right) \quad , \quad x' = \gamma(x - vt)$$

$$y' = y \quad , \quad z' = z$$

Recall the following definitions from Chapter 2:

$$\beta = \frac{v}{c} \quad , \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$$

Also recall that O' moves with constant velocity

$$\vec{v}$$

relative to O.

The Lorentz transformation is convenient when you know the coordinates (t, x, y, z) in O and wish to find the coordinates (t', x', y', z') in O'. What if you know (t', x', y', z') and wish to find (t, x, y, z)? In that case, use the following inverse transformation:

$$t = \gamma\left(t' + \frac{v}{c^2}x'\right) \quad , \quad x = \gamma(x' + vt')$$

$$y = y' \quad , \quad z = z'$$

In some problems, we don't measure instantaneous values of the coordinates, but instead measure intervals (such as a time interval between two events, or the endpoints of a rod). In terms of intervals, the Lorentz transformation is:

$$\Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right) \quad , \quad \Delta x' = \gamma(\Delta x - v\Delta t)$$

$$\Delta y' = \Delta y \quad , \quad \Delta z' = \Delta z$$

Note that $\Delta t = t_2 - t_1$, $\Delta t' = t_2' - t_1'$, $\Delta x = x_2 - x_1$, $\Delta x' = x_2' - x_1'$, etc. The inverse Lorentz transformation in terms of time intervals is:

$$\Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right) \quad , \quad \Delta x = \gamma(\Delta x' + v\Delta t')$$

$$\Delta y = \Delta y' \quad , \quad \Delta z = \Delta z'$$

Time Dilation and Length Contraction in the Lorentz Transformation

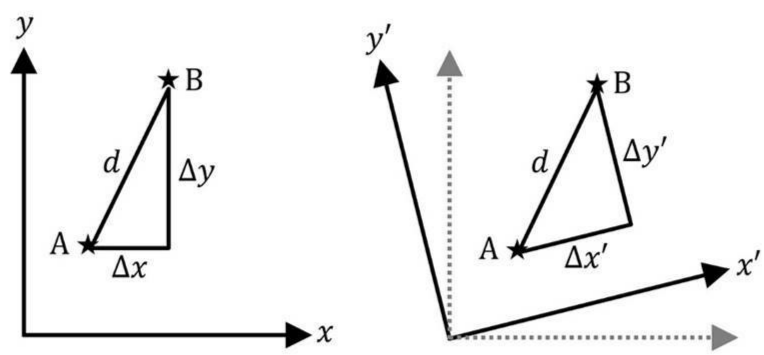
Time dilation and length contraction are actually built into the Lorentz transformation. Let's consider time dilation first. Suppose that observer O' measures the proper time:

$\Delta t' = \Delta t_0$. In order for observer O' to measure the proper time, O' must see the events occur at the same place, meaning that $\Delta x' = 0$. (Put another way, the clock used by O' is at rest relative to the events.) Plug $\Delta x' = 0$ into $\Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right)$ to get $\Delta t = \gamma\Delta t' = \gamma\Delta t_0$, which is the equation for **time dilation** from Chapter 2.

Now let's consider length contraction. Suppose that observer O′ measures the proper length: $\Delta x' = L_0$. For simplicity, suppose that one endpoint lies at $x'_A = 0$ and the other endpoint lies at $x'_B = L_0$. (Observe that $\Delta x' = x'_B - x'_A = L_0 - 0 = L_0$.) In order for O′ to measure the proper length, O′ must be at rest relative to the distance. This means that the coordinates $x'_A = 0$ and $x'_B = L_0$ aren't changing. Since observer O is moving relative to O′ (it's mutual: it's similarly true that O′ is moving relative to O), when we determine x_A and x_B we must make these measurements at the same time relative to O (since the values of x_A and x_B are constantly changing due to the relative motion between O and O′). It's convenient to measure x_A and x_B at $t_A = t_B = 0$. Since we originally setup our coordinate system with the two origins coincident at $t = 0$, this guarantees that $x_A = x'_A = 0$. According to the Lorentz transformation, $x'_B = \gamma(x_B - vt_B)$. At $t_A = t_B = 0$, this reduces to $x'_B = \gamma x_B$. Recall that $x'_B = L_0$, such that $L_0 = \gamma x_B$, which leads to $x_B = \frac{L_0}{\gamma}$. Since $x_A = 0$, observer O measures the length of the rod to be $\Delta x = x_B - x_A = \frac{L_0}{\gamma} - 0 = \frac{L_0}{\gamma} = L$, which is the equation for **length contraction** from Chapter 2. (At other times, you still get the same result for $\Delta x = x_B - x_A$, since the equations for x_A and x_B both involve the term $-\gamma vt$, which will cancel out when subtracting $x_B - x_A$.)

Rotational Invariance in Galilean Relativity

In Galilean relativity, the distance between two points is invariant. When we say **invariant**, we mean that it has the same value (it is a constant) regardless of which coordinate system you use to measure the distance. For example, consider the distance between points A and B in the diagram below. If you measure the horizontal distance between these points, $\Delta x = x_B - x_A$, and the vertical distance between these points, $\Delta y = y_B - y_A$, you could find the distance between A and B using the Pythagorean theorem: $\Delta x^2 + \Delta y^2 = d$. Now consider the rotated coordinate system (x', y'). The distance equals $\Delta x'^2 + \Delta y'^2 = d$ in the rotated coordinate system. In Galilean relativity, the distance between points A and B must be the same in either coordinate system: For example, the length of a rod doesn't change when it is rotated. We conclude that $\Delta x^2 + \Delta y^2 = \Delta x'^2 + \Delta y'^2$ in Galilean relativity. If we extend our argument to three dimensions, we find that $\Delta x^2 + \Delta y^2 + \Delta z^2 = \Delta x'^2 + \Delta y'^2 + \Delta z'^2$. This means that the quantity $\Delta x^2 + \Delta y^2 + \Delta z^2$ is a Galilean invariant.



TIP FOR READING EQUATIONS

Some equations and paragraphs with equations appear larger in landscape mode.

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

x PORTRAIT

(looks small on a small device)

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

✓ LANDSCAPE

Lorentz Invariance

As we have learned, different observers don't always measure the same lengths in special relativity. Due to length contraction, the quantity $\Delta x^2 + \Delta y^2 + \Delta z^2$ isn't a Lorentz invariant. It turns out that $c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$ is invariant under Lorentz transformations. Note that $c\Delta t$ is the distance that light travels in time Δt relative to O. To see that the quantity $c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$ is a Lorentz invariant, apply the Lorentz transformation equations.

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\gamma^2\left(\Delta t' + \frac{v}{c^2}\Delta x'\right)^2 - \gamma^2(\Delta x' + v\Delta t')^2 - \Delta y'^2 - \Delta z'^2$$

$$= c^2\gamma^2\left(\Delta t'^2 + 2\frac{v}{c^2}\Delta t'\Delta x' + \frac{v^2}{c^4}\Delta x'^2\right) - \gamma^2\left(\Delta x'^2 + 2v\Delta t'\Delta x' + v^2\Delta t'^2\right) - \Delta y'^2 - \Delta z'^2$$

$$= c^2\gamma^2\Delta t'^2 + 2v\gamma^2\Delta t'\Delta x' + \frac{v^2}{c^2}\gamma^2\Delta x'^2 - \gamma^2\Delta x'^2 - 2v\gamma^2\Delta t'\Delta x' - v^2\gamma^2\Delta t'^2 - \Delta y'^2 - \Delta z'^2$$

$$= \gamma^2\Delta t'^2(c^2 - v^2) + \gamma^2\Delta x'^2\left(\frac{v^2}{c^2} - 1\right) - \Delta y'^2 - \Delta z'^2$$

Note that $\gamma^2 = \frac{1}{1-\left(\frac{v}{c}\right)^2} = \frac{1}{1-\frac{v^2}{c^2}} = \frac{1}{\frac{c^2}{c^2}-\frac{v^2}{c^2}} = \frac{c^2}{c^2-v^2}$. We will use both $\gamma^2 = \frac{c^2}{c^2-v^2}$ and $\gamma^2 = \frac{1}{1-\frac{v^2}{c^2}}$.

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = \frac{c^2}{c^2 - v^2}\Delta t'^2(c^2 - v^2) + \frac{1}{1 - \frac{v^2}{c^2}}\Delta x'^2\left(\frac{v^2}{c^2} - 1\right) - \Delta y'^2 - \Delta z'^2$$

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2$$

Note that $1-\frac{v^2}{c^2}=-\left(\frac{v^2}{c^2}-1\right)$, such that $\frac{\left(\frac{v^2}{c^2}-1\right)}{1-\frac{v^2}{c^2}}=-1$. We see that $c^2\Delta t^2-\Delta x^2-\Delta y^2-\Delta z^2$ is indeed Lorentz invariant, since we have shown that it equals $c^2\Delta t'^2-\Delta x'^2-\Delta y'^2-\Delta z'^2$. Any inertial observer who measures $c^2\Delta t^2-\Delta x^2-\Delta y^2-\Delta z^2$ will get the same value. One way to interpret Lorentz invariance is that the speed of light is the same for all observers. Suppose that an observer proceeds to measure the speed of light in vacuum. During time interval Δt, the distance that light travels from (x_1, y_1, z_1) to (x_2, y_2, z_2) is $\Delta x^2+\Delta y^2+\Delta z^2$, such that the speed of light equals $c=\frac{\sqrt{\Delta x^2+\Delta y^2+\Delta z^2}}{\Delta t}$. Square both sides to get $c^2=\frac{\Delta x^2+\Delta y^2+\Delta z^2}{\Delta t^2}$. Multiply by Δt^2 and subtract $\Delta x^2+\Delta y^2+\Delta z^2$ to get $c^2\Delta t^2-\Delta x^2-\Delta y^2-\Delta z^2=0$ (note that this only equals zero when measuring the speed of light). Lorentz invariance shows that this expression will equal zero for any inertial observer, which equates to saying that the speed of light is the same for all inertial observers.

Tensor Notation

All advanced work in relativity is done using tensor notation. The reason for adopting tensor notation has to do with convenience. For example, the equation $R^{\sigma}_{\mu\nu\rho} = \partial_\nu \Gamma^{\sigma}_{\mu\rho} - \partial_\rho \Gamma^{\sigma}_{\mu\nu}$ looks like a single equation when written with tensor notation, but would instead involve writing 256 separate equations if you insist upon writing the same thing without using tensor notation. Thus, anyone who does much work in relativity quickly learns the value of tensor notation.

Lowercase Greek letters (like α, μ, or ν) are used to represent **spacetime** indices, where 0 corresponds to ct and 1-3 correspond to the **spatial** coordinates (x, y, z). In relativity, there are two types of indices:

- **contravariant** indices are up, like A^μ and $g^{\mu\nu}$.
- **covariant** indices are down, like A_μ and $g_{\mu\nu}$.

The notation A^μ could refer to A^0 (the time component of A), A^1 (the x-component of A), A^2 (the y-component of A), or A^3 (the z-component of A). A first-rank tensor, like A^μ or B_α, is called a 4-vector. A **4-vector** is the generalization of a vector to spacetime. A 4-vector has 4 components: one time component and three spatial components. There are two kinds of 4-vectors: **contravariant** with the index up (like A^μ) and **covariant** with the index down (like A_μ).

A particularly useful 4-vector is the **position-time 4-vector**. The contravariant and covariant position-time 4-vectors have the following components:

$$\{x^\mu\} = \{ct, x, y, z\} \quad , \quad \{x_\mu\} = \{ct, -x, -y, -z\}$$

As we will see, the relative minus signs between time and space for the covariant 4-vector help to construct Lorentz invariance.

It is so common in relativity for equations to include sums that almost all books and papers on the subject follow **Einstein summation notation**, which means that summation over an index is **implied** when the same index is **repeated** (in a term). For example, $A^\mu B_\mu$ implies a sum over the index μ:

$$A^\mu B_\mu = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$$

In contrast, the expression $A^\mu B_\nu$ doesn't imply any summation because none of its indices are repeated. The expression $A^\mu B_\nu$ could be $A^1 B_3$, $A^2 B_2$, or any of 16 possibilities. Unless the values of the indices are given, you have no way of knowing which. Compare with $A^\mu B_\mu$, which equals a sum of four specific terms (according to the equation above).

With Einstein summation notation, the contravariant and covariant forms of the position-time 4-vectors can be used to express **Lorentz invariance** concisely as:

$$x^{\mu}x_{\mu} = x'^{\mu}x'_{\mu}$$

Since there is an implied summation over the index μ on both sides, the above equation is shorthand for:

$$x^0x_0 + x^1x_1 + x^2x_2 + x^3x_3 = x'^0x'_0 + x'^1x'_1 + x'^2x'_2 + x'^3x'_3$$

Plugging in the values for each component, the above equation becomes:

$$c^2t^2 - x^2 - y^2 - z^2 = c^2t'^2 - x'^2 - y'^2 - z'^2$$

The above equation expresses Lorentz invariance, but it's more concise to write $x^{\mu}x_{\mu} = x'^{\mu}x'_{\mu}$.

The **Lorentz transformation** may be expressed with tensor notation as:

$$x'^{\mu} = \Lambda^{\mu}_{\nu}x^{\nu}$$

where the components Λ^{μ}_{ν} of the second-rank tensor Λ can be arranged in a matrix as:

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

What may look like one equation, $x'^{\mu} = \Lambda^{\mu}_{\nu}x^{\nu}$, is actually a set of four equations: a different equation for each possible value of the index μ. On the right-hand side of the equation, there is an implied sum over the repeated index ν. For example, when $\mu = 1$, the equation becomes $x'^0 = \Lambda^0_{\nu}x^{\nu} = \Lambda^0_0x^0 + \Lambda^0_1x^1 + \Lambda^0_2x^2 + \Lambda^0_3x^3$. Recall that $x'^0 = ct'$, $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$. We know this because $\{x^{\mu}\} = \{ct, x, y, z\}$. The equation for x'^0 is $ct' = \gamma ct - \gamma\beta x + 0y + 0z$ or $ct' = \gamma ct - \gamma\beta x$. Using $\beta = \frac{v}{c}$, we get $ct' = \gamma ct - \gamma\frac{v}{c}x$ or $t' = \gamma\left(t - \frac{v}{c^2}x\right)$.

Notation

Latin indices (like i, j, and k) represent **spatial** indices (1, 2, 3), whereas **Greek** indices (like μ, ν, and ρ) represent **spacetime** indices (0, 1, 2, 3). For example, the quantity x_i could be x_1, x_2, or x_3 (but not x_0), whereas the quantity x_μ could be x_0, x_1, x_2, or x_3. The distinction between **contravariant** and **covariant** indices is only made in the context of spacetime indices (represented by Greek letters).

The Kronecker Delta

The **Kronecker delta** symbol, δ_{ij} (not to be confused with the delta function, which means something else entirely), equals 1 when the indices are the same and 0 otherwise:

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

For example, $\delta_{12} = 0$ and $\delta_{22} = 1$. When an expression involves a Kronecker delta, the only nonzero terms arise when the Kronecker delta's indices are equal. This means that we may effectively remove the Kronecker delta symbol and replace one of its indices with the other. For example, consider the expression $A_i\delta_{ij}$. Whenever $i \neq j$, we get zero, such that the only nonzero term is $A_i\delta_{ij} = A_j$. As another example, note that $A_iB_j\delta_{ij} = A_iB_i$ (which represents the dot product between two vectors, $A_1B_1 + A_2B_2 + A_3B_3$). The expression $A_iB_j\delta_{ij}$ involves two implied sums (one over i and another over j), which is the sum of 9 terms. However, 6 of these terms are zero, since the Kronecker delta equals zero when $i \neq j$. The three nonzero terms are included in the single implied sum A_iB_i. Going from $A_iB_j\delta_{ij}$ to A_iB_i, we removed the Kronecker detla and replaced j with i.

The Levi-Civita Symbol

The three-dimensional **Levi-Civita** symbol is defined as follows:

$$\epsilon_{ijk} = \begin{cases} 1 \text{ for } \epsilon_{123}, \epsilon_{231}, \text{ and } \epsilon_{312} \\ -1 \text{ for } \epsilon_{132}, \epsilon_{213}, \text{ and } \epsilon_{321} \\ 0 \text{ if any two indices are equal} \end{cases}$$

Note that the Levi-Civita symbol changes sign with the interchange of any two indices. For example, $\epsilon_{132} = -\epsilon_{123}$ and $\epsilon_{321} = -\epsilon_{312}$. One example of where the Levi-Civita symbols is useful is in forming the vector product. For example, if $\vec{\mathrm{C}} = \vec{\mathrm{A}} \times \vec{\mathrm{B}}$, then $C_k = \epsilon_{ijk} A_i B_j$, with an implied double sum over the repeated indices i and j. Although this double sum includes 9 terms, only 2 of these terms are nonzero. For example, for $C_3 = \epsilon_{ij3} A_i B_j$, the only nonzero terms arise when $i = 1$ and $j = 2$ or when $i = 2$ and $j = 1$: $C_3 = \epsilon_{123} A_1 B_2 + \epsilon_{213} A_2 B_1 = (1) A_1 B_2 + (-1) A_2 B_1 = A_1 B_2 - A_2 B_1 = A_x B_y - A_y B_x = C_z$. Note that ϵ_{ij3} equals zero if i or j equals 3 (or if i equals j).

There is also a four-dimensional **Levi-Civita** symbol:

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} -1 \text{ for even permutations of } \epsilon^{0123} \\ 1 \text{ for odd permutations of } \epsilon^{0123} \\ 0 \text{ if any two indices are equal} \end{cases} \qquad \epsilon_{\mu\nu\rho\sigma} = \begin{cases} 1 \text{ for even permutations of } \epsilon_{0123} \\ -1 \text{ for odd permutations of } \epsilon_{0123} \\ 0 \text{ if any two indices are equal} \end{cases}$$

An even permutation is obtained by swapping indices an even number of times. For example, ϵ^{0312} can be formed from ϵ^{0123} with exactly two swaps: First, interchange the 2 and 3 to get ϵ^{0132} and then interchange the 1 and 3 to get ϵ^{0312}. An odd permutation involves swapping indices an odd number of times. For example, ϵ^{0132} is a single swap different from ϵ^{0123}.

It's correct in both three and four dimensions to count the number of swaps. Note that it's **incorrect** in four dimensions to describe this as "cyclic order." Although ϵ_{123}, ϵ_{213}, and ϵ_{312} follow a cyclic order in three dimensions (1 before 2 before 3 before 1 before 2 etc.), in four dimensions (that is, spacetime) the Levi-Civita symbol **doesn't** follow a cyclic order.

The Metric Tensor

The metric tensor is fundamental to tensor notation for both special and general relativity. In special relativity, the **metric tensor** is defined as:

$$g^{\mu\nu} = \begin{cases} 1 \text{ if } \mu = \nu = 0 \\ -1 \text{ if } \mu = \nu > 0 \\ 0 \text{ if } \mu \neq \nu \end{cases}$$

The nonzero elements are $g^{00} = 1$, $g^{11} = -1$, $g^{22} = -1$, and $g^{33} = -1$. If both indices are lowered instead of raised, the elements of the metric tensor are the same:

$$g_{\mu\nu} = \begin{cases} 1 \text{ if } \mu = \nu = 0 \\ -1 \text{ if } \mu = \nu > 0 \\ 0 \text{ if } \mu \neq \nu \end{cases}$$

However, if one index is raised and the other is lowered, you instead get the four-dimensional Kronecker delta: $g^{\mu}_{\nu} = \delta^{\mu}_{\nu}$, which equals 1 if $\mu = \nu$ and 0 if $\mu \neq \nu$. The metric tensor can be used to relate the covariant and contravariant forms of the position-time 4-vector:

$$x_\mu = g_{\mu\nu} x^\nu \quad , \quad x^\mu = g^{\mu\nu} x_\nu$$

For example, $g_{\mu\nu} x^\nu$ includes an implied sum over the index ν: 3 of the 4 terms will be zero. The only nonzero term arises when $\mu = \nu$. Recall that the components of the position-time 4-vector are $\{x^\mu\} = \{ct, x, y, z\}$ and $\{x_\mu\} = \{ct, -x, -y, -z\}$. What $g_{\mu\nu} x^\nu$ effectively does is change the sign of the spatial components only, effectively changing the contravariant form into the covariant form (while $g^{\mu\nu} x_\nu$ does vice-versa). The covariant and contravariant forms of any spacetime 4-vector can be similarly related via the metric tensor:

$$A_\mu = g_{\mu\nu} A^\nu \quad , \quad A^\mu = g^{\mu\nu} A_\nu$$

Lorentz invariance can be expressed using the metric tensor as follows:

$$g_{\mu\nu} x^\mu x^\nu = g_{\mu\nu} x'^\mu x'^\nu$$

The Scalar Product between 4-Vectors

The four-dimensional **scalar product** between two 4-vectors is:

$$A^\mu B_\mu = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$$

Since $B^\mu = g^{\mu\nu} B_\nu$, it follows that $B^0 = B_0$, $B^1 = -B_1$, $B^2 = -B_2$, and $B^3 = -B_3$, such that:

$$A^\mu B_\mu = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3 = A^0 B^0 - \vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$$

Note that $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$ is the usual three-dimensional scalar product between ordinary vectors in 3D space: $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}} = A_i B_i = A_1 B_1 + A_2 B_2 + A_3 B_3$.

If we let in $B = A$, we get the scalar product of a 4-vector with itself:

$$A^\mu A_\mu = (A^0)^2 - \vec{\mathbf{A}} \cdot \vec{\mathbf{A}}$$

Note that $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}} = (A^1)^2 + (A^2)^2 + (A^3)^2$ is the square of the magnitude: $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}} = \left\|\vec{\mathbf{A}}\right\|^2$ in 3D space. (What can get confusing is that books on vector algebra commonly write $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}}$ as A^2, yet books on relativity commonly write $A^\mu A_\mu$ as A^2, but the quantities $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}}$ and $A^\mu A_\mu$ aren't the same. One way around this is to write $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}} = \left\|\vec{\mathbf{A}}\right\|^2$ versus $A^\mu A_\mu = A^2$, or at least write $\vec{\mathbf{A}}^2$ with an arrow for $\left\|\vec{\mathbf{A}}\right\|^2$ so that you can distinguish between the 3D and 4D quantities.)

In 3D space, $\vec{\mathbf{A}} \cdot \vec{\mathbf{A}} = \left\|\vec{\mathbf{A}}\right\|^2$ is always positive (except for the null vector, in which case it is zero). However, in 4D spacetime, $A^\mu A_\mu = A^2$ can actually be negative.

- When $A^\mu A_\mu = A^2 > 0$, the 4-vector is considered to be **time-like**.
- When $A^\mu A_\mu = A^2 < 0$, the 4-vector is considered to be **space-like**.
- When $A^\mu A_\mu = A^2 = 0$, the 4-vector is considered to be **light-like**.

Note: For books or instructors that define A_0 to be negative and A_i to be positive (contrary to the convention adopted in this book), the above conditions are reversed.

Tensor Relations and Identities

The following relations and identities are sometimes handy when working with tensors:

$$\delta_{ij}\epsilon_{ijk} = 0$$

$$\epsilon_{ijk}\epsilon_{ijk} = 6$$

$$\epsilon_{ijm}\epsilon_{ijn} = 2\delta_{mn}$$

$$\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$$

$$g^{\mu\nu}g_{\mu\nu} = 4$$

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta} = -6\delta^{\alpha}_{\beta}$$

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}\right)$$

Symbols and SI Units

Symbol	Name	SI Units
v	relative speed between two observers	m/s
c	speed of light in vacuum	m/s
t	time measured by O	s
t'	time measured by O′	s
Δt	time interval measured by O	s
$\Delta t'$	time interval measured by O′	s
(x, y, z)	position coordinates measured by O	m
(x', y', z')	position coordinates measured by O′	m

$\Delta x, \Delta y, \Delta z$	$x_2 - x_1, y_2 - y_1, z_2 - z_1$	m
$\Delta x', \Delta y', \Delta z'$	$x'_2 - x'_1, y'_2 - y'_1, z'_2 - z'_1$	m
(x^0, x^1, x^2, x^3)	(ct, x, y, z)	m
(x_0, x_1, x_2, x_3)	$(ct, -x, -y, -z)$	m
β	fraction of the speed of light	unitless
γ	time dilation factor	unitless
$i, j, k, m, n, p, q \cdots$	spatial indices (1,2,3)	N/A
$\mu, \nu, \rho, \sigma, \alpha, \beta \cdots$	spacetime indices (0,1,2,3)	N/A

x_i	a spatial coordinate: x_1, x_2, or x_3	m
x^μ	a contravariant spacetime coordinate (a component of the position-time 4-vector): x^0, x^1, x^2, or x^3	m
x_μ	a covariant spacetime coordinate (a component of the position-time 4-vector): x_0, x_1, x_2, or x_3	m
x'^μ	a contravariant spacetime coordinate measured by O'	m
x'_μ	a covariant spacetime coordinate measured by O'	m

A_i	a spatial component of a 3D vector: A_1, A_2, or A_3	it depends
A^μ	a contravariant component of a 4D spacetime vector: A^0, A^1, A^2, or A^3	it depends
A_μ	a covariant component of a 4D spacetime vector: A_0, A_1, A_2, or A_3	it depends
$\vec{\mathrm{A}}$	a vector in 3D space	it depends
Λ^μ_ν	a component of the Lorentz transformation matrix	unitless
δ_{ij}	the Kronecker delta in 3D space	unitless

$\epsilon^{\mu\nu\rho\sigma}$	the 4D Levi-Civita symbol with contravariant indices	unitless
$\epsilon_{\mu\nu\rho\sigma}$	the 4D Levi-Civita symbol with covariant indices	unitless
$g^{\mu\nu}$	a component of the metric tensor with contravariant indices	unitless
$g_{\mu\nu}$	a component of the metric tensor with covariant indices	unitless

$\vec{\mathrm{A}} \cdot \vec{\mathrm{B}}$	the scalar product between two 3D vectors	it depends
$A^{\mu} B_{\mu}$	the scalar product between two 4-vectors	it depends
$\lVert\vec{\mathrm{A}}\rVert$	the magnitude of a 3D vector	it depends

Note: The following Greek letters are uppercase lambda (Λ), lowercase delta (δ), a variation of lowercase epsilon (ϵ), lowercase mu (μ), lowercase nu (ν), lowercase rho (ρ), lowercase sigma (σ), lowercase alpha (α), and lowercase beta (β).

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s

Strategy for Problems Involving the Lorentz Transformation

To solve a problem that involves the Lorentz transformation, follow these steps:

- Setup a coordinate system with $+x$ along the relative velocity. Identify observers O and O′, where O′ moves with constant velocity $\vec{v}$ relative to O. Apply the equations for the Lorentz transformation.

$$t' = \gamma\left(t - \frac{v}{c^2}x\right) \quad , \quad x' = \gamma(x - vt) \quad , \quad y' = y \quad , \quad z' = z$$

$$t = \gamma\left(t' + \frac{v}{c^2}x'\right) \quad , \quad x = \gamma(x' + vt') \quad , \quad y = y' \quad , \quad z = z'$$

$$\Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right) \quad , \quad \Delta x' = \gamma(\Delta x - v\Delta t) \quad , \quad \Delta y' = \Delta y \quad , \quad \Delta z' = \Delta z$$

$$\Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right) \quad , \quad \Delta x = \gamma(\Delta x' + v\Delta t') \quad , \quad \Delta y = \Delta y' \quad , \quad \Delta z = \Delta z'$$

$$\beta = \frac{v}{c} \quad , \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$$

- To determine whether or not a specified quantity is Lorentz invariant, use the Lorentz transformation equations to convert from unprimed coordinates to primed coordinates (or vice-versa) and compare with the original quantity. For example, the quantity $c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$ is Lorentz invariant because the above equations can be used to transform it to $c^2\Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2$ (after applying algebra, as we did earlier in this chapter), which has the same form as the original quantity.
- If the problem involves tensor notation, see the following strategy.
- If the problem involves a Galilean transformation, use the equations for a Galilean transformation:

$$t' = t \quad , \quad x' = x - vt \quad , \quad y' = y \quad , \quad z' = z$$

- If a problem involves relative velocity, see Chapter 4. If a problem involves relativistic momentum or energy, see Chapter 5.

TIP FOR READING EQUATIONS

Some equations and paragraphs with equations appear larger in landscape mode.

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

x PORTRAIT

(looks small on a small device)

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

✓ LANDSCAPE

Strategy for Problems Involving Tensor Notation

To solve a problem that involves tensor notation, follow these steps:

- Are the indices (in subscripts or superscripts) **Latin** like i, j, and k, or are they **Greek** like μ, ν, and ρ? Latin indices may equal 1, 2, or 3 (corresponding to x, y, and z). Greek indices may equal 0, 1, 2, or 3 (where 0 corresponds to ct).
- Are any indices repeated in the same term? If so, there is an **implied summation** over any **repeated index.** For example, the quantity $g^{\mu\nu}A_\nu$ involves a sum over ν. This means that $g^{\mu\nu}A_\nu = g^{\mu 0}A_0 + g^{\mu 1}A_1 + g^{\mu 2}A_2 + g^{\mu 3}A_3$.
- Note that $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$, whereas $x_0 = ct$, $x_1 = -x$, $x_2 = -y$, and $x_3 = -z$.

- Is there a **Kronecker delta**, like δ_{ij} or δ^{μ}_{ν} (but possibly with different letters for the indices)? If so, the Kronecker delta equals 1 if the two indices are equal and 0 if they are different. For example, $\delta_{33} = 1$ and $\delta_{21} = 0$. If a Kronecker delta is multiplying other tensors, you may remove the Kronecker delta provided that you replace one of its indices with the other. For example, $A_i\delta_{ij} = A_j$ and $x_i x_j \delta_{ij} = x_i x_i = x_j x_j$.
- Is there a **Levi-Civita** symbol, like ϵ_{ijk}, $\epsilon^{\mu\nu\rho\sigma}$, or $\epsilon_{\mu\nu\rho\sigma}$ (but possibly with different letters for the indices)? If so, note that $\epsilon_{123} = 1$, $\epsilon^{0123} = -1$, and $\epsilon_{0123} = 1$. An even number of swaps of the indices results in the same value, whereas an odd number of swaps changes the sign. For example, $\epsilon_{213} = -1$ and $\epsilon_{231} = 1$. Note that if an index is repeated, the Levi-Civita symbol equals zero, as in $\epsilon_{232} = 0$.
- Is there a **metric tensor**, like $g^{\mu\nu}$ or $g_{\mu\nu}$? If so, note that $g^{00} = g_{00} = 1$, $g^{11} = g_{11} = -1$, $g^{22} = g_{22} = -1$, and $g^{33} = g_{33} = -1$. The off-diagonal elements are zero, like $g^{13} = g_{13} = 0$ and $g^{02} = g_{02} = 0$. If a metric tensor is multiplying a 4-vector with a matching index, you may remove the metric tensor provided that you raise or lower the index of the 4-vector (switching it from covariant to contravariant, or vice-versa) and replace the index of the 4-vector with the other index of the metric tensor. For example, $g_{\mu\nu}A^{\nu} = A_{\mu}$ and $g^{\mu\nu}x_{\mu} = x^{\nu}$.

- Note that the **scalar product** between 3D vectors is $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}} = A_i B_i = A_1 B_1 + A_2 B_2 + A_3 B_3$, whereas the scalar product between 4D spacetime vectors is $A^\mu B_\mu = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3 = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3 = A^0 B^0 - \vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$.
- To determine whether a quantity is space-like, time-like, or light-like, compute $A^2 = A^\mu A_\mu = (A^0)^2 - \vec{\mathbf{A}} \cdot \vec{\mathbf{A}}$. The quantity is time-like if $A^2 > 0$, space-like if $A^2 < 0$, and light-like if $A^2 = 0$ (using our notation, where A_0 is positive and A_i is negative).
- Is there an uppercase lambda, Λ? If so, the **Lorentz transformation** is given by $\Lambda_0^0 = \Lambda_1^1 = \gamma$, $\Lambda_1^0 = \Lambda_0^1 = -\gamma\beta$, $\Lambda_2^2 = \Lambda_3^3 = 1$, and all other elements equal zero.
- Are there primed coordinates, like x'^μ or x'_μ (but possibly with a different index)? If so, the **Lorentz transformation** is $x'^\mu = \Lambda_\nu^\mu x^\nu$. In tensor notation, Lorentz invariance is expressed as $x^\mu x_\mu = x'^\mu x'_\mu$.

- It may help to apply one or more of the following relations:

$$\delta_{ij}\epsilon_{ijk} = 0$$

$$\epsilon_{ijk}\epsilon_{ijk} = 6$$

$$\epsilon_{ijm}\epsilon_{ijn} = 2\delta_{mn}$$

$$\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$$

$$g^{\mu\nu}g_{\mu\nu} = 4$$

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta} = -6\delta_\beta^\alpha$$

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta_\sigma^\rho\delta_\beta^\alpha - \delta_\beta^\rho\delta_\sigma^\alpha\right)$$

Example: A 4-vector is defined to have the following components. Show whether or not this 4-vector is Lorentz invariant.

$$\{\eta^\mu\} = \{\gamma c, \gamma v_x, \gamma v_y, \gamma v_z\}$$

The above equation specifies the contravariant components of η.

$$\eta^0 = \gamma c \quad , \quad \eta^1 = \gamma v_x \quad , \quad \eta^2 = \gamma v_y \quad , \quad \eta^3 = \gamma v_z$$

Find the covariant components by negating the sings of the spatial components, according to $\eta_\mu = g_{\mu\nu}\eta^\nu$. For example, $\eta_1 = g_{1\nu}\eta^\nu = g_{10}\eta^0 + g_{11}\eta^1 + g_{12}\eta^2 + g_{13}\eta^3 = g_{11}\eta^1 = -\eta^1 = -\gamma v_x$ (since $g_{10} = g_{12} = g_{13} = 0$ and $g_{11} = -1$).

$$\eta_0 = \gamma c \quad , \quad \eta_1 = -\gamma v_x \quad , \quad \eta_2 = -\gamma v_y \quad , \quad \eta_3 = -\gamma v_z$$

Calculate $\eta^\mu \eta_\mu$.

$$\eta^\mu \eta_\mu = \eta^0\eta_0 + \eta^1\eta_1 + \eta^2\eta_2 + \eta^3\eta_3$$

$$= (\gamma c)(\gamma c) + (\gamma v_x)(-\gamma v_x) + (\gamma v_y)(-\gamma v_y) + (\gamma v_z)(-\gamma v_z)$$

$$= \gamma^2 c^2 - \gamma^2 v_x^2 - \gamma^2 v_z^2 - \gamma^2 v_z^2 = \gamma^2 c^2 - \gamma^2 (v_x^2 + v_z^2 + v_z^2)$$

Recall that the magnitude of a 3D vector is related to its components by $v^2 = v_x^2 + v_z^2 + v_z^2$.

$$\eta^\mu \eta_\mu = \gamma^2 c^2 - \gamma^2 v^2 = \gamma^2 (c^2 - v^2)$$

Note that we could have obtained the same result via $\eta^\mu \eta_\mu = (\eta^0)^2 - \vec{\eta} \cdot \vec{\eta} = (\eta^0)^2 - (\eta^1)^2 - (\eta^2)^2 - (\eta^3)^2 = \gamma^2 c^2 - \gamma^2 v_x^2 - \gamma^2 v_z^2 - \gamma^2 v_z^2$. Recall that $\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}$. Square both sides to get $\gamma^2 = \frac{1}{1-\frac{v^2}{c^2}}$. Multiply the numerator and denominator by c^2 to get $\gamma^2 = \frac{c^2}{c^2-v^2}$.

$$\eta^\mu \eta_\mu = \gamma^2 (c^2 - v^2) = \frac{c^2}{c^2 - v^2}(c^2 - v^2) = \boxed{c^2}$$

Since $\eta^\mu \eta_\mu = c^2$ and since the speed of light is the same for all observers, it follows that $\eta^\mu \eta_\mu$ is Lorentz invariant, meaning that $\eta'^\mu \eta'_\mu = \eta^\mu \eta_\mu = c^2$.

Example: For each expression below, indicate how many sums are implied (if any) and how many terms the expression represents.

(A) $x_i x_j$

The expression $x_i x_j$ represents a $\boxed{\text{single term}}$. This expression doesn't have any implied sums because neither index (i or j) is repeated. Since the values of i and j are unknown, $x_i x_j$ could be $x_1 x_3$, $x_2 x_2$, or any of 9 possibilities (since i and j each equal 1, 2, or 3).

(B) $x_i x_i$

The expression $x_i x_i$ has $\boxed{1}$ implied sum (since the index i is repeated) with $\boxed{3}$ terms: $x_i x_i = x_1 x_1 + x_2 x_2 + x_3 x_3 = x_1^2 + x_2^2 + x_3^2$.

(C) $\delta_{ab}\delta_{ab}$

The expression $\delta_{ab}\delta_{ab}$ has $\boxed{2}$ implied sums (since the indices a and b are each repeated) with $\boxed{9}$ terms. This is every permutation where a equals 1, 2, or 3 and where b equals 1, 2, or 3: $\delta_{ab}\delta_{ab} = \delta_{11}\delta_{11} + \delta_{12}\delta_{12} + \delta_{13}\delta_{13} + \delta_{21}\delta_{21} + \cdots$.

(D) $A^\mu B^\nu C_\mu D_\nu$

The expression $A^\mu B^\nu C_\mu D_\nu$ has $\boxed{2}$ implied sums (since the indices μ and ν are each repeated) with $\boxed{16}$ terms. This is every permutation where μ equals 0, 1, 2, or 3 and where ν equals 0, 1, 2, or 3: $A^\mu B^\nu C_\mu D_\nu = A^0 B^0 C_0 D_0 + A^0 B^1 C_0 D_1 + A^0 B^2 C_0 D_2 + A^0 B^3 C_0 D_3 + A^1 B^0 C_1 D_0 + \cdots$. Recall that Greek indices (like μ or ν) equal 0, 1, 2, or 3, whereas Latin indices (like i or j) equal 1, 2, or 3. The Greek indices are spacetime indices, whereas the Latin indices are purely spatial indices (no time).

Example: Rewrite the equation below longhand.

$$x^\mu = g^{\mu\nu} x_\nu$$

Note that there is an implied summation over the repeated index ν, but that there isn't any summation over the index μ. Why not? Because ν is repeated in the same term, whereas μ isn't repeated in the same term (μ appears only once on two different terms). What appears to be one equation is actually 4 separate equations: one equation for each possible value of μ.

$$\boxed{x^0 = g^{0\nu} x_\nu = g^{00} x_0 + g^{01} x_1 + g^{02} x_2 + g^{03} x_3}$$

$$\boxed{x^1 = g^{1\nu} x_\nu = g^{10} x_0 + g^{11} x_1 + g^{12} x_2 + g^{13} x_3}$$

$$\boxed{x^2 = g^{2\nu} x_\nu = g^{20} x_0 + g^{21} x_1 + g^{22} x_2 + g^{23} x_3}$$

$$\boxed{x^3 = g^{3\nu} x_\nu = g^{30} x_0 + g^{31} x_1 + g^{32} x_2 + g^{33} x_3}$$

We could continue our solution, simplifying it by plugging in $g^{00} = 1$, $g^{01} = 0$, $g^{02} = 0$, etc. However, the problem merely asked us to write the expression out longhand, which we have already done.

Example: What, precisely, does each symbol below equal or represent?

(A) x^0

If you think the answer is one (that x^0 equals 1), in the context of relativity you're wrong! Why? Because this is a superscript, not an exponent. If you wanted to raise x to the power of zero, since $x^1 = x$ (where this is a superscript, not an exponent), it would look like this instead: $(x^1)^0 = 1$. Once you realize that x^0 involves a superscript instead of an exponent, this problem is easy. This is the first component of the contravariant form of the position-time 4-vector: $\boxed{x^0 = ct}$.

(B) δ_{22}

The Kronecker delta equals one when both indices are the same and zero otherwise. In this case, the indices are the same: $\boxed{\delta_{22} = 1}$.

(C) ϵ_{1023}

This is the 4D Levi-Civita symbol with covariant indices. With covariant indices, $\epsilon_{0123} = 1$. Note that ϵ_{1023} is an odd permutation of ϵ_{0123}: One swap of indices (interchange the 0 and 1) turns ϵ_{0123} into ϵ_{1023}. Since ϵ_{1023} is an odd permutation, the sign changes: $\boxed{\epsilon_{1023} = -1}$.

(D) $A_i\delta_{2i}$

There is an implied sum over the repeated index i:

$$A_i\delta_{2i} = A_1\delta_{21} + A_2\delta_{22} + A_3\delta_{23}$$

The Kronecker delta equals one when both indices are the same and zero otherwise. Thus, $\delta_{21} = 0$, $\delta_{22} = 1$, and $\delta_{23} = 0$. The only surviving term is:

$$A_i\delta_{2i} = A_2\delta_{22} = \boxed{A_2}$$

(E) $B_\mu g^{\mu 3}$

There is an implied sum over the repeated index μ. Since μ is a Greek letter, this sum is over the values 0, 1, 2, and 3 (in contrast to part D, which used a Latin index i instead of a Greek index).

$$B_\mu g^{\mu 3} = B_0 g^{03} + B_1 g^{13} + B_2 g^{23} + B_3 g^{33}$$

Recall that g represents the metric tensor: $g^{03} = g^{13} = g^{23} = 0$ and $g^{33} = -1$. The only nonzero term is:

$$B_\mu g^{\mu 3} = B_3(-1) = -B_3$$

The covariant and contravariant components of a 4-vector are related by changing the signs of the spatial indices. Thus, $B_3 = -B^3$.

$$B_\mu g^{\mu 3} = \boxed{B^3}$$

It's helpful to remember that the metric tensor has the effect of raising or lowering indices.

Example: Simplify the following expressions.

(A) $K_a \delta_{ab}$

Since there is an implied sum over the repeated index a, and since δ_{ab} is only nonzero when $a = b$, it follows that $K_a \delta_{ab} = \boxed{K_b}$. (If you write out the sum longhand, it is $K_a \delta_{ab} = K_1 \delta_{1b} + K_2 \delta_{2b} + K_3 \delta_{3b}$. What we're saying is that two of these Kronecker deltas equal zero and the other one equals one. Whatever the value of b is, that's the only term that is nonzero, such that the sum equals K_b.)

(B) $g^{\mu\nu} x_\mu x_\nu$

Apply the rule that $g^{\mu\nu} x_\mu = x^\nu$. The metric tensor has the effect of raising or lowering indices. Thus, $g^{\mu\nu} x_\mu x_\nu = \boxed{x^\nu x_\nu}$. Note that we could alternatively write $g^{\mu\nu} x_\nu = x^\mu$ to get $g^{\mu\nu} x_\mu x_\nu = x^\mu x_\mu$, which is the same answer. It doesn't matter whether we write $x^\mu x_\mu$ or $x^\nu x_\nu$, since the index is being summed over: $x^\mu x_\mu$ and $x^\nu x_\nu$ both equal $x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3$. (If you write out $g^{\mu\nu} x_\mu x_\nu$ longhand, you get 16 terms from the 2 double sums. 12 of these 16 terms are zero, since $g^{\mu\nu}$ equals zero whenever $\mu \neq \nu$. The nonzero terms are $g^{00} x_0 x_0 + g^{11} x_1 x_1 + g^{22} x_2 x_2 + g^{33} x_3 x_3 = x_0^2 - x_1^2 - x_2^2 - x_3^2$, since $g^{00} = 1$ and $g^{11} = g^{22} = g^{33} = -1$. Note that this is the same answer as $x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3$ because $x_0 = x^0$, $x_1 = -x^1$, $x_2 = -x^2$, and $x_3 = -x^3$.)

(C) $\epsilon_{ijk}\epsilon_{mnk}A_iB_jA_mB_n$

Apply the following identity:

$$\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$$

Compare $\epsilon_{ijk}\epsilon_{mnk}$ with $\epsilon_{imn}\epsilon_{ipq}$. They share the same structure: one index (k) is repeated in $\epsilon_{ijk}\epsilon_{mnk}$, and one index ($i$) is repeated in $\epsilon_{imn}\epsilon_{ipq}$. However, the two expressions are not quite identical. The main difference is the position of the repeated index. In $\epsilon_{ijk}\epsilon_{mnk}$, the k is repeated in the third position of each Levi-Civita symbol, but in $\epsilon_{imn}\epsilon_{ipq}$, the i is repeated in the first position of each Levi-Civita symbol. That's an easy fix: $\epsilon_{ijk} = -\epsilon_{ikj} = \epsilon_{kij}$ and $\epsilon_{mnk} = -\epsilon_{mkn} = \epsilon_{kmn}$ because the Levi-Civita symbol changes sign with the interchange of any two indices. The given expression can thus be written

$$\epsilon_{kij}\epsilon_{kmn}A_iB_jA_mB_n$$

Now $\epsilon_{kij}\epsilon_{kmn}$ has the same structure as $\epsilon_{imn}\epsilon_{ipq}$. Examine the identity $\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$. What does $\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$ really say? It says, either $m = p$ and $n = q$ or else $m = q$ and $n = p$, where these are the 2nd and 3rd indices of ϵ_{imn} and ϵ_{ipq}. If we apply the same logic to $\epsilon_{kij}\epsilon_{kmn}$, we get

$$\epsilon_{kij}\epsilon_{kmn} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}$$

Note that i, j, m, and n are the 2nd and 3rd indices of ϵ_{kij} and ϵ_{kmn}. Substitute the above identity into the original expression.

$$\epsilon_{ijk}\epsilon_{mnk}A_iB_jA_mB_n = (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})A_iB_jA_mB_n$$
$$= \delta_{im}\delta_{jn}A_iB_jA_mB_n - \delta_{in}\delta_{jm}A_iB_jA_mB_n$$

Each term includes a quadruple sum over the repeated indices i, j, m, and n. Most of the 81 terms are zero. The only nonzero terms arise when $i = m$ and $j = n$ (for the first expression), and $i = n$ and $j = m$ (for the second expression). We're applying the following rule: We may remove the Kronecker delta provided that we replace one of its indices with the other index in the remaining expression.

$$\epsilon_{ijk}\epsilon_{mnk}A_iB_jA_mB_n = A_mB_nA_mB_n - A_nB_mA_mB_n$$
$$= (A_mA_m)(B_nB_n) - (A_mB_m)(A_nB_n) = \boxed{A^2B^2 - 2\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}}$$

In the last step, we used the fact that $A^2 = A_x^2 + A_z^2 + A_z^2 = A_mA_m$, $B^2 = B_x^2 + B_z^2 + B_z^2 = B_nB_n$, and $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}} = A_xB_x + A_yB_y + A_zB_z = A_mB_m = A_nB_n$.

Example: Show that $\delta_{ij}\epsilon_{ijk} = 0$.

This is a simple logic problem:

- δ_{ij} equals 1 if $i = j$ and 0 if $i \neq j$.
- ϵ_{ijk} equals 0 if $i = j$ (or $j = k$ or $k = i$) and 1 or -1 if $i \neq j \neq k$.

If $i = j$, then $\epsilon_{ijk} = 0$ due to repeated indices. However, if $i \neq j$, then $\delta_{ij} = 0$ due to unequal indices. No matter what, either δ_{ij} or ϵ_{ijk} must be zero. Therefore, their product must equal zero: $\delta_{ij}\epsilon_{ijk} = \boxed{0}$.

Example: Show that $\delta^{\mu}_{\mu} = 4$.

There is an implied sum over the repeated index μ. Since μ is a Greek letter, the sum is over 0, 1, 2, and 3. For the Kronecker delta, $\delta^0_0 = \delta^1_1 = \delta^2_2 = \delta^3_3 = 1$ (unlike $g^{\mu\nu}$ and $g_{\mu\nu}$).

$$\delta^{\mu}_{\mu} = \delta^0_0 + \delta^1_1 + \delta^2_2 + \delta^3_3 = 1 + 1 + 1 + 1 = \boxed{4}$$

Example: Show that $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\alpha} = -24$.

There are 4 implied sums over the 4 repeated indices. Since the indices are Greek letters, the sums are over 0, 1, 2, and 3. Most of the 256 terms equal zero since the Levi-Civita symbol is zero when indices are repeated (like $\epsilon^{0112} = 0$ and $\epsilon_{1233} = 0$). There are 24 nonzero terms corresponding to the 24 permutations of the digits 0123:

0123	0132	0213	0231	0312	0321
1023	1032	1203	1230	1302	1320
2013	2031	2103	2130	2301	2310
3012	3021	3102	3120	3201	3210

The expression $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\alpha} = \epsilon^{0123}\epsilon_{0123} + \epsilon^{0132}\epsilon_{0132} + \epsilon^{0213}\epsilon_{0213} + \cdots$ includes 24 terms on the right-hand side, corresponding to the digits in the table above. Since the 4D Levi-Civita symbol defines $\epsilon^{0123} = -1$ and $\epsilon_{0123} = 1$, for any of the digits listed above, the two Levi-Civita symbols will have opposite sign: Both symbols involve the same number of interchanges of their indices (since their indices are in the same order). For example, ϵ^{0213} and ϵ_{0213} involve two interchanges of indices compared to ϵ^{0123} and ϵ_{0123} (first swap the 13 to get 0231 from 0213, then swap the 21 to get 0123 from 0213). The first few terms of the sum include $\epsilon^{0123}\epsilon_{0123} = (-1)(1) = -1$, $\epsilon^{0132}\epsilon_{0132} = (1)(-1) = -1$, $\epsilon^{0213}\epsilon_{0213} = (1)(-1) = -1$, and $\epsilon^{0231}\epsilon_{0231} = (-1)(1) = -1$. All 24 terms in the sum equal -1.

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\alpha} = \epsilon^{0123}\epsilon_{0123} + \epsilon^{0132}\epsilon_{0132} + \epsilon^{0213}\epsilon_{0213} + \cdots = (-1)(24) = \boxed{-24}$$

Example: Show that $\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$.

There is only one index (i) that is repeated in the same term ($\epsilon_{imn}\epsilon_{ipq}$), so there is one sum on the left-hand side. There are no sums on the right-hand side (since none of the indices are repeated in the same terms: note that $\delta_{mp}\delta_{nq}$ and $\delta_{mq}\delta_{np}$ are two different terms). Since i is a Latin index, the sum on the left-hand side is over 1, 2, and 3.

$$\epsilon_{imn}\epsilon_{ipq} = \epsilon_{1mn}\epsilon_{1pq} + \epsilon_{2mn}\epsilon_{2pq} + \epsilon_{3mn}\epsilon_{3pq}$$

The problem is that there are 81 possible permutations of the indices m, n, p, and q in the expression $\epsilon_{1mn}\epsilon_{1pq}$, since each index can be a 1, 2, or 3. We need to show that $\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$ for all possible values of m, n, p, and q. We will approach this systematically:

- Case 1: What if m, n, p, and q all have different values?
- Case 2: $m = n$ or $p = q$.
- Case 3: $m = p$ or $n = q$.
- Case 4: $m = q$ or $n = p$.

Case 1: What if m, n, p, and q all have different values?

- Actually, this case isn't possible. Why not? Because m, n, p, and q can each equal 1, 2, or 3. There are only 3 values to choose from, but there are 4 indices. Therefore, at least one pair of indices must be equal. For example, if $m = 1$, $n = 2$, and $p = 3$, the index q must be equal to m, n, or p. All 4 indices can't have different values.

Case 2: $m = n$ or $p = q$.

- If $m = n$, then $\epsilon_{imn} = 0$. This checks out: $\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np} = \delta_{np}\delta_{nq} - \delta_{nq}\delta_{np} = 0$.
- If $p = q$, then $\epsilon_{ipq} = 0$. This checks out: $\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np} = \delta_{mq}\delta_{nq} - \delta_{mq}\delta_{nq} = 0$.
- Thus, if $m = n$ or $p = q$, both sides of $\epsilon_{imn}\epsilon_{ipq} = \delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$ equal zero.

Case 3: $m = p$ or $n = q$.

- If $m = p$, then $\epsilon_{1mn}\epsilon_{1pq} + \epsilon_{2mn}\epsilon_{2pq} + \epsilon_{3mn}\epsilon_{3pq} = \epsilon_{1pn}\epsilon_{1pq} + \epsilon_{2pn}\epsilon_{2pq} + \epsilon_{3pn}\epsilon_{3pq}$. (In this case, don't sum over the repeated index.)
  - If $p = 1$, then $\epsilon_{1pn}\epsilon_{1pq} + \epsilon_{2pn}\epsilon_{2pq} + \epsilon_{3pn}\epsilon_{3pq} = 0 + \epsilon_{21n}\epsilon_{21q} + \epsilon_{31n}\epsilon_{31q}$. This equals 0 unless $n = q$. (In this case, don't sum over the repeated index.)
  - If $p = 2$, then $\epsilon_{1pn}\epsilon_{1pq} + \epsilon_{2pn}\epsilon_{2pq} + \epsilon_{3pn}\epsilon_{3pq} = \epsilon_{12n}\epsilon_{12q} + 0 + \epsilon_{32n}\epsilon_{32q}$. This equals 0 unless $n = q$. (In this case, don't sum over the repeated index.)
  - If $p = 3$, then $\epsilon_{1pn}\epsilon_{1pq} + \epsilon_{2pn}\epsilon_{2pq} + \epsilon_{3pn}\epsilon_{3pq} = \epsilon_{13n}\epsilon_{13q} + \epsilon_{23n}\epsilon_{23q} + 0$. This equals 0 unless $n = q$. (In this case, don't sum over the repeated index.)
- We have just shown that if $m = p$, the only way $\epsilon_{imn}\epsilon_{ipq}$ won't equal 0 is if $n = q$. It can similarly be shown that if $n = q$, the only way $\epsilon_{imn}\epsilon_{ipq}$ won't equal 0 is if $m = p$.

- If both $m = p$ and $n = q$, then $\epsilon_{imn}\epsilon_{ipq} = (\epsilon_{1pq})^2 + (\epsilon_{2pq})^2 + (\epsilon_{3pq})^2$. We don't need to worry about $p = q$ because we covered that in Case 2. For $p \neq q$, $(\epsilon_{1pq})^2 + (\epsilon_{2pq})^2 + (\epsilon_{3pq})^2 = 1$ because exactly one term will be nonzero.
- Here is a recap: If both $m = p$ and $n = q$, then $\epsilon_{imn}\epsilon_{ipq}$ equals 1, but if only one of the equalities $m = p$ or $n = q$ is true, then $\epsilon_{imn}\epsilon_{ipq}$ equals 0.
- Now let's examine $\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$. The first term, $\delta_{mp}\delta_{nq}$, only equals 1 if both $m = p$ and $n = q$, which is in exact agreement with the previous bullet point. The second term, $\delta_{mq}\delta_{np}$, only equals 1 if both $m = q$ and $n = p$. We don't need to worry about this second term in Case 3, since if $m = p$ or $n = q$ (as required in Case 3), then the condition $m = q$ and $n = p$ would turn this into Case 2 instead of Case 3.

Case 4: $m = q$ or $n = p$.

- If $m = q$, then $\epsilon_{1mn}\epsilon_{1pq} + \epsilon_{2mn}\epsilon_{2pq} + \epsilon_{3mn}\epsilon_{3pq} = \epsilon_{1qn}\epsilon_{1pq} + \epsilon_{2qn}\epsilon_{2pq} + \epsilon_{3qn}\epsilon_{3pq}$. (In this case, don't sum over the repeated index.)
    - If $q = 1$, then $\epsilon_{1qn}\epsilon_{1pq} + \epsilon_{2qn}\epsilon_{2pq} + \epsilon_{3qn}\epsilon_{3pq} = 0 + \epsilon_{21n}\epsilon_{2p1} + \epsilon_{31n}\epsilon_{3p1}$. This equals 0 unless $n = p$. (In this case, don't sum over the repeated index.)
    - If $q = 2$, then $\epsilon_{1qn}\epsilon_{1pq} + \epsilon_{2qn}\epsilon_{2pq} + \epsilon_{3qn}\epsilon_{3pq} = \epsilon_{12n}\epsilon_{1p2} + 0 + \epsilon_{32n}\epsilon_{3p2}$. This equals 0 unless $n = p$. (In this case, don't sum over the repeated index.)
    - If $q = 3$, then $\epsilon_{1qn}\epsilon_{1pq} + \epsilon_{2qn}\epsilon_{2pq} + \epsilon_{3qn}\epsilon_{3pq} = \epsilon_{13n}\epsilon_{1p3} + \epsilon_{23n}\epsilon_{2p3} + 0$. This equals 0 unless $n = p$. (In this case, don't sum over the repeated index.)
- We have just shown that if $m = q$, the only way $\epsilon_{imn}\epsilon_{ipq}$ won't equal 0 is if $n = p$. It can similarly be shown that if $n = p$, the only way $\epsilon_{imn}\epsilon_{ipq}$ won't equal 0 is if $m = q$.

- If both $m = q$ and $n = p$, then $\epsilon_{imn}\epsilon_{ipq} = -(\epsilon_{1pq})^2 - (\epsilon_{2pq})^2 - (\epsilon_{3pq})^2$. Why? First set $\epsilon_{imn} = \epsilon_{iqp}$ (but don't sum over p or q because they aren't repeated in $\epsilon_{imn}\epsilon_{ipq}$). Next, swap indices to write $\epsilon_{iqp} = -\epsilon_{ipq}$. Thus, $\epsilon_{imn}\epsilon_{ipq}$ equals $-\epsilon_{ipq}$ times ϵ_{ipq}, which is the same as $-(\epsilon_{ipq})^2$ with a sum over i (because i is repeated in $\epsilon_{imn}\epsilon_{ipq}$). We don't need to worry about $p = q$ because we covered that in Case 2. For $p \neq q$, $-(\epsilon_{1pq})^2 - (\epsilon_{2pq})^2 - (\epsilon_{3pq})^2 = -1$ because exactly one term will be nonzero.
- Here is a recap: If both $m = q$ and $n = p$, then $\epsilon_{imn}\epsilon_{ipq}$ equals -1, but if only one of the equalities $m = q$ or $n = p$ is true, then $\epsilon_{imn}\epsilon_{ipq}$ equals 0.
- Now let's examine $\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}$. The second term, $\delta_{mq}\delta_{np}$, equals -1 if both $m = q$ and $n = p$, which is in exact agreement with the previous bullet point. The first term, $\delta_{mp}\delta_{nq}$, equals 1 if both $m = p$ and $n = q$. We don't need to worry about this first term in Case 4, since if $m = q$ or $n = p$ (as required in Case 4), then the condition $m = p$ and $n = q$ would turn this into Case 2 instead of Case 4.

Combining all four cases together, we have shown that:

$$\epsilon_{imn}\epsilon_{ipq} = \boxed{\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}}$$

Chapter 3 Problems

1. Relativistic 4-momentum is defined to have the following components. Show whether or not relativistic 4-momentum is Lorentz invariant.

$$\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$$

$$E = \gamma m_0 c^2 \quad , \quad \vec{\mathbf{p}} = \gamma m_0 \vec{\boldsymbol{u}}$$

Notes: m_0 represents rest mass, $\vec{\boldsymbol{u}}$ represents velocity, and $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$.

Want help? Check the solution at the end of the chapter.

Answer: 1. It is.

2. For each expression below, indicate how many sums are implied (if any) and how many terms the expression represents.

(A) $P_k Q_k$ (B) $A_m B_n \delta_{mn}$

(C) $\epsilon_{abc} x_a x_b$ (D) $g^{\alpha\beta} A_\alpha B_\beta$

(E) $\epsilon^{\kappa\lambda\sigma\tau} x_\kappa x_\lambda x_\sigma$ (F) $F^\mu G^\nu H_\mu$

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) 1, 3 (B) 2, 9

(C) 2, 9 (D) 2, 16 (E) 3, 64 (F) 1, 4

3. Rewrite each equation below longhand.

$$\text{(A) } A_j = B_i \delta_{ij}$$

$$\text{(B) } T = U^\theta V_\theta$$

$$\text{(C) } L_{ij} = M_i N_j$$

Want help? Check the solution at the end of the chapter.

Answers: The answers to this problem can be found at the end of the chapter.

4. For each symbol below, indicate precisely what it equals or represents.

(A) x^0	(B) $x^1 x_2$	(C) x^3
(D) δ_{11}	(E) δ_{23}	(F) δ_{33}
(G) ϵ_{112}	(H) ϵ_{312}	(I) ϵ_{213}
(J) ϵ^{1320}	(K) ϵ_{3021}	(L) ϵ_{1231}
(M) g^{00}	(N) g_{11}	(O) g^{22}
(P) g_{31}	(Q) $x_k \delta_{2k}$	(R) $x^\gamma g_{0\gamma}$

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) ct (B) $-xy$ (C) z

(D) 1 (E) 0 (F) 1 (G) 0 (H) 1 (I) –1

(J) –1 (K) 1 (L) 0 (M) 1 (N) –1

(O) –1 (P) 0 (Q) y (R) ct

5. Simplify each expression below.

(A) $A_i B_m \delta_{ij} \delta_{mn}$

(B) $\epsilon_{ijk} \epsilon_{ijn} \delta_{kn}$

(C) $g^{\mu\alpha} g^{\nu\beta} g_{\mu\nu} P_\alpha Q_\beta$

(D) $\epsilon^{\rho\sigma\mu\nu} \epsilon_{\alpha\beta\mu\nu} K_\rho L_\sigma K^\alpha L^\beta$

Want help? Check the solution at the end of the chapter.

Answers:

5. (A) $A_j B_n$ (B) 6 (C) $P^\mu Q_\mu$

(D) $-2K_\alpha K^\alpha L_\beta L^\beta + 2L_\alpha K^\alpha K_\beta L^\beta$

6. Derive each of the following relations.

(A) $\delta_{ii} = 3$

(B) $\epsilon_{ijk}\epsilon_{ijk} = 6$

(C) $\epsilon_{ijm}\epsilon_{ijn} = 2\delta_{mn}$

(D) $g^{\mu\nu} g_{\mu\nu} = 4$

(E) $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta} = -6\delta^{\alpha}_{\beta}$

(F) $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}\right)$

Want help? Check the solution at the end of the chapter.

Answers: The answers to this problem can be found at the end of the chapter.

Solutions to Chapter 3

1. According to the given equation, the contravariant components of the 4-momentum are:

$$p^0 = \frac{E}{c} \quad , \quad p^1 = p_x \quad , \quad p^2 = p_y \quad , \quad p^3 = p_z$$

Find the covariant components by negating the sings of the spatial components, according to $p_\mu = g_{\mu\nu} p^\nu$. For example, $p_1 = g_{1\nu} p^\nu = g_{10} p^0 + g_{11} p^1 + g_{12} p^2 + g_{13} p^3 = g_{11} p^1 = -p^1 = -p_x$ (since $g_{10} = g_{12} = g_{13} = 0$ and $g_{11} = -1$).

$$p_0 = \frac{E}{c} \quad , \quad p_1 = -p_x \quad , \quad p_2 = -p_y \quad , \quad p_3 = -p_z$$

Calculate $p^\mu p_\mu$.

$$p^\mu p_\mu = p^0 p_0 + p^1 p_1 + p^2 p_2 + p^3 p_3 = \frac{E^2}{c^2} + p_x(-p_x) + p_y(-p_y) + p_z(-p_z)$$
$$= \frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = \frac{E^2}{c^2} - p^2$$

Recall that the magnitude of a 3D vector is related to its components by $p^2 = p_x^2 + p_z^2 + p_z^2$. Plug the given equations, $E = \gamma m_0 c^2$ and $\vec{\mathbf{p}} = \gamma m_0 \vec{\boldsymbol{u}}$, into the previous equation.

$$p^\mu p_\mu = \frac{E^2}{c^2} - p^2 = \frac{(\gamma m_0 c^2)^2}{c^2} - (\gamma m_0 u)^2 = \frac{\gamma^2 m_0^2 c^4}{c^2} - \gamma^2 m_0^2 u^2$$
$$= \gamma^2 m_0^2 c^2 - \gamma^2 m_0^2 u^2 = \gamma^2 m_0^2 (c^2 - u^2)$$

Recall that $\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{u}{c}\right)^2}}$. Square both sides to get $\gamma^2 = \frac{1}{1-\frac{u^2}{c^2}}$. Multiply the numerator and denominator by c^2 to get $\gamma^2 = \frac{c^2}{c^2-u^2}$.

$$p^\mu p_\mu = \gamma^2 m_0^2 (c^2 - u^2) = \frac{c^2}{c^2 - u^2} m_0^2 (c^2 - u^2) = \boxed{m_0^2 c^2}$$

Since $p^\mu p_\mu = m_0^2 c^2$ and since the speed of light is the same for all observers, it follows that $p^\mu p_\mu$ is Lorentz invariant, meaning that $p'^\mu p'_\mu = p^\mu p_\mu = m_0^2 c^2$. Note that rest-mass ($m_0$) is a constant (it is the same for all observers), unlike relativistic mass (see Chapter 5).

2. (A) The expression $P_k Q_k$ has $\boxed{1}$ implied sum (since the index k is repeated) with $\boxed{3}$ terms:

$$P_k Q_k = P_1 Q_1 + P_2 Q_2 + P_3 Q_3$$

(B) The expression $A_m B_n \delta_{mn}$ has $\boxed{2}$ implied sums (since the indices m and n are each repeated) with $3^2 = \boxed{9}$ terms. This is every permutation where m equals 1, 2, or 3 and where n equals 1, 2, or 3:

$$A_m B_n \delta_{mn} = A_1 B_1 \delta_{11} + A_1 B_2 \delta_{12} + A_1 B_3 \delta_{13} + A_2 B_1 \delta_{21} + A_2 B_2 \delta_{22} + A_2 B_3 \delta_{23} + A_3 B_1 \delta_{31} + A_3 B_2 \delta_{32} + A_3 B_3 \delta_{33}$$

(C) The expression $\epsilon_{abc} x_a x_b$ has $\boxed{2}$ implied sums (since the indices a and b are each repeated) with $3^2 = \boxed{9}$ terms. This is every permutation where a equals 1, 2, or 3 and where b equals 1, 2, or 3:

$$\epsilon_{abc} x_a x_b = \epsilon_{11c} x_1 x_1 + \epsilon_{12c} x_1 x_2 + \epsilon_{13c} x_1 x_3 + \epsilon_{21c} x_2 x_1 + \epsilon_{22c} x_2 x_2 + \epsilon_{23c} x_2 x_3 + \epsilon_{31c} x_3 x_1 + \epsilon_{32c} x_3 x_2 + \epsilon_{33c} x_3 x_3$$

Note that the value of c is unknown: It could be 1, 2, or 3 (but it is **not** summed over because the index c isn't repeated in the same term).

(D) The expression $g^{\alpha\beta} A_\alpha B_\beta$ has $\boxed{2}$ implied sums (since the indices α and β are each repeated) with $4^2 = \boxed{16}$ terms. This is every permutation where α equals 0, 1, 2, or 3 and where β equals 0, 1, 2, or 3:

$$g^{\alpha\beta} A_\alpha B_\beta = g^{00} A_0 B_0 + g^{01} A_0 B_1 + g^{02} A_0 B_2 + g^{03} A_0 B_3 + g^{10} A_1 B_0 + g^{11} A_1 B_1 + g^{12} A_1 B_2 + g^{13} A_1 B_3 + g^{20} A_2 B_0 + g^{21} A_2 B_1 + g^{22} A_2 B_2 + g^{23} A_2 B_3 + g^{30} A_3 B_0 + g^{31} A_3 B_1 + g^{32} A_3 B_2 + g^{33} A_3 B_3$$

Recall that Greek indices (like μ or ν) equal 0, 1, 2, or 3, whereas Latin indices (like i or j) equal 1, 2, or 3. The Greek indices are spacetime indices, whereas the Latin indices are purely spatial indices (no time).

(E) The expression $\epsilon^{\kappa\lambda\sigma\tau}x_\kappa x_\lambda x_\sigma$ has $\boxed{3}$ implied sums (since the indices κ, λ, and σ are each repeated) with $4^3 = \boxed{64}$ terms. This is every permutation where κ equals 0, 1, 2, or 3, where λ equals 0, 1, 2, or 3, and where σ equals 0, 1, 2, or 3:

$$\begin{aligned}\epsilon^{\kappa\lambda\sigma\tau}x_\kappa x_\lambda x_\sigma = {} & \epsilon^{000\tau}x_0x_0x_0 + \epsilon^{001\tau}x_0x_0x_1 + \epsilon^{002\tau}x_0x_0x_2 + \epsilon^{003\tau}x_0x_0x_3 + \epsilon^{010\tau}x_0x_1x_0 \\
& + \epsilon^{011\tau}x_0x_1x_1 + \epsilon^{012\tau}x_0x_1x_2 + \epsilon^{013\tau}x_0x_1x_3 + \epsilon^{020\tau}x_0x_2x_0 + \epsilon^{021\tau}x_0x_2x_1 \\
& + \epsilon^{022\tau}x_0x_2x_2 + \epsilon^{023\tau}x_0x_2x_3 + \epsilon^{030\tau}x_0x_3x_0 + \epsilon^{031\tau}x_0x_3x_1 + \epsilon^{032\tau}x_0x_3x_2 \\
& + \epsilon^{033\tau}x_0x_3x_3 + \epsilon^{100\tau}x_1x_0x_0 + \epsilon^{101\tau}x_1x_0x_1 + \epsilon^{102\tau}x_1x_0x_2 + \epsilon^{103\tau}x_1x_0x_3 \\
& + \epsilon^{110\tau}x_1x_1x_0 + \epsilon^{111\tau}x_1x_1x_1 + \epsilon^{112\tau}x_1x_1x_2 + \epsilon^{113\tau}x_1x_1x_3 + \epsilon^{120\tau}x_1x_2x_0 \\
& + \epsilon^{121\tau}x_1x_2x_1 + \epsilon^{122\tau}x_1x_2x_2 + \epsilon^{123\tau}x_1x_2x_3 + \epsilon^{130\tau}x_1x_3x_0 + \epsilon^{131\tau}x_1x_3x_1 \\
& + \epsilon^{132\tau}x_1x_3x_2 + \epsilon^{133\tau}x_1x_3x_3 + \epsilon^{200\tau}x_2x_0x_0 + \epsilon^{201\tau}x_2x_0x_1 + \epsilon^{202\tau}x_2x_0x_2 \\
& + \epsilon^{203\tau}x_2x_0x_3 + \epsilon^{210\tau}x_2x_1x_0 + \epsilon^{211\tau}x_2x_1x_1 + \epsilon^{212\tau}x_2x_1x_2 + \epsilon^{213\tau}x_2x_1x_3 \\
& + \epsilon^{220\tau}x_2x_2x_0 + \epsilon^{221\tau}x_2x_2x_1 + \epsilon^{222\tau}x_2x_2x_2 + \epsilon^{223\tau}x_2x_2x_3 + \epsilon^{230\tau}x_2x_3x_0 \\
& + \epsilon^{231\tau}x_2x_3x_1 + \epsilon^{232\tau}x_2x_3x_2 + \epsilon^{233\tau}x_2x_3x_3 + \epsilon^{300\tau}x_3x_0x_0 + \epsilon^{301\tau}x_3x_0x_1 \\
& + \epsilon^{302\tau}x_3x_0x_2 + \epsilon^{303\tau}x_3x_0x_3 + \epsilon^{310\tau}x_3x_1x_0 + \epsilon^{311\tau}x_3x_1x_1 + \epsilon^{312\tau}x_3x_1x_2 \\
& + \epsilon^{313\tau}x_3x_1x_3 + \epsilon^{320\tau}x_3x_2x_0 + \epsilon^{321\tau}x_3x_2x_1 + \epsilon^{322\tau}x_3x_2x_2 + \epsilon^{323\tau}x_3x_2x_3 \\
& + \epsilon^{330\tau}x_3x_3x_0 + \epsilon^{331\tau}x_3x_3x_1 + \epsilon^{332\tau}x_3x_3x_2 + \epsilon^{333\tau}x_3x_3x_3\end{aligned}$$

Note that the value of τ is unknown: It could be 0, 1, 2, or 3 (but it is **not** summed over because the index τ isn't repeated in the same term).

(F) The expression $F^\mu G^\nu H_\mu$ has $\boxed{1}$ implied sum (since the index μ is repeated) with $\boxed{4}$ terms:

$$F^\mu G^\nu H_\mu = F^0 G^\nu H_0 + F^1 G^\nu H_1 + F^2 G^\nu H_2 + F^3 G^\nu H_3$$

Note that the value of ν is unknown: It could be 0, 1, 2, or 3 (but it is **not** summed over because the index ν isn't repeated in the same term).

3. (A) Note that there is an implied summation over the repeated index i, but that there isn't any summation over the index j. Why not? Because i is repeated in the same term, whereas j isn't repeated in the same term (j appears only once in two different terms). What appears to be one equation is actually 3 separate equations: one equation for each possible value of j.

$$\boxed{A_1 = B_i\delta_{i1} = B_1\delta_{11} + B_2\delta_{21} + B_3\delta_{31}}$$

$$\boxed{A_2 = B_i\delta_{i2} = B_1\delta_{12} + B_2\delta_{22} + B_3\delta_{32}}$$

$$\boxed{A_3 = B_i\delta_{i3} = B_1\delta_{13} + B_2\delta_{23} + B_3\delta_{33}}$$

We could continue our solution, simplifying it by plugging in $\delta_{11} = 1$, $\delta_{21} = 0$, $\delta_{31} = 0$, etc. However, the problem merely asked us to write the expression out longhand.

(B) There is an implied summation over the repeated index θ. Since θ is the lowercase Greek letter theta, the sum is over 0, 1, 2, and 3 (this is a spacetime index).

$$\boxed{T = U^\theta V_\theta = U^0 V_0 + U^1 V_1 + U^2 V_2 + U^3 V_3}$$

(C) There are **no** implied summations because no index is repeated in the same term. (The indices i and j appear in different terms on separate sides of the equal sign.) What appears to be one equation is actually 9 separate equations: one for each possible permutation of i and j.

$$\boxed{L_{11} = M_1 N_1 \quad , \quad L_{12} = M_1 N_2 \quad , \quad L_{13} = M_1 N_3}$$

$$\boxed{L_{21} = M_2 N_1 \quad , \quad L_{22} = M_2 N_2 \quad , \quad L_{23} = M_2 N_3}$$

$$\boxed{L_{31} = M_3 N_1 \quad , \quad L_{32} = M_3 N_2 \quad , \quad L_{33} = M_3 N_3}$$

4. (A) $x^0 = \boxed{ct}$. This isn't an exponent of zero. This is the first component of the contravariant form of the position-time 4-vector.

(B) $x^1 x_2 = (x)(-y) = \boxed{-xy}$. The covariant form of the position-time 4-vector has minus signs in its spatial components, unlike the contravariant form. Thus, $x_2 = -y$ whereas $x^1 = x$.

(C) $x^3 = \boxed{z}$. This component has a contravariant index.

(D) $\delta_{11} = \boxed{1}$ since both indices are equal.

(E) $\delta_{23} = \boxed{0}$ since the indices are different.

(F) $\delta_{33} = \boxed{1}$ since both indices are equal.

(G) $\epsilon_{112} = \boxed{0}$ since an index is repeated.

(H) $\epsilon_{312} = \boxed{1}$. Start with $\epsilon_{123} = 1$. Swap the 2 and 3 to get $\epsilon_{132} = -1$. Then swap the 1 and 3 to get $\epsilon_{312} = 1$.

(I) $\epsilon_{213} = \boxed{-1}$. Start with $\epsilon_{123} = 1$. Swap the 1 and 2 to get $\epsilon_{213} = -1$. (Each time you swap the order of two adjacent indices in the Levi-Civita symbol, change the sign.)

(J) $\epsilon^{1320} = \boxed{-1}$. Start with $\epsilon^{0123} = -1$ (the 4D form with contravariant indices has a minus sign). You can go from ϵ^{0123} to ϵ^{1320} with 4 swaps (see below). If you change the sign four times, it will remain negative.

$$\epsilon^{0123} = -\epsilon^{0132} = \epsilon^{1032} = -\epsilon^{1302} = \epsilon^{1320} = -1$$

(K) $\epsilon_{3021} = \boxed{1}$. Start with $\epsilon_{0123} = 1$ (the 4D form with covariant indices has a plus sign). You can go from ϵ_{0123} to ϵ_{3021} with 4 swaps (see below). If you change the sign four times, it will remain positive.

$$\epsilon_{0123} = -\epsilon_{0132} = \epsilon_{0312} = -\epsilon_{3012} = \epsilon_{3021} = 1$$

(L) $\epsilon_{1231} = \boxed{0}$ since an index is repeated.

(M) $g^{00} = \boxed{1}$. This is the (0,0) element of the metric tensor.

(N) $g_{11} = \boxed{-1}$. The spatial components of the metric tensor with equal indices are negative.

(O) $g^{22} = \boxed{-1}$. For the metric tensor, whether both indices are contravariant or both indices are covariant makes no difference. (However, if it is mixed, like g_2^2, it equals the Kronecker delta instead.)

(P) $g_{31} = \boxed{0}$ since the indices are different.

(Q) $x_k \delta_{2k} = x_2 = \boxed{y}$. There is an implied sum over the repeated index k, and only $\delta_{22} = 1$ yields a nonzero term. You may remove a Kronecker delta provided that you replace one index with the other in the remaining expression (in this case, we replaced the k with 2 in x_k). Note: Since k isn't a Greek letter, this **<u>isn't</u>** a covariant 4-vector, which is why $x_2 = +y$.

(R) $x^\gamma g_{0\gamma} = x_0 = \boxed{ct}$. There is an implied sum over the repeated index γ. Since γ is a Greek letter (lowercase gamma), the sum is over 0, 1, 2, and 3. Only $g_{00} = 1$ yields a nonzero term. The metric tensor has the effect of raising or lowering indices. You may remove the metric tensor provided that you replace one index with the other in the remaining expression, and also raise or lower its index, as appropriate (in this case, x^γ was lowered to x_0, replacing γ with 0.)

5. (A) A_i, B_m, δ_{ij}, and δ_{mn} are just numbers. For example, δ_{ij} either equals 1 or it equals 0. Since these are all just numbers, order doesn't matter when we multiply them. Therefore, $A_i B_m \delta_{ij} \delta_{mn} = A_i \delta_{ij} B_m \delta_{mn}$. There are 2 implied sums (over i and m). The only nonzero terms arise when $i = j$ and $m = n$. Thus, $A_i \delta_{ij} = A_j$ and $B_m \delta_{mn} = B_n$. (You may remove a Kronecker delta provided that you replace one index with the other in the remaining expression).

$$A_i B_m \delta_{ij} \delta_{mn} = \boxed{A_j B_n}$$

(B) There are 4 implied sums (over i, j, k, and n). Whenever $k \neq n$, that term will equal zero according to the Kronecker delta (δ_{kn}). For any nonzero term, k must equal n. This allows us to remove the Kronecker delta, provided that we replace k with n (or vice-versa) in the remaining expression.

$$\epsilon_{ijk} \epsilon_{ijn} \delta_{kn} = \epsilon_{ijk} \epsilon_{ijk} = \boxed{6}$$

In the last step, we applied the identity $\epsilon_{ijk} \epsilon_{ijk} = 6$ (see page 44). If you want to show that $\epsilon_{ijk} \epsilon_{ijk} = 6$, note that this involves a triple sum. The only nonzero terms arise when i, j, and k are all different. There are 6 permutations where that happens: $\epsilon_{ijk} \epsilon_{ijk} = (\epsilon_{123})^2 + (\epsilon_{132})^2 + (\epsilon_{213})^2 + (\epsilon_{231})^2 + (\epsilon_{312})^2 + (\epsilon_{321})^2 = 1^2 + (-1)^2 + (-1)^2 + 1^2 + 1^2 + (-1)^2 = 6$.

(C) There are 4 implied sums (over μ, α, ν, and β). Whenever $\mu \neq \alpha$, $\nu \neq \beta$, or $\mu \neq \nu$, that term will equal zero according to the corresponding metric tensor ($g^{\mu\alpha}$, $g^{\nu\beta}$, or $g_{\mu\nu}$), since off-diagonal elements of the metric tensor are zero (like g^{13}, g^{20}, and g_{21}). For any nonzero term, μ must equal α, ν must equal β, and μ must equal ν. This allows us to remove the metric tensors, provided that we replace α with μ, β with ν, and ν with μ (or vice-versa) in the remaining expression, and also provided that we raise or lower any index that we replace. The reason for raising or lowering the index is that the metric tensor has negative spatial components ($g^{11} = g_{11} = -1$, $g^{22} = g_{22} = -1$, and $g^{33} = g_{33} = -1$) and a positive time component ($g^{00} = g_{00} = 1$), while contravariant and covariant 4-vectors differ in the signs of their spatial components (for example, $x^3 = z$ whereas $x_3 = -z$, such that $x^3 = g^{33}x_3$).

$$g^{\mu\alpha}g^{\nu\beta}g_{\mu\nu}P_\alpha Q_\beta = g^{\mu\alpha}P_\alpha g^{\nu\beta}Q_\beta g_{\mu\nu} = P^\mu Q^\nu g_{\mu\nu} = \boxed{P^\mu Q_\mu}$$

We could alternatively express our answer as $P_\nu Q^\nu$, since $P^0Q_0 + P^1Q_1 + P^2Q_2 + P^3Q_3$ and $P_0Q^0 + P_1Q^1 + P_2Q^2 + P_3Q^3$ are both equal to $P^0Q^0 - P^1Q^1 - P^2Q^2 - P^3Q^3$.

(D) Apply the identity $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta^\rho_\sigma \delta^\alpha_\beta - \delta^\rho_\beta \delta^\alpha_\sigma\right)$, which can be found on the bottom of page 44. Compare the given Levi-Civita symbols, $\epsilon^{\rho\sigma\mu\nu}\epsilon_{\alpha\beta\mu\nu}$, to the Levi-Civita symbols in the identity, $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$. Although they have the same structure, the names and order of the indices differ. First of all, note that $\epsilon^{\rho\sigma\mu\nu}\epsilon_{\alpha\beta\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\alpha\beta}$ (where in each case the indices were reordered with 4 swaps). So we are really comparing $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\alpha\beta}$ to $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$. Now the first two indices are identical in each pair of Levi-Civita symbols. We don't really need to examine the last two indices closely and set up a correspondence. It's easier to examine the identity $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta^\rho_\sigma \delta^\alpha_\beta - \delta^\rho_\beta \delta^\alpha_\sigma\right)$ and understand the logic behind it. On the right-hand side, the first set of Kronecker deltas groups the corresponding third and fourth indices from the Levi-Civita symbols, while the second set of Kronecker deltas groups the third index with the fourth (and vice-versa) from the Levi-Civita symbols. If we apply that logic to the given Levi-Civita symbols, we get $\epsilon^{\rho\sigma\mu\nu}\epsilon_{\alpha\beta\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\alpha\beta} = -2\left(\delta^\rho_\alpha \delta^\sigma_\beta - \delta^\rho_\beta \delta^\sigma_\alpha\right)$.

$$\epsilon^{\rho\sigma\mu\nu}\epsilon_{\alpha\beta\mu\nu}K_\rho L_\sigma K^\alpha L^\beta = -2\left(\delta^\rho_\alpha \delta^\sigma_\beta - \delta^\rho_\beta \delta^\sigma_\alpha\right)K_\rho L_\sigma K^\alpha L^\beta$$
$$= -2\delta^\rho_\alpha \delta^\sigma_\beta K_\rho L_\sigma K^\alpha L^\beta + 2\delta^\rho_\beta \delta^\sigma_\alpha K_\rho L_\sigma K^\alpha L^\beta = -2K_\alpha L_\beta K^\alpha L^\beta + 2K_\beta L_\alpha K^\alpha L^\beta$$
$$= -2K_\alpha K^\alpha L_\beta L^\beta + 2L_\alpha K^\alpha K_\beta L^\beta$$

In the middle line, we applied the argument from Part B: We may remove a Kronecker delta, provided that we replace one of its indices with the other in the remaining expression. Thus, $\delta^\rho_\alpha K_\rho = K_\alpha$, $\delta^\sigma_\beta L_\sigma = L_\beta$, $\delta^\rho_\beta K_\rho = K_\beta$, and $\delta^\sigma_\alpha L_\sigma = L_\alpha$. (Unlike the metric tensor, the Kronecker delta neither raises nor lowers indices.)

6. (A) Taking the implied sum over a repeated index to apply even when the index is repeated within the subscripts of the same symbol, we get:

$$\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 1 + 1 + 1 = \boxed{3}$$

(B) There are 3 implied sums (over i, j, and k). However, most of the 64 terms are zero since the Levi-Civita symbol equals zero when two indices are the same (like $\epsilon_{112} = 0$ and $\epsilon_{212} = 0$). Nonzero terms only arise when all three indices are different. There are 6 nonzero terms:

$$\epsilon_{ijk}\epsilon_{ijk} = \epsilon_{123}\epsilon_{123} + \epsilon_{132}\epsilon_{132} + \epsilon_{213}\epsilon_{213} + \epsilon_{231}\epsilon_{231} + \epsilon_{312}\epsilon_{312} + \epsilon_{321}\epsilon_{321}$$

$$(1)(1) + (-1)(-1) + (-1)(-1) + (1)(1) + (1)(1) + (-1)(-1)$$

$$= 1 + 1 + 1 + 1 + 1 + 1 = \boxed{6}$$

(C) There are 2 implied sums (over i and j). There are no sums over m and n since these indices aren't repeated in the same term. The values of m and n are unknown.

$$\epsilon_{ijm}\epsilon_{ijn} = \epsilon_{11m}\epsilon_{11n} + \epsilon_{12m}\epsilon_{12n} + \epsilon_{13m}\epsilon_{13n} + \epsilon_{21m}\epsilon_{21n} + \epsilon_{22m}\epsilon_{22n} + \epsilon_{23m}\epsilon_{23n} + \epsilon_{31m}\epsilon_{31n} + \epsilon_{32m}\epsilon_{32n} + \epsilon_{33m}\epsilon_{33n}$$

The terms with two equal indices equal zero. For example, $\epsilon_{11m} = 0$.

$$\epsilon_{ijm}\epsilon_{ijn} = \epsilon_{12m}\epsilon_{12n} + \epsilon_{13m}\epsilon_{13n} + \epsilon_{21m}\epsilon_{21n} + \epsilon_{23m}\epsilon_{23n} + \epsilon_{31m}\epsilon_{31n} + \epsilon_{32m}\epsilon_{32n}$$

On the right-hand side, the values of i and j are different. (For example, ϵ_{12m} corresponds to $i = 1$ and $j = 2$.) If either m or n equals i or j, the term will be zero. There are two different cases to consider:

- If $m \neq n$, every term will be zero because there is no way for all 4 indices (i, j, m, and n) to all be different (since there are only three values—1, 2, and 3—to choose from). Therefore, when $m \neq n$, $\epsilon_{ijm}\epsilon_{ijn}$ must be zero.
- If $m = n$, the right-hand side of the above equation becomes
$$(\epsilon_{12m})^2 + (\epsilon_{13m})^2 + (\epsilon_{21m})^2 + (\epsilon_{23m})^2 + (\epsilon_{31m})^2 + (\epsilon_{32m})^2$$
No matter which value m equals, exactly two terms in the previous expression will be nonzero. For example, for $m = 1$, the nonzero terms are $(\epsilon_{23m})^2 + (\epsilon_{32m})^2$. Thus, when $m = n$, $\epsilon_{ijm}\epsilon_{ijn}$ must equal 2 since the two nonzero terms give $1^2 + (-1)^2 = 2$.

The Kronecker delta δ_{mn} equals zero when $m \neq n$ and equals 1 when $m = n$. If we multiply the Kronecker delta by 2, we can satisfy the two previous bullet points perfectly.

$$\epsilon_{ijm}\epsilon_{ijn} = \boxed{2\delta_{mn}}$$

(D) There are 2 implied sums (over μ and ν). Since μ and ν are lowercase Greek indices, the sums are over 0, 1, 2, and 3. However, 12 of the 16 terms are zero, since the metric tensor is only nonzero when the two indices are equal (for example, $g^{13} = 0$). The nonzero terms are:

$$g^{\mu\nu}g_{\mu\nu} = g^{00}g_{00} + g^{11}g_{11} + g^{22}g_{22} + g^{33}g_{33}$$

Recall that $g^{00} = g_{00} = 1$, $g^{11} = g_{11} = -1$, $g^{22} = g_{22} = -1$, and $g^{33} = g_{33} = -1$.

$$g^{\mu\nu}g_{\mu\nu} = (1)(1) + (-1)(-1) + (-1)(-1) + (-1)(-1) = 1 + 1 + 1 + 1 = \boxed{4}$$

(E) There are 3 implied sums (over μ, ν, and ρ). There are no sums over α and β since these indices aren't repeated in the same term. The values of α and β are unknown. Most of the terms are zero since the Levi-Civita symbol is zero whenever two indices are equal (for example, $\epsilon^{0131} = 0$ and $\epsilon_{1223} = 0$). The nonzero terms are:

$$\begin{aligned}\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta} = {} & \epsilon^{012\alpha}\epsilon_{012\beta} + \epsilon^{021\alpha}\epsilon_{021\beta} + \epsilon^{102\alpha}\epsilon_{102\beta} + \epsilon^{120\alpha}\epsilon_{120\beta} + \epsilon^{201\alpha}\epsilon_{201\beta} + \epsilon^{210\alpha}\epsilon_{210\beta} \\ & + \epsilon^{013\alpha}\epsilon_{013\beta} + \epsilon^{031\alpha}\epsilon_{031\beta} + \epsilon^{103\alpha}\epsilon_{103\beta} + \epsilon^{130\alpha}\epsilon_{130\beta} + \epsilon^{301\alpha}\epsilon_{301\beta} + \epsilon^{310\alpha}\epsilon_{310\beta} \\ & + \epsilon^{023\alpha}\epsilon_{023\beta} + \epsilon^{032\alpha}\epsilon_{032\beta} + \epsilon^{203\alpha}\epsilon_{203\beta} + \epsilon^{230\alpha}\epsilon_{230\beta} + \epsilon^{302\alpha}\epsilon_{302\beta} + \epsilon^{320\alpha}\epsilon_{320\beta} \\ & + \epsilon^{123\alpha}\epsilon_{123\beta} + \epsilon^{132\alpha}\epsilon_{132\beta} + \epsilon^{213\alpha}\epsilon_{213\beta} + \epsilon^{231\alpha}\epsilon_{231\beta} + \epsilon^{312\alpha}\epsilon_{312\beta} + \epsilon^{321\alpha}\epsilon_{321\beta}\end{aligned}$$

There are two different cases to consider:

- If $\alpha \neq \beta$, every term will be zero because there is no way for all 5 indices (μ, ν, ρ, α, and β) to all be different (since there are only 4 values—0, 1, 2, and 3—to choose from). Therefore, when $\alpha \neq \beta$, $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta}$ must be zero.
- If $\alpha = \beta$, exactly 6 terms in the previous expression will be nonzero. For example, for $\alpha = 1$, the nonzero terms are:

  $$\begin{aligned}&\epsilon^{0231}\epsilon_{0231} + \epsilon^{0321}\epsilon_{0321} + \epsilon^{2031}\epsilon_{2031} + \epsilon^{2301}\epsilon_{2301} + \epsilon^{3021}\epsilon_{3021} + \epsilon^{3201}\epsilon_{3201} \\ &= (-1)(1) + (1)(-1) + (1)(-1) + (-1)(1) + (-1)(1) + (1)(-1) \\ &= -1 - 1 - 1 - 1 - 1 - 1 = -6\end{aligned}$$

  Recall that $\epsilon^{0123} = -1$ and $\epsilon_{0123} = 1$, and that each changes sign with the swap of any two indices. The main idea is that every nonzero term will always equal -1 because $\epsilon^{\mu\nu\rho\alpha}$ and $\epsilon_{\mu\nu\rho\alpha}$ will always have opposite signs. Thus, when $\alpha = \beta$, $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta}$ must equal -6 since there will be 6 nonzero terms (each equal to -1).

The Kronecker delta δ^{α}_{β} equals zero when $\alpha \neq \beta$ and equals 1 when $\alpha = \beta$. If we multiply the Kronecker delta by -6, we can satisfy the two previous bullet points perfectly.

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\rho\beta} = \boxed{-6\delta^{\alpha}_{\beta}}$$

(F) There are two indices (μ and ν) that are repeated in the same term ($\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$), so there are two implied sums. Since μ and ν are lowercase Greek indices, the sums are over 0, 1, 2, and 3.

$$\begin{aligned}\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = {} & \epsilon^{01\rho\alpha}\epsilon_{01\sigma\beta} + \epsilon^{02\rho\alpha}\epsilon_{02\sigma\beta} + \epsilon^{03\rho\alpha}\epsilon_{03\sigma\beta} + \epsilon^{10\rho\alpha}\epsilon_{10\sigma\beta} + \epsilon^{12\rho\alpha}\epsilon_{12\sigma\beta} \\ & + \epsilon^{13\rho\alpha}\epsilon_{13\sigma\beta} + \epsilon^{20\rho\alpha}\epsilon_{20\sigma\beta} + \epsilon^{21\rho\alpha}\epsilon_{21\sigma\beta} + \epsilon^{23\rho\alpha}\epsilon_{23\sigma\beta} + \epsilon^{30\rho\alpha}\epsilon_{30\sigma\beta} \\ & + \epsilon^{31\rho\alpha}\epsilon_{31\sigma\beta} + \epsilon^{32\rho\alpha}\epsilon_{32\sigma\beta}\end{aligned}$$

Case 1: What if ρ, α, σ, and β all have different values?

- They can't all be different (and still yield nonzero terms) because two indices in each term are already taken (from μ and ν) in the equation above. For example, if $\rho = 0$, $\alpha = 1, \sigma = 2$, and $\beta = 3$, you won't be able to find a single nonzero term in the equation above. If you try the term $\epsilon^{23\rho\alpha}\epsilon_{23\sigma\beta}$, for example, you get $\epsilon^{2301}\epsilon_{2323} = 0$. Therefore, if ρ, α, σ, and β all have different values, $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ must equal zero. This agrees with the right-hand side of $\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}$, since all four Kronecker delta's will be zero if ρ, α, σ, and β are all different.

Case 2: $\rho = \alpha$ or $\sigma = \beta$.

- If $\rho = \alpha$ or $\sigma = \beta$, either $\epsilon^{\mu\nu\rho\alpha}$ or $\epsilon_{\mu\nu\sigma\beta}$ will equal zero. If $\rho = \alpha$, then $\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}$ $= \delta^{\alpha}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\alpha}_{\beta}\delta^{\alpha}_{\sigma} = 0$. If $\sigma = \beta$, then $\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma} = \delta^{\rho}_{\beta}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\beta} = 0$. Thus, if $\rho = \alpha$ or $\sigma = \beta$, both sides of $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2\left(\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}\right)$ equal zero.

Case 3: $\rho = \sigma$ or $\alpha = \beta$.

- If $\rho = \sigma$, the only way $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ won't equal 0 is if $\alpha = \beta$. Conversely, if $\alpha = \beta$, the only way $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ won't equal 0 is if $\rho = \sigma$. Why? The 4 indices on each Levi-Civita symbol must all be different in order to get a nonzero term. The indices μ and ν are summed over, and are in the same position in each Levi-Civita symbol. For the Levi-Civita symbols to be nonzero, μ and ν must be different, ρ and α must be different from each other and from μ and ν, and σ and β must be different from each other and from μ and ν. There are 6 subscripts, but there are only 4 values to choose from (0, 1, 2, and 3). This requires two pairs to be equal. In case 3, one pair is either $\rho = \sigma$ or $\alpha = \beta$. The only way to make either choice yield nonzero terms is if both $\rho = \sigma$ and $\alpha = \beta$.
- If both $\rho = \sigma$ and $\alpha = \beta$, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2$ provided that $\rho \neq \alpha$. (We don't need to worry about $\rho = \alpha$ now because we already covered that in Case 2.) To see that $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2$, see the previous expression for $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$. Exactly two terms will be nonzero. For example, if $\rho = \sigma = 2$ and $\alpha = \beta = 3$, the two nonzero terms will be $\epsilon^{0123}\epsilon_{0123} + \epsilon^{1023}\epsilon_{1023} = (-1)(1) + (1)(-1) = -1 - 1 = -2$. Every nonzero term is $(-1)(1)$ or $(1)(-1)$ because $\epsilon^{\mu\nu\rho\alpha}$ and $\epsilon_{\mu\nu\rho\alpha}$ always have opposite signs.

- Here is a recap: If both $\rho = \sigma$ and $\alpha = \beta$, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ equals -2, but if only one of the equalities $\rho = \sigma$ or $\alpha = \beta$ is true, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ equals 0.
- Now let's examine $-2\left(\delta^\rho_\sigma\delta^\alpha_\beta - \delta^\rho_\beta\delta^\alpha_\sigma\right)$. The first term, $\delta^\rho_\sigma\delta^\alpha_\beta$, only equals 1 if both $\rho = \sigma$ and $\alpha = \beta$, which is in exact agreement with the previous bullet point. The second term, $\delta^\rho_\beta\delta^\alpha_\sigma$, only equals 1 if both $\rho = \beta$ and $\alpha = \sigma$. We don't need to worry about this second term in Case 3, since if $\rho = \sigma$ or $\alpha = \beta$ (as required in Case 3), then the condition $\rho = \beta$ and $\alpha = \sigma$ would turn this into Case 2 instead of Case 3.

Case 4: $\rho = \beta$ or $\alpha = \sigma$.

- If $\rho = \beta$, the only way $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ won't equal 0 is if $\alpha = \sigma$. Conversely, if $\alpha = \sigma$, the only way $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ won't equal 0 is if $\rho = \beta$. Why? The 4 indices on each Levi-Civita symbol must all be different in order to get a nonzero term. The indices μ and ν are summed over, and are in the same position in each Levi-Civita symbol. For the Levi-Civita symbols to be nonzero, μ and ν must be different, ρ and α must be different from each other and from μ and ν, and σ and β must be different from each other and from μ and ν. There are 6 subscripts, but there are only 4 values to choose from (0, 1, 2, and 3). This requires two pairs to be equal. In case 4, one pair is either $\rho = \beta$ or $\alpha = \sigma$. The only way to make either choice yield nonzero terms is if both $\rho = \beta$ and $\alpha = \sigma$.
- If both $\rho = \beta$ and $\alpha = \sigma$, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = 2$ provided that $\sigma \neq \beta$. (We don't need to worry about $\sigma = \beta$ now because we already covered that in Case 2.) To see that $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = -2$, see the previous expression for $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$. Exactly two terms will be nonzero. For example, if $\rho = \beta = 2$ and $\alpha = \sigma = 3$, the two nonzero terms will be $\epsilon^{0123}\epsilon_{0132} + \epsilon^{1023}\epsilon_{1032} = (-1)(-1) + (1)(1) = 1 + 1 = 2$. Every nonzero term is $(-1)(-1)$ or $(1)(1)$ because $\epsilon^{\mu\nu\rho\alpha}$ and $\epsilon_{\mu\nu\alpha\rho}$ always the same sign. (Recall that $\epsilon^{\mu\nu\rho\alpha}$ and $\epsilon_{\mu\nu\rho\alpha}$ have opposite signs, and note that $\epsilon_{\mu\nu\alpha\rho} = -\epsilon_{\mu\nu\rho\alpha}$, such that $\epsilon^{\mu\nu\rho\alpha}$ and $\epsilon_{\mu\nu\alpha\rho}$ have the same sign.)

- Here is a recap: If both $\rho = \beta$ and $\alpha = \sigma$, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ equals 2, but if only one of the equalities $\rho = \beta$ or $\alpha = \sigma$ is true, then $\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta}$ equals 0.
- Now let's examine $-2\left(\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}\right)$. The second term, $\delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}$, only equals 1 if both $\rho = \beta$ and $\alpha = \sigma$, which is in exact agreement with the previous bullet point. (Note that the two minus signs make a plus sign when you distribute the leading minus sign to the second term). The first term, $\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta}$, only equals 1 if both $\rho = \sigma$ and $\alpha = \beta$. We don't need to worry about this first term in Case 4, since if $\rho = \beta$ and $\alpha = \sigma$ (as required in Case 4), then the condition $\rho = \sigma$ and $\alpha = \beta$ would turn this into Case 2 instead of Case 4.

Combining all four cases together, we have shown that:

$$\epsilon^{\mu\nu\rho\alpha}\epsilon_{\mu\nu\sigma\beta} = \boxed{-2\left(\delta^{\rho}_{\sigma}\delta^{\alpha}_{\beta} - \delta^{\rho}_{\beta}\delta^{\alpha}_{\sigma}\right)}$$

4 RELATIVE VELOCITY

Relevant Terminology

Inertial reference frame – a frame that travels with constant velocity.

Relative Velocity

Suppose that observer O claims to be at rest and that observer O' moves with constant velocity

$$\vec{v}$$

relative to observer O, where the *x*-axis is oriented along the velocity. Suppose also that observer O' measures the *x*-component of the velocity of an object to be u'_x, where

$$u'_x = \frac{dx'}{dt'}$$

Recall from Chapter 3 that $\Delta x' = \gamma(\Delta x - v\,\Delta t)$ and $\Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)$. For infinitesimal intervals, we may replace the deltas (Δx and Δt) with differential elements (dx and dt), such that $dx' = \gamma(dx - v\,dt)$ and $dt' = \gamma\left(dt - \frac{v}{c^2}dx\right)$.

$$u'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - v\,dt)}{\gamma\left(dt - \frac{v}{c^2}dx\right)} = \frac{dx - v\,dt}{dt - \frac{v}{c^2}dx}$$

Divide the numerator and denominator each by dt.

$$u'_x = \frac{\frac{dx}{dt} - v\frac{dt}{dt}}{\frac{dt}{dt} - \frac{v}{c^2}\frac{dx}{dt}} = \frac{\frac{dx}{dt} - v}{1 - \frac{v}{c^2}\frac{dx}{dt}} = \frac{u_x - v}{1 - \frac{u_x v}{c^2}}$$

Note that

$$u_x = \frac{dx}{dt}$$

is the x-component of the velocity of the object measured by observer O. The equation that relates u'_x to u_x is the Lorentz transformation for velocity. If two different observers measure the velocity of an object, the Lorentz velocity transformation equation can be used to relate their measurements.

In the limit that $u_x, v \ll c$ (meaning that the speeds are small compared to the speed of light), note that the denominator becomes one: $1 - \frac{u_x v}{c^2} \approx 1 - 0 = 1$, such that the relative velocity becomes $u'_x \approx u_x - v$, which agrees with the equation for Galilean relativity from Chapter 1. In the opposite extreme where $u_x \approx c$ (meaning that the relative speed is nearly equal to the speed of light), we get $u'_x \approx \frac{c-v}{1-\frac{cv}{c^2}} = \frac{c-v}{1-\frac{v}{c}} = \frac{c-v}{\frac{c}{c}-\frac{v}{c}} = \frac{c-v}{\frac{c-v}{c}} = \frac{1}{1/c} = c$, which corresponds to the second postulate of special relativity (see Chapter 1): If two different inertial observers measure the speed of light in vacuum, both observers will measure it to be $c = 2.9979 \times 10^8$ m/s.

Symbols and SI Units

Symbol	Name	SI Units
v	relative speed between two observers	m/s
c	speed of light in vacuum	m/s
u_x	the x-component of the velocity of an object measured by O	m/s
u'_x	the x-component of the velocity of an object measured by O′	m/s
x	x-coordinate measured by O	m
x'	x-coordinate measured by O′	m
t	time measured by O	s
t'	time measured by O′	s

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s

Relative Velocity Strategy

To solve a one-dimensional relative velocity problem, follow these steps:

To solve a one-dimensional relative velocity problem, follow these steps:

- Setup a coordinate system with $+x$ along the velocity of O' relative to O, where O claims to be at rest.
- Apply the relative velocity equation.

$$u_x' = \frac{u_x - v}{1 - \frac{u_x v}{c^2}}$$

  - u_x is the x-component of the velocity of the object relative to O.
  - u'_x is the x-component of the velocity of the object relative to O'.
  - v is the speed of O' relative to O.
- Note: If you know u'_x and v, but wish to find u_x, the inverted equation is

$$u_x = \frac{u'_x + v}{1 + \frac{u_x' v}{c^2}}$$

Example: A spaceship passes earth with a speed of $0.7c$ relative to earth. The spaceship fires a rocket in its forward direction with a speed of $0.5c$ relative to the spaceship. What is the speed of the rocket relative to the earth?

Identify the observers and the object, and setup a coordinate system.

- Take the earth to be at rest. The earth is observer O.
- The spaceship is observer O'. Orient $+x$ along the spaceship's velocity.
- The object is the rocket.

Identify the given information.

- The relative speed between O and O' is $v = 0.7c$.
- The velocity of the rocket relative to the spaceship (observer O') is $u'_x = 0.5c$.

We're solving for the speed of the rocket relative to earth, which is u_x. Choose the second equation from the strategy.

$$u_x = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}} = \frac{0.5c + 0.7c}{1 + \frac{(0.5c)(0.7c)}{c^2}}$$

$$= \frac{1.2c}{1 + 0.35} = \frac{1.2c}{1.35} \approx \boxed{0.89c}$$

Example: Derive the equation for u_x in terms of v and u'_x.

Begin with the equation for relative velocity.

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}}$$

This problem is asking us to solve for u_x. Multiply both sides of the equation by $1 - \frac{u_x v}{c^2}$.

$$u'_x \left(1 - \frac{u_x v}{c^2}\right) = u_x - v$$

Distribute the u'_x.

$$u'_x - \frac{u'_x u_x v}{c^2} = u_x - v$$

Add v to both sides and also add $\frac{u'_x u_x v}{c^2}$ to both sides.

$$u'_x + v = u_x + \frac{u'_x u_x v}{c^2}$$

Factor out u_x.

$$u'_x + v = u_x \left(1 + \frac{u'_x v}{c^2}\right)$$

Divide both sides of the equation by $1 + \frac{u'_x v}{c^2}$.

$$\boxed{u_x = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}}}$$

Chapter 4 Problems

1. A spaceship passes earth with a speed of $0.6c$ relative to earth. The spaceship fires a rocket in its forward direction. The rocket has a speed of $0.8c$ relative to earth. What is the speed of the rocket relative to the spaceship?

Want help? Check the solution at the end of the chapter.

Answer: 1. $0.38c$

2. Relative to an observer on earth, one spaceship is heading towards the earth with a speed of $0.9c$ while a second spaceship is heading towards the earth from the opposite direction with a speed of $0.5c$. What is the velocity of each spaceship relative to the other spaceship?

Want help? Check the solution at the end of the chapter.

Answers: 2. $-0.97c$, $-0.97c$

Solutions to Chapter 4

1. Identify the observers and the object, and setup a coordinate system.

- Take the earth to be at rest. The earth is observer O.
- The spaceship is observer O'. Orient $+x$ along the spaceship's velocity.
- The rocket is the object.

Identify the given information. **Note**: This problem is different from the example. This problem gives you the speed of the rocket relative to the earth, whereas the example gave you the speed of the rocket relative to the spaceship. In the example, we were given v and u'_x and were solving for u_x, but in this problem we are given v and u_x and are solving for u'_x.

- The relative speed between O and O' is $v = 0.6c$.
- The velocity of the rocket relative to the earth (observer O) is $u_x = 0.8c$.

We're solving for the speed of the rocket relative to the spaceship (observer O'), which is u'_x. Choose the first equation from the strategy.

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} = \frac{0.8c - 0.6c}{1 - \frac{(0.8c)(0.6c)}{c^2}}$$

$$= \frac{0.2c}{1 - 0.48} = \frac{0.2c}{0.52} \approx \boxed{0.38c}$$

2. Identify the observers and the object, and setup a coordinate system.

- Take the earth to be at rest. The earth is observer O.
- Take the first spaceship to be observer O'. Orient $+x$ along the first spaceship's velocity.
- The second spaceship is the object.

Identify the given information.

- The relative speed between O and O' is $v = 0.9c$.
- The velocity of the second spaceship relative to the earth (observer O) is $u_x = -0.5c$. Why is it negative? Because it is heading in the opposite direction of the first spaceship.

We're solving for the velocity of the second spaceship relative to the first spaceship (observer O'), which is u'_x. Choose the first equation from the strategy.

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} = \frac{-0.5c - 0.9c}{1 - \frac{(-0.5c)(0.9c)}{c^2}} = \frac{-1.4c}{1 - (-0.45)}$$

$$= \frac{-1.4c}{1 + 0.45} = \frac{-1.4c}{1.45} \approx \boxed{-0.97c}$$

If you take the earth to be at rest and the second spaceship to be observer O', and if you orient $+x$ along the second spaceship's velocity, then $v = 0.5c$, $u_x = -0.9c$, and you get

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} = \frac{-0.9c - 0.5c}{1 - \frac{(-0.9c)(0.5c)}{c^2}} = \frac{-1.4c}{1 - (-0.45)}$$

$$= \frac{-1.4c}{1 + 0.45} = \frac{-1.4c}{1.45} \approx \boxed{-0.97c}$$

5 RELATIVISTIC MOMENTUM AND ENERGY

Relevant Terminology

Momentum – mass times velocity.

Inertia – the natural tendency of any object to maintain constant momentum.

Mass – a measure of inertia.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Momentum and Energy

Relativistic **momentum** ($\vec{p}$) equals relativistic mass (m) times velocity ($\vec{u}$), but this equation is often expressed in terms of rest mass (m_0), where $m = \gamma m_0$:

$$\vec{p} = m\vec{u} = \gamma m_0 \vec{u} = \frac{m_0 \vec{u}}{\sqrt{1 - \frac{u^2}{c^2}}}$$

You must be careful when using the equation for relativistic momentum:

- We use $\vec{u}$ for the velocity of the object, since we have already used $\vec{v}$ for the relative velocity between two different observers. Note that a different observer will measure $\vec{u}'$ rather than $\vec{u}$, where $\vec{u}'$ and $\vec{u}$ are related by the equation from Chapter 4.
- In this context, we use u (not v) in the equation $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$. Remember, $\vec{u}$ is the velocity of the object relative to O, whereas $\vec{v}$ is the velocity of O$'$ relative to O.
- Most problems specify the rest mass (m_0), **not** the relativistic mass (m_0).

Relativistic **kinetic energy** (K) equals:

$$K = (\gamma - 1)m_0c^2 = \gamma m_0c^2 - m_0c^2 = mc^2 - m_0c^2 = \frac{m_0c^2}{\sqrt{1 - \frac{u^2}{c^2}}} - m_0c^2$$

A particle at rest has **rest energy** (E_0):

$$E_0 = m_0c^2$$

The **total energy** combines kinetic energy and rest energy together. (However, if there is also potential energy, you must account for that, too.)

$$E = K + E_0 = mc^2 = \gamma m_0c^2 = \frac{m_0c^2}{\sqrt{1 - \frac{u^2}{c^2}}}$$

The equation $E = mc^2$ expresses an equivalence between energy and relativistic mass. Since relativistic mass is proportional to energy, there really is no reason to work with relativistic mass. Thus, in most problems, rest mass (m_0) is given rather than relativistic mass (m), and most solutions work with energy (E), momentum (p), and rest mass (m_0), but <u>**not**</u> relativistic mass (m). It is often useful to express the total energy in terms of momentum:

$$E^2 = p^2c^2 + m_0^2c^4$$

The general form of Newton's second law states that the net external force acting on an object equals the instantaneous rate at which the momentum changes in time. The general form of Newton's second law holds in the context of relativity.

$$\sum \vec{\mathbf{F}}_{ext} = \frac{d\vec{\mathbf{p}}}{dt}$$

Why Does $\boldsymbol{E^2 = p^2c^2 + m_0^2c^4}$?

We will begin with the equation for relativistic momentum.

$$p = \gamma m_0 u = \frac{m_0 u}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Square both sides of the equation.

$$p^2 = \frac{m_0^2 u^2}{1 - \frac{u^2}{c^2}}$$

Multiply the numerator and denominator by c^2.

$$p^2 = \frac{m_0^2 u^2 c^2}{c^2 - u^2}$$

Multiply both sides of the equation by c^2.

$$p^2 c^2 = \frac{m_0^2 u^2 c^4}{c^2 - u^2}$$

Add $m_0^2c^4$ to both sides of the equation.

$$p^2c^2 + m_0^2c^4 = \frac{m_0^2u^2c^4}{c^2 - u^2} + m_0^2c^4$$

Make a common denominator.

$$p^2c^2 + m_0^2c^4 = \frac{m_0^2u^2c^4}{c^2 - u^2} + m_0^2c^4\frac{c^2 - u^2}{c^2 - u^2} = \frac{m_0^2u^2c^4 + m_0^2c^4(c^2 - u^2)}{c^2 - u^2}$$

$$p^2c^2 + m_0^2c^4 = \frac{m_0^2u^2c^4 + m_0^2c^6 - m_0^2c^4u^2}{c^2 - u^2} = \frac{m_0^2c^6}{c^2 - u^2}$$

On the right-hand side, divide the numerator and denominator each by c^2.

$$p^2c^2 + m_0^2c^4 = \frac{m_0^2c^4}{1 - \frac{u^2}{c^2}} = \gamma^2m_0^2c^4 = m^2c^4 = E^2$$

We have just shown that $p^2c^2 + m_0^2c^4 = E^2$.

Why Does $K = \gamma m_0 c^2 - m_0 c^2$?

According to the **work-energy theorem**, the change in kinetic energy equals the net work. Recall that work done equals the integral of force over displacement. The net external force corresponds to the net work done. For one-dimensional motion along the x-axis:

$$\Delta K = K - K_i = W_{net} = \int_{x_i}^{x} \left(\sum F_{ext}\right) dx = \int_{x_i}^{x} \frac{dp}{dt} dx = \int_{p_i}^{p} \frac{dx}{dt} dp = \int_{p_i}^{p} u\, dp$$

Above, we used **Newton's second law**, $\sum F_{ext} = \frac{dp}{dt}$, and the definition of velocity, $u = \frac{dx}{dt}$. We also used the chain rule to write $\frac{dp}{dt} dx = \frac{dpdx}{dt} = \frac{dx}{dt} dp = u\, dp$. Relativistic momentum equals

$$p = \gamma m_0 u = \frac{m_0 u}{\sqrt{1 - \frac{u^2}{c^2}}} = m_0 u \left(1 - \frac{u^2}{c^2}\right)^{-1/2}$$

Take a derivative with respect to u. Apply the product rule.

$$\frac{dp}{du} = \frac{d}{du}\left[m_0 u\left(1 - \frac{u^2}{c^2}\right)^{-1/2}\right] = \left(1 - \frac{u^2}{c^2}\right)^{-1/2} \frac{d}{du}(m_0 u) + m_0 u \frac{d}{du}\left(1 - \frac{u^2}{c^2}\right)^{-1/2}$$

$$= \left(1 - \frac{u^2}{c^2}\right)^{-1/2} m_0 + m_0 u\left(-\frac{1}{2}\right)\left(1 - \frac{u^2}{c^2}\right)^{-3/2}\left(-\frac{2u}{c^2}\right) = \frac{m_0}{\left(1 - \frac{u^2}{c^2}\right)^{1/2}} + \frac{\frac{m_0 u^2}{c^2}}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}}$$

Make a common denominator. Note that $\left(1 - \frac{u^2}{c^2}\right)^{1/2}\left(1 - \frac{u^2}{c^2}\right) = \left(1 - \frac{u^2}{c^2}\right)^{3/2}$.

$$\frac{dp}{du} = \frac{m_0\left(1 - \frac{u^2}{c^2}\right)}{\left(1 - \frac{u^2}{c^2}\right)^{1/2}\left(1 - \frac{u^2}{c^2}\right)} + \frac{\frac{m_0 u^2}{c^2}}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} = \frac{m_0 - \frac{m_0 u^2}{c^2} + \frac{m_0 u^2}{c^2}}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} = \frac{m_0}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}}$$

It follows that

$$dp = \frac{m_0}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} du$$

Substitute the above expression into the integral.

$$K - K_i = \int_{p_i}^{p} u\, dp = \int_{u_i}^{u} \frac{m_0 u}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} du = m_0 \int_{u_i}^{u} u \left(1 - \frac{u^2}{c^2}\right)^{-3/2} du$$

Make the substitution $q = 1 - \frac{u^2}{c^2}$ and $dq = -\frac{2u}{c^2} du$, such that $u\, du = -\frac{c^2}{2} dq$.

$$K - K_i = -\frac{m_0 c^2}{2} \int_{q_i}^{q} q^{-3/2}\, dq = -\frac{m_0 c^2}{2} \left[-2q^{-1/2}\right]_{q_i}^{q} = m_0 c^2 \left[\left(1 - \frac{u^2}{c^2}\right)^{-1/2}\right]_{u_i}^{u}$$

If we let the object begin from rest, then $u_i = 0$ and $K_i = 0$. Evaluate the anti-derivative over the limits in order to complete the definite integral.

$$K = m_0 c^2 \left[\left(1 - \frac{u^2}{c^2}\right)^{-1/2}\right]_0^{u} = m_0 c^2 \left(1 - \frac{u^2}{c^2}\right)^{-1/2} - m_0 c^2 \left(1 - \frac{0^2}{c^2}\right)^{-1/2}$$

$$K = m_0 c^2 \left(1 - \frac{u^2}{c^2}\right)^{-1/2} - m_0 c^2 (1)^{-1/2} = \frac{m_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}} - m_0 c^2 = \gamma m_0 c^2 - m_0 c^2$$

We have just derived the equation for relativistic kinetic energy.

Mass and Energy

Examine the equations relating the total energy (E), relativistic mass (m), and rest mass (m_0).

$$E^2 = p^2c^2 + m_0^2c^4$$

$$E = mc^2 = \gamma m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}}$$

If you set $p = 0$ for an object at rest, you find that an object has **rest energy** (E_0).

$$E_0 = m_0 c^2$$

This shows that even an object that is at rest has energy that can be harnessed. The idea behind this is applied in nuclear reactions, for example.

A few particles found in nature have zero rest mass. A common example is the photon (a particle of light). Yet, such particles still carry energy and momentum.

$$E = pc \text{ (if } m_0 = 0\text{)}$$

For all objects, moving or otherwise, Einstein's famous equation (involving relativistic mass) expresses an equivalence between mass and energy.

$$E = mc^2 = \gamma m_0 c^2$$

Note that although rest mass is conserved in chemical reactions (like $2H_2 + O_2 \to 2H_2O$), rest mass isn't conserved in general. For example, when an electron and positron (which is an anti-electron) approach one another and annihilate to produce a pair of photons (we use the lowercase Greek letter gamma, γ, to represent the photon) in $e^- + e^+ \to \gamma + \gamma$, rest mass isn't conserved. The initial particles (the electron and positron) both have mass, while the final particles (the photons) both have zero rest mass. This is an extreme example of a reaction where rest mass isn't conserved (but where energy is conserved). Another example is nuclear fusion in the sun. In the nuclear reaction $4H + 2e^- \to He + 2\nu + 6\gamma$, four hydrogen atoms (and two electrons) fuse together to form helium (plus two neutrinos and six photons, where neutrinos have very little mass). The difference in mass between the two sides of the reaction accounts for the energy released by the reaction perfectly through $E_0 = m_0 c^2$.

Relativistic 4-Momentum

The **relativistic 4-momentum** combines energy (divided by the speed of light) and momentum into a single 4-vector (it may help to review tensor notation in Chapter 3):

$$\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$$

When energy and momentum are both conserved for a process, we can combine the two conservation laws together into **conservation of 4-momentum**. (As we will explore next, this is common for relativistic collisions.) Recall that the covariant form of a 4-vector is obtained by negating the signs of the spatial components (Chapter 3).

$$\{p_\mu\} = \left\{\frac{E}{c}, -p_x, -p_y, -p_z\right\}$$

The 4D scalar product of a 4-momentum vector with itself is (see the solution to Problem 1 from Chapter 3):

$$p^\mu p_\mu = p^0 p_0 + p^1 p_1 + p^2 p_2 + p^3 p_3 = \frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = \frac{E^2}{c^2} - p^2$$

Note that the magnitude of a 3D vector has the form $p^2 = p_x^2 + p_y^2 + p_z^2$. Substitute the equation $E^2 = p^2c^2 + m_0^2c^4$ into the equation above.

$$p^\mu p_\mu = \frac{p^2c^2 + m_0^2c^4}{c^2} - p^2 = p^2 + m_0^2c^2 - p^2$$

$$p^\mu p_\mu = m_0^2c^2$$

The above equation often comes in handy in the context of 4-momentum.

Relativistic Collisions

In classical collisions between macroscopic objects moving at speeds that are slow compared to the speed of light, both mass and momentum are conserved, but kinetic energy is only conserved for elastic collisions. Relativistic collisions are different. In relativistic collisions between particles, **both energy and momentum are conserved**, but kinetic energy and rest mass are only conserved for *elastic* collisions. Note that rest mass is **not** conserved in general for a relativistic collision.

Since the (total) energy and momentum are both conserved for relativistic collisions, we may combine these together and conserve 4-momentum. For example, for a collision of the form $A + B \to C + D$, the conservation laws $E_A + E_B = E_C + E_D$ and $\vec{\mathbf{p}}_A + \vec{\mathbf{p}}_B = \vec{\mathbf{p}}_C + \vec{\mathbf{p}}_D$ (which is a vector equation) may be combined together to make $p_A^\mu + p_B^\mu = p_C^\mu + p_D^\mu$ (which is really four equations in one, since it is true for each possible value of the spacetime index μ). It is sometimes helpful to apply the following relations (where $p^2 = p_x^2 + p_y^2 + p_z^2$):

$$p^\mu p_\mu = m_0^2 c^2$$

$$E^2 = p^2 c^2 + m_0^2 c^4$$

The Lab and CM Frames

Two different common inertial reference frames may be helpful for analyzing a collision.

- In the **lab** frame, one object is initially at rest. For example, if $A + B \rightarrow C + D$ and object B is at rest, $\vec{\mathbf{p}}_{B,Lab} = 0$. In this example, object B is called the target.
- In the **center-of-momentum** (CM) frame, the total 3-momentum ($\vec{\mathbf{p}}_{tot}$) of the system is zero. For a collision of the form $A + B \rightarrow C + D$, $\vec{\mathbf{p}}_{A,CM} + \vec{\mathbf{p}}_{B,CM} = \vec{\mathbf{p}}_{C,CM} + \vec{\mathbf{p}}_{D,CM} = 0$.

A particular collision problem may be simpler in one frame than in the other. If you work in the CM frame and wish to convert your answer to the lab frame or vice-versa, it may be helpful to apply the following relations. (It may also be helpful to apply the equations from Chapter 4.)

$$p^{\mu}_{Lab} p_{\mu,Lab} = p^{\mu}_{CM} p_{\mu,CM} = m_0^2 c^2$$

$$E^2_{Lab} = p^2_{Lab} c^2 + m_0^2 c^4 \quad , \quad E^2_{CM} = p^2_{CM} c^2 + m_0^2 c^4$$

Note that the equations above apply to individual particles. Add subscripts as needed for each particle (for example, $p^{\mu}_{A,Lab}$ versus $p^{\mu}_{B,Lab}$ and $E_{A,Lab}$ versus $E_{B,Lab}$).

Conservation Laws Versus Invariance

Note the distinction between conservation and invariance.

- A quantity is **conserved** if it is the same before and after a collision.
- A quantity is **invariant** if it is the same in different inertial reference frames.

Following are a few examples:

- Energy is **conserved** for the reaction $A + B \rightarrow C + D$ because $E_A + E_B = E_C + E_D$, but energy isn't invariant because $E_A' + E_B'$ is generally different from $E_A + E_B$. (Energy is conserved in any inertial reference frame, meaning that $E'_A + E'_B = E'_C + E'_D$, but since energy isn't invariant, the value of the total energy is different in each reference frame.)
- Rest mass is **invariant** because the rest mass of a particle is the same in any reference frame, but mass isn't conserved in general because there exist many reactions where the total initial rest mass is different from the total final rest mass. For example, in pair annihilation ($e^- + e^+ \rightarrow \gamma + \gamma$), the total initial rest mass equals $2m_e$, whereas the total final rest mass is zero (since photons have no rest mass).
- The quantity $p^\mu p_\mu$ is both **conserved** and **invariant**. It is invariant because $p'^\mu p'_\mu = p^\mu p_\mu = m_0^2 c^2$, and it is conserved because $p^\mu_{tot,f} = p^\mu_{tot,i}$ (the total initial 4-momentum equals the total final 4-momentum).
- Velocity is neither conserved nor invariant.

Electron Volts

When working with individual particles, the SI unit of energy—which is the Joule (J)—isn't very suitable (because the energy of an individual particle is typically very small compared to one Joule). In this context, it is common to work with **electron Volts** (eV) instead of Joules (J). Recall from electricity and magnetism that work (W) equals charge (q) times potential difference (ΔV): $W = q\Delta V$. Also recall that the SI unit of work (W) is the Joule (J), the SI unit of charge (q) is the Coulomb (C), and the SI unit of potential difference (ΔV) is the Volt (V). From the equation $W = q\Delta V$, it follows that a Joule equals a Coulomb times a Volt: 1 J = 1 C · V. When we apply the same equation, $W = q\Delta V$, to determine the work need to accelerate a proton (or electron, since their charges are equal and opposite) through a potential difference of one Volt (1 V), using the charge of a proton ($e = 1.6021766 \times 10^{-19}$ C; note that e stands for the charge of a proton and is positive, such that the charge of an electron is – e) we find that one electron Volt is related to a Joule by:

$$1 \text{ eV} = 1.6021766 \times 10^{-19} \text{ J}$$

It's common to see a prefix added to electron Volts, such as keV or MeV.

Symbols and SI Units

Symbol	Name	SI Units
m_0	rest mass	kg
m	relativistic mass	kg
$\vec{\mathrm{u}}$	velocity	m/s
u	the magnitude of the velocity	m/s
c	speed of light in vacuum	m/s
p^μ	a component of the contravariant 4-momentum	$kg \cdot m/s$
p_μ	a component of the covariant 4-momentum	$kg \cdot m/s$
p'^μ	contravariant 4-momentum in another reference frame	$kg \cdot m/s$
p'_μ	covariant 4-momentum in another reference frame	$kg \cdot m/s$

$\vec{\mathbf{p}}$	momentum vector in 3D space	$\text{kg} \cdot \text{m/s}$
p	magnitude of the momentum vector in 3D space	$\text{kg} \cdot \text{m/s}$
$\vec{\mathbf{F}}$	force	N
W	work	J
K	kinetic energy	J
K_i	initial kinetic energy	J
E	total energy	J
E'	total energy measured in another inertial reference frame	J
E_0	rest energy	J

t	time	s
γ	time dilation factor (also used to represent a photon)	unitless
e^-	electron	N/A
e^+	positron (anti-electron)	N/A

Note: The symbol γ is the lowercase Greek letter gamma.

Metric Prefixes

Prefix	Value
kilo (k)	$k = 10^3$
mega (M)	$M = 10^6$
giga (G)	$G = 10^9$
tera (T)	$T = 10^{12}$

Constants

Quantity	Value
speed of light in vacuum	$c = 2.99792458 \times 10^{8}$ m/s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C

Notes Regarding Units

Since work (W) is related to force ($\vec{\mathbf{F}}$) and displacement ($\vec{\mathbf{s}}$) by $W = \int_i^f \vec{\mathbf{F}} \cdot d\vec{\mathbf{s}}$, a Joule (J) equals a Newton (N) times a meter (m): 1 J = 1 Nm. Since momentum ($\vec{\mathbf{p}}$) is related to rest mass (m_0) and velocity ($\vec{\mathbf{u}}$) by $\vec{\mathbf{p}} = \gamma m_0 \vec{\mathbf{u}}$, the SI units of momentum are $[\vec{\mathbf{p}}] = \text{kg m/s}$ (since $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$ is unitless). According to Newton's second law, $\sum \vec{\mathbf{F}}_{ext} = \frac{d\vec{\mathbf{p}}}{dt}$, we may express a Joule as:

$1\,\text{J} = 1\frac{\text{kg m}^2}{\text{s}^2}$, since $[\vec{\mathbf{p}}] = \text{kg m/s} = \text{Ns}$ such that $1\,\text{N} = 1\frac{\text{kg m}}{\text{s}^2}$ and since 1 J = 1 Nm.

When working with tiny particles, the following units are common, where $c = 2.99792458 \times 10^8$ m/s is the speed of light in vacuum and where $1\,\text{eV} = 1.6021766 \times 10^{-19}$ J. The following are **not** SI units. The notation $[x]$ is shorthand for "the units of x."

$$[W] = [E] = \text{eV}$$

$$[m] = \frac{\text{eV}}{c^2}$$

Note that $1\,\frac{\text{eV}}{c^2} = \frac{1.6021766\times10^{-19}\,\text{J}}{(2.99792458\times10^{8}\,\text{m/s})^2} = 1.7826619 \times 10^{-36}\,\text{kg}$.

Strategy for Relating Energy, Momentum, and Speed

To relate speed and energy for a relativistic object, follow these steps:

- Identify the known quantities and the desired unknown.

  m_0 = rest mass , m = relativistic mass , $\vec{\mathbf{u}}$ = velocity , u = speed

  $\vec{\mathbf{p}}$ = momentum , K = kinetic energy , E_0 = rest energy , E = total energy
- Choose the appropriate equation based on what you know and what you're finding.

$$\vec{\mathbf{p}} = m\vec{\mathbf{u}} = \gamma m_0\vec{\mathbf{u}} = \frac{m_0\vec{\mathbf{u}}}{\sqrt{1-\frac{u^2}{c^2}}} \quad , \quad \gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$$

$$K = (\gamma - 1)m_0c^2 = \gamma m_0c^2 - m_0c^2 = mc^2 - m_0c^2 = \frac{m_0c^2}{\sqrt{1-\frac{u^2}{c^2}}} - m_0c^2$$

$$E_0 = m_0c^2 \quad , \quad E^2 = p^2c^2 + m_0^2c^4$$

$$E = K + E_0 = mc^2 = \gamma m_0c^2 = \frac{m_0c^2}{\sqrt{1-\frac{u^2}{c^2}}}$$

- When the total energy is conserved, $E_f = E_i$. (Note that E_0 represents rest energy, **<u>not</u>** initial energy.)
- Apply algebra to solve for the desired unknown.
- If you're working with electron Volts (eV), note that 1 eV = $1.6021766 \times 10^{-19}$ J.

Relativistic Collision Strategy

To solve a problem that involves a relativistic collision, follow these steps:

- Label the particles. For example, if two initial particles interact and produce two new particles, the reaction has the form $A + B \rightarrow C + D$. As another example, if one particle decays into two lighter particles, the reaction has the form $A \rightarrow B + C$.
- Choose a reference frame.
  - In the **lab** frame, a moving particle collides with a stationary particle. If the stationary particle is B, for example, then $\vec{\mathbf{p}}_B = 0$ and $\{p_B^\mu\} = \left\{\frac{E_B}{c}, 0,0,0\right\}$. Also, $E_B^2 = p_B^2 c^2 + m_{B0}^2 c^4$ would simplify to $E_B = m_{B0}c^2$ such that $\frac{E_B}{c} = m_{B0}c$.
  - In the **CM** frame, for a reaction like $A + B \rightarrow C + D$, the incident particles have equal and opposite momenta: $\vec{\mathbf{p}}_B = -\vec{\mathbf{p}}_A$. For a decay like $A \rightarrow B + C$, the initial particle A decays from rest: $\vec{\mathbf{p}}_A = 0$.
- Setup a coordinate system. Choose the x-axis to be along the direction of one of the (moving) incident particles. For example, choose $\{p_A^\mu\} = \left\{\frac{E_A}{c}, p_A, 0,0\right\}$, which means that $p_{Ax} = p_A$, $p_{Ay} = 0$, and $p_{Az} = 0$.
- Write down an expression for the 4-momentum of each particle using appropriate subscripts, such as $\{p_C^\mu\} = \left\{\frac{E_C}{c}, p_C \cos\theta\,, p_C \sin\theta\,, 0\right\}$. This example corresponds to a process that occurs in the xy plane with $p_{Cx} = p_C \cos\theta$, $p_{Cy} = p_C \sin\theta$, and $p_{Cz} = 0$. The examples that follow will help to illustrate this step.

- Apply **conservation of 4-momentum.** For a reaction of the form $A + B \to C + D$, this means $p_A^\mu + p_B^\mu = p_C^\mu + p_D^\mu$, which applies for each component separately: $E_A + E_B = E_C + E_D$, $p_{Ax} + p_{Bx} = p_{Cx} + p_{Dx}$, $p_{Ay} + p_{By} = p_{Cy} + p_{Dy}$, and $p_{Az} + p_{Bz} = p_{Cz} + p_{Dz}$.
- It may help to apply the equations $p^\mu p_\mu = m_0^2 c^2$ and $E^2 = p^2c^2 + m_0^2c^4$ (where $p^2 = p_x^2 + p_y^2 + p_z^2$) to one or more particles. For example, $p_D^\mu p_{D\mu} = m_{D0}^2 c^2$ (where $\{p_{D\mu}\} = \left\{\frac{E_D}{c}, -p_{Dx}, -p_{Dy}, -p_{Dz}\right\}$) and $E_D^2 = p_D^2 c^2 + m_{D0}^2 c^4$ (where $p_D^2 = p_{Dx}^2 + p_{Dy}^2 + p_{Dz}^2$).
- It may help to apply the equations $p = \gamma m_0 u$ or $E = \gamma m_0 c^2$, where $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$.
- If one of the particles is a **photon** (γ), its rest mass is zero. For example, if particle D is a photon, then $m_{D0} = 0$ and $E_D = p_D c$.
- If two of the particles have the same rest mass, set their rest masses equal. An example of this is the process $e^- + e^+ \to \gamma + \gamma$, where the electron and positron (which is an anti-electron, and is nothing like a proton) have equal rest mass: $m_{B0} = m_{A0} = m_e$.
- Apply algebra to solve for the desired unknown(s).
- If you need to relate quantities from the lab frame and CM frame, it may be helpful to apply the following relations (or equations from Chapter 4).

$$p_{Lab}^\mu p_{\mu,Lab} = p_{CM}^\mu p_{\mu,CM} = m_0^2 c^2 \quad , \quad E_{Lab}^2 = p_{Lab}^2 c^2 + m_0^2 c^4 \quad , \quad E_{CM}^2 = p_{CM}^2 c^2 + m_0^2 c^4$$

TIP FOR READING EQUATIONS

Some equations and paragraphs with equations appear larger in landscape mode.

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

X

PORTRAIT

(looks small on a small device)

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

✓

LANDSCAPE

Example: A proton has a mass of 1.67262×10^{-27} kg.

(A) Express the mass of a proton in units of $\frac{\text{MeV}}{c^2}$ to three significant figures.

Note that $1 \text{ eV} = 1.6021766 \times 10^{-19}$ J, the speed of light in vacuum is $c = 2.99792458 \times 10^8$ m/s, and the prefix mega (M) stands for $\text{M} = 10^6$. We can use this information to establish a conversion factor from $\frac{\text{MeV}}{c^2}$ to kg.

$$1 \frac{\text{MeV}}{c^2} = \frac{(10^6)(1.6021766 \times 10^{-19} \text{ J})}{(2.99792458 \times 10^8 \text{ m/s})^2} = \frac{1.6021766 \times 10^{-13} \text{ J}}{8.9875518 \times 10^{16} \text{ m}^2/\text{s}^2} = 1.7826619 \times 10^{-30} \text{ kg}$$

Divide both sides by $1.7826619 \times 10^{-30}$ in order to find the conversion factor from kg to $\frac{\text{MeV}}{c^2}$.

$$1 \text{ kg} = \frac{1}{1.7826619 \times 10^{-30}} \frac{\text{MeV}}{c^2} = 5.6095887 \times 10^{29} \frac{\text{MeV}}{c^2}$$

Apply this conversion factor to the mass of the proton.

$$m_p = 1.67262 \times 10^{-27} \text{ kg} = 1.67262 \times 10^{-27} \times 5.6095887 \times 10^{29} \frac{\text{MeV}}{c^2} = \boxed{938 \frac{\text{MeV}}{c^2}}$$

(Although the problem only asked for the answer to three significant figures, it's customary to keep additional digits for intermediate steps in order to limit round-off error. That's why we didn't round the numbers $1.7826619 \times 10^{-30}$ and 5.6095887×10^{29}.)

(B) How much energy must be supplied in order to accelerate a proton from rest to a speed of $0.9c$?

Express conservation of energy for the process: $E_{rest} + E_{supplied} = E_f$. Use the appropriate equations to determine E_{rest} and E_f. Recall from Part A that $m_p = 938\frac{\text{MeV}}{c^2}$.

$$E_{rest} = m_0 c^2 = 938\frac{\text{MeV}}{c^2}c^2 = 938 \text{ MeV}$$

$$E_f = \frac{m_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}} = \frac{938 \text{ MeV}}{\sqrt{1 - \frac{(0.9c)^2}{c^2}}} = \frac{938 \text{ MeV}}{\sqrt{1 - \frac{0.81c^2}{c^2}}} = \frac{938 \text{ MeV}}{\sqrt{1 - 0.81}} = \frac{938 \text{ MeV}}{\sqrt{0.19}} = 2152 \text{ MeV}$$

Plug these numbers into the equation for conservation of energy.

$$938 \text{ MeV} + E_{supplied} = 2152 \text{ MeV}$$

$$E_{supplied} = 2152 \text{ MeV} - 938 \text{ MeV} = 1214 \text{ MeV} = \boxed{1.2 \text{ GeV}}$$

Recall that the prefix giga (G) stands for $\text{G} = 10^9$ while mega (M) stands for $\text{M} = 10^6$.

(C) What potential difference is needed in order to accelerate a proton from rest to a speed of $0.9c$?

Since a proton has the same charge (apart from the sign) as an electron, each electron Volt of energy corresponds to one Volt of potential difference: $\Delta V = 1.2 \text{ GV} = \boxed{1.2 \times 10^9 \text{ V}}$. Note that electron Volt (eV) is a unit of energy (E), while Volt is a unit of potential difference (ΔV). Recall that electrical work equals charge times potential difference: $W_e = q\Delta V$.

Example: When carbon-14 undergoes radioactive beta decay into nitrogen-14, it releases an electron (e^-) and an antineutrino ($\bar{\nu}_e$) according to the reaction below. (An electron is called a beta particle. The release of the electron characterizes this as beta decay.) A neutrino and its antiparticle (the antineutrino) have very little mass: You may treat $\bar{\nu}_e$ as if it has zero rest mass to very good approximation. In the nuclide symbols for carbon-14 ($^{14}_{6}C$) and nitrogen-14 ($^{14}_{7}N$), the bottom left number is the atomic number (which equals the number of protons) and the top left number is the atomic mass number (which equals the number of protons plus the number of neutrons).

$$^{14}_{6}C \rightarrow {}^{14}_{7}N + e^- + \bar{\nu}_e$$

The average atomic mass of carbon-14 and nitrogen-14 are:

$$m_C = 14.003242 \text{ u} \quad , \quad m_N = 14.003074 \text{ u}$$

The rest mass of a proton, neutron, and electron are:

$$m_p = 1.672621898 \times 10^{-27} \text{ kg} \quad , \quad m_n = 1.674927471 \times 10^{-27} \text{ kg}$$

$$m_e = 9.10938356 \times 10^{-31} \text{ kg}$$

Note: The symbol u represents atomic mass units. The conversion from atomic mass units to kilograms is $1 \text{ u} = 1.6605390 \times 10^{-27}$ kg.

(A) Express the rest mass of a carbon-14 atom in units of $\frac{\text{GeV}}{c^2}$.

First use the conversion factor $1 \text{ u} = 1.6605390 \times 10^{-27}$ kg to convert from u to kilograms.

$$m_C = 14.003242 \text{ u} = 14.003242 \times 1.6605390 \times 10^{-27} \text{ kg} = 2.3252929 \times 10^{-26} \text{ kg}$$

Now use the conversion factor $1 \text{ kg} = 5.6095887 \times 10^{26} \frac{\text{GeV}}{c^2}$ to convert from kilograms to $\frac{\text{GeV}}{c^2}$. (We found $1 \text{ kg} = 5.6095887 \times 10^{29} \frac{\text{MeV}}{c^2}$ in the previous example. However, in this example, we want $\frac{\text{GeV}}{c^2}$, where $\text{G} = 10^9$, instead of $\frac{\text{MeV}}{c^2}$, where $\text{M} = 10^6$. Note that $\text{M} = 10^{-3}\text{G}$.)

$$m_C = 2.3252929 \times 10^{-26} \times 5.6095887 \times 10^{26} \frac{\text{GeV}}{c^2} = \boxed{13.043937 \frac{\text{GeV}}{c^2}}$$

(B) Add up the rest masses of the particles that make up a carbon-14 atom. Compare with the answer to Part A. Explain any discrepancy.

Carbon-14 consists of 6 protons and 8 neutrons. (The atomic number of carbon, 6, equals the number of protons in its nucleus. The atomic mass number of the carbon-14 isotope, 14, equals the number of protons plus neutrons. Subtract 6 from 14 to determine that the carbon-14 isotope has 8 neutrons in its nucleus.) Add up the masses of 6 protons, 8 neutrons, and 6 electrons (since a neutral atom has the same number of protons and electrons).

$$m_{tot} = 6(1.672621898 \times 10^{-27} \text{ kg}) + 8(1.674927471 \times 10^{-27} \text{ kg}) + 6(9.10938356 \times 10^{-31} \text{ kg}) = 2.34406168 \times 10^{-26} \text{ kg}$$

Use the conversion factor $1 \text{ kg} = 5.6095887 \times 10^{26} \frac{\text{GeV}}{c^2}$.

$$m_{tot} = 2.34406168 \times 10^{-26} \times 5.6095887 \times 10^{26} \frac{\text{GeV}}{c^2} = \boxed{13.149222 \frac{\text{GeV}}{c^2}}$$

The mass of a carbon-14 atom, $13.044 \frac{\text{GeV}}{c^2}$, is less than the sum of its parts, $13.149 \frac{\text{GeV}}{c^2}$. Why? The nuclear binding energy that holds the protons and neutrons together is equivalent to the mass difference according to $E_0 = m_{tot}c^2$. (Note that the electrons are fairly negligible, as their mass only affects the fourth significant figure in the calculation.)

(C) How much energy is released during a single carbon-14 decay into nitrogen-14?

First convert the mass of nitrogen-14 into units of $\frac{\text{GeV}}{c^2}$.

$$m_N = 14.003074 \text{ u} = 14.003074 \times 1.6605390 \times 10^{-27} \text{ kg} = 2.3252651 \times 10^{-26} \text{ kg}$$

$$m_N = 2.3252651 \times 10^{-26} \times 5.6095887 \times 10^{26} \frac{\text{GeV}}{c^2} = 13.043781 \frac{\text{GeV}}{c^2}$$

The difference in mass between carbon-14 and nitrogen-14 is

$$m_{\text{diff}} = m_C - m_N = 13.043937 \frac{\text{GeV}}{c^2} - 13.043781 \frac{\text{GeV}}{c^2} = 0.000156 \frac{\text{GeV}}{c^2}$$

Since rest energy equals $E_0 = m_0 c^2$, multiply m_{diff} by c^2 to determine the energy released during a single decay. (The speed of light squared will cancel out.) Note that $\text{G} = 10^3 \text{M}$.

$$E_0 = m_0 c^2 = 0.000156 \frac{\text{GeV}}{c^2} c^2 = 0.000156 \text{ GeV} = \boxed{0.156 \text{ MeV}} = \boxed{156 \text{ keV}}$$

Example: A negatively charged kaon (K^-) decays into a muon (μ^-) and an antineutrino ($\bar{\nu}_\mu$). (A muon is similar to an electron, except that a muon has much more mass than an electron.)

$$K^- \to \mu^- + \bar{\nu}_\mu$$

You may treat $\bar{\nu}_\mu$ as if it has zero rest mass to very good approximation. The rest masses of a charged kaon and muon are:

$$m_{\text{kaon}} = 493.68 \frac{\text{MeV}}{c^2} \quad , \quad m_{\text{muon}} = 105.66 \frac{\text{MeV}}{c^2}$$

(A) Write down the 4-momentum for each particle in the rest frame of the kaon.

The 4-momentum of a particle has the structure $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$. We choose to setup our coordinate system with $+x$ along the muon's velocity. Recall that $E^2 = p^2c^2 + m_0^2c^4$.

- The kaon is at rest, so it has zero momentum: $\left\{p^\mu_{\text{kaon}}\right\} = \left\{\frac{E_{\text{kaon}}}{c}, 0,0,0\right\}$. It only has rest energy: $E_{\text{kaon}} = m_{\text{kaon}}c^2$. Therefore, its 4-momentum is $\boxed{\left\{p^\mu_{\text{kaon}}\right\} = \{m_{\text{kaon}}c, 0,0,0\}}$.
- The muon is headed along the $+x$-axis, so it only has an x-component of momentum: $p_x = p$, $p_y = 0$, and $p_z = 0$. Therefore, $\boxed{\left\{p^\mu_{\text{muon}}\right\} = \left\{\frac{E_{\text{muon}}}{c}, p, 0,0\right\}}$.
- The antineutrino is approximately massless, such that $E_\nu = p_\nu c$. From conservation of momentum, it should be clear that the $\vec{\mathbf{p}}_\nu = -\vec{\mathbf{p}}_{\text{muon}}$, which means $p_x = -p$, $p_y = 0$, and $p_z = 0$. Also, $p_\nu = p_{\text{muon}} = p$. Therefore, $\boxed{\left\{p^\mu_\nu\right\} = \{p, -p, 0,0\}}$.

(B) In the rest frame of the kaon, derive an equation for the momentum of the muon in terms of the rest masses of the particles.

All of the equations from Part A apply, since that also involved the rest frame of the kaon.

Express conservation of 4-momentum for the decay $K^- \to \mu^- + \bar{\nu}_\mu$.

$$p^\mu_{\text{kaon}} = p^\mu_{\text{muon}} + p^\mu_\nu$$

$$\{m_{\text{kaon}}c, 0,0,0\} = \left\{\frac{E_{\text{muon}}}{c}, p, 0,0\right\} + \{p, -p, 0,0\}$$

The above equation is true for each component. The zeroth component corresponds to conservation of energy divided by c:

$$m_{\text{kaon}}c = \frac{E_{\text{muon}}}{c} + p$$

The first component gives $0 = p + (-p)$, which represents that the muon and antineutrino travel in opposite directions in the rest frame of the kaon. (This is the same as the CM frame.) Solve for the muon's energy in the above equation.

$$m_{\text{kaon}}c^2 = E_{\text{muon}} + pc$$

$$E_{\text{muon}} = m_{\text{kaon}}c^2 - pc$$

If we apply the equation $E^2 = p^2c^2 + m_0^2c^4$ to the muon, we get

$$E^2_{\text{muon}} = p^2c^2 + m^2_{\text{muon}}c^4$$

since $p_{\text{muon}} = p_\nu = p$. Substitute $E_{\text{muon}} = m_{\text{kaon}}c^2 - pc$ into the previous equation.

$$(m_{\text{kaon}}c^2 - pc)^2 = p^2c^2 + m^2_{\text{muon}}c^4$$

$$m^2_{\text{kaon}}c^4 - 2m_{\text{kaon}}pc^3 + p^2c^2 = p^2c^2 + m^2_{\text{muon}}c^4$$

$$m^2_{\text{kaon}}c^4 - 2m_{\text{kaon}}pc^3 = m^2_{\text{muon}}c^4$$

Solve for the magnitude of the momentum (p).

$$m^2_{\text{kaon}}c^4 - m^2_{\text{muon}}c^4 = 2m_{\text{kaon}}pc^3$$

$$p = p_{\text{muon}} = \frac{(m^2_{\text{kaon}} - m^2_{\text{muon}})c^4}{2m_{\text{kaon}}c^3} = \boxed{\frac{m^2_{\text{kaon}} - m^2_{\text{muon}}}{2m_{\text{kaon}}}c}$$

(C) In the rest frame of the kaon, derive an equation for the energy of the muon in terms of the rest masses of the particles.

Substitute the previous expression for p into the equation $E_{\text{muon}} = m_{\text{kaon}}c^2 - pc$ from Part (B).

$$E_{\text{muon}} = m_{\text{kaon}}c^2 - pc = m_{\text{kaon}}c^2 - \frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2 = m_{\text{kaon}}c^2\frac{2m_{\text{kaon}}}{2m_{\text{kaon}}} - \frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2$$

$$E_{\text{muon}} = \frac{2m_{\text{kaon}}^2}{2m_{\text{kaon}}}c^2 - \frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2 = \frac{2m_{\text{kaon}}^2 - \left(m_{\text{kaon}}^2 - m_{\text{muon}}^2\right)}{2m_{\text{kaon}}}c^2$$

$$E_{\text{muon}} = \frac{2m_{\text{kaon}}^2 - m_{\text{kaon}}^2 + m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2 = \boxed{\frac{m_{\text{kaon}}^2 + m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2}$$

(D) In the rest frame of the kaon, determine the speed of the muon.

We have expressions for p_{muon} and E_{muon} from Parts B and C. Apply the following equations from Chapter 4.

$$p_{\text{muon}} = \gamma m_{\text{muon}}u_{\text{muon}}$$
$$E_{\text{muon}} = \gamma m_{\text{muon}}c^2$$

Divide these two equations.

$$\frac{p_{\text{muon}}}{E_{\text{muon}}} = \frac{u_{\text{muon}}}{c^2}$$

$$u_{\text{muon}} = \frac{p_{\text{muon}}}{E_{\text{muon}}}c^2 = \frac{\frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{2m_{\text{kaon}}}c}{\frac{m_{\text{kaon}}^2 + m_{\text{muon}}^2}{2m_{\text{kaon}}}c^2}c^2 = \frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{m_{\text{kaon}}^2 + m_{\text{muon}}^2}\frac{c^3}{c^2} = \frac{m_{\text{kaon}}^2 - m_{\text{muon}}^2}{m_{\text{kaon}}^2 + m_{\text{muon}}^2}c$$

$$u_{\text{muon}} = \frac{493.68^2 - 105.66^2}{493.68^2 + 105.66^2}c = \frac{232{,}556}{254{,}884}c = \boxed{0.91c}$$

Example: An electron (e^-) and a positron (e^+)—the antiparticle of the electron—produce a pair of photons (γ and γ). This process is called pair annihilation.

$$e^- + e^+ \to \gamma + \gamma$$

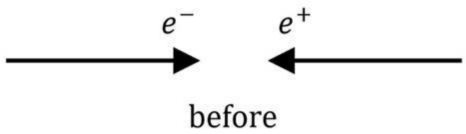
Photons have zero rest mass. The rest mass of the electron and positron are each

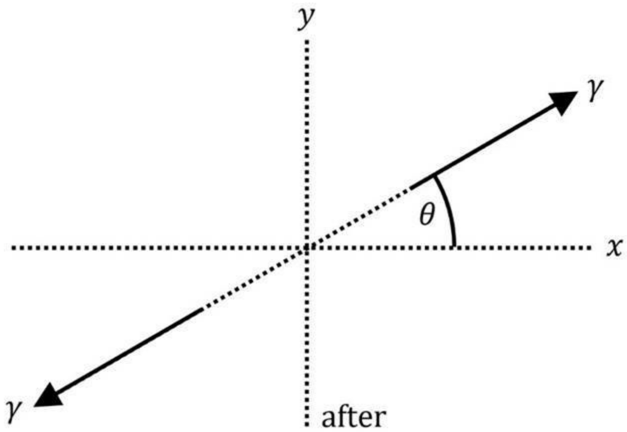
$$m_e = 0.511 \frac{\text{MeV}}{c^2}$$

(A) Write down the 4-momentum for each particle in the CM frame.

The 4-momentum of a particle has the structure $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$. We choose to setup our coordinate system with $+x$ along the electron's velocity and with the photons lying in the xy plane. Recall that $E^2 = p^2c^2 + m_0^2c^4$. In the CM frame, the incident particles have equal and opposite momentum: $\vec{\mathbf{p}}_{e^-} = -\vec{\mathbf{p}}_{e^+}$.

- The electron is headed along the $+x$-axis, so it only has an x-component of momentum: $p_x = p_e$, $p_y = 0$, and $p_z = 0$. Therefore, $\boxed{\{p^\mu_{e^-}\} = \left\{\frac{E_{e^-}}{c}, p_e, 0,0\right\}}$.
- In the CM frame, $\vec{\mathbf{p}}_{e^-} = -\vec{\mathbf{p}}_{e^+}$, which means $p_x = -p_e$, $p_y = 0$, and $p_z = 0$. Also, $p_{e^+} = p_{e^-} = p_e$. Since $m_{e^+} = m_{e^-} = m_e$ and $p_{e^+} = p_{e^-} = p_e$, it follows that $E_{e^+} = E_{e^-} = E_e$. Therefore, $\boxed{\{p^\mu_{e^+}\} = \left\{\frac{E_e}{c}, -p_e, 0,0\right\}}$.
- The photons have zero rest mass, such that $E_\gamma = p_\gamma c$. In the CM frame, the photons are oppositely directed, but possibly at an angle compared to the incident electron (see the figure that follows), which means $p_x = p_\gamma \cos\theta$, $p_y = p_\gamma \sin\theta$, and $p_z = 0$ for one photon and $p_x = -p_\gamma \cos\theta$, $p_y = -p_\gamma \sin\theta$, and $p_z = 0$ for the second photon. Therefore, $\boxed{\{p^\mu_{\gamma 1}\} = \{p_\gamma, p_\gamma \cos\theta, p_\gamma \sin\theta, 0\}}$ and $\boxed{\{p^\mu_{\gamma 2}\} = \{p_\gamma, -p_\gamma \cos\theta, -p_\gamma \sin\theta, 0\}}$.





(B) Express conservation of 4-momentum in the CM frame.

For the process $e^- + e^+ \to \gamma + \gamma$, conservation of 4-momentum is:

$$p^{\mu}_{e^-} + p^{\mu}_{e^+} = p^{\mu}_{\gamma 1} + p^{\mu}_{\gamma 2}$$

$$\boxed{\left\{\frac{E_e}{c}, p_e, 0,0\right\} + \left\{\frac{E_e}{c}, -p_e, 0,0\right\} = \left\{p_\gamma, p_\gamma \cos\theta, p_\gamma \sin\theta, 0\right\} + \left\{p_\gamma, -p_\gamma \cos\theta, -p_\gamma \sin\theta, 0\right\}}$$

The above equation is true for each component. The zeroth component corresponds to conservation of energy divided by c. Recall from Part A that $E_\gamma = p_\gamma c$.

$$\frac{E_e}{c} + \frac{E_e}{c} = p_\gamma + p_\gamma \quad \to \quad 2\frac{E_e}{c} = 2p_\gamma \quad \to \quad E_e = p_\gamma c = E_\gamma$$

The first and second components, $p_e + (-p_e) = p_\gamma \cos\theta + \left(-p_\gamma \cos\theta\right)$ and $0 + 0 = p_\gamma \sin\theta + \left(-p_\gamma \sin\theta\right)$, represent the x- and y-components of conservation of momentum.

(C) Relate the momentum of the electron/positron and photons in the CM frame.

Apply the equation $E^2 = p^2c^2 + m_0^2c^4$ to the electron (which is the same for the positron, since we reasoned in Part A that $p_{e^+} = p_{e^-} = p_e$ and $E_{e^+} = E_{e^-} = E_e$).

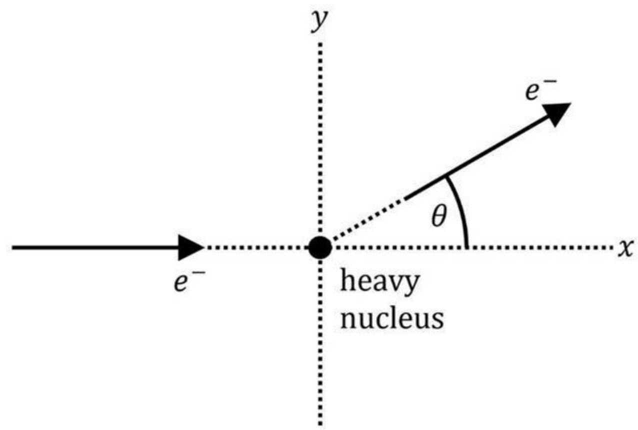
$$E_e^2 = p_e^2c^2 + m_e^2c^4$$

In Part B, we found that $E_e = E_\gamma = p_\gamma c$. Substitute $E_e = p_\gamma c$ into the previous equation.

$$p_\gamma^2c^2 = p_e^2c^2 + m_e^2c^4$$

$$\boxed{p_\gamma^2 = p_e^2 + m_e^2c^2}$$

Example: An electron scatters off of a heavy nucleus.



(A) Write down the 4-momentum for each particle in the lab frame with the nucleus as the target.

The 4-momentum of a particle has the structure $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$. We choose to setup our coordinate system with $+x$ along the electron's initial velocity. Recall that $E^2 = p^2c^2 + m_0^2c^4$. In the lab frame, the target is at rest.

- The electron is initially headed along the $+x$-axis, so it only has an x-component of momentum: $p_x = p_{ei}, p_y = 0$, and $p_z = 0$. Therefore, $\boxed{\{p_{ei}^\mu\} = \left\{\frac{E_{ei}}{c}, p_{ei}, 0,0\right\}}$.
- In the lab frame, the heavy nucleus is at rest, which means $p_x = p_y = p_z = p_{hi} = 0$. Therefore, $\{p_{hi}^\mu\} = \left\{\frac{E_{hi}}{c}, 0,0,0\right\}$. From $E_{hi}^2 = p_{hi}^2c^2 + m_h^2c^4$, it follows that $E_{hi} = m_{hi}c^2$. Thus, $\boxed{\{p_{hi}^\mu\} = \{m_h c, 0,0,0\}}$.
- After the collision, the electron travels at an angle θ relative to the $+x$-axis: $p_x = p_{ef}\cos\theta, p_y = p_{ef}\sin\theta$, and $p_z = 0$. Therefore, $\boxed{\{p_{ef}^\mu\} = \left\{\frac{E_{ef}}{c}, p_{ef}\cos\theta, p_{ef}\sin\theta, 0\right\}}$.
- The heavy nucleus experiences recoil: $p_x = p_{hf}\cos\varphi$, $p_y = -p_{hf}\sin\varphi$, and $p_z = 0$. Therefore, $\boxed{\{p_{hf}^\mu\} = \left\{\frac{E_{hf}}{c}, p_{hf}\cos\varphi, -p_{hf}\sin\varphi, 0\right\}}$.

Note that we have used the subscript i for initial (as in p_{ei}) and f for final (as in p_{ef}).

(B) Express conservation of 4-momentum in the lab frame.

Combine the results from Part A.

$$p^{\mu}_{ei} + p^{\mu}_{hi} = p^{\mu}_{ef} + p^{\mu}_{hf}$$

$$\boxed{\left\{\frac{E_{ei}}{c}, p_{ei}, 0,0\right\} + \{m_h c, 0,0,0\} = \left\{\frac{E_{ef}}{c}, p_{ef}\cos\theta, p_{ef}\sin\theta, 0\right\} + \left\{\frac{E_{hf}}{c}, p_{hf}\cos\varphi, -p_{hf}\sin\varphi, 0\right\}}$$

Chapter 5 Problems

1. A particle with a mass of 3.5654×10^{-30} kg travels with a speed of $0.6c$.

(A) Determine the rest mass of the particle in units of MeV/c^2.

(B) Determine the momentum of the particle in units of MeV/c.

(C) Determine the rest energy of the particle in units of MeV.

(D) Determine the kinetic energy of the particle in units of MeV.

(E) Determine the total energy of the particle in units of MeV.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 2.0000 MeV/c^2 (B) 1.5 MeV/c

(C) 2.0000 MeV (D) 0.50 MeV (E) 2.5 MeV

2. Deuterium—which is also called “heavy hydrogen”—is an isotope of hydrogen consisting of one proton, one neutron, and one electron. A deuterium nucleus is called a deuteron. Thus, a deuteron consists of one proton and one neutron bound together (since the nucleus doesn’t include electrons). The rest mass of a proton and neutron are:

$$m_p = 1.67262 \times 10^{-27}\text{ kg}$$

$$m_n = 1.67493 \times 10^{-27}\text{ kg}$$

The rest mass of a deuteron is:

$$m_d = 3.34358 \times 10^{-27}\text{ kg}$$

(A) Compare the rest mass of a deuteron with the sum of its parts. Explain any discrepancy.

(B) What minimum energy is required to separate the proton and neutron of a deuteron?

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) 3.34755×10^{-27} kg; it’s greater than m_d due to nuclear binding energy

(B) 2.23 MeV

3. Two monkeys each have special guns that launch tiny balls of dough with speeds of $0.8c$ relative to the gun. Each tiny ball of dough has a rest mass of 3.0 mg. When the monkeys fire the guns towards one another, the two tiny balls of dough collide in midair. After the collision, the two tiny balls of dough stick together.

(A) Determine the rest mass of each tiny ball of dough in units of TeV/c^2.

(B) Determine the initial momentum of each tiny ball of dough in units of TeV/c.

(C) Determine the rest energy of each tiny ball of dough in units of TeV.

(D) Determine the initial kinetic energy of each tiny ball of dough in units of TeV.

(E) Determine the initial total energy of each tiny ball of dough in units of TeV.

(F) Determine the speed of the composite object after the collision relative to the monkeys.

(G) Determine the rest mass of the composite object after the collision.

(H) Compare your answer to Part G with your answer to Part A. Explain any discrepancy.

Want help? Check the solution at the end of the chapter.

Answers: 3. (A) 1.7×10^{18} TeV/c^2

(B) 2.3×10^{18} TeV/c

(C) 1.7×10^{18} TeV

(D) 1.1×10^{18} TeV

(E) 2.8×10^{18} TeV

(F) 0

(G) 5.6×10^{18} TeV/c^2

(H) kinetic energy is converted into rest mass

4. A neutral kaon (K^0) decays into a pair of charged pions (π^+ and π^-).

$$K^0 \rightarrow \pi^+ + \pi^-$$

The rest masses of a neutral kaon and charged pion are:

$$m_K = 497.65 \,\frac{\text{MeV}}{c^2} \quad , \quad m_\pi = 139.57 \,\frac{\text{MeV}}{c^2}$$

(A) Write down the 4-momentum for each particle in the rest frame of the kaon.

(B) In the rest frame of the kaon, derive an equation for the energy of each pion in terms of the rest masses of the particles.

(C) In the rest frame of the kaon, derive an equation for the momentum of each pion in terms of the rest masses of the particles.

(D) In the rest frame of the kaon, determine the speed of each pion.

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) $\{m_K c, 0,0,0\}, \left\{\frac{E_\pi}{c}, p_\pi, 0,0\right\}, \left\{\frac{E_\pi}{c}, -p_\pi, 0,0\right\}$

(B) $\frac{m_K c^2}{2}$ (C) $c\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}$ (D) $0.83c$

5. A photon (γ) collides with an electron (e^-) according to the following process:

$$\gamma + e^- \rightarrow \gamma + e^-$$

(A) Write down the 4-momentum for each particle in the lab frame.

(B) Apply conservation of 4-momentum to derive the following equation for the lab frame, where γ_i and γ_f represent the initial and final photons and θ is the angle between γ_i and γ_f.

$$\frac{1}{p_{\gamma f}} - \frac{1}{p_{\gamma i}} = \frac{1 - \cos\theta}{m_e c}$$

Want help? Check the solutions at the end of the chapter.

Answers: 5. (A) $\{p_{\gamma i}, p_{\gamma i}, 0, 0\}$, $\{m_e c, 0, 0, 0\}$, $\{p_{\gamma f}, p_{\gamma f}\cos\theta, p_{\gamma f}\sin\theta, 0\}$, $\left\{\frac{E_{ef}}{c}, p_{ef}\cos\varphi, -p_{ef}\sin\varphi, 0\right\}$

Want help? Check the solution at the end of the chapter.

Solutions to Chapter 5

1. (A) Note that 1 eV = $1.6021766 \times 10^{-19}$ J, $c = 2.99792458 \times 10^{8}$ m/s, and M = 10^6.

$$1\frac{\text{MeV}}{c^2} = \frac{(10^6)(1.6021766 \times 10^{-19}\text{ J})}{(2.99792458 \times 10^{8}\text{ m/s})^2} = \frac{1.6021766 \times 10^{-13}\text{ J}}{8.98755179 \times 10^{16}\text{ m}^2/\text{s}^2} = 1.7826619 \times 10^{-30}\text{ kg}$$

Divide both sides by $1.7826619 \times 10^{-24}$.

$$1\text{ kg} = \frac{1}{1.7826619 \times 10^{-30}}\frac{\text{MeV}}{c^2} = 5.6095887 \times 10^{29}\frac{\text{MeV}}{c^2}$$

Apply this conversion factor to the mass of the particle.

$$m_0 = 3.5654 \times 10^{-30}\text{ kg} = 3.5654 \times 10^{-30} \times 5.6095887 \times 10^{29}\frac{\text{MeV}}{c^2} = \boxed{2.0000\frac{\text{MeV}}{c^2}}$$

(B) First determine the constant γ using $u = 0.6c$.

$$\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{(0.6c)^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{0.36c^2}{c^2}}} = \frac{1}{\sqrt{1-0.36}} = \frac{1}{\sqrt{0.64}} = \frac{1}{0.8} = 1.25$$

Use the equation for relativistic momentum. In terms of speed (u) and the magnitude of the momentum (p), the vector equation $\vec{\mathbf{p}} = \gamma m_0 \vec{\mathbf{u}}$ can be expressed as $p = \gamma m_0 u$. (Recall that the difference between a vector quantity like, $\vec{\mathbf{p}}$, and its magnitude, p, is that the vector includes direction.) Note that we know the rest mass ($m_0 = 2.0000\frac{\text{MeV}}{c^2}$ from Part A) of the particle, **not** the relativistic mass (m). (Almost all masses given in problems are rest masses.)

$$p = \gamma m_0 u = (1.25)\left(2\ \frac{\text{MeV}}{c^2}\right)(0.6c) = \boxed{1.5\ \frac{\text{MeV}}{c}}$$

(C) Use the equation for rest energy and the rest mass from Part A: c^2 will cancel out.

$$E_0 = m_0 c^2 = \left(2\ \frac{\text{MeV}}{c^2}\right) c^2 = \boxed{2.0000\ \text{MeV}}$$

(D) Use the equation for relativistic kinetic energy. Recall from Part B that $\gamma = 1.25$.

$$K = (\gamma - 1) m_0 c^2 = (1.25 - 1)\left(2\ \frac{\text{MeV}}{c^2}\right) c^2 = \boxed{0.50\ \text{MeV}}$$

(E) Add the kinetic energy and rest energy to determine the total energy.

$$E = K + E_0 = 0.50\ \text{MeV} + 2.0\ \text{MeV} = \boxed{2.5\ \text{MeV}}$$

One alternative is to use the equation $E = \gamma m_0 c^2 = (1.25)\left(2\ \frac{\text{MeV}}{c^2}\right)(c^2) = 2.5\ \text{MeV}$.

2. (A) A deuteron consists of a proton and a neutron. If you add the mass of one proton to the mass of one neutron, you get:

$$m_p + m_n = 1.67262 \times 10^{-27}\text{ kg} + 1.67493 \times 10^{-27}\text{ kg} = \boxed{3.34755 \times 10^{-27}\text{ kg}}$$

This is actually greater than the mass of a deuteron, $m_d = 3.34358 \times 10^{-27}$ kg. The reason that $m_p + m_n > m_d$ is that the nuclear binding energy is equivalent to the difference in mass according to $E_0 = m_{\text{diff}}c^2$.

(B) First determine the mass difference between $m_p + m_n$ and m_d.

$$m_{\text{diff}} = (m_p + m_n) - m_d = 3.34755 \times 10^{-27}\text{ kg} - 3.34358 \times 10^{-27}\text{ kg}$$

$$m_{\text{diff}} = 0.00397 \times 10^{-27}\text{ kg} = 3.97 \times 10^{-30}\text{ kg}$$

Convert this mass difference to units of $\frac{\text{MeV}}{c^2}$. Use the conversion factor 1 kg $= 5.6095887 \times 10^{29}\frac{\text{MeV}}{c^2}$ from Problem 1.

$$m_{\text{diff}} = 3.97 \times 10^{-30}\text{ kg} = 3.97 \times 10^{-30} \times 5.6095887 \times 10^{29}\frac{\text{MeV}}{c^2} = 2.23\frac{\text{MeV}}{c^2}$$

Use the equation for rest energy: c^2 will cancel out.

$$E_0 = m_0c^2 = \left(2.23\,\frac{\text{MeV}}{c^2}\right)c^2 = \boxed{2.23\text{ MeV}}$$

3. (A) First convert from milligrams (mg) to kilograms (kg), noting that m = 10^{-3} and k = 10^3.

$$m_0 = 3.0 \text{ mg} = 3.0 \times 10^{-3} \text{ g} = 3.0 \times 10^{-6} \text{ kg}$$

Recall from Problem 1 that 1 kg = $5.6095887 \times 10^{29} \frac{\text{MeV}}{c^2}$. In this problem, we want $\frac{\text{TeV}}{c^2}$ instead of $\frac{\text{MeV}}{c^2}$. According to the chart on page 68, T = 10^{12} and M = 10^6, such that M = 10^{-6} T (since $10^6 = 10^{-6}10^{12}$). It follows that 1 kg = $5.6095887 \times 10^{23} \frac{\text{TeV}}{c^2}$

$$m_0 = 3 \times 10^{-6} \text{ kg} = 3 \times 10^{-6} \times 5.6095887 \times 10^{23} \frac{\text{TeV}}{c^2} = \boxed{1.7 \times 10^{18} \frac{\text{TeV}}{c^2}}$$

(B) First determine the constant γ_i using $u_i = 0.8c$.

$$\gamma_i = \frac{1}{\sqrt{1 - \frac{u_i^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{(0.8c)^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{0.64c^2}{c^2}}} = \frac{1}{\sqrt{1 - 0.64}} = \frac{1}{\sqrt{0.36}} = \frac{1}{0.6} = \frac{1}{3/5} = \frac{5}{3}$$

Use the equation for relativistic momentum.

$$p_i = \gamma_i m_0 u_i = \left(\frac{5}{3}\right)\left(1.7 \times 10^{18} \frac{\text{TeV}}{c^2}\right)(0.8c) = \boxed{2.3 \times 10^{18} \frac{\text{TeV}}{c}}$$

(C) Use the equation for rest energy and the rest mass from Part A: c^2 will cancel out.

$$E_0 = m_0 c^2 = \left(1.7 \times 10^{18} \frac{\text{TeV}}{c^2}\right) c^2 = \boxed{1.7 \times 10^{18} \text{ TeV}}$$

(D) Use the equation for relativistic kinetic energy. Recall from Part B that $\gamma_i = \frac{5}{3}$.

$$K_i = (\gamma_i - 1) m_0 c^2 = \left(\frac{5}{3} - 1\right) m_0 c^2 = \left(\frac{2}{3}\right)\left(1.7 \times 10^{18} \frac{\text{MeV}}{c^2}\right) c^2 = \boxed{1.1 \times 10^{18} \text{ TeV}}$$

(E) Add the initial kinetic energy and rest energy to determine the total initial energy.

$$E_i = K_i + E_0 = 1.1 \times 10^{18} \text{ TeV} + 1.7 \times 10^{18} \text{ TeV} = \boxed{2.8 \times 10^{18} \text{ TeV}}$$

(F) Write down the 4-momentum, $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$, of each object in the CM frame. Setup a coordinate system with $+x$ along one tiny ball of dough, such that $p_{1x} = p_i$, $p_{1y} = 0$, and $p_{1z} = 0$ for one ball and $p_{2x} = -p_i$, $p_{2y} = 0$, and $p_{2z} = 0$ for the other ball.

$$\{p_1^\mu\} = \left\{\frac{E_i}{c}, p_i, 0{,}0\right\} = \left\{2.8 \times 10^{18} \frac{\text{TeV}}{c}, 2.3 \times 10^{18} \frac{\text{TeV}}{c}, 0{,}0\right\}$$

$$\{p_2^\mu\} = \left\{\frac{E_i}{c}, -p_i, 0{,}0\right\} = \left\{2.8 \times 10^{18} \frac{\text{TeV}}{c}, -2.3 \times 10^{18} \frac{\text{TeV}}{c}, 0{,}0\right\}$$

Since 4-momentum is conserved, simply add these together to find the 4-momentum of the composite object after the collision.

$$p_1^\mu + p_2^\mu = p_c^\mu$$

$$\{p_c^\mu\} = \left\{\frac{E_i}{c} + \frac{E_i}{c}, p_i - p_i, 0{,}0\right\} = \left\{\frac{2E_i}{c}, 0{,}0{,}0\right\} = \left\{5.6 \times 10^{18} \frac{\text{TeV}}{c}, 0{,}0{,}0\right\} = \left\{\frac{E_c}{c}, p_c, 0{,}0\right\}$$

The momentum of the composite object is $p_c = 0$. Therefore, $u_c = 0$.

(G) From conservation of 4-momentum in Part F, we see that the energy of the composite object is $E_c = 2E_i = 5.6 \times 10^{18}$ TeV. Apply the equation $E^2 = p^2c^2 + m_0^2c^4$ to get $E_c^2 = p_c^2c^2 + m_c^2c^4 = 0 + m_c^2c^4$ or $E_c = m_c c^2$, such that:

$$m_c = \frac{E_c}{c^2} = \frac{5.6 \times 10^{18} \text{ TeV}}{c^2} = \boxed{5.6 \times 10^{18} \frac{\text{TeV}}{c^2}}$$

(H) In Part G, we found the rest mass of the composite object to be $m_c = 5.6 \times 10^{18} \frac{\text{TeV}}{c^2}$ after the collision. In Part A, we found the rest mass of each tiny ball of dough individually to be $m_0 = 1.7 \times 10^{18} \frac{\text{TeV}}{c^2}$. Observe that rest mass isn't conserved. The final mass, $5.6 \times 10^{18} \frac{\text{TeV}}{c^2}$, is more than the sum of the initial masses, $1.7 \times 10^{18} \frac{\text{TeV}}{c^2} + 1.7 \times 10^{18} \frac{\text{TeV}}{c^2} = 3.4 \times 10^{18} \frac{\text{TeV}}{c^2}$. Why? The initial kinetic energy was converted into rest energy, resulting in an increase in the total rest mass.

4. (A) Write down the 4-momentum, $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$, of each object in the rest frame of the initial kaon. Setup a coordinate system with $+x$ along the velocity of the π^+, such that $p_{\pi^+ x} = p_\pi$, $p_{\pi^+ y} = 0$, and $p_{\pi^+ z} = 0$ for π^+ and $p_{\pi^- x} = -p_\pi$, $p_{\pi^- y} = 0$, and $p_{\pi^- z} = 0$ for π^-.

- The kaon is at rest, so it has zero momentum: $\{p^\mu_{K^0}\} = \left\{\frac{E_K}{c}, 0,0,0\right\}$. It only has rest energy: $E_K = m_K c^2$. Therefore, its 4-momentum is $\{p^\mu_{K^0}\} = \boxed{\{m_K c, 0,0,0\}}$.
- The pions' momenta are along $\pm x$: $\{p^\mu_{\pi^+}\} = \boxed{\left\{\frac{E_\pi}{c}, p_\pi, 0,0\right\}}$ and $\{p^\mu_{\pi^-}\} = \boxed{\left\{\frac{E_\pi}{c}, -p_\pi, 0,0\right\}}$.

(B) Express conservation of 4-momentum for the decay $K^0 \to \pi^+ + \pi^-$.

$$p^\mu_{K^0} = p^\mu_{\pi^+} + p^\mu_{\pi^-}$$

$$\{m_K c, 0,0,0\} = \left\{\frac{E_\pi}{c}, p_\pi, 0,0\right\} + \left\{\frac{E_\pi}{c}, -p_\pi, 0,0\right\}$$

The above equation is true for each component. The zeroth component corresponds to conservation of energy divided by c:

$$m_K c = \frac{E_\pi}{c} + \frac{E_\pi}{c} = \frac{2E_\pi}{c}$$

The first component gives $0 = p_\pi + (-p_\pi)$, which represents that the charged pions travel in opposite directions in the rest frame of the kaon. (This is the same as the CM frame.) Solve for the energy of each pion in the above equation.

$$\boxed{E_\pi = \frac{m_K c^2}{2}}$$

(C) If we apply the equation $E^2 = p^2c^2 + m_0^2c^4$ to either pion, we get

$$E_\pi^2 = p_\pi^2c^2 + m_\pi^2c^4$$

Substitute the equation $E_\pi = \frac{m_{K0}c^2}{2}$ from Part B into the previous equation.

$$\left(\frac{m_K c^2}{2}\right)^2 = p_\pi^2c^2 + m_\pi^2c^4$$

$$\frac{m_K^2c^4}{4} = p_\pi^2c^2 + m_\pi^2c^4$$

Solve for the magnitude of each pion's momentum (p_π).

$$\frac{m_K^2c^4}{4} - m_\pi^2c^4 = p_\pi^2c^2$$

$$\frac{m_K^2c^2}{4} - m_\pi^2c^2 = p_\pi^2$$

$$p_\pi = \sqrt{\frac{m_K^2c^2}{4} - m_\pi^2c^2} = \sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)c^2} = \boxed{c\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}}$$

(D) We have expressions for p_π and E_π from Parts B and C. Apply the following equations from Chapter 4.

$$p_\pi = \gamma m_\pi u_\pi$$
$$E_\pi = \gamma m_\pi c^2$$

Divide these two equations.

$$\frac{p_\pi}{E_\pi} = \frac{u_\pi}{c^2} \quad \rightarrow \quad u_\pi = \frac{p_\pi}{E_\pi} c^2$$

Now plug in the expressions from Parts B and C.

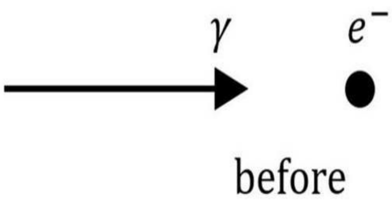
$$u_\pi = \frac{p_\pi}{E_\pi} c^2 = \frac{c\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}}{\frac{m_K c^2}{2}} c^2 = \frac{\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}}{\frac{m_K c^2}{2}} \frac{c^3}{c^2} = \frac{\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}}{m_K/2} c = \frac{2c}{m_K}\sqrt{\left(\frac{m_K^2}{4} - m_\pi^2\right)}$$

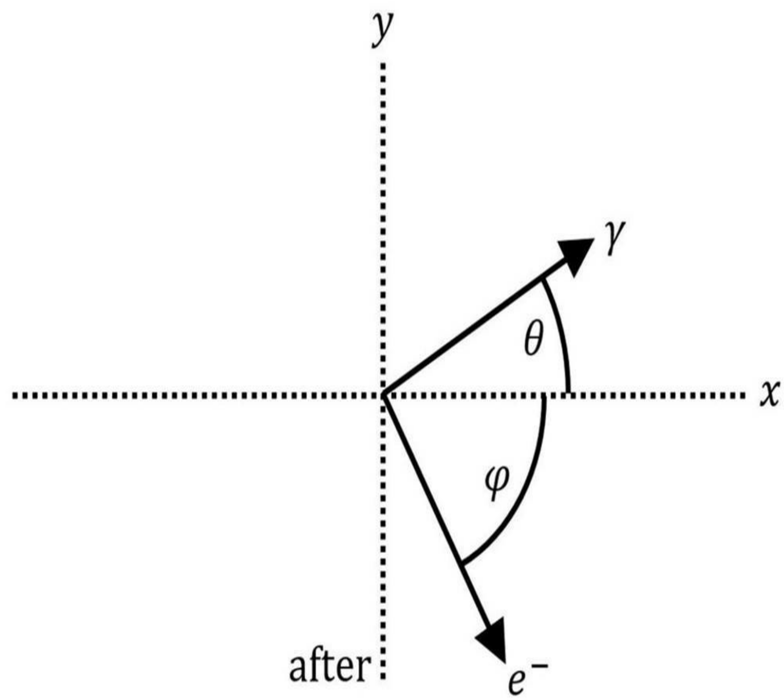
Note that division by 1/2 equates to multiplying by 2: $\frac{1}{1/2} = 1 \div \frac{1}{2} = 1 \times \frac{2}{1} = 2$.

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

5. (A) Write down the 4-momentum, $\{p^\mu\} = \left\{\frac{E}{c}, p_x, p_y, p_z\right\}$, of each object. Setup a coordinate system with $+x$ along the velocity of the initial photon (γ), such that $p_{\gamma x} = p_{\gamma i}$, $p_{\gamma y} = 0$, and $p_{\gamma z} = 0$.





- The initial photon's momentum is along $+x$, and since the photon is massless, $E_{\gamma i}^2 = p_{\gamma i}^2 c^2$ and $E_{\gamma i} = p_{\gamma i} c$: $\{p_{\gamma i}^\mu\} = \boxed{\{p_{\gamma i}, p_{\gamma i}, 0,0\}}$.
- In the lab frame, the initial electron is at rest (it is the target). It only has rest energy: $E_{ei} = m_e c^2$. Therefore, $\{p_{ei}^\mu\} = \boxed{\{m_e c, 0,0,0\}}$.
- Apply trig to determine the x- and y-components of the final photon's momentum in the diagram above: $\{p_{\gamma f}^\mu\} = \boxed{\{p_{\gamma f}, p_{\gamma f}\cos\theta\,, p_{\gamma f}\sin\theta\,, 0\}}$.
- Apply trig to determine the x- and y-components of the final electron's momentum in the diagram above: $\{p_{ef}^\mu\} = \boxed{\left\{\frac{E_{ef}}{c}, p_{ef}\cos\varphi\,, -p_{ef}\sin\varphi\,, 0\right\}}$.

(B) Express conservation of 4-momentum for the process $\gamma + e^- \to \gamma + e^-$.

$$\{p^{\mu}_{\gamma i}\} + \{p^{\mu}_{ei}\} = \{p^{\mu}_{\gamma f}\} + \{p^{\mu}_{ef}\}$$

$$\{p_{\gamma i}, p_{\gamma i}, 0,0\} + \{m_e c, 0,0,0\} = \{p_{\gamma f}, p_{\gamma f}\cos\theta\,, p_{\gamma f}\sin\theta\,, 0\} + \left\{\frac{E_{ef}}{c}, p_{ef}\cos\varphi\,, -p_{ef}\sin\varphi\,, 0\right\}$$

The above equation is true for each component. The zeroth component corresponds to conservation of energy divided by c:

$$p_{\gamma i} + m_e c = p_{\gamma f} + \frac{E_{ef}}{c}$$

The first and second components correspond to conservation of momentum.

$$p_{\gamma i} + 0 = p_{\gamma f}\cos\theta + p_{ef}\cos\varphi$$
$$0 + 0 = p_{\gamma f}\sin\theta - p_{ef}\sin\varphi$$

Look at the equation that we are trying to derive. That equation doesn't have φ in it. Isolate $p_{ef}\cos\varphi$ and $p_{ef}\sin\varphi$ in each equation above.

$$p_{ef}\cos\varphi = p_{\gamma i} - p_{\gamma f}\cos\theta$$
$$p_{ef}\sin\varphi = p_{\gamma f}\sin\theta$$

Square each equation. Recall from algebra that $(A - B)^2 = A^2 - 2AB + B^2$.

$$p_{ef}^2\cos^2\varphi = p_{\gamma i}^2 - 2p_{\gamma i}p_{\gamma f}\cos\theta + p_{\gamma f}^2\cos^2\theta$$
$$p_{ef}^2\sin^2\varphi = p_{\gamma f}^2\sin^2\theta$$

Add these equations together. Recall the trig identity $\cos^2\varphi + \sin^2\varphi = 1$.

$$p_{ef}^2\cos^2\varphi + p_{ef}^2\sin^2\varphi = p_{\gamma i}^2 - 2p_{\gamma i}p_{\gamma f}\cos\theta + p_{\gamma f}^2\cos^2\theta + p_{\gamma f}^2\sin^2\theta$$
$$p_{ef}^2(\cos^2\varphi + \sin^2\varphi) = p_{\gamma i}^2 - 2p_{\gamma i}p_{\gamma f}\cos\theta + p_{\gamma f}^2(\cos^2\theta + \sin^2\theta)$$
$$p_{ef}^2 = p_{\gamma i}^2 - 2p_{\gamma i}p_{\gamma f}\cos\theta + p_{\gamma f}^2$$

Solve for E_{ef} in the conservation of energy equation.

$$p_{\gamma i}c + m_e c^2 = p_{\gamma f}c + E_{ef}$$
$$E_{ef} = p_{\gamma i}c + m_e c^2 - p_{\gamma f}c$$

Square the previous equation.

$$E_{ef}^2 = p_{\gamma i}^2 c^2 + 2m_e p_{\gamma i} c^3 + m_e^2 c^4 - 2p_{\gamma i} p_{\gamma f} c^2 - 2m_e p_{\gamma f} c^3 + p_{\gamma f}^2 c^2$$

Apply the equation $E^2 = p^2 c^2 + m_0^2 c^4$ to the final electron:

$$E_{ef}^2 = p_{ef}^2 c^2 + m_e^2 c^4$$

Substitute the previous equation into the equation before it.

$$p_{ef}^2 c^2 + m_e^2 c^4 = p_{\gamma i}^2 c^2 + 2m_e p_{\gamma i} c^3 + m_e^2 c^4 - 2p_{\gamma i} p_{\gamma f} c^2 - 2m_e p_{\gamma f} c^3 + p_{\gamma f}^2 c^2$$

Subtract $m_e^2 c^4$ from both sides of the equation and divide the entire equation by c^2.

$$p_{ef}^2 = p_{\gamma i}^2 + 2m_e p_{\gamma i} c - 2p_{\gamma i} p_{\gamma f} - 2m_e p_{\gamma f} c + p_{\gamma f}^2$$

Plug the equation $p_{ef}^2 = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2$ (from before) into the previous equation.

$$p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2 = p_{\gamma i}^2 + 2m_e p_{\gamma i} c - 2p_{\gamma i} p_{\gamma f} - 2m_e p_{\gamma f} c + p_{\gamma f}^2$$

Subtract $p_{\gamma i}^2$ and $p_{\gamma f}^2$ from both sides of the equation and divide the entire equation by 2.

$$-p_{\gamma i} p_{\gamma f} \cos\theta = m_e p_{\gamma i} c - p_{\gamma i} p_{\gamma f} - m_e p_{\gamma f} c$$

Add $p_{\gamma i} p_{\gamma f}$ to both sides of the equation and factor. Then divide both sides by $p_{\gamma i} p_{\gamma f}$.

$$p_{\gamma i} p_{\gamma f}(1 - \cos\theta) = m_e p_{\gamma i} c - m_e p_{\gamma f} c$$

$$1 - \cos\theta = \frac{m_e c}{p_{\gamma f}} - \frac{m_e c}{p_{\gamma i}}$$

Now divide both sides of the equation by $m_e c$.

$$\boxed{\frac{1 - \cos\theta}{m_e c} = \frac{1}{p_{\gamma f}} - \frac{1}{p_{\gamma i}}}$$

This is the equation for Compton scattering. We will see this equation again in Chapter 8.

6 BLACKBODY RADIATION

Relevant Terminology

Blackbody – an ideal absorber of thermal radiation, meaning that it absorbs 100% of the thermal radiation that is incident upon it.

Thermal radiation – thermal energy (heat) that is transferred in the form of electromagnetic radiation.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Work – work is done when there is not only a force acting on an object, but when the force also contributes toward the displacement of an object.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Power – the rate at which work is done or the rate at which energy is transferred.

Intensity – power per unit area.

Temperature – a measure of the average kinetic energy of the molecules of a substance.

Emission rate – the rate at which an object emits thermal radiation.

Absorption rate – the rate at which an object absorbs thermal radiation.

Emissivity – a measure of how efficiently an object emits (or absorbs) thermal radiation.

Quantum – a fixed elemental unit corresponding to the minimum possible value that can be measured for a quantity that comes in discrete bundles like energy or angular momentum.

Quantized – limited to integer multiples of a quantum unit. A quantity like energy or angular momentum that is quantized is discrete (rather than continuous).

Thermal Radiation

Objects constantly absorb and emit radiation in the form of electromagnetic waves.

- When an object **absorbs** radiation, this causes the object's temperature to increase.
- When an object **emits** radiation, this causes the object's temperature to decrease.
- When the absorption and emission rates become equal, the object is in a state of **radiative equilibrium** and the object's temperature remains constant.

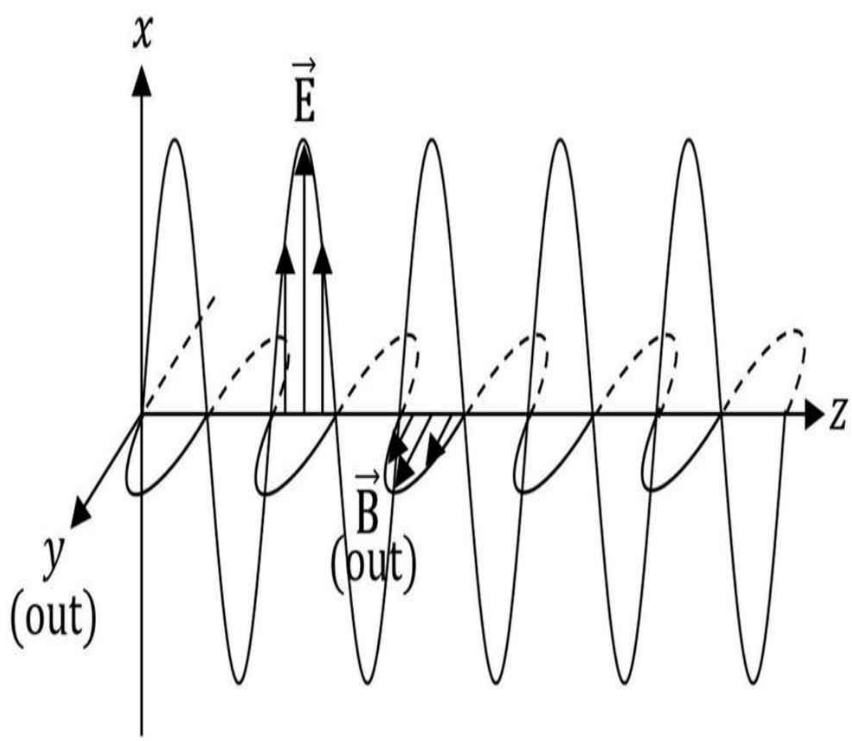
The emissivity (ϵ) provides a measure of how efficiently an object emits (or absorbs) thermal radiation.

- $\epsilon = 1$ for a perfect absorber/emitter (called a blackbody).
- $\epsilon = 0$ for a perfect reflector.

Electromagnetic Waves

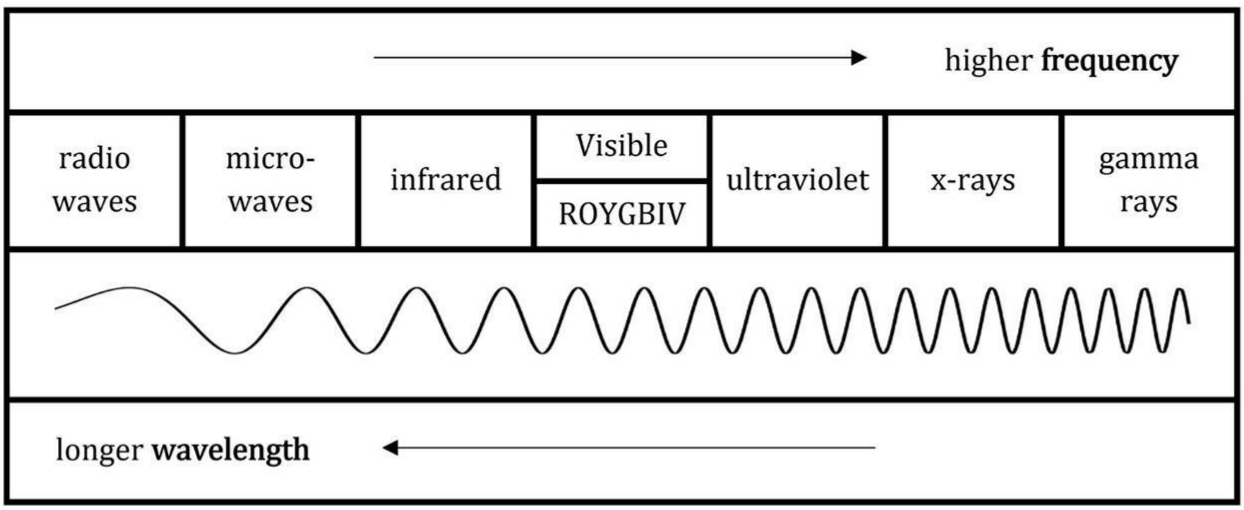
An electromagnetic wave (also called light) involves oscillating electric and magnetic fields. As the electromagnetic wave in the figure below propagates to the right (along $+z$), the electric field ($\vec{\mathrm{E}}$) oscillates up and down (along $\pm x$) while the magnetic field ($\vec{\mathrm{B}}$) oscillates into and out of the page (along $\pm y$). The wavelength (λ) is equal to the distance between consecutive crests, and the frequency (f) is equal to the number of cycles completed per second. The wavelength and frequency are inversely related via the speed of light (c) according to:

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



Electromagnetic Spectrum

The electromagnetic spectrum consists of radio waves, microwaves, infrared, visible light, ultraviolet, x-rays, and gamma rays. The visible spectrum is a just a very narrow slice (much narrower than it would appear by looking at the chart below – the blocks do **not** really have equal width when drawn to scale) of the full electromagnetic spectrum. Wavelength (λ) and frequency (f) share a reciprocal relationship: $\lambda = \frac{c}{f}$ and $f = \frac{c}{\lambda}$. Higher frequency corresponds to shorter wavelength, while lower frequency corresponds to longer wavelength.



The Spectrum of Thermal Radiation

Condensed matter (solids and liquids) states emit a continuous spectrum of thermal radiation. Their thermal radiation spectra depend very strongly on temperature, but only a little on the composition of the material.

- At low temperatures, most solids and liquids don’t emit enough thermal radiation (in the visible spectrum) to see them. At night, with the lights off, they are invisible. (You can see them during the day or with the lights on because in that case you are seeing reflected light, which is different. With thermal radiation, we’re concerned with light that a body emits, not light that it reflects.)

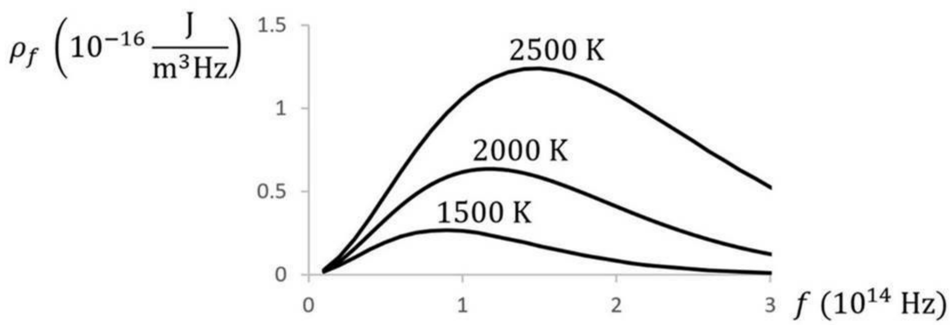
- As the temperature of a body increases, at first it becomes hot to touch, but doesn’t appear hot. That is, you still can’t see it with your eyes at night with the lights off. If you use infrared goggles, however, then you can see light from the thermal radiation that it emits. In this temperature range, most of the thermal radiation is infrared.

- Once a body’s temperature increases high enough, it emits enough visible light to see it even at night with the lights off. (This is emitted light, not reflected light). When it first becomes visible, it appears a dull red. (One way to see this is to dip a metal rod into a fire and wait for it to heat up sufficiently.)

- As the temperature increases further, the body changes color. From low temperature to high temperature, we see dull red, dark red, bright red, orange, yellow/orange, yellow/white, and white (or blue/white).

Wavelength, Frequency, and Temperature

Note that the spectrum of thermal radiation shows a relationship between temperature and color. Different colors of light correspond to different wavelengths (or frequencies) of electromagnetic radiation. Solids and liquids at higher temperatures emit thermal radiation with a higher peak frequency (corresponding to a shorter peak wavelength, since $\lambda = \frac{c}{f}$). We will see this dependence in **Wien's displacement law**. Note that you see the same relationship whether you dip a metal rod into a fire or look at stars through a telescope: Red stars are cooler and white stars are hotter.

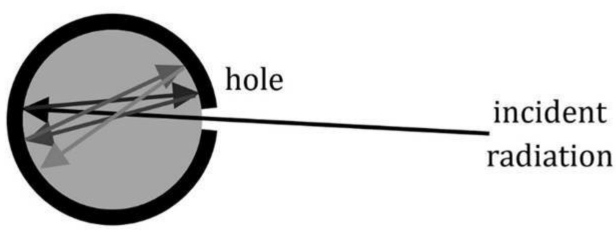


Kirchhoff's Law for Radiation

Suppose that two bodies can only exchange energy via thermal radiation. When they attain **radiative equilibrium**, the absorption rate (a) of each body equals its emission rate (e). That is, $a_1 = e_1$ and $a_2 = e_2$. It follows that $e_1/a_1 = e_2/a_2 = 1$, which is Kirchhoff's law for radiation. This law shows that good absorbers are also good emitters.

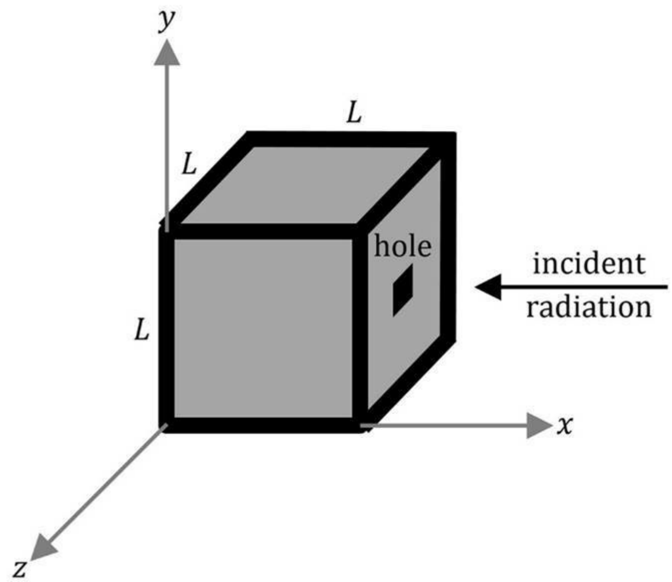
Blackbodies

A **blackbody** is a perfect absorber of thermal radiation, absorbing 100% of the radiation that is incident upon it. The reason for studying the spectrum of thermal radiation emitted by a blackbody is that the spectrum is completely independent of the composition of the material. Two common examples of approximate blackbodies include a furnace and a star. Consider a furnace with thick, highly insulated walls and a small hole in the door. Such a furnace has the structure of a cavity. Any thermal radiation incident upon the small hole enters the cavity and is reflected so many times inside of the cavity such that the incident thermal radiation is almost entirely absorbed by the cavity. The temperature of the cavity walls increases as the cavity absorbs thermal radiation, such that the cavity walls emit thermal radiation. Some of this emitted thermal radiation escapes through the hole, such that the cavity not only absorbs thermal radiation, it also emits thermal radiation. (It is a blackbody because it doesn't *reflect* thermal radiation.) From observations made outside of the cavity, the hole itself has the characteristic behavior of a blackbody. The thermal radiation emitted by the blackbody through the hole depends only on two quantities: temperature (T) and frequency (f).



Blackbody Radiation

Blackbody radiation is independent of the composition of the material as well as the shape of the body. Theoretically, since blackbody radiation doesn't depend on the shape of the body, this allows us to choose whichever shape makes the calculation simplest. For simplicity, let us assume that the blackbody is a cavity with a small hole in the wall (like our example of a furnace with a small hole in the door), and let us choose a cube with edge length L as the shape of the cavity.



Cavity Radiation

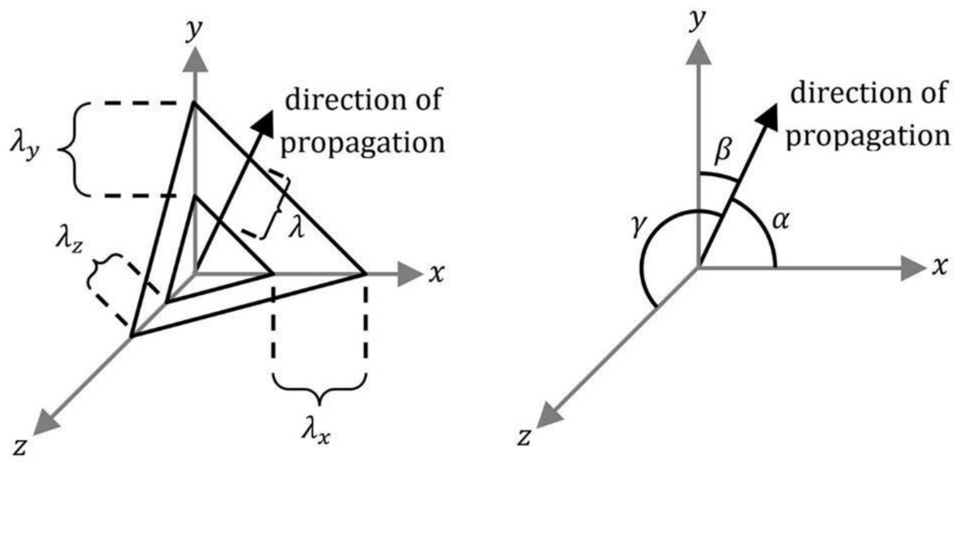
The thermal radiation inside of the blackbody cavity is electromagnetic radiation. We choose a cavity in the shape of a cube with edge length L with conductive walls, and we setup our coordinate system with the origin at one corner and the x-, y-, and z-axes along the edges. Since the walls are conductive, the electric field must be zero at the walls of the cavity.

$$E(x=0)=0 \quad , \quad E(y=0)=0 \quad , \quad E(z=0)=0$$
$$E(x=L)=0 \quad , \quad E(y=L)=0 \quad , \quad E(z=L)=0$$

With these boundary conditions, the electromagnetic waves form standing waves inside of the cavity with nodes at each wall. The nodes are represented by the following equations.

$$\frac{2L}{\lambda_x}=n_x \quad , \quad \frac{2L}{\lambda_y}=n_y \quad , \quad \frac{2L}{\lambda_z}=n_z \quad , \quad \{n_x,n_y,n_z\}\in\{1,2,3,\cdots\}$$

The three-dimensional electromagnetic standing waves are analogous to the standing waves that form on a string that is fixed at both ends (treated in first-year physics textbooks).



The diagram on the previous page shows an electromagnetic wave propagating in an arbitrary direction (that is, not along one of the coordinate axes). The arrow indicates the direction of propagation. The two triangles represent two parallel planes separated by one wavelength (λ) along the direction of propagation. The distances λ_x, λ_y, and λ_z are **not** components of a vector: Based on how these distances are defined geometrically, they are actually longer than λ (whereas the components of a vector are shorter than its magnitude). The angles between the direction of propagation and the x-, y-, and z-axes are α, β, and γ. These angles are related to λ_x, λ_y, and λ_z by:

$$\lambda_x = \frac{\lambda}{\cos\alpha} \quad , \quad \lambda_y = \frac{\lambda}{\cos\beta} \quad , \quad \lambda_z = \frac{\lambda}{\cos\gamma}$$

Take the reciprocal of each equation to see that

$$\frac{1}{\lambda_x} = \frac{\cos\alpha}{\lambda} \quad , \quad \frac{1}{\lambda_y} = \frac{\cos\beta}{\lambda} \quad , \quad \frac{1}{\lambda_z} = \frac{\cos\gamma}{\lambda}$$

Substitute these equations into the equations for the nodes (on the previous page).

$$\frac{2L}{\lambda}\cos\alpha = n_x \quad , \quad \frac{2L}{\lambda}\cos\beta = n_y \quad , \quad \frac{2L}{\lambda}\cos\gamma = n_z \quad , \quad \{n_x, n_y, n_z\} \in \{1,2,3,\cdots\}$$

Square both sides of each equation and add the equations together to get

$$\frac{4L^2}{\lambda^2}(\cos^2\alpha + \cos^2\beta + \cos^2\gamma) = n_x^2 + n_y^2 + n_z^2$$

Note that $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$. (One way to see this is to consider the velocity vector, $\vec{\boldsymbol{v}}$, along the direction of propagation. The components of the velocity vector are $v_x = v\cos\alpha$, $v_y = v\cos\beta$, and $v_z = v\cos\gamma$. Unlike wavelength, velocity is a vector: Note that v_x, v_y, and v_z are components of a vector, whereas λ_x, λ_y, and λ_z are not. Since v_x, v_y, and v_z are components of a vector, they satisfy $v_x^2 + v_y^2 + v_z^2 = v^2$, which is the three-dimensional generalization of the Pythagorean theorem. We thus get $v^2\cos^2\alpha + v^2\cos^2\beta + v^2\cos^2\gamma = v^2$, which reduces to $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.)

$$\frac{4L^2}{\lambda^2} = n_x^2 + n_y^2 + n_z^2$$

Squareroot both sides of the equation.

$$\frac{2L}{\lambda} = \sqrt{n_x^2 + n_y^2 + n_z^2}$$

Wavelength (λ) times frequency (f) equals wave speed, and the speed of an electromagnetic wave is the speed of light (c): $\lambda f = c$. Divide both sides by λ to get $f = \frac{c}{\lambda}$. Divide both sides by c to get $\frac{f}{c} = \frac{1}{\lambda}$. We can write the equation above in terms of frequency instead of wavelength if we make the substitution $\frac{f}{c} = \frac{1}{\lambda}$.

$$\frac{2Lf}{c} = \sqrt{n_x^2 + n_y^2 + n_z^2}$$

TIP FOR READING EQUATIONS

Some equations and paragraphs with equations appear larger in landscape mode.

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

✗ **PORTRAIT**

(looks small on a small device)

$$u_\pi = \frac{2c}{497.65}\sqrt{\frac{(497.65)^2}{4} - (139.57)^2} = \frac{c}{248.825}\sqrt{61{,}914 - 19{,}480}$$

$$u_\pi = \frac{c}{248.825}\sqrt{42{,}434} = \frac{206c}{248.825} = \boxed{0.83c}$$

✓ **LANDSCAPE**

The Law of Equipartition of Energy (Classical Result)

In the kinetic theory of gases, the **law of equipartition of energy** states that for a system of gas molecules in thermal equilibrium, the average kinetic energy per degree of freedom is

$$\frac{\overline{KE}}{n_d} = \frac{k_B T}{2}$$

where $\overline{KE}$ represents the average kinetic energy of a molecule, n_d is the number of degrees of freedom, $k_B = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant, and T is the absolute temperature in Kelvin (K). A common example of the law of equipartition of energy is found in the kinetic theory of ideal gases. For example, a monatomic ideal gas has three degrees of freedom ($n_d = 3$) and an internal energy of $U = \frac{3}{2} N k_B T$ (where N is the number of gas molecules), while a diatomic ideal gas has five degrees of freedom ($n_d = 5$) and an internal energy of $U = \frac{5}{2} N k_B T$. (A monatomic gas molecule has three degrees of freedom because space is three-dimensional. A diatomic gas molecule, like H_2 or O_2, has two additional degrees of freedom corresponding to two different ways that a diatomic molecule, shaped like a dumbbell, can rotate.)

The law of equipartition of energy doesn't just apply to gases: It extends to any system with a large number of entities (where all of the entities are of the same kind). The electromagnetic standing waves in a blackbody cavity form such a system. (Technically, this is a sort of gas: It is a "photon gas," but when the blackbody radiation problem was being investigated in the history of physics, the concept of a photon was as of yet unknown.)

Energy of a Single Electromagnetic Standing Wave in a Cavity (Classical)

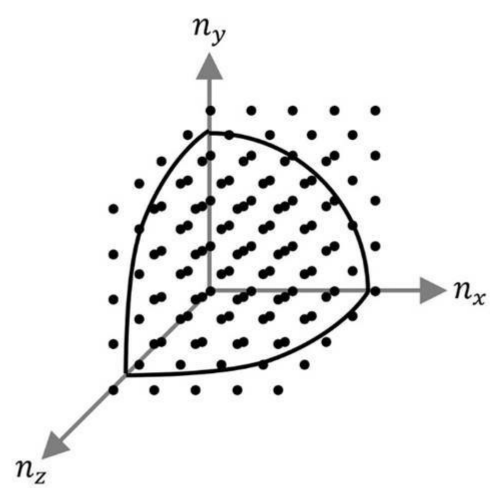
Consider one electromagnetic standing wave in a blackbody cavity with particular values of n_x, n_y, and n_z. For its given values of n_x, n_y, and n_z, a single degree of freedom ($n_d = 1$) remains, corresponding to the amplitude of the standing wave. This single degree of freedom ($n_d = 1$) results in an average kinetic energy of $\overline{KE} = \frac{k_B T}{2}$ at thermal equilibrium. The average total energy ($\bar{\mathcal{E}}$) is twice the average kinetic energy because an electromagnetic standing wave oscillates with simple harmonic motion: $\bar{\mathcal{E}} = 2\overline{KE} = 2\left(\frac{k_B T}{2}\right) = k_B T$. To see why $\bar{\mathcal{E}} = 2\overline{KE}$ for a system with simple harmonic motion, consider a spring with potential energy $PE = \frac{1}{2}kx^2$ and kinetic energy $KE = \frac{1}{2}mv_x^2$, where $x = A\sin(\omega_0 t)$ and $v_x = \frac{dx}{dt} = \frac{d}{dt}A\sin(\omega_0 t) = A\omega_0\cos(\omega_0 t)$. The total energy of the spring is $\mathcal{E} = PE + KE = \frac{1}{2}kA^2\sin^2(\omega_0 t) + \frac{1}{2}mA^2\omega_0^2\cos^2(\omega_0) = \frac{1}{2}kA^2$ since $\omega_0 = \sqrt{\frac{k}{m}}$ (square this to get $\omega_0^2 = \frac{k}{m}$). The average kinetic energy is $\overline{KE} = \frac{1}{2}m\overline{v_x^2} = \frac{1}{2}mA^2\omega_0^2\overline{\cos^2(\omega_0 t)} = \frac{1}{4}kA^2 = \frac{\mathcal{E}}{2}$. Thus, $\bar{\mathcal{E}} = 2\overline{KE}$. (We used the fact that the average value of the cosine function over one cycle is one-half, $\overline{\cos^2\theta} = \frac{1}{2}$, and again used $\omega_0^2 = \frac{k}{m}$.)

The Number of Electromagnetic Standing Waves in a Cavity

Let N represent the number of electromagnetic standings waves in the cavity with a frequency less than or equal to f. (As we will explore later, this will help us determine the total energy of all of the standing waves per frequency.) From the equation at the bottom of page 88, the frequency of an electromagnetic standing wave is given by

$$f = \frac{c}{2L}\sqrt{n_x^2 + n_y^2 + n_z^2}$$

where n_x, n_y, and n_z are positive integers. Each lattice point illustrated below corresponds to a different possible value of (n_x, n_y, n_z). For a typical cavity in a laboratory, n_x, n_y, and n_z are large numbers. For example, since $c = 2.99792458 \times 10^8$ m/s, for a cube with an edge length of about one meter and a frequency of visible light on the order of 10^{14} Hz, the value of $\sqrt{n_x^2 + n_y^2 + n_z^2}$ is about 10^6.



One-eighth of a sphere is illustrated above. It's only an eighth of a sphere since n_x, n_y, and n_z can't be negative. One-eighth of a sphere with radius r has volume $V_s = \frac{1}{8}\left(\frac{4}{3}\pi r^3\right) = \frac{1}{6}\pi r^3$. Note that the radius of the sphere above is $r = \sqrt{n_x^2 + n_y^2 + n_z^2}$. The frequency can thus be expressed in terms of the radius as $f = \frac{c}{2L} r$. Solving for the radius, $r = \frac{2L}{c} f$. The volume of the eighth of the sphere is $V_s = \frac{1}{6}\pi r^3 = \frac{1}{6}\pi\left(\frac{2L}{c} f\right)^3 = \frac{4}{3}\pi\frac{L^3}{c^3} f^3$. The volume of the cavity is $V_c = L^3$, such that $V_s = \frac{4}{3}\pi\frac{V_c}{c^3} f^3$. The number of lattice points enclosed by the sphere equals the volume of the sphere times the density of lattice points. The density of lattice points equals the number of lattice points per unit volume, which is one. (The unit cube has 8 lattice points at its corners, but only one-eighth of each point is actually "inside" of the unit cube, and 8 times $\frac{1}{8}$ equals one. The density of lattice points is a standard crystal structure calculation.) The number of lattice points enclosed by the sphere thus equals $V_s(1) = \frac{4}{3}\pi\frac{V_c}{c^3} f^3$.

The number of electromagnetic standing waves in the cavity with a frequency less than or equal to f is actually twice the number of lattice points. Why? Each electromagnetic standing wave has two different polarization states. (For example, for an electromagnetic wave that is circularly polarized, there are two different ways that the electric field can rotate: If the wave travels along $+z$, the electric field can rotate clockwise or counterclockwise in the xy plane.) Therefore, we must multiply the number of lattice points, $\frac{4}{3}\pi\frac{V_c}{c^3}f^3$, by 2 in order to determine the number of electromagnetic standing waves in the cavity with a frequency less than or equal to f.

$$N = 2\left(\frac{4}{3}\pi\frac{V_c}{c^3}f^3\right) = \frac{8}{3}\pi\frac{V_c}{c^3}f^3$$

That's the number of electromagnetic standing waves with a frequency less than or equal to f. It would be more useful to know how many have a frequency equal to f (instead of knowing how many have a frequency between 0 and f). We can determine this by taking a derivative of N with respect to f.

$$\frac{dN}{df} = \frac{d}{df}\left(\frac{8}{3}\pi\frac{V_c}{c^3}f^3\right) = \left(\frac{8}{3}\pi\frac{V_c}{c^3}\right)\frac{d}{df}(f^3) = \left(\frac{8}{3}\pi\frac{V_c}{c^3}\right)(3f^2) = \frac{8\pi V_c}{c^3}f^2$$

Since $\frac{dN}{df} = \frac{8\pi V_c}{c^3}f^2$, we may write:

$$dN = \frac{8\pi V_c}{c^3}f^2 df$$

Here, dN is the (infinitesimal) number of electromagnetic standing waves in the cavity with a frequency equal to f (meaning that they have a frequency between f and $f + df$, which is an infinitesimal range df).

There is an alternative way of arriving at the previous equation. Instead of asking how many lattice points lie in the sphere, we could have asked how many lattice points lie in a thin spherical shell of thickness dr. Since the surface area of a sphere is $4\pi r^2$, the surface area of one-eighth of a spherical shell is $\frac{1}{8}4\pi r^2 = \frac{1}{2}\pi r^2$. The volume of an infinitesimally thin spherical shell of thickness dr is then $dV_s = \frac{1}{2}\pi r^2 dr$. Recall that $r = \frac{2L}{c}f$, such that $dr = \frac{2L}{c}df$. Plug these expressions into dV_s to see that $dV_s = \frac{1}{2}\pi\left(\frac{2L}{c}f\right)^2\frac{2L}{c}df = \frac{1}{2}\pi\left(\frac{4L^2}{c^2}f^2\right)\frac{2L}{c}df = \frac{4\pi L^3}{c^3}f^2 df$. Since $V_c = L^3$, this becomes $dV_s = \frac{4\pi V_c}{c^3}f^2 df$. Multiply by the density of lattice points (which equals one) to get the number of lattice points, and then multiply by two for the two different polarization states of the electromagnetic waves: $dN = dV_s(1)(2) = \frac{8\pi V_c}{c^3}f^2 df$. This is exactly the same result that we obtained with the previous method.

The Rayleigh-Jeans Formula (Classical Prediction)

We define ρ_f as energy per unit volume per unit frequency. The classical prediction for ρ_f for electromagnetic standing waves in the cavity is given by:

$$\rho_f \, df = \frac{(\text{average energy of one standing wave})(\text{number of standing waves in } df)}{\text{volume of the cavity}}$$

Recall that the average energy of one standing wave is $\overline{\mathcal{E}} = 2\overline{KE} = k_B T$ (see the bottom of page 89) and that the number of standing waves with a frequency between f and $f + df$ is $dN = \frac{8\pi V_c}{c^3} f^2 df$ (see page 91).

$$\rho_f \, df = \frac{(k_B T)\left(\frac{8\pi V_c}{c^3} f^2 df\right)}{V_c} = \frac{8\pi}{c^3} k_B T f^2 df$$

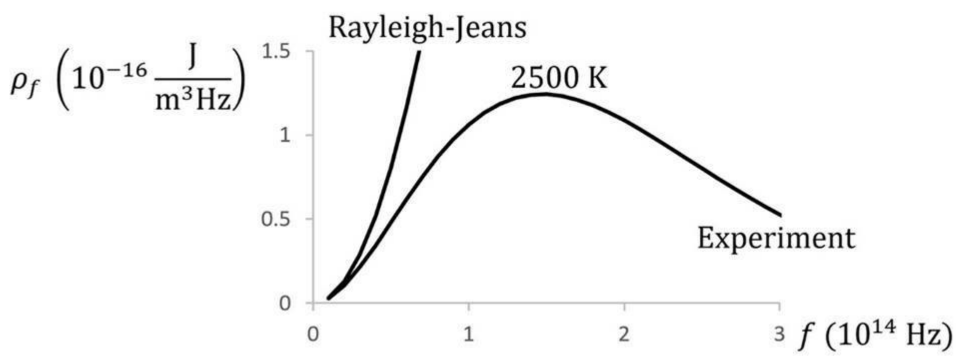
Compare the two sides of the above equation to see that ρ_f is:

$$\rho_f = \frac{8\pi}{c^3} k_B T f^2$$

This is the **Rayleigh-Jeans** formula for blackbody radiation. Note that this equation **doesn't** agree with experimental results (as we will explore next). Classical physics fails to explain blackbody radiation. We must apply principles of quantum physics instead of classical physics in order to explain experimental data relating to blackbody radiation.

The Ultraviolet Catastrophe

The graph below shows experimental data at 2500 K compared to the Rayleigh-Jeans formula. Although there is agreement at low frequencies, the Rayleigh-Jeans prediction from classical physics is in complete disagreement with experiment at high frequencies. This is termed the **ultraviolet catastrophe**, since the discrepancy is quite noticeable for ultraviolet frequencies. Not only is there disagreement between classical theory and experiment, but the Rayleigh-Jeans prediction itself is unreasonable: Since ρ_f is proportional to f^2 in the Rayleigh-Jeans formula, it predicts that the energy grows to infinity as the frequency increases. Experiment shows that ρ_f remains finite.



Planck's Solution to the Ultraviolet Catastrophe

Max Planck solved the blackbody radiation problem by postulating that the electromagnetic standing waves can't have an arbitrary value of energy, but that their energy is **quantized**. This means that energy is a discrete quantity, rather than a continuous variable.

$$\mathcal{E}_n = nhf \quad , \quad n = 0,1,2,3,\cdots$$

Energy comes in discrete bundles. The minimum possible nonzero value of the energy of an electromagnetic standing wave is hf, where $h = 6.626 \times 10^{-34}$ J·s is **Planck's constant**. The energy must equal an integer multiple of hf, such as hf, $2hf$, $3hf$, $4hf$, and so on. It would be possible for the energy to equal $711hf$, but not $42.5hf$. The smallest nonzero value for the energy, equal to hf, is called a **quantum**. This simple postulate—that the energy of the electromagnetic standing waves is **quantized**—solves the ultraviolet catastrophe.

Average Energy of Electromagnetic Standing Waves (Classical Physics)

To understand how the quantization of energy solves the blackbody radiation problem, we will first examine the classical result for the average energy of the electromagnetic standing waves in the cavity, and then we will redo the calculation using Planck's postulate.

In statistical thermodynamics, the average energy is given by the following formula:

$$\bar{\mathcal{E}} = \frac{\int_{\mathcal{E}=0}^{\infty} \mathcal{E} e^{-\mathcal{E}/k_B T} \, d\mathcal{E}}{\int_{\mathcal{E}=0}^{\infty} e^{-\mathcal{E}/k_B T} \, d\mathcal{E}}$$

First evaluate the bottom integral:

$$\int_{\mathcal{E}=0}^{\infty} e^{-\mathcal{E}/k_B T} \, d\mathcal{E} = -k_B T \left[e^{-\mathcal{E}/k_B T}\right]_{\mathcal{E}=0}^{\infty} = -k_B T \left[\frac{1}{e^{\mathcal{E}/k_B T}}\right]_{\mathcal{E}=0}^{\infty} = -k_B T \left(\frac{1}{e^{\infty}} - \frac{1}{e^{0}}\right) = -k_B T(0-1) = k_B T$$

For the top integral, integrate by parts using $u = \mathcal{E}$ and $dv = e^{-\mathcal{E}/k_B T} d\mathcal{E}$ such that $du = d\mathcal{E}$ and $v = -k_B T e^{-\mathcal{E}/k_B T}$.

$$\int_{\mathcal{E}=0}^{\infty} \mathcal{E} e^{-\mathcal{E}/k_B T} \, d\mathcal{E} = \int_{i}^{f} u \, dv = [uv]_i^f - \int_{i}^{f} v \, du = \left[-k_B T \mathcal{E} e^{-\mathcal{E}/k_B T}\right]_{\mathcal{E}=0}^{\infty} - \int_{\mathcal{E}=0}^{\infty} -k_B T e^{-\mathcal{E}/k_B T} \, d\mathcal{E}$$

$$= -k_B T \left(\lim_{\mathcal{E}\to\infty} \frac{\mathcal{E}}{e^{\mathcal{E}/k_B T}} - \lim_{\mathcal{E}\to 0} \frac{\mathcal{E}}{e^{\mathcal{E}/k_B T}}\right) + k_B T \int_{\mathcal{E}=0}^{\infty} e^{-\mathcal{E}/k_B T} \, d\mathcal{E} = -k_B T \left(\lim_{\mathcal{E}\to\infty} \frac{\frac{d}{d\mathcal{E}} \mathcal{E}}{\frac{d}{d\mathcal{E}} e^{\mathcal{E}/k_B T}} - \frac{0}{e^{0}}\right) + k_B T (k_B T)$$

$$= -k_B T \left(\frac{1}{\frac{1}{k_B T} e^{\infty}} - \frac{0}{1} \right) + k_B^2 T^2 = -k_B T(0 - 0) + k_B^2 T^2 = k_B^2 T^2$$

Note that we applied l'Hôpital's rule to the first limit. The average energy thus equals:

$$\overline{\mathcal{E}} = \frac{\int_{\mathcal{E}=0}^{\infty} \mathcal{E} e^{-\mathcal{E}/k_B T} \, d\mathcal{E}}{\int_{\mathcal{E}=0}^{\infty} e^{-\mathcal{E}/k_B T} \, d\mathcal{E}} = \frac{k_B^2 T^2}{k_B T} = k_B T \quad \text{(classical prediction)}$$

Average Energy of Electromagnetic Standing Waves (Quantum Physics)

The previous result is the classical result from the law of equipartition of energy (discussed earlier). Now we will see how Planck's quantization of energy corrects the calculation.

According to Planck, energy is quantized: $\mathcal{E}_n = nhf$ where n is a nonnegative integer. Since energy can't have any possible value, we shouldn't be doing an integral; we should instead be doing a sum.

$$\bar{\mathcal{E}} = \frac{\sum_{n=0}^{\infty} \mathcal{E}_n \, e^{-\mathcal{E}_n/k_B T}}{\sum_{n=0}^{\infty} e^{-\mathcal{E}_n/k_B T}} = \frac{\sum_{n=0}^{\infty} nhf \, e^{-nhf/k_B T}}{\sum_{n=0}^{\infty} e^{-nhf/k_B T}} = hf \frac{\sum_{n=0}^{\infty} n \, e^{-nhf/k_B T}}{\sum_{n=0}^{\infty} e^{-nhf/k_B T}}$$

This looks much simpler if we make the following definition: $\alpha = \frac{hf}{k_B T}$.

$$\bar{\mathcal{E}} = hf \frac{\sum_{n=0}^{\infty} n \, e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}}$$

The trick to evaluating this is to consider the following derivative (involving a natural log). Recall from calculus that $\frac{d}{dx} \ln x = \frac{1}{x}$. Apply the chain rule: $\frac{d}{dx} \ln u = \frac{d}{du}\frac{du}{dx} \ln u = \frac{1}{u}\frac{du}{dx}$.

$$\frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha} = \frac{\frac{d}{d\alpha}\sum_{n=0}^{\infty} e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} = \frac{\sum_{n=0}^{\infty} \frac{d}{d\alpha} e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} = \frac{-\sum_{n=0}^{\infty} n e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}}$$

Substitute the previous equation into the equation for the average energy.

$$\bar{\varepsilon} = hf\left(-\frac{d}{d\alpha}\ln\sum_{n=0}^{\infty} e^{-n\alpha}\right) = -hf\frac{d}{d\alpha}\ln\sum_{n=0}^{\infty} e^{-n\alpha}$$

Recall from calculus the Taylor series expansion of $(1-x)^{-1}$ for $|x| < 1$.

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + x^4 + \cdots$$

Write out the previous sum and compare it with the Taylor series expansion above.

$$\sum_{n=0}^{\infty} e^{-n\alpha} = 1 + e^{-\alpha} + e^{-2\alpha} + e^{-3\alpha} + e^{-4\alpha} + \cdots$$

The two series are identical if we let $x = e^{-\alpha}$. We may thus write the average energy as:

$$\bar{\varepsilon} = -hf\frac{d}{d\alpha}\ln(1-e^{-\alpha})^{-1}$$

Once again, apply the chain rule when taking the derivative: $\frac{d}{dx}\ln u = \frac{d}{du}\frac{du}{dx}\ln u = \frac{1}{u}\frac{du}{dx}$.

$$\bar{\varepsilon} = -\frac{hf}{(1-e^{-\alpha})^{-1}}\frac{d}{d\alpha}(1-e^{-\alpha})^{-1} = -\frac{hf}{(1-e^{-\alpha})^{-1}}(-1)(1-e^{-\alpha})^{-2}[-(-e^{-\alpha})]$$

$$= hf\frac{(1-e^{-\alpha})^{-2}}{(1-e^{-\alpha})^{-1}}(e^{-\alpha}) = hf(1-e^{-\alpha})^{-1}(e^{-\alpha}) = \frac{hfe^{-\alpha}}{1-e^{-\alpha}} = \frac{hf}{e^{\alpha}-1}$$

We multiplied by $\frac{e^{\alpha}}{e^{\alpha}}$, noting that $e^{\alpha}e^{-\alpha} = 1$. Recall that $\alpha = \frac{hf}{k_B T}$. The average energy is:

$$\bar{\varepsilon} = \frac{hf}{e^{hf/k_B T} - 1}$$

Energy Density per Frequency for a Blackbody Cavity (Quantum Prediction)

Now we can correct the Rayleigh-Jeans calculation from page 92. All we will do is replace the classical average energy $\overline{\mathcal{E}} = k_B T$ with the quantum average energy

$$\overline{\mathcal{E}} = \frac{hf}{e^{hf/k_B T} - 1}$$

in the equation for ρ_f for electromagnetic standing waves in the cavity (see page 92).

$$\rho_f \, df = \frac{(\text{average energy of one standing wave})(\text{number of standing waves in } df)}{\text{volume of the cavity}}$$

$$\rho_f \, df = \frac{\left(\frac{hf}{e^{hf/k_B T} - 1}\right)\left(\frac{8\pi V_c}{c^3} f^2 df\right)}{V_c} = \frac{8\pi}{c^3} \frac{h}{e^{hf/k_B T} - 1} f^3 df$$

Compare the two sides of the above equation to see that ρ_f is:

$$\rho_f = \frac{8\pi}{c^3} \frac{h}{e^{hf/k_B T} - 1} f^3$$

This equation agrees with experimental data for blackbody radiation.

Energy Density and Intensity

Note that there are two similar quantities related to the energy of a blackbody cavity:

- The **energy density**, u, is the energy per unit volume inside of the cavity: $u = U/V$. The energy density has SI units of J/m^3.

- The **intensity**, I, is the power per unit area escaping through the cavity hole: $I = P/A$. The intensity has SI units of W/m^2.

Energy Density Per Unit Frequency and Spectral Radiance

Note that there are two quantities which are very similar to the energy density ($u = \frac{U}{V}$) and intensity ($I = \frac{P}{A}$), but which are defined per unit frequency:

- ρ_f is the **energy density per unit frequency**. It has units of $\frac{\text{J}}{\text{m}^3\,\text{Hz}}$.
- R_f is the **spectral radiance** (intensity per unit frequency). It has units of $\frac{\text{W}}{\text{m}^2\,\text{Hz}}$.

The **energy density** (u) is related to the energy density per unit frequency (ρ_f) by:

$$u = \int_{f=0}^{\infty} \rho_f \, df$$

The **intensity** (I) is related to the spectral radiance (R_f) by:

$$I = \int_{f=0}^{\infty} R_f \, df$$

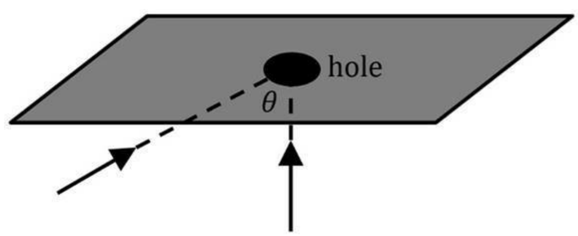
Recall that power is the instantaneous rate at which work is done: $P = \frac{dW}{dt}$. Since energy is the ability to do work, the instantaneous rate at which energy is transferred through the hole also corresponds to power. The SI unit of power is $\mathrm{W} = \frac{\mathrm{J}}{\mathrm{s}}$, such that the SI units of intensity may be written as $\frac{\mathrm{W}}{\mathrm{m}^2} = \frac{\mathrm{J}}{\mathrm{m}^2\mathrm{s}}$. Comparing the units of intensity $\left(\frac{\mathrm{J}}{\mathrm{m}^2\mathrm{s}}\right)$ with the units of energy density $\left(\frac{\mathrm{J}}{\mathrm{m}^3}\right)$, you can see that the units of intensity equal the units of energy density times the units of speed: $\frac{\mathrm{J}}{\mathrm{m}^2\mathrm{s}} = \frac{\mathrm{J}}{\mathrm{m}^3} \cdot \frac{\mathrm{m}}{\mathrm{s}}$. The units of ρ_f and R_f are similarly related: $\frac{\mathrm{J}}{\mathrm{m}^2\mathrm{s}\cdot\mathrm{Hz}} = \frac{\mathrm{J}}{\mathrm{m}^3\cdot\mathrm{Hz}} \cdot \frac{\mathrm{m}}{\mathrm{s}}$. The quantities ρ_f and R_f are proportional, and the proportionality factor equals one-fourth the speed of light in vacuum:

$$R_f = \frac{c}{4}\rho_f$$

Note that different textbooks use different symbols (and even different names) for energy density and intensity. See the section devoted to notation and terminology on pages 105-106.

Why Does $R_f = \frac{c}{4}\rho_f$?

Imagine that the cavity is shaped like a cube with a small circular hole in the horizontal top side, as illustrated below. In thermal equilibrium, the electromagnetic radiation inside the cavity will be homogeneous (uniform throughout) and isotropic (the same in all directions). Thus, there will be electromagnetic waves incident upon the hole from all directions in equal measure. Consider an electromagnetic wave propagating straight upward towards the hole, and then consider another electromagnetic wave propagating at an angle towards the hole. The hole is effectively smaller in the latter case by a factor of $\cos\theta$, where θ is the angle that the electromagnetic wave's velocity makes with the vertical.



There are two factors of one-half involved:

- On average, one-half of the electromagnetic waves will be heading upward and one-half of the electromagnetic waves will be heading downward. Thus, only one-half of the electromagnetic waves have a chance of escaping through the hole at any moment.
- A second factor of one-half comes from the effective size of the hole, on average. The effective size of the hole is proportional to $\cos\theta$.

The overall factor is found by integrating over solid angle, where $d\Omega = \sin\theta\, d\theta d\varphi$ in spherical coordinates. We will find the ratio of the anisotropy (angular dependence) of the radiation heading towards the hole to the ratio of the anisotropy of the distribution of radiation within the cavity. In the numerator, we integrate from $\theta = 0$ to $\pi/2$ because only half of the waves are heading upward (the waves from $\pi/2$ to π are heading downward, with zero chance of escaping through the hole). For the top integral, let $q = \cos\theta$, for which $dq = -\sin\theta\, d\theta$. The new limits are from $q_i = \cos 0 = 1$ to $q_f = \cos\left(\frac{\pi}{2}\right) = 0$.

$$\frac{\int \cos\theta\, \Omega}{\int d\Omega} = \frac{\int_{\varphi=0}^{2\pi}\int_{\theta=0}^{\pi/2} \cos\theta \sin\theta\, d\theta\, d\varphi}{\int_{\varphi=0}^{2\pi}\int_{\theta=0}^{\pi} \sin\theta\, d\theta d\varphi} = \frac{2\pi \int_{q=1}^{0} -q\, dq}{2\pi[-\cos\theta]_{\theta=0}^{\pi}} = \frac{-\left[\frac{q^2}{2}\right]_{q=1}^{0}}{-[\cos\theta]_{\theta=0}^{\pi}} = \frac{\left[\frac{q^2}{2}\right]_{q=1}^{0}}{[\cos\theta]_{\theta=0}^{\pi}}$$

$$= \frac{\frac{0^2}{2} - \frac{1^2}{2}}{\cos\pi - \cos 0} = \frac{-1/2}{-1-1} = \frac{-1/2}{-2} = \frac{1}{4}$$

The factor of c comes from how fast the radiation escapes through the hole, while the factor of $\frac{1}{4}$ comes from the ratio of solid angles (as shown in the calculation above).

Dependence on Frequency and Wavelength

In terms of frequency, the energy density per unit frequency and spectral radiance are:

$$\rho_f = \frac{8\pi}{c^3}\frac{h}{e^{hf/k_BT}-1}f^3 \quad , \quad R_f = \frac{c}{4}\rho_f = \frac{2\pi}{c^2}\frac{h}{e^{hf/k_BT}-1}f^3$$

Sometimes, it is useful to work with energy density per unit wavelength (ρ_λ) instead of energy density per unit frequency (ρ_f). There is a corresponding form of the spectral radiance that is intensity per unit wavelength (R_λ) instead of intensity per unit frequency (R_f). The total energy density over all frequencies or all wavelengths comes out the same either way:

$$u = \int_{f=0}^{\infty} \rho_f \, df = \int_{\lambda=0}^{\infty} \rho_\lambda \, d\lambda \quad , \quad I = \int_{f=0}^{\infty} R_f \, df = \int_{\lambda=0}^{\infty} R_\lambda \, d\lambda$$

Therefore, the quantities per unit frequency and per unit wavelength are related by

$$\rho_f df = -\rho_\lambda d\lambda \quad , \quad R_f df = -R_\lambda d\lambda$$

Why is there a minus sign? Because $\lambda f = c$: As f increases, λ decreases, such that $\frac{df}{d\lambda} < 0$.

$$\rho_\lambda = -\rho_f \frac{df}{d\lambda} \quad , \quad R_\lambda = -R_f \frac{df}{d\lambda}$$

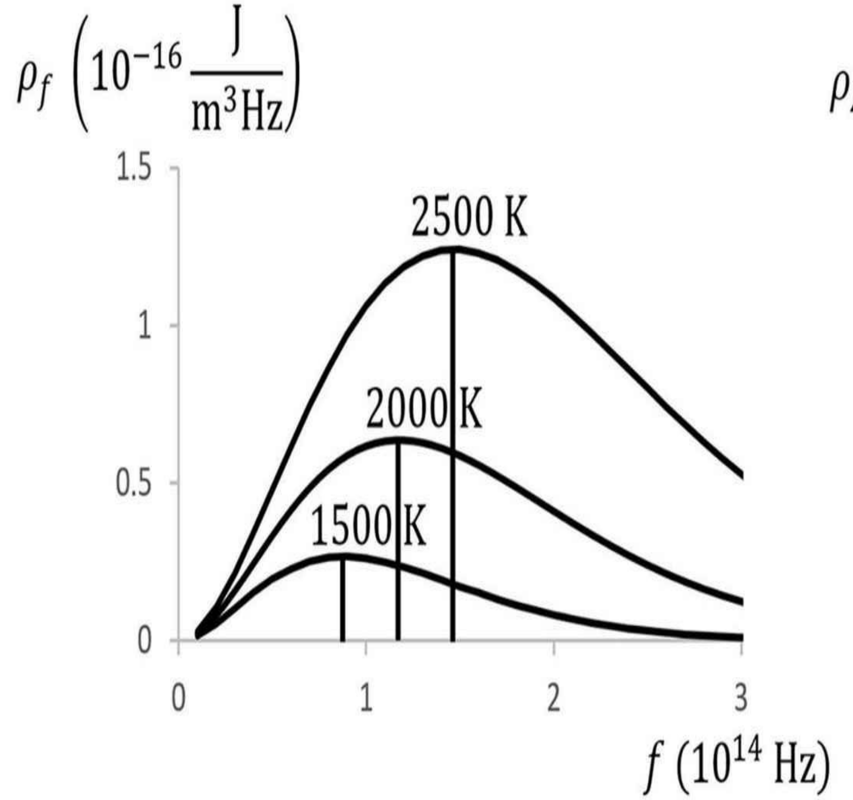
Since wavelength and frequency are related to the speed of light by $f = \frac{c}{\lambda}$, it follows that

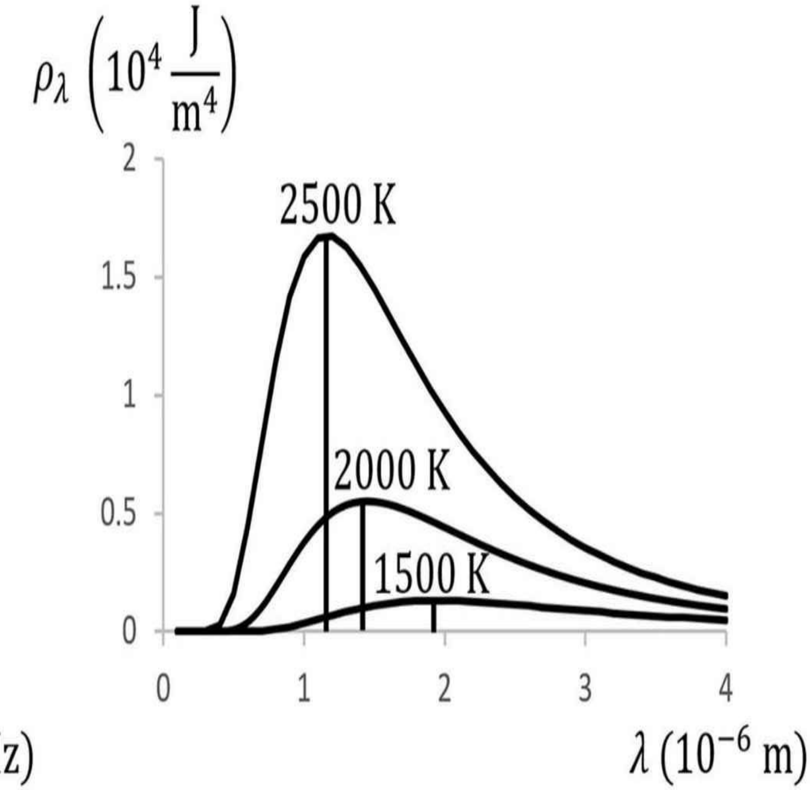
$$\frac{df}{d\lambda} = \frac{d}{d\lambda}\left(\frac{c}{\lambda}\right) = -\frac{c}{\lambda^2}$$

If we make the substitutions $f = \frac{c}{\lambda}$ and $\frac{df}{d\lambda} = -\frac{c}{\lambda^2}$, we get the following equations:

$$\rho_\lambda = -\rho_f \frac{df}{d\lambda} = \frac{c\rho_f}{\lambda^2} = \frac{8\pi}{\lambda^5} \frac{hc}{e^{hc/\lambda k_B T} - 1} \quad , \quad R_\lambda = -R_f \frac{df}{d\lambda} = \frac{cR_f}{\lambda^2} = \frac{2\pi}{\lambda^5} \frac{hc^2}{e^{hc/\lambda k_B T} - 1} = \frac{c}{4}\rho_\lambda$$

Note that the units of ρ_λ and R_λ are $\frac{\text{J}}{\text{m}^4}$ and $\frac{\text{W}}{\text{m}^3}$, whereas the units of ρ_f and R_f are $\frac{\text{J}}{\text{m}^3\text{Hz}}$ and $\frac{\text{W}}{\text{m}^2\text{Hz}}$.





Wien's Displacement Law

The vertical lines on the right graph above indicate the wavelength (λ_m) that maximizes the energy density per unit wavelength. (It also maximizes the corresponding spectral radiance, R_λ.) It is incorrect to think of λ_m as a "maximum wavelength" because the maximum wavelength is actually infinite. Rather, λ_m is the wavelength for which ρ_λ (and thus R_λ) is maximum.

According to **Wien's displacement law**, λ_m and temperature are inversely proportional: As temperature increases, λ_m decreases. Note that T is the absolute temperature in Kelvin (K).

$$\lambda_m T = \text{const.}$$

(You see inverted behavior on the left graph above because frequency and wavelength are inversely proportional: $f = \frac{c}{\lambda}$. However, it's a little more complicated than this: See Problem 5.)

Stefan's Law

Another thing that you can see visually in the graphs above is that the area under the curve depends dramatically on temperature. Since the area under a curve represents the definite integral of the function, area corresponds to $u = \int_{f=0}^{\infty} \rho_f \, df = \int_{\lambda=0}^{\infty} \rho_\lambda \, d\lambda$. Since $R_f = \frac{c}{4}\rho_f$, graphs of R_f rather than ρ_f exhibit similar behavior, but in that case area corresponds to $I = \int_{f=0}^{\infty} R_f \, df = \int_{\lambda=0}^{\infty} R_\lambda \, d\lambda$.

Stefan's law states that the total power (per unit area) radiated is proportional to the fourth power of the absolute temperature in Kelvin (K):

$$I = \sigma T^4$$

where $\sigma = 5.67 \times 10^{-8} \frac{\mathrm{W}}{\mathrm{m^2 K^4}}$ is the **Stefan-Boltzmann constant.**

Deriving Stefan's Law

We will use the equation for spectral radiance from page 97.

$$R_f = \frac{2\pi}{c^2}\frac{h}{e^{hf/k_BT}-1}f^3$$

The integral of the spectral radiance (R_f) over frequency equals the total power per unit area or intensity (I).

$$I = \int_{f=0}^{\infty} R_f\, df = \frac{2\pi h}{c^2}\int_{f=0}^{\infty}\frac{f^3}{e^{hf/k_BT}-1}df$$

Make the substitution $x = \frac{hf}{k_BT}$, for which $dx = \frac{h}{k_BT}df$. It follows that $f = \frac{k_BTx}{h}$ and $df = \frac{k_BT}{h}dx$.

$$I = \frac{2\pi h}{c^2}\int_{x=0}^{\infty}\frac{\left(\frac{k_BTx}{h}\right)^3}{e^x-1}\left(\frac{k_BT}{h}\right)dx = \left(\frac{2\pi h}{c^2}\right)\left(\frac{k_B^3T^3}{h^3}\right)\left(\frac{k_BT}{h}\right)\int_{x=0}^{\infty}\frac{x^3}{e^x-1}dx = \frac{2\pi k_B^4T^4}{c^2h^3}\int_{x=0}^{\infty}\frac{x^3}{e^x-1}dx$$

This isn't an integral that you learn how to solve with standard techniques of integration in a calculus course. Rather, it's one of a variety of special functions covered in mathematical physics. Consider the integral form of the Riemann zeta function, $\zeta(z)$, given below, which involves the gamma function, $\Gamma(z)$, which is also known as the factorial function, where $z > 1$.

$$\zeta(z) = \frac{1}{\Gamma(z)} \int_{x=0}^{\infty} \frac{x^{z-1}}{e^x - 1} dx \quad , \quad \Gamma(z) = \int_{x=0}^{\infty} e^{-x} x^{z-1}\, dx$$

Our integral involves the special case corresponding to $z = 4$.

$$\zeta(4) = \frac{1}{\Gamma(4)} \int_{x=0}^{\infty} \frac{x^{4-1}}{e^x - 1} dx = \frac{1}{\Gamma(4)} \int_{x=0}^{\infty} \frac{x^3}{e^x - 1} dx \quad , \quad \Gamma(4) = \int_{x=0}^{\infty} e^{-x} x^{4-1}\, dx = \int_{x=0}^{\infty} e^{-x} x^3\, dx$$

The Riemann zeta function can alternatively be expressed as a sum (for $z > 1$):

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$$

For $z = 4$, it can be shown that in this case the series converges to $\frac{\pi^4}{90}$.

$$\zeta(4) = \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \frac{1}{5^4} + \cdots = \frac{\pi^4}{90}$$

Now we can replace $\zeta(4)$ with $\frac{\pi^4}{90}$ to solve for our integral.

$$\zeta(4) = \frac{1}{\Gamma(4)} \int_{x=0}^{\infty} \frac{x^3}{e^x - 1} dx$$

$$\int_{x=0}^{\infty} \frac{x^3}{e^x - 1} dx = \zeta(4)\Gamma(4) = \frac{\pi^4}{90}\Gamma(4)$$

For an argument that is a positive integer, the gamma function is related to the factorial: $\Gamma(n) = (n-1)!$. Set $n = 4$ for our case: $\Gamma(4) = (4-1)! = 3! = (3)(2)(1) = 6$. Plug $\Gamma(4) = 3! = 6$ into the previous equation.

$$\int_{x=0}^{\infty} \frac{x^3}{e^x - 1} dx = \left(\frac{\pi^4}{90}\right)(3!) = \left(\frac{\pi^4}{90}\right)(6) = \frac{\pi^4}{15}$$

Replacing our integral with $\frac{\pi^4}{15}$, we get

$$I = \frac{2\pi k_B^4 T^4}{c^2 h^3} \int_{x=0}^{\infty} \frac{x^3}{e^x - 1} dx = \frac{2\pi k_B^4 T^4}{c^2 h^3}\left(\frac{\pi^4}{15}\right) = \frac{2\pi^5 k_B^4 T^4}{15 c^2 h^3} = \sigma T^4$$

where the **Stefan-Boltzmann constant** is

$$\sigma = \frac{2\pi^5 k_B^4}{15 c^2 h^3} = \frac{2\pi^5 (1.38 \times 10^{-23}\ \text{J/K})^4}{15(2.9979 \times 10^8\ \text{m/s})^2 (6.626 \times 10^{-34}\ \text{J}\cdot\text{s})^3} = 5.67 \times 10^{-8} \frac{\text{W}}{\text{m}^2\text{K}^4}$$

Deriving Wien's Displacement Law

We will use the equation for spectral radiance from the bottom of page 97.

$$R_\lambda = \frac{2\pi}{\lambda^5}\frac{hc^2}{e^{hc/\lambda k_B T} - 1}$$

Wien's displacement law concerns the value of λ that maximizes R_λ (or alternatively ρ_λ). Recall from calculus that the way to find a relative maximum is to set the first derivative of a function equal to zero. Take a derivative of R_λ with respect to λ.

$$\frac{dR_\lambda}{d\lambda} = \frac{d}{d\lambda}\left(\frac{2\pi}{\lambda^5}\frac{hc^2}{e^{hc/\lambda k_B T} - 1}\right) = 2\pi hc^2 \frac{d}{d\lambda}\left(\frac{1}{\lambda^5}\frac{1}{e^{hc/\lambda k_B T} - 1}\right) = 2\pi hc^2 \frac{d}{d\lambda}\left[\lambda^{-5}\left(e^{hc/\lambda k_B T} - 1\right)^{-1}\right]$$

Apply the product rule: $\frac{d}{d\lambda}(pq) = q\frac{dp}{d\lambda} + p\frac{dq}{d\lambda}$ with $p = \lambda^{-5}$ and $q = \left(e^{hc/\lambda k_B T} - 1\right)^{-1}$.

$$\frac{dR_\lambda}{d\lambda} = 2\pi hc^2(\lambda^{-5})\frac{d}{d\lambda}\left(e^{hc/\lambda k_B T} - 1\right)^{-1} + 2\pi hc^2\left(e^{hc/\lambda k_B T} - 1\right)^{-1}\frac{d}{d\lambda}(\lambda^{-5})$$

Apply the chain rule with $u = e^{hc/\lambda k_B T} - 1$ and $g = u^{-1}$ to write $\frac{d}{d\lambda}\left(e^{hc/\lambda k_B T} - 1\right)^{-1} = \frac{dg}{du}\frac{du}{d\lambda}$.

Note that $\frac{dg}{du} = \frac{d}{du}(u^{-1}) = -u^{-2}$ and $\frac{du}{d\lambda} = \frac{d}{d\lambda}\left(e^{hc/\lambda k_B T} - 1\right) = -\frac{hc}{\lambda^2 k_B T}e^{hc/\lambda k_B T}$.

$$\frac{dR_\lambda}{d\lambda} = -2\pi hc^2(\lambda^{-5})\left(e^{hc/\lambda k_B T} - 1\right)^{-2}\left(-\frac{hc}{\lambda^2 k_B T}e^{hc/\lambda k_B T}\right) + 2\pi hc^2\left(e^{hc/\lambda k_B T} - 1\right)^{-1}(-5\lambda^{-6})$$

To find the extreme values of R_λ, set the first derivative equal to zero: $\frac{dR_\lambda}{d\lambda} = 0$.

$$-2\pi hc^2(\lambda^{-5})\left(e^{hc/\lambda k_B T} - 1\right)^{-2}\left(-\frac{hc}{\lambda^2 k_B T}e^{hc/\lambda k_B T}\right) + 2\pi hc^2\left(e^{hc/\lambda k_B T} - 1\right)^{-1}(-5\lambda^{-6}) = 0$$

Divide both sides of the equation by $2\pi hc^2$. Note that $\left(e^{hc/\lambda k_B T} - 1\right)^{-n} = \frac{1}{\left(e^{hc/\lambda k_B T} - 1\right)^n}$.

$$-(\lambda^{-5})\frac{1}{(e^{hc/\lambda k_B T} - 1)^2}\left(-\frac{hc}{\lambda^2 k_B T}e^{hc/\lambda k_B T}\right) + \frac{1}{e^{hc/\lambda k_B T} - 1}(-5\lambda^{-6}) = 0$$

Add $-(\lambda^{-5})\frac{1}{\left(e^{hc/\lambda k_BT}-1\right)^2}\left(-\frac{hc}{\lambda^2 k_BT}e^{hc/\lambda k_BT}\right)$ to both sides of the equation.

$$\frac{1}{e^{hc/\lambda k_BT}-1}(-5\lambda^{-6}) = (\lambda^{-5})\frac{1}{(e^{hc/\lambda k_BT}-1)^2}\left(-\frac{hc}{\lambda^2 k_BT}e^{hc/\lambda k_BT}\right)$$

Multiply both sides of the equation by λ^6.

$$\frac{1}{e^{hc/\lambda k_BT}-1}(-5) = \frac{\lambda}{(e^{hc/\lambda k_BT}-1)^2}\left(-\frac{hc}{\lambda^2 k_BT}e^{hc/\lambda k_BT}\right)$$

Multiply both sides of the equation by $\left(e^{hc/\lambda k_BT}-1\right)^2$.

$$-5\left(e^{hc/\lambda k_BT}-1\right) = \lambda\left(-\frac{hc}{\lambda^2 k_BT}e^{hc/\lambda k_BT}\right)$$

Multiply both sides by negative one.

$$5\left(e^{hc/\lambda k_BT}-1\right) = \lambda\left(\frac{hc}{\lambda^2 k_BT}e^{hc/\lambda k_BT}\right)$$

Distribute the 5 and the λ.

$$5e^{hc/\lambda k_BT}-5 = \frac{hc}{\lambda k_BT}e^{hc/\lambda k_BT}$$

Make the substitution $x = \frac{hc}{\lambda k_B T}$.

$$5e^x - 5 = xe^x$$

Multiply both sides by e^{-x}. Note that $e^x e^{-x} = 1$.

$$5 - 5e^{-x} = x$$

This is a transcendental equation. If you apply a numerical technique (such as Newton's method) to find the roots of the equation $5 - 5e^{-x} - x = 0$, you should get $x \approx 4.96511$. Substitute $x \approx 4.96511$ into the equation $x = \frac{hc}{\lambda k_B T}$. We will replace λ with λ_m since λ_m is the wavelength that maximizes R_λ. (If you wish to prove that it is maximum, take a second derivative of R_λ with respect to λ, evaluate it at the value of λ_m that we obtain below, and check that the sign is negative. It is simpler to examine the graph or invoke some physics.)

$$4.96511 \approx \frac{hc}{\lambda_m k_B T}$$

$$\lambda_m T \approx \frac{hc}{4.96511 k_B} \approx 0.201405 \frac{hc}{k_B} = 0.201405 \frac{(6.626 \times 10^{-34}\ \text{J·s})\left(2.9979 \times 10^8\ \frac{\text{m}}{\text{s}}\right)}{(1.38 \times 10^{-23}\ \text{J/K})}$$

$$\lambda_m T \approx 0.00290\ \text{K·m}$$

Symbols and SI Units

Symbol	Name	SI Units
λ	wavelength	m
λ_m	the wavelength that maximizes ρ_λ (or R_λ)	m
f	frequency	Hz
c	speed of light in vacuum	m/s
T	absolute temperature	K
PE	potential energy	J
KE	kinetic energy	J
$\overline{KE}$	average kinetic energy	J

U	internal energy	J
$\mathcal{E}$	energy	J
$\bar{\mathcal{E}}$	average energy	J
$\mathcal{E}_n$	energy corresponding to n	J
h	Planck's constant	J·s or J/Hz
k_B	Boltzmann's constant	J/K
σ	Stefan-Boltzmann constant	$\frac{\mathrm{W}}{\mathrm{m^2K^4}}$

ρ_f	energy density per unit frequency	$\frac{\text{J}}{\text{m}^3\text{Hz}}$ or $\frac{\text{J}\cdot\text{s}}{\text{m}^3}$
ρ_λ	energy density per unit wavelength	$\frac{\text{J}}{\text{m}^4}$
R_f	spectral radiance: intensity per unit frequency	$\frac{\text{W}}{\text{m}^2\text{Hz}}$ or $\frac{\text{J}}{\text{m}^2}$
R_λ	spectral radiance: intensity per unit wavelength	$\frac{\text{W}}{\text{m}^3}$
u	energy density	$\frac{\text{J}}{\text{m}^3}$
I	intensity	$\frac{\text{W}}{\text{m}^2}$

P	power	W
ϵ	emissivity	unitless
e_1, e_2	emission rates of objects 1 and 2	W
a_1, a_2	absorption rates of objects 1 and 2	W
$\vec{\mathbf{E}}$	electric field	N/C or V/m
E	magnitude of electric field	N/C or V/m
$\vec{\mathbf{B}}$	magnetic field	T

L	length	m
A	area	m^2
V	volume	m^3
V_s	volume of a sphere	m^3
V_c	volume of the cavity	m^3
N	number of electromagnetic standing waves	unitless
n_d	number of degrees of freedom	unitless
n	a nonnegative integer	unitless
n_x, n_y, n_z	positive integers	unitless

n_x, n_y, n_z	positive integers	unitless
α, β, γ	direction cosines (the angle with x-, y-, and z-axes)	rad
Ω	solid angle	steradians
v	speed	m/s
A	amplitude	it depends
ω_0	angular frequency	rad/s
m	mass	kg
k	spring constant	N/m
$\zeta(z)$	Riemann zeta function	unitless
$\Gamma(z)$	gamma function (also called the factorial function)	unitless

Note: The following Greek letters are lowercase lambda (λ), lowercase sigma (σ), uppercase omega (Ω), lowercase omega (ω), lowercase rho (ρ), lowercase zeta (ζ), uppercase gamma (Γ), lowercase gamma (γ), lowercase alpha (α), lowercase beta (β), and two variations of epsilon ($\mathcal{E}$ and ϵ).

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^{8}\ \mathrm{m/s}$
Planck's constant	$h = 6.626 \times 10^{-34}\ \mathrm{J{\cdot}s}$
Boltzmann's constant	$k_B = 1.38 \times 10^{-23}\ \mathrm{J/K}$
Stefan-Boltzmann constant	$\sigma = 5.67 \times 10^{-8}\ \dfrac{\mathrm{W}}{\mathrm{m^2K^4}}$

Metric Prefixes and the Angstrom

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Prefix	Name	Power of 10
c	centi	10^{-2}
m	milli	10^{-3}
μ	micro	10^{-6}
n	nano	10^{-9}

Absolute Temperature

Equations of thermodynamics and statistical mechanics that involve temperature require using absolute temperature in **Kelvin** (K). If a problem gives you a temperature in degrees Celsius (°C) or degrees Fahrenheit (°F), first convert the temperature to Kelvin (K) using the following equations. Note that Kelvin is special: It's <u>**not**</u> preceded by a degree (°) symbol.

$$T_K = T_C + 273.15 \quad , \quad T_C = \frac{9}{5} T_F + 32$$

Notes Regarding Units

Recall that a Watt (W) is a Joule (J) per second: $1\text{ W} = 1\frac{\text{J}}{\text{s}}$. Since energy density ($u$) equals energy ($U$) per unit volume, $u = \frac{U}{V}$, its units are $\frac{\text{J}}{\text{m}^3}$. Since intensity ($I$) equals power ($P$) per unit area, $I = \frac{P}{A}$, its units are $\frac{\text{W}}{\text{m}^2}$. From the equations $u = \int_{f=0}^{\infty} \rho_f \, df = \int_{f=0}^{\infty} \rho_\lambda \, d\lambda$, it follows that ρ_f has units of $\frac{\text{J}}{\text{m}^3\text{Hz}}$ or $\frac{\text{J}\cdot\text{s}}{\text{m}^3}$ and that ρ_λ has units of $\frac{\text{J}}{\text{m}^4}$. From the equations $I = \int_{f=0}^{\infty} R_f \, df = \int_{f=0}^{\infty} R_\lambda \, d\lambda$ (or from $R_f = \frac{c}{4}\rho_f$ and $R_\lambda = \frac{c}{4}\rho_\lambda$), it follows that R_f has units of $\frac{\text{W}}{\text{m}^2\text{Hz}}$ or $\frac{\text{J}}{\text{m}^2}$ and that R_λ has units of $\frac{\text{W}}{\text{m}^3}$.

From Planck's law, $\mathcal{E}_n = nhf$, it follows that Planck's constant has units of $\frac{\text{J}}{\text{Hz}}$ or J·s.

Since $k_B T$ has units of energy, it follows that Boltzmann's constant has units of J/K (which are the same SI units as entropy from statistical thermodynamics).

The Stefan-Boltzmann constant is related to other constants via $\sigma = \frac{2\pi^5 k_B^4}{15c^2h^3}$. Its SI units work out to $\frac{\text{W}}{\text{m}^2\text{K}^4}$ because $I = \sigma T^4$.

Differences in Notation and Terminology

If you read a variety of modern physics textbooks or attend lectures by a variety of modern physics professors, you'll discover that the notation for blackbody radiation can vary wildly from one textbook or course to another. Unfortunately, the notation for this topic isn't at all standard. Even worse, many of the original textbooks on the subject tend to adopt notation that is fairly easy to misinterpret (but that's just the notation: many of the original modern physics textbooks are quite valuable sources of content knowledge).

The greatest source of confusion for students who are new to this topic has to do with the notation for distribution functions. When you take first-year calculus, you might write an equation like $g = \int_{x=0}^{\infty} f(x)\,dx$, but it would seem rather confusing to write $f = \int_{x=0}^{\infty} f(x)\,dx$. Study this equation closely until you can see what would likely cause confusion. Regarding distribution functions, this is exactly what some modern physics textbooks do:

- Some textbooks write $R_T = \int_{\nu=0}^{\infty} R_T(\nu)\,d\nu$, where R_T and $R_T(\nu)$ are different quantities with different units. Such textbooks call R_T the **radiance** (or emittance, etc.) and $R_T(\nu)$ the **spectral radiance** (or spectral emittance, etc.). What they call R_T, we call I: It is the **intensity** (which is the same as power per unit area) in units of $\frac{\mathrm{W}}{\mathrm{m}^2}$. What they call $R_T(\nu)$, we call $R_f(f)$. It is the **spectral radiance** (or spectral emittance) in units of $\frac{\mathrm{W}}{\mathrm{m}^2\mathrm{Hz}}$ or $\frac{\mathrm{J}}{\mathrm{m}^2}$. Another problem with that notation is that $R_T(\nu)$ and $R_T(\lambda)$ are also different quantities with different units: We call these R_f and R_λ with units $\frac{\mathrm{W}}{\mathrm{m}^2\mathrm{Hz}}$ or $\frac{\mathrm{J}}{\mathrm{m}^2}$ and $\frac{\mathrm{W}}{\mathrm{m}^3}$.
- Other textbooks use yet different notation for these same quantities, such as S_ν or B_ν for the **spectral radiance** (or spectral emittance) and $R(T)$ for **radiance** (or intensity).

- There are similar issues for the corresponding energy density quantities. For example, some textbooks write $\rho_T = \int_{\nu=0}^{\infty} \rho_T(\nu)\, d\nu$, where ρ_T is the **energy density** in units of $\frac{\text{J}}{\text{m}^3}$ while $\rho_T(\nu)$ is energy density per unit frequency in units of $\frac{\text{J}}{\text{m}^3\text{Hz}}$ or $\frac{\text{J}\cdot\text{s}}{\text{m}^3}$, whereas we use u for **energy density** in $\frac{\text{J}}{\text{m}^3}$ and ρ_f for energy density per unit frequency in units of $\frac{\text{J}}{\text{m}^3\text{Hz}}$.
- Some books call ρ_f the spectral energy density, and use the symbol u_ν or ρ_ν instead.
- In books on astronomy or astrophysics, it is common to work with **luminosity** (L) and **flux** (F), where luminosity is basically the power (P) in Watts and the flux is the intensity in $\frac{\text{W}}{\text{m}^2}$. In this context, F_λ is the spectral flux density in $\frac{\text{W}}{\text{m}^3}$ (what we called R_λ).

Another common difference in notation is that some books use f for **frequency** and other books use ν instead. Both symbols are used so often throughout the physics community and literature that it's worth being familiar with both. If you work with a multi-volume textbook that covers first-year physics, the modern physics chapters are likely to use f for frequency in order to be consistent with earlier chapters. When you learn about waves, f is common for frequency because it would be easy for first-year physics students to confuse the Greek letter nu (ν) with speed (v), especially after some sloppy handwriting. In more advanced texts, the symbol ν is common for frequency, but there are some exceptions.

Where we have used L for edge length, some texts prefer a, so that the cavity volume is a^3 instead of L^3 (for a cavity in the shape of a cube).

We chose to write a script $\mathcal{E}$ symbol for energy instead of E such that the symbol for energy ($\mathcal{E}$) would look different from the symbol for the magnitude of electric field (E). We did that for *this chapter only* (since in other chapters we won't be working with electric field).

Strategy for Solving Thermal Radiation Problems

How you solve a problem involving thermal radiation depends on which kind of problem it is:

- You will probably need to apply one or more of the following equations.

$$\lambda_m T = \text{const.} \quad \text{(Wien's law)} \quad , \quad I = \sigma T^4 \quad \text{(Stefan's law)}$$

$$\rho_f = \frac{8\pi}{c^3}\frac{h}{e^{hf/k_BT}-1}f^3 \quad , \quad R_f = \frac{c}{4}\rho_f = \frac{2\pi}{c^2}\frac{h}{e^{hf/k_BT}-1}f^3$$

$$\rho_\lambda = -\rho_f\frac{df}{d\lambda} = \frac{c\rho_f}{\lambda^2} = \frac{8\pi}{\lambda^5}\frac{hc}{e^{hc/\lambda k_BT}-1} \quad , \quad R_\lambda = -R_f\frac{df}{d\lambda} = \frac{cR_f}{\lambda^2} = \frac{2\pi}{\lambda^5}\frac{hc^2}{e^{hc/\lambda k_BT}-1} = \frac{c}{4}\rho_\lambda$$

$$\rho_f = \frac{8\pi}{c^3}k_BTf^2 \quad \text{(Rayleigh-Jeans)}$$

$$u = \frac{U}{V} = \int_{f=0}^{\infty}\rho_f\,df = \int_{\lambda=0}^{\infty}\rho_\lambda\,d\lambda$$

$$I = \frac{P}{A} = \int_{f=0}^{\infty}R_f\,df = \int_{\lambda=0}^{\infty}R_\lambda\,d\lambda$$

$$N = \frac{8}{3}\pi\frac{V_c}{c^3}f^3 \quad , \quad dN = \frac{8\pi V_c}{c^3}f^2df \quad , \quad \bar{\mathcal{E}} = \frac{hf}{e^{hf/k_BT}-1}$$

$$c = \lambda f \quad , \quad \mathcal{E}_n = nhf \quad , \quad n = 0,1,2,3,\cdots \quad , \quad \frac{e_1}{a_1} = \frac{e_2}{a_2} = 1$$

where $h = 6.626 \times 10^{-34}$ J·s, $\sigma = 5.67 \times 10^{-8}\,\frac{\text{W}}{\text{m}^2\text{K}^4}$, and $k_B = 1.38 \times 10^{-23}$ J/K. The quantities ρ_f and ρ_λ are energy density per frequency or per wavelength, R_f and R_λ are the corresponding spectral radiances, u is energy density, and I is intensity.

- **Wien's displacement law** ($\lambda_m T = \text{const.}$) is useful for problems involving temperature and the wavelength that maximizes the spectral radiance (or the energy density per wavelength). For qualitative questions involving color, note that the acronym ROY G. BIV (red, orange, yellow, green, blue, indigo, violet) orders the colors of visible light from long wavelength to short wavelength. For quantitative problems, it may help to express Wien's displacement law in the form $\lambda_{1m} T_1 = \lambda_{2m} T_2$ if there are two objects, or as $\lambda_m T \approx 0.00290$ K·m if there is a single object (where K and m are SI units).
- **Stefan's law** is useful for problems involving power (or intensity) and temperature. Note that $P_e = \sigma\epsilon A T^4$ is the instantaneous rate at which thermal energy is emitted by an object at temperature T, whereas $P_a = \sigma\epsilon A T_{env}^4$ is the instantaneous rate at which thermal energy is absorbed by the object's surroundings at temperature T_{env}, where ϵ is the emissivity and A is the surface area. The net power radiated is $P_{net} = P_e - P_a$. At radiative equilibrium, $P_e = P_a$ (sometimes written as $e = a$). Note that the surface area of a sphere is $4\pi R^2$ and a cylinder is $2\pi RL + 2\pi R^2$ (the body and two ends). For a blackbody, $\epsilon = 1$ such that $P = \sigma A T^4$ and $I = \frac{P}{A} = \sigma T^4$. If there are two objects that can only exchange energy via thermal radiation, at equilibrium $\frac{e_1}{a_1} = \frac{e_2}{a_2} = 1$.

- If a problem asks you to derive an equation, review the derivations from this chapter:
  - Wien's displacement law is derived on pages 100-101.
  - Stefan's law is derived on pages 99-100.
  - The relation $R_f = \frac{c}{4}\rho_f$ was derived on pages 96-97.
  - Planck's solution was derived on pages 94-95.
  - The Rayleigh-Jeans formula was derived on pages 87-92.

Example: One heated metal rod glows orange while another heated metal rod glows yellow. The rods are otherwise identical. Which rod is hotter?

Apply Wien's displacement law: $\lambda_m T = \text{const.}$, which may be expressed as $\lambda_{om} T_o = \lambda_{ym} T_y$ (using o for orange and y for yellow). Rewrite this as $\frac{\lambda_{om}}{\lambda_{ym}} = \frac{T_y}{T_o}$. The acronym ROY G. BIV orders the colors of visible light from longer wavelength to shorter wavelength. Therefore, orange (O) light has longer wavelength than yellow (Y) light: $\lambda_{om} > \lambda_{ym}$, which means that $\frac{\lambda_{om}}{\lambda_{ym}} > 1$. It follows that $\frac{T_y}{T_o} > 1$, which means that $T_y > T_o$. The $\boxed{\text{yellow}}$ rod is hotter than the orange rod. (The equation $\lambda_m T = \text{const.}$ shows that λ_m and T are inversely related. In order for the product $\lambda_m T$ to be constant, a shorter λ_m must correspond to a higher T and a longer λ_m must correspond to a lower T. Since orange light has longer wavelength than yellow light, it follows that the yellow rod is hotter.)

Example: The temperature of a blackbody is increased from 1000 K to 2000 K.

(A) By what factor does λ_m change?

Apply Wien's displacement law: $\lambda_m T = \text{const.}$, which may be expressed as $\lambda_{im} T_i = \lambda_{fm} T_f$ (using i for initial and f for final). Divide both sides by T_f.

$$\lambda_{fm} = \frac{T_i}{T_f}\lambda_{im} = \frac{1000}{2000}\lambda_{im} = \boxed{\frac{\lambda_{im}}{2}}$$

The final λ_m is one-half of its initial value. The value of λ_m is reduced by a factor of two.

(B) By what factor does the intensity of radiation change?

Apply Stefan's law: For a blackbody, $I = \sigma T^4$. Initially $I_i = \sigma T_i^4$ and finally $I_f = \sigma T_f^4$. Divide these two equations.

$$\frac{I_f}{I_i} = \frac{\sigma T_f^4}{\sigma T_i^4} = \left(\frac{T_f}{T_i}\right)^4 = \left(\frac{2000}{1000}\right)^4 = 2^4 = \boxed{16}$$

$$I_f = 16 I_i$$

The final intensity is 16 times the initial intensity of radiation.

Example: Earth's radius is 6.378×10^6 m, the sun's radius is 6.957×10^8 m, earth's average orbital radius is 1.496×10^{11} m, and the intensity of sunlight reaching the earth is $1367 \frac{\text{W}}{\text{m}^2}$.

(A) What is the total power radiated from the surface of the sun?

Intensity equals power per unit area, $I = \frac{P}{A}$, so power equals $P = IA$. Note that the intensity and area must correspond to the same distance from the center of the sun. Since the problem gives us the intensity at the earth, we need to find the surface area of a sphere that extends to earth's orbital radius. The surface area of a sphere is $A = 4\pi r^2$.

Since this problem has multiple parts, we will use the subscripts *es* for quantities that apply specifically to the earth-sun distance.

$$P = I_{es}A_{es} = I_{es}(4\pi r_{es}^2) = 4\pi r_{es}^2 I_{es} = 4\pi(1.496 \times 10^{11})^2(1367) = \boxed{3.845 \times 10^{26} \text{ W}}$$

(B) What is the intensity of sunlight at the surface of the sun?

We will use the same equation again, but this time with the radius of the sun. We will use the subscripts ss for the surface of the sun.

$$I_{ss} = \frac{P}{A_{ss}} = \frac{P}{4\pi r_{ss}^2} = \frac{3.845 \times 10^{26}}{4\pi(6.957 \times 10^8)^2} = \boxed{6.322 \times 10^7 \ \frac{\text{W}}{\text{m}^2}}$$

(C) What fraction of the sun's radiation is received by the earth?

From the sun's perspective, the earth appears to be a circular disc with area $A_d = \pi r_e^2$. Just a fraction of the sun's light shines towards this disc (where r_e is earth's radius). This area represents the sunlight that reaches the earth. The total amount of sunlight reaching earth's orbital radius is much greater: $A_{es} = 4\pi r_{es}^2$. Divide these two areas to find the fraction of the sun's radiation that is received by the earth.

$$\frac{A_d}{A_{es}} = \frac{\pi r_e^2}{4\pi r_{es}^2} = \frac{1}{4}\left(\frac{r_e}{r_{es}}\right)^2 = \frac{1}{4}\left(\frac{6.378 \times 10^6}{1.496 \times 10^{11}}\right)^2 = \boxed{4.5 \times 10^{-10}}$$

(D) What is the surface temperature of the sun?

Apply Stefan's law with $\sigma = 5.67 \times 10^{-8} \frac{\text{W}}{\text{m}^2\text{K}^4}$, treating the sun as a blackbody.

$$I_{ss} = \sigma T_{ss}^4$$

$$T_{ss} = \left(\frac{I_{ss}}{\sigma}\right)^{1/4} = \left(\frac{6.322 \times 10^7}{5.67 \times 10^{-8}}\right)^{1/4} = \boxed{5779 \text{ K}}$$

(E) For which wavelength is the sun's spectral radiance a maximum?

Apply Wien's displacement law: $\lambda_m T \approx 0.00290$ K·m.

$$\lambda_m \approx \frac{0.00290}{5779} = \boxed{5.02 \times 10^{-7} \text{ m}} = \boxed{502 \text{ nm}} = 5020 \text{ Å}$$

Example: Examine the graph depicting the ultraviolet catastrophe. Although the Rayleigh-Jeans prediction disagrees with experiment for ultraviolet and higher frequencies, the Rayleigh-Jeans prediction does agree with experiment for very low frequencies.

(A) Show mathematically that the Rayleigh-Jeans formula agrees with Planck's formula at low frequencies.

Compare the Rayleigh-Jeans formula for blackbody radiation to Planck's correction.

$$\rho_f^{\text{R-J}} = \frac{8\pi}{c^3} k_B T f^2 \quad , \quad \rho_f^{\text{Planck}} = \frac{8\pi}{c^3} \frac{h}{e^{hf/k_B T} - 1} f^3$$

What we need to do is show that these expressions agree (to good approximation) for low frequencies. Planck's formula involves the ratio $\frac{hf}{k_B T}$ (in the exponent in the denominator). For low frequencies, $hf \ll k_B T$ such that $\frac{hf}{k_B T} \ll 1$. It's convenient to define $x = \frac{hf}{k_B T}$. Recall from calculus that the Taylor series expansion of e^x about $x = 0$ is:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

For low frequencies, $x = \frac{hf}{k_B T} \ll 1$. For very small values of x, the first two terms provide a first approximation. Substitute $e^{hf/k_B T} \approx 1 + \frac{hf}{k_B T}$ into Planck's formula.

$$\rho_f^{\text{Planck}}(\text{for small } f) = \frac{8\pi}{c^3} \frac{h}{e^{hf/k_B T} - 1} f^3 \approx \frac{8\pi}{c^3} \frac{h}{1 + \frac{hf}{k_B T} - 1} f^3 = \frac{8\pi}{c^3} \frac{h}{\frac{hf}{k_B T}} f^3$$

To divide by a fraction, multiply by its reciprocal.

$$\rho_f^{\text{Planck}}(\text{for small } f) \approx \frac{8\pi}{c^3} (hf^3) \frac{k_B T}{hf} = \boxed{\frac{8\pi}{c^3} k_B T f^2} = \rho_f^{\text{R-J}}(\text{for small } f)$$

(B) Show that at high frequencies, Planck's formula exhibits the desired behavior whereas the Rayleigh-Jeans formula does not.

To investigate high frequencies, take the limit that f approaches infinity.

$$\rho_f^{\text{Planck}}(\text{as } f \to \infty) = \lim_{f\to\infty} \frac{8\pi}{c^3} \frac{h}{e^{hf/k_BT} - 1} f^3 = \frac{8\pi h}{c^3} \lim_{f\to\infty} \frac{f^3}{e^{hf/k_BT} - 1}$$

We need to apply l'Hôpital's rule (from calculus) three times.

$$\rho_f^{\text{Planck}}(\text{as } f \to \infty) = \frac{8\pi h}{c^3} \frac{\lim\limits_{f\to\infty} \frac{d^3}{df^3} f^3}{\lim\limits_{f\to\infty} \frac{d^3}{df^3}\left(e^{\frac{hf}{k_BT}} - 1\right)} = \frac{8\pi h}{c^3} \frac{6}{\frac{h^3}{k_B^3 T^3} \lim\limits_{f\to\infty} e^{\frac{hf}{k_BT}}} = \boxed{0}$$

As $f \to \infty$, in Planck's formula ρ_f approaches zero, whereas in the Rayleigh-Jeans formula ρ_f clearly increases to $\boxed{\infty}$. As shown on page 92, only Planck's formula exhibits the expected behavior at high frequencies.

Chapter 6 Problems

1. For each of the following cases, indicate which object is hotter.

(A) Compare the surface temperature of a red giant to a yellow main sequence star.

(B) Compare a blackbody with a wavelength of maximum spectral radiance equal to 524 nm to a blackbody with a wavelength of maximum spectral radiance equal to 576 nm.

(C) Compare a blackbody with a wavelength of maximum spectral radiance in the infrared region to a blackbody with a wavelength of maximum spectral radiance that is ultraviolet.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) $T_y > T_r$

(B) $T_{524} > T_{576}$

(C) $T_{UV} > T_{IR}$

2. The wavelength of maximum spectral radiance is 4800 Angstroms for one star and 6400 Angstroms for a second star.

(A) Find the ratio of the surface temperature of the second star to the first star.

(B) Find the corresponding ratio of the intensity of radiation at the surface of each star.

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) 3/4 = 0.75

(B) 81/256 ~ 0.32

3. Sirius A is 8.611 lightyears from earth. Starlight from Sirius A has an intensity of 1.166×10^{-7} W/m^2 when it reaches earth. Sirius A's spectral radiance is a maximum for a wavelength of 301.0 nm.

(A) What is the surface temperature of Sirius A?

(B) What is the total power radiated from the surface of Sirius A?

(C) What is the intensity at the surface of Sirius A?

(D) What is the radius of Sirius A?

Want help? Check the solution at the end of the chapter.

Answers: 3. (A) 9635 K

(B) 9.711×10^{27} W

(C) 4.89×10^{8} W/m^2

(D) 1.26×10^{9} m

4. The Rayleigh-Jeans prediction for blackbody radiation according to classical physics is:

$$\rho_f = \frac{8\pi}{c^3} k_B T f^2$$

(A) Apply the equation $\rho_f df = -\rho_\lambda d\lambda$ to derive an equation for ρ_λ for classical physics.

(B) Derive equations for R_f and R_λ according to classical physics.

(C) Show that the Rayleigh-Jeans formula disagrees with Wien's displacement law.

(D) Show that the Rayleigh-Jeans formula disagrees with Stefan's law.

(E) Sketch a graph showing ρ_λ from Part A compared with ρ_λ using Planck's formula.

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) $\rho_\lambda = \frac{8\pi}{\lambda^4} k_B T$ (B) $R_f = \frac{2\pi}{c^2} k_B T f^2$, $R_\lambda = \frac{2\pi}{\lambda^4} c k_B T$
(C) $\lambda_m = \infty$ (D) $I = \infty$

5. On pages 100-101, we started with Planck's formula for R_λ to derive Wien's law in terms of λ_m.

(A) Start with Planck's formula for R_f to derive Wien's law in terms of f_m instead of λ_m.

(B) Why is it incorrect to simply replace λ_m with $\frac{c}{f_m}$ in the equation $\lambda_m T \approx 0.00290$ K·m in order to derive Wien's law in terms of f_m?

Want help? Check the solution at the end of the chapter.

Answers: 5. (A) $\frac{f_m}{T} = 5.88 \times 10^{10} \frac{\text{Hz}}{\text{K}}$
(B) the shape of the spectral radiance depends on the parametrization

Solutions to Chapter 6

1. Apply Wien's displacement law: $\lambda_m T = \text{const}$. Write this as $\lambda_{1m} T_1 = \lambda_{2m} T_2$ or $\frac{\lambda_{1m}}{\lambda_{2m}} = \frac{T_2}{T_1}$.

(A) Red light has longer wavelength than yellow light: $\lambda_{rm} > \lambda_{ym}$ such that $\frac{\lambda_{rm}}{\lambda_{ym}} > 1$. Since $\frac{\lambda_{rm}}{\lambda_{ym}} = \frac{T_y}{T_r}$, it follows that $\frac{T_y}{T_r} > 1$ or $\boxed{T_y > T_r}$. The yellow main sequence star is hotter.

(B) $\lambda_{576m} = 576$ nm is greater than $\lambda_{524m} = 524$ nm: $\lambda_{576m} > \lambda_{524m}$ such that $\frac{\lambda_{576m}}{\lambda_{524m}} > 1$. Since $\frac{\lambda_{576m}}{\lambda_{524m}} = \frac{T_{524}}{T_{576}}$, it follows that $\frac{T_{524}}{T_{576}} > 1$ or $\boxed{T_{524} > T_{576}}$.

(C) Infrared light has longer wavelength than visible red light ($\lambda_{IRm} > \lambda_{rm}$), red light has longer wavelength than violet light ($\lambda_{rm} > \lambda_{vm}$), and violet light has longer wavelength than ultraviolet light ($\lambda_{vm} > \lambda_{UVm}$), according to the chart on page 84. It follows that $\lambda_{IRm} > \lambda_{UVm}$, such that $\frac{\lambda_{IRm}}{\lambda_{UVm}} > 1$. Since $\frac{\lambda_{IRm}}{\lambda_{UVm}} = \frac{T_{UV}}{T_{IR}}$, it follows that $\frac{T_{UV}}{T_{IR}} > 1$ or $\boxed{T_{UV} > T_{IR}}$.

2. Treat the stars as approximate blackbodies.

(A) Apply Wien's displacement law: $\lambda_m T = \text{const.}$ Write this as $\lambda_{4800m}T_{4800} = \lambda_{6400m}T_{6400}$.

$$\frac{T_{6400}}{T_{4800}} = \frac{\lambda_{4800m}}{\lambda_{6400m}} = \frac{4800}{6400} = \frac{4800 \div 1600}{6400 \div 1600} = \boxed{\frac{3}{4}} = \boxed{0.75}$$

The star with λ_m equal to 6400 Å is 3/4 as hot as the star with λ_m equal to 4800 Å.

(B) Apply Stefan's law: For a blackbodies, $I_{4800} = \sigma T_{4800}^4$ and $I_{6400} = \sigma T_{6400}^4$.

$$\frac{I_{6400}}{I_{4800}} = \frac{\sigma T_{6400}^4}{\sigma T_{4800}^4} = \left(\frac{T_{6400}}{T_{4800}}\right)^4 = \left(\frac{3}{4}\right)^4 = \frac{3^4}{4^4} = \boxed{\frac{81}{256}} \approx \boxed{0.32}$$

The star with λ_m equal to 6400 Å has 81/256 the surface intesity as the star with λ_m equal to 4800 Å.

3. First identify the information given in the problem:

- $r_{es} = 8.611$ ly is the distance between earth and Sirius A in lightyears.
- $I_{es} = 1.166 \times 10^{-7} \frac{\text{W}}{\text{m}^2}$ is the intensity of Sirius A's starlight when it reaches earth.
- $\lambda_m = 301 \text{ nm} = 301 \times 10^{-9}$ m since the prefix nano (n) stands for 10^{-9}.

(A) Apply Wien's displacement law: $\lambda_m T \approx 0.00290$ K·m.

$$T \approx \frac{0.00290}{\lambda_m} = \frac{0.00290}{301 \times 10^{-9}} = \boxed{9635 \text{ K}}$$

(B) Convert $r_{es} = 8.611$ ly from lightyears to meters, given that one lightyear is the distance that light travels in one year. The speed of light in vacuum is $c = 2.9979 \times 10^8$ m/s. To find the number of seconds in one year, multiply by 365 days/year, 24 hours/day, and 3600 s/hr.

$$t_{yr} = 1 \text{ year} \times \frac{365 \text{ days}}{1 \text{ year}} \times \frac{24 \text{ hours}}{1 \text{ day}} \times \frac{3600 \text{ s}}{1 \text{ hr}} = 3.1536 \times 10^7 \text{ s}$$

$$1\,ly = c\,t_{y4} = (2.9979 \times 10^8)(3.1536 \times 10^7) = 9.4542 \times 10^{15} \text{ m}$$

$$r_{es} = 8.611 \text{ ly} = (8.611)(9.4542 \times 10^{15}) = 8.141 \times 10^{16} \text{ m}$$

Intensity equals power per unit area, $I = \frac{P}{A}$, so power equals $P = IA$. Note that the intensity and area must correspond to the same distance from the center of Sirius A. Since the problem gives us the intensity at the earth, we need to find the surface area of a sphere that extends to earth's orbital radius. The surface area of a sphere is $A = 4\pi r^2$. We use the subscripts es for quantities that apply specifically to the distance between earth and Sirius A.

$$P = I_{es}A_{es} = I_{es}(4\pi r_{es}^2) = 4\pi r_{es}^2 I_{es} = 4\pi(8.141 \times 10^{16})^2(1.166 \times 10^{-7}) = \boxed{9.711 \times 10^{27} \text{ W}}$$

(C) Apply Stefan's law with $\sigma = 5.67 \times 10^{-8} \frac{\text{W}}{\text{m}^2\text{K}^4}$, treating Sirius A as a blackbody. We use the subscripts ss for the surface of Sirius A. Recall from Part A that $T \approx 9635$ K.

$$I_{ss} = \sigma T_{ss}^4 \approx (5.67 \times 10^{-8})(9635)^4 = \boxed{4.89 \times 10^8 \frac{\text{W}}{\text{m}^2}}$$

(D) Use the same equation from Part B, but this time with the intensity at the surface of Sirius A. Recall from Parts B and C that $P = 9.711 \times 10^{27}$ W and $I_{ss} = 4.89 \times 10^{8} \frac{\text{W}}{\text{m}^2}$.

$$P = I_{ss} A_{ss}$$

$$A_{ss} = \frac{P}{I_{ss}}$$

Recall that the surface area of a sphere is $A_{ss} = 4\pi r_{ss}^2$.

$$4\pi r_{ss}^2 = \frac{P}{I_{ss}}$$

$$r_{ss}^2 = \frac{P}{4\pi I_{ss}}$$

$$r_{ss} = \sqrt{\frac{P}{4\pi I_{ss}}} = \sqrt{\frac{9.711 \times 10^{27}}{4\pi(4.89 \times 10^{8})}} = \boxed{1.26 \times 10^{9} \text{ m}}$$

4. (A) Since ρ_f and ρ_λ don't even have the same units (see page 102), it would be **incorrect** if the **only** thing you did was substitute $f = \frac{c}{\lambda}$. Instead, you must apply the technique shown on page 97. Begin with $\rho_f df = -\rho_\lambda d\lambda$. Since $f = \frac{c}{\lambda}$, it follows that $\frac{df}{d\lambda} = \frac{d}{d\lambda}\left(\frac{c}{\lambda}\right) = -\frac{c}{\lambda^2}$.

$$\rho_\lambda = -\rho_f \frac{df}{d\lambda} = -\rho_f\left(-\frac{c}{\lambda^2}\right) = \frac{c}{\lambda^2}\rho_f$$

Plug in the equation given in the problem. Make the substitution $f = \frac{c}{\lambda}$.

$$\rho_\lambda = \frac{c}{\lambda^2}\rho_f = \frac{c}{\lambda^2}\left(\frac{8\pi}{c^3}k_B T f^2\right) = \frac{c}{\lambda^2}\left[\frac{8\pi}{c^3}k_B T\left(\frac{c}{\lambda}\right)^2\right] = \frac{c}{\lambda^2}\left(\frac{8\pi}{c^3}k_B T\frac{c^2}{\lambda^2}\right) = \boxed{\frac{8\pi}{\lambda^4}k_B T}$$

(B) Apply the equations $R_f = \frac{c}{4}\rho_f$ and $R_\lambda = \frac{c}{4}\rho_\lambda$ (see pages 96 and 107).

$$R_f = \frac{c}{4}\rho_f = \frac{c}{4}\left(\frac{8\pi}{c^3}k_B T f^2\right) = \boxed{\frac{2\pi}{c^2}k_B T f^2} \quad , \quad R_\lambda = \frac{c}{4}\rho_\lambda = \frac{c}{4}\left(\frac{8\pi}{\lambda^4}k_B T\right) = \boxed{\frac{2\pi}{\lambda^4}c k_B T}$$

(C) The problem is asking you to follow the strategy that we applied on page 100 to derive Wien's law, but to begin with the Rayleigh-Jeans formula instead of Planck's formula. Begin with the Rayleigh-Jeans equation for R_λ from Part B.

$$R_\lambda = \frac{2\pi}{\lambda^4} c k_B T$$

To find the value of λ that maximizes R_λ, take a derivative of R_λ with respect to λ.

$$\frac{dR_\lambda}{d\lambda} = \frac{d}{d\lambda}\left(\frac{2\pi}{\lambda^4} c k_B T\right) = 2\pi c k_B T \frac{d}{d\lambda}\left(\frac{1}{\lambda^4}\right) = 2\pi c k_B T \frac{d}{d\lambda}(\lambda^{-4}) = 2\pi c k_B T(-4\lambda^{-5}) = -\frac{8\pi c k_B T}{\lambda^5}$$

To find the value of λ that maximizes R_λ, set $\frac{dR_\lambda}{d\lambda}$ equal to zero.

$$\frac{dR_\lambda}{d\lambda} = -\frac{8\pi c k_B T}{\lambda^5} = 0$$

What value of λ would make the above equation equal to zero? The answer is $\boxed{\lambda_m = \infty}$ since $\lim_{\lambda \to \infty} \frac{1}{\lambda^5} = 0$. According to the Rayleigh-Jeans formula, λ_m is infinite. This disagrees with Wien's law, $\lambda_m T \approx 0.00290$ K·m.

(D) The problem is asking you to follow the strategy that we applied on page 99 to derive Stefan's law, but to begin with the Rayleigh-Jeans formula instead of Planck's formula. Begin with the Rayleigh-Jeans equation for R_f from Part B.

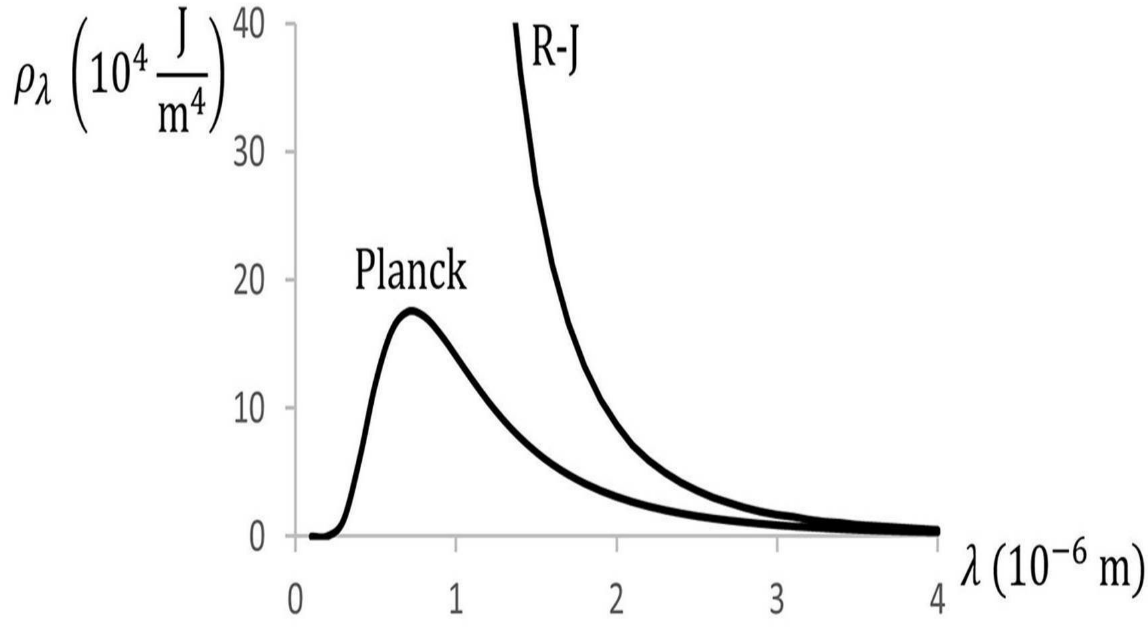
$$R_f = \frac{2\pi}{c^2} k_B T f^2$$

Integrate R_f over f to find the intensity.

$$I = \int_{f=0}^{\infty} R_f \, df = \int_{f=0}^{\infty} \frac{2\pi}{c^2} k_B T f^2 \, df = \frac{2\pi}{c^2} k_B T \int_{f=0}^{\infty} f^2 \, df = \frac{2\pi}{c^2} k_B T \left[\frac{f^3}{3}\right]_{f=0}^{\infty} = \boxed{\infty}$$

According to the Rayleigh-Jeans formula, the intensity is infinite. This disagrees with Stefan's law, $I = \sigma T^4$.

(E) Plot $\frac{8\pi}{\lambda^4} k_B T$ (Rayleigh-Jeans) and $\frac{8\pi}{\lambda^5} \frac{hc}{e^{hc/\lambda k_B T} - 1}$ (Planck). Our graph is for $T = 4000$ K.



5. Planck's formula for R_f is:

$$R_f = \frac{2\pi}{c^2}\frac{h}{e^{hf/k_BT}-1}f^3$$

(A) Follow the strategy for deriving Wien's law on pages 100-101. Take a derivative of R_f with respect to f.

$$\frac{dR_f}{df} = \frac{d}{df}\left(\frac{2\pi}{c^2}\frac{h}{e^{hf/k_BT}-1}f^3\right) = \frac{2\pi h}{c^2}\frac{d}{df}\left(\frac{f^3}{e^{hf/k_BT}-1}\right) = \frac{2\pi h}{c^2}\frac{d}{df}\left[f^3\left(e^{hf/k_BT}-1\right)^{-1}\right]$$

Apply the product rule: $\frac{d}{df}(pq) = q\frac{dp}{df} + p\frac{dq}{df}$ with $p = f^3$ and $q = \left(e^{hf/k_BT}-1\right)^{-1}$.

$$\frac{dR_f}{df} = \frac{2\pi h}{c^2}(f^3)\frac{d}{df}\left(e^{hf/k_BT}-1\right)^{-1} + \frac{2\pi h}{c^2}\left(e^{hf/k_BT}-1\right)^{-1}\frac{d}{df}(f^3)$$

$$= \frac{2\pi h f^3}{c^2}\frac{d}{df}\left(e^{hf/k_BT}-1\right)^{-1} + \frac{2\pi h}{c^2}\left(e^{hf/k_BT}-1\right)^{-1}(3f^2)$$

$$= \frac{2\pi h f^3}{c^2}\frac{d}{df}\left(e^{hf/k_BT}-1\right)^{-1} + \frac{6\pi h f^2}{c^2}\left(e^{hf/k_BT}-1\right)^{-1}$$

Apply the chain rule with $u = e^{hf/k_BT} - 1$ and $g = u^{-1}$ to write $\frac{d}{df}\left(e^{hf/k_BT} - 1\right)^{-1} = \frac{dg}{du}\frac{du}{df}$.

Note that $\frac{dg}{du} = \frac{d}{du}(u^{-1}) = -u^{-2}$ and $\frac{du}{df} = \frac{d}{df}\left(e^{hf/k_BT} - 1\right) = \frac{h}{k_BT}e^{hf/k_BT}$.

$$\frac{dR_f}{df} = -\frac{2\pi hf^3}{c^2}\left(e^{hf/k_BT} - 1\right)^{-2}\left(\frac{h}{k_BT}e^{hf/k_BT}\right) + \frac{6\pi hf^2}{c^2}\left(e^{hf/k_BT} - 1\right)^{-1}$$

$$= -\frac{2\pi h^2f^3}{c^2k_BT}\left(e^{hf/k_BT} - 1\right)^{-2}\left(e^{hf/k_BT}\right) + \frac{6\pi hf^2}{c^2}\left(e^{hf/k_BT} - 1\right)^{-1}$$

To find the extreme values of R_f, set the first derivative equal to zero: $\frac{dR_f}{df} = 0$.

$$-\frac{2\pi h^2f^3}{c^2k_BT}\left(e^{hf/k_BT} - 1\right)^{-2}\left(e^{hf/k_BT}\right) + \frac{6\pi hf^2}{c^2}\left(e^{hf/k_BT} - 1\right)^{-1} = 0$$

Note that $\left(e^{hf/k_BT} - 1\right)^{-n} = \frac{1}{\left(e^{hf/k_BT}-1\right)^n}$.

$$-\frac{2\pi h^2f^3}{c^2k_BT}\frac{\left(e^{hf/k_BT}\right)}{(e^{hf/k_BT} - 1)^2} + \frac{6\pi hf^2}{c^2}\frac{1}{e^{hf/k_BT} - 1} = 0$$

Multiply both sides of the equation by c^2 and divide by $2\pi h$.

$$-\frac{hf^3}{k_BT}\frac{\left(e^{hf/k_BT}\right)}{(e^{hf/k_BT}-1)^2}+\frac{3f^2}{e^{hf/k_BT}-1}=0$$

Add $\frac{hf^3}{k_BT}\frac{\left(e^{hf/k_BT}\right)}{\left(e^{hf/k_BT}-1\right)^2}$ to both sides of the equation.

$$\frac{3f^2}{e^{hf/k_BT}-1}=\frac{hf^3}{k_BT}\frac{\left(e^{hf/k_BT}\right)}{(e^{hf/k_BT}-1)^2}$$

Divide both sides of the equation by f^2 and multiply both sides by $\left(e^{hf/k_BT}-1\right)^2$.

$$3\left(e^{hf/k_BT}-1\right)=\frac{hf}{k_BT}e^{hf/k_BT}$$

Distribute the 3.

$$3e^{hf/k_BT}-3=\frac{hf}{k_BT}e^{hf/k_BT}$$

Make the substitution $x=\frac{hf}{k_BT}$.

$$3e^x-3=xe^x$$

Multiply both sides by e^{-x}. Note that $e^xe^{-x}=1$.

$$3-3e^{-x}=x$$

This is a transcendental equation. Apply a numerical technique (such as Newton's method) to find the roots of the equation $3-3e^{-x}-x=0$. You should get $x\approx 2.82144$. Substitute $x\approx 2.82144$ into the equation $x=\frac{hf}{k_BT}$. We will replace f with f_m since f_m is the frequency that maximizes R_f.

$$2.82144 \approx \frac{hf_m}{k_B T} \quad \rightarrow \quad \frac{f_m}{T} \approx 2.82144 \frac{k_B}{h} \approx 2.82144 \frac{1.38 \times 10^{-23} \text{ J/K}}{6.626 \times 10^{-34} \text{ J·s}}$$

$$\frac{f_m}{T} \approx \boxed{\frac{5.88 \times 10^{10}}{\text{K·s}}} \approx \boxed{5.88 \times 10^{10} \frac{\text{Hz}}{\text{K}}}$$

Note that $\frac{\text{J}}{\text{K}} \div \text{J·s} = \frac{\text{J}}{\text{K}} \div \frac{\text{J·s}}{1} = \frac{\text{J}}{\text{K}} \times \frac{1}{\text{J·s}} = \frac{\text{J}}{\text{K·J·s}} = \frac{1}{\text{K·s}} = \frac{\text{Hz}}{K}$ since $1 \text{ Hz} = \frac{1}{\text{s}}$. (To divide by a fraction, multiply by its reciprocal. The reciprocal of J·s is $\frac{1}{\text{J·s}}$.)

(B) First, let's try replacing λ_m with $\frac{c}{f_m}$ in the equation $\lambda_m T \approx 0.00290$ K·m so that you can see that it doesn't work. You would get $\frac{c}{f_m} T \approx 0.00290$ K·m. Multiply both sides by $\frac{f_m}{T}$ and divide both sides by 0.00290 to get $\frac{c}{0.00290 \text{ K·m}} \approx \frac{f_m}{T}$. Plug in $c = 2.9979 \times 10^8$ m/s to get $\frac{f_m}{T} \approx \frac{2.9979 \times 10^8 \text{ m/s}}{0.00290 \text{ K·m}} \approx 1.03 \times 10^{11} \frac{\text{Hz}}{\text{K}}$, which is 75% larger than $\frac{f_m}{T} \approx 5.88 \times 10^{10} \frac{\text{Hz}}{\text{K}}$ from Part B (since $1.03 \times 10^{11} = 10.3 \times 10^{10}$ and since $\frac{10.3 - 5.88}{5.88} 100\% = 75\%$). The question is asking why we don't get the same answer as Part A when we replace λ_m with $\frac{c}{f_m}$ in the equation $\lambda_m T \approx 0.00290$ K·m.

Although $\lambda f = c$ for electromagnetic waves, in this case we're not looking at a single electromagnetic wave with wavelength λ_m or f_m, so λ_m and f_m don't correspond through the usual formula $\lambda f = c$. That is, $\lambda_m f_m \neq c$. The quantities λ_m and f_m are defined through two different parametrizations:

- λ_m is the value of λ that maximizes R_λ, which is the spectral radiance per unit λ.
- f_m is the value of f that maximizes R_f, which is the spectral radiance per unit f.
- Note that R_λ and R_f also have different units: $\frac{\text{W}}{\text{m}^3}$ and $\frac{\text{W}}{\text{m}^2\text{Hz}}$. They are different quantities.

Mathematically, the difference arises from the starting point, which is $R_f df = -R_\lambda d\lambda$. To go from R_f to R_λ, it isn't enough to write $f = \frac{c}{\lambda}$: You must also account for the derivative. Solve for R_λ to see that $R_\lambda = -R_f \frac{df}{d\lambda} = -R_f \frac{d}{d\lambda}\left(\frac{c}{\lambda}\right) = -R_f\left(-\frac{c}{\lambda^2}\right) = \frac{cR_f}{\lambda^2}$. What we're really saying is that the value of f that maximizes R_f isn't the same value of f that maximizes $R_\lambda = \frac{cR_f}{\lambda^2}$.

7 THE PHOTOELECTRIC EFFECT

Relevant Terminology

Photoelectron – an electron that is ejected when a photon strikes a metallic surface.

Photon – a particle of light.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Monochromatic – a single, well-defined wavelength (or frequency).

Work – work is done when there is not only a force acting on an object, but when the force also contributes toward the displacement of an object.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Potential energy – work that can be done by changing position. All forms of potential energy are stored energy.

Electric potential – electric potential energy per unit charge.

Potential difference – the difference in electric potential between two points; also called the voltage. It is the work per unit charge needed to move a test charge between two points.

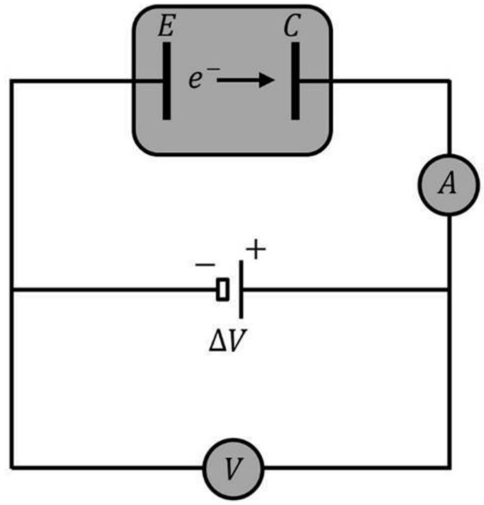
Stopping potential – a limiting electric potential difference for which no photoelectric effect is observed.

Current – the instantaneous rate of flow of charge across a conductor.

Work function – the minimum energy with which an electron is bound to the metal.

Quantum – a fixed elemental unit corresponding to the minimum possible value that can be measured for a quantity that comes in discrete bundles like energy or angular momentum.

Quantized – limited to integer multiples of a quantum unit. A quantity like energy or angular momentum that is quantized is discrete (rather than continuous).


E
C
e⁻
A
−
+
ΔV
V

The Photoelectric Effect

When light with sufficient frequency is incident upon a metallic surface, electrons (called **photoelectrons**) can be ejected from the surface. This is known as the photoelectric effect. One way to observe the photoelectric effect is with a circuit similar to the schematic diagram shown on the previous page, where:

- A DC power supply provides a potential difference

  $$(\Delta V)$$

  across an evacuated tube.

- The tube contains two metal plates: an emitter (E) and a collector (C).

- A voltmeter measures the potential difference

  $$(\Delta V)$$

  across the plates.

- An ammeter measures the current (I) through the power supply.

- Monochromatic light (having a single, well-defined wavelength) shines on plate E.

- When the **photoelectric effect** occurs, **photoelectrons** ejected from the emitter (plate E) flow across the gap in the tube to the collector (plate C).

- When the circuit is in complete darkness, the current (I) is zero.

- When photoelectrons jump across the gap from plate E to plate C, a nonzero current (I) results.

- The power supply potential difference

  $$(\Delta V)$$

  —sometimes called the voltage—may be adjusted. When

  $$\Delta V$$

  is positive, the negative terminal connects to plate E and the positive terminal connects to plate C. When

  $$\Delta V$$

  is negative, the polarity is reversed.

The following experimental observations have been established regarding the photoelectric effect:

- For a given metallic surface at plate E, a minimum **cutoff frequency** (f_c) is needed in order for the incident light shining at plate E to produce the photoelectric effect. When the incident light has a frequency less than f_c, no photoelectrons are ejected from plate E and the current (I) is zero. The value of f_c depends on the metal used at plate E.

• Even if $f > f_c$, there is a **stopping potential**, V_s, for which the current drops to zero. When

$$\Delta V \leq -V_s$$

(where the negative sign indicates that the polarity of the DC power supply is reversed compared to the previous schematic diagram), no photoelectrons are ejected from plate E.

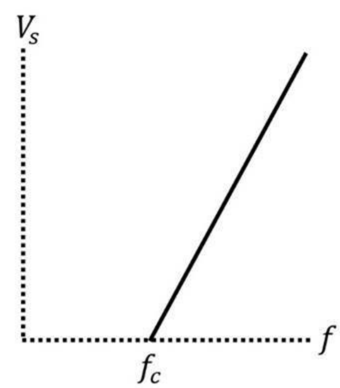
• When $f > f_c$ and

$$\Delta V > -V_s$$

the photoelectric effect is observed **immediately**: Photo-electrons travel from plate E and collect at plate C without any noticeable delay. The photoelectric effect doesn't take time to build up.

• Increasing the intensity (I) of light shining on plate E has **no** effect on the maximum kinetic energy (K_m) of the photoelectrons. However, increasing the intensity does increase the **number** of photoelectrons that are ejected from plate E.

• Increasing the **frequency** (f) of the light shining on plate E does increase the maximum kinetic energy (K_m) of the photoelectrons.



Classical Predictions for the Photoelectric Effect

Classical physics was unable to explain the following features of the photoelectric effect:

• Why is there a **cutoff frequency** (f_c)? According to classical physics, the photoelectric effect was expected to occur for all frequencies, provided that the incident light had sufficient intensity (I).

• Why isn't there a noticeable **time delay**? Classical physics expected that it would take time for the photoelectric effect to build up before photoelectrons would be ejected from the metallic surface.

• Why doesn't the **intensity** (I) of the incident light affect the maximum kinetic energy (K_m) of the photoelectrons? Since light is an electromagnetic wave, as the intensity of light increases, the oscillating electric field

$$(\vec{\mathbf{E}})$$

increases, and electric force equals charge times electric field

$$(\vec{\mathbf{F}}_e = -e\vec{\mathbf{E}})$$

where e is the charge of a proton (such that $-e$ is the charge of an electron). Yet this increased force didn't increase the maximum kinetic energy of the photoelectrons. Why not?

• Why does **frequency** (f) matter? Classical physics couldn't explain why increasing the frequency of the incident light increased the maximum kinetic energy (K_m) of the photoelectrons.

Quantum Explanation of the Photoelectric Effect

Albert Einstein formulated a quantum explanation for the photoelectric effect which also interprets Planck's solution to the blackbody radiation problem (Chapter 6). Recall Planck's formula from Chapter 6: $E_n = nhf$, where $h = 6.626 \times 10^{-34}$ J·s is **Planck's constant** and n is a positive integer. Planck's formula shows that electromagnetic waves inside a blackbody cavity have their energy **quantized**, meaning that they come in discrete bundles. A **quantum** of energy is the minimum value of the energy, which is hf. Why is the energy quantized? Why does the energy of electromagnetic waves have a minimum value? Einstein's solution to the photoelectric effect naturally answers these questions.

Einstein realized that electromagnetic waves (of all kinds, not just those found in blackbody cavities) are made up of particles which we call **photons**. One photon is a single particle of light with energy

$$E_\gamma = hf$$

The minimum possible energy in a beam of light is

$$E_\gamma = hf$$

since one is the least number of photons that you can have. A quantum of energy is the energy of one photon. The total energy in a beam of monochromatic light is $E = nhf$, where n is the total number photons in the beam. The reason that a beam of light can't have an energy of $4.5hf$, for example, is that you can't have half a photon.

The energy of a single photon is proportional to the frequency of the light:

$$E_\gamma = hf$$

Light with higher frequency effectively carries photons with more energy. For example, since blue light has higher frequency than red light, a single blue photon carries more energy than a single red photon.

Einstein showed that the concept of a photon—a particle of light with energy

$$E_\gamma = hf$$

—could explain the photoelectric effect. There are two key concepts: The energy of a single photon is proportional to the frequency of the light, and the photons in the beam of light interact with atoms in the metallic surface on a one-to-one basis (that is, a single photon from the beam interacts with a single atom).

• Why is there a **cutoff frequency** (f_c)? A single photon interacts with a single atom. If the energy of the photon,

$$E_\gamma = hf$$

exceeds the binding energy of an electron, the atom releases the electron (creating a photoelectron) when it absorbs the photon. Since a photon's energy is proportional to frequency, a minimum frequency ensures that the photons will have enough energy to free electrons from the metallic surface.

• Why isn't there a noticeable **time delay**? Since photons interact with atoms on a one-to-one basis, as soon as the first photons in the beam of light strike the metallic surface, photoelectrons are instantaneously ejected from the plate.

• Why doesn't the **intensity** (I) of the incident light affect the maximum kinetic energy (K_m) of the photoelectrons? Higher intensity means that the beam of light contains more photons. However, the energy of each photoelectron depends on the energy of each photon, which depends on frequency, not intensity, according to

$$E_\gamma = hf$$

• Why does **frequency** (f) matter? Increasing the frequency of the incident light results in higher energy photons, since

$$E_\gamma = hf$$

When the photons interact with atoms in the metallic surface, the excess energy produces photoelectrons with additional kinetic energy. Thus, increasing the frequency of the light shining on plate E increases the maximum kinetic energy (K_m) of the photoelectrons.

Photoelectric Effect Equations

The energy of a single **photon** is:

$$E_\gamma = hf$$

From conservation of energy, the **kinetic energy** of a photoelectron equals the difference in energy between the photon's energy and the work needed to remove the electron from the metallic surface.

$$K = E_\gamma - W_e = hf - W_e$$

The work, W_e, is a different value for different electrons because some electrons are more tightly or loosely bound to their atoms. For a given metal, there exists a **work function**, W_0, which represents the energy of the electron that is most loosely bound to an atom. The work function is the minimum amount of work that must be done to remove an electron from the metallic surface. (The value of work corresponding to the work function won't remove *any* electrons, but there *are* electrons in the metal that *can* be removed by this value of work.) The kinetic energy of a photoelectron is **maximum** when W_e happens to equal W_0.

$$K_m = E_\gamma - W_0 = hf - W_0$$

For more tightly bound electrons, $W_e > W_0$ and $K < K_m$.

There is a **cutoff frequency** (f_c) because K_m will be zero if hf equals W_0. In order to eject an electron from the metallic surface, hf must exceed W_0. To determine the cutoff frequency, set K_m equal to zero: This means setting hf_c equal to W_0:

$$K_m = 0 \quad \rightarrow \quad hf_c = W_0$$

The work done moving an electron from plate E to place C equals $-e\Delta V$ when the polarity of the DC power supply is as drawn in the figure on page 115 (since the charge of an electron, $-e$, is negative). When ΔV equals $-V_s$ (corresponding to the **stopping potential**, which is negative because the polarity is reversed for the stopping potential, as mentioned on page 116), K_m is equal to eV_s (the two minus signs from $-e$ and $-V_s$ make a plus sign).

$$K_m = eV_s$$

Substitute the previous equation into the equation $K_m = hf - W_0$.

$$eV_s = hf - W_0$$

The **stopping** potential thus equals

$$V_s = \frac{hf}{e} - \frac{W_0}{e}$$

Note that h, f, and e are all constants. For a given metallic surface, W_0 is also constant. The only variables in the above equation are V_s and f. The structure of the previous equation is thus $y = mx + b$, where $y = V_s$, $x = f$, $m = \frac{h}{e}$, and $b = -\frac{W_0}{e}$. This is the equation for a straight line with a slope of $\frac{h}{e}$ and a y-intercept of $-\frac{W_0}{e}$. Such a graph is shown on page 117. If you set y equal to zero, you can see that the x-intercept (not to be confused with the y-intercept) is $x_{int} = -\frac{b}{m} = -\frac{-W_0/e}{h/e} = \frac{W_0}{h} = f_c$ (since earlier we found that $W_0 = hf_c$).

Symbols and SI Units

Symbol	Name	SI Units
f	frequency	Hz
f_c	cutoff frequency	Hz
λ	wavelength	m
c	speed of light in vacuum	m/s
h	Planck's constant	J·s or J/Hz
n	a positive integer	unitless
E_γ	energy of one photon	J
K	kinetic energy of a photoelectron	J
K_m	maximum kinetic energy of a photoelectron	J

W_e	the work needed to remove an electron from the metal	J
W_0	work function	J
$\vec{\mathbf{E}}$	electric field	N/C or V/m
$\vec{\mathbf{F}}_e$	electric force	N
$e, -e$	the charge of a proton / the charge of an electron	C
V_s	stopping potential	V
ΔV	potential difference	V
I	current or intensity (these are two different quantities with different units)	A or $\frac{\mathrm{W}}{\mathrm{m}^2}$
N	number of photons per square meter per second	$\frac{1}{\mathrm{m}^2\mathrm{s}}$

m	mass or slope of a straight line	kg, it depends
b , x_{int}	y- and x-intercepts	it depends
u_m	the maximum speed of a photoelectron	m/s

Note: The lowercase Greek letters λ and γ are lambda and gamma, respectively.

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^{8}$ m/s
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C

Metric Prefixes

Prefix	Value
kilo (k)	$k = 10^3$
mega (M)	$M = 10^6$
giga (G)	$G = 10^9$
tera (T)	$T = 10^{12}$

Electron Volts

When working with individual particles, the SI unit of energy—which is the Joule (J)—isn't very suitable (because the energy of an individual particle is typically very small compared to one Joule). In this context, it is common to work with **electron Volts** (eV) instead of Joules (J). Recall from electricity and magnetism that work (W) equals charge (q) times potential difference (ΔV): $W = q\Delta V$. Also recall that the SI unit of work (W) is the Joule (J), the SI unit of charge (q) is the Coulomb (C), and the SI unit of potential difference (ΔV) is the Volt (V). From the equation $W = q\Delta V$, it follows that a Joule equals a Coulomb times a Volt: $1\text{ J} = 1\text{ C}\cdot\text{V}$. When we apply the same equation, $W = q\Delta V$, to determine the work need to accelerate a proton (or electron, since their charges are equal and opposite) through a potential difference of one Volt (1 V), using the charge of a proton ($e = 1.6021766 \times 10^{-19}$ C; note that e stands for the charge of a proton and is positive, such that the charge of an electron is $-e$) we find that one electron Volt is related to a Joule by:

$$1\text{ eV} = 1.6021766 \times 10^{-19}\text{ J}$$

It's common to see a prefix added to electron Volts, such as keV or MeV.

Strategy for Problems Involving the Photoelectric Effect

If a problem involves the photoelectric effect, follow these steps:

- Identify the given information and the desired unknown(s). Here are a few tips:
  - Use f_c when the problem clearly gives you or asks for the **cutoff** frequency. Otherwise, use f without the subscript 'c.'
  - Use V_s when the problem clearly gives you or asks for the **stopping** potential ($V_s > 0$ and $\Delta V = -V_s$ in this case). Otherwise, use ΔV for potential difference.
  - The units can sometimes help. For example, wavelength (λ) may be expressed in m, cm, nm, or Å, whereas frequency (f) may be expressed in Hz or s^{-1}.
  - K, K_m, E_γ, W, and W_e may be expressed in Joules (J) or electron Volts (eV), but V and ΔV are expressed in Volts (V). Note that an electron Volt is a unit of work or energy, whereas a Volt is a unit of electric potential.
- Choose the relevant equation(s) based on what you know and what you're looking for.

$$K = hf - W_e \quad , \quad K_m = hf - W_0 \quad , \quad E_\gamma = hf$$

$$W_0 = hf_c \quad , \quad K_m = eV_s \quad , \quad eV_s = hf - W_0 \quad , \quad \lambda f = c$$

- Solve for the desired unknown(s).

- To find the maximum speed of a photoelectron, recall the formulas from Chapter 5.
- If a problem gives (or asks for) **electron Volts**, note that $1 \text{ eV} = 1.6021766 \times 10^{-19}$ J.
- There are three types of problems that relate to the equation $V_s = \frac{hf}{e} - \frac{W_0}{e}$, which is the equation for a straight line with a slope of $m = \frac{h}{e}$, a y-intercept of $b = -\frac{W_0}{e}$, and an x-intercept of $x_{int} = f_c$.
    - A problem could give you a graph similar to the one shown on page 117. If the graph includes numbers along the f- and V_s-axes, use the graph to determine the slope and intercepts.
    - A problem could give you an equation of the form $V_s = mf + b$ with numerical values in place of m and b. If so, compare the given equation with $V_s = mf + b$ in order to determine the slope and y-intercept.
    - A problem could give you two data points, (f_1, V_{1s}) and (f_2, V_{2s}) with numbers instead of symbols. If so, find the slope: $m = \frac{V_{2s} - V_{1s}}{f_2 - f_1}$. Use the equation $V_s = mf + b$ with either pair of given numbers to solve for b.

- To find the **number** of photons (N) in a beam of light per square meter per second (in units of $\frac{1}{\text{m}^2\text{s}}$), use the following formula (where I is the intensity in units of $\frac{\text{W}}{\text{m}^2}$):

$$I = NE_\gamma$$

- For qualitative questions regarding different colors of visible light or different types of electromagnetic waves, consult the chart on page 84 in Chapter 6. The acronym ROY G. BIV orders the wavelengths of visible light from longest to shortest: red, orange, yellow, green, blue, indigo, violet.

Example: Between a blue photon and a violet photon, which carries more energy?

The acronym ROY G. BIV orders the visible colors from longer wavelengths to shorter wavelengths. Thus, violet (V) light has shorter wavelength than blue (B) light. Since $f = \frac{c}{\lambda}$, light with shorter wavelength has higher frequency. Thus, violet light has higher frequency than blue light. Since the energy of a photon is $E_\gamma = hf$, photons with higher frequency carry more energy. Thus, a violet photon carries more energy than a blue photon: $\boxed{E_v > E_b}$.

Example: Light with a wavelength of 200 nm shines on a metal with a work function of 4.0 eV.

First identify the given information:

- The wavelength is $\lambda = 200$ nm. Note that 1 nm = 10^{-9} m.
- The work function is $W_0 = 4.0$ eV.

(A) What is the energy of each incident photon in electron Volts?

Find the frequency from the wavelength and the speed of light ($c = 2.9979 \times 10^8$ m/s).

$$f = \frac{c}{\lambda} = \frac{2.9979 \times 10^8}{200 \times 10^{-9}} = 1.50 \times 10^{15} \text{ Hz}$$

The energy of a photon equals Planck's constant ($h = 6.626 \times 10^{-34}$ J·s) times the frequency.

$$E_\gamma = hf = (6.626 \times 10^{-34})(1.50 \times 10^{15}) = 9.94 \times 10^{-19} \text{ J}$$

Note that one electron Volt equals 1 eV = $1.6021766 \times 10^{-19}$ J. Divide by $1.6021766 \times 10^{-19}$ in order to find the conversion factor from Joules to electron Volts:

$$1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18} \text{ eV}$$

The energy of each photon in electron Volts is

$$E_\gamma = 9.94 \times 10^{-19} \times 6.2415092 \times 10^{18} = \boxed{6.20 \text{ eV}}$$

(B) What is the cutoff frequency for observing the photoelectric effect?

The work function equals Planck's constant times the cutoff frequency: $W_0 = hf_c$. Divide by Planck's constant.

$$f_c = \frac{W_0}{h}$$

We have a problem: We know the work function in electron Volts ($W_0 = 4.0$ eV) and Planck's constant in Joules times seconds ($h = 6.626 \times 10^{-34}$ J·s). The units of energy need to match in order to find the cutoff frequency in Hertz. Convert the work function to Joules using 1 eV = $1.6021766 \times 10^{-19}$ J.

$$W_0 = 4.0 \text{ eV} = 4.0 \times 1.6021766 \times 10^{-19} = 6.4 \times 10^{-19} \text{ J}$$

Now we can find the cutoff frequency in Hertz.

$$f_c = \frac{W_0}{h} = \frac{6.4 \times 10^{-19}}{6.626 \times 10^{-34}} = \boxed{9.7 \times 10^{14} \text{ Hz}}$$

(C) What is the stopping potential?

Use the equation $eV_s = hf - W_0$ with $e = 1.6021766 \times 10^{-19}$ C. Be sure to use the frequency of the incident light, $f = 1.50 \times 10^{15}$ Hz, and not the cutoff frequency (f_c). Also, be sure to express Planck's constant and the work function both in terms of Joules (or both in terms of electron Volts). Recall from Part B that $W_0 = 6.4 \times 10^{-19}$ J.

$$eV_s = hf - W_0$$

$$V_s = \frac{hf - W_0}{e} = \frac{(6.626 \times 10^{-34})(1.50 \times 10^{15}) - 6.4 \times 10^{-19}}{1.6021766 \times 10^{-19}} = \boxed{2.2 \text{ V}}$$

(D) What is the kinetic energy of the fastest photoelectrons in electron Volts?

The maximum kinetic energy of a photoelectron equals the charge of a proton (which is the same as the charge of an electron, except for the sign) times the stopping potential. To get this in electron Volts (eV), simply multiply the stopping potential (found in Part C) by one. (If you multiply by $e = 1.6021766 \times 10^{-19}$ C to get K_m in Joules, you would then divide by $1.6021766 \times 10^{-19}$ in order to convert from Joules to electron Volts. The $1.6021766 \times 10^{-19}$ simply cancels out.)

$$K_m = eV_s = (1 \text{ e})(2.2 \text{ V}) = \boxed{2.2 \text{ eV}}$$

Example: Suppose that you measure stopping potential as a function of frequency and plot the data.

(A) How could you determine the value of Planck's constant from the graph?

The graph will be a straight line given by the equation $V_s = \frac{hf}{e} - \frac{W_0}{e}$. Note that h, f, and e are all constants. Compare this equation to the general form $y = mx + b$ for a straight line to see that f is x, V_s is y, and the slope is $m = \frac{h}{e}$. Read off the (f, V_s) values for two points on the graph and determine the slope from the following equation.

$$m = \frac{V_{2s} - V_{1s}}{f_2 - f_1}$$

Since $m = \frac{h}{e}$, it follows that $h = me$. Multiply the numerical value of the slope (in Volts per Hz, or equivalently Volts times seconds) by $e = 1.6021766 \times 10^{-19}$ C.

(B) How could you determine the cutoff frequency from the graph?

The cutoff frequency (f_c) is the frequency for which the stopping potential equals zero. Setting V_s equal to zero in the equation $V_s = \frac{hf}{e} - \frac{W_0}{e}$ corresponds to finding the f-intercept of the graph. Therefore, the f-intercept of the graph equals the cutoff frequency.

Example: A particular beam of monochromatic light with a wavelength of 400 nm has one billion photons per square meter per second. What is the intensity of the beam of light?

Combine the formulas $I = NE_\gamma$, $E_\gamma = hf$, and $f = \frac{c}{\lambda}$. Note that 1 billion = 1,000,000,000 = 10^9.

$$I = NE_\gamma = N(hf) = Nh\left(\frac{c}{\lambda}\right) = \frac{hcN}{\lambda} = \frac{(6.626 \times 10^{-34})(2.9979 \times 10^8)(10^9)}{400 \times 10^{-9}} = \boxed{5.0 \times 10^{-10}\,\frac{\text{W}}{\text{m}^2}}$$

Chapter 7 Problems

1. For each question below, indicate which photon carries more energy.

(A) Compare an orange photon to a yellow photon.

(B) Compare a 400-nm photon to a 500-nm photon.

(C) Compare a 5.0×10^{14}-Hz photon to a 6.0×10^{14}-Hz photon.

(D) Compare an infrared photon to an ultraviolet photon.

(E) Compare an ultraviolet photon to a photon in a gamma ray.

(F) Compare a photon in an x-ray to a photon in a microwave.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) yellow

(B) 400 nm

(C) 6.0×10^{14} Hz

(D) UV

(E) gamma ray

(F) x-ray

2. Consider a variety of metal plates made from the following common metals: copper (Cu), aluminum (Al), iron (Fe), silver (Ag), and lead (Pb). Each of these metals has a work function in the range 4.0 eV < W_0 < 5.0 eV. **Note**: These values only apply to Part A below.

(A) If you shine visible light on any of these metals, would photoelectrons result? Explain.

(B) For wavelengths of visible light

$$380 \text{ nm} < \lambda < 750 \text{ nm}$$

what numerical values of work functions in electron Volts would be needed to observe the photoelectric effect?

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) no; f_c is ultraviolet

(B) W_0 < 1.65 eV to W_0 < 3.26 eV

3. A He-Ne laser produces red light with a wavelength of 633 nm and an intensity of 1.5 mW. How many photons per square meter per second are in the laser beam?

Want help? Check the solution at the end of the chapter.

Answer: (3) $4.78 \times 10^{15}/m^2/s$

4. Light with a frequency of 900 THz shines on a metal with a work function of 3.00 eV. The mass of an electron is 9.11×10^{-31} kg.

(A) What is the energy of each incident photon in electron Volts?

(B) What is the cutoff frequency for observing the photoelectric effect?

(C) What is the cutoff wavelength for observing the photoelectric effect?

(D) What is the kinetic energy of the fastest photoelectrons in electron Volts?

(E) What is the stopping potential?

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) 3.72 eV (B) 725 THz (C) 414 nm (D) 0.72 eV (E) 0.72 V

5. Light with a wavelength of 300 nm shines on a metal. The fastest photoelectrons produced are observed to have a kinetic energy of 2.00 eV.

(A) What is the stopping potential?

(B) What is the work function for the metal in electron Volts?

(C) What is the cutoff wavelength for observing the photoelectric effect?

Want help? Check the solution at the end of the chapter.

Answers: 5. (A) 2.00 V (B) 2.13 eV (C) 582 nm

6. For a given metal, the stopping potential is 1.75 V when the incident light has a wavelength of 250 nm.

(A) What is the work function for the metal?

(B) What is the stopping potential when the wavelength is reduced to 125 nm?

(C) What is the cutoff wavelength for observing the photoelectric effect?

(D) What is the maximum kinetic energy of the photoelectrons when the wavelength is 250 nm?

(E) What is the maximum kinetic energy of the photoelectrons when the wavelength is 125 nm?

Want help? Check the solution at the end of the chapter.

Answers: 6. (A) 3.21 eV (B) 6.71 V (C) 386 nm (D) 1.75 eV (E) 6.71 eV

Solutions to Chapter 7

1. The energy of a photon equals $E_\gamma = hf$, where $h = 6.626 \times 10^{-34}$ J·s is Planck's constant. Photons with higher frequency carry more energy.

(A) The acronym ROY G. BIV orders the visible colors from longer wavelengths to shorter wavelengths. Thus, yellow (Y) light has shorter wavelength than orange (O) light. Since $f = \frac{c}{\lambda}$, light with shorter wavelength has higher frequency. Thus, yellow light has higher frequency than orange light. According to $E_\gamma = hf$, a yellow photon carries more energy than an orange photon: $\boxed{E_y > E_o}$.

(B) A wavelength (λ) of 400 nm is shorter than a wavelength of 500 nm. Since $f = \frac{c}{\lambda}$, light with shorter wavelength has higher frequency. Thus, a 400-nm photon has higher frequency than a 500-nm photon. According to $E_\gamma = hf$, a 400-nm photon carries more energy than a 500-nm photon: $\boxed{E_{400} > E_{500}}$.

(C) A frequency (f) of 6.0×10^{14} Hz is higher than a frequency of 5.0×10^{14} Hz. According to $E_\gamma = hf$, a 6.0×10^{14}-Hz photon carries more energy than a 5.0×10^{14}-Hz photon: $\boxed{E_6 > E_5}$.

(D) According to the chart on page 84 in Chapter 6, ultraviolet light has higher frequency than infrared light. According to $E_\gamma = hf$, an ultraviolet photon carries more energy than an infrared photon: $\boxed{E_{UV} > E_{IR}}$.

(E) According to the chart on page 84 in Chapter 6, gamma rays have higher frequency than ultraviolet light. According to $E_\gamma = hf$, a gamma ray carries more energy than an ultraviolet photon: $\boxed{E_g > E_{UV}}$.

(F) According to the chart on page 84 in Chapter 6, x-rays have higher frequency than microwaves. According to $E_\gamma = hf$, an x-ray carries more energy than a microwave photon: $\boxed{E_x > E_\mu}$.

2. (A) Calculate the cutoff frequency using $f_c = \frac{W_0}{h}$. Convert the work function to Joules via $1 \text{ eV} = 1.6021766 \times 10^{-19}$ J so that its units are consistent with $h = 6.626 \times 10^{-34}$ J·s.

$$W_{0,low} = 4.0 \text{ eV} = 4.0 \times 1.6021766 \times 10^{-19} = 6.4 \times 10^{-19} \text{ J}$$

$$W_{0,high} = 5.0 \text{ eV} = 5.0 \times 1.6021766 \times 10^{-19} = 8.0 \times 10^{-19} \text{ J}$$

The cutoff frequencies corresponding to $W_{0,low} = 6.4 \times 10^{-19}$ J and $W_{0,high} = 8.0 \times 10^{-19}$ J are:

$$f_{c,low} = \frac{W_{0,low}}{h} = \frac{6.4 \times 10^{-19}}{6.626 \times 10^{-34}} = \boxed{9.7 \times 10^{14} \text{ Hz}}$$

$$f_{c,high} = \frac{W_{0,high}}{h} = \frac{8.0 \times 10^{-19}}{6.626 \times 10^{-34}} = 12 \times 10^{14} \text{ Hz} = \boxed{1.2 \times 10^{15} \text{ Hz}}$$

The cutoff frequency (f_c) represents the minimum frequency needed in order to observe the photoelectric effect. The frequencies of visible light lie in the approximate range 4.00×10^{14} Hz to 7.89×10^{14} Hz (corresponding to wavelengths of 380 nm to 750 nm according to $\lambda = \frac{c}{f}$).

Since $f_{c,low}$ and $f_{c,high}$ are both well above the frequencies of visible light (they are ultraviolet), visible light won't result in the photoelectric effect.

(B) Whereas the cutoff frequency (f_c) represents the minimum frequency needed in order to observe the photoelectric effect, the cutoff wavelength (λ_c) represents the maximum wavelength that can result in the photoelectric effect (since $\lambda = \frac{c}{f}$). That is, in order to observe the photoelectric effect, we need $f > f_c$ which is equivalent to $\lambda < \lambda_c$. Therefore, if the cutoff wavelength is at least 380 nm, we would be able to observe the photoelectric effect for at least *some* wavelength of visible light. If the cutoff wavelength is 750 nm, we would be able to observe the photoelectric effect for *all* frequencies of visible light. Use $\lambda_{c,violet} = 380$ nm and $\lambda_{c,red} =$ 750 nm (for extreme violet and extreme red) to find the corresponding cutoff frequencies. Recall that the speed of light in vacuum is $c = 2.9979 \times 10^8$ m/s.

$$f_{c,violet} = \frac{c}{\lambda_{c,violet}} = \frac{2.9979 \times 10^8}{380 \times 10^{-9}} = 7.89 \times 10^{14} \text{ Hz}$$

$$f_{c,red} = \frac{c}{\lambda_{c,red}} = \frac{2.9979 \times 10^8}{750 \times 10^{-9}} = 4.00 \times 10^{14} \text{ Hz}$$

Use these cutoff frequencies to find the corresponding values of the work function in Joules.

$$W_{0,violet} = hf_{c,violet} = (6.626 \times 10^{-34})(7.89 \times 10^{14}) = 5.23 \times 10^{-19}\text{ J}$$

$$W_{0,red} = hf_{c,red} = (6.626 \times 10^{-34})(4.00 \times 10^{14}) = 2.65 \times 10^{-19}\text{ J}$$

To convert from Joules to electron Volts, use $1\text{ J} = \frac{1\text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}\text{ eV}$.

$$W_{0,violet} = 5.23 \times 10^{-19}\text{ J} \times 6.2415092 \times 10^{18} = \boxed{3.26\text{ eV}}$$

$$W_{0,red} = 2.65 \times 10^{-19}\text{ J} \times 6.2415092 \times 10^{18} = \boxed{1.65\text{ eV}}$$

If $W_0 < 3.26$ eV, we would observe the photoelectric effect for *some* frequency of visible light (extreme violet only), whereas if $W_0 < 1.65$ eV we would observe the photoelectric effect for *all* frequencies of visible light (red thru violet).

3. The intensity of the laser beam ($I = 1.5$ mW $= 0.0015$ W) equals the number of photons per square meter per second times the energy per photon ($E_\gamma = hf$): $I = NE_\gamma$. We need to determine the energy of each photon, which means first finding the frequency. The frequency can be found from the wavelength ($\lambda = 633$ nm) using $c = 2.9979 \times 10^8$ m/s.

$$f = \frac{c}{\lambda} = \frac{2.9979 \times 10^8}{633 \times 10^{-9}} = 4.74 \times 10^{14} \text{ Hz}$$

Use Planck's constant ($h = 6.626 \times 10^{-34}$ J·s) to find the energy of each photon.

$$E_\gamma = hf = (6.626 \times 10^{-34})(4.74 \times 10^{14}) = 3.14 \times 10^{-19} \text{ J}$$

Since $I = NE_\gamma$, it follows that $N = \frac{I}{E_\gamma}$.

$$N = \frac{I}{E_\gamma} = \frac{0.0015}{3.14 \times 10^{-19}} = \boxed{4.78 \times 10^{15} \text{ photons/m}^2\text{/s}}$$

4. Since Planck's constant ($h = 6.626 \times 10^{-34}$ J·s) involves a Joule, whereas the work function and energies are given in or requested in electron Volts (eV), we can make it simpler to deal with the units in this problem by expressing Planck's constant in terms of electron Volts. To convert from Joules to electron Volts, use $1\text{ J} = \frac{1\text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV.

$$h = 6.626 \times 10^{-34} \times 6.2415092 \times 10^{18} = 4.136 \times 10^{-15} \text{ eV·s}$$

We will use this value of Planck's constant in units of eV·s for this problem.

(A) Multiply the frequency by Planck's constant. Recall that the prefix Tera equals $\text{T} = 10^{12}$.

$$E_\gamma = hf = (4.136 \times 10^{-15} \text{ eV·s})(900 \times 10^{12} \text{ Hz}) = 3.72 \text{ eV}$$

(B) Divide the work function by Planck's constant to find the cutoff frequency.

$$f_c = \frac{W_0}{h} = \frac{3.00 \text{ eV}}{4.136 \times 10^{-15} \text{ eV·s}} = \boxed{7.25 \times 10^{14} \text{ Hz}} = \boxed{725 \text{ THz}}$$

(C) Divide the speed of light in vacuum by the cutoff frequency to find the cutoff wavelength.

$$\lambda_c = \frac{c}{f_c} = \frac{2.9979 \times 10^8}{7.25 \times 10^{14}} = \boxed{4.14 \times 10^{-7} \text{ m}} = \boxed{414 \text{ nm}}$$

(D) The maximum kinetic energy of the photoelectrons is:

$$K_m = hf - W_0 = E_\gamma - W_0 = 3.72 \text{ eV} - 3.00 \text{ eV} = \boxed{0.72 \text{ eV}}$$

(E) The maximum kinetic energy is related to the stopping potential by $K_m = eV_s$. Divide by the charge of a proton (the same as an electron, but positive). The e from eV will cancel.

$$V_s = \frac{K_m}{e} = \frac{0.72 \text{ eV}}{e} = \boxed{0.72 \text{ V}}$$

5. As we did in the solution to Problem 4, we will express Planck's constant as $h = 6.626 \times 10^{-34} \times 6.2415092 \times 10^{18} = 4.136 \times 10^{-15}$ eV·s for this problem.

(A) The maximum kinetic energy of the photoelectrons is $K_m = 2.00$ eV. The maximum kinetic energy is related to the stopping potential by $K_m = eV_s$. Divide by the charge of a proton (the same as an electron, but positive). The e from eV will cancel.

$$V_s = \frac{K_m}{e} = \frac{2.00 \text{ eV}}{e} = \boxed{2.00 \text{ V}}$$

(B) The maximum kinetic energy of the photoelectrons is related to the work function via $K_m = hf - W_0$. Add W_0 to both sides to get $K_m + W_0 = hf$. Subtract K_m from both sides to get $W_0 = hf - K_m$. Find the frequency from the wavelength using $c = 2.9979 \times 10^8$ m/s.

$$f = \frac{c}{\lambda} = \frac{2.9979 \times 10^8}{300 \times 10^{-9}} = 9.99 \times 10^{14} \text{ Hz}$$

Now we can use the equation $W_0 = hf - K_m$.

$$W_0 = hf - K_m = (4.136 \times 10^{-15} \text{ eV·s})(9.99 \times 10^{14} \text{ Hz}) - 2.00 \text{ eV} = \boxed{2.13 \text{ eV}}$$

(C) Divide the work function by Planck's constant to find the cutoff frequency.

$$f_c = \frac{W_0}{h} = \frac{2.13 \text{ eV}}{4.136 \times 10^{-15} \text{ eV·s}} = 5.15 \times 10^{14} \text{ Hz}$$

Divide the speed of light in vacuum by the cutoff frequency to find the cutoff wavelength.

$$\lambda_c = \frac{c}{f_c} = \frac{2.9979 \times 10^8}{5.15 \times 10^{14}} = \boxed{5.82 \times 10^{-7} \text{ m}} = \boxed{582 \text{ nm}}$$

6. Note that this problem involves two different stopping potentials ($V_{1s} = 1.75$ V and V_{2s}) and two different wavelengths ($\lambda_1 = 250$ nm and $\lambda_2 = 125$ nm). Stopping potential is related to frequency via $V_s = \frac{hf}{e} - \frac{W_0}{e}$. Since $f = \frac{c}{\lambda}$, we may write $V_s = \frac{hc}{e\lambda} - \frac{W_0}{e}$. Write this equation for each case:

$$V_{1s} = \frac{hc}{e\lambda_1} - \frac{W_0}{e} \quad , \quad V_{2s} = \frac{hc}{e\lambda_2} - \frac{W_0}{e}$$

(A) Plug the given values of the stopping potential and wavelength into the first equation.

$$1.75 \text{ V} = \frac{hc}{e(250 \times 10^{-9} \text{ m})} - \frac{W_0}{e}$$

Recall that $h = 6.626 \times 10^{-34} \times 6.2415092 \times 10^{18} = 4.136 \times 10^{-15}$ eV·s (see the solution to Problem 4) and $c = 2.9979 \times 10^8$ m/s.

$$1.75 \text{ V} = \frac{(4.136 \times 10^{-15} \text{ eV·s})(2.9979 \times 10^8 \text{ m/s})}{e(250 \times 10^{-9} \text{ m})} - \frac{W_0}{e}$$

$$1.75 \text{ V} = \frac{1.24 \times 10^{-6} \text{ eV·m}}{e(250 \times 10^{-9} \text{ m})} - \frac{W_0}{e}$$

$$1.75 \text{ V} = \frac{4.96 \text{ eV}}{e} - \frac{W_0}{e}$$

Multiply both sides of the equation by the charge of a proton. (Note that e is the charge of a proton, such that $-e$ is the charge of an electron.) Note that e times a Volt makes eV.

$$1.75 \text{ eV} = 4.96 \text{ eV} - W_0$$

$$W_0 + 1.75 \text{ eV} = 4.96 \text{ eV}$$

$$W_0 = \boxed{3.21 \text{ eV}}$$

(B) Now use the second equation from the top of the page.

$$V_{2s} = \frac{hc}{e\lambda_2} - \frac{W_0}{e}$$

$$V_{2s} = \frac{(4.136 \times 10^{-15} \text{ eV·s})(2.9979 \times 10^8 \text{ m/s})}{e(125 \times 10^{-9} \text{ m})} - \frac{3.21 \text{ eV}}{e}$$

$$V_{2s} = 9.92 \text{ V} - 3.21 \text{ V} = \boxed{6.71 \text{ V}}$$

(C) To find the cutoff wavelength, set $V_s = 0$ in the equation $V_s = \frac{hc}{e\lambda} - \frac{W_0}{e}$ to get $0 = \frac{hc}{e\lambda_c} - \frac{W_0}{e}$. Add $\frac{W_0}{e}$ to both sides to get $\frac{W_0}{e} = \frac{hc}{e\lambda_c}$. Cross multiply: $e\lambda_c W_0 = ehc$. Thus, $\lambda_c = \frac{hc}{W_0}$.

$$\lambda_c = \frac{hc}{W_0} = \frac{(4.136 \times 10^{-15} \text{ eV·s})(2.9979 \times 10^8 \text{ m/s})}{3.21 \text{ eV}} = \boxed{3.86 \times 10^{-7} \text{ m}} = \boxed{386 \text{ nm}}$$

(D) The maximum kinetic energy is related to the stopping potential by $K_m = eV_s$.

$$K_{1m} = eV_{1s} = e(1.75 \text{ V}) = \boxed{1.75 \text{ eV}}$$

(E) The maximum kinetic energy is related to the stopping potential by $K_m = eV_s$.

$$K_{2m} = eV_{1s} = e(6.71 \text{ V}) = \boxed{6.71 \text{ eV}}$$

Note that $V_{2s} = 6.71$ V goes with $\lambda_2 = 125$ nm whereas $V_{1s} = 1.75$ V goes with $\lambda_1 = 250$ nm.

8 THE COMPTON EFFECT

Relevant Terminology

Photon – a particle of light.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Quantum – a fixed elemental unit corresponding to the minimum possible value that can be measured for a quantity that comes in discrete bundles like energy or angular momentum.

Quantized – limited to integer multiples of a quantum unit. A quantity like energy or angular momentum that is quantized is discrete (rather than continuous).

Compton Scattering

In Compton scattering, a **photon** (γ) scatters off an **electron** (e^-) that is initially at rest.

$$\gamma + e^- \rightarrow \gamma + e^-$$

Setup a coordinate system with $+x$ along the velocity of the initial photon (γ). Momentum and energy are both conserved for relativistic collisions between elementary particles (see Chapter 5). According to the law of **conservation of energy**,

$$E_{\gamma i} + E_{ei} = E_{\gamma f} + E_{ef}$$

The law of **conservation of momentum** gives two equations: one for x- and y-components of momentum. The initial photon is heading entirely along $+x$ such that $p_{\gamma i,x} = p_{\gamma i}$ and $p_{\gamma i,y} = 0$. The electron is initially at rest: $p_{ei,x} = p_{ei,y} = 0$. The components of the final momenta involve trig with the angles θ and φ shown in the figure above.

$$p_{\gamma i,x} + p_{ei,x} = p_{\gamma f,x} + p_{ef,x} \quad \rightarrow \quad p_{\gamma i} + 0 = p_{\gamma f} \cos\theta + p_{ef} \cos\varphi$$

$$p_{\gamma i,y} + p_{ei,y} = p_{\gamma f,y} + p_{ef,y} \quad \rightarrow \quad 0 + 0 = p_{\gamma f} \sin\theta - p_{ef} \sin\varphi$$

The energy and momentum of each particle are related through the equation $E^2 = p^2c^2 + m_0^2c^4$ from Chapter 5. When we apply this equation to the photons, we get $E_{\gamma i} = p_{\gamma i}c$ and $E_{\gamma f} = p_{\gamma f}c$ because the photons are massless. For the electrons, $E_{ei} = \sqrt{p_{ei}^2c^2 + m_e^2c^4} = m_ec^2$ since the electron is initially at rest ($p_{ei} = 0$) and $E_{ef} = \sqrt{p_{ef}^2c^2 + m_e^2c^4}$. Plug these equations into the equation for conservation of energy.

$$p_{\gamma i}c + m_ec^2 = p_{\gamma f}c + \sqrt{p_{ef}^2c^2 + m_e^2c^4}$$

Subtract $p_{\gamma f}c$ from both sides of the equation.

$$p_{\gamma i}c + m_ec^2 - p_{\gamma f}c = \sqrt{p_{ef}^2c^2 + m_e^2c^4}$$

Square both sides of the equation. Recall from algebra that $\left(\sqrt{x}\right)^2 = x$ and that $(a+b+c)^2 = (a+b+c)(a+b+c) = a^2 + 2ab + b^2 + 2bc + c^2 + 2ac$.

$$p_{\gamma i}^2 c^2 + 2m_e p_{\gamma i} c^3 + m_e^2 c^4 - 2m_e p_{\gamma f} c^3 + p_{\gamma f}^2 c^2 - 2p_{\gamma i} p_{\gamma f} c^2 = p_{ef}^2 c^2 + m_e^2 c^4$$

Note that $m_e^2 c^4$ cancels. Divide both sides by c^2.

$$p_{\gamma i}^2 + 2m_e p_{\gamma i} c - 2m_e p_{\gamma f} c + p_{\gamma f}^2 - 2p_{\gamma i} p_{\gamma f} = p_{ef}^2$$

Isolate $p_{ef} \cos\varphi$ and $p_{ef} \sin\varphi$ in the equations from conservation of momentum.

$$p_{ef} \cos\varphi = p_{\gamma i} - p_{\gamma f} \cos\theta$$

$$p_{ef} \sin\varphi = p_{\gamma f} \sin\theta$$

Square each equation. Recall from algebra that $(A-B)^2 = A^2 - 2AB + B^2$.

$$p_{ef}^2 \cos^2\varphi = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2 \cos^2\theta$$

$$p_{ef}^2 \sin^2\varphi = p_{\gamma f}^2 \sin^2\theta$$

Add these equations together. Recall the trig identity $\cos^2\varphi + \sin^2\varphi = 1$.

$$p_{ef}^2 \cos^2\varphi + p_{ef}^2 \sin^2\varphi = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2 \cos^2\theta + p_{\gamma f}^2 \sin^2\theta$$

$$p_{ef}^2 (\cos^2\varphi + \sin^2\varphi) = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2 (\cos^2\theta + \sin^2\theta)$$

$$p_{ef}^2 = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2$$

Now we have two different equations that both equal p_{ef}^2. Set these equations equal.

$$p_{\gamma i}^2 + 2m_e p_{\gamma i} c - 2m_e p_{\gamma f} c + p_{\gamma f}^2 - 2p_{\gamma i} p_{\gamma f} = p_{\gamma i}^2 - 2p_{\gamma i} p_{\gamma f} \cos\theta + p_{\gamma f}^2$$

Note that $p_{\gamma i}^2$ and $p_{\gamma f}^2$ both cancel. Divide both sides of the equation by 2.

$$m_e p_{\gamma i} c - m_e p_{\gamma f} c - p_{\gamma i} p_{\gamma f} = -p_{\gamma i} p_{\gamma f} \cos\theta$$

Add $p_{\gamma i} p_{\gamma f}$ to both sides of the equation and factor. Then divide both sides by $p_{\gamma i} p_{\gamma f}$.

$$m_e p_{\gamma i} c - m_e p_{\gamma f} c = p_{\gamma i} p_{\gamma f} (1 - \cos\theta)$$

$$\frac{m_e c}{p_{\gamma f}} - \frac{m_e c}{p_{\gamma i}} = 1 - \cos\theta$$

Now divide both sides of the equation by $m_e c$.

$$\frac{1}{p_{\gamma f}} - \frac{1}{p_{\gamma i}} = \frac{1 - \cos\theta}{m_e c}$$

This is the equation for **Compton scattering**. (This was actually the solution to Problem 5 in Chapter 5.) On the following page, we will rewrite this equation in terms of wavelength rather than momentum.

Photon Energy, Momentum, Frequency, and Wavelength

The energy and momentum of a particle are related by $E^2 = p^2c^2 + m_0^2c^4$ (see Chapter 5). Since a photon has zero rest mass, the energy and momentum of a photon are related by

$$E_\gamma = p_\gamma c$$

where p_γ is the (nonnegative) magnitude of the photon's momentum. In Chapter 7 (on the photoelectric effect), we learned that a photon's energy is proportional its frequency, where $h = 6.626 \times 10^{-34}$ J·s is Planck's constant.

$$E_\gamma = hf$$

Set the two equations for a photon's energy equal to one another in order to relate a photon's momentum to its frequency.

$$p_\gamma = \frac{hf}{c}$$

A photon's wavelength and frequency are related by $\lambda f = c$. Divide both sides by c to get $\frac{\lambda f}{c} = 1$ and then divide both sides by λ to see that $\frac{f}{c} = \frac{1}{\lambda}$. Replace $\frac{f}{c}$ by $\frac{1}{\lambda}$ in the equation above.

$$p_\gamma = \frac{h}{\lambda}$$

The Compton Effect

Substitute $p_\gamma = \frac{h}{\lambda}$ into the Compton scattering equation at the bottom of page 130. Note that $\frac{1}{p_\gamma} = \frac{\lambda}{h}$.

$$\frac{1}{p_{\gamma f}} - \frac{1}{p_{\gamma i}} = \frac{1 - \cos\theta}{m_e c}$$

$$\frac{\lambda_f}{h} - \frac{\lambda_i}{h} = \frac{1 - \cos\theta}{m_e c}$$

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$

The **wavelength** of the incident photon is shifted during **Compton scattering**, where θ is the scattering angle between the incident and final photon shown in the figure on page 129. The wavelength of the scattered photon is longer than the wavelength of the incident photon (except for a grazing collision for which $\theta = 0°$ exactly, in which case the incident photon isn't deflected at all). Note that $0 \leq \theta \leq 180°$, where $180°$ is for a perfectly head-on collision.

Photons Scattering with Bound Electrons

Consider a beam of photons (such as an x-ray) incident upon a slab of material (such as lead or graphite). In this case, an incident photon may interact with a bound electron (that is, an electron that is originally bound to an atom in the material). In this case, a plot of the intensity of the scattered photons (x-rays) as a function of wavelength shows two peaks:

- One peak corresponds to the wavelength of the scattered photon as predicted for the **Compton effect:** $\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$. This peak occurs when an incident photon interacts with an electron that is loosely bound to the atom.
- A second peak corresponds to the wavelength of the incident photon (λ_i). When an incident photon interacts with an electron that is tightly bound to the atom, such that the electron isn't freed from the atom, the photon effectively scatters off the atom as a whole rather than just the electron. If you replace the mass of an electron, m_e, with the mass of an entire atom in the Compton effect equation, the shift in wavelength will be negligible (virtually no shift at all).

Symbols and SI Units

Symbol	Name	SI Units
θ	angle between the incident and final photons	° or rad
φ	angle between the incident photon and final electron	° or rad
e	elementary charge	C
E	energy	J
p	momentum	$\frac{\text{kg}\cdot\text{m}}{\text{s}}$
m_e	rest mass of an electron	kg
c	the speed of light in vacuum	m/s
h	Planck's constant	J·s or J/Hz
λ	wavelength	m
λ_i	wavelength of the incident photon	m
λ_f	wavelength of the final photon	m
f	frequency	Hz

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^{8}$ m/s
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C

Electron Volts and Angstroms

The conversion from electron Volts (eV) to Joules (J) is:

$$1\ \text{eV} = 1.6021766 \times 10^{-19}\ \text{J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1\ \text{J} = \frac{1\ \text{eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}\ \text{eV}$$

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Strategy for Problems Involving the Compton Effect

If a problem involves the Compton effect, follow these steps:

- The equation for the Compton effect is
$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$
where λ_i and λ_f are the wavelengths of the incident and final photons, θ is the angle between the incident and final photons, $h = 6.626 \times 10^{-34}$ J·s, $c = 2.9979 \times 10^8$ m/s, and $m_e = 9.10938356 \times 10^{-31}$ kg.
- If the photon scatters off a particle other than an electron, replace m_e with the mass of the specified particle. If the mass is in units of in $\frac{\text{eV}}{c^2}$, note that $c = 2.9979 \times 10^8$ m/s and 1 eV = $1.6021766 \times 10^{-19}$ J (see page 69 in Chapter 5).
- One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.
- To find the kinetic energy of the recoiling electron, $K_{ef} = E_{ef} - m_e c^2$ (Chapter 5).
- Fractional energy loss can be expressed as $\frac{|\Delta E|}{E_i} = \frac{|E_f - E_i|}{E_i}$. Recall that $E_\gamma = hf$ and $f = \frac{c}{\lambda}$.
- If a problem asks you to derive the equation for the Compton effect, this was done on pages 129-131. If you need to derive another equation involving Compton scattering or a process that is similar to Compton scattering, study pages 129-131. It may also help to review the collisions from Chapter 5.

Example: A photon with an initial wavelength of 0.0200 nm scatters off a stationary electron at an angle of 60.0° compared to its initial direction.

(A) What is the wavelength of the scattered photon?

Use the Compton effect equation with $\theta = 60.0°$, $\lambda_i = 0.0200 \text{ nm} = 0.0200 \times 10^{-9}$ m, $h = 6.626 \times 10^{-34}$ J·s, $c = 2.9979 \times 10^8$ m/s, and $m_e = 9.10938356 \times 10^{-31}$ kg.

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$

$$\lambda_f - 0.0200 \times 10^{-9} = \frac{6.626 \times 10^{-34}}{(9.10938356 \times 10^{-31})(2.9979 \times 10^8)}(1 - \cos 60.0°)$$

$$\lambda_f - 0.0200 \times 10^{-9} = 2.426 \times 10^{-12}(1 - 0.500)$$

$$\lambda_f - 0.0200 \times 10^{-9} = 1.21 \times 10^{-12}$$

$$\lambda_f = 1.21 \times 10^{-12} + 0.0200 \times 10^{-9}$$

$$\lambda_f = 0.00121 \times 10^{-9} + 0.0200 \times 10^{-9} = \boxed{0.0212 \text{ nm}}$$

(B) What is the photon's fractional energy loss?

Divide the change in the photon's energy by the photon's initial energy. Recall that the energy of a photon is $E_\gamma = hf$ and that frequency is related to wavelength by $f = \frac{c}{\lambda}$.

$$\frac{|\Delta E_\gamma|}{E_{\gamma i}} = \frac{|E_{\gamma f} - E_{\gamma i}|}{E_{\gamma i}} = \frac{|hf_f - hf_i|}{hf_i} = \frac{|f_f - f_i|}{f_i} = \frac{\left|\frac{c}{\lambda_f} - \frac{c}{\lambda_i}\right|}{\frac{c}{\lambda_i}} = \frac{\left|\frac{1}{\lambda_f} - \frac{1}{\lambda_i}\right|}{\frac{1}{\lambda_i}} = \frac{\left|\frac{1}{0.0212} - \frac{1}{0.0200}\right|}{\frac{1}{0.0200}} = \boxed{0.057}$$

(C) What is the kinetic energy of the recoiling electron in electron Volts?

Conservation of energy for Compton scattering can be expressed as:

$$E_{\gamma i} + E_{ei} = E_{\gamma f} + E_{ef}$$

According to Chapter 5, the total energy of an electron is the sum of its kinetic energy and its rest energy: $E_e = K_e + m_e c^2$. The electron is initially at rest such that $E_{ei} = m_e c^2$. For the final electron, $E_{ef} = K_{ef} + m_e c^2$. For the photons, $E_\gamma = hf = \frac{hc}{\lambda}$ (since $f = \frac{c}{\lambda}$). Conservation of energy can therefore be expressed as:

$$\frac{hc}{\lambda_i} + m_e c^2 = \frac{hc}{\lambda_f} + K_{ef} + m_e c^2$$

Subtract $m_e c^2$ from both sides of the equation. It cancels out.

$$\frac{hc}{\lambda_i} = \frac{hc}{\lambda_f} + K_{ef}$$

$$K_{ef} = \frac{hc}{\lambda_i} - \frac{hc}{\lambda_f} = hc\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\right) = (6.626 \times 10^{-34})(2.9979 \times 10^8)\left(\frac{1}{0.0200 \times 10^{-9}} - \frac{1}{0.0212 \times 10^{-9}}\right)$$

$$K_{ef} = (6.626 \times 10^{-34})(2.9979 \times 10^8)(2.83 \times 10^9) = 5.62 \times 10^{-16}\text{ J}$$

Multiply by $1\text{ J} = \frac{1\text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV in order to convert from Joules (J) to electron Volts (eV).

$$K_{ef} = 5.62 \times 10^{-16} \times 6.2415092 \times 10^{18} = 3508\text{ eV} = \boxed{3.51\text{ keV}}$$

Example: Derive the following relationship between θ and φ for Compton scattering, where these angles are defined in the figure on page 129.

$$\cot\left(\frac{\theta}{2}\right) = \left(1 + \frac{h}{m_e c \lambda_i}\right) \tan\varphi$$

The scattering angles are used in the equations for conservation of momentum (page 129).

$$p_{\gamma i,x} + p_{ei,x} = p_{\gamma f,x} + p_{ef,x} \quad \rightarrow \quad p_{\gamma i} + 0 = p_{\gamma f}\cos\theta + p_{ef}\cos\varphi$$
$$p_{\gamma i,y} + p_{ei,y} = p_{\gamma f,y} + p_{ef,y} \quad \rightarrow \quad 0 + 0 = p_{\gamma f}\sin\theta - p_{ef}\sin\varphi$$

Isolate $p_{ef}\cos\varphi$ and $p_{ef}\sin\varphi$ in these equations.

$$p_{ef}\cos\varphi = p_{\gamma i} - p_{\gamma f}\cos\theta$$
$$p_{ef}\sin\varphi = p_{\gamma f}\sin\theta$$

Divide the bottom equation by the top equation. Recall from trig that $\frac{\sin\varphi}{\cos\varphi} = \tan\varphi$.

$$\tan\varphi = \frac{p_{\gamma f}\sin\theta}{p_{\gamma i} - p_{\gamma f}\cos\theta}$$

Recall that $E_\gamma = p_\gamma c = hf$ and $f = \frac{c}{\lambda}$ such that $p_\gamma = \frac{hf}{c} = \frac{h}{\lambda}$ (see page 131). Note that Planck's constant cancels. We will multiply by $\frac{\lambda_i \lambda_f}{\lambda_i \lambda_f}$ (which is the same as multiplying by one).

$$\tan\varphi = \frac{\frac{h}{\lambda_f}\sin\theta}{\frac{h}{\lambda_i} - \frac{h}{\lambda_f}\cos\theta} = \frac{\frac{1}{\lambda_f}\sin\theta}{\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\cos\theta} = \frac{\left(\frac{1}{\lambda_f}\sin\theta\right)\lambda_i\lambda_f}{\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\cos\theta\right)\lambda_i\lambda_f} = \frac{\lambda_i \sin\theta}{\lambda_f - \lambda_i\cos\theta}$$

Use the Compton effect equation, which may be expressed as $\lambda_f = \lambda_i + \frac{h}{m_e c}(1 - \cos\theta)$.

$$\tan\varphi = \frac{\lambda_i \sin\theta}{\lambda_i + \frac{h}{m_e c}(1 - \cos\theta) - \lambda_i\cos\theta} = \frac{\lambda_i \sin\theta}{\frac{h}{m_e c}(1 - \cos\theta) + \lambda_i(1 - \cos\theta)}$$

We factored λ_i from $\lambda_i - \lambda_i\cos\theta$. In the next step, we will factor out $(1 - \cos\theta)$.

$$\tan\varphi = \frac{\lambda_i \sin\theta}{\left(\frac{h}{m_e c} + \lambda_i\right)(1 - \cos\theta)} = \frac{\sin\theta}{\left(\frac{h}{m_e c\lambda_i} + 1\right)(1 - \cos\theta)}$$

We divided the numerator and denominator each by λ_i. In the next step, multiply both sides of the equation by $\left(\frac{h}{m_e c \lambda_i} + 1\right)$.

$$\left(\frac{h}{m_e c \lambda_i} + 1\right) \tan\varphi = \frac{\sin\theta}{1 - \cos\theta}$$

Recall the trig identity $\cot\left(\frac{\theta}{2}\right) = \frac{\sin\theta}{1-\cos\theta}$.

$$\boxed{\left(\frac{h}{m_e c \lambda_i} + 1\right) \tan\varphi = \cot\left(\frac{\theta}{2}\right)}$$

Note that this is equivalent to the equation that we were asked to derive.

Chapter 8 Problems

1. A photon with an initial wavelength of 0.1250 Angstroms scatters off a stationary electron.

(A) Which scattering angle maximizes the wavelength of the scattered photon?

(B) What is the maximum possible wavelength of the scattered photon?

(C) Which scattering angle minimizes the wavelength of the scattered photon?

(D) What is the minimum possible wavelength of the scattered photon?

(E) What is the maximum possible fractional energy loss for the photon?

(F) What is the maximum possible kinetic energy of the recoiling electron in keV?

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 180°

(B) 0.1735 Angstroms

(C) 0°

(D) 0.1250 Angstroms

(E) 28%

(F) 27.72 keV

2. A photon scatters off a stationary electron. Measurements show that the photon's energy is 50.0 keV before scattering and 49.0 keV after scattering.

(A) What is the photon's fractional energy loss?

(B) What is the kinetic energy of the recoiling electron in electron Volts?

(C) What are the wavelengths of the initial and final photons?

(D) What is the angle between the initial photon and the scattered photon?

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) 2.0%

(B) 1.0 keV

(C) 0.0248 nm, 0.0253 nm

(D) 37°

3. Derive the following relationship between the kinetic energy of the recoiling electron (K_{ef}) and the energy of the incident photon ($E_{\gamma i}$) for Compton scattering. As shown in the figure on page 129, θ is the angle between the incident photon and the scattered photon.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{1}{1 + \frac{m_e c \lambda_i}{2h} \csc^2\left(\frac{\theta}{2}\right)}$$

Want help? Check the solution at the end of the chapter.

Solutions to Chapter 8

1. Since one Angstrom (Å) is equal to 10^{-10} m, the wavelength of the incident photon equals $\lambda_i = 0.1250$ Å $= 0.1250 \times 10^{-10}$ m $= 1.250 \times 10^{-11}$ m.

(A) The equation for the Compton effect is $\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$, which can be written as $\lambda_f = \lambda_i + \frac{h}{m_e c}(1 - \cos\theta)$. Since λ_i was given in the problem and since, h, m_e, and c are all constants, the only variable to consider is the scattering angle (θ). We need to find the value of θ that maximizes $(1 - \cos\theta)$. Recall from trig that $-1 \leq \cos\theta \leq 1$. When $\cos\theta = -1$, this maximizes $1 - \cos\theta$, which will equal $1 - (-1) = 1 + 1 = 2$. Take the inverse cosine of both sides of $\cos\theta = -1$ to get $\theta = \cos^{-1}(-1) = 180°$. A scattering angle of $\boxed{\theta = 180°}$ will maximize the wavelength (λ_f) of the scattered photon.

(B) Plug the answer to Part A ($\theta = 180°$) into the equation for the Compton effect. Recall that $h = 6.626 \times 10^{-34}$ J·s, $c = 2.9979 \times 10^8$ m/s, and $m_e = 9.10938356 \times 10^{-31}$ kg.

$$\lambda_f = \lambda_i + \frac{h}{m_e c}(1 - \cos\theta)$$

$$\lambda_{f,max} = 1.250 \times 10^{-11} + \frac{6.626 \times 10^{-34}}{(9.10938356 \times 10^{-31})(2.9979 \times 10^8)}(1 - \cos 180.0°)$$

$$\lambda_{f,max} = 1.250 \times 10^{-11} + 2.426 \times 10^{-12}[1 - (-1)] = 1.250 \times 10^{-11} + 2.426 \times 10^{-12}(2)$$

$$\lambda_{f,max} = 1.250 \times 10^{-11} + 4.852 \times 10^{-12} = 1.250 \times 10^{-11} + 0.4852 \times 10^{-11}$$

$$\lambda_{f,max} = 1.735 \times 10^{-11} = 0.1735 \times 10^{-10} = \boxed{0.1735\ \text{Å}}$$

(C) This is similar to Part A, except that now we wish to minimize $(1 - \cos\theta)$ rather than maximize it. Recall that $-1 \leq \cos\theta \leq 1$. When $\cos\theta = 1$, this minimizes $1 - \cos\theta$, which will equal $1 - 1 = 0$. Take the inverse cosine of both sides of $\cos\theta = 1$ to get $\theta = \cos^{-1}(1) = 0°$. A scattering angle of $\boxed{\theta = 0°}$ will minimize the wavelength (λ_f) of the scattered photon.

(D) Plug the answer to Part C ($\theta = 0°$) into the equation for the Compton effect.

$$\lambda_f = \lambda_i + \frac{h}{m_e c}(1 - \cos\theta)$$

$$\lambda_{f,min} = 1.250 \times 10^{-11} + \frac{6.626 \times 10^{-34}}{(9.10938356 \times 10^{-31})(2.9979 \times 10^8)}(1 - \cos 0°)$$

$$\lambda_{f,min} = 1.250 \times 10^{-11} + 2.426 \times 10^{-12}(1 - 1) = 1.250 \times 10^{-11} + 2.426 \times 10^{-12}(0)$$

$$\lambda_{f,min} = 1.250 \times 10^{-11} = 0.1250 \times 10^{-10} = \boxed{0.1250\ \text{Å}}$$

(E) Divide the change in the photon's energy by the photon's initial energy. Recall that the energy of a photon is $E_\gamma = hf$ and that frequency is related to wavelength by $f = \frac{c}{\lambda}$.

$$\frac{|\Delta E_\gamma|}{E_{\gamma i}} = \frac{|E_{\gamma f} - E_{\gamma i}|}{E_{\gamma i}} = \frac{|hf_f - hf_i|}{hf_i} = \frac{|f_f - f_i|}{f_i} = \frac{\left|\frac{c}{\lambda_f} - \frac{c}{\lambda_i}\right|}{\frac{c}{\lambda_i}} = \frac{\left|\frac{1}{\lambda_f} - \frac{1}{\lambda_i}\right|}{\frac{1}{\lambda_i}}$$

The maximum energy loss corresponds to the maximum wavelength of the scattered photon. Use the answer to Part B: $\lambda_{f,max} = 0.1735$ Å. Recall that $\lambda_i = 0.1250$ Å.

$$\left(\frac{|\Delta E_\gamma|}{E_{\gamma i}}\right)_{max} = \frac{\left|\frac{1}{0.1735} - \frac{1}{0.1250}\right|}{\frac{1}{0.1250}} = \boxed{0.28}$$

If you prefer a percentage, multiply the decimal 0.28 by 100% to get $\boxed{28\%}$.

(F) Conservation of energy for Compton scattering can be expressed as:

$$E_{\gamma i} + E_{ei} = E_{\gamma f} + E_{ef}$$

According to Chapter 5, the total energy of an electron is the sum of its kinetic energy and its rest energy: $E_e = K_e + m_e c^2$. The electron is initially at rest such that $E_{ei} = m_e c^2$. For the final electron, $E_{ef} = K_{ef} + m_e c^2$. For the photons, $E_\gamma = hf = \frac{hc}{\lambda}$ (since $f = \frac{c}{\lambda}$). Conservation of energy can therefore be expressed as:

$$\frac{hc}{\lambda_i} + m_e c^2 = \frac{hc}{\lambda_f} + K_{ef} + m_e c^2$$

Subtract $m_e c^2$ from both sides of the equation. It cancels out.

$$\frac{hc}{\lambda_i} = \frac{hc}{\lambda_f} + K_{ef}$$

$$K_{ef} = \frac{hc}{\lambda_i} - \frac{hc}{\lambda_f} = hc\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\right)$$

The kinetic energy of the recoiling electron is maximum when the wavelength of the scattered photon is maximum. Use the answer to Part B: $\lambda_{f,max} = 0.1735$ Å. Recall that $\lambda_i = 0.1250$ Å.

$$K_{ef,max} = (6.626 \times 10^{-34})(2.9979 \times 10^8)\left(\frac{1}{0.1250 \times 10^{-10}} - \frac{1}{0.1735 \times 10^{-10}}\right)$$

$$K_{ef,max} = (6.626 \times 10^{-34})(2.9979 \times 10^8)(2.236 \times 10^{10}) = 4.442 \times 10^{-15} \text{ J}$$

Multiply by $1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV in order to convert from Joules (J) to electron Volts (eV).

$$K_{ef,max} = 4.442 \times 10^{-15} \times 6.2415092 \times 10^{18} = 27{,}724.8 \text{ eV} = \boxed{27.72 \text{ keV}}$$

2. We are given $E_{\gamma i} = 50.0$ keV and $E_{\gamma f} = 49.0$ keV.

(A) Divide the change in the photon's energy by the photon's initial energy.

$$\frac{|\Delta E_\gamma|}{E_{\gamma i}} = \frac{|E_{\gamma f} - E_{\gamma i}|}{E_{\gamma i}} = \frac{|49.0 - 50.0|}{50.0} = \boxed{0.020}$$

If you prefer a percentage, multiply the decimal 0.020 by 100% to get $\boxed{2.0\%}$.

(B) Conservation of energy for Compton scattering can be expressed as:

$$E_{\gamma i} + E_{ei} = E_{\gamma f} + E_{ef}$$

According to Chapter 5, the total energy of an electron is the sum of its kinetic energy and its rest energy: $E_e = K_e + m_e c^2$. The electron is initially at rest such that $E_{ei} = m_e c^2$. For the final electron, $E_{ef} = K_{ef} + m_e c^2$.

$$E_{\gamma i} + m_e c^2 = E_{\gamma f} + K_{ef} + m_e c^2$$

Subtract $m_e c^2$ from both sides of the equation. It cancels out.

$$E_{\gamma i} = E_{\gamma f} + K_{ef}$$

$$K_{ef} = E_{\gamma i} - E_{\gamma f} = 50.0 - 49.0 = \boxed{1.0 \text{ keV}} = 1000 \text{ eV}$$

(C) We know the energy of each photon and need to find wavelength. Recall that the energy of a photon equals $E_\gamma = hf$. Frequency is related to wavelength by $f = \frac{c}{\lambda}$. Combine these two equations to get $E_\gamma = hf = h\left(\frac{c}{\lambda}\right) = \frac{hc}{\lambda}$. Multiply both sides by λ to get $\lambda E_\gamma = hc$. Divide both sides by E_γ to get $\lambda = \frac{hc}{E_\gamma}$. Note that $E_{\gamma i}$ = 50.0 keV = 50,000 eV and $E_{\gamma f}$ = 49.0 keV = 49,000 eV. We have a problem: We know the energies in electron Volts (eV), but Planck's constant (h = 6.626×10^{-34} J·s) is expressed in terms of Joules (J). Recall that 1 eV = $1.6021766 \times 10^{-19}$ J.

$$\lambda_i = \frac{hc}{E_{\gamma i}} = \frac{(6.626 \times 10^{-34})(2.9979 \times 10^{8})}{50{,}000 \times 1.6021766 \times 10^{-19}} = \boxed{2.48 \times 10^{-11}\text{ m}} = \boxed{0.248\text{ Å}} = \boxed{0.0248\text{ nm}}$$

$$\lambda_f = \frac{hc}{E_{\gamma f}} = \frac{(6.626 \times 10^{-34})(2.9979 \times 10^{8})}{49{,}000 \times 1.6021766 \times 10^{-19}} = \boxed{2.53 \times 10^{-11}\text{ m}} = \boxed{0.253\text{ Å}} = \boxed{0.0253\text{ nm}}$$

(D) Use the equation for the Compton effect.

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$

$$2.53 \times 10^{-11} - 2.48 \times 10^{-11} = \frac{6.626 \times 10^{-34}}{(9.10938356 \times 10^{-31})(2.9979 \times 10^{8})}(1 - \cos\theta)$$

$$0.05 \times 10^{-11} = 2.426 \times 10^{-12}(1 - \cos\theta)$$

Divide both sides of the equation by 2.426×10^{-12}.

$$\frac{0.05 \times 10^{-11}}{2.426 \times 10^{-12}} = (1 - \cos\theta)$$

$$0.21 = 1 - \cos\theta$$

$$0.21 + \cos\theta = 1$$

$$\cos\theta = 1 - 0.21 = 0.79$$

$$\theta = \cos^{-1}(0.79) = \boxed{37°}$$

Note: The best practice is to keep a few extra digits throughout the calculation and round the final answer after the calculation is complete. (The 0.05×10^{-11} could be more precise.)

3. Conservation of energy for Compton scattering can be expressed as:

$$E_{\gamma i} + E_{ei} = E_{\gamma f} + E_{ef}$$

According to Chapter 5, the total energy of an electron is the sum of its kinetic energy and its rest energy: $E_e = K_e + m_e c^2$. The electron is initially at rest such that $E_{ei} = m_e c^2$. For the final electron, $E_{ef} = K_{ef} + m_e c^2$.

$$E_{\gamma i} + m_e c^2 = E_{\gamma f} + K_{ef} + m_e c^2$$

Subtract $m_e c^2$ from both sides of the equation. It cancels out.

$$E_{\gamma i} = E_{\gamma f} + K_{ef}$$

Subtract $E_{\gamma f}$ from both sides of the equation.

$$E_{\gamma i} - E_{\gamma f} = K_{ef}$$

Divide both sides of the equation by $E_{\gamma i}$.

$$\frac{E_{\gamma i}}{E_{\gamma i}} - \frac{E_{\gamma f}}{E_{\gamma i}} = \frac{K_{ef}}{E_{\gamma i}}$$

$$1 - \frac{E_{\gamma f}}{E_{\gamma i}} = \frac{K_{ef}}{E_{\gamma i}}$$

We can simply swap sides and write the following. (If left equals right, then right equals left.)

$$\frac{K_{ef}}{E_{\gamma i}} = 1 - \frac{E_{\gamma f}}{E_{\gamma i}}$$

Note that $E_\gamma = hf = \frac{hc}{\lambda}$ since $f = \frac{c}{\lambda}$.

$$\frac{K_{ef}}{E_{\gamma i}} = 1 - \frac{hc/\lambda_f}{hc/\lambda_i}$$

To divide fractions, multiply by the reciprocal of the bottom fraction.

$$\frac{K_{ef}}{E_{\gamma i}} = 1 - \frac{hc}{\lambda_f}\frac{\lambda_i}{hc} = 1 - \frac{\lambda_i}{\lambda_f}$$

Make a common denominator.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{\lambda_f}{\lambda_f} - \frac{\lambda_i}{\lambda_f} = \frac{\lambda_f - \lambda_i}{\lambda_f}$$

Apply the equation for the Compton effect: $\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{h}{\lambda_f m_e c}(1 - \cos\theta)$$

Apply the equation for the Compton effect again, this time as $\lambda_f = \lambda_i + \frac{h}{m_e c}(1 - \cos\theta)$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{\frac{h}{m_e c}(1 - \cos\theta)}{\lambda_i + \frac{h}{m_e c}(1 - \cos\theta)}$$

Multiply the numerator and denominator each by $\frac{m_e c}{h}$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{1 - \cos\theta}{\frac{m_e c \lambda_i}{h} + 1 - \cos\theta}$$

Recall the trig identity $\sin^2\left(\frac{\theta}{2}\right) = \frac{1 - \cos\theta}{2}$, which shows that $1 - \cos\theta = 2\sin^2\left(\frac{\theta}{2}\right)$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{2\sin^2\left(\frac{\theta}{2}\right)}{\frac{m_e c \lambda_i}{h} + 2\sin^2\left(\frac{\theta}{2}\right)}$$

Divide the numerator and denominator each by $2\sin^2\left(\frac{\theta}{2}\right)$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{1}{\dfrac{m_e c \lambda_i}{2h \sin^2\left(\frac{\theta}{2}\right)} + 1}$$

Recall from trig that the cosecant function is the reciprocal of the sine function: $\csc x = \frac{1}{\sin x}$.

$$\frac{K_{ef}}{E_{\gamma i}} = \frac{1}{\frac{m_e c \lambda_i}{2h} \csc^2\left(\frac{\theta}{2}\right) + 1} = \boxed{\frac{1}{1 + \frac{m_e c \lambda_i}{2h} \csc^2\left(\frac{\theta}{2}\right)}}$$

Addition is commutative: The order of the terms doesn't matter in the denominator.

9 BOHR'S MODEL

Relevant Terminology

Photon – a particle of light.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Work – work is done when there is not only a force acting on an object, but when the force also contributes toward the displacement of an object.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Ionization energy – the energy needed to remove the electron(s) from an atom.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Potential energy – work that can be done by changing position. All forms of potential energy are stored energy.

Electric potential – electric potential energy per unit charge.

Angular momentum – moment of inertia times angular velocity.

Quantum – a fixed elemental unit corresponding to the minimum possible value that can be measured for a quantity that comes in discrete bundles like energy or angular momentum.

Quantized – limited to integer multiples of a quantum unit. A quantity like energy or angular momentum that is quantized is discrete (rather than continuous).

Atomic number – the number of protons in the nucleus of an atom.

Bohr’s Model of the Atom

Niels Bohr developed a simple quantum model of the hydrogen atom, which provides a good approximation for predicting energy levels and frequencies for spectral lines. According to Bohr’s model:

- The electron travels in a circular orbit where the centripetal force is given by **Coulomb’s law**: The magnitude of the force is $F = \frac{ke^2}{r^2}$, where e is the charge of the proton and $-e$ is the charge of the electron.
- The electron can only travel in an orbit where its **angular momentum** (L) equals an integer multiple of $\hbar$, where $\hbar = \frac{h}{2\pi}$. That is, $L = n\hbar$, where $n \in \{1,2,3,\cdots\}$.
- The electron doesn’t radiate electromagnetic energy while traveling in a circular orbit. The electron’s energy remains constant while it travels in the same orbit.
- If the atom emits or absorbs a photon, the electron jumps to a different orbit, where the **energy** difference is related to the **frequency** of the photon by $|E_f - E_i| = hf$.

Equations for Bohr's Model

The force with which the electron is attracted to the positively charged nucleus is given by **Coulomb's law**, where $e = 1.6021766 \times 10^{-19}$ C is the value of elementary charge. For a hydrogen-like atom with a single electron and atomic number Z (equal to the number of protons in the nucleus), the magnitude of the **force** is:

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

where $k = 8.99 \times 10^9 \frac{\text{N·m}^2}{\text{C}^2}$ is Coulomb's constant. Recall from electricity and magnetism that Coulomb's constant is related to the permittivity of free space ($\epsilon_0 = 8.85 \times 10^{-12} \frac{\text{C}^2}{\text{N·m}^2}$) via $k = \frac{1}{4\pi\epsilon_0}$. The electron travels in a circular orbit with **centripetal acceleration**.

$$a_c = \frac{v^2}{r}$$

According to **Newton's second law**, the net force exerted on the electron equals its mass times its acceleration: $F = m_e a_c$. Combine this with the previous equations to see that:

$$\frac{Zke^2}{r^2} = \frac{m_e v^2}{r}$$

Multiply both sides of the above equation by r and divide by m_e to get:

$$v^2 = \frac{Zke^2}{m_e r}$$

The **potential energy** of the electron is related to the force by $\vec{\mathbf{F}} = -\vec{\nabla}U$, where $\vec{\nabla}$ is the gradient operator. If we place the origin at the nucleus, $\vec{\mathbf{F}} = -\frac{\partial U}{\partial r}$, for which:

$$U = -\frac{Zke^2}{r}$$

The potential energy is negative for bound states. For larger values of r, the quantity $\frac{Zke^2}{r}$ gets smaller, meaning that $U = -\frac{Zke^2}{r}$ gets less negative (which means that that U increases). The potential energy (U) in Joules (J) or electron Volts (eV) is related to **electric potential** (V) in Volts (V) by the value of the electron's charge: $U = -eV$.

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

The usual formula for the **kinetic energy** of an electron is:

$$K = \frac{1}{2} m_e v^2$$

Plug $v^2 = \frac{Zke^2}{m_e r}$ (from earlier) into the kinetic energy equation to get:

$$K = \frac{Zke^2}{2r}$$

The **total energy** of the electron equals (since $\frac{1}{2} - 1 = -\frac{1}{2}$):

$$E = K + U = \frac{Zke^2}{2r} - \frac{Zke^2}{r} = -\frac{Zke^2}{2r}$$

The **angular momentum** (L) of the electron is:

$$L = m_e v r$$

The equations up to this point are from classical physics. From this point on, we will introduce Bohr's postulates. According to Bohr's model, the angular momentum of the electron is **quantized**:

$$L_n = n\hbar$$

where $\hbar = \frac{h}{2\pi}$, $h = 6.626 \times 10^{-34}$ J·s is **Planck's constant**, and n is a positive integer $(1,2,3,\cdots)$. From the previous equation onward, we will include a subscript n on any quantity that is limited to discrete values for Bohr's model. Combine the two previous equations to get:

$$m_e v_n r_n = n\hbar$$

Solve for the **speed** of the electron.

$$v_n = \frac{n\hbar}{m_e r_n}$$

Plug this into the equation for speed from the previous page: $v_n^2 = \frac{Zke^2}{m_e r_n}$.

$$\frac{n^2\hbar^2}{m_e^2 r_n^2} = \frac{Zke^2}{m_e r_n}$$

Solve for the **radius** of the electron's orbit.

$$r_n = \frac{n^2\hbar^2}{Zkm_e e^2}$$

When Z = 1 (for hydrogen) and $n = 1$ (for the ground state), we get the **Bohr radius** (a_0).

$$a_0 = \frac{\hbar^2}{km_e e^2} = 0.529\ \text{Å}$$

In terms of the Bohr radius, the radii of the allowed orbits are:

$$r_n = \frac{a_0 n^2}{Z}$$

Substitute $\frac{1}{r_n} = \frac{Zkm_e e^2}{n^2\hbar^2}$ (see above) into the previous equation for the electron's **speed**.

$$v_n = \frac{Zke^2}{n\hbar}$$

Since $\frac{1}{r_n} = \frac{Z}{a_0 n^2}$ and $v_n = \frac{n\hbar}{m_e r_n}$ (see above), the electron's speed can also be expressed as:

$$v_n = \frac{Z\hbar}{m_e a_0 n}$$

Substitute $\frac{1}{r_n} = \frac{Zkm_e e^2}{n^2\hbar^2}$ into the equation for the electron's **energy** at the top of the page.

$$E_n = -\frac{Z^2k^2m_e e^4}{2n^2\hbar^2}$$

Alternatively, plug $\frac{1}{r_n} = \frac{Z}{a_0 n^2}$ into $E = -\frac{Zke^2}{2r}$ (from the top of the previous page) to get:

$$E_n = -\frac{Z^2ke^2}{2a_0 n^2}$$

When $Z = 1$ (for hydrogen) and $n = 1$ (for the ground state), we get the **ground state energy** ($E_{g,H}$) for hydrogen.

$$E_{g,H} = -\frac{ke^2}{2a_0} = -2.18 \times 10^{-18}\text{ J} = -13.6\text{ eV}$$

In terms of the ground state energy for hydrogen, the electron's energy is:

$$E_n = \frac{Z^2 E_{g,H}}{n^2}$$

The extreme value $n \to \infty$ corresponds to a free electron with energy $E_\infty = 0$. If the atom **absorbs a photon**, the electron jumps to a higher energy level according to:

$$E_f - E_i = hf$$

If the electron drops down a lower energy level, the atom **emits a photon** according to:

$$E_i - E_f = hf$$

Note that f is the **frequency** of the photon (**<u>not</u>** the frequency of the electron's orbital motion). For the case where the electron drops down to a lower energy level:

$$f = \frac{E_i - E_f}{h} = \frac{Z^2 E_{g,H}}{h}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) = -\frac{Z^2 k e^2}{2a_0 h}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

The previous equation is often expressed in terms of **wavelength** using $f = \frac{c}{\lambda}$ (such that $\frac{1}{\lambda} = \frac{f}{c}$).

$$\frac{1}{\lambda} = \frac{f}{c} = \frac{E_i - E_f}{hc} = \frac{Z^2 E_{g,H}}{hc}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) = -\frac{Z^2 k e^2}{2a_0 hc}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

For $Z = 1$ (for hydrogen), the constants out front form **Rydberg's constant** (R_H).

$$R_H = \frac{ke^2}{2a_0 hc} = 1.10 \times 10^7 \text{m}^{-1}$$

Recall that $a_0 = \frac{\hbar^2}{km_e e^2}$. Plug $\frac{1}{a_0} = \frac{km_e e^2}{\hbar^2}$ into the equation above and recall that $\hbar = \frac{h}{2\pi}$ to get:

$$R_H = \frac{k^2 m_e e^4}{2h\hbar^2 c} = \frac{k^2 m_e e^4}{2(2\pi\hbar)\hbar^2 c} = \frac{k^2 m_e e^4}{4\pi\hbar^3 c}$$

Note that $h = 2\pi\hbar$ since $\hbar = \frac{h}{2\pi}$. In terms of Rydberg's constant, the reciprocal **wavelength** is:

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

Different values of n_f (if a photon is emitted; if instead it is absorbed, it is n_i) correspond to different series of **spectral lines**.

$n_f = 1$	$n_f = 2$	$n_f = 3$	$n_f = 4$	$n_f = 5$	$n_f = 6$
Lyman	Balmer	Paschen	Brackett	Pfund	Humphreys

Symbols and SI Units

Symbol	Name	SI Units
F	force	N
E	total energy	J
$E_{g,H}$	the ground state energy for hydrogen	J
K	kinetic energy	J
U	potential energy	J
V	electric potential	V
L	angular momentum	$\frac{kg \cdot m^2}{s}$
m_e	mass of an electron	kg
r	radius	m

Z	atomic number (number of protons in the nucleus)	unitless
e	elementary charge	C
k	Coulomb's constant	$\frac{\text{N·m}^2}{\text{C}^2}$ or $\frac{\text{kg·m}^3}{\text{C}^2\text{·s}^2}$
ϵ_0	permittivity of free space	$\frac{\text{C}^2}{\text{N·m}^2}$ or $\frac{\text{C}^2\text{·s}^2}{\text{kg·m}^3}$
h	Planck's constant	J·s or J/Hz
$\hbar$	h-bar	J·s or J/Hz
n	a positive integer	unitless
f	frequency of light	Hz
λ	wavelength	m

c	speed of light in vacuum	m/s
v	speed of the electron	m/s
a_c	centripetal acceleration	$\mathrm{m/s^2}$
a_0	Bohr radius	m
R_H	the Rydberg constant	$\frac{1}{\mathrm{m}}$
$\vec{\nabla}$	gradient operator	$\frac{1}{\mathrm{m}}$

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s
h-bar	$\hbar = 1.055 \times 10^{-34}$ J·s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C
Coulomb's constant	$k = 8.99 \times 10^9 \frac{\text{N·m}^2}{\text{C}^2}$
the permittivity of free space	$\epsilon_0 = 8.85 \times 10^{-12} \frac{\text{C}^2}{\text{N·m}^2}$

the Rydberg constant	$R_H = 1.10 \times 10^7 \text{m}^{-1}$
the Bohr radius	$a_0 = 0.529$ Å
the ground state energy for hydrogen	$E_{g,H} = -13.6$ eV

Electron Volts and Angstroms

The conversion from electron Volts (eV) to Joules (J) is:

$$1\ \text{eV} = 1.6021766 \times 10^{-19}\ \text{J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1\ \text{J} = \frac{1\ \text{eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}\ \text{eV}$$

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Important Distinction

Note the distinction between Planck's constant (h) and the related quantity $\hbar = \frac{h}{2\pi}$.

Strategy for Solving Problems Relating to Bohr's Model

To solve a problem that involves Bohr's model, follow these steps:

- If the problem gives you numerical values, identify the corresponding symbols. The charts on the previous pages define the symbols used in this chapter.

- If the problem asks you to derive an equation in symbols, make a list of the symbols that you want to have in your final answer. Any symbol not on this list you will want to eliminate using algebraic substitutions.

- Identify the desired unknown(s) that you are solving for.

- Find equations that involve the desired unknown(s) and the given values. Equations that relate to Bohr's model are listed and described earlier in this chapter.

- Apply algebra to solve for the desired unknown.

- If you are solving a problem with numerical values, you have two options regarding the units:

  - You can convert all values to SI units. If you're given an energy in electron Volts (eV), use the conversion 1 eV = $1.6021766 \times 10^{-19}$ J.
  - You can convert all Joules to electron Volts using 1 J = 6.2415092×10^{18} eV. Note that Planck's constant includes a Joule, such that $h = 4.136 \times 10^{-15}$ eV·s. Note that mass can be expressed in units of $\frac{\text{eV}}{c^2}$, where $c = 2.9979 \times 10^8$ m/s and $1\ \frac{\text{eV}}{c^2} = \frac{1.6021766 \times 10^{-19}\text{ J}}{(2.99792458 \times 10^8\text{ m/s})^2} = 1.7826619 \times 10^{-36}$ kg.

Example: Consider the lithium ion Li^{2+}, where the 2+ indicates that the atom has lost two electrons compared to neutral lithium.

(A) What is the radius of the ground state for Li^{2+}?

Consult a periodic table to see that the atomic number of lithium (Li) is $Z = 3$. For the ground state, $n = 1$. Use $r_n = \frac{a_0 n^2}{Z}$ to find r_1 from Z and n. Note that $a_0 = 0.529$ Å $= 0.529 \times 10^{-10}$ m.

$$r_1 = \frac{a_0(1)^2}{Z} = \frac{a_0(1)^2}{3} = \frac{a_0}{3} = \frac{0.529 \text{ Å}}{3} = \boxed{0.177 \text{ Å}} = \boxed{1.77 \times 10^{-11} \text{ m}}$$

(B) What is the ionization energy for Li^{2+}?

Although neutral lithium would have 3 protons and 3 electrons (and a few neutrons), the Li^{2+} ion has 3 protons and 1 electron (and a few neutrons). The ionization energy is thus the energy needed to remove the remaining electron from the ground state of Li^{2+}. Since a free electron has zero energy, the answer equals the absolute value of the ground state energy (this positive energy must be supplied to remove the electron): $E_{ioniz.} = |E_1|$. Use $E_n = \frac{Z^2 E_{g,H}}{n^2}$ to find E_1 from $Z = 3$ and $n = 1$. Note that $E_{g,H} = -13.6$ eV is the ground state energy for H.

$$E_{ioniz.,Li^{2+}} = |E_1| = \left|\frac{Z^2 E_{g,H}}{n^2}\right|_{n=1,Z=3} = \left|\frac{(3)^2 E_{g,H}}{(1)^2}\right| = 9|E_{g,H}| = 9|-13.6 \text{ eV}| = \boxed{122 \text{ eV}}$$

Example: An electron in a hydrogen atom drops down from the $n = 3$ orbital to the ground state. What is the wavelength of the emitted photon?

The atomic number of hydrogen is $Z = 1$. The electron is initially in the $n_i = 3$ orbital and drops down to the $n_f = 1$ orbital (that's the ground state). Use $\frac{1}{\lambda} = -Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$ to find the wavelength of the emitted photon from $Z, n_i,$ and n_f. Note that $R_H = 1.10 \times 10^7 \text{m}^{-1}$.

$$\frac{1}{\lambda} = -Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) = -(1)^2 R_H \left(\frac{1}{3^2} - \frac{1}{1^2}\right) = -R_H \left(\frac{1}{9} - \frac{1}{1}\right) = -R_H \left(\frac{1}{9} - \frac{9}{9}\right) = -R_H \left(-\frac{8}{9}\right)$$

$$\frac{1}{\lambda} = \frac{8}{9} R_H$$

$$\lambda = \frac{9}{8R_H} = \frac{9}{8(1.10 \times 10^7)} = \boxed{1.02 \times 10^{-7} \text{ m}} = \boxed{102 \text{ nm}}$$

Example: Derive the following formula for the period of the electron's motion in Bohr's model.

$$T_n = \frac{2\pi n^3 \hbar^3}{Z^2 k^2 m_e e^4}$$

We already derived an equation for the speed of the electron on page 141.

$$v_n = \frac{Zke^2}{n\hbar}$$

The period (T) is the time it takes for the electron to travel around its circular orbit once. Since the electron travels with constant speed, its speed equals the circumference ($C = 2\pi r_n$) divided by the period (T_n).

$$v_n = \frac{2\pi r_n}{T_n}$$

Set the two previous equations equal to one another.

$$\frac{2\pi r_n}{T_n} = \frac{Zke^2}{n\hbar}$$

Cross multiply.

$$2\pi r_n n\hbar = Zke^2 T_n$$

Recall an equation for r_n from page 141.

$$r_n = \frac{n^2\hbar^2}{Zkm_e e^2}$$

Substitute this expression into the previous equation.

$$\frac{2\pi n^3\hbar^3}{Zkm_e e^2} = Zke^2 T_n$$

Divide both sides of the equation by Zke^2.

$$T_n = \boxed{\frac{2\pi n^3\hbar^3}{Z^2k^2m_e e^4}}$$

Numerically, this works out to $T_n = (1.52 \times 10^{-16})\frac{n^3}{Z^2}$ sec.

Chapter 9 Problems

1. Determine each of the following quantities in the units specified for the $n = 4$ state of the hydrogen atom. Note: The mass of an electron is 9.109×10^{-31} kg.

(A) the orbital radius of the electron in Angstroms

(B) the electron's electric potential in V

(C) the speed of the electron in terms of c

(D) the electron's total energy in eV

(E) the acceleration of the electron in m/s^2

(F) the net force on the electron in N

(G) the electron's angular momentum in eV·s

(H) the electron's potential energy in eV

(I) the electron's linear momentum in eV/c

(J) the electron's kinetic energy in eV

(K) the period of the electron's motion in s

(L) the electron's frequency in Hz

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 8.46 Å (B) 1.70 V (C) $0.00183c$ (D) −0.850 eV (E) 3.55×10^{20} m/s^2 (F) 3.23×10^{-10} N (G) 2.63×10^{-15} eV·s (H) −1.70 eV (I) $935 \frac{\text{eV}}{c}$ (J) 0.856 eV (K) 9.70×10^{-15} s (L) 1.03×10^{14} Hz

2. Determine the wavelength of the photon for each of the following transitions for an electron in a hydrogen atom. Also indicate whether the photon is emitted or absorbed in each case.

(A) from $n = 3$ to $n = 2$

(B) from $n = 2$ to $n = 3$

(C) from $n = 4$ to $n = 7$

(D) from $n = 5$ to the ground state

(E) from $n = 6$ to $n = 2$

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) 655 nm, emitted

(B) 655 nm, absorbed

(C) 2160 nm, absorbed

(D) 94.7 nm, emitted

(E) 409 nm, emitted

Solutions to Chapter 9

1. Note that $Z = 1$ (for H), $n = 4$, $m_e = 9.109 \times 10^{-31}$ kg, $h = 6.626 \times 10^{-34}$ J·s, $\hbar = \frac{h}{2\pi}$, $c = 2.9979 \times 10^8$ m/s, $k = 8.99 \times 10^9 \frac{\text{N}\cdot\text{m}^2}{\text{C}^2}$, $e = 1.6021766 \times 10^{-19}$ C, $a_0 = 0.529$ Å $= 0.529 \times 10^{-10}$ m, and $E_{g,H} = -13.6$ eV. To convert from eV to J, use 1 eV = $1.6021766 \times 10^{-19}$ J, and to convert from J to eV, use 1 J = $\frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV.

(A) One formula for the orbital radius is $r_n = \frac{a_0 n^2}{Z}$.

$$r_4 = \frac{a_0(4)^2}{1} = 16a_0 = 16(0.529 \text{ Å}) = \boxed{8.46 \text{ Å}}$$

(B) The formula for the electron's electric potential (not to be confused with potential energy) is $V_n = \frac{Zke}{r_n}$. Note that one Angstrom equals 1 Å = 10^{-10} m: $r_4 = 8.46$ Å $= 8.46 \times 10^{-10}$ m.

$$V_4 = \frac{Zke}{r_4} = \frac{(1)(8.99 \times 10^9)(1.6021766 \times 10^{-19})}{8.46 \times 10^{-10} \text{ m}} = \boxed{1.70 \text{ V}}$$

(C) One formula for the speed of the electron is $v_n = \frac{n\hbar}{m_e r_n}$, where $\hbar = \frac{h}{2\pi} = 1.055 \times 10^{-34}$ J·s.

$$v_4 = \frac{4\hbar}{m_e r_4} = \frac{4(1.055 \times 10^{-34})}{(9.109 \times 10^{-31})(8.46 \times 10^{-10})} = 5.48 \times 10^5 \text{ m/s}$$

To express this in terms of the speed of light, divide by $c = 2.9979 \times 10^8$ m/s.

$$v_4 = \frac{5.48 \times 10^5}{2.9979 \times 10^8} c = \boxed{0.00183c}$$

Note: If you use the equation $v_n^2 = \frac{Zke^2}{m_e r_n}$, you will need to apply a squareroot in order to solve for the speed.

(D) One formula for the electron's total energy is $E_n = \frac{Z^2 E_{g,H}}{n^2}$, where $E_{g,H} = -13.6$ eV.

$$E_4 = \frac{Z^2 E_{g,H}}{4^2} = \frac{(1)^2(-13.6)}{16} = -0.850 \text{ eV}$$

(E) The electron has centripetal acceleration from its circular motion. Recall that $r_4 = 8.46\ \text{Å} = 8.46 \times 10^{-10}$ m and $v_4 = 5.48 \times 10^5$ m/s.

$$a_c = \frac{v_n^2}{r_n} = \frac{(5.48 \times 10^5)^2}{8.46 \times 10^{-10}} = \boxed{3.55 \times 10^{20}\ \text{m/s}^2}$$

Why is the acceleration so large in SI units? The electron has a high speed and a small orbit, and its acceleration describes how the direction of its velocity changes.

(F) According to Newton's second law, net force equals mass times acceleration.

$$F_{net} = m_e a_c = (9.109 \times 10^{-31})(3.55 \times 10^{20}) = \boxed{3.23 \times 10^{-10}\ \text{N}}$$

You can check your answer by comparing it with Coulomb's law for the force between the electron and the proton: $F_{net} = \frac{ke^2}{r_4^4}$. Be sure to square both e and r_4.

(G) If you use $L_n = n\hbar$, be sure to use $\hbar = \frac{h}{2\pi} = 1.055 \times 10^{-34}$ J·s and not $h = 6.626 \times 10^{-34}$ J·s.

$$L_n = n\hbar = (4)(1.055 \times 10^{-34}) = 4.22 \times 10^{-34} \text{ J·s}$$

Recall the conversion factor $1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV.

$$L_n = 4.22 \times 10^{-34} \times 6.2415092 \times 10^{18} = \boxed{2.63 \times 10^{-15} \text{ eV·s}}$$

You could alternatively use $L_n = m_e v_n r_n$, but if you use $m_e = 9.109 \times 10^{-31}$ kg, you will still need to convert from Joules (J) to electron Volts (eV).

(H) The simplest way to get the electron's potential energy (U) in electron Volts (eV) from its electric potential ($V_4 = 1.70$ V, found in Part B) is to multiply by the charge of an electron, which is $-e$.

$$U_4 = -eV_4 = (-e)(1.70) = \boxed{-1.70 \text{ eV}}$$

The long way is to use the formula $U_4 = -\frac{Zke^2}{r_4}$ to get -2.73×10^{-19} J and then multiply by $1 \text{ J} = 6.2415092 \times 10^{18}$ eV (but if you do it both ways it helps to check your answers).

(I) The correct formula for the electron's linear momentum is $p_4 = \gamma m_e v_4$, where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ (see Chapter 5, where we used u instead of v for the object's speed). However, since $v_4 = 0.00183c$ (found in Part C) is small compared to the speed of light, in this case we could use the equation $p_4 = m_e v_4$ from classical physics: $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-(0.00183)^2}} \approx 1$. In Chapter 5, we learned that $1\ \frac{\text{eV}}{c^2} = \frac{1.6021766 \times 10^{-19}\ \text{J}}{(2.99792458 \times 10^8\ \text{m/s})^2} = 1.7826619 \times 10^{-36}$ kg (see page 69). We can thus convert mass from kilograms to $\frac{\text{eV}}{c^2}$ using $1\ \text{kg} = \frac{1}{1.7826619 \times 10^{-36}} = 5.6095887 \times 10^{35}\ \frac{\text{eV}}{c^2}$.

$$m_e = 9.109 \times 10^{-31}\ \text{kg} = 9.109 \times 10^{-31} \times 5.6095887 \times 10^{35}\ \frac{\text{eV}}{c^2} = 5.11 \times 10^5\ \frac{\text{eV}}{c^2}$$

$$p_4 = \gamma m_e v_4 \approx (1)\left(5.11 \times 10^5\ \frac{\text{eV}}{c^2}\right)(0.00183c) = \boxed{935\ \frac{\text{eV}}{c}}$$

If you use $m_e = 9.109 \times 10^{-31}$ kg and $v_4 = 5.48 \times 10^5$ m/s to get $p_4 = 4.99 \times 10^{-24}\ \frac{\text{kg}\cdot\text{m}^2}{\text{s}^2}$, you will need to divide by $e = 1.6021766 \times 10^{-19}$ C and multiply by $c = 2.9979 \times 10^8$ m/s in order to get units of $\frac{\text{eV}}{c}$.

(J) As we noted in Part I, the relativistic effects are negligible since $v_4 = 0.00183c$ is small compared to the speed of light, so we may use the equation for kinetic energy from classical physics instead of the equation from special relativity (Chapter 5) to good approximation. (In fact, if relativistic effects were significant for an electron in a hydrogen atom, we would need to return all the way back to the bottom of page 140 and adjust the equation for kinetic energy back there.) Recall from Part (I) that $m_e = 5.11 \times 10^5 \frac{\text{eV}}{c^2}$.

$$K_4 = \frac{1}{2} m_e v_4^2 = \frac{1}{2}\left(5.11 \times 10^5 \frac{\text{eV}}{c^2}\right)(0.00183c)^2 \boxed{= 0.856 \text{ eV}}$$

If instead you use $m_e = 9.109 \times 10^{-31}$ kg and $v_4 = 5.48 \times 10^5$ m/s to get $K_4 = 1.37 \times 10^{-19}$ J, you will need to multiply by $1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}$ eV. Yet another way to obtain the answer is to write the total energy as $E_4 = U_4 + K_4$, such that $K_4 = E_4 - U_4$. Use the answers from Parts D and H ($E_4 = -0.850$ eV and $U_4 = -1.70$ eV):

$$K_4 = E_4 - U_4 = -0.850 - (-1.70) = -0.850 + 1.70 = 0.850 \text{ eV}$$

This differs slightly from our previous answer due to round-off error. If you keep additional digits throughout your calculation (which is the proper way to do it), it will minimize your round-off error (but remember to round your *final* answer to the appropriate number of significant figures).

(K) Recall that period (T) is defined as the time that it takes for the electron to complete one revolution. Since the electron travels with constant speed, its speed equals the distance that it travels in one revolution divided by the period. In one revolution, the distance traveled is the circumference: $C_4 = 2\pi r_4$.

$$v_4 = \frac{C_4}{T_4} = \frac{2\pi r_4}{T_4}$$

Multiply both sides of the equation by the period.

$$v_4 T_4 = 2\pi r_4$$

Divide both sides of the equation by the speed.

$$T_4 = \frac{2\pi r_4}{v_4}$$

Recall from Part A that $r_4 = 8.46$ Å $= 8.46 \times 10^{-10}$ m and Part C that $v_4 = 5.48 \times 10^5$ m/s.

$$T_4 = \frac{2\pi(8.46 \times 10^{-10})}{5.48 \times 10^5} = \boxed{9.70 \times 10^{-15}\text{ s}}$$

(L) Recall that frequency is the reciprocal of the period.

$$f_4 = \frac{1}{T_4} = \frac{1}{9.72 \times 10^{-15}} = \boxed{1.03 \times 10^{14}\text{ Hz}}$$

2. The Rydberg constant equals $R_H = 1.10 \times 10^7 \text{m}^{-1}$. For hydrogen, $Z = 1$.

(A) Use $n_i = 3$ and $n_f = 2$.

$$\frac{1}{\lambda} = \left|-Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)\right| = \left|-(1)^2 R_H \left(\frac{1}{3^2} - \frac{1}{2^2}\right)\right| = \left|-R_H \left(\frac{1}{9} - \frac{1}{4}\right)\right|$$

Make a common denominator.

$$\frac{1}{\lambda} = \left|-R_H \left(\frac{4}{36} - \frac{9}{36}\right)\right| = \left|-R_H \left(-\frac{5}{36}\right)\right| = \frac{5}{36} R_H$$

Take the reciprocal of both sides.

$$\lambda = \frac{36}{5R_H} = \frac{36}{5(1.10 \times 10^7)} = \boxed{6.55 \times 10^{-7} \text{ m}} = \boxed{655 \text{ nm}}$$

Since the electron dropped down from a higher level ($n_i = 3$) to a lower level ($n_f = 2$), the photon is $\boxed{\text{emitted}}$.

(B) If you swap the values of n_i and n_f, you will still obtain $\lambda = \boxed{655 \text{ nm}}$ just like we did in Part A. The only difference is that the photon must now be $\boxed{\text{absorbed}}$ since the electron jumped up from a lower level ($n_i = 2$) to a higher level ($n_f = 3$).

(C) Use $n_i = 4$ and $n_f = 7$.

$$\frac{1}{\lambda} = \left|-Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)\right| = \left|-(1)^2 R_H \left(\frac{1}{4^2} - \frac{1}{7^2}\right)\right| = \left|-R_H \left(\frac{1}{16} - \frac{1}{49}\right)\right|$$

$$\frac{1}{\lambda} = \left|-R_H \left(\frac{49}{784} - \frac{16}{784}\right)\right| = \left|-R_H \left(\frac{33}{784}\right)\right| = \left|-\frac{33}{784}\right| R_H$$

$$\lambda = \frac{784}{33R_H} = \frac{784}{33(1.10 \times 10^7)} = \boxed{2.16 \times 10^{-6}\text{ m}} = \boxed{2160\text{ nm}}$$

Since the electron jumped up from a lower level ($n_i = 4$) to a higher level ($n_f = 7$), the photon is $\boxed{\text{absorbed}}$.

(D) Use $n_i = 5$ and $n_f = 1$.

$$\frac{1}{\lambda} = \left|-Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)\right| = \left|-(1)^2 R_H \left(\frac{1}{5^2} - \frac{1}{1^2}\right)\right| = \left|-R_H \left(\frac{1}{25} - \frac{1}{1}\right)\right|$$

$$\frac{1}{\lambda} = \left|-R_H \left(\frac{1}{25} - \frac{25}{25}\right)\right| = \left|-R_H \left(-\frac{24}{25}\right)\right| = \frac{24}{25} R_H$$

$$\lambda = \frac{25}{24R_H} = \frac{25}{24(1.10 \times 10^7)} = \boxed{9.47 \times 10^{-8}\text{ m}} = \boxed{94.7\text{ nm}}$$

Since the electron dropped down from a higher level ($n_i = 5$) to a lower level ($n_f = 1$), the photon is $\boxed{\text{emitted}}$.

(E) Use $n_i = 6$ and $n_f = 2$.

$$\frac{1}{\lambda} = \left| -Z^2 R_H \left(\frac{1}{n_i^2} - \frac{1}{n_f^2} \right) \right| = \left| -(1)^2 R_H \left(\frac{1}{6^2} - \frac{1}{2^2} \right) \right| = \left| -R_H \left(\frac{1}{36} - \frac{1}{4} \right) \right|$$

$$\frac{1}{\lambda} = \left| -R_H \left(\frac{1}{36} - \frac{9}{36} \right) \right| = \left| -R_H \left(-\frac{8}{36} \right) \right| = \frac{2}{9} R_H$$

$$\lambda = \frac{9}{2R_H} = \frac{9}{2(1.10 \times 10^7)} = \boxed{4.09 \times 10^{-7} \text{ m}} = \boxed{409 \text{ nm}}$$

Since the electron dropped down from a higher level ($n_i = 6$) to a lower level ($n_f = 2$), the photon is $\boxed{\text{emitted}}$.

10 THE DE BROGLIE RELATION

Relevant Terminology

Photon – a particle of light.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Velocity – a combination of speed and direction.

Momentum – mass times velocity.

Wave-Particle Duality of Light

Is light **wave**-like or **particle**-like in nature? The answer is **both**!

- Young's double-slit experiment demonstrates convincingly that light exhibits **wave-like** properties. The interference pattern that appears on the screen is characteristic of waves (like water waves and sound waves) with a definite wavelength.

- The photoelectric effect demonstrates that light interacts with matter with **particle-like** properties. In this case light behaves like a beam of photons, where each photon has energy hf and interacts with an atom on a one-to-one basis.

Light has a **dual nature**: It exhibits both wave-like and particle-like properties. The wave-like and particle-like properties complement one another. Whether the wave-like nature or the particle-like nature of light (or some extent of each) is revealed depends on the nature of the experiment being conducted.

- If you shine a monochromatic beam of light through a narrow slit, the beam of light reveals its **wave**-like nature, forming a diffraction pattern on a screen.

- If you shine a beam of ultraviolet light on a metal and measure the cutoff frequency for producing photoelectrons, the incident light reveals itself as a beam of **photons**, interacting with atoms on a one-to-one basis.

- If you direct a beam of x-rays on a sample and the x-rays interact with electrons via Compton scattering, the detectors measure the incident and scattered wavelengths of the x-rays through light's **wave**-like properties, while the explanation for how the scattered wavelengths are related to the incident wavelength involves light's **particle**-like properties (as the photons interacted with electrons in atoms).

Wave-Particle Duality of Matter

Since light—which was previously believed to be purely wave-like—turns out to also exhibit particle-like properties, you might wonder if **matter**—which consists of particles like protons and electrons—also **exhibits wave-like behavior**. The answer is **yes**. If you direct a beam of **electrons** through a double-slit, they form an **interference** pattern on a screen.

The de Broglie Relation

If particles of **matter**—like electrons—exhibit **wave**-like behavior, what is their wavelength? The answer to this question is related to another question: If particles of light—like photons—exhibit particle-like behavior, what is their momentum?

We will first address the second question: What is the momentum of a photon? The energy and momentum of a particle are related by $E^2 = p^2c^2 + m_0^2c^4$ (see Chapter 5). Since a photon has zero rest mass, the energy and **momentum** of a photon are related by

$$E_\gamma = p_\gamma c$$

In Chapter 7 (on the photoelectric effect), we learned that a photon's **energy** is proportional its frequency, where $h = 6.626 \times 10^{-34}$ J·s is **Planck's constant.**

$$E_\gamma = hf$$

Set the two equations for a photon's energy equal to one another in order to relate a photon's momentum to its **frequency**.

$$p_\gamma c = hf \quad \rightarrow \quad p_\gamma = \frac{hf}{c}$$

A photon's **wavelength** and frequency are related by $\lambda f = c$. Divide both sides by c to get $\frac{\lambda f}{c} = 1$ and then divide both sides by λ to see that $\frac{f}{c} = \frac{1}{\lambda}$. Replace $\frac{f}{c}$ by $\frac{1}{\lambda}$ in the equation above.

$$p_\gamma = \frac{h}{\lambda}$$

This shows that the momentum of a photon is inversely proportional to its wavelength.

Now we will address the first question: What is the wavelength of a particle of matter, like an electron? According to de Broglie, the same equation relates the momentum and wavelength of all particles, whether they are particles of matter or particles of radiation.

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

For a particle traveling with a speed in the neighborhood of the speed of light, we would write the momentum as $p = \gamma m_0 v = \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}$ (see Chapter 5, where we used u instead of v).

For a particle that travels much slower than the speed of light, $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \approx 1$, such that the equation for momentum is approximately the equation from classical physics: $p = m_0 v$.

$$m_0 v = \frac{h}{\lambda}$$

The wavelength of a particle of matter is inversely proportional to its speed:

$$\lambda = \frac{h}{m_0 v}$$

As we will explore in later chapters, matter is a **probability** wave.

Symbols and SI Units

Symbol	Name	SI Units
h	Planck's constant	J·s or J/Hz
f	frequency	Hz
λ	wavelength	m
v	speed	m/s
c	the speed of light in vacuum	m/s
m_0	rest mass	kg
p	momentum	$\frac{\text{kg}\cdot\text{m}}{\text{s}}$
p_γ	momentum of a photon	$\frac{\text{kg}\cdot\text{m}}{\text{s}}$

E	energy	J
E_γ	energy of a photon	J
γ	time dilation factor (or a photon)	unitless

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s

Electron Volts and Angstroms

The conversion from electron Volts (eV) to Joules (J) is:

$$1 \text{ eV} = 1.6021766 \times 10^{-19} \text{ J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18} \text{ eV}$$

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Notes Regarding Units

The SI units of momentum (p) follow from the equation $\vec{\mathbf{p}} = m\vec{v}$: The SI units of momentum equal $\frac{\text{kg}\cdot\text{m}}{\text{s}}$. For elementary particles, SI units aren't ideal. It's common to instead express the units of momentum as electron Volts divided by the speed of light in vacuum: $\frac{\text{eV}}{c}$. Since an electron Volt is related to a Joule by $1\text{ eV} = 1.6021766 \times 10^{-19}\text{ J}$ and the speed of light in vacuum is $c = 2.9979 \times 10^{8}\text{ m/s}$, the two common units of momentum are related by:

$$1\,\frac{\text{eV}}{c} = \frac{1.6021766 \times 10^{-19}\text{ J}}{2.9979 \times 10^{8}\text{ m/s}} = 5.3443 \times 10^{-28}\,\frac{\text{kg}\cdot\text{m}}{\text{s}}$$

(Recall from the equation for kinetic energy, $K = \frac{1}{2}mv^2$, that a Joule is equal to $1\text{ J} = \text{kg}\frac{\text{m}^2}{\text{s}^2}$.)

Strategy for Problems Involving Wave-Particle Duality

If a problem involves the dual nature of radiation (like light) or the dual nature of matter (like a proton or electron), follow these steps:

- Apply the **de Broglie relation** to relate the particle's wavelength to its momentum.

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

  Note that $h = 6.626 \times 10^{-34}$ J·s is Planck's constant.

- For a particle of matter, the **momentum** equals (see Chapter 5):

$$p = \gamma m_0 v = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \quad , \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

  However, if the particle travels slow compared to the speed of light, then

$$p \approx m_0 v$$

- For a **photon**, the energy equals

$$E_\gamma = hf$$

- For all particles, **energy** and momentum are related by

$$E^2 = p^2 c^2 + m_0^2 c^4$$

  The above equation simplifies for a **photon**, since it has zero rest mass:

$$E_\gamma = p_\gamma c$$

- Recall that **kinetic energy** equals $K = E - m_0c^2$ (Chapter 5). However, if the particle travels slow compared to the speed of light, then $K \approx \frac{1}{2}m_0v^2$.
- To find the de Broglie **frequency**, see the note in Problem 2.
- For conceptual questions, note that diffraction and interference are standard wave phenomena, whereas the one-to-one interaction between a photon and an electron is a particle-like interaction.
- If a problem involves Heisenberg's uncertainty principle, see Chapter 11.

Example: Find the momentum and wavelength for each of the following particles. The rest mass of an electron is 9.109×10^{-31} kg.

(A) An electron travels 20 km/s.

Since 20 km/s = 20,000 m/s is small compared to the speed of light:

$$p \approx m_0 v = (9.109 \times 10^{-31})(20{,}000) = \boxed{1.8 \times 10^{-26} \frac{\text{kg}\cdot\text{m}}{\text{s}}}$$

$$p = \frac{h}{\lambda} \quad \rightarrow \quad \lambda = \frac{h}{p} \approx \frac{6.626 \times 10^{-34}}{1.8 \times 10^{-26}} = \boxed{3.7 \times 10^{-8}\text{ m}} = \boxed{37\text{ nm}}$$

(B) An electron travels $0.6c$.

Since $0.6c$ is close to the speed of light:

$$p = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{(9.109 \times 10^{-31})(0.6 \times 2.9979 \times 10^8)}{\sqrt{1 - (0.6)^2}} = \boxed{2.0 \times 10^{-22} \frac{\text{kg}\cdot\text{m}}{\text{s}}}$$

$$\lambda = \frac{h}{p} \approx \frac{6.626 \times 10^{-34}}{2.0 \times 10^{-22}} = \boxed{3.3 \times 10^{-12}\text{ m}} = \boxed{0.033\text{ Å}}$$

(C) A photon has a frequency of 4.0×10^{16} Hz.

Since a photon has zero rest mass:

$$p_\gamma = \frac{E_\gamma}{c} = \frac{hf}{c} = \frac{(6.626 \times 10^{-34})(4.0 \times 10^{16})}{2.9979 \times 10^8} = \boxed{8.8 \times 10^{-26} \frac{\text{kg}\cdot\text{m}}{\text{s}}}$$

$$\lambda = \frac{h}{p} \approx \frac{6.626 \times 10^{-34}}{8.8 \times 10^{-26}} = \boxed{7.5 \times 10^{-9}\text{ m}} = \boxed{7.5\text{ nm}}$$

Example: For Compton scattering (Chapter 8), $\gamma + e^- \to \gamma + e^-$, the wavelengths of the initial and scattered photons are related by $\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$. The quantity $\frac{h}{m_e c}$ is called the Compton wavelength: $\lambda_C = \frac{h}{m_e c}$. At what speed would an electron need to travel in order for its de Broglie wavelength to equal its Compton wavelength?

Set the electron's Compton wavelength equal to its de Broglie wavelength: $\lambda = \lambda_C$.

$$\frac{h}{p} = \frac{h}{m_e c}$$

Cross multiply. Planck's constant cancels and we get $m_e c = p$. Since the speed of the electron is near the speed of light, we must be careful to use the relativistic equation for momentum.

$$m_e c = \frac{m_e v}{\sqrt{1 - \frac{v^2}{c^2}}}$$

The mass cancels. Multiply by $\sqrt{1 - \frac{v^2}{c^2}}$ and divide by c to get $\sqrt{1 - \frac{v^2}{c^2}} = \frac{v}{c}$. Square both sides to get $1 - \frac{v^2}{c^2} = \frac{v^2}{c^2}$. Add $\frac{v^2}{c^2}$ to both sides: $1 = 2\frac{v^2}{c^2}$. Solve this to get $v = \boxed{\frac{c}{\sqrt{2}}} \approx \boxed{0.71c}$.

Chapter 10 Problems

1. Find the momentum and wavelength for each of the following particles. The rest mass of an electron is 9.109×10^{-31} kg.

(A) An electron has a kinetic energy of 100 eV.

(B) An electron has a kinetic energy of 100 keV.

(C) A photon has a kinetic energy of 100 keV.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) 10.0 keV/c, 1.23 Angstroms

(B) 335 keV/c, 0.0370 Angstroms

(C) 100 keV/c, 0.124 Angstroms

2. According to de Broglie, particles of matter have a frequency equal to:

$$f = \frac{E}{h}$$

Show that the frequency of a particle of matter is related to its de Broglie wavelength by:

$$f = \sqrt{\frac{c^2}{\lambda^2} + \frac{m_0^2 c^4}{h^2}}$$

Want help? Check the solution at the end of the chapter.

Solutions to Chapter 10

1. Note that $m_e = 9.109 \times 10^{-31}$ kg, $c = 2.9979 \times 10^{8}$ m/s, and $h = 6.626 \times 10^{-34}$ J·s. The conversion from electron Volts (eV) to Joules (J) is 1 eV = $1.6021766 \times 10^{-19}$ J.

(A) Do you know whether or not the electron's speed is close to the speed of light? If not, the prudent thing to do is to use relativistic equations for momentum and energy (Chapter 5). At the end of our solution, we will compute the speed (and then we will see whether or not we needed to use relativistic equations in our solution). First find the electron's total energy.

$$E = K + m_e c^2 = 100 \times 1.6021766 \times 10^{-19} + (9.109 \times 10^{-31})(2.9979 \times 10^{8})^2$$

$$E = 8.1882 \times 10^{-14} \text{ J}$$

We kept a couple of extra digits in order to help reduce round-off error in our calculations. (It's a good habit to avoid rounding until you reach a *final* answer.) Now we can find the electron's momentum.

$$E^2 = p^2 c^2 + m_e^2 c^4$$

Bring $m_e^2 c^4$ to the other side to get $E^2 - m_e^2 c^4 = p^2 c^2$. Then divide both sides by c^2.

$$\frac{E^2}{c^2} - m_e^2 c^2 = p^2$$

$$p = \sqrt{\frac{E^2}{c^2} - m_e^2 c^2} = \sqrt{\frac{(8.1882 \times 10^{-14})^2}{(2.9979 \times 10^{8})^2} - (9.109 \times 10^{-31})^2 (2.9979 \times 10^{8})^2}$$

This calculation is quite sensitive to round-off error. If you round 8.1882 to 8.19, you get a value of 7.84 instead of 5.35. This is because the two terms are very nearly equal.

$$p = \boxed{5.35 \times 10^{-24} \frac{\text{kg·m}}{\text{s}}}$$

Convert your answer to units of $\frac{\text{eV}}{c}$ by multiplying by $c = 2.9979 \times 10^8$ m/s and dividing by $e = 1.6021766 \times 10^{-19}$ C (see page 152).

$$p = 5.35 \times 10^{-24} \frac{2.9979 \times 10^8}{1.6021766 \times 10^{-19}} = \boxed{1.00 \times 10^4 \frac{\text{eV}}{c}} = \boxed{10.0 \frac{\text{keV}}{c}}$$

Use the de Broglie relation in order to find the wavelength. Multiply $p = \frac{h}{\lambda}$ by λ to get $p\lambda = h$ and divide by p to get $\lambda = \frac{h}{p}$. Be sure to use matching units.

$$\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34}}{5.35 \times 10^{-24}} = \boxed{1.23 \times 10^{-10} \text{ m}} = \boxed{1.23 \text{ Å}}$$

Note: In Part A, if you use $K \approx \frac{1}{2} m_e v^2$, you get $v \approx \sqrt{\frac{2K}{m_e}} = \sqrt{\frac{(2)(100 \times 1.6021766 \times 10^{-19})}{9.109 \times 10^{-31}}} = 5.93 \times 10^6$ m/s, which is 2% of the speed of light. The classical formula for momentum would give $p \approx m_e v = 5.40 \times 10^{-24} \frac{\text{kg·m}}{\text{s}}$, which is close to the actual value of $p = 5.35 \times 10^{-24} \frac{\text{kg·m}}{\text{s}}$.

(B) In Part B, the velocity turns out to be relativistic, so it would be **incorrect** to use the formula $K \approx \frac{1}{2} m_e v^2$. We will use the same equations from Part A. Note that 100 keV is 1000 times greater than 100 eV.

$$E = K + m_e c^2 = 100{,}000 \times 1.6021766 \times 10^{-19} + (9.109 \times 10^{-31})(2.9979 \times 10^8)^2$$

$$E = 9.7888 \times 10^{-14}\ \text{J}$$

$$p = \sqrt{\frac{E^2}{c^2} - m_e^2 c^2} = \sqrt{\frac{(9.7888 \times 10^{-14})^2}{(2.9979 \times 10^8)^2} - (9.109 \times 10^{-31})^2 (2.9979 \times 10^8)^2}$$

$$p = \boxed{1.79 \times 10^{-22}\ \frac{\text{kg·m}}{\text{s}}}$$

$$p = 1.79 \times 10^{-22} \frac{2.9979 \times 10^8}{1.6021766 \times 10^{-19}} = \boxed{3.35 \times 10^5\ \frac{\text{eV}}{c}} = \boxed{335\ \frac{\text{keV}}{c}}$$

$$\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34}}{1.79 \times 10^{-22}} = \boxed{3.70 \times 10^{-12}\ \text{m}} = \boxed{0.0370\ \text{Å}}$$

In Part B, the speed of the electron is $v = \frac{pc^2}{E} = \frac{pc}{E} c = \frac{(1.79 \times 10^{-22})(2.9979 \times 10^8)}{9.7888 \times 10^{-14}} c = 0.548c$, or 55% of the speed of light in vacuum. Since the speed of the electron is in the neighborhood of the speed of light, the classical equations $p_{class} = mv$ and $K_{class} = \frac{1}{2} mv^2$ would result in significantly incorrect answers. (To see why $v = \frac{pc^2}{E}$, see Part D on page 75 in Chapter 5.)

(C) The photon is massless. Set the rest mass equal to zero in the equations from Part B.

$$E_\gamma = K_\gamma + 0 = 100{,}000 \times 1.6021766 \times 10^{-19}$$

$$E = 1.60 \times 10^{-14}\ \text{J}$$

$$p_\gamma = \frac{E_\gamma}{c} = \frac{1.60 \times 10^{-14}}{2.9979 \times 10^{8}} = \boxed{5.34 \times 10^{-23}\ \frac{\text{kg}\cdot\text{m}}{\text{s}}}$$

$$p_\gamma = 5.34 \times 10^{-23}\ \frac{2.9979 \times 10^{8}}{1.6021766 \times 10^{-19}} = \boxed{9.99 \times 10^{4}\ \frac{\text{eV}}{c}} = \boxed{99.9\ \frac{\text{keV}}{c}}$$

Note that the correct answer is actually $\boxed{100\ \frac{\text{keV}}{c}}$, since the kinetic energy was specified as 100 keV in the problem. When we converted from eV to $\frac{\text{kg}\cdot\text{m}}{\text{s}}$ and then back to $\frac{\text{eV}}{c}$, the value of 100 changed to 99.9 due to round-off error.

$$\lambda = \frac{h}{p_\gamma} = \frac{6.626 \times 10^{-34}}{5.34 \times 10^{-23}} = \boxed{1.24 \times 10^{-11}\ \text{m}} = \boxed{0.124\ \text{Å}}$$

2. The energy and momentum of a particle are related by $E^2 = p^2c^2 + m_0^2c^4$. Substitute this expression for energy into the given equation for frequency.

$$f = \frac{E}{h} = \frac{\sqrt{p^2c^2 + m_0^2c^4}}{h} = \frac{\sqrt{p^2c^2 + m_0^2c^4}}{\sqrt{h^2}} = \sqrt{\frac{p^2c^2 + m_0^2c^4}{h^2}} = \sqrt{\frac{p^2c^2}{h^2} + \frac{m_0^2c^4}{h^2}}$$

Apply the de Broglie relation: $p = \frac{h}{\lambda}$.

$$f = \sqrt{\frac{h^2c^2}{\lambda^2h^2} + \frac{m_0^2c^4}{h^2}} = \boxed{\sqrt{\frac{c^2}{\lambda^2} + \frac{m_0^2c^4}{h^2}}}$$

11 HEISENBERG’S UNCERTAINTY PRINCIPLE

Relevant Terminology

Photon – a particle of light.

Wavelength – the horizontal distance between two consecutive crests in a wave.

Frequency – the number of oscillations completed per second.

Velocity – a combination of speed and direction.

Momentum – mass times velocity.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Uncertainty in Measurements

Even in classical physics, there is inherent uncertainty in any measurement. However, if you have an unlimited budget, ample time, a well-trained research team, ideal facilities, and the best instrumentation that you can imagine, in classical physics there is no limit, in principle, to how precisely you can make your measurements.

In quantum physics, it turns out that there is a limit to how precisely measurements can be made. Specifically, **the more precisely you measure position, the more uncertainty there is in momentum, and vice-versa**. (We will explore the uncertainty principle in the next section.)

Consider some of the standard problems from an introductory, first-semester physics class. For example, suppose that a ball is thrown from a height of 20 m above level ground at an angle of 30° above the horizontal with an initial speed of 40 m/s. In classical physics, we use this given information to predict where the ball will land on the ground and how fast it will be moving just before impact.

In quantum mechanics, not only can't you know both exactly where the ball will land and how fast it will be moving just before impact—but **you can't even ask the question**!

In quantum mechanics, you can't know both exactly where an object is and exactly what the object's momentum is in the beginning of the problem, let alone predict what the position and momentum will be at some later time.

Instead, you would have to ask a question like this in quantum mechanics. There is an electron somewhere in the region $a < x_0 < b$ with an initial speed in the range $c < v_{x0} < d$. What is the probability that the electron will be in the region $e < x < f$ some specified time later?

Heisenberg's Uncertainty Principle

If you simultaneously measure the x-coordinate of an object's **position** and the x-component of the object's **momentum**, the corresponding **uncertainties** (Δx and Δp_x) must satisfy:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

There are corresponding inequalities regarding the y- and z-components. Recall that $\hbar = \frac{h}{2\pi}$, where $h = 6.626 \times 10^{-34}$ J·s is **Planck's constant**. This shows that it is *impossible* to know an object's *exact* position and *exact* momentum at the *same* moment. The more precisely you know one, the more uncertainty there is in the other.

Why Does $\Delta x \Delta p_x \geq \frac{\hbar}{2}$?

Imagine that you have a single elementary particle, like the electron. How would you "see" this electron in order to make a measurement? One way to go about making a very precise measurement of the electron's position and momentum may involve having a single photon interact with the electron. The photon's momentum is $p_\gamma = \frac{h}{\lambda}$ (Chapter 10). The uncertainty in the electron's momentum will be at least this much: $\Delta p_x \geq \frac{h}{\lambda}$, where λ is the wavelength of the photon (since the interaction, or collision, involving the electron and photon involves an exchange of momentum). The uncertainty in the electron's position will be at least $\Delta x \geq \lambda$ based on the wave-like nature of both the photon and the electron (Chapter 10). If we put these together, we get $\Delta x \Delta p_x \geq (\lambda)\left(\frac{h}{\lambda}\right)$, which simplifies to $\Delta x \Delta p_x \geq h$. More careful analysis shows that $\Delta x \Delta p_x \geq \frac{\hbar}{2}$.

The Energy-Time Uncertainty Principle

An alternative form of Heisenberg's uncertainty principle involves the **energy** (E) of a system and the **time interval** over which the energy is measured (Δt).

$$\Delta E \Delta t \geq \frac{\hbar}{2}$$

Regarding why this is, one way to is recall that all particles exhibit wave-like behavior with $E = hf$ (see Chapter 10, Problem 2), such that the uncertainty in energy is $\Delta E \geq h\Delta f$ and that the uncertainty in time is $\Delta t \geq \frac{1}{\Delta f}$, for which $\Delta E \Delta t \geq (h\Delta f)\left(\frac{1}{\Delta f}\right) = h$. Alternatively, the energy of a free particle in 1D is $E = \frac{p_x^2}{2m}$, such that $\Delta E \approx \frac{p_x}{m}\Delta p_x = v_x \Delta p_x$ (since $p_x = mv_x$), and the uncertainty in position is $\Delta x = v_x \Delta t$, such that $\Delta E \Delta t = (v_x \Delta p_x)\left(\frac{\Delta x}{v_x}\right) = \Delta p_x \Delta x$. Note that $dE = \frac{d}{dp_x}\left(\frac{p_x^2}{2m}\right) dp_x = \frac{2p_x}{2m} dp_x = \frac{p_x}{m} dp_x$ (an implicit derivative), and $dE \approx \Delta E$.

Symbols and SI Units

Symbol	Name	SI Units
h	Planck's constant	J·s or J/Hz
$\hbar$	h-bar	J·s or J/Hz
Δx	the uncertainty in the x coordinate of position	m
p_x	the x-component of momentum	$\frac{\text{kg·m}}{\text{s}}$
Δp_x	the uncertainty in p_x	$\frac{\text{kg·m}}{\text{s}}$
E	energy of a system	J
ΔE	the uncertainty in E	J
Δt	the time interval over which the energy is measured	s

λ	wavelength	m
f	frequency	Hz
m	mass	kg
v_x	the x-component of velocity	m/s

Constants

Quantity	Value
speed of light in vacuum	$c = 2.9979 \times 10^8$ m/s
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s

Electron Volts

The conversion from electron Volts (eV) to Joules (J) is:

$$1 \text{ eV} = 1.6021766 \times 10^{-19} \text{ J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18} \text{ eV}$$

Strategy for Problems Involving Heisenberg's Uncertainty Principle

If a problem involves Heisenberg's uncertainty principle, follow these steps:

- The **uncertainty** in a **position** coordinate (such as Δx) is related to the **uncertainty** in the corresponding component of **momentum** (such as Δp_x) by:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

where $\hbar = \frac{h}{2\pi}$ and $h = 6.626 \times 10^{-34}$ J·s.

- Alternatively, some problems involve the uncertainty in the **energy** of a system (ΔE) and the **time interval** over which this measurement is made (Δt).

$$\Delta E \Delta t \geq \frac{\hbar}{2}$$

- Ensure that the units match. For example, if ΔE is given in electron Volts (eV), first convert it to Joules if you are using $h = 6.626 \times 10^{-34}$ J·s.
- It may help to recall the **de Broglie relation** (Chapter 10).

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

Note that the de Broglie relation involves Planck's constant (h), whereas Heisenberg's uncertainty principle involves h-bar ($\hbar = \frac{h}{2\pi}$).

- **Energy** is related to **frequency** by (see Chapter 10, Problem 2):

$$E = hf$$

- You may need to apply equations from a first-year physics course. For example:
  - $p_x = mv_x$ is the equation for the x-component of momentum (if nonrelativistic).
  - $K = \frac{1}{2}mv^2 = \frac{p^2}{2m}$ is the equation for kinetic energy (if nonrelativistic).
  - $\Delta x = v_x t$ if the x-component of velocity is constant.
  - $\Delta y = v_{y0}t - \frac{1}{2}gt^2$ for an object in free fall in a uniform gravitational field.
  - $E = U + K$ relates total energy, potential energy, and kinetic energy.
- For an object that travels close to the speed of light, $E^2 = p^2c^2 + m_0c^4$, $p = \gamma m_0 v$, $K = E - m_0c^2$, and $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$.
- For a **photon**, $E_\gamma = p_\gamma c = hf$, where $c = 2.9979 \times 10^8$ m/s.
- If a problem may involve Schrödinger's equation or expectation values, see Chapters 13-16.
- If a problem involves eigenvalues, eigenvectors, or matrices, that is beyond the scope of this book. Seek a book on quantum mechanics that involves linear algebra.

Example: An experiment is designed to simultaneously measure the position and speed of an electron. The experiment determines that the electron lies in a volume with a diameter of one Angstrom. The mass of an electron is 9.109×10^{-31} kg. What is the minimum possible uncertainty in the electron's speed? Assume that the electron's speed is nonrelativistic.

According to Heisenberg's uncertainty principle, $\Delta x \Delta p_x \geq \frac{\hbar}{2}$. Since $\Delta p_x = m\Delta v_x$, we may write this as $\Delta x(m\Delta v_x) \geq \frac{\hbar}{2}$ or $\Delta x \Delta v_x \geq \frac{\hbar}{2m}$. Recall that $1\ \text{Å} = 10^{-10}$ m and $\hbar = \frac{h}{2\pi}$.

$$\Delta v_x \geq \frac{\hbar}{2m\Delta x} = \frac{h}{4\pi m\Delta x} = \frac{6.626 \times 10^{-34}}{4\pi(9.109 \times 10^{-31})(10^{-10})} = \boxed{5.79 \times 10^5 \text{ m/s}} = \boxed{579 \text{ km/s}}$$

Example: A tiny steel ball is released from rest from a height H above horizontal ground on a planet which, unlike earth, doesn't rotate on its axis (so there is no Coriolis force). Due to Heisenberg's uncertainty principle, the ball won't land on the exact point directly under its initial position. Show that the ball is expected to miss by a distance of:

$$\Delta x \geq \left(\frac{\hbar}{m}\right)^{1/2} \left(\frac{H}{2g}\right)^{1/4}$$

Recall from first-year physics that $H = \frac{1}{2}gt^2$ for a ball dropped from rest. Solve for time to get $t = \sqrt{\frac{2H}{g}}$. Setup a coordinate system with the origin on the ground directly underneath the ball. In classical physics, x and p_x both equal zero, but according to quantum physics, the ball is expected to gain momentum of at least Δp_x and stray off course by at least Δx. The ball doesn't fall straight down, but instead follows the path of a projectile, where $\Delta x = v_x t$ for the horizontal component of the motion. Momentum is mass times velocity: $\Delta p_x = mv_x = m\frac{\Delta x}{t}$. Plug $t = \sqrt{\frac{2H}{g}}$ into $\Delta p_x = m\frac{\Delta x}{t}$ to get $\Delta p_x = m\Delta x\sqrt{\frac{g}{2H}}$. According to Heisenberg's uncertainty principle, $\Delta x \Delta p_x \geq \frac{\hbar}{2}$. Substitute $\Delta p_x = m\Delta x\sqrt{\frac{g}{2H}}$ into this inequality.

$$\Delta x \left(m\Delta x \sqrt{\frac{g}{2H}} \right) \geq \frac{\hbar}{2}$$

$$\Delta x^2 \geq \frac{\hbar}{2m}\sqrt{\frac{2H}{g}} = \frac{\hbar}{m}\sqrt{\frac{H}{2g}}$$

Note that $\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$ since $\sqrt{2}\sqrt{2} = 2$. Squareroot both sides of the previous equation.

$$\Delta x \geq \boxed{\frac{\hbar^{1/2}H^{1/4}}{2^{1/4}m^{1/2}g^{1/4}}} = \boxed{\left(\frac{\hbar}{m}\right)^{1/2}\left(\frac{H}{2g}\right)^{1/4}}$$

Chapter 11 Problems

1. An experiment is designed to simultaneously measure the position and speed of an electron. The experiment determines that the electron has a speed of 25 km/s with an accuracy of 0.2%. The mass of an electron is 9.109×10^{-31} kg. What is the minimum possible uncertainty in the electron's position?

Want help? Check the solution at the end of the chapter.

Answers: (1) 1.2 μm

2. An experiment is designed to simultaneously measure the position and momentum of a particle. The experiment determines that the uncertainty in the x-coordinate of the particle equals $\Delta x = \frac{a}{2\pi}$, where a is a constant. Derive an equation for the minimum uncertainty in the x-component of the particle's momentum.

Want help? Check the solution at the end of the chapter.

Answers: (2) $\Delta p_x^{min} = \frac{\pi\hbar}{a}$

Solutions to Chapter 11

1. According to Heisenberg's uncertainty principle, $\Delta x \Delta p_x \geq \frac{\hbar}{2}$. Since $\Delta p_x = m\Delta v_x$, we may write this as $\Delta x(m\Delta v_x) \geq \frac{\hbar}{2}$ or $\Delta x \Delta v_x \geq \frac{\hbar}{2m}$. Recall that $h = 6.626 \times 10^{-34}$ J·s and $\hbar = \frac{h}{2\pi}$. The speed of the electron is $v_x = 25$ km/s = 25,000 m/s (since the prefix kilo stands for 1000). The uncertainty in the electron's speed is 0.2% of this (where $0.2\% = \frac{0.2}{100} = 0.002$).

$$\Delta v_x = 0.2\% v_x = 0.002 v_x = (0.002)(25{,}000) = 50 \text{ m/s}$$

Plug the uncertainty in the speed into Heisenberg's uncertainty principle.

$$\Delta x \geq \frac{\hbar}{2m\Delta v_x} = \frac{h}{4\pi m \Delta v_x} = \frac{6.626 \times 10^{-34}}{4\pi(9.109 \times 10^{-31})(50)} = \boxed{1.2 \times 10^{-6} \text{ m/s}} = \boxed{1.2 \text{ μm/s}}$$

Recall that the prefix micro (μ) stands for one millionth: μ = 10^{-6}.

---

2. Substitute the given equation, $\Delta x = \frac{a}{2\pi}$, into Heisenberg's uncertainty principle.

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \quad \to \quad \left(\frac{a}{2\pi}\right)\Delta p_x \geq \frac{\hbar}{2} \quad \to \quad \Delta p_x \geq \frac{\hbar}{2}\frac{2\pi}{a} \quad \to \quad \boxed{\Delta p_x \geq \frac{\pi\hbar}{a}}$$

12 DIFFERENTIAL EQUATIONS

Differential Equations in Physics

For almost every topic in physics, differential equations come about very naturally. In fact, did you know that **Newton's second law** is actually a second-order differential equation? When Newton's second law is written $\vec{\mathbf{F}}_{net} = m\vec{\mathbf{a}}$, it might not seem like a differential equation, but when you recall that acceleration is a derivative of velocity with respect to time, $\vec{\mathbf{a}} = \frac{d\vec{\mathbf{v}}}{dt}$, and the second derivative of position with respect to time, $\vec{\mathbf{a}} = \frac{d^2\vec{\mathbf{r}}}{dt^2}$, this becomes clear:

$$\vec{\mathbf{F}}_{net} = m\frac{d^2\vec{\mathbf{r}}}{dt^2}$$

For example, if the net force acting on an object is given by Hooke's law and if the motion is one-dimensional, Newton's second law looks like this:

$$-kx = m\frac{d^2x}{dt^2}$$

Conservation of energy, which applies to virtually every topic in physics, is also a differential equation because speed is the magnitude of velocity and velocity is a derivative of position with respect to time, $v = \|\vec{\mathbf{v}}\| = \left\|\frac{d\vec{\mathrm{r}}}{dt}\right\|$. As an example, for a freely falling object released from rest in a uniform gravitational field, conservation of energy looks like this:

$$mgy_0 = mgy + \frac{1}{2}m\left(\frac{dy}{dt}\right)^2$$

Wave equations are inherently differential equations. The linear wave equation is a partial differential equation (where ∂ is the symbol used for partial derivatives):

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}$$

As will explore in the remaining chapters, quantum mechanics involves solving **Schrödinger's equation**, which is a second-order differential equation. For example, for a one-dimensional potential, Schrödinger's time-dependent equation is:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}$$

For some problems, it is helpful to write equations that relate differential elements, and by doing so we essentially begin the solution by writing a differential equation. As an example, a simple yet common model of atmospheric pressure begins by considering the differential pressure dP exerted on a differential layer of the atmosphere of thickness dy and mass m at height y, which leads to the following differential equation:

$$dP = -mgn_V dy$$

In this chapter, we will explore some essential differential equations, properties, and a few techniques for solving them. For further study, consult a textbook on differential equations.

First-Order, Separable Differential Equations

A differential equation is **first-order** if there are no second-order (or higher) derivatives. A first-order differential equation is **separable** if it is possible, by applying algebra, to put each variable on its own side of the equation. For example, consider the following equation.

$$\frac{dy}{dx} = xy$$

This is a first-order differential equation. Presently, the two variables (x and y) each appear on both sides of the equation. Yet, this differential equation is separable. Multiply both sides of the equation by dx (or if you want to be fancy, apply the chain rule instead) and divide both sides of the equation by y.

$$\frac{dy}{y} = x\,dx$$

Not all differential equations can be separated in this way, but when they can, the technique of separating variables is an efficient way to solve them. Once you separate the variables in a first-order differential equation, you can simply integrate both sides.

$$\int \frac{dy}{y} = \int x\,dx$$

Second-Order, Separable Differential Equations

Many differential equations in physics involve a second derivative. For example, consider the **second-order** differential equation below, where c is a constant.

$$\frac{d^2y}{dt^2} = -c$$

Note that you **can't** separate a second-order differential equation the same way that you can separate a first-order differential equation. That is, you can't multiply by dt^2 in the equation above: When the differential element is part of a second derivative, the differentials can't be split up like that. Some simpler second-order differential equations can be reduced to first-order differential equations. For example, for the second-order differential equation above, which involves y and t, we can rewrite the equation as a first-order differential equation by making the substitution $v_y = \frac{dy}{dt}$. Note that $\frac{dv_y}{dt} = \frac{d}{dt}\left(\frac{dy}{dt}\right) = \frac{d^2y}{dt^2}$.

$$\frac{dv_y}{dt} = -c$$

The differential equation is now a first-order differential equation. Separate the new variables (v_y and t) by bringing dt over to the right side (we can do this now that it is first order).

$$dv_y = -c\,dt$$

$$\int dv_y = -\int c\,dt$$

When a substitution of this form works, it is quite helpful.

Applying the Chain Rule to Differential Equations

In physics, some differential equations can be solved by applying the **chain rule**. For example, consider the second-order differential equation below.

$$\frac{d^2x}{dt^2} = x$$

First, we'll make the substitution $v_x = \frac{dx}{dt}$, such that $\frac{dv_x}{dt} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{d^2x}{dt^2}$, similar to what we did on the previous page (except that now we have x in place of y).

$$\frac{dv_x}{dt} = x$$

We have a problem: There are three variables (x, t, and v_x) instead of just two variables (like there were in the previous examples). There is no way to separate three different variables on just two sides of an equation. However, in this case, we can apply the chain rule in order to rewrite the equation in terms of just two variables. This is actually a pretty common and handy trick. Using the **chain rule**, we may write:

$$\frac{dv_x}{dt} = \frac{dv_x}{dx}\frac{dx}{dt} = v_x\frac{dv_x}{dx}$$

Note that $\frac{dx}{dt} = v_x$. The chain rule lets us replace $\frac{dv_x}{dt}$ with $v_x \frac{dv_x}{dx}$. With this substitution, our first-order differential equation transforms into:

$$v_x \frac{dv_x}{dx} = x$$

Move the dx over to the right in order to separate variables.

$$v_x dv_x = x\,dx$$

$$\int v_x\,dv_x = \int x\,dx$$

Partial Differential Equations

When a quantity is a function of multiple variables, like $f(x,t)$, the differential equation may involve **partial derivatives**, like the example below where c is a constant.

$$\frac{\partial^2 f(x,t)}{\partial x^2} = c\frac{\partial^2 f(x,t)}{\partial t^2}$$

Some simpler partial differential equations can be solved when the function turns out to be **separable**. For example, we can try writing $f(x,t) = g(x)h(t)$ for the equation above.

$$\frac{\partial^2}{\partial x^2}[g(x)h(t)] = c\frac{\partial^2}{\partial t^2}[g(x)h(t)]$$

Note that $h(t)$ doesn't depend on x, such that $\frac{\partial^2}{\partial x^2}h(t) = 0$. Similarly, $\frac{\partial^2}{\partial t^2}g(x) = 0$.

$$h(t)\frac{d^2 g(x)}{dx^2} = c\, g(x)\frac{d^2 h(t)}{dt^2}$$

Since $g(x)$ is a function of x only, the derivative is no longer a 'partial' derivative. Divide the previous equation by $g(x)h(t)$ in order to separate the variables.

$$\frac{1}{g(x)}\frac{d^2g(x)}{dx^2} = \frac{c}{h(t)}\frac{d^2h(t)}{dt^2}$$

The equation above states that a function of x equals a function of t. The only way that both sides can be equal for all possible values of x and t is if each side equals a constant. If we call that constant a, we get two new differential equations:

$$\frac{1}{g(x)}\frac{d^2g(x)}{dx^2} = a \quad , \quad \frac{c}{h(t)}\frac{d^2h(t)}{dt^2} = a$$

By assuming that the original function may be separated as $f(x,t) = g(x)h(t)$, we were able to transform the original partial differential equation into two ordinary differential equations. Not all partial differential equations can be solved this way, but many can. We didn't know if $f(x,t)$ could be separated into the product of two functions, $g(x)h(t)$, but we tried it and discovered that the differential equation could be solved this way. If it didn't work out, we would need to return to the original differential equation and try something else instead.

Boundary Conditions

As we have seen, the solution to a differential equation leads to an integral. Every indefinite integral involves a **constant of integration**. The solution to a first-order differential equation involves one constant of integration, while the solution to a second-order differential equation involves two constant of integration. These constants can be determined from the **boundary conditions** specified in the problem. (Alternatively, you can build these boundary conditions into the solution by doing definite integrals rather than indefinite integrals.) As an example, consider the simple case of **uniform acceleration**, meaning that $a_x = a$ is a constant.

$$\frac{d^2x}{dt^2} = a$$

$$\frac{dv_x}{dt} = a$$

$$dv_x = a\,dt$$

$$\int dv_x = a\int dt$$

$$v_x = at + c_1$$

$$\frac{dx}{dt} = at + c_1$$

$$dx = at\,dt + c_1 dt$$

$$\int dx = a\int t\,dt + c_1\int dt$$

$$x = \frac{1}{2}at^2 + c_1 t + c_2$$

There are two **constants of integration**, c_1 and c_2. We can determine these constant from the **boundary conditions**. In this example, if you set $t = 0$, you can see that these constants equal the initial values of the velocity and position. Of course, the notation $v_x(t)$ means that v_x is a function of time (it doesn't mean to multiply v_x by time), and similarly with $x(t)$.

$$v_x(t) = at + c_1 \quad \rightarrow \quad v_x(0) = c_1 = v_{x0}$$

$$x(t) = \frac{1}{2}at^2 + c_1 t + c_2 \quad \rightarrow \quad x(0) = c_2 = x_0$$

Note that we could have obtained the same results by performing definite integrals rather than indefinite integrals. We will assume that the motion begins at $t = 0$, so that $v_x(t) = v_{x0}$ and $x(0) = x_0$ are the initial velocity and position, respectively.

$$\int_{v_x=v_{x0}}^{v_x} dv_x = a \int_{t=0}^{t} dt$$

$$v_x - v_{x0} = at$$

$$\frac{dx}{dt} = v_{x0} + at$$

$$dx = v_{x0}dt + at\,dt$$

$$\int_{x=x_0}^{x} dx = v_{x0} \int_{t=0}^{t} dt + a \int_{t=0}^{t} t\,dt$$

$$x - x_0 = v_{x0}t + \frac{1}{2}at^2$$

Desirable Properties of Differential Equations and their Solutions

For a first-order differential equation, the function should be **smooth** and **continuous** such that its first derivative is well-defined. For a second-order differential equation, the function and its first derivative should both be **smooth** and **continuous** such that the function's second derivative is well-defined. The function should also be **single-valued**. For some applications in physics (such as Schrödinger's equation), it is also desirable for the function to be finite. The solution to a differential equation should be both **unique** and **complete**. What does this mean? Suppose that you solved a differential equation, obtaining $y(x) = \sin(x)$ as a solution. That wouldn't be a unique solution if $y(x) = \cos(x)$ also solved the same problem. Suppose that you solved a differential equation, obtaining $y = 0$ as a solution. Your solution would be incomplete if $y = x^2 + 3x$ also solved the equation: In this case, you only found the special case where x happens to be zero.

For many second-order differential equations, a complete solution has the following form:

$$y(x) = A\,f(x) + B\,g(x)$$

We will see a few common examples of this on the following page.

A Few Common Special Cases

The simplest case is when the second derivative equals a constant.

$$\frac{d^2y}{dx^2} = c \quad \to \quad y = \frac{1}{2}cx^2 + Ax + B$$

Exponential growth and **exponential decay** are characteristic of systems where the second derivative is proportional to the function. This is because $\frac{d}{dx}e^{ax} = ae^{ax}$.

$$\frac{d^2y}{dx^2} = a^2y \quad \to \quad y = Ae^{ax} + Be^{-ax}$$

Sinusoidal behavior is characteristic of systems where the second derivative is proportional to the negative of the function. This is because $\frac{d}{dx}\sin ax = a\cos ax$ and $\frac{d}{dx}\cos ax = -a\sin ax$.

$$\frac{d^2y}{dx^2} = -a^2y \quad \to \quad y = A\sin ax + B\cos ax$$

Note that the solution may alternatively be expressed as:

$$y = x_m \sin(ax + \varphi)$$

Recall that $\sin(\theta_1 + \theta_2) = \sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1$. If we let $\theta_1 = ax$ and $\theta_2 = \varphi$, then we get $x_m\sin(ax + \varphi) = x_m\sin ax\cos\varphi + x_m\sin\varphi\cos ax$. Let $A = x_m\cos\varphi$ and $B = x_m\sin\varphi$ to see that $x_m\sin(ax + \varphi) = A\sin ax + B\cos ax$. Yet another way to write the solution is:

$$y = Ce^{iax} + De^{-iax}$$

where i is an imaginary number (see the following section).

Complex Numbers

Complex numbers are common in quantum mechanics. Let's review a few useful properties.

When the **imaginary number** i is squared, the result is minus one:

$$i^2 = -1$$

A **complex number** can be expressed in terms of real and imaginary parts:

$$z = x + iy$$

To find the **complex conjugate** of a complex number, negate the imaginary part.

$$z^* = x - iy$$

A complex number times its complex conjugate yields a **real** answer.

$$zz^* = (x + iy)(x - iy) = x^2 + ixy - ixy - i^2y^2 = x^2 - (-1)y^2 = x^2 + y^2$$

Recall the following identities from complex analysis:

$$i^2 = -1 \quad , \quad i^3 = -i \quad , \quad i^4 = 1 \quad , \quad \frac{1}{i} = -i$$

$$e^{ix} = \cos x + i \sin x \quad , \quad e^{-ix} = \cos x - i \sin x$$

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad , \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

Strategy for Solving Differential Equations

To solve a differential equation, follow these steps:

- If the differential equation is **first-order** (it doesn't have higher than first derivatives), first see if the differential equation is **separable**. Try to apply algebra in order to put each variable on its own side of the equation, like we showed earlier and as you will see in the examples.

- If a first-order differential equation has **three variables** (instead of two), you may be able to apply the **chain rule** to rewrite the equation in terms of just two variables, like we showed earlier and as you will see in one of the examples.

- If a differential equation involves a **second derivative**, try to rewrite it as a first-order differential equation by making a substitution, like we showed earlier and as you will see in the examples.

- If a second-order differential equation has one of the following structures, you should recognize it as a common special case (discussed earlier in this chapter).

  - For an equation of the form $\frac{d^2y}{dx^2} = a^2y$, the solution is $y = Ae^{ax} + Be^{-ax}$.
  - For an equation of the form $\frac{d^2y}{dx^2} = -a^2y$, you may express the solution in the form $y = A\sin ax + B\cos ax$, $y = Ce^{iax} + De^{-iax}$, or $y = x_m \sin(ax + \varphi)$.

- If a differential equation involves **partial derivatives**, try writing the function as the product of functions of each variables. For example, if you are given of function of x and y, try $f(x,y) = g(x)h(y)$, but instead if you are given a function of x and t, try $f(x,t) = g(x)h(t)$. We saw this on pages 163-164 and will see it again in an example.

- If the problem specifies any **boundary conditions**, use these to determine the constants of integration. Alternatively, you may perform definite integrals. Review the pages on boundary conditions and see the examples.

- If the differential equation involves complex numbers, review the page on complex numbers.

- If a differential equation can't be solved by one of the techniques that we reviewed in this chapter, consult a textbook on the subject of differential equations.

- If a problem asks you to show that an equation solves a differential equation, simply plug the expression into the differential equation and see if both sides are equal.

• **Check your answer**: Solving a differential equation sometimes involves much work, but it’s easy to check your answer. Simply plug your solution into the original differential equation and check that both sides are equal.

Example: Solve the following differential equation, given that $y(3) = 2$.

$$\frac{dy}{dx} = \frac{4x}{y}$$

Multiply both sides of the equation by $y\,dx$ in order to separate variables.

$$y\,dy = 4x\,dx$$

Now that the variables have been separated, integrate both sides.

$$\int y\,dy = \int 4x\,dx$$

Include a constant of integration with one of the integrals.

$$\frac{y^2}{2} = 2x^2 + c$$

Multiply both sides of the equation by 2.

$$y^2 = 4x^2 + 2c$$

Since $2c$ is a constant, we may define a new constant $a = 2c$.

$$y^2 = 4x^2 + a$$

Squareroot both sides of the equation. Note that $\sqrt{y^2} = \pm y$ because $(-y)^2 = y^2$.

$$y = \pm\sqrt{4x^2 + a}$$

Apply the boundary condition: Since $y(3) = 2$, plug in $x = 3$ and $y = 2$. Solve for a.

$$2 = \sqrt{4(3)^2 + a}$$

$$2 = \sqrt{4(9) + a}$$

$$2 = \sqrt{36 + a}$$

Square both sides of the equation.

$$2^2 = 36 + a$$

$$4 = 36 + a$$

$$a = 4 - 36 = -32$$

Set a equal to -32 in the equation that we found for y.

$$y = \boxed{\sqrt{4x^2 - 32}} = \sqrt{4(x^2 - 8)} = \sqrt{4}\sqrt{(x^2 - 8)} = \boxed{2\sqrt{x^2 - 8}}$$

Check the answer: Plug $y = \sqrt{4x^2 - 32}$ (or $y = 2\sqrt{x^2 - 8}$) into the original equation.

$$\frac{dy}{dx} = \frac{4x}{y}$$

$$\frac{d\sqrt{4x^2 - 32}}{dx} = \frac{4x}{\sqrt{4x^2 - 32}}$$

Apply the chain rule with $f = u^{1/2}$ and $u = 4x^2 - 32$ to write $\frac{d\sqrt{4x^2-32}}{dx} = \frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$.

$$\frac{\frac{1}{2}(8x)}{\sqrt{4x^2 - 32}} = \frac{4x}{\sqrt{4x^2 - 32}}$$

$$\frac{4x}{\sqrt{4x^2 - 32}} = \frac{4x}{\sqrt{4x^2 - 32}}$$

The answer checks out.

Example: Solve the following differential equation, given that $f(0) = 1$ and $f'(0) = 6$.

$$\frac{d^2 f}{du^2} = 3\frac{df}{du}$$

Make the substitution $g = \frac{df}{du}$ such that $\frac{dg}{du} = \frac{d}{du}\left(\frac{df}{du}\right) = \frac{d^2 f}{du^2}$.

$$\frac{dg}{du} = 3g$$

Divide both sides by g and multiply by du in order to separate variables.

$$\frac{dg}{g} = 3\,du$$

Now that the variables have been separated, integrate both sides.

$$\int \frac{dg}{g} = 3\int du$$

Include a constant of integration with one of the integrals.

$$\ln g = 3u + c$$

Exponentiate both sides of the equation. Recall the identity $e^{\ln g} = g$.

$$g = e^{3u+c} = e^{3u}e^{c} = ae^{3u}$$

We applied the rule $e^{x}e^{y} = e^{x+y}$ and defined a new constant $a = e^{c}$. Recall that $g = \frac{df}{du}$.

$$\frac{df}{du} = ae^{3u}$$

Multiply both sides by du to separate variables again.

$$df = ae^{3u}du$$

$$\int df = a\int e^{3u}\,du$$

We get a second constant of integration with this new integral.

$$f = \frac{ae^{3u}}{3} + b$$

Apply the boundary conditions. Since $f'(0) = 6$ and $f' = \frac{df}{du} = g$, plug in $u = 0$ and $g = 6$.

$$f' = \frac{df}{du} = g = ae^{3u} \quad \rightarrow \quad 6 = ae^{3(0)} = ae^{0} = a(1) = a$$

Recall that $e^0 = 1$. Since $f(0) = 1$, plug in $u = 0$ and $f = 1$. Solve for b. Recall that $a = 6$.

$$f = \frac{ae^{3u}}{3} + b \quad \to \quad 1 = \frac{6e^{3(0)}}{3} + b \quad \to \quad 1 = 2 + b \quad \to \quad b = 1 - 2 = -1$$

Set $a = 6$ and $b = -1$ in the equation that we found for f.

$$f = \boxed{2e^{3u} - 1}$$

Check the answer: Plug $f = 2e^{3u} - 1$ into the original equation.

$$\frac{d^2 f}{du^2} = 3\frac{df}{du} \quad \to \quad \frac{d^2}{du^2}(2e^{3u} - 1) = 3\frac{d}{du}(2e^{3u} - 1)$$

$$18e^{3u} = 18e^{3u}$$

The answer checks out.

Example: Solve the following differential equation, given that $z(0) = 1$ and $z'(1) = 2$.

$$\frac{d^2z}{dy^2} = \frac{1}{\sqrt{z}}$$

Make the substitution $u = \frac{dz}{dy}$ such that $\frac{du}{dy} = \frac{d}{dy}\left(\frac{dz}{dy}\right) = \frac{d^2z}{dy^2}$. Note that $\frac{1}{\sqrt{z}} = \frac{1}{z^{1/2}} = z^{-1/2}$.

$$\frac{du}{dy} = z^{-1/2}$$

Since we have three variables, see if you can eliminate one by applying the chain rule.

$$\frac{du}{dy} = \frac{du}{dz}\frac{dz}{dy} = u\frac{du}{dz}$$

Recall that $u = \frac{dz}{dy}$. Plug $\frac{du}{dy} = u\frac{du}{dz}$ into the equation $\frac{du}{dy} = z^{-1/2}$.

$$u\frac{du}{dz} = z^{-1/2}$$

Multiply by dz in order to separate variables. Include a constant of integration.

$$u\,du = z^{-1/2}dz \quad \rightarrow \quad \int u\,du = \int z^{-1/2}\,dz \quad \rightarrow \quad \frac{u^2}{2} = 2z^{1/2} + c_1 = 2\sqrt{z} + c_1$$

Let's apply the second boundary condition, $z'(1) = 2$, now. Plug in $y = 1$ and $u = \frac{dz}{dy} = 2$.

$$\frac{u^2}{2} = 2\sqrt{z} + c_1 \quad \rightarrow \quad \frac{2^2}{2} = 2\sqrt{1} + c_1 \quad \rightarrow \quad 2 = 2 + c_1 \quad \rightarrow \quad c_1 = 2 - 2 = 0$$

Now we know that $c_1 = 0$. Note that $\sqrt{u^2} = \pm u$ and $\sqrt{z^{1/2}} = \left(z^{1/2}\right)^{1/2} = z^{1/4}$.

$$\frac{u^2}{2} = 2z^{1/2} \quad \rightarrow \quad u^2 = 4z^{1/2} \quad \rightarrow \quad \sqrt{u^2} = \sqrt{4z^{1/2}} \quad \rightarrow \quad \pm u = 2z^{1/4} \quad \rightarrow \quad u = \pm 2z^{1/4}$$

Recall that $u = \frac{dz}{dy}$.

$$\frac{dz}{dy} = \pm 2z^{1/4}$$

Multiply by dy and divide by $z^{1/4}$. Note that $\frac{1}{z^{1/4}} = z^{-1/4}$.

$$\frac{dz}{z^{1/4}} = \pm 2\, dy \quad \rightarrow \quad \int z^{-1/4}\, dz = \pm 2 \int dy \quad \rightarrow \quad \frac{4z^{3/4}}{3} = \pm 2y + c_2$$

$$\rightarrow \quad z = \left[\frac{3}{4}(\pm 2y + c_2)\right]^{\frac{4}{3}} \quad \rightarrow \quad z = \left(\pm\frac{3y}{2} + \frac{3c_2}{4}\right)^{\frac{4}{3}} \quad \rightarrow \quad z = \left(\pm\frac{3y}{2} + a\right)^{\frac{4}{3}}$$

Since $\frac{3c_2}{4}$ is a constant, we defined a new constant: $a = \frac{3c_2}{4}$. Apply the first boundary condition, $z(0) = 1$. Plug in $y = 0$ and $z = 1$. Solve for a.

$$z = \left(\pm\frac{3y}{2} + a\right)^{\frac{4}{3}} \quad \rightarrow \quad 1 = (0 + a)^{4/3} \quad \rightarrow \quad 1 = a^{4/3} \quad \rightarrow \quad (1)^{3/4} = a \quad \rightarrow \quad \pm 1 = a$$

Set $a = \pm 1$ in the equation that we found for z. The solution has 4 possible forms from the $\pm$'s.

$$z = \boxed{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{4}{3}}}$$

Check the answer: Plug $z = \left(\pm\frac{3y}{2} \pm 1\right)^{\frac{4}{3}}$ into the original equation.

The original equation, $\frac{d^2z}{dy^2} = \frac{1}{\sqrt{z}}$, involves a second derivative of z with respect to y. Begin by taking a first derivative. Apply the chain rule with $z = u^{4/3}$ and $u = \pm\frac{3y}{2} \pm 1$.

$$\frac{dz}{dy} = \frac{dz}{du}\frac{du}{dy} = \frac{d}{du}\left(u^{4/3}\right)\frac{d}{dy}\left(\pm\frac{3y}{2} \pm 1\right) = \left(\frac{4u^{1/3}}{3}\right)\left(\pm\frac{3}{2}\right) = \pm 2u^{1/3} = \pm 2\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{1}{3}}$$

Now take a second derivative, applying the chain rule again.

$$\frac{d^2z}{dy^2} = \frac{d}{dy}\left(\frac{dz}{dy}\right) = \pm\frac{d}{dy}2\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{1}{3}} = \pm 2\frac{d}{dy}\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{1}{3}} = \pm 2\frac{d}{dy}u^{1/3} = \pm 2\frac{d}{du}\left(u^{1/3}\right)\frac{du}{dy}$$

$$= \pm 2\frac{d}{du}\left(u^{1/3}\right)\frac{d}{dy}\left(\pm\frac{3y}{2} \pm 1\right) = \pm 2\left(\frac{u^{-2/3}}{3}\right)\left(\pm\frac{3}{2}\right) = u^{-2/3} = \left(\pm\frac{3y}{2} \pm 1\right)^{-\frac{2}{3}} = \frac{1}{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{2}{3}}}$$

Note that $\pm 2\left(\pm\frac{3}{2}\right) = 3$ because $2\left(\frac{3}{2}\right) = 3$ and $(-2)\left(-\frac{3}{2}\right)$. If you follow where these $\pm$ signs came from, you should see that both signs must be the same (both positive or both negative). Recall the rule $x^{-n} = \frac{1}{x^n}$. Plug this second derivative into the original differential equation.

$$\frac{d^2z}{dy^2} = \frac{1}{\sqrt{z}} \quad \rightarrow \quad \frac{d^2}{dz^2}\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{4}{3}} = \frac{1}{\sqrt{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{4}{3}}}} \quad \rightarrow \quad \frac{1}{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{2}{3}}} = \frac{1}{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{2}{3}}}$$

The answer checks out. We applied the rule $\sqrt{x^m} = (x^m)^{1/2} = x^{m/2}$ to write $\sqrt{\left(\pm\frac{3y}{2} \pm 1\right)^{\frac{4}{3}}} = \left(\pm\frac{3y}{2} \pm 1\right)^{\frac{2}{3}}$.

Example: Solve the following differential equation. Include no more than two constants of integration in your final answer.

$$\frac{\partial^2 z(x,y)}{\partial x^2} = 3\frac{\partial z(x,y)}{\partial y}$$

When there are partial derivatives (and three or more variables), try writing the function as a product of functions of each individual variable. In this case, we will try $z(x,y) = f(x)g(y)$.

$$\frac{\partial^2}{\partial x^2} f(x)g(y) = 3\frac{\partial}{\partial y} f(x)g(y)$$

Note that $g(y)$ doesn't depend on x, such that $\frac{\partial^2}{\partial x^2} g(y) = 0$. Similarly, $\frac{\partial}{\partial y} f(x) = 0$.

$$g(y)\frac{d^2 f(x)}{dx^2} = 3f(x)\frac{dg(y)}{dy}$$

Since $f(x)$ is a function of x only and $g(y)$ is a function of y only, the derivatives are no longer 'partial' derivatives. Divide both sides of the equation by $f(x)g(y)$.

$$\frac{1}{f(x)}\frac{d^2 f(x)}{dx^2} = \frac{3}{g(y)}\frac{dg(y)}{dy}$$

The equation above states that a function of x equals a function of y. The only way that both sides can be equal for all possible values of x and y is if each side equals a constant. If we call that constant κ (Greek letter kappa), we get two new differential equations:

$$\frac{1}{f(x)}\frac{d^2f(x)}{dx^2} = \kappa \quad , \quad \frac{3}{g(y)}\frac{dg(y)}{dy} = \kappa$$

Apply algebra to isolate the derivative in each equation.

$$\frac{d^2f(x)}{dx^2} = \kappa\, f(x) \quad , \quad \frac{dg(y)}{dy} = \frac{\kappa}{3}g(y)$$

The first equation is characteristic of exponential growth or decay: See page 166, noting that we have κ in place of a^2. That is, $\kappa = a^2$ or $\sqrt{\kappa} = a$.

$$f(x) = Ae^{\sqrt{\kappa}x} + Be^{-\sqrt{\kappa}x}$$

The second equation is also characteristic of exponential growth: This equation states that the first derivative of the function is proportional to the function. You can easily verify that the following solution satisfies the second equation (by plugging it into the equation).

$$g(y) = Ce^{\kappa y/3}$$

The solution is the product of $f(x)$ and $g(y)$.

$$z(x,y) = f(x)g(y) = \left(Ae^{\sqrt{\kappa}x} + Be^{-\sqrt{\kappa}x}\right)\left(Ce^{\kappa y/3}\right) = \boxed{De^{\sqrt{\kappa}x+\kappa y/3} + Ee^{-\sqrt{\kappa}x+\kappa y/3}}$$

We applied the rule $e^u e^v = e^{u+v}$ and defined two new constants, $D = AC$ and $E = BC$.

Check the answer: Plug $z(x,y) = De^{\sqrt{\kappa}x+\kappa y/3} + Ee^{-\sqrt{\kappa}x+\kappa y/3}$ into the original equation.

$$\frac{\partial^2 z(x,y)}{\partial x^2} = 3\frac{\partial z(x,y)}{\partial y} \quad \rightarrow \quad \frac{\partial^2}{\partial x^2}\left(De^{\sqrt{\kappa}x+\kappa y/3} + Ee^{-\sqrt{\kappa}x+\kappa y/3}\right) = 3\frac{\partial}{\partial y}\left(De^{\sqrt{\kappa}x+\kappa y/3} + Ee^{-\sqrt{\kappa}x+\kappa y/3}\right)$$

Apply the rule $e^{u+v} = e^u e^v$.

$$\frac{\partial^2}{\partial x^2}\left(De^{\sqrt{\kappa}x}e^{\kappa y/3} + Ee^{-\sqrt{\kappa}x}e^{\kappa y/3}\right) = 3\frac{\partial}{\partial y}\left(De^{\sqrt{\kappa}x}e^{\kappa y/3} + Ee^{-\sqrt{\kappa}x}e^{\kappa y/3}\right)$$

Note that these are partial derivatives. When taking a partial derivative with respect to the variable x, treat the independent variable y as if it were a constant, and vice-versa.

$$De^{\kappa y/3}\frac{\partial^2}{\partial x^2}\left(e^{\sqrt{\kappa}x}\right) + Ee^{\kappa y/3}\frac{\partial^2}{\partial x^2}\left(e^{-\sqrt{\kappa}x}\right) = 3De^{\sqrt{\kappa}x}\frac{\partial}{\partial y}\left(e^{\kappa y/3}\right) + 3Ee^{-\sqrt{\kappa}x}\frac{\partial}{\partial y}\left(e^{\kappa y/3}\right)$$

$$De^{\kappa y/3}\left[\left(\sqrt{\kappa}\right)^2 e^{\sqrt{\kappa}x}\right] + Ee^{\kappa y/3}\left[(-1)^2\left(\sqrt{\kappa}\right)^2 e^{-\sqrt{\kappa}x}\right] = 3De^{\sqrt{\kappa}x}\left(\frac{\kappa e^{\kappa y/3}}{3}\right) + 3Ee^{-\sqrt{\kappa}x}\left(\frac{\kappa e^{\kappa y/3}}{3}\right)$$

$$D\kappa e^{\sqrt{\kappa}x}e^{\kappa y/3} + E\kappa e^{-\sqrt{\kappa}x}e^{\kappa y/3} = D\kappa e^{\sqrt{\kappa}x}e^{\kappa y/3} + E\kappa e^{-\sqrt{\kappa}x}e^{\kappa y/3}$$

The answer checks out.

Example: Show that $y = A\sin ax + B\cos ax$ and $y = Ce^{iax} + De^{-iax}$ are equivalent solutions to $\frac{d^2y}{dx^2} = -a^2y$, and determine how the constants are related.

It's easy to see that both equations solve the differential equation by plugging them in:

$$\frac{d^2}{dx^2}(A\sin ax + B\cos ax) = -a^2(A\sin ax + B\cos ax)$$
$$\frac{d}{dx}(aA\cos ax - aB\sin ax) = -a^2A\sin ax - a^2B\cos ax$$
$$-a^2A\sin ax - a^2B\cos ax = -a^2A\sin ax - a^2B\cos ax$$
$$\frac{d^2}{dx^2}\left(Ce^{iax} + De^{-iax}\right) = -a^2\left(Ce^{iax} + De^{-iax}\right)$$
$$\frac{d}{dx}\left(iaCe^{iax} - iaDe^{-iax}\right) = -a^2Ce^{iax} - a^2De^{-iax}$$
$$i^2a^2Ce^{iax} + i^2a^2De^{-iax} = -a^2Ce^{iax} - a^2De^{-iax}$$
$$-a^2Ce^{iax} - a^2De^{-iax} = -a^2Ce^{iax} - a^2De^{-iax}$$

We used the identity $i^2 = -1$. Both solutions check out. To determine how the constants are related, apply the identities $e^{ix} = \cos x + i\sin x$ and $e^{-ix} = \cos x - i\sin x$ (see page 166). In this problem, we have ax in place of x: Compare e^{iax} with e^{ix}.

$$y = Ce^{iax} + De^{-iax} = C(\cos ax + i\sin ax) + D(\cos ax - i\sin ax)$$
$$= C\cos ax + iC\sin ax + D\cos ax - iD\sin ax$$
$$= (C + D)\cos ax + i(C - D)\sin ax$$

Compare the above equation with $y = A\sin ax + B\cos ax$ to see that:

$$\boxed{A = C + D} \quad , \quad \boxed{B = i(C - D)}$$

Chapter 12 Problems

1. The charge (Q) stored on a discharging capacitor in an RC circuit satisfies the following equation. The charge stored initially equals Q_m at $t = 0$.

$$\frac{Q}{RC} = -\frac{dQ}{dt}$$

(A) Derive an equation for the charge stored at time t.

(B) Derive an equation for the current (I) at time t, given that $I = -\frac{dQ}{dt}$ (the minus sign is present because the charge stored on the plates of the capacitor is decreasing).

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) $Q(t) = Q_m e^{-t/RC}$ (B) $I(t) = \frac{Q_m}{RC} e^{-t/RC}$

2. While a toy car accelerates along a straight line, its acceleration equals:

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2} = \beta t$$

where β is a constant. The initial position and velocity of the car are x_0 and v_{x0} when $t = 0$.

(A) Derive an equation for the velocity (v_x) of the toy car as a function of time.

(B) Derive an equation for the position (x) of the toy car as a function of time.

Want help? Check the solution at the end of the chapter.

Answers: 2. (A) $v_x(t) = v_{x0} + \frac{\beta t^2}{2}$ (B) $x(t) = x_0 + v_{x0}t + \frac{\beta t^3}{6}$

3. A spring oscillates back and forth according to the following equation, where k is the spring constant and m is the suspended mass.

$$m\frac{d^2x}{dt^2} = -kx$$

The initial conditions are $x(0) = x_m$ and $v_x(0) = 0$ where $v_x = \frac{dx}{dt}$.

(A) Make the substitution $v_x = \frac{dx}{dt}$ and apply the chain rule to eliminate t from this equation.

(B) Integrate the differential equation obtained in Part A and solve for v_x in terms of x.

(C) Make the substitution $v_x = \frac{dx}{dt}$ and integrate again to solve for x in terms of t.

Want help? Check the solution at the end of the chapter.

Answers: 3. (A) $mv_x\frac{dv_x}{dx} = -kx$ (B) $v_x(x) = \sqrt{\frac{k}{m}(x_m^2 - x^2)}$ (C) $x(t) = x_m \cos\left(\sqrt{\frac{k}{m}}t\right)$

4. An object of mass m falls straight downward according to the following equation, where g represents a constant gravitational field and b is a constant relating to air resistance.

$$m\frac{d^2y}{dt^2} + b\frac{dy}{dt} + mg = 0$$

The initial values are $y_0 = y(0) = h$ and $v_{y0} = v_y(0) = 0$, where $v_y = \frac{dy}{dt}$.

(A) Derive an equation for v_y as a function of time.

(B) Derive an equation for y as a function of time.

(C) Show that if h is sufficiently tall, the object reaches a terminal speed of $v_t = \frac{mg}{b}$.

Want help? Check the solution at the end of the chapter.

Answers: 4. (A) $v_y(t) = \frac{mg}{b}\left(e^{-bt/m} - 1\right)$ (B) $y(t) = \frac{m^2g}{b^2}\left(1 - e^{-bt/m}\right) - \frac{mgt}{b} + h$

5. For a particular rectangular wave guide with width a and height b, the z-component of the electric field (E_z) satisfies the following equation, where k, ω, and c are constants.

$$\frac{\partial^2 E_z(x,y)}{\partial x^2} + \frac{\partial^2 E_z(x,y)}{\partial y^2} = \left(k^2 - \frac{\omega^2}{c^2}\right) E_z(x,y)$$

(A) Make the substitution $E_z(x,y) = f(x)g(y)$ and derive the following equations.

$$\frac{d^2 f(x)}{dx^2} = -c_1^2 f(x) \quad , \quad \frac{d^2 g(y)}{dy^2} = -c_2^2 g(y)$$

(B) Show that $E_z(x,y) = E_m \cos\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{b}\right)$ is a solution to the second-order partial differential equation above, where $k^2 = \frac{\omega^2}{c^2} - \frac{n_x^2 \pi^2}{a^2} - \frac{n_y^2 \pi^2}{b^2}$.

Want help? Check the solution at the end of the chapter.

Solutions to Chapter 12

1. Separate variables (t and Q): Multiply both sides by dt and divide by Q.

$$\frac{dt}{RC} = -\frac{dQ}{Q}$$

Now that variables have been separated, integrate both sides, where R and C are constants.

$$\frac{1}{RC}\int dt = -\int \frac{dQ}{Q}$$

We will call the constant of integration a.

$$\frac{t}{RC} = -\ln Q + a$$

Add $\ln Q$ to both sides and subtract $\frac{t}{RC}$.

$$\ln Q = a - \frac{t}{RC}$$

Exponentiate both sides of the equation. Note that $e^{\ln Q} = Q$.

$$Q = e^{a-t/RC}$$

Apply the identity $e^x e^{-y} = e^{x-y}$.

$$Q = e^a e^{-t/RC}$$

The boundary condition is $Q(0) = Q_m$: Plug in $t = 0$ and $Q = Q_m$.

$$Q_m = e^a e^0 = e^a(1) = e^a$$

Thus, we see that $Q_m = e^a$. Substitute this into the equation for Q.

$$Q(t) = \boxed{Q_m e^{-t/RC}}$$

Check your answer: Plug the solution for $Q(t)$ into the original differential equation.

$$\frac{Q}{RC} = -\frac{dQ}{dt} \quad \rightarrow \quad \frac{Q_m e^{-t/RC}}{RC} = -\frac{d}{dt}Q_m e^{-t/RC} \quad \rightarrow \quad \frac{Q_m e^{-t/RC}}{RC} = \frac{Q_m e^{-t/RC}}{RC}$$

(B) Current (I) equals a derivative of charge (Q) with respect to time, apart from an overall minus sign that is explained in the problem.

$$I(t) = -\frac{dQ}{dt} = -\frac{d}{dt}\left(Q_m e^{-t/RC}\right) = (-Q_m)\left(-\frac{1}{RC}\right)e^{-t/RC} = \boxed{\frac{Q_m}{RC}e^{-t/RC}}$$

2. (A) Begin with the first-order differential equation involving velocity.

$$\frac{dv_x}{dt} = \beta t$$

Separate variables (v_x and t): Multiply both sides by dt.

$$dv_x = \beta t \, dt$$

Now that variables have been separated, integrate both sides, where β is a constant.

$$\int dv_x = \beta \int t \, dt$$

We will call the constant of integration c_1.

$$v_x = \frac{\beta t^2}{2} + c_1$$

Apply the boundary condition that $v_x(0) = v_{x0}$: Plug in $t = 0$ and $v_x = v_{x0}$.

$$v_{x0} = c_1$$

Substitute this into the equation for v_x.

$$v_x(t) = \frac{\beta t^2}{2} + v_{x0} = \boxed{v_{x0} + \frac{\beta t^2}{2}}$$

Check your answer: Plug the solution for $v_x(t)$ into the original differential equation.

$$\frac{dv_x}{dt} = \beta t \quad \to \quad \frac{dv_x}{dt}\left(v_{x0} + \frac{\beta t^2}{2}\right) = \beta t \quad \to \quad \beta t = \beta t$$

(B) Make the substitution $v_x = \frac{dx}{dt}$.

$$\frac{dx}{dt} = v_{x0} + \frac{\beta t^2}{2}$$

Separate variables again: Multiply both sides by dt.

$$dx = v_{x0} dt + \frac{\beta t^2}{2} dt$$

Integrate both sides, where v_{x0} and β are constants.

$$\int dx = v_{x0} \int dt + \frac{\beta}{2} \int t^2 \, dt$$

We will call the constant of integration c_2.

$$x = v_{x0} t + \frac{\beta t^3}{6} + c_2$$

Apply the boundary condition that $x(0) = x_0$: Plug in $t = 0$ and $x = x_0$.

$$x_0 = c_2$$

Substitute this into the equation for x.

$$x(t) = v_{x0} t + \frac{\beta t^3}{6} + x_0 = \boxed{x_0 + v_{x0} t + \frac{\beta t^3}{6}}$$

Check your answer: Plug the solution for $x(t)$ into the differential equation involving x.

$$\frac{d^2x}{dt^2} = \beta t \quad \to \quad \frac{d^2}{dt^2}\left(x_0 + v_{x0} t + \frac{\beta t^3}{6}\right) = \beta t \quad \to \quad \beta t = \beta t$$

3. (A) Make the substitution $v_x = \frac{dx}{dt}$ such that $\frac{dv_x}{dt} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{d^2x}{dt^2}$.

$$m\frac{dv_x}{dt} = -kx$$

Apply the chain rule to eliminate t.

$$\frac{dv_x}{dt} = \frac{dv_x}{dx}\frac{dx}{dt} = v_x\frac{dv_x}{dx}$$

Substitute $v_x\frac{dv_x}{dx}$ in place of $\frac{dv_x}{dt}$ in the equation $m\frac{dv_x}{dt} = -kx$.

$$\boxed{mv_x\frac{dv_x}{dx} = -kx}$$

(B) Now that there are just two variables (x and v_x), separate variables: Multiply by dx.

$$mv_x\, dv_x = -kx\, dx$$

Now that variables have been separated, integrate both sides where m and k are constants.

$$m\int v_x\, dv_x = -k\int x\, dx$$

We will call the constant of integration c_1.

$$\frac{mv_x^2}{2} = -\frac{kx^2}{2} + c_1$$

Apply the boundary conditions $v_x(0) = 0$ and $x(0) = x_m$. Be careful here: This means that $v_x = 0$ at $t = 0$ (not at x equal to zero). Plug in $v_x = 0$ and $x = x_m$.

$$0 = -\frac{kx_m^2}{2} + c_1 \quad \to \quad c_1 = \frac{kx_m^2}{2}$$

Substitute this into the equation for v_x.

$$\frac{mv_x^2}{2} = -\frac{kx^2}{2} + \frac{kx_m^2}{2} = \frac{kx_m^2}{2} - \frac{kx^2}{2} = \frac{k}{2}(x_m^2 - x^2)$$

Isolate v_x. Multiply both sides by 2 and divide by m.

$$v_x^2 = \frac{k}{m}(x_m^2 - x^2)$$

Squareroot both sides of the equation.

$$v_x(x) = \boxed{\sqrt{\frac{k}{m}(x_m^2 - x^2)}}$$

(C) Make the substitution $v_x = \frac{dx}{dt}$.

$$\frac{dx}{dt} = \sqrt{\frac{k}{m}(x_m^2 - x^2)}$$

Separate variables (x and t). Multiply both sides by dt and divide by the squareroot.

$$\frac{dx}{\sqrt{\frac{k}{m}(x_m^2 - x^2)}} = dt$$

Now that variables have been separated, integrate both sides. Note that x_m is a constant.

$$\int \frac{dx}{\sqrt{\frac{k}{m}(x_m^2 - x^2)}} = \int dt$$

Note that $\sqrt{AB} = \sqrt{A}\sqrt{B}$. We will call the constant of integration c_2.

$$\frac{1}{\sqrt{k/m}} \int \frac{dx}{\sqrt{x_m^2 - x^2}} = t + c_2$$

Note that $\frac{1}{\sqrt{k/m}} = \sqrt{\frac{m}{k}}$. Make the substitution $x = x_m \cos\theta$ and $dx = -x_m \sin\theta\, d\theta$. Recall that $\sin^2\theta + \cos^2\theta = 1$ such that $1 - \cos^2\theta = \sin^2\theta$.

$$\sqrt{\frac{m}{k}} \int \frac{-x_m \sin\theta\, d\theta}{\sqrt{x_m^2 - x_m^2 \cos^2\theta}} = t + c_2 \quad \to \quad -x_m \sqrt{\frac{m}{k}} \int \frac{\sin\theta\, d\theta}{\sqrt{x_m^2 (1 - \cos^2\theta)}} = t + c_2$$

$$-x_m \sqrt{\frac{m}{k}} \int \frac{\sin\theta\, d\theta}{\sqrt{x_m^2 \sin^2\theta}} = t + c_2 \quad \to \quad -x_m \sqrt{\frac{m}{k}} \int \frac{\sin\theta\, d\theta}{x_m \sin\theta} = t + c_2$$

$$-\sqrt{\frac{m}{k}} \int d\theta = t + c_2 \quad \to \quad -\sqrt{\frac{m}{k}}\, \theta = t + c_2$$

Take the inverse cosine of both sides of $x = x_m \cos\theta$ to see that $\theta = \cos^{-1}\left(\frac{x}{x_m}\right)$.

$$-\sqrt{\frac{m}{k}} \cos^{-1}\left(\frac{x}{x_m}\right) = t + c_2$$

Apply the boundary condition $x(0) = x_m$. Plug in $x = x_m$ and $t = 0$.

$$-\sqrt{\frac{m}{k}} \cos^{-1}\left(\frac{x_m}{x_m}\right) = 0 + c_2 \quad \to \quad -\sqrt{\frac{m}{k}} \cos^{-1}(1) = c_2$$

Recall from trig that $\cos 0 = 1$ such that $\cos^{-1}(1) = 0$.

$$-\sqrt{\frac{m}{k}}(0) = c_2 \quad \rightarrow \quad c_2 = 0$$

Plug $c_2 = 0$ into the equation involving x.

$$-\sqrt{\frac{m}{k}}\cos^{-1}\left(\frac{x}{x_m}\right) = t$$

Multiply both sides by $-\sqrt{\frac{k}{m}}$. Note that $\sqrt{\frac{k}{m}}\sqrt{\frac{m}{k}} = \sqrt{\frac{km}{mk}} = \sqrt{1} = 1$.

$$\cos^{-1}\left(\frac{x}{x_m}\right) = -\sqrt{\frac{k}{m}}\,t$$

Take the cosine of both sides. Note that $\cos\left[\cos^{-1}\left(\frac{x}{x_m}\right)\right] = \frac{x}{x_m}$.

$$\frac{x}{x_m} = \cos\left(-\sqrt{\frac{k}{m}}\,t\right) = \cos\left(\sqrt{\frac{k}{m}}\,t\right)$$

Recall that cosine is an even function, such that $\cos(-\varphi) = \cos\varphi$. Multiply both sides by x_m.

$$x(t) = \boxed{x_m \cos\left(\sqrt{\frac{k}{m}}\,t\right)}$$

Check your answer: Plug the solution for $x(t)$ into the original differential equation.

$$m\frac{d^2x}{dt^2} = -kx \quad \rightarrow \quad m\frac{d^2}{dt^2}x_m \cos\left(\sqrt{\frac{k}{m}}\,t\right) = -kx_m \cos\left(\sqrt{\frac{k}{m}}\,t\right)$$

$$-mx_m\left(\sqrt{\frac{k}{m}}\right)^2 \cos\left(\sqrt{\frac{k}{m}}\,t\right) = -kx_m \cos\left(\sqrt{\frac{k}{m}}\,t\right)$$

$$-kx_m \cos\left(\sqrt{\frac{k}{m}}\,t\right) = -kx_m \cos\left(\sqrt{\frac{k}{m}}\,t\right)$$

Note that $m\left(\sqrt{\frac{k}{m}}\right)^2 = m\left(\frac{k}{m}\right) = k$.

4. (A) Make the substitution $v_y = \frac{dy}{dt}$ such that $\frac{dv_y}{dt} = \frac{d}{dt}\left(\frac{dy}{dt}\right) = \frac{d^2y}{dt^2}$.

$$m\frac{dv_y}{dt} + bv_y + mg = 0$$

Separate variables (v_y and t). Multiply both sides by dt and move two terms to the right.

$$m\, dv_y = -bv_y\, dt - mg\, dt = -(bv_y + mg)dt$$

Now divide both sides of the equation by $(bv_y + mg)$.

$$\frac{m}{bv_y + mg} dv_y = -dt$$

Now that variables have been separated, integrate both sides where m and b are constants.

$$m \int \frac{dv_y}{bv_y + mg} = -\int dt$$

Factor out a b in the denominator. We will call the constant of integration c_1.

$$m \int \frac{dv_y}{b\left(v_y + \frac{mg}{b}\right)} = -t + c_1$$

Multiply both sides of the equation by b and divide by m.

$$\int \frac{dv_y}{v_y + \frac{mg}{b}} = -\frac{b}{m}t + \frac{b}{m}c_1$$

Since b and m are constants, we may define a new constant: $c_2 = \frac{b}{m}c_1$. Make the substitutions $u = v_y + \frac{mg}{b}$ and $du = dv_y$.

$$\int \frac{du}{u} = -\frac{b}{m}t + c_2$$

$$\ln u = -\frac{b}{m}t + c_2$$

Exponentiate both sides. Recall the identity $e^{\ln u} = u$.

$$u = e^{-\frac{bt}{m}+c_2} = e^{-bt/m}e^{c_2} = c_3 e^{-bt/m}$$

Note that $e^a e^b = e^{a+b}$. We defined a new constant: $c_3 = e^{c_2}$. Recall that $u = v_y + \frac{mg}{b}$.

$$v_y + \frac{mg}{b} = c_3 e^{-bt/m} \quad \to \quad v_y = c_3 e^{-bt/m} - \frac{mg}{b}$$

Apply the boundary condition $v_y(0) = 0$. Plug in $v_y = 0$ and $t = 0$.

$$0 = c_3 e^0 - \frac{mg}{b} \quad \rightarrow \quad c_3(1) = \frac{mg}{b} \quad \rightarrow \quad c_3 = \frac{mg}{b}$$

Plug $c_3 = \frac{mg}{b}$ into the equation for v_y.

$$v_y(t) = \frac{mg}{b} e^{-bt/m} - \frac{mg}{b} = \boxed{\frac{mg}{b}\left(e^{-bt/m} - 1\right)} = \boxed{-\frac{mg}{b}\left(1 - e^{-bt/m}\right)}$$

(B) Make the substitution $v_y = \frac{dy}{dt}$.

$$\frac{dy}{dt} = \frac{mg}{b}\left(e^{-bt/m} - 1\right)$$

Separate variables again: Multiply both sides by dt.

$$dy = \frac{mg}{b}\left(e^{-bt/m} - 1\right)dt$$

Now that variables have been separated, integrate both sides.

$$\int dy = \frac{mg}{b}\int \left(e^{-bt/m} - 1\right) dt = \frac{mg}{b}\int e^{-bt/m}\, dt - \frac{mg}{b}\int dt$$

We will call the constant of integration c_4.

$$y = \frac{mg}{b}\left(-\frac{m}{b}\right)e^{-bt/m} - \frac{mgt}{b} + c_4 = -\frac{m^2 g}{b^2}e^{-bt/m} - \frac{mgt}{b} + c_4$$

Apply the boundary condition $y(0) = h$. Plug in $y = h$ and $t = 0$.

$$h = -\frac{m^2 g}{b^2}e^0 - \frac{mg(0)}{b} + c_4 = -\frac{m^2 g}{b^2}(1) - 0 + c_4 = -\frac{m^2 g}{b^2} + c_4 \quad \rightarrow \quad c_4 = \frac{m^2 g}{b^2} + h$$

Plug $c_4 = \frac{m^2 g}{b^2} + h$ into the equation for y.

$$y(t) = -\frac{m^2 g}{b^2}e^{-bt/m} - \frac{mgt}{b} + \frac{m^2 g}{b^2} + h = \boxed{\frac{m^2 g}{b^2}\left(1 - e^{-bt/m}\right) - \frac{mgt}{b} + h}$$

Check your answer: Plug the solution for $y(t)$ into the original differential equation.

$$m\frac{d^2}{dt^2}\left[\frac{m^2g}{b^2}\left(1-e^{-bt/m}\right)-\frac{mgt}{b}+h\right]+b\frac{d}{dt}\left[\frac{m^2g}{b^2}\left(1-e^{-bt/m}\right)-\frac{mgt}{b}+h\right]+mg=0$$

$$m\frac{d^2}{dt^2}\left[\frac{m^2g}{b^2}-\frac{m^2g}{b^2}e^{-bt/m}-\frac{mgt}{b}+h\right]+b\frac{d}{dt}\left[\frac{m^2g}{b^2}-\frac{m^2g}{b^2}e^{-bt/m}-\frac{mgt}{b}+h\right]+mg=0$$

$$m\left[-\frac{m^2g}{b^2}\left(-\frac{b}{m}\right)^2e^{-bt/m}\right]+b\left[-\frac{m^2g}{b^2}\left(-\frac{b}{m}\right)e^{-bt/m}-\frac{mg}{b}\right]+mg=0$$

$$m\left(-\frac{m^2g}{b^2}\frac{b^2}{m^2}e^{-bt/m}\right)+b\left(\frac{mg}{b}e^{-bt/m}-\frac{mg}{b}\right)+mg=0$$

$$-mg\,e^{-bt/m}+mge^{-bt/m}-mg+mg=0 \quad \to \quad 0=0$$

(C) Take the limit that $t\to\infty$ in the answer to Part (A). Note that $e^{-x}=\frac{1}{e^x}\to 0$ as $x\to\infty$.

$$v_{yt}=\lim_{t\to\infty}v_y=\lim_{t\to\infty}\left[-\frac{mg}{b}\left(1-e^{-bt/m}\right)\right]=-\frac{mg}{b}\left(1-\lim_{t\to\infty}e^{-bt/m}\right)=-\frac{mg}{b}(1-0)=-\frac{mg}{b}$$

Speed is the magnitude of velocity: $v_t=|v_{yt}|=\boxed{\frac{mg}{b}}$.

5. (A) Make the substitution $E_z(x,y) = f(x)g(y)$.

$$\frac{\partial^2 f(x)g(y)}{\partial x^2} + \frac{\partial^2 f(x)g(y)}{\partial y^2} = \left(k^2 - \frac{\omega^2}{c^2}\right) f(x)g(y)$$

Since $g(y)$ doesn't depend on x, we may treat it as a constant when taking a partial derivative with respect to x: $\frac{\partial^2 f(x)g(y)}{\partial x^2} = g(y)\frac{\partial^2 f(x)}{\partial x^2}$. Similarly, $\frac{\partial^2 f(x)g(y)}{\partial y^2} = f(x)\frac{\partial^2 g(y)}{\partial y^2}$.

$$g(y)\frac{d^2 f(x)}{dx^2} + f(x)\frac{d^2 g(y)}{dy^2} = \left(k^2 - \frac{\omega^2}{c^2}\right) f(x)g(y)$$

The derivatives are no longer 'partial.' Divide both sides of the equation by $f(x)g(y)$.

$$\frac{1}{f(x)}\frac{d^2 f(x)}{dx^2} + \frac{1}{g(y)}\frac{d^2 g(y)}{dy^2} = k^2 - \frac{\omega^2}{c^2}$$

Separate variables (x and y). Move $\frac{1}{g(y)}\frac{d^2 g(y)}{dy^2}$ to the other side.

$$\frac{1}{f(x)}\frac{d^2 f(x)}{dx^2} = k^2 - \frac{\omega^2}{c^2} - \frac{1}{g(y)}\frac{d^2 g(y)}{dy^2}$$

The equation above states that a function of x equals a function of y. The only way that both sides can be equal for all possible values of x and y is if each side equals a constant. If we call that constant $-c_1^2$, we get two new differential equations:

$$\frac{1}{f(x)}\frac{d^2 f(x)}{dx^2} = -c_1^2 \quad , \quad k^2 - \frac{\omega^2}{c^2} - \frac{1}{g(y)}\frac{d^2 g(y)}{dy^2} = -c_1^2$$

Apply algebra to isolate the derivative in each equation.

$$\boxed{\frac{d^2 f(x)}{dx^2} = -c_1^2 f(x)} \quad , \quad \boxed{\frac{d^2 g(y)}{dy^2}} = \left(k^2 - \frac{\omega^2}{c^2} + c_1^2\right) g(y) \boxed{= -c_2^2 g(y)}$$

We defined a new constant: $c_2^2 = -\left(k^2 - \frac{\omega^2}{c^2} - c_1^2\right) = -k^2 + \frac{\omega^2}{c^2} + c_1^2$.

(B) Substitute $E_z(x,y) = E_m \cos\left(\frac{n_x \pi x}{a}\right)\sin\left(\frac{n_y \pi y}{b}\right)$ into the given differential equation.

$$\frac{\partial^2}{\partial x^2}\left[E_m \cos\left(\frac{n_x \pi x}{a}\right)\sin\left(\frac{n_y \pi y}{b}\right)\right] + \frac{\partial^2}{\partial y^2}\left[E_m \cos\left(\frac{n_x \pi x}{a}\right)\sin\left(\frac{n_y \pi y}{b}\right)\right]$$

$$= \left(k^2 - \frac{\omega^2}{c^2}\right) E_m \cos\left(\frac{n_x \pi x}{a}\right)\sin\left(\frac{n_y \pi y}{b}\right)$$

When we take a partial derivative with respect to x, we treat the independent variable y as if it were a constant, and vice-versa.

$$-\left(\frac{n_x\pi}{a}\right)^2 E_m \cos\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right) - \left(\frac{n_y\pi}{b}\right)^2 E_m \cos\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right)$$

$$= \left(k^2 - \frac{\omega^2}{c^2}\right) E_m \cos\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right)$$

Recall from calculus that $\frac{d^2}{dx^2}\cos(c_1 x) = \frac{d}{dx}[-c_1\sin(c_1 x)] = -c_1^2\cos(c_1 x)$ and $\frac{d^2}{dy^2}\sin(c_2 y) = \frac{d}{dy}[c_2\cos(c_2 y)] = -c_2^2\sin(c_2 y)$. Note that $E_m\cos\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right)$ cancels out.

$$-\frac{n_x^2\pi^2}{a^2} - \frac{n_y^2\pi^2}{b^2} = k^2 - \frac{\omega^2}{c^2}$$

Substitute the expression, $k^2 = \frac{\omega^2}{c^2} - \frac{\pi^2 n_x^2}{a^2} - \frac{\pi^2 n_y^2}{b^2}$, which was given in the problem.

$$-\frac{n_x^2\pi^2}{a^2} - \frac{n_y^2\pi^2}{b^2} = \frac{\omega^2}{c^2} - \frac{\pi^2 n_x^2}{a^2} - \frac{\pi^2 n_y^2}{b^2} - \frac{\omega^2}{c^2}$$

$$-\frac{n_x^2\pi^2}{a^2} - \frac{n_y^2\pi^2}{b^2} = -\frac{\pi^2 n_x^2}{a^2} - \frac{\pi^2 n_y^2}{b^2}$$

This completes our proof.

13 SCHRÖDINGER'S EQUATION

Relevant Terminology

Frequency – the number of oscillations completed per second.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Potential energy – work that can be done by changing position. All forms of potential energy are stored energy.

Momentum – mass times velocity.

Wave function – a quantity that helps to determine probabilities relating to matter waves.

Eigenvalues – a set of values for which a differential equation has a nonzero solution.

Schrödinger's Equation for One-dimensional Motion

For one-dimensional motion along $\pm x$, Schrödinger's **time-dependent** equation is:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}$$

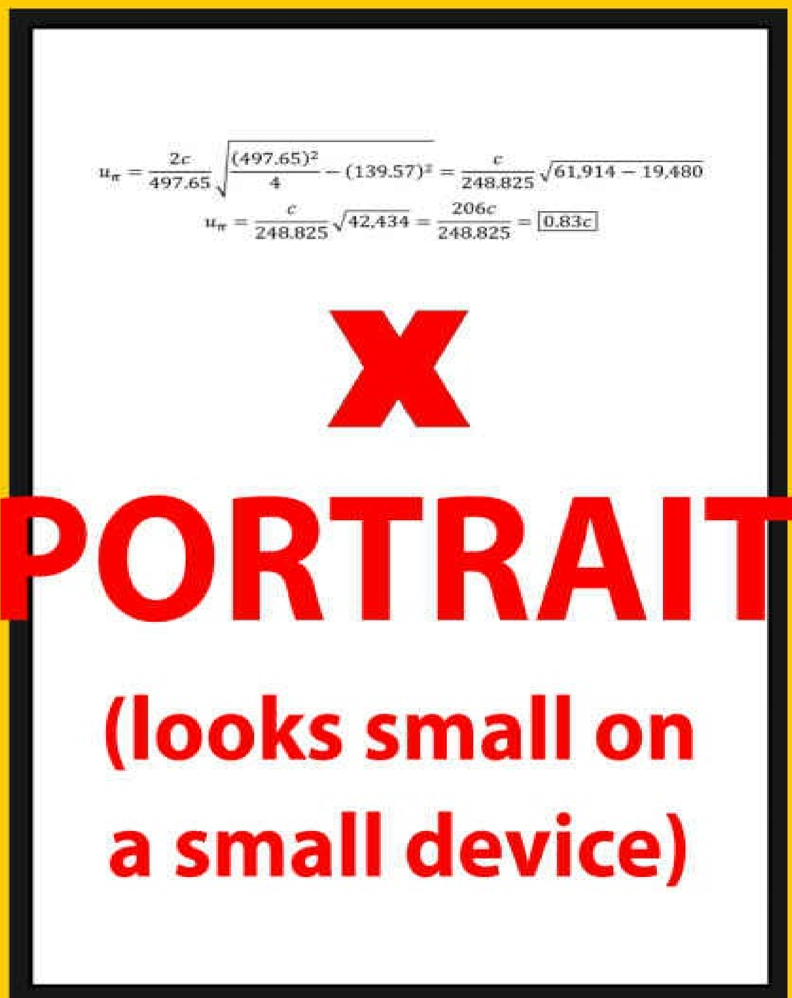
where the uppercase Greek letter psi, $\Psi(x,t)$, is called the **wave function**, $\hbar = \frac{h}{2\pi}$, $h = 6.626 \times 10^{-34}$ J·s is Planck's constant, and m is the mass of the particle. For a given potential energy, $V(x,t)$, Schrödinger's equation can be solved to determine the wave function, $\Psi(x,t)$. As we will explore in Chapter 14, the wave function can be used to calculate probabilities and expectation values, for example. By writing $\Psi(x,t) = \psi(x)\phi(t)$, we can separate variables in Schrödinger's time-dependent equation (like we learned in Chapter 12), which leads to the following **time-independent** form of Schrödinger's equation (as we will see on the next page), provided that the potential energy is independent of time:

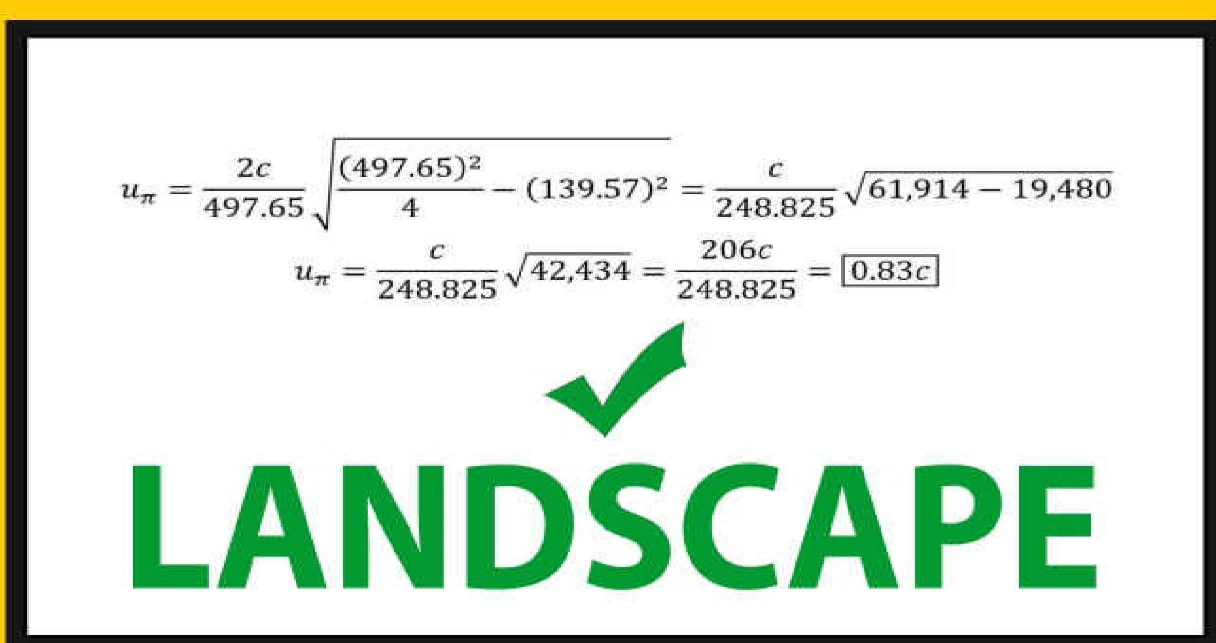
$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

where E is the energy.

TIP FOR READING EQUATIONS

Some equations and paragraphs with equations appear larger in landscape mode.





The Wave Function

According to **Heisenberg's uncertainty principle** (Chapter 11), we can't know exactly where a particle is and its exact momentum at the same instant. What we can do is solve Schrödinger's equation for the wave function, $\Psi(x,t)$, and use the wave function to calculate the probability of the particle having a particular range of positions or range of momentum. The wave function times its **complex conjugate** (page 166), $\Psi^*\Psi$, serves as the **probability density** (Chapter 14).

Separating Variables in Schrödinger's Equation

Let's see how Schrödinger's time-independent equation comes from Schrödinger's time-dependent equation. We begin by writing down Schrödinger's **time-dependent** equation.

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2}+V(x,t)\Psi(x,t)=i\hbar\frac{\partial\Psi(x,t)}{\partial t}$$

In Chapter 12, we saw that a partial differential equation can sometimes be solved by writing a function of two variables as the product of functions of each variable. In this case, we will write $\Psi(x,t)=\psi(x)\phi(t)$. Substitute this into Schrödinger's equation.

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)\phi(t)}{\partial x^2}+V(x,t)\psi(x)\phi(t)=i\hbar\frac{\partial\psi(x)\phi(t)}{\partial t}$$

When we take a partial derivative with respect to x, we treat the independent variable t as if it were a constant, and vice-versa. Therefore, $\frac{\partial^2\psi(x)\phi(t)}{\partial x^2}=\phi(t)\frac{d^2\psi(x)}{dx^2}$ and $\frac{\partial\psi(x)\phi(t)}{\partial t}=\psi(x)\frac{d\phi(t)}{dt}$.

$$-\frac{\hbar^2\phi(t)}{2m}\frac{d^2\psi(x)}{dx^2}+V(x,t)\psi(x)\phi(t)=i\hbar\psi(x)\frac{d\phi(t)}{dt}$$

Divide both sides of the equation by $\psi(x)\phi(t)$.

$$-\frac{\hbar^2}{2m\psi(x)}\frac{d^2\psi(x)}{dx^2}+V(x,t)=\frac{i\hbar}{\phi(t)}\frac{d\phi(t)}{dt}$$

If the potential energy is independent of time, $V(x,t) = V(x)$, then we can separate variables.

$$-\frac{\hbar^2}{2m\psi(x)}\frac{d^2\psi(x)}{dx^2} + V(x) = \frac{i\hbar}{\phi(t)}\frac{d\phi(t)}{dt}$$

Now we have a function of x only equal to a function of t only. The only way that both sides can be equal for all possible values of x and t is if each side equals a constant. If we call that constant E (since it turns out to equal the energy), we get two new differential equations:

$$-\frac{\hbar^2}{2m\psi(x)}\frac{d^2\psi(x)}{dx^2} + V(x) = E \quad , \quad \frac{i\hbar}{\phi(t)}\frac{d\phi(t)}{dt} = E$$

If we multiply the left equation by $\psi(x)$, we get the **time-independent** Schrödinger equation.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

The right equation can be solved immediately. Multiply both sides by $\phi(t)$.

$$i\hbar\frac{d\phi(t)}{dt} = E\phi(t)$$

Divide both sides by $i\hbar$. Note that $\frac{1}{i} = \frac{1}{i}\frac{i}{i} = \frac{i}{i^2} = \frac{i}{-1} = -i$ (see page 166).

$$\frac{d\phi(t)}{dt} = -\frac{iE}{\hbar}\phi(t)$$

It is easy to verify (by substitution into the above equation) that the solution to this equation is:

$$\phi(t) = e^{-iEt/\hbar}$$

Therefore, the solution to the **time-dependent** Schrödinger equation is:

$$\Psi(x,t) = \psi(x)\phi(t) = \psi(x)e^{-iEt/\hbar}$$

Operators, Eigenfunctions, and Eigenvalues

The **Hamiltonian** operator ($\hat{H}$) is defined as:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$$

In terms of the Hamiltonian operator, Schrödinger's time-independent equation is:

$$\hat{H}\psi(x) = E\psi(x)$$

This does **not** state that energy equals the Hamiltonian: The wave function doesn't 'cancel' because the Hamiltonian includes a second derivative. Rather, the allowed energies are the **eigenvalues** of the Hamiltonian operator, and the wave function is the corresponding eigen-function. For a given problem, we solve Schrödinger's second-order differential equation in order to determine the eigenvalues (energy levels) and corresponding eigenfunctions, $\psi(x)$.

For one-dimensional motion along $\pm x$, the **momentum** operator ($\hat{p}$) is defined as:

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}$$

The Hamiltonian operator can thus be expressed in terms of the momentum operator.

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$$

since

$$\frac{\hat{p}^2}{2m} = \frac{1}{2m}(-i\hbar)^2\frac{\partial^2}{\partial x^2} = \frac{1}{2m}(-1)^2 i^2\hbar^2\frac{\partial^2}{\partial x^2} = \frac{1}{2m}(1)(-1)\hbar^2\frac{\partial^2}{\partial x^2} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) = \hat{H}$$

Recall that $i^2 = -1$ (see page 166). Note that kinetic energy can be expressed in terms of momentum as $\hat{K} = \frac{\hat{p}^2}{2m}$. Thus, we see that Schrödinger's equation corresponds to the classical form of the total energy: $\hat{H} = \hat{K} + V$ (kinetic energy plus potential energy).

The **energy** operator ($\hat{E}$) is defined as:

$$\hat{E} = i\hbar\frac{\partial}{\partial t}$$

In terms of the energy operator, Schrödinger's time-dependent equation is:

$$\hat{H}\Psi(x,t) = \hat{E}\Psi(x,t)$$

Recall that $\phi(t) = e^{-iEt/\hbar}$ (page 180). Note that $\varphi(t)$ is an energy eigenfunction.

$$\hat{E}\phi(t) = E\phi(t)$$

since

$$i\hbar\frac{\partial}{\partial t}e^{-iEt/\hbar} = Ee^{-iEt/\hbar}$$

We can alternatively write $\phi(t) = e^{-i\omega t}$ since $E = hf = \hbar\omega$ (see the de Broglie relation in Problem 2 of Chapter 10), where $\omega = 2\pi f$ is the angular frequency and $\hbar = \frac{h}{2\pi}$.

Angular Frequency and Wave Number

Recall that **frequency** (f) is the reciprocal of **period** (T), $f = \frac{1}{T}$, and that **angular frequency** (ω) is defined as:

$$\omega = \frac{2\pi}{T} = 2\pi f$$

Wave number (k) is similarly defined in terms of **wavelength** (λ) as:

$$k = \frac{2\pi}{\lambda}$$

Recall the **de Broglie** relations from Chapter 10: $p = \frac{h}{\lambda}$ and $E = hf$. We can alternatively write these relations as:

$$p = \frac{h}{\lambda} = \frac{hk}{2\pi} = \hbar k \quad , \quad E = hf = \frac{h\omega}{2\pi} = \hbar\omega$$

Schrödinger's Equation in Three Dimensions

For three-dimensional motion, the momentum operator is defined in terms of the gradient (∇) operator:

$$\hat{p} = -i\hbar\nabla$$

where the **gradient** is defined in terms of the Cartesian coordinates as

$$\nabla\psi(x,y,z) = \hat{\mathbf{i}}\frac{\partial\psi(x,y,z)}{\partial x} + \hat{\mathbf{j}}\frac{\partial\psi(x,y,z)}{\partial y} + \hat{\mathbf{k}}\frac{\partial\psi(x,y,z)}{\partial z}$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are unit vectors along the x-, y-, and z-axes. Schrödinger's equation involves the square of the momentum operator, which in turn involves the **Laplacian** operator:

$$\nabla^2\psi(x,y,z) = \frac{\partial^2\psi(x,y,z)}{\partial x^2} + \frac{\partial^2\psi(x,y,z)}{\partial y^2} + \frac{\partial^2\psi(x,y,z)}{\partial z^2}$$

For motion in **three dimensions**, Schrödinger's time-independent equation is (as usual, this form of the equation applies when the potential energy isn't a function of time):

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,y,z)}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2\psi(x,y,z)}{\partial y^2} - \frac{\hbar^2}{2m}\frac{\partial^2\psi(x,y,z)}{\partial z^2} + V(x,y,z)\psi(x,y,z) = E\psi(x,y,z)$$

For some problems, it is more natural to work with spherical coordinates:

$$\nabla^2\psi(r,\theta,\varphi) = \frac{1}{\mathrm{r}^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial\psi(r,\theta,\varphi)}{\partial r}\right] + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial\psi(r,\theta,\varphi)}{\partial\theta}\right] + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi(r,\theta,\varphi)}{\partial\varphi^2}$$

In **spherical coordinates**, Schrödinger's time-independent equation is:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(r,\theta,\varphi) + V(r,\theta,\varphi)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi)$$

Symbols and SI Units

Symbol	Name	SI Units
h	Planck's constant	J·s or J/Hz
$\hbar$	h-bar	J·s or J/Hz
Ψ	wave function for position and time	unitless
ψ	wave function for position	unitless
ϕ	wave function for time	unitless
∇	gradient operator	1/m
∇^2	Laplacian operator	$1/m^2$
i	imaginary number	unitless

V	potential energy	J
E	energy	J
$\hat{E}$	energy operator	J
$\hat{H}$	Hamiltonian operator	J
p	momentum	$\frac{\text{kg} \cdot \text{m}}{\text{s}}$
$\hat{p}$	momentum operator	$\frac{\text{kg} \cdot \text{m}}{\text{s}}$
$\hat{K}$	kinetic energy operator	J
x, y, z	Cartesian coordinates	m
$\hat{\mathbf{i}}, \hat{\mathbf{i}}, \hat{\mathbf{k}}$	Cartesian unit vectors	unitless

r	distance from the origin	m
θ	angle with the z-axis	rad
φ	angle counterclockwise from the x-axis (viewed from $+z$) after projecting onto the xy plane	rad
t	time	s
T	period	s
f	frequency	Hz
ω	angular frequency	rad/s
λ	wavelength	m
k	wave number (or decay constant)	1/m

Note: The symbols Ψ and ψ are the uppercase and lowercase Greek letter psi, ϕ and φ are two variations of the lowercase Greek letter phi, θ is the lowercase Greek letter theta, ω is the lowercase Greek letter omega, and λ is the lowercase Greek letter lambda.

Constants

Quantity	Value
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s
h-bar	$\hbar = 1.055 \times 10^{-34}$ J·s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C

Electron Volts and Angstroms

The conversion from electron Volts (eV) to Joules (J) is:

$$1 \text{ eV} = 1.6021766 \times 10^{-19} \text{ J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1 \text{ J} = \frac{1 \text{ eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18} \text{ eV}$$

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Important Distinction

In earlier chapters, we used U for potential energy and V for electric potential. However, in the context of Schrödinger's equation, it is fairly common to use the symbol V for potential energy. Therefore, in this chapter we are using V for potential energy (instead of U).

Strategy for Applying Schrödinger's Equation

If a problem involves applying Schrödinger's equation, follow these steps:

- Use Schrödinger's time-independent equation (unless V is a function of time).

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad \text{for one-dimensional motion}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z) + V(x,y,z)\psi(x,y,z) = E\psi(x,y,z) \quad \text{for three-dimensional motion}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi(r,\theta,\varphi) + V(r,\theta,\varphi)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi) \quad \text{in spherical coordinates}$$

- Plug the specified **potential energy** into Schrödinger's equation.
- If appropriate, divide the problem into multiple regions. This applies to a square well or to a rectangular barrier, for example.
- Solve the second-order differential equation for each region. It may help to review Chapter 12. Note the following common special cases:

$$\frac{d^2\psi(x)}{dx^2} = k^2\psi(x) \quad \rightarrow \quad \psi(x) = Ae^{kx} + Be^{-kx}$$

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x) \quad \rightarrow \quad \psi(x) = Ae^{ikx} + Be^{-ikx}$$

- For a problem with **multiple regions**, an incident sinusoidal wave is e^{ikx} (if traveling to the right), a reflected sinusoidal wave is e^{-ikx} (if traveling to the left), exponential decay to the right is $e^{-\kappa x}$, and reflected exponential decay to the left is $e^{\kappa x}$. In each region, the coefficients (A, B, C, D, etc.) and constants (k_I, k_{II}, etc.) may be different, as shown in the examples.
- Apply appropriate **boundary conditions**, such as:
  - $\psi(x)$ must be finite because $\int \psi^*(x)\psi(x)\,dx$ represents probability (as we will see in Chapter 14). Check the **finiteness** of $\psi(x)$ as $x \to \pm\infty$, for example.
  - $\psi(x)$ and $\frac{d\psi(x)}{dx}$ must be **continuous** across any boundaries. For example, if regions I and II meet at $x = a$, set $\psi_I(a) = \psi_{II}(a)$ and $\left.\frac{d\psi_I}{dx}\right|_{x=a} = \left.\frac{d\psi_{II}}{dx}\right|_{x=a}$.
- It may help to apply the de Broglie relations:

$$p = \frac{h}{\lambda} = \frac{hk}{2\pi} = \hbar k \quad , \quad E = hf = \frac{h\omega}{2\pi} = \hbar\omega$$

- The **time-dependent** wave function is $\Psi(x,t) = \psi(x)\phi(t) = \psi(x)e^{-iEt/\hbar}$, according to page 180 (provided that V isn't a function of time: for a time-dependent potential energy, it is necessary to solve Schrödinger's time-dependent equation).

- In Chapter 14, we will learn how to **normalize** the wave function: $\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = 1$. This ensures that there is a 100% chance of finding the particle *somewhere*.
- To find probabilities, expectation values, variance, or uncertainty, see Chapter 14.

Example: Solve Schrödinger's time-dependent equation for a free particle traveling in one dimension along $\pm x$.

Begin by solving Schrödinger's time-independent equation. For a free particle, the potential energy is zero: $V(x) = 0$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x)$$

Multiply both sides by $-2m$ and divide both sides by $\hbar^2$.

$$\frac{d^2\psi(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi(x)$$

This equation is like the second special case on page 185, where:

$$k^2 = \frac{2mE}{\hbar^2}$$

$$k = \frac{\sqrt{2mE}}{\hbar}$$

The general solution to this differential equation is:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

The time-dependent solution can now be found (according to page 180):

$$\Psi(x,t) = \psi(x)\phi(t) = \psi(x)e^{-iEt/\hbar} = Ae^{ikx}e^{-iEt/\hbar} + Be^{-ikx}e^{-iEt/\hbar}$$

Since $E = hf = \hbar\omega$ (according to the de Broglie relation), we may write the solution as:

$$\Psi(x,t) = Ae^{ikx}e^{-i\omega t} + Be^{-ikx}e^{-i\omega t} = Ae^{i(kx-\omega t)} + Be^{-i(kx+\omega t)}$$

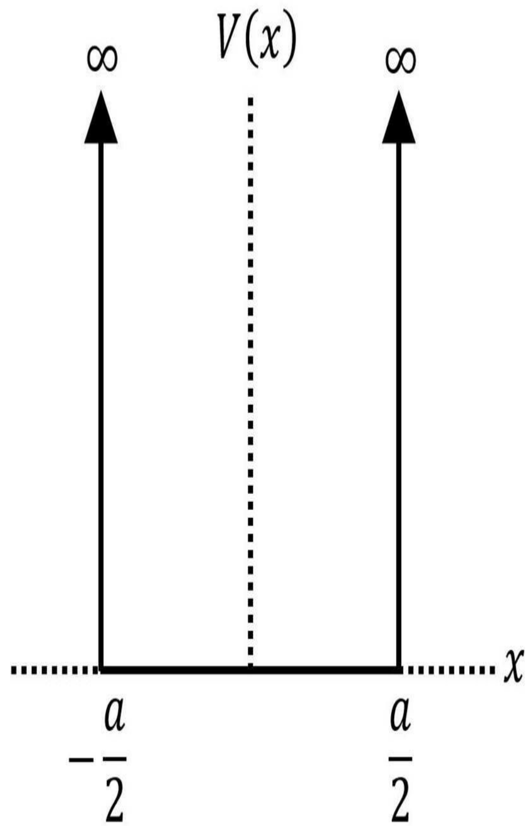
The first term, which is a function of $kx - \omega t$, is characteristic of a wave propagating in the $+x$-direction with a speed equal to $v = \frac{\omega}{k} = \frac{2\pi/T}{2\pi/\lambda} = \frac{\lambda}{T} = \lambda f$, while the second term corresponds to a wave propagating in the $-x$-direction with the same speed.

$$\boxed{\Psi(x,t) = \begin{cases} Ae^{i(kx-\omega t)} & \text{particle traveling along } +x \\ Be^{-i(kx+\omega t)} & \text{particle traveling along } -x \end{cases}}$$

Example: A particle travels in one dimension along $\pm x$ in an infinite square well potential given by:

$$V(x) = \begin{cases} \infty & \text{for } |x| > \dfrac{a}{2} \\ 0 & \text{for } |x| < \dfrac{a}{2} \end{cases}$$

Determine the energy eigenvalues and the corresponding wave functions.



There are three regions of interest:

$$\text{region I} \qquad x < -\frac{a}{2}$$

$$\text{region II} \quad -\frac{a}{2} < x < \frac{a}{2}$$

$$\text{region III} \qquad x > \frac{a}{2}$$

In region II, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = E\psi_{II}(x)$$

In the previous example, we saw that the solution to this differential equation is

$$\psi_{II}(x) = Ae^{ikx} + Be^{-ikx}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}$$

For this problem, it is convenient to write the exponentials in terms of sine and cosine using $e^{i\theta} = \cos\theta + i\sin\theta$ and $e^{-i\theta} = \cos\theta - i\sin\theta$ (see page 166).

$$\psi_{II}(x) = A\cos(kx) + iA\sin(kx) + B\cos(kx) - iB\sin(kx)$$
$$= (A+B)\cos(kx) + i(A-B)\sin(kx)$$

If we define two new constants, $C = A + B$ and $D = i(A - B)$, this becomes:

$$\psi_{II}(x) = C\cos(kx) + D\sin(kx)$$

In regions I and III, $V(x) = \infty$ and $\psi_I(x) = \psi_{III}(x) = 0$ because there is zero probability of finding the particle where the potential energy is infinite. Since the wave function must be continuous across the boundary, set the wave functions equal at the points where they meet:

$$\psi_I\left(-\frac{a}{2}\right) = \psi_{II}\left(-\frac{a}{2}\right) \quad \rightarrow \quad 0 = C\cos\left(-\frac{ka}{2}\right) + D\sin\left(-\frac{ka}{2}\right)$$
$$\psi_{II}\left(\frac{a}{2}\right) = \psi_{III}\left(\frac{a}{2}\right) \quad \rightarrow \quad C\cos\left(\frac{ka}{2}\right) + D\sin\left(\frac{ka}{2}\right) = 0$$

Recall from trig that cosine is an even function of its argument, $\cos(-\theta) = \cos\theta$, whereas sine is an odd function, $\sin(-\theta) = -\sin\theta$. The pair of equations above may thus be written:

$$C\cos\left(\frac{ka}{2}\right) - D\sin\left(\frac{ka}{2}\right) = 0$$

$$C\cos\left(\frac{ka}{2}\right) + D\sin\left(\frac{ka}{2}\right) = 0$$

Add these two equations to get:

$$2C\cos\left(\frac{ka}{2}\right) = 0$$

Subtract the first equation from the second to get:

$$2D\sin\left(\frac{ka}{2}\right) = 0$$

Which values of C, D, and k satisfy both of the equations above? We don't want to set both C and D equal to zero because then the wave function would be exactly zero everywhere (so there would be zero probability of finding it anywhere, which makes no sense). This leaves two ways to satisfy the pair of equations above:

$$D = 0 \quad \text{and} \quad \cos\left(\frac{ka}{2}\right) = 0$$

$$C = 0 \quad \text{and} \quad \sin\left(\frac{ka}{2}\right) = 0$$

Note that the sine function equals zero if the argument equals $\pi, 2\pi, 3\pi, 4\pi$, etc., whereas the cosine function equals zero if the argument equals $\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}$, etc. Since the argument is $\frac{ka}{2}$, we can summarize this concisely as follows:

$$D = 0 \quad \text{and} \quad k_n = \frac{n\pi}{a} \text{ for } n = 1,3,5,7,\cdots$$

$$C = 0 \quad \text{and} \quad k_n = \frac{n\pi}{a} \text{ for } n = 2,4,6,8,\cdots$$

That is, the wave function must have one of the following forms:

$$\boxed{\psi_n(x) = \begin{cases} C\cos\left(\dfrac{n\pi x}{a}\right) & \text{where } n = 1,3,5,7,\cdots \\ D\sin\left(\dfrac{n\pi x}{a}\right) & \text{where } n = 2,4,6,8,\cdots \end{cases}}$$

The energy eigenvalues, E_n, corresponding to these wave functions, $\psi_n(x)$, can be found by equating the expressions $k_n = \frac{\sqrt{2mE_n}}{\hbar}$ (see the previous page) and $k_n = \frac{n\pi}{a}$.

$$\frac{\sqrt{2mE_n}}{\hbar} = \frac{n\pi}{a}$$

Multiply both sides of the equation by $\hbar$.

$$\sqrt{2mE_n} = \frac{n\pi\hbar}{a}$$

Square both sides of the equation and divide both sides by $2m$.

$$\boxed{E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}} \quad , \quad n = 1,2,3,4,5,6,7,8,\cdots$$

The ground state (or zero-point) energy for the one-dimensional infinite square well is:

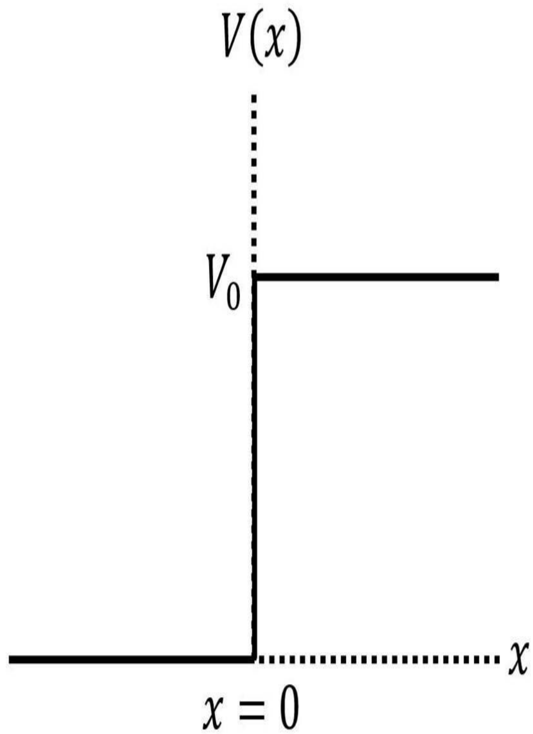
$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2}$$

Note: We excluded $n = 0$ as a possible solution because in that case the wave function ψ would equal $D \sin\left(\frac{0\pi x}{a}\right) = D \sin 0 = 0$ everywhere, which wouldn't make sense: there would be zero probability of finding the particle *anywhere*, since $\int_{x=a}^{b} \psi^* \psi \, dx$ serves as a measure of probability, as we will learn in Chapter 14. We could determine the coefficients C and D by normalizing the wave function (Chapter 14 shows you how to normalize a wave function).

Example: A particle travels in one dimension along $\pm x$ with a step potential given by:

$$V(x) = \begin{cases} V_0 & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

Determine the wave function and reflection coefficient, $R = \frac{B^*B}{A^*A}$, for a particle with $0 < E < V_0$.



There are two regions of interest:

$$\text{region I} \quad x < 0$$
$$\text{region II} \quad x > 0$$

In region I, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} = E\psi_I(x)$$

In the previous examples, we saw that the solution to this differential equation is

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x}$$

where

$$k_1 = \frac{\sqrt{2mE}}{\hbar}$$

In region II, $V(x) = V_0$ equals a positive constant and Schrödinger's equation is.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} + V_0\psi_{II}(x) = E\psi_{II}(x)$$

Subtract $V_0\psi_{II}(x)$ from both sides of the equation and factor out the $\psi_{II}(x)$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = (E - V_0)\psi_{II}(x)$$

Multiply both sides by $-\frac{2m}{\hbar^2}$ and distribute the minus sign to write $-(E - V_0) = V_0 - E$.

$$\frac{d^2\psi_{II}(x)}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi_{II}(x)$$

Since $0 < E < V_0$ in this problem, $V_0 - E > 0$ and both sides of the equation are positive, in contrast to the differential equation for region I. This corresponds to the first special case on page 185, where

$$\psi_{II}(x) = Ce^{k_2 x} + De^{-k_2 x}$$

and

$$k_2 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

In region II, as x grows to infinity, $Ce^{k_2 x}$ becomes infinite. Since $\int \psi^*(x)\psi(x)\,dx$ represents probability and since probabilities must be finite, the coefficient C must equal zero:

$$C = 0 \quad \rightarrow \quad \psi_{II}(x) = De^{-k_2 x}$$

Since the wave function must be continuous across the boundary, set the wave functions equal at the point where they meet (which is $x = 0$):

$$\psi_I(0) = \psi_{II}(0) \quad \rightarrow \quad Ae^0 + Be^{-0} = De^{-0} \quad \rightarrow \quad A + B = D$$

The first derivative of the wave function must also be continuous across the boundary.

$$\left.\frac{d\psi_I}{dx}\right|_{x=0} = \left.\frac{d\psi_{II}}{dx}\right|_{x=0} \quad \to \quad \left.\frac{d}{dx}\left(Ae^{ik_1x} + Be^{-ik_1x}\right)\right|_{x=0} = \left.\frac{d}{dx}\left(De^{-k_2x}\right)\right|_{x=0}$$

$$\left.\left(ik_1Ae^{ik_1x} - ik_1Be^{-ik_1x}\right)\right|_{x=0} = -k_2De^{-k_2x}\big|_{x=0} \quad \to \quad ik_1A - ik_1B = -k_2D$$

$$ik_1(A - B) = -k_2D \quad \to \quad D = -\frac{ik_1}{k_2}(A - B) = \frac{ik_1}{k_2}(B - A)$$

Since $D = A + B$ and $D = \frac{ik_1}{k_2}(B - A)$, we may set these expressions equal:

$$A + B = \frac{ik_1}{k_2}(B - A) \quad \to \quad k_2A + k_2B = ik_1B - ik_1A$$

$$ik_1A + k_2A = ik_1B - k_2B \quad \to \quad (ik_1 + k_2)A = (ik_1 - k_2)B \quad \to \quad B = \frac{ik_1 + k_2}{ik_1 - k_2}A$$

Substitute this expression into the equation $D = A + B$.

$$D = A + B = A + \frac{ik_1 + k_2}{ik_1 - k_2}A = \frac{ik_1 - k_2}{ik_1 - k_2}A + \frac{ik_1 + k_2}{ik_1 - k_2}A = \frac{2ik_1}{ik_1 - k_2}A$$

The wave functions for each region are:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} = \boxed{Ae^{ik_1x} + \frac{ik_1 + k_2}{ik_1 - k_2} Ae^{-ik_1x}}$$

$$\psi_{II}(x) = De^{-k_2x} = \boxed{\frac{2ik_1}{ik_1 - k_2} Ae^{-k_2x}}$$

The term Ae^{ik_1x} represents a particle traveling to the right towards the step potential, the term Be^{-ik_1x} represents a particle traveling to the left (having been reflected from the step potential), and De^{-k_2x} represents the classically forbidden region $x > 0$ where $E < V_0$. The **reflection coefficient** is:

$$R = \frac{B^*B}{A^*A} = \left(\frac{B}{A}\right)^* \frac{B}{A}$$

Recall that $\frac{B}{A} = \frac{ik_1+k_2}{ik_1-k_2}$. To find the complex conjugate, change i to $-i$. That is, $\left(\frac{B}{A}\right)^* = \frac{-ik_1+k_2}{-ik_1-k_2}$.

$$R = \left(\frac{-ik_1 + k_2}{-ik_1 - k_2}\right)\left(\frac{ik_1 + k_2}{ik_1 - k_2}\right) = \frac{-i^2k_1^2 - ik_1k_2 + ik_1k_2 + k_2^2}{-i^2k_1^2 + ik_1k_2 - ik_1k_2 + k_2^2} = \frac{k_1^2 + k_2^2}{k_1^2 + k_2^2} = \boxed{1}$$

Recall that $i^2 = -1$. Since $R = 1$, a particle incident upon the step potential with $E < V_0$ has 100% probability of being reflected.

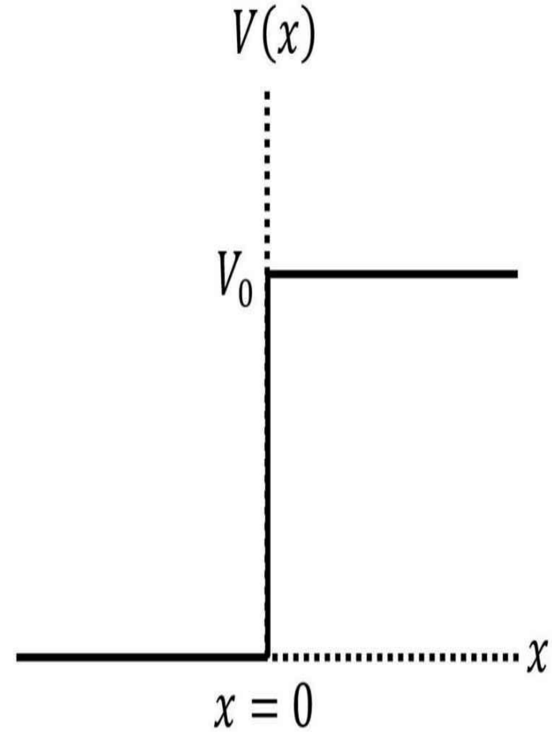
Chapter 13 Problems

1. A particle travels in one dimension along $\pm x$ with a step potential given by:

$$V(x) = \begin{cases} V_0 & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

Assume that the particle begins with $x_i < 0$ and is initially heading to the right.

(A) Determine the wave function in each region for a particle with $E > V_0$.



(B) Determine the reflection coefficient. (See the bottom of page 191.)

(C) Determine the corresponding transmission coefficient, where $T = \frac{k_2 C^* C}{k_1 A^* A}$.

(D) Show that $R + T = 1$.

Want help? Check the solution at the end of the chapter.

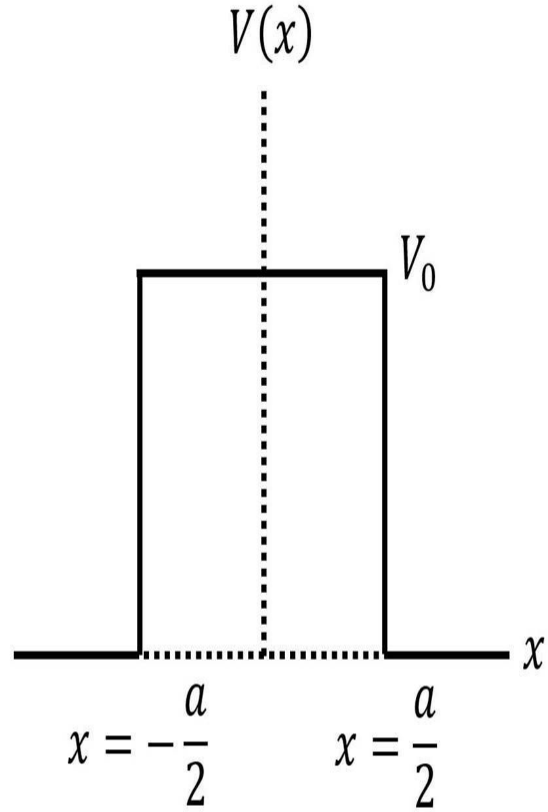
Answers: 1. (A) $\psi_I(x) = Ae^{ik_1 x} + \frac{k_1 - k_2}{k_1 + k_2} Ae^{-ik_1 x}$, $\psi_{II}(x) = \frac{2k_1}{k_1 + k_2} Ae^{ik_2 x}$

(B) $R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2$ (C) $T = \frac{4k_1 k_2}{(k_1 + k_2)^2}$

2. A particle travels in one dimension along $\pm x$ with a barrier potential given by:

$$V(x) = \begin{cases} 0 & \text{for } x < -\frac{a}{2} \\ V_0 & \text{for } -\frac{a}{2} < x < \frac{a}{2} \\ 0 & \text{for } x > \frac{a}{2} \end{cases}$$

Assume that the particle begins with $x_i < -\frac{a}{2}$ and is initially heading to the right.



(A) Write down the wave function in each region for a particle with $0 < E < V_0$ and solve for the ratios of the coefficients: $\frac{F}{A}, \frac{F}{B}, \frac{F}{C}$, and $\frac{F}{D}$ (don't use E for a coefficient, since that's energy).

Hint: In region III, the coefficient of e^{-ik_3x} equals zero based on the given initial conditions.

(B) Determine the transmission coefficient. It may help to review the instructions for Part C of Problem 1.

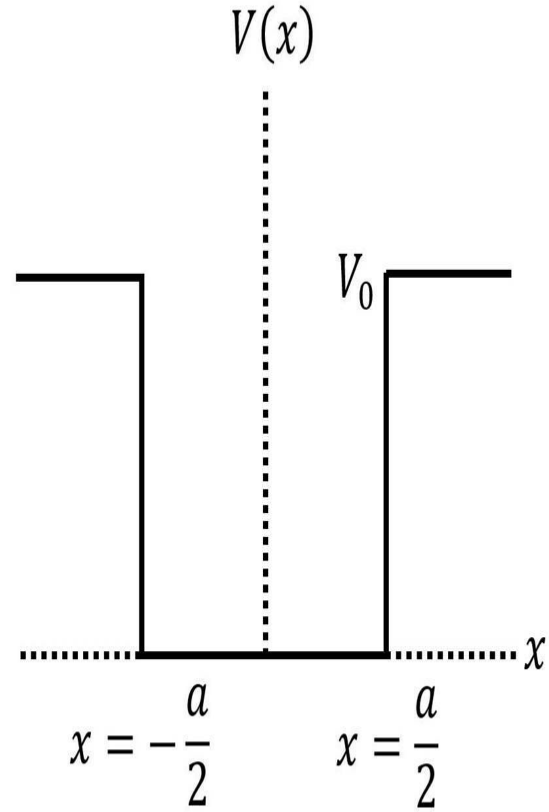
Want help? Check the solution at the end of the chapter.

Answers: 2. (A) Check the end of the chapter. (B) $T = \frac{1}{1+\frac{V_0^2 \sinh^2(2k_2a)}{4E(V_0-E)}}$

---

3. A particle travels in one dimension along $\pm x$ in a finite square well potential given by:

$$V(x) = \begin{cases} V_0 & \text{for } x < -\frac{a}{2} \\ 0 & \text{for } -\frac{a}{2} < x < \frac{a}{2} \\ V_0 & \text{for } x > \frac{a}{2} \end{cases}$$



(A) Write down the wave function in each region for a particle for which $0 < E < V_0$.

(B) Two of the six coefficients equal zero. Which coefficients equal zero and why?

(C) Apply the boundary conditions to obtain 4 equations for the 4 remaining coefficients.

(D) Solve the system of equations in Part C to derive the following equation.

$$(k_1 + ik_2)^2 e^{ik_2 a} = (k_1 - ik_2)^2 e^{-ik_2 a}$$

(E) Show that the equation from Part D has two possible solutions:

$$k_1 = -k_2 \cot\left(\frac{k_2 a}{2}\right) \quad , \quad k_1 = k_2 \tan\left(\frac{k_2 a}{2}\right)$$

(F) Show that the equations in Part E can alternatively be expressed as:

$$\frac{a}{\hbar}\sqrt{\frac{m(V_0 - E)}{2}} = -\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\cot\left(\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\right) \quad , \quad \frac{a}{\hbar}\sqrt{\frac{m(V_0 - E)}{2}} = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}\tan\left(\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\right)$$

(G) Consider the following variables.

$$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}} \quad , \quad v = \frac{a}{\hbar}\sqrt{\frac{mV_0}{2}} \quad , \quad w_1 = u \tan u \quad , \quad w_2 = -u \cot u \quad , \quad y = \sqrt{v^2 - u^2}$$

Use a computer to graph both y and w_1 as functions of u on the same plot. Make a second graph with y and w_2 as functions of u on the same plot. Explore what happens to each set of graphs as you try the following values of v: 2, 4, 6, 8, 10. Sketch each graph for the case $v = 10$ in the space below. Describe your observations for varying the value of v.

(H) Show that the number of energy eigenvalues depends on the value of v as follows.

- For $w_1 = y$, if $(n-1)\pi \le v < n\pi$, there will be n energy eigenvalues.
- For $w_2 = y$, if $(n-1)\pi \le v + \frac{\pi}{2} < n\pi$, there will be n energy eigenvalues.

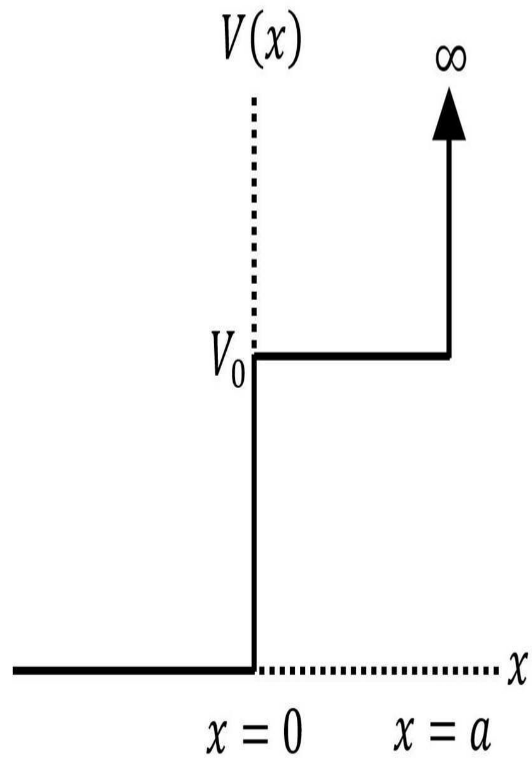
Want help? Check the solution at the end of the chapter.

Answers: The answers to this problem can be found at the end of the chapter.

4. A particle travels in one dimension along $\pm x$ with a barrier potential given by:

$$V(x) = \begin{cases} 0 & \text{for } x < 0 \\ V_0 & \text{for } 0 < x < a \\ \infty & \text{for } x > a \end{cases}$$

Determine the wave function in each region for a particle with $0 < E < V_0$ in terms of a single unknown coefficient.



Want help? Check the solution at the end of the chapter.

Answers: The answers to this problem can be found at the end of the chapter.

Solutions to Chapter 13

1. There are two regions of interest:

$$\text{region I} \quad x < 0$$
$$\text{region II} \quad x > 0$$

(A) In region I, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} = E\psi_I(x)$$

Multiply both sides by $-2m$ and divide both sides by $\hbar^2$.

$$\frac{d^2\psi_I(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi_I(x)$$

This equation is like the second special case on page 185, where:

$$k_1^2 = \frac{2mE}{\hbar^2}$$

$$k_1 = \frac{\sqrt{2mE}}{\hbar}$$

The general solution to this differential equation is:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x}$$

In region II, $V(x) = V_0$ equals a positive constant and Schrödinger's equation is.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} + V_0\psi_{II}(x) = E\psi_{II}(x)$$

Subtract $V_0\psi_{II}(x)$ from both sides of the equation and factor out the $\psi_{II}(x)$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = (E - V_0)\psi_{II}(x)$$

Multiply both sides by $-\frac{2m}{\hbar^2}$ (but unlike the example on pages 190-191, don't distribute the minus sign; leave it out front instead).

$$\frac{d^2\psi_{II}(x)}{dx^2} = -\frac{2m}{\hbar^2}(E - V_0)\psi_{II}(x)$$

Note that $E > V_0$ in this problem, which makes the solution fundamentally different from the example on pages 190-191. Since $E > V_0$, the quantity $E - V_0 > 0$ and the two sides of the equation have opposite sign, similar to the differential equation for region I. This corresponds to the second special case on page 185, where

$$\psi_{II}(x) = Ce^{ik_2x} + De^{-ik_2x}$$

and

$$k_2^2 = \frac{2m(E - V_0)}{\hbar^2}$$

$$k_2 = \frac{\sqrt{2m(E - V_0)}}{\hbar}$$

Since the wave functions involve terms of the form $e^{\pm ikx}$ in both regions, no terms grow to infinity as x approaches positive or negative infinity (unlike the example on pages 190-191 where ψ_{II} involved $e^{\pm kx}$, which is exponential rather than oscillatory behavior: look closely, that i makes a big difference). However, one of the four coefficients does equal zero, but for a different reason than in the example on pages 190-191: The coefficient D must equal zero because the term De^{-ik_2x} represents a wave function that has reflected *after* entering region II (which is impossible based on the initial conditions stated in the problem).

$$D = 0 \quad \rightarrow \quad \psi_{II}(x) = Ce^{ik_2 x}$$

Since the wave function must be continuous across the boundary, set the wave functions equal at the point where they meet (which is $x = 0$):

$$\psi_I(0) = \psi_{II}(0) \quad \rightarrow \quad Ae^0 + Be^{-0} = Ce^0 \quad \rightarrow \quad A + B = C$$

The first derivative of the wave function must also be continuous across the boundary.

$$\left.\frac{d\psi_I}{dx}\right|_{x=0} = \left.\frac{d\psi_{II}}{dx}\right|_{x=0} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ae^{ik_1 x} + Be^{-ik_1 x}\right)\right|_{x=0} = \left.\frac{d}{dx}\left(Ce^{ik_2 x}\right)\right|_{x=0}$$

$$\left.\left(ik_1 Ae^{ik_1 x} - ik_1 Be^{-ik_1 x}\right)\right|_{x=0} = \left.ik_2 Ce^{ik_2 x}\right|_{x=0}$$

$$ik_1 A - ik_1 B = ik_2 C \quad \rightarrow \quad k_1 A - k_1 B = k_2 C \quad \rightarrow \quad C = \frac{k_1}{k_2}(A - B)$$

Since $C = A + B$ and $C = \frac{k_1}{k_2}(A - B)$, we may set these expressions equal:

$$A + B = \frac{k_1}{k_2}(A - B) \quad \rightarrow \quad k_2 A + k_2 B = k_1 A - k_1 B$$

$$k_2 B + k_1 B = k_1 A - k_2 A \quad \rightarrow \quad (k_1 + k_2)B = (k_1 - k_2)A \quad \rightarrow \quad B = \frac{k_1 - k_2}{k_1 + k_2} A$$

Substitute this expression into the equation $C = A + B$.

$$C = A + B = A + \frac{k_1 - k_2}{k_1 + k_2}A = \frac{k_1 + k_2}{k_1 + k_2}A + \frac{k_1 - k_2}{k_1 + k_2}A = \frac{2k_1}{k_1 + k_2}A$$

The wave functions for each region are:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} = \boxed{Ae^{ik_1x} + \frac{k_1 - k_2}{k_1 + k_2}Ae^{-ik_1x}}$$

$$\psi_{II}(x) = Ce^{ik_2x} = \boxed{\frac{2k_1}{k_1 + k_2}Ae^{ik_2x}}$$

(B) The term Ae^{ik_1x} represents a particle traveling to the right towards the step potential, the term Be^{-ik_1x} represents a particle traveling to the left (having been reflected from the step potential), and Ce^{ik_2x} represents the particle traveling to the right after crossing the point $x = 0$ (which is a classically allowed region since $E > V_0$ in this problem). The reflection coefficient is:

$$R = \frac{B^*B}{A^*A} = \left(\frac{B}{A}\right)^* \frac{B}{A}$$

Recall from Part A that $\frac{B}{A} = \frac{k_1 - k_2}{k_1 + k_2}$. In this problem, $\frac{B}{A}$ is real: $\left(\frac{B}{A}\right)^* = \frac{B}{A}$.

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)\left(\frac{k_1 - k_2}{k_1 + k_2}\right) = \boxed{\left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2} = \boxed{\frac{k_1^2 - 2k_1k_2 + k_2^2}{k_1^2 + 2k_1k_2 + k_2^2}}$$

(C) The corresponding transmission coefficient is:

$$T = \frac{k_2 C^* C}{k_1 A^* A} = \frac{k_2}{k_1}\left(\frac{C}{A}\right)^* \frac{C}{A}$$

Recall from Part A that $\frac{C}{A} = \frac{2k_1}{k_1 + k_2}$.

$$T = \frac{k_2}{k_1}\left(\frac{2k_1}{k_1 + k_2}\right)\left(\frac{2k_1}{k_1 + k_2}\right) = \left(\frac{k_2}{k_1}\right)\left(\frac{4k_1^2}{k_1^2 + 2k_1k_2 + k_2^2}\right) = \boxed{\frac{4k_1k_2}{k_1^2 + 2k_1k_2 + k_2^2}} = \boxed{\frac{4k_1k_2}{(k_1 + k_2)^2}}$$

(D) Add the reflection and transmission coefficients together.

$$R + T = \frac{k_1^2 - 2k_1k_2 + k_2^2}{k_1^2 + 2k_1k_2 + k_2^2} + \frac{4k_1k_2}{k_1^2 + 2k_1k_2 + k_2^2} = \frac{k_1^2 + 2k_1k_2 + k_2^2}{k_1^2 + 2k_1k_2 + k_2^2} = \boxed{1}$$

Are you wondering why $T = \frac{k_2 C^* C}{k_1 A^* A}$ includes the ratio $\frac{k_2}{k_1}$, whereas $R = \frac{B^* B}{A^* A}$ does not? Actually, R includes the ratio $\frac{k_1}{k_1} = 1$. The difference is that the reflected wave function $Be^{-ik_1 x}$ and the incident wave function $Ae^{ik_1 x}$ involve the same de Broglie wavelength (since they both involve the same wave number, k_1), whereas the transmitted wave function $Ce^{ik_2 x}$ involves a different wave number, k_2 (corresponding to a different de Broglie wavelength). Since the speed of the particle equals $v = \lambda f = \frac{2\pi f}{k} = \frac{\omega}{k}$ (see page 182), the different wavelength corresponds to a different speed. (You might be reminded of Snell's law in optics, but recall that this problem is purely one-dimensional: The particle's direction doesn't change like it does in refraction.)

2. There are three regions of interest:

$$\text{region I} \qquad x < -\frac{a}{2}$$

$$\text{region II} \quad -\frac{a}{2} < x < \frac{a}{2}$$

$$\text{region III} \qquad x > \frac{a}{2}$$

(A) In regions I and III, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} = E\psi_I(x) \quad , \quad -\frac{\hbar^2}{2m}\frac{d^2\psi_{III}(x)}{dx^2} = E\psi_{III}(x)$$

Multiply both sides by $-2m$ and divide both sides by $\hbar^2$.

$$\frac{d^2\psi_I(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi_I(x) \quad , \quad \frac{d^2\psi_{III}(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi_{III}(x)$$

These equations are like the second special case on page 185, where:

$$k_1^2 = k_3^2 = \frac{2mE}{\hbar^2}$$

$$k_1 = k_3 = \frac{\sqrt{2mE}}{\hbar}$$

The general solution to these differential equations is:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} \quad , \quad \psi_{III}(x) = Fe^{ik_1x} + Ge^{-ik_1x}$$

Note that $k_3 = k_1$ since $V(x) = 0$ in both of these regions. We're reserving coefficients C and D for region II. We're not using E (since that's energy), so the next two letters are F and G. In region II, $V(x) = V_0$ equals a positive constant and Schrödinger's equation is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} + V_0\psi_{II}(x) = E\psi_{II}(x)$$

Subtract $V_0\psi_{II}(x)$ from both sides of the equation and factor out the $\psi_{II}(x)$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = (E - V_0)\psi_{II}(x)$$

Multiply both sides by $-\frac{2m}{\hbar^2}$ and distribute the minus sign to write $-(E - V_0) = V_0 - E$.

$$\frac{d^2\psi_{II}(x)}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi_{II}(x)$$

Since $0 < E < V_0$ in this problem, $V_0 - E > 0$ and both sides of the equation are positive, in contrast to the differential equations for regions I and III. (In this regard, this problem is similar to the example on pages 190-191, but unlike the previous problem.) This corresponds to the first special case on page 185, where

$$\psi_{II}(x) = Ce^{k_2 x} + De^{-k_2 x}$$

and

$$k_2 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

Unlike the example on pages 190-191, neither C nor D equal zero because x is limited to $-\frac{a}{2} < x < \frac{a}{2}$ in region II. However, we can set $G = 0$ (similar to what we did in the solution to the previous problem) because the term $Ge^{-ik_2 x}$ represents a wave function that has reflected *after* entering region III (which is impossible based on the initial conditions stated in the problem).

$$G = 0 \quad \to \quad \psi_{III}(x) = Fe^{ik_1 x}$$

Since the wave function must be continuous across the boundary, set the wave functions equal at the points where they meet ($x = -\frac{a}{2}$ and $x = \frac{a}{2}$):

$$\psi_I\left(-\frac{a}{2}\right) = \psi_{II}\left(-\frac{a}{2}\right) \quad \rightarrow \quad Ae^{-ik_1a/2} + Be^{ik_1a/2} = Ce^{-k_2a/2} + De^{k_2a/2}\ (Eq.1)$$

$$\psi_{II}\left(\frac{a}{2}\right) = \psi_{III}\left(\frac{a}{2}\right) \quad \rightarrow \quad Ce^{k_2a/2} + De^{-k_2a/2} = Fe^{ik_1a/2}\ (Eq.2)$$

The first derivative of the wave function must also be continuous across the boundaries.

$$\left.\frac{d\psi_I}{dx}\right|_{x=-\frac{a}{2}} = \left.\frac{d\psi_{II}}{dx}\right|_{x=-\frac{a}{2}} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ae^{ik_1x} + Be^{-ik_1x}\right)\right|_{x=-\frac{a}{2}} = \left.\frac{d}{dx}\left(Ce^{k_2x} + De^{-k_2x}\right)\right|_{x=-\frac{a}{2}}$$

$$\left.\left(ik_1Ae^{ik_1x} - ik_1Be^{-ik_1x}\right)\right|_{x=-\frac{a}{2}} = \left.\left(k_2Ce^{k_2x} - k_2De^{-k_2x}\right)\right|_{x=-\frac{a}{2}}$$

$$ik_1Ae^{-ik_1a/2} - ik_1Be^{ik_1a/2} = k_2Ce^{-k_2a/2} - k_2De^{k_2a/2}\ (Eq.3)$$

$$\left.\frac{d\psi_{II}}{dx}\right|_{x=\frac{a}{2}} = \left.\frac{d\psi_{III}}{dx}\right|_{x=\frac{a}{2}} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ce^{k_2x} + De^{-k_2x}\right)\right|_{x=\frac{a}{2}} = \left.\frac{d}{dx}\left(Fe^{ik_1x}\right)\right|_{x=\frac{a}{2}}$$

$$\left.\left(k_2Ce^{k_2x} - k_2De^{-k_2x}\right)\right|_{x=\frac{a}{2}} = \left.ik_1Fe^{ik_1x}\right|_{x=\frac{a}{2}}$$

$$k_2Ce^{k_2a/2} - k_2De^{-k_2a/2} = ik_1Fe^{ik_1a/2}\ (Eq.4)$$

The boundary conditions have given us four equations involving five coefficients: A, B, C, D, and F. We can apply algebra to solve for the ratios $\frac{F}{A}, \frac{F}{B}, \frac{F}{C}$, and $\frac{F}{D}$. For convenience, we have numbered the equations of this solution. Multiply Eq.'s 1 and 2 by k_2.

$$k_2 A e^{-ik_1 a/2} + k_2 B e^{ik_1 a/2} = k_2 C e^{-k_2 a/2} + k_2 D e^{k_2 a/2} \ (Eq. 5)$$

$$k_2 C e^{k_2 a/2} + k_2 D e^{-k_2 a/2} = k_2 F e^{ik_1 a/2} \ (Eq. 6)$$

Add Eq.'s 4 and 6 together. Note that $k_2 D e^{-k_2 a/2}$ cancels out.

$$2k_2 C e^{k_2 a/2} = (ik_1 + k_2) F e^{ik_1 a/2}$$

$$\frac{F}{C} = \frac{2k_2}{ik_1 + k_2} \frac{e^{k_2 a/2}}{e^{ik_1 a/2}} = \boxed{\frac{2k_2}{ik_1 + k_2} e^{(k_2 - ik_1)a/2}} = \boxed{\frac{2k_2}{ik_1 + k_2} e^{k_2 a/2} e^{-ik_1 a/2}}$$

We applied the rule $\frac{e^y}{e^z} = e^{y-z}$. Subtract Eq. 6 from Eq. 4. This time $k_2 C e^{k_2 a/2}$ cancels out.

$$-2k_2 D e^{-k_2 a/2} = (ik_1 - k_2) F e^{ik_1 a/2}$$

$$\frac{F}{D} = \frac{-2k_2}{ik_1 - k_2} \frac{e^{-k_2 a/2}}{e^{ik_1 a/2}} = \boxed{-\frac{2k_2}{ik_1 - k_2} e^{-(k_2 + ik_1)a/2}} = \boxed{-\frac{2k_2}{ik_1 - k_2} e^{-k_2 a/2} e^{-ik_1 a/2}}$$

Take the reciprocal of the previous two equations. Apply the rule $\frac{1}{e^y} = e^{-y}$.

$$\frac{C}{F} = \frac{ik_1 + k_2}{2k_2} \frac{1}{e^{k_2 a/2} e^{-ik_1 a/2}} = \frac{ik_1 + k_2}{2k_2} e^{-k_2 a/2} e^{ik_1 a/2} \ (Eq. 7)$$

$$\frac{D}{F} = -\frac{ik_1 - k_2}{2k_2} \frac{1}{e^{-k_2 a/2} e^{-ik_1 a/2}} = -\left(\frac{ik_1 - k_2}{2k_2}\right) e^{k_2 a/2} e^{ik_1 a/2} \ (Eq. 8)$$

Substitute Eq.'s 7 and 8 into Eq.'s 3 and 5.

$$ik_1 A e^{-ik_1 a/2} - ik_1 B e^{ik_1 a/2} = \frac{ik_1 + k_2}{2} F e^{-k_2 a} e^{ik_1 a/2} + \frac{ik_1 - k_2}{2} F e^{k_2 a} e^{ik_1 a/2} \ (Eq. 9)$$

$$k_2 A e^{-ik_1 a/2} + k_2 B e^{ik_1 a/2} = \frac{ik_1 + k_2}{2} F e^{-k_2 a} e^{ik_1 a/2} - \left(\frac{ik_1 - k_2}{2}\right) F e^{k_2 a} e^{ik_1 a/2} \ (Eq. 10)$$

Multiply Eq.'s 9 and 10 by $e^{-ik_1 a/2}$. Note that $e^{ik_1 a/2} e^{-ik_1 a/2} = 1$ and $e^{-ik_1 a/2} e^{-ik_1 a/2} = e^{-ik_1 a}$.

$$ik_1 A e^{-ik_1 a} - ik_1 B = \frac{ik_1 + k_2}{2} F e^{-k_2 a} + \frac{ik_1 - k_2}{2} F e^{k_2 a} \ (Eq. 11)$$

$$k_2 A e^{-ik_1 a} + k_2 B = \frac{ik_1 + k_2}{2} F e^{-k_2 a} + \frac{-ik_1 + k_2}{2} F e^{k_2 a} \ (Eq. 12)$$

Note that $-\left(\frac{ik_1-k_2}{2}\right)=\frac{-ik_1+k_2}{2}$. Multiply Eq. 11 by k_2 and Eq. 12 by ik_1. Recall that $i^2=-1$.

$$ik_1k_2Ae^{-ik_1a}-ik_1k_2B=\frac{ik_1k_2+k_2^2}{2}Fe^{-k_2a}+\frac{ik_1k_2-k_2^2}{2}Fe^{k_2a}\ (Eq.13)$$

$$ik_1k_2Ae^{-ik_1a}+ik_1k_2B=\frac{-k_1^2+ik_1k_2}{2}Fe^{-k_2a}+\frac{k_1^2+ik_1k_2}{2}Fe^{k_2a}\ (Eq.14)$$

It is convenient to factor out ik_1k_2, k_1^2, and k_2^2 on the right-hand side of Eq.'s 13-14.

$$ik_1k_2Ae^{-ik_1a}-ik_1k_2B=ik_1k_2F\frac{e^{-k_2a}+e^{k_2a}}{2}+k_2^2F\frac{e^{-k_2a}-e^{k_2a}}{2}\ (Eq.14)$$

$$ik_1k_2Ae^{-ik_1a}+ik_1k_2B=ik_1k_2F\frac{e^{-k_2a}+e^{k_2a}}{2}+k_1^2F\frac{-e^{-k_2a}+e^{k_2a}}{2}\ (Eq.15)$$

Recall the following hyperbolic cosine ($\cosh y$) and hyperbolic sine ($\sinh y$) identities:

$$\cosh y=\frac{e^y+e^{-y}}{2}\quad,\quad\sinh y=\frac{e^y-e^{-y}}{2}$$

Identify the hyperbolic cosine and hyperbolic sine functions in Eq.'s 14-15.

$$ik_1k_2Ae^{-ik_1a} - ik_1k_2B = ik_1k_2F\cosh(k_2a) - k_2^2F\sinh(k_2a) \ (Eq.16)$$

$$ik_1k_2Ae^{-ik_1a} + ik_1k_2B = ik_1k_2F\cosh(k_2a) + k_1^2F\sinh(k_2a) \ (Eq.17)$$

Note that $\frac{e^{-y}-e^{y}}{2} = -\sinh y$. Add Eq.'s 16 and 17 together. Note that ik_1k_2B cancels out.

$$2ik_1k_2Ae^{-ik_1a} = 2ik_1k_2F\cosh(k_2a) + (k_1^2 - k_2^2)F\sinh(k_2a) \ (Eq.18)$$

$$\frac{F}{A} = \frac{2ik_1k_2e^{-ik_1a}}{2ik_1k_2\cosh(k_2a) + (k_1^2 - k_2^2)\sinh(k_2a)} = \frac{e^{-ik_1a}}{\cosh(k_2a) + \frac{k_1^2 - k_2^2}{2ik_1k_2}\sinh(k_2a)}$$

We divided the numerator and denominator each by $2ik_1k_2$. Note that $\frac{1}{i} = \frac{i}{i^2} = \frac{i}{-1} = -i$.

$$\frac{F}{A} = \boxed{\frac{e^{-ik_1a}}{\cosh(k_2a) - i\frac{k_1^2 - k_2^2}{2k_1k_2}\sinh(k_2a)}} \ (Eq.19)$$

Subtract Eq. 16 from Eq. 17. This time $ik_1k_2Ae^{-ik_1a}$ and $ik_1k_2F\cosh(k_2a)$ cancel out.

$$2ik_1k_2B = (k_1^2 + k_2^2)F\sinh(k_2a) \ (Eq.20)$$

$$\frac{F}{B} = \boxed{\frac{2ik_1k_2}{(k_1^2 + k_2^2)\sinh(k_2a)}} \ (Eq.21)$$

Notational differences: If you are comparing this solution to the solution written by another book or instructor, note the following.

- Some books label the coefficients differently. For example, some books use F and G for region II and C and D in region III. In that case, you would need to swap C and F and swap D and G in order to compare the solutions.
- Some books define a barrier to extend from $x = -a$ to $x = a$ instead of from $x = -\frac{a}{2}$ to $x = \frac{a}{2}$. In that case, there will be a factor of 2 with the a in their final answers. (Yet another option is to define a barrier to extend from $x = 0$ to $x = a$.)
- Some books solve the barrier problem for $E > V_0$, for which the answers and solution are somewhat different. Note that we have solved the problem for $0 < E < V_0$.

(B) The term Ae^{ik_1x} represents a particle traveling to the right towards the barrier potential, and $Fe^{ik_3x} = Fe^{ik_1x}$ represents the particle traveling to the right after passing through the barrier potential (which is a classically forbidden region since $E < V_0$ in this problem). In this problem, the transmission coefficient is (see the solution to Problem 1):

$$T = \frac{k_3 F^* F}{k_1 A^* A} = \frac{k_1}{k_1}\left(\frac{F}{A}\right)^* \frac{F}{A} = \left(\frac{F}{A}\right)^* \frac{F}{A}$$

Recall from Part A that $k_3 = k_1$. To find the complex conjugate, $\left(\frac{F}{A}\right)^*$, replace i with $-i$.

$$T = \left[\frac{e^{ik_1 a}}{\cosh(k_2 a) + i\frac{k_1^2 - k_2^2}{2k_1 k_2}\sinh(k_2 a)}\right]\frac{e^{-ik_1 a}}{\cosh(k_2 a) - i\frac{k_1^2 - k_2^2}{2k_1 k_2}\sinh(k_2 a)}$$

Note that $e^{ik_1 a}e^{-ik_1 a} = 1$. Recall that $i^2 = -1$. Note that the cross-terms cancel out.

$$T = \frac{1}{\cosh^2(k_2 a) + \left(\frac{k_1^2 - k_2^2}{2k_1 k_2}\right)^2 \sinh^2(k_2 a)} = \frac{1}{\cosh^2(k_2 a) + \frac{k_1^4 - 2k_1^2 k_2^2 + k_2^4}{4k_1^2 k_2^2}\sinh^2(k_2 a)}$$

Apply the identity $\cosh^2 y - \sinh^2 y = 1$. (These are hyperbolic functions, **not** ordinary trig functions, which is why this identity is different from the identity $\sin^2\theta + \cos^2\theta = 1$.)

$$T = \frac{1}{1 + \sinh^2 y + \frac{k_1^4 - 2k_1^2 k_2^2 + k_2^4}{4k_1^2 k_2^2}\sinh^2(k_2 a)} = \frac{1}{1 + \left[1 + \frac{k_1^4 - 2k_1^2 k_2^2 + k_2^4}{4k_1^2 k_2^2}\right]\sinh^2(k_2 a)}$$

$$T = \frac{1}{1 + \left[\frac{4k_1^2 k_2^2}{4k_1^2 k_2^2} + \frac{k_1^4 - 2k_1^2 k_2^2 + k_2^4}{4k_1^2 k_2^2}\right]\sinh^2(k_2 a)} = \frac{1}{1 + \left[\frac{k_1^4 + 2k_1^2 k_2^2 + k_2^4}{4k_1^2 k_2^2}\right]\sinh^2(k_2 a)}$$

$$T = \frac{1}{1 + \left(\frac{k_1^2 + k_2^2}{2k_1 k_2}\right)^2 \sinh^2(k_2 a)}$$

Recall from Part A that

$$k_1 = k_3 = \frac{\sqrt{2mE}}{\hbar} \quad , \quad k_2 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

We can use these expressions to rewrite $\left(\frac{k_1^2+k_2^2}{2k_1 k_2}\right)^2$ in terms of E and V_0.

$$\left(\frac{k_1^2 + k_2^2}{2k_1 k_2}\right)^2 = \left[\frac{\frac{2mE}{\hbar^2} + \frac{2m(V_0 - E)}{\hbar^2}}{2\frac{\sqrt{2mE}}{\hbar}\frac{\sqrt{2m(V_0 - E)}}{\hbar}}\right]^2 = \left[\frac{\frac{2mE + 2mV_0 - 2mE}{\hbar^2}}{\frac{2(2m)\sqrt{E(V_0 - E)}}{\hbar^2}}\right]^2 = \left[\frac{\frac{2mV_0}{\hbar^2}}{\frac{4m\sqrt{E(V_0 - E)}}{\hbar^2}}\right]^2$$

The way to divide by a fraction is to multiply by its reciprocal.

$$\left(\frac{k_2^2 + k_1^2}{2k_1 k_2}\right)^2 = \left(\frac{2mV_0}{\hbar^2}\frac{\hbar^2}{4m\sqrt{E(V_0 - E)}}\right)^2 = \left(\frac{2mV_0}{4m\sqrt{E(V_0 - E)}}\right)^2 = \frac{4m^2V_0^2}{16m^2E(V_0 - E)} = \frac{V_0^2}{4E(V_0 - E)}$$

Substitute the previous expression into the equation for the transmission coefficient.

$$\boxed{T = \frac{1}{1 + \frac{V_0^2}{4E(V_0 - E)}\sinh^2(k_2 a)}}$$

Thus, there is a nonzero probability that the particle will appear on the opposite side of the barrier potential. This phenomenon is called **tunneling**.

3. There are three regions of interest:

$$\text{region I} \qquad x < -\frac{a}{2}$$

$$\text{region II} \quad -\frac{a}{2} < x < \frac{a}{2}$$

$$\text{region III} \qquad x > \frac{a}{2}$$

(A) In region II, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = E\psi_{II}(x)$$

Multiply both sides by $-2m$ and divide both sides by $\hbar^2$.

$$\frac{d^2\psi_{II}(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi_{II}(x)$$

This equation is like the second special case on page 185, where:

$$k_2^2 = \frac{2mE}{\hbar^2}$$

$$k_2 = \frac{\sqrt{2mE}}{\hbar}$$

The general solution to this differential equations is:

$$\boxed{\psi_{II}(x) = Ce^{ik_2x} + De^{-ik_2x}}$$

In regions I and III, $V(x) = V_0$ equals a positive constant and Schrödinger's equation is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} + V_0\psi_I(x) = E\psi_I(x) \quad , \quad -\frac{\hbar^2}{2m}\frac{d^2\psi_{III}(x)}{dx^2} + V_0\psi_{III}(x) = E\psi_{III}(x)$$

Subtract $V_0\psi(x)$ from both sides of the equation and factor out the $\psi(x)$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} = (E - V_0)\psi_I(x) \quad , \quad -\frac{\hbar^2}{2m}\frac{d^2\psi_{III}(x)}{dx^2} = (E - V_0)\psi_{III}(x)$$

Multiply both sides by $-\frac{2m}{\hbar^2}$ and distribute the minus sign to write $-(E - V_0) = V_0 - E$.

$$\frac{d^2\psi_I(x)}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi_I(x) \quad , \quad \frac{d^2\psi_{III}(x)}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi_{III}(x)$$

Since $0 < E < V_0$ in this problem, $V_0 - E > 0$ and both sides of the equation are positive, in contrast to the differential equation for region II. This corresponds to the first special case on page 185, where

$$k_1 = k_3 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

and

$$\boxed{\psi_I(x) = Ae^{k_1 x} + Be^{-k_1 x}} \quad , \quad \boxed{\psi_{III}(x) = Fe^{k_1 x} + Ge^{-k_1 x}}$$

(B) In region III, as x grows to infinity, $Fe^{k_2 x}$ becomes infinite. Similarly, in region I, as x grows to negative infinity, $Be^{-k_1 x}$ becomes infinite. Since $\int \psi^*(x)\psi(x)\,dx$ represents probability and since probabilities must be finite, the coefficients F and B must equal zero:

$$\boxed{F = 0} \quad \to \quad \psi_{III}(x) = Ge^{-k_1 x}$$

$$\boxed{B = 0} \quad \to \quad \psi_I(x) = Ae^{k_1 x}$$

(C) Since the wave function must be continuous across the boundary, set the wave functions equal at the points where they meet ($x = -\frac{a}{2}$ and $x = \frac{a}{2}$):

$$\psi_I\left(-\frac{a}{2}\right) = \psi_{II}\left(-\frac{a}{2}\right) \quad \rightarrow \quad \boxed{Ae^{-k_1 a/2} = Ce^{-ik_2 a/2} + De^{ik_2 a/2}} \; (Eq. 1)$$

$$\psi_{II}\left(\frac{a}{2}\right) = \psi_{III}\left(\frac{a}{2}\right) \quad \rightarrow \quad \boxed{Ce^{ik_2 a/2} + De^{-ik_2 a/2} = Ge^{-k_1 a/2}} \; (Eq. 2)$$

The first derivative of the wave function must also be continuous across the boundaries.

$$\left.\frac{d\psi_I}{dx}\right|_{x=-\frac{a}{2}} = \left.\frac{d\psi_{II}}{dx}\right|_{x=-\frac{a}{2}} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ae^{k_1 x}\right)\right|_{x=-\frac{a}{2}} = \left.\frac{d}{dx}\left(Ce^{ik_2 x} + De^{-ik_2 x}\right)\right|_{x=-\frac{a}{2}}$$

$$k_1 Ae^{k_1 x}\Big|_{x=-\frac{a}{2}} = \left.\left(ik_2 Ce^{ik_2 x} - ik_2 De^{-ik_2 x}\right)\right|_{x=-\frac{a}{2}}$$

$$\boxed{k_1 Ae^{-k_1 a/2} = ik_2 Ce^{-ik_2 a/2} - ik_2 De^{ik_2 a/2}} \; (Eq. 3)$$

$$\left.\frac{d\psi_{II}}{dx}\right|_{x=\frac{a}{2}} = \left.\frac{d\psi_{III}}{dx}\right|_{x=\frac{a}{2}} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ce^{ik_2 x} + De^{-ik_2 x}\right)\right|_{x=\frac{a}{2}} = \left.\frac{d}{dx}\left(Ge^{-k_1 x}\right)\right|_{x=\frac{a}{2}}$$

$$\left.\left(ik_2 Ce^{ik_2 x} - ik_2 De^{-ik_2 x}\right)\right|_{x=\frac{a}{2}} = -k_1 Ge^{-k_1 x}\Big|_{x=\frac{a}{2}}$$

$$\boxed{ik_2 Ce^{ik_2 a/2} - ik_2 De^{-ik_2 a/2} = -k_1 Ge^{-k_1 a/2}} \; (Eq. 4)$$

(D) The boundary conditions have given us four equations involving four coefficients: A, C, D, and G. We will attempt to solve this system of equations. Multiply Eq.'s 1 and 2 by ik_2.

$$ik_2Ae^{-k_1a/2} = ik_2Ce^{-ik_2a/2} + ik_2De^{ik_2a/2} \; (Eq.5)$$

$$ik_2Ce^{ik_2a/2} + ik_2De^{-ik_2a/2} = ik_2Ge^{-k_1a/2} \; (Eq.6)$$

Add Eq.'s 3 and 5 together. Note that $ik_2De^{ik_2a/2}$ cancels out.

$$(k_1 + ik_2)Ae^{-k_1a/2} = 2ik_2Ce^{-ik_2a/2} \quad \rightarrow \quad C = \frac{k_1 + ik_2}{2ik_2}Ae^{-k_1a/2}e^{ik_2a/2} \; (Eq.7)$$

Subtract Eq. 5 from Eq. 3. This time $ik_2Ce^{-ik_2a/2}$ cancels out.

$$(k_1 - ik_2)Ae^{-k_1a/2} = -2ik_2De^{ik_2a/2} \quad \rightarrow \quad D = \frac{k_1 - ik_2}{-2ik_2}Ae^{-k_1a/2}e^{-ik_2a/2} \; (Eq.8)$$

Add Eq.'s 4 and 6 together. Note that $ik_2De^{-ik_2a/2}$ cancels out.

$$-2ik_2Ce^{ik_2a/2} = (k_1 - ik_2)Ge^{-k_1a/2} \quad \rightarrow \quad C = \frac{k_1 - ik_2}{-2ik_2}Ge^{-k_1a/2}e^{-ik_2a/2} \; (Eq.9)$$

Subtract Eq. 4 from Eq. 6. This time $ik_2Ce^{ik_2a/2}$ cancels out.

$$2ik_2De^{-ik_2a/2} = (k_1 + ik_2)Ge^{-k_1a/2} \quad \rightarrow \quad D = \frac{k_1 + ik_2}{2ik_2}Ge^{-k_1a/2}e^{ik_2a/2} \; (Eq.10)$$

Set Eq.'s 7 and 9 equal to each other.

$$\frac{k_1 + ik_2}{2ik_2} A e^{-k_1 a/2} e^{ik_2 a/2} = \frac{k_1 - ik_2}{-2ik_2} G e^{-k_1 a/2} e^{-ik_2 a/2}$$

$$(k_1 + ik_2) A e^{ik_2 a/2} = -(k_1 - ik_2) G e^{-ik_2 a/2} \quad \rightarrow \quad (k_1 + ik_2) A e^{ik_2 a} = -(k_1 - ik_2) G \ (Eq. 11)$$

We applied the rule $\frac{e^y}{e^{-y}} = e^{y-(-y)} = e^{y+y} = e^{2y}$. Set Eq.'s 8 and 10 equal to each other.

$$\frac{k_1 - ik_2}{-2ik_2} A e^{-k_1 a/2} e^{-ik_2 a/2} = \frac{k_1 + ik_2}{2ik_2} G e^{-k_1 a/2} e^{ik_2 a/2}$$

$$-(k_1 - ik_2) A e^{-ik_2 a/2} = (k_1 + ik_2) G e^{ik_2 a/2} \quad \rightarrow \quad -(k_1 - ik_2) A e^{-ik_2 a} = (k_1 + ik_2) G \ (Eq. 12)$$

We again applied the rule $\frac{e^y}{e^{-y}} = e^{y-(-y)} = e^{y+y} = e^{2y}$. Divide Eq. 11 by Eq. 12.

$$-\frac{k_1 + ik_2}{k_1 - ik_2} \frac{e^{ik_2 a}}{e^{-ik_2 a}} = -\frac{k_1 - ik_2}{k_1 + ik_2} \quad \rightarrow \quad \boxed{(k_1 + ik_2)^2 e^{ik_2 a} = (k_1 - ik_2)^2 e^{-ik_2 a}} \ (Eq. 13)$$

(E) Squareroot both sides of the equation. Note that $\sqrt{e^y} = (e^y)^{1/2} = e^{y/2}$ since $\left(e^{y/2}\right)^2 = e^{y/2}e^{y/2} = e^y$ and that $\sqrt{u^2} = \pm u$ because $(-u)^2 = u^2$. We must consider both possible roots.

$$(k_1 + ik_2)e^{ik_2a/2} = \pm(k_1 - ik_2)e^{-ik_2a/2} \; (Eq. 14)$$

For the positive root, when we distribute and regroup the terms, we get:

$$k_1\left(e^{ik_2a/2} - e^{-ik_2a/2}\right) = -ik_2\left(e^{ik_2a/2} + e^{-ik_2a/2}\right) \; (Eq. 15a)$$

For the negative root, when we distribute and regroup the terms, we get:

$$k_1\left(e^{ik_2a/2} + e^{-ik_2a/2}\right) = -ik_2\left(e^{ik_2a/2} - e^{-ik_2a/2}\right) \; (Eq. 15b)$$

Apply the identities $\cos z = \frac{e^{iz}+e^{-iz}}{2}$ and $\sin z = \frac{e^{iz}-e^{-iz}}{2i}$ to Eq.'s 15a and 15b.

$$k_1 2i \sin\left(\frac{k_2a}{2}\right) = -ik_2 2\cos\left(\frac{k_2a}{2}\right) \quad \text{or} \quad k_1 2\cos\left(\frac{k_2a}{2}\right) = -ik_2 2i\sin\left(\frac{k_2a}{2}\right) \; (Eq. 16)$$

Divide both sides of the first equation by $2i$ and the second equation by 2. Note that $i^2 = -1$.

$$k_1 \sin\left(\frac{k_2a}{2}\right) = -k_2\cos\left(\frac{k_2a}{2}\right) \quad \text{or} \quad k_1\cos\left(\frac{k_2a}{2}\right) = k_2\sin\left(\frac{k_2a}{2}\right) \; (Eq. 17)$$

Recall from trig that $\frac{\sin\theta}{\cos\theta} = \tan\theta$ and $\frac{\cos\theta}{\sin\theta} = \cot\theta$.

$$\boxed{k_1 = -k_2\cot\left(\frac{k_2a}{2}\right)} \quad \text{or} \quad \boxed{k_1 = k_2\tan\left(\frac{k_2a}{2}\right)} \; (Eq. 18)$$

(F) Recall from Part A that:

$$k_1 = k_3 = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \quad , \quad k_2 = \frac{\sqrt{2mE}}{\hbar}$$

Substitute these expressions for k_1 and k_2 into Eq. 18.

$$\frac{\sqrt{2m(V_0 - E)}}{\hbar} = -\frac{\sqrt{2mE}}{\hbar}\cot\left(\frac{a}{2}\frac{\sqrt{2mE}}{\hbar}\right) \quad , \quad \frac{\sqrt{2m(V_0 - E)}}{\hbar} = \frac{\sqrt{2mE}}{\hbar}\tan\left(\frac{a}{2}\frac{\sqrt{2mE}}{\hbar}\right) \quad (Eq. 19)$$

Multiply both sides of the equation by $\frac{a}{2}$.

$$\frac{a\sqrt{2m(V_0 - E)}}{2\hbar} = -\frac{a\sqrt{2mE}}{2\hbar}\cot\left(\frac{a}{2}\frac{\sqrt{2mE}}{\hbar}\right) \quad , \quad \frac{a\sqrt{2m(V_0 - E)}}{2\hbar} = \frac{a\sqrt{2mE}}{2\hbar}\tan\left(\frac{a}{2}\frac{\sqrt{2mE}}{\hbar}\right) \quad (Eq. 20)$$

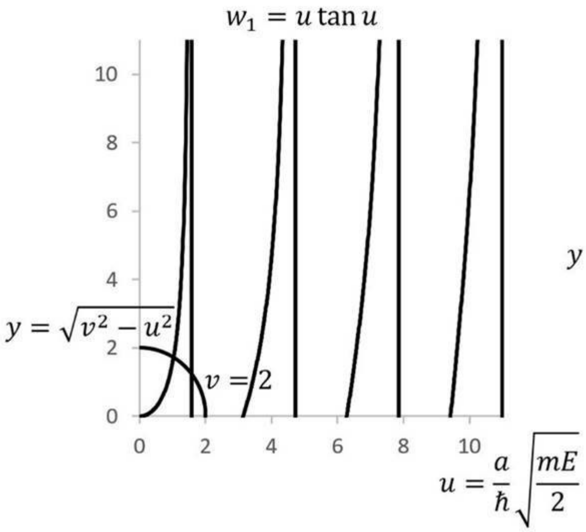
Note that $\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{\sqrt{2^2}} = \sqrt{\frac{2}{2^2}} = \sqrt{\frac{1}{2}}$. Therefore, $\frac{a\sqrt{2mE}}{2\hbar} = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$ and $\frac{a\sqrt{2m(V_0-E)}}{2\hbar} = \frac{a}{\hbar}\sqrt{\frac{m(V_0-E)}{2}}$.

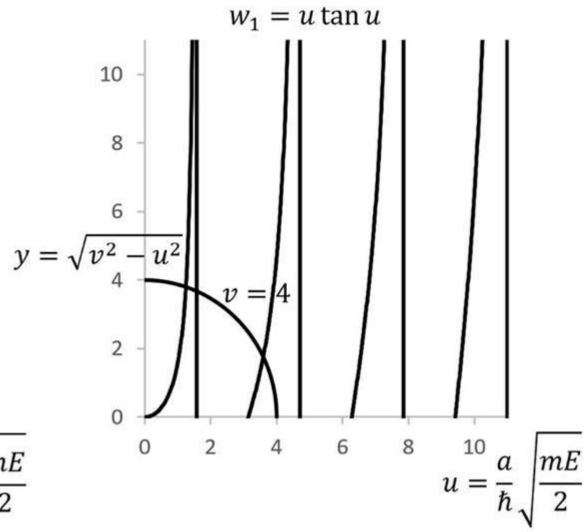
$$\boxed{\frac{a}{\hbar}\sqrt{\frac{m(V_0 - E)}{2}} = -\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\cot\left(\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\right)} \quad , \quad \boxed{\frac{a}{\hbar}\sqrt{\frac{m(V_0 - E)}{2}} = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}\tan\left(\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\right)} \quad (Eq. 21)$$

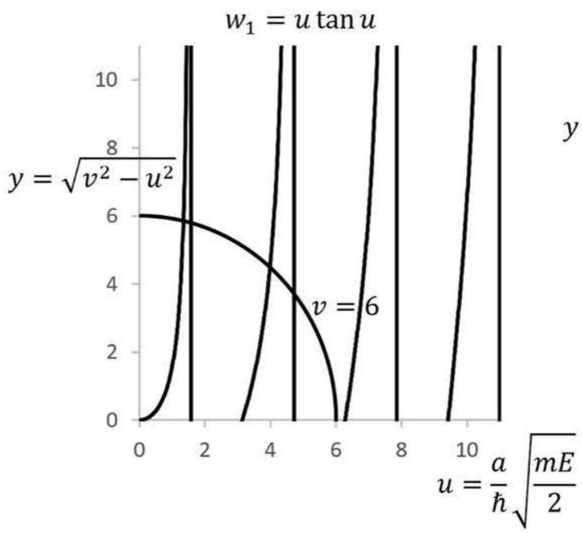
(G) Note the following:

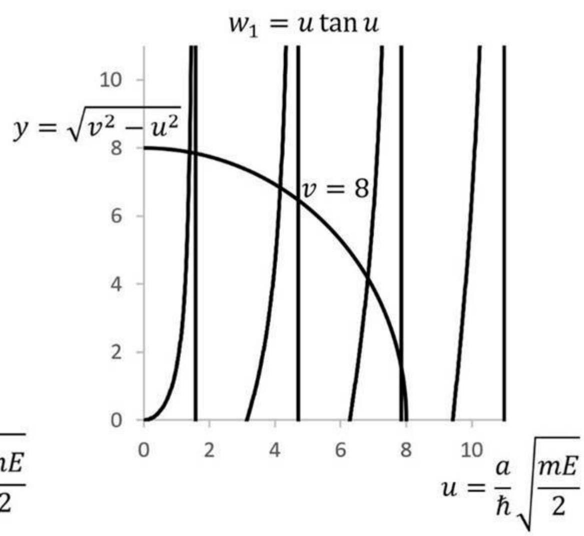
$$y = \sqrt{v^2 - u^2} = \sqrt{\left(\frac{a}{\hbar}\sqrt{\frac{mV_0}{2}}\right)^2 - \left(\frac{a}{\hbar}\sqrt{\frac{mE}{2}}\right)^2} = \sqrt{\frac{a^2 mV_0}{\hbar^2 2} - \frac{a^2 mE}{\hbar^2 2}} = \frac{a}{\hbar}\sqrt{\frac{m(V_0 - E)}{2}}$$

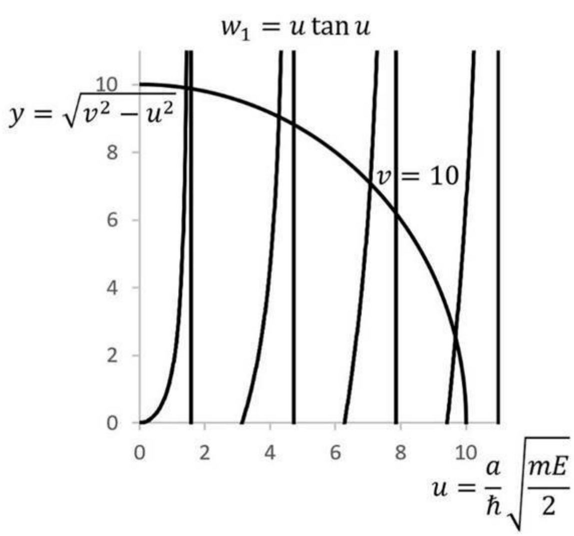
Therefore, the given equations, $w_1 = y = u\tan u$ and $w_2 = y = -u\cot u$, are simply Eq. 21 written in a different form.

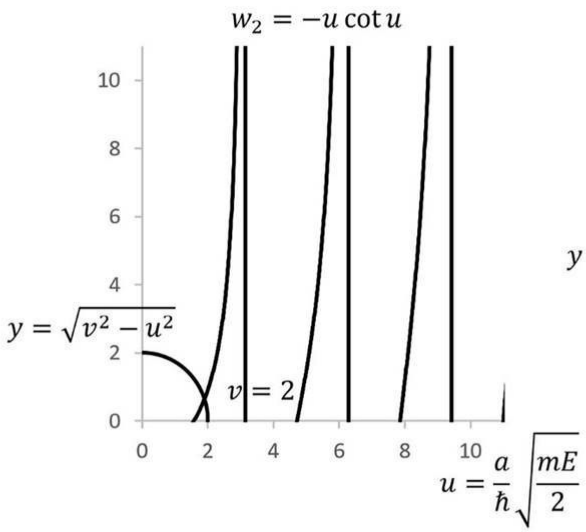

$w_1 = u \tan u$
$y = \sqrt{v^2 - u^2}$
$v = 2$
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$


$w_1 = u \tan u$
$y = \sqrt{v^2 - u^2}$
$v = 4$
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$

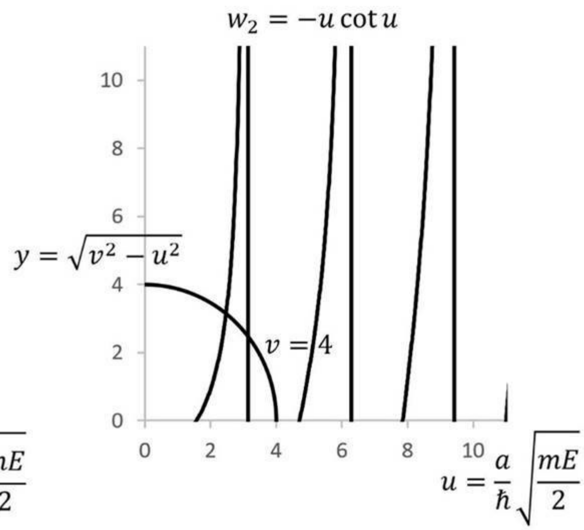

$w_1 = u \tan u$
$y = \sqrt{v^2 - u^2}$
$v = 6$
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$


$w_1 = u \tan u$
$y = \sqrt{v^2 - u^2}$
$v = 8$
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$

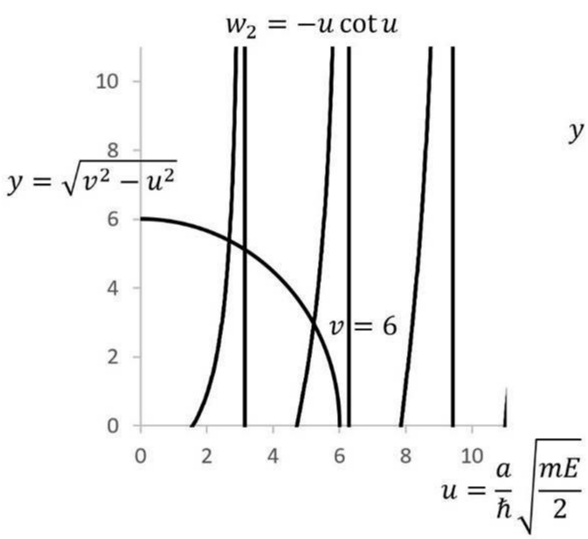

$w_1 = u \tan u$
$y = \sqrt{v^2 - u^2}$
$v = 10$
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$



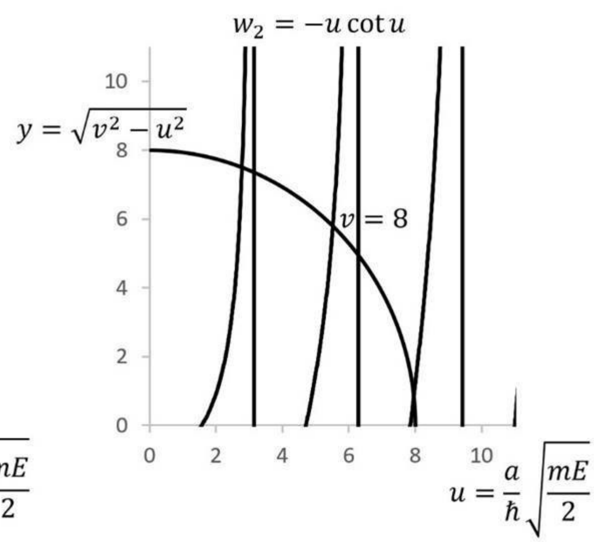
$w_2 = -u \cot u$
$y = \sqrt{v^2 - u^2}$
$v = 2$
0 2 4 6 8 10
0 2 4 6 8 10
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$



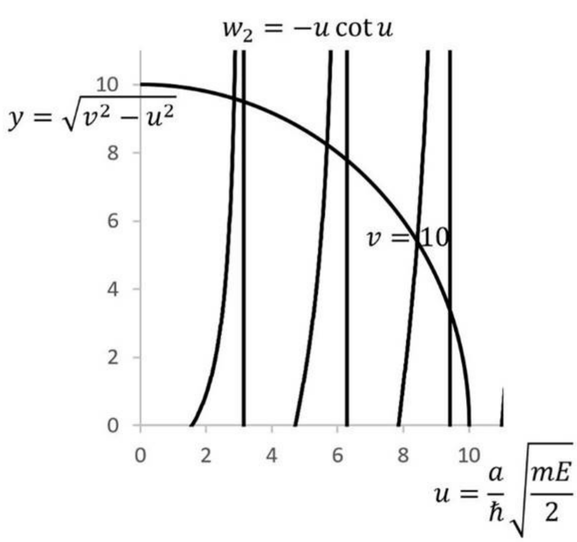
$w_2 = -u \cot u$
$y = \sqrt{v^2 - u^2}$
$v = 4$
0 2 4 6 8 10
0 2 4 6 8 10
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$



$w_2 = -u \cot u$
$y = \sqrt{v^2 - u^2}$
$v = 6$
0 2 4 6 8 10
0 2 4 6 8 10
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$



$w_2 = -u \cot u$
$y = \sqrt{v^2 - u^2}$
$v = 8$
0 2 4 6 8 10
0 2 4 6 8 10
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$



$w_2 = -u \cot u$
$y = \sqrt{v^2 - u^2}$
$v = 10$
0 2 4 6 8 10
0 2 4 6 8 10
$u = \frac{a}{\hbar}\sqrt{\frac{mE}{2}}$

Note that Eq. 21 is a transcendental equation (you can't solve for E algebraically by isolating an unknown). Part G is showing you how to graph the solutions to Eq. 21 on a computer. The points where y intersects with w_1 or w_2 are the solutions to Eq. 21. The graphs of y and w_1 and of y and w_2 are shown on the previous pages for v equal to 2, 4, 6, 8, and 10.

Observe the following from the previous sets of graphs:

- $y = \sqrt{v^2 - u^2}$ is the equation of a circle with a radius of v (since $y^2 + u^2 = v^2$). The effect of increasing v from 2 to 10 is to increase the radius of this circle.
- $w_1 = u \tan u$ has vertical asymptotes at $u = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}$, etc. These are shown as vertical solid lines in the previous diagrams.
- $w_2 = -u \cot u$ has vertical asymptotes at $u = \pi, 2\pi, 3\pi$, etc. These are shown as vertical solid lines in the previous diagrams.
- The number of times that the circle, $y = \sqrt{v^2 - u^2}$, intersects the graph of $w_1 = u \tan u$ or the graph of $w_2 = -u \cot u$ (but **don't** count where the circle intersects the vertical asymptotes) equals the number of solutions to the equation $u \tan u = \sqrt{v^2 - u^2}$ or the equation $-u \cot u = \sqrt{v^2 - u^2}$.

- For $v = 2$, there is 1 solution to $u \tan u = \sqrt{v^2 - u^2}$ and $-u \cot u = \sqrt{v^2 - u^2}$.
- For $v = 4$, there are 2 solutions to $u \tan u = \sqrt{v^2 - u^2}$ and there is 1 solution to $-u \cot u = \sqrt{v^2 - u^2}$.
- For $v = 6$, there are 2 solutions to $u \tan u = \sqrt{v^2 - u^2}$ and $-u \cot u = \sqrt{v^2 - u^2}$.
- For $v = 8$, there are 3 solutions to $u \tan u = \sqrt{v^2 - u^2}$ and $-u \cot u = \sqrt{v^2 - u^2}$.
- For $v = 10$, there are 4 solutions to $u \tan u = \sqrt{v^2 - u^2}$ and there are 3 solutions to $-u \cot u = \sqrt{v^2 - u^2}$.

(H) The number of energy eigenvalues equals the number of solutions to the equation $u \tan u = \sqrt{v^2 - u^2}$ or the equation $-u \cot u = \sqrt{v^2 - u^2}$, which (as we noted in Part G) equals the number of points where the circle $y = \sqrt{v^2 - u^2}$ intersects the graph of $w_1 = u \tan u$ or the graph of $w_2 = -u \cot u$ (**not** counting where the circle intersects the vertical asymptotes). The number of points of intersection depends on when the graph of $w_1 = u \tan u$ and the graph of $w_2 = -u \cot u$ become positive.

- $w_1 = u \tan u$ becomes positive at $u = \pi, 2\pi, 3\pi$, etc. (It is negative just before these.)
- $w_2 = -u \cot u$ becomes positive at $u = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}$, etc. (It is negative just before these.)

How many energy eigenvalues (corresponding to points of intersection) are there?

- For the case $w_1 = u \tan u = y = \sqrt{v^2 - u^2}$:
  - If $v < \pi$, there is 1 solution to $u \tan u = \sqrt{v^2 - u^2}$.
  - If $\pi \le v < 2\pi$, there are 2 solutions to $u \tan u = \sqrt{v^2 - u^2}$.
  - If $2\pi \le v < 3\pi$, there are 3 solutions to $u \tan u = \sqrt{v^2 - u^2}$.
  - If $3\pi \le v < 4\pi$, there are 4 solutions to $u \tan u = \sqrt{v^2 - u^2}$.

- For the case $w_2 = -u \cot u = y = \sqrt{v^2 - u^2}$:
  - If $v < \frac{\pi}{2}$, there are no solutions to $-u \cot u = \sqrt{v^2 - u^2}$.
  - If $\frac{\pi}{2} \leq v < \frac{3\pi}{2}$, there is 1 solution to $-u \cot u = \sqrt{v^2 - u^2}$.
  - If $\frac{3\pi}{2} \leq v < \frac{5\pi}{2}$, there are 2 solutions to $-u \cot u = \sqrt{v^2 - u^2}$.
  - If $\frac{5\pi}{2} \leq v < \frac{7\pi}{2}$, there are 3 solutions to $-u \cot u = \sqrt{v^2 - u^2}$.

We can summarize this more concisely as follows.

- For $w_1 = y$, if $(n-1)\pi \leq v < n\pi$, there will be n energy eigenvalues. For example, for $n = 3$, there are 3 solutions when $2\pi \leq v < 3\pi$.
- For $w_2 = y$, if $(n-1)\pi \leq v - \frac{\pi}{2} < n\pi$, there will be n energy eigenvalues. For example, for $n = 2$, there are 2 solutions when $\pi \leq v - \frac{\pi}{2} < 2\pi$. Add $\frac{\pi}{2}$ to each side of the inequalities to see that $\pi + \frac{\pi}{2} \leq v < 2\pi + \frac{\pi}{2}$, which equates to $\frac{3\pi}{2} \leq v < \frac{5\pi}{2}$. You can see that this agrees with the previous set of bullet points.

4. There are three regions of interest:

$$\begin{array}{ll} \text{region I} & x < 0 \\ \text{region II} & 0 < x < a \\ \text{region III} & x > a \end{array}$$

In region I, $V(x) = 0$ and Schrödinger's time-independent equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I(x)}{dx^2} = E\psi_I(x)$$

Multiply both sides by $-2m$ and divide both sides by $\hbar^2$.

$$\frac{d^2\psi_I(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi_I(x)$$

This equation is like the second special case on page 185, where:

$$k_1^2 = \frac{2mE}{\hbar^2}$$

$$k_1 = \frac{\sqrt{2mE}}{\hbar}$$

The general solution to this differential equations is:

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x}$$

In region II, $V(x) = V_0$ equals a positive constant and Schrödinger's equation is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} + V_0\psi_{II}(x) = E\psi_{II}(x)$$

Subtract $V_0\psi(x)$ from both sides of the equation and factor out $\psi_{II}(x)$.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} = (E - V_0)\psi_{II}(x)$$

Multiply both sides by $-\frac{2m}{\hbar^2}$ and distribute the minus sign to write $-(E - V_0) = V_0 - E$.

$$\frac{d^2\psi_{II}(x)}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi_{II}(x)$$

Since $0 < E < V_0$ in this problem, $V_0 - E > 0$ and both sides of the equation are positive, in contrast to the differential equation for region I. This corresponds to the first special case on page 185, where

$$k_2 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

and

$$\psi_{II}(x) = Ce^{k_2 x} + De^{-k_2 x}$$

In region III, $V(x) = \infty$ and $\psi_{III}(x) = 0$ because there is zero probability of finding the particle where the potential energy is infinite. Since the wave function must be continuous across the boundary, set the wave functions equal at the points where they meet ($x = 0$ and $x = a$):

$$\psi_I(0) = \psi_{II}(0) \quad \rightarrow \quad A + B = C + D \ (Eq. 1)$$

$$\psi_{II}(a) = \psi_{III}(a) \quad \rightarrow \quad Ce^{k_2 a} + De^{-k_2 a} = 0 \ (Eq. 2)$$

The first derivative of the wave function must also be continuous across the point $x = 0$.

$$\left.\frac{d\psi_I}{dx}\right|_{x=0} = \left.\frac{d\psi_{II}}{dx}\right|_{x=0} \quad \rightarrow \quad \left.\frac{d}{dx}\left(Ae^{ik_1 x} + Be^{-ik_1 x}\right)\right|_{x=0} = \left.\frac{d}{dx}\left(Ce^{k_2 x} + De^{-k_2 x}\right)\right|_{x=0}$$

$$\left.\left(ik_1 Ae^{ik_1 x} - ik_1 Be^{-ik_1 x}\right)\right|_{x=0} = \left(k_2 Ce^{k_2 x} - k_2 De^{-k_2 x}\right)|_{x=0}$$

$$ik_1 A - ik_1 B = k_2 C - k_2 D \ (Eq. 3)$$

Multiply Eq. 1 by ik_1.

$$ik_1 A + ik_1 B = ik_1 C + ik_1 D \ (Eq. 4)$$

Add Eq.'s 3 and 4 together. Note that $ik_1 B$ cancels out.

$$2ik_1 A = (ik_1 + k_2)C + (ik_1 - k_2)D \ (Eq. 5)$$

Subtract Eq. 3 from Eq. 4. This time $ik_1 A$ cancels out.

$$2ik_1 B = (ik_1 - k_2)C + (ik_1 + k_2)D \ (Eq. 6)$$

Multiply Eq. 2 by $e^{k_2 a}$. Note that $e^{k_2 a}e^{k_2 a} = e^{2k_2 a}$ and $e^{k_2 a}e^{-k_2 a} = 1$.

$$Ce^{2k_2 a} + D = 0 \quad \rightarrow \quad -Ce^{2k_2 a} = D \ (Eq. 7)$$

Substitute Eq. 7 into Eq.'s 5 and 6.

$$2ik_1 A = (ik_1 + k_2)C - (ik_1 - k_2)Ce^{2k_2 a} \quad \rightarrow \quad C = \frac{2ik_1 A}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2 a}} \ (Eq. 8)$$

$$2ik_1 B = (ik_1 - k_2)C - (ik_1 + k_2)Ce^{2k_2 a} \quad \rightarrow \quad C = \frac{2ik_1 B}{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2 a}} \ (Eq. 9)$$

Set the two expressions for C from Eq.'s 8 and 9 equal to one another.

$$\frac{2ik_1 A}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2 a}} = \frac{2ik_1 B}{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2 a}} \ (Eq. 10)$$

Divide both sides of the equation by $2ik_1$.

$$\frac{A}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2 a}} = \frac{B}{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2 a}} \ (Eq. 11)$$

Multiply both sides of the equation by $[ik_1 - k_2 - (ik_1 + k_2)e^{2k_2 a}]$.

$$B = \frac{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2a}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} A \; (Eq. 12)$$

Plug the expression for B into Eq. 9. Note that $[ik_1 - k_2 - (ik_1 + k_2)e^{2k_2a}]$ cancels out.

$$C = \frac{2ik_1}{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2a}} \frac{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2a}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} A$$

$$C = \frac{2ik_1A}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} \; (Eq. 13)$$

Plug the expression for C into Eq. 7.

$$D = -Ce^{2k_2a} = \frac{-2ik_1Ae^{2k_2a}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} \; (Eq. 14)$$

Substitute Eq.'s 12, 13, and 14 into the solutions for the wave function in each region.

$$\psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} = Ae^{ik_1x} + \frac{ik_1 - k_2 - (ik_1 + k_2)e^{2k_2a}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} Ae^{-ik_1x} \; (Eq. 15)$$

$$\psi_{II}(x) = Ce^{k_2x} + De^{-k_2x} = \frac{2ik_1Ae^{k_2x}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} - \frac{2ik_1Ae^{2k_2a}e^{-k_2x}}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}}$$

$$= \frac{2ik_1A(e^{k_2x} - e^{2k_2a}e^{-k_2x})}{ik_1 + k_2 - (ik_1 - k_2)e^{2k_2a}} \; (Eq. 16)$$

There is an alternative way to express the answer. First, we will distribute.

$$\psi_I(x) = Ae^{ik_1x} + \frac{ik_1 - k_2 - ik_1e^{2k_2a} - k_2e^{2k_2a}}{ik_1 + k_2 - ik_1e^{2k_2a} + k_2e^{2k_2a}} Ae^{-ik_1x} \; (Eq. 17)$$

$$\psi_{II}(x) = \frac{2ik_1A\left(e^{k_2x} - e^{2k_2a}e^{-k_2x}\right)}{ik_1 + k_2 - ik_1e^{2k_2a} + k_2e^{2k_2a}} \; (Eq. 18)$$

Next, multiply the numerator and denominator each by e^{-k_2a}.

$$\psi_I(x) = Ae^{ik_1x} + \frac{ik_1e^{-k_2a} - k_2e^{-k_2a} - ik_1e^{k_2a} - k_2e^{k_2a}}{ik_1e^{-k_2a} + k_2e^{-k_2a} - ik_1e^{k_2a} + k_2e^{k_2a}} Ae^{-ik_1x} \; (Eq. 19)$$

$$\psi_{II}(x) = \frac{2ik_1A\left(e^{-k_2a}e^{k_2x} - e^{k_2a}e^{-k_2x}\right)}{ik_1e^{-k_2a} + k_2e^{-k_2a} - ik_1e^{k_2a} + k_2e^{k_2a}} \; (Eq. 20)$$

Apply the hyperbolic identities $\cosh y = \frac{e^y + e^{-y}}{2}$ and $\sinh y = \frac{e^y - e^{-y}}{2}$.

$$\psi_I(x) = Ae^{ik_1x} + \frac{-2ik_1 \sinh(k_2a) - 2k_2 \cosh(k_2a)}{-2ik_1 \sinh(k_2a) + 2k_2 \cosh(k_2a)} Ae^{-ik_1x}$$

$$= \boxed{Ae^{ik_1x} + \frac{ik_1 \sinh(k_2a) + k_2 \cosh(k_2a)}{ik_1 \sinh(k_2a) - k_2 \cosh(k_2a)} Ae^{-ik_1x}} \; (Eq. 21)$$

$$\psi_{II}(x) = \frac{2ik_1A \sinh(k_2x - k_2a)}{-2ik_1 \sinh(k_2a) + 2k_2 \cosh(k_2a)} = \frac{ik_1A \sinh(k_2x - k_2a)}{-ik_1 \sinh(k_2a) + k_2 \cosh(k_2a)}$$

$$= \boxed{\frac{-ik_1A \sinh[k_2(x - a)]}{ik_1 \sinh(k_2a) - k_2 \cosh(k_2a)}} \; (Eq. 22)$$

Note that it is possible to express these answers in a variety of different forms. If you obtain a seemingly different answer, it isn't necessarily incorrect just because it appears different at first. It may take some algebra to make a proper comparison.

14 PROBABILITIES AND EXPECTATION VALUES

Relevant Terminology

Frequency – the number of oscillations completed per second.

Energy – the ability to do work, meaning that a force is available to contribute towards the displacement of an object.

Kinetic energy – work that can be done by changing speed. Moving objects have kinetic energy. Hence, kinetic energy is considered to be energy of motion.

Potential energy – work that can be done by changing position. All forms of potential energy are stored energy.

Momentum – mass times velocity.

Wave function – a quantity that helps to determine probabilities relating to matter waves.

Normalization

As we will explore later, $\psi^*\psi$ serves as a probability density. In order for integrals of the form $\int_{x=a}^{b} \psi^*(x)\psi(x)\,dx$ to serve as a meaningful measure of probability, the wave function must be **normalized**. For a particle traveling in one dimension along $\pm x$, the normalization condition is:

$$\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = 1$$

Recall that the **complex conjugate**, ψ^*, is found by replacing i with $-i$ in ψ (page 166). Although the wave function, ψ, may be complex, it's a property of complex algebra that the quantity $\psi^*\psi$ is always real. For example, consider $z = x + iy$, for which $z^*z = (x - iy)(x + iy) = x^2 - ixy + ixy + y^2 = x^2 + y^2$. The time-dependent wave function, $\Psi(x,t) = \psi(x)\phi(t) = \psi(x)e^{-iEt/\hbar}$ provided that $V(x,t) = V(x)$ (see Chapter 13, page 180), is also normalized.

$$\int_{x=-\infty}^{\infty} \Psi^*(x,t)\Psi(x,t)\,dx = 1$$

Normalization ensures that there is a 100% chance of finding the particle *somewhere*.

For a three-dimensional problem, the normalization integral is over all space:

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty}\int_{z=-\infty}^{\infty} \psi^*(x,y,z)\psi(x,y,z)\, dxdydz = 1$$

In spherical coordinates, the differential volume element is $dV = r^2 \sin\theta\, drd\theta d\varphi$:

$$\int_{r=0}^{\infty}\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)\, r^2 \sin\theta\, drd\theta d\varphi = 1$$

Operators

Recall from Chapter 13 the following operators:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) \quad \text{(the Hamiltonian operator)}$$

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \quad \text{(the momentum operator in 1D along } \pm x\text{)}$$

$$\hat{p} = -i\hbar\nabla \quad \text{(the momentum operator in 3D)}$$

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \quad \text{(the energy operator)}$$

$$\nabla = \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z} \quad \text{(the gradient operator)}$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are the Cartesian unit vectors.

Expectation Values

The **expectation value** for a given quantity provides a measure of the average value of that quantity. For example, for a particle traveling in one dimension along $\pm x$, the expectation value for x, labeled as $\bar{x}$ (it has a bar over it), is:

$$\bar{x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x\, \Psi(x,t)\, dx$$

The expectation value for any other quantity is obtained the same way by placing it between $\Psi^*(x,t)$ and $\Psi(x,t)$. For example, the expectation value for x^2 is:

$$\overline{x^2} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x^2\, \Psi(x,t)\, dx$$

To find the expectation value for an **operator**, insert the operator between $\Psi^*(x,t)$ and $\Psi(x,t)$.

$$\overline{p_x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{p}_x\, \Psi(x,t)\, dx = -i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \frac{\partial}{\partial x}\, \Psi(x,t)\, dx$$

$$\bar{E} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{E}\, \Psi(x,t)\, dx = i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \frac{\partial}{\partial t}\, \Psi(x,t)\, dx$$

The Root-Mean-Square Deviation

The **variance** (Δx) or the root-mean-square deviation provides a measure of the **uncertainty**.

$$\Delta x = \sqrt{\overline{x^2} - \bar{x}^2}$$

$$\Delta p = \sqrt{\overline{p^2} - \bar{p}^2}$$

Probabilities

For one-dimensional motion along $\pm x$, the quantity $\Psi^*(x,t)\Psi(x,t)$ is called the **probability density**. Specifically, the probability that the particle will be found in the interval $a \leq x \leq b$ at the instant t is given by the following definite integral.

$$P(a \leq x \leq b) = \int_{x=a}^{b} \Psi^*(x,t)\Psi(x,t)\,dx = \int_{x=a}^{b} \psi^*(x)\psi(x)\,dx$$

Since $\Psi(x,t) = \psi(x)\phi(t) = \psi(x)e^{-iEt/\hbar}$ provided that V is independent of time (according to Chapter 13), $\Psi^*(x,t) = \psi(x)e^{iEt/\hbar}$, such that $\Psi^*\Psi = \psi^* e^{iEt/\hbar}\psi e^{-iEt/\hbar} = \psi^*\psi$.

Gaussian Integrals

For some systems in quantum mechanics, the wave function involves an expression of the form e^{-ax^2} (characteristic of a Gaussian distribution). The anti-derivative of e^{-ax^2} doesn't exist, yet it is possible to perform a definite integral over e^{-ax^2} if the limits of integration are $\pm\infty$ or zero. The following definite integrals often come in handy in quantum mechanics.

$$\int_{x=0}^{\infty} e^{-ax^2}\,dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} \quad , \quad \int_{x=-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}$$

$$\int_{x=0}^{\infty} x^2 e^{-ax^2}\,dx = \frac{1}{4a}\sqrt{\frac{\pi}{a}} \quad , \quad \int_{x=-\infty}^{\infty} x^2 e^{-ax^2}\,dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

$$\int_{x=0}^{\infty} x^4 e^{-ax^2}\,dx = \frac{3}{8a^2}\sqrt{\frac{\pi}{a}} \quad , \quad \int_{x=-\infty}^{\infty} x^4 e^{-ax^2}\,dx = \frac{3}{4a^2}\sqrt{\frac{\pi}{a}}$$

$$\int_{x=0}^{\infty} x^{2n} e^{-ax^2}\,dx = \frac{(2n-1)!!}{2^{n+1}a^n}\sqrt{\frac{\pi}{a}} \quad , \quad \int_{x=-\infty}^{\infty} x^{2n} e^{-ax^2}\,dx = \frac{(2n-1)!!}{2^n a^n}\sqrt{\frac{\pi}{a}}$$

Note: The double factorial (!!) means **<u>every other</u>** integer. Example: $9!! = (9)(7)(5)(3)(1)$.

$$\int_{x=0}^{\infty} x e^{-ax^2}\, dx = \frac{1}{2a} \quad , \quad \int_{x=-\infty}^{\infty} x e^{-ax^2}\, dx = 0$$

$$\int_{x=0}^{\infty} x^3 e^{-ax^2}\, dx = \frac{1}{2a^2} \quad , \quad \int_{x=-\infty}^{\infty} x^3 e^{-ax^2}\, dx = 0$$

$$\int_{x=0}^{\infty} x^5 e^{-ax^2}\, dx = \frac{1}{a^3} \quad , \quad \int_{x=-\infty}^{\infty} x^5 e^{-ax^2}\, dx = 0$$

$$\int_{x=0}^{\infty} x^{2n+1} e^{-ax^2}\, dx = \frac{n!}{2a^{n+1}} \quad , \quad \int_{x=-\infty}^{\infty} x^{2n+1} e^{-ax^2}\, dx = 0$$

Why Does $\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$?

Begin by calling this integral H.

$$H = \int_{x=-\infty}^{\infty} e^{-ax^2}\, dx$$

The 'trick' is to square both sides, using a different symbol for the second integral so that we may write it as a double integral. We will then transform to 2D polar coordinates, where $x = r\cos\theta$, $y = r\sin\theta$, and $dA = dxdy = r\,drd\theta$. Note that limits from $r = 0$ to ∞ and from $\theta = 0$ to 2π cover the entire xy plane (just like the original limits). Also note that $e^{-ax^2}e^{-ay^2}$ $= e^{-ax^2-ay^2} = e^{-a(x^2+y^2)}$.

$$H^2 = \left(\int_{x=-\infty}^{\infty} e^{-ax^2}\, dx\right)\left(\int_{y=-\infty}^{\infty} e^{-ay^2}\, dy\right) = \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} e^{-a(x^2+y^2)}\, dxdy$$

Note that $x^2 + y^2 = r^2\cos^2\theta + r^2\sin^2\theta = r^2$ because $\cos^2\theta + \sin^2\theta = 1$.

$$H^2 = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} e^{-ar^2}\, r\, dr d\theta$$

Make the substitution $u = r^2$, such that $du = 2r\, dr$ and $dr = \frac{du}{2r}$.

$$H^2 = \int_{u=0}^{\infty} \int_{\theta=0}^{2\pi} e^{-au}\, r\, \frac{du}{2r} d\theta = \frac{1}{2} \int_{u=0}^{\infty} \int_{\theta=0}^{2\pi} e^{-au}\, du d\theta = -\frac{1}{2} \left[\frac{e^{-au}}{a}\right]_{u=0}^{\infty} [\theta]_{\theta=0}^{2\pi}$$

$$H^2 = -\frac{1}{2}\left(\lim_{u\to\infty} \frac{e^{-au}}{a} - \frac{e^0}{a}\right)(2\pi - 0) = -\pi\left(0 - \frac{1}{a}\right) = \frac{\pi}{a}$$

Note that $\lim_{u\to\infty} \frac{e^{-au}}{a} = 0$ because $e^{-au} = \frac{1}{e^{au}}$ gets smaller and smaller as u grows larger. Take the squareroot of both sides of the equation.

$$H = \int_{x=-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$$

Where Do the Other Gaussian Integrals Come From?

The integral $\int_{x=0}^{\infty} xe^{-ax^2}\,dx$ is simpler: Make the substitution $u = x^2$, such that $du = 2xdx$.

$$\int_{x=0}^{\infty} xe^{-ax^2}\,dx = \int_{x=0}^{\infty} xe^{-au}\frac{du}{2x} = \frac{1}{2}\int_{x=0}^{\infty} e^{-au}\,du = -\frac{1}{2}\left[\frac{e^{-au}}{a}\right]_{u=0}^{\infty}$$

$$= -\frac{1}{2}\left(\lim_{u\to\infty}\frac{e^{-au}}{a} - \frac{e^{0}}{a}\right) = -\frac{1}{2}\left(0 - \frac{1}{a}\right) = \frac{1}{2a}$$

Integrals of the form $\int_{x=-\infty}^{\infty} xe^{-ax^2}\,dx = 0$ or $\int_{x=-\infty}^{\infty} x^3e^{-ax^2}\,dx = 0$ (with an odd power of x) are zero because e^{-ax^2} is an even function. These integrands are odd functions (since an odd function times an even function makes an odd function), and an odd function integrated over symmetric limits equals zero (the areas cancel out).

A simple, efficient way to obtain the other Gaussian integrals is to apply a partial derivative with respect to a. When taking a partial derivative with respect to a, treat x as if it were a constant. For example:

$$\frac{\partial}{\partial a}\int_{x=-\infty}^{\infty} e^{-ax^2}\,dx = \frac{\partial}{\partial a}\sqrt{\frac{\pi}{a}} \quad \to \quad \int_{x=-\infty}^{\infty} (-x^2)e^{-ax^2}\,dx = \sqrt{\pi}\left(-\frac{1}{2a^{3/2}}\right)$$

$$\int_{x=-\infty}^{\infty} x^2 e^{-ax^2}\,dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

$$\frac{\partial}{\partial a}\int_{x=0}^{\infty} xe^{-ax^2}\,dx = \frac{\partial}{\partial a}\frac{1}{2a} \quad \to \quad \int_{x=0}^{\infty} x(-x^2)e^{-ax^2}\,dx = -\frac{1}{2a^2}$$

$$\int_{x=0}^{\infty} x^3 e^{-ax^2}\,dx = \frac{1}{2a^2}$$

Wave Functions for the Simple Harmonic Oscillator and the Hydrogen Atom

For the one-dimensional **simple harmonic oscillator**, $V(x) = \frac{1}{2}Cx^2$, Schrödinger's equation gives the following energy eigenvalues and corresponding wave functions.

$$E_n = \left(n + \frac{1}{2}\right)hf \quad , \quad n = 0,1,2,3\cdots \quad , \quad \xi = \frac{C^{1/4}m^{1/4}}{\hbar^{1/2}}x$$

$$\psi_0 = A_0 e^{-\xi^2/2} \quad , \quad \psi_1 = A_1 \xi e^{-\xi^2/2} \quad , \quad \psi_2 = A_2(1 - 2\xi^2)e^{-\xi^2/2}$$

$$\psi_3 = A_3(3\xi - 2\xi^3)e^{-\xi^2/2} \quad , \quad \psi_4 = A_4(3 - 12\xi^2 + 4\xi^4)e^{-\xi^2/2} \quad , \quad \cdots$$

For a **hydrogen-like atom** (with a single electron), $V(r) = -\frac{Zke^2}{r}$, Schrödinger's equation gives the following energy eigenvalues and corresponding wave functions. (In contrast to Chapter 9, in Chapters 13-14 the symbol V represents potential energy, **not** electric potential.)

$$E_n = -\frac{Z^2 k\mu e^4}{2\hbar^2 n^2} \quad , \quad n = 1,2,3,\cdots \quad , \quad \mu = \frac{m_{nucleus} m_e}{m_{nucleus} + m_e} \quad , \quad a_0 = \frac{\hbar^2}{k\mu e^2}$$

$$\psi_{1,0,0} = A_{100} e^{-Zr/a_0} \quad , \quad \psi_{2,0,0} = A_{200}\left(2 - \frac{Zr}{a_0}\right)e^{-Zr/2a_0}$$

$$\psi_{2,1,0} = A_{210} r \cos\theta \, e^{-Zr/2a_0} \quad , \quad \psi_{2,1,\pm 1} = A_{211} r \sin\theta \, e^{-Zr/2a_0} e^{\pm i\varphi} \quad \cdots$$

Symbols and SI Units

Symbol	Name	SI Units
Ψ	wave function for position and time	unitless
ψ	wave function for position	unitless
ϕ	wave function for time	unitless
h	Planck's constant	J·s or J/Hz
$\hbar$	h-bar	J·s or J/Hz
E	energy	J
V	potential energy (or volume)	J (or m^3)
H	Hamiltonian	J
p	momentum	$\frac{kg \cdot m}{s}$

m	mass	kg
∇	gradient	1/m
$\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$	Cartesian unit vectors	unitless
$\bar{x}$	expectation value for x	m
Δx	variance (root-mean-square deviation) or uncertainty in x	m
$\bar{p}$	expectation value for momentum	$\frac{\text{kg·m}}{\text{s}}$
Δp	variance (root-mean-square deviation) or uncertainty in p	$\frac{\text{kg·m}}{\text{s}}$
P	probability	unitless

r	distance from the origin	m
θ	polar angle (with the z-axis in spherical coordinates), angle counterclockwise from $+x$ in 2D polar coordinates	rad
φ	azimuthal angle (counterclockwise from $+x$ after projecting onto the xy plane in spherical coordinates)	rad
C	spring constant for Hooke's law	N/m
f	frequency	Hz
n	an integer	unitless
A_n	normalization constant	it depends
ξ	dimensionless simple harmonic oscillator coordinate	unitless

Z	atomic number (number of protons in the nucleus)	unitless
k	Coulomb's constant	$\frac{\text{N}\cdot\text{m}^2}{\text{C}^2}$ or $\frac{\text{kg}\cdot\text{m}^3}{\text{C}^2\cdot\text{s}^2}$
e	the charge of a proton (the absolute value of the charge of an electron)	C
μ	reduced mass	kg
a_0	Bohr radius	m

Note: The symbols Ψ and ψ are the uppercase and lowercase Greek letter psi, ϕ and φ are two variations of the lowercase Greek letter phi, θ is the lowercase Greek letter theta, μ is the lowercase Greek letter mu, and ξ is the lower-case Greek letter xi.

Constants

Quantity	Value
Planck's constant	$h = 6.626 \times 10^{-34}$ J·s
h-bar	$\hbar = 1.055 \times 10^{-34}$ J·s
charge of a proton	$e = 1.6021766 \times 10^{-19}$ C
charge of an electron	$-e = -1.6021766 \times 10^{-19}$ C

Electron Volts and Angstroms

The conversion from electron Volts (eV) to Joules (J) is:

$$1\ \text{eV} = 1.6021766 \times 10^{-19}\ \text{J}$$

The conversion from Joules (J) to electron Volts (eV) is:

$$1\ \text{J} = \frac{1\ \text{eV}}{1.6021766 \times 10^{-19}} = 6.2415092 \times 10^{18}\ \text{eV}$$

One **Angstrom** (Å) equals 10^{-10} m, which equates to 0.1 nm.

Strategy

If a problem involves normalization, probability, expectation values, or variance (or root-mean-square values), choose the relevant formula:

- To **normalize** a wave function, perform the appropriate integral. This allows you to solve for a constant.

$$\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = 1 \quad \text{(1D motion along } \pm x\text{)}$$

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty}\int_{z=-\infty}^{\infty} \psi^*(x,y,z)\psi(x,y,z)\,dxdydz = 1 \quad \text{(3D Cartesian coordinates)}$$

$$\int_{r=0}^{\infty}\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \psi^*(r,\theta,\varphi)\psi(r,\theta,\varphi)\,r^2\sin\theta\,drd\theta d\varphi = 1 \quad \text{(3D spherical coordinates)}$$

- The **probability** that a particle will be found in the interval $a \le x \le b$ at the instant t is given by the following definite integral (provided that V isn't time-dependent).

$$P(a \le x \le b) = \int_{x=a}^{b} \Psi^*(x,t)\Psi(x,t)\,dx = \int_{x=a}^{b} \psi^*(x)\psi(x)\,dx$$

- The **expectation value** for a particular quantity, A, is defined as a definite integral.

$$\bar{A} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, A\, \Psi(x,t)\, dx$$

Some common examples include:

$$\bar{x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x\, \Psi(x,t)\, dx \quad , \quad \overline{x^2} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x^2\, \Psi(x,t)\, dx$$

$$\overline{p_x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{p}_x\, \Psi(x,t)\, dx = -i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t) \frac{\partial}{\partial x} \Psi(x,t)\, dx$$

$$\bar{E} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{E}\, \Psi(x,t)\, dx = i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t) \frac{\partial}{\partial t} \Psi(x,t)\, dx$$

- The **variance** (or root-mean-square deviation) provides a measure of uncertainty.

$$\Delta x = \sqrt{\overline{x^2} - \bar{x}^2} \quad , \quad \Delta p = \sqrt{\overline{p^2} - \bar{p}^2}$$

Recall Heisenberg's uncertainty principle (Chapter 11): $\Delta x \Delta p_x \geq \frac{\hbar}{2}$.

- For a **simple harmonic oscillator** or a **hydrogen-like atom**, see page 205. For a step potential or square well potential, see Chapter 13.
- If the problem involves a **Gaussian** function, e^{-ax^2}, see page 203.

Example: A particle is confined to move within the interval $0 \le x \le L$, where L is a constant. Its wave function is known to be:

$$\psi(x) = A \sin\left(\frac{2\pi x}{L}\right)$$

(A) Determine the coefficient A.

We can solve for the coefficient A by normalizing the wave function. Since the particle only exists in the interval $0 \le x \le L$ in this problem, we may adjust the limits of integration. (In the intervals $-\infty < x < 0$ and $L < x < \infty$, the wave function ψ equals zero.)

$$\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\, dx = \int_{x=0}^{L} A \sin\left(\frac{2\pi x}{L}\right) A \sin\left(\frac{2\pi x}{L}\right) dx = A^2 \int_{x=0}^{L} \sin^2\left(\frac{2\pi x}{L}\right) dx = 1$$

Recall the trig identity $\sin^2\theta = \frac{1-\cos(2\theta)}{2}$. Let $\theta = \frac{2\pi x}{L}$ such that $2\theta = \frac{4\pi x}{L}$.

$$A^2 \int_{x=0}^{L} \frac{1-\cos\left(\frac{4\pi x}{L}\right)}{2} dx = \frac{A^2}{2} \int_{x=0}^{L} dx - \frac{A^2}{2} \int_{x=0}^{L} \cos\left(\frac{4\pi x}{L}\right) dx = 1$$

$$\frac{A^2}{2}[x]_{x=0}^{L} - \frac{A^2}{2}\left(\frac{L}{4\pi}\right)\left[\sin\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L} = 1$$

$$\frac{A^2}{2}(L-0) - \frac{A^2 L}{8\pi}[\sin(4\pi) - \sin(0)] = 1$$

$$\frac{A^2 L}{2} - \frac{A^2 L}{8\pi}(0-0) = 1$$

$$\frac{A^2 L}{2} = 1$$

$$\boxed{A = \sqrt{\frac{2}{L}}}$$

(B) Determine the expectation value for the particle's position.

Perform the integral for $\bar{x}$. Recall from Chapter 13 that $\Psi(x,t) = \psi(x)e^{-iEt/\hbar}$: Thus $\Psi^* x \Psi = \psi^* e^{iEt/\hbar} x \psi e^{-iEt/\hbar} = \psi^* x \psi$. Just as in Part A, the limits adapt to the particle's confinement.

$$\bar{x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x\, \Psi(x,t)\, dx = \int_{x=-\infty}^{\infty} \psi^*(x)\, x\, \psi(x)\, dx$$

$$\bar{x} = \int_{x=0}^{L} A \sin\left(\frac{2\pi x}{L}\right) x\, A \sin\left(\frac{2\pi x}{L}\right) dx = A^2 \int_{x=0}^{L} x \sin^2\left(\frac{2\pi x}{L}\right) dx$$

Recall the trig identity $\sin^2\theta = \frac{1-\cos(2\theta)}{2}$. Let $\theta = \frac{2\pi x}{L}$ such that $2\theta = \frac{4\pi x}{L}$.

$$\bar{x} = A^2 \int_{x=0}^{L} \frac{x - x\cos\left(\frac{4\pi x}{L}\right)}{2} dx = \frac{A^2}{2} \int_{x=0}^{L} x\, dx - \frac{A^2}{2} \int_{x=0}^{L} x \cos\left(\frac{4\pi x}{L}\right) dx$$

Use integration by parts for the second integral with $u = x$ and $dv = \cos\left(\frac{4\pi x}{L}\right)dx$, such that $du = dx$ and $v = \frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right)$. Recall that $\int_i^f u\,dv = [uv]_i^f - \int_i^f v\,du$.

$$\bar{x} = \frac{A^2}{2}\left[\frac{x^2}{2}\right]_{x=0}^{L} - \frac{A^2}{2}\left[x\frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L} + \frac{A^2}{2}\int_{x=0}^{L}\frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right)dx$$

$$\bar{x} = \frac{A^2}{2}\left(\frac{L^2}{2} - \frac{0^2}{2}\right) - \frac{A^2}{2}\left[L\frac{L}{4\pi}\sin\left(\frac{4\pi L}{L}\right) - 0\frac{L}{4\pi}\sin\left(\frac{4\pi 0}{L}\right)\right] + \frac{A^2}{2}\left(\frac{L}{4\pi}\right)^2\left[-\cos\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L}$$

$$\bar{x} = \frac{A^2}{2}\left(\frac{L^2}{2}\right) - \frac{A^2}{2}\left[\frac{L^2}{4\pi}\sin(4\pi)\right] - \frac{A^2}{2}\left(\frac{L}{4\pi}\right)^2[\cos(4\pi) - \cos(0)]$$

$$\bar{x} = \frac{A^2L^2}{4} - \frac{A^2}{2}\left[\frac{L^2}{4\pi}(0)\right] - \frac{A^2}{2}\left(\frac{L}{4\pi}\right)^2(1-1) = \frac{A^2L^2}{4} - 0 - 0 = \frac{A^2L^2}{4}$$

Recall from Part A that $A = \sqrt{\frac{2}{L}}$.

$$\bar{x} = \left(\sqrt{\frac{2}{L}}\right)^2\frac{L^2}{4} = \frac{2}{L}\frac{L^2}{4} = \boxed{\frac{L}{2}}$$

This answer should make sense. The particle is basically traveling back and forth inside an infinite square well in the interval $0 \leq x \leq L$: On average, the particle is at the center of the infinite square well, $\bar{x} = \frac{L}{2}$.

(C) Determine the expectation value for the particle's momentum.

Perform the integral for $\overline{p_x}$. Similar to Part B, $\Psi^* \hat{p}_x \Psi = -i\hbar \Psi^* \frac{\partial}{\partial x} \Psi = -i\hbar\, \psi^* e^{iEt/\hbar} \frac{\partial}{\partial x} \psi e^{-iEt/\hbar}$ $= -i\hbar \psi^* \frac{\partial}{\partial x} \psi$ and the limits adapt to the particle's confinement. (Recall from Chapter 13 that in 1D the momentum operator is $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$.)

$$\overline{p_x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{p}_x\, \Psi(x,t)\, dx = -i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \frac{\partial}{\partial x}\, \Psi(x,t)\, dx$$

$$\overline{p_x} = -i\hbar \int_{x=-\infty}^{\infty} \psi^*(x)\, \frac{\partial}{\partial x}\, \psi(x)\, dx = -i\hbar \int_{x=0}^{L} A \sin\left(\frac{2\pi x}{L}\right) \frac{\partial}{\partial x} A \sin\left(\frac{2\pi x}{L}\right) dx$$

$$\overline{p_x} = -i\hbar A^2 \left(\frac{2\pi}{L}\right) \int_{x=0}^{L} \sin\left(\frac{2\pi x}{L}\right) \cos\left(\frac{2\pi x}{L}\right) dx$$

Recall the trig identity $\sin(2\theta) = 2\sin\theta\cos\theta$, for which $\sin\theta\cos\theta = \frac{1}{2}\sin(2\theta)$. Let $\theta = \frac{2\pi x}{L}$ such that $2\theta = \frac{4\pi x}{L}$.

$$\overline{p_x} = -\frac{2\pi i\hbar A^2}{L}\int_{x=0}^{L}\frac{1}{2}\sin\left(\frac{4\pi x}{L}\right)dx = -\frac{\pi i\hbar A^2}{L}\left(\frac{L}{4\pi}\right)\left[-\cos\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L}$$

$$\overline{p_x} = -\frac{i\hbar A^2}{4}[-\cos(4\pi) + \cos(0)] = -\frac{i\hbar A^2}{4}(-1+1) = -\frac{i\hbar A^2}{4}(0) = \boxed{0}$$

Since the particle is basically traveling back and forth inside an infinite square well in the interval $0 \le x \le L$, on average the particle's momentum is zero, $\bar{p}_x = 0$ (it has just as much momentum to the right as it has to the left, on average).

(D) Determine the variance in the particle's position.

Since the variance in the particle's position is defined as $\Delta x = \sqrt{\overline{x^2} - \bar{x}^2}$, we will first need to perform the integral for $\overline{x^2}$ (which we will see turns out to be different from $\bar{x}^2$).

$$\overline{x^2} = \int_{x=-\infty}^{\infty}\Psi^*(x,t)\,x^2\,\Psi(x,t)\,dx = \int_{x=-\infty}^{\infty}\psi^*(x)\,x^2\,\psi(x)\,dx$$

$$\overline{x^2} = \int_{x=0}^{L} A\sin\left(\frac{2\pi x}{L}\right) x^2 A\sin\left(\frac{2\pi x}{L}\right) dx = A^2 \int_{x=0}^{L} x^2 \sin^2\left(\frac{2\pi x}{L}\right) dx$$

Recall the trig identity $\sin^2\theta = \frac{1-\cos(2\theta)}{2}$. Let $\theta = \frac{2\pi x}{L}$ such that $2\theta = \frac{4\pi x}{L}$.

$$\overline{x^2} = A^2 \int_{x=0}^{L} \frac{x^2 - x^2\cos\left(\frac{4\pi x}{L}\right)}{2} dx = \frac{A^2}{2}\int_{x=0}^{L} x^2\,dx - \frac{A^2}{2}\int_{x=0}^{L} x^2\cos\left(\frac{4\pi x}{L}\right) dx$$

Use integration by parts for the second integral with $u = x^2$ and $dv = \cos\left(\frac{4\pi x}{L}\right) dx$, such that $du = 2x\,dx$ and $v = \frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right)$. Recall that $\int_i^f u\,dv = [uv]_i^f - \int_i^f v\,du$.

$$\overline{x^2} = \frac{A^2}{2}\left[\frac{x^3}{3}\right]_{x=0}^{L} - \frac{A^2}{2}\left[x^2\frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L} + \frac{A^2}{2}\int_{x=0}^{L}\frac{L}{4\pi}\sin\left(\frac{4\pi x}{L}\right) 2x\,dx$$

$$\overline{x^2} = \frac{A^2}{2}\left(\frac{L^3}{3} - \frac{0^3}{3}\right) - \frac{A^2}{2}\left[L^2\frac{L}{4\pi}\sin\left(\frac{4\pi L}{L}\right) - 0^2\frac{L}{4\pi}\sin\left(\frac{4\pi 0}{L}\right)\right] + \frac{A^2 L}{4\pi}\int_{x=0}^{L} x\sin\left(\frac{4\pi x}{L}\right) dx$$

$$\overline{x^2} = \frac{A^2}{2}\left(\frac{L^3}{3}\right) - \frac{A^2}{2}\left[\frac{L^3}{4\pi}\sin(4\pi)\right] + \frac{A^2 L}{4\pi}\int_{x=0}^{L} x\sin\left(\frac{4\pi x}{L}\right)dx$$

$$\overline{x^2} = \frac{A^2 L^3}{6} - \frac{A^2}{2}\left[\frac{L^3}{4\pi}(0)\right] + \frac{A^2 L}{4\pi}\int_{x=0}^{L} x\sin\left(\frac{4\pi x}{L}\right)dx = \frac{A^2 L^3}{6} + \frac{A^2 L}{4\pi}\int_{x=0}^{L} x\sin\left(\frac{4\pi x}{L}\right)dx$$

Use integration parts again, this time with $u = x$ and $dv = \sin\left(\frac{4\pi x}{L}\right)dx$, such that $du = dx$ and $v = -\frac{L}{4\pi}\cos\left(\frac{4\pi x}{L}\right)$.

$$\overline{x^2} = \frac{A^2 L^3}{6} + \frac{A^2 L}{4\pi}\left[x\left(-\frac{L}{4\pi}\right)\cos\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L} - \frac{A^2 L}{4\pi}\int_{x=0}^{L}\left(-\frac{L}{4\pi}\right)\cos\left(\frac{4\pi x}{L}\right)dx$$

$$\overline{x^2} = \frac{A^2 L^3}{6} - \frac{A^2 L^2}{16\pi^2}\left[x\cos\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L} + \frac{A^2 L^2}{16\pi^2}\int_{x=0}^{L}\cos\left(\frac{4\pi x}{L}\right)dx$$

$$\overline{x^2} = \frac{A^2 L^3}{6} - \frac{A^2 L^2}{16\pi^2}\left[L\cos\left(\frac{4\pi L}{L}\right) - 0\cos\left(\frac{4\pi 0}{L}\right)\right] + \frac{A^2 L^2}{16\pi^2}\left(\frac{L}{4\pi}\right)\left[\sin\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L}$$

$$\overline{x^2} = \frac{A^2L^3}{6} - \frac{A^2L^2}{16\pi^2} L\cos(4\pi) + \frac{A^2L^3}{64\pi^3}[\sin(4\pi) - \sin(0)]$$

$$\overline{x^2} = \frac{A^2L^3}{6} - \frac{A^2L^3}{16\pi^2}(1) + \frac{A^2L^3}{64\pi^3}(0-0) = \frac{A^2L^3}{6} - \frac{A^2L^3}{16\pi^2} = \frac{A^2L^3}{2}\left(\frac{1}{3} - \frac{1}{8\pi^2}\right)$$

In the last step, we factored out $\frac{A^2L^3}{2}$. Recall from Part A that $A = \sqrt{\frac{2}{L}}$.

$$\overline{x^2} = \left(\sqrt{\frac{2}{L}}\right)^2 \frac{L^3}{2}\left(\frac{1}{3} - \frac{1}{8\pi^2}\right) = \frac{2}{L}\frac{L^3}{2}\left(\frac{1}{3} - \frac{1}{8\pi^2}\right) = L^2\left(\frac{1}{3} - \frac{1}{8\pi^2}\right) \approx 0.32\,L^2$$

We're not finished yet. We still need to find the variance in the particle's position. Recall from Part B that $\bar{x} = \frac{L}{2}$.

$$\Delta x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{L^2\left(\frac{1}{3} - \frac{1}{8\pi^2}\right) - \left(\frac{L}{2}\right)^2} = \sqrt{\frac{L^2}{3} - \frac{L^2}{8\pi^2} - \frac{L^2}{4}} = L\sqrt{\frac{1}{3} - \frac{1}{8\pi^2} - \frac{1}{4}}$$

$$\Delta x = L\sqrt{\frac{1}{3} - \frac{1}{4} - \frac{1}{8\pi^2}} = L\sqrt{\frac{4}{12} - \frac{3}{12} - \frac{1}{8\pi^2}} = \boxed{L\sqrt{\frac{1}{12} - \frac{1}{8\pi^2}}} \approx \boxed{0.27\,L}$$

(E) Determine the variance in the particle's momentum.

Since the variance in the particle's momentum is defined as $\Delta p_x = \sqrt{\overline{p_x^2} - \overline{p_x}^2}$, we will first need to perform the integral for $\overline{p_x^2}$ (which we will see turns out to be different from $\overline{p_x}^2$).

$$\overline{p_x^2} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{p}_x^2\, \Psi(x,t)\, dx = (-i\hbar)^2 \int_{x=-\infty}^{\infty} \Psi^*(x,t) \frac{\partial^2}{\partial x^2} \Psi(x,t)\, dx$$

$$\overline{p_x^2} = (-1)^2 i^2 \hbar^2 \int_{x=-\infty}^{\infty} \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x)\, dx = (1)(-1)\hbar^2 \int_{x=0}^{L} A \sin\left(\frac{2\pi x}{L}\right) \frac{\partial^2}{\partial x^2} A \sin\left(\frac{2\pi x}{L}\right) dx$$

$$\overline{p_x^2} = -\hbar^2 A^2 \int_{x=0}^{L} \sin\left(\frac{2\pi x}{L}\right)\left[-\left(\frac{2\pi}{L}\right)^2 \sin\left(\frac{2\pi x}{L}\right)\right] dx = \frac{4\pi^2\hbar^2 A^2}{L^2} \int_{x=0}^{L} \sin^2\left(\frac{2\pi x}{L}\right) dx$$

Recall the trig identity $\sin^2\theta = \frac{1-\cos(2\theta)}{2}$. Let $\theta = \frac{2\pi x}{L}$ such that $2\theta = \frac{4\pi x}{L}$.

$$\overline{p_x^2} = \frac{4\pi^2\hbar^2 A^2}{L^2} \int_{x=0}^{L} \frac{1-\cos\left(\frac{4\pi x}{L}\right)}{2} dx = \frac{2\pi^2\hbar^2 A^2}{L^2} \int_{x=0}^{L} dx - \frac{2\pi^2\hbar^2 A^2}{L^2} \int_{x=0}^{L} \cos\left(\frac{4\pi x}{L}\right) dx$$

$$\overline{p_x^2} = \frac{2\pi^2\hbar^2A^2}{L^2}[x]_{x=0}^{L} - \frac{2\pi^2\hbar^2A^2}{L^2}\left(\frac{L}{4\pi}\right)\left[\sin\left(\frac{4\pi x}{L}\right)\right]_{x=0}^{L}$$

$$\overline{p_x^2} = \frac{2\pi^2\hbar^2A^2}{L^2}(L-0) - \frac{\pi\hbar^2A^2}{2L}[\sin(4\pi) - \sin(0)]$$

$$\overline{p_x^2} = \frac{2\pi^2\hbar^2A^2}{L^2}L - \frac{\pi\hbar^2A^2}{2L}(0-0) = \frac{2\pi^2\hbar^2A^2}{L}$$

Recall from Part A that $A = \sqrt{\frac{2}{L}}$.

$$\overline{p_x^2} = \frac{2\pi^2\hbar^2}{L}\left(\sqrt{\frac{2}{L}}\right)^2 = \frac{2\pi^2\hbar^2}{L}\frac{2}{L} = \frac{4\pi^2\hbar^2}{L^2} \approx 39\frac{\hbar^2}{L^2}$$

We're not finished yet. We still need to find the variance in the particle's momentum. Recall from Part C that $\overline{p_x} = 0$.

$$\Delta p_x = \sqrt{\overline{p_x^2} - \overline{p_x}^2} = \sqrt{\frac{4\pi^2\hbar^2}{L^2} - (0)^2} = \sqrt{\frac{4\pi^2\hbar^2}{L^2}} = \boxed{\frac{2\pi\hbar}{L}} \approx \boxed{6.3\frac{\hbar}{L}}$$

(F) Show that the answers to Parts D and E satisfy Heisenberg's uncertainty principle.

Recall Heisenberg's uncertainty principle from Chapter 11: $\Delta x \Delta p_x \geq \frac{\hbar}{2}$. Calculate the product $\Delta x \Delta p_x$ and compare it with $\frac{\hbar}{2}$. Use the answers from Parts D and E.

$$\Delta x \Delta p_x = \left(L\sqrt{\frac{1}{12} - \frac{1}{8\pi^2}} \right)\left(\frac{2\pi\hbar}{L}\right) = 2\pi\hbar\sqrt{\frac{1}{12} - \frac{1}{8\pi^2}} \approx (6.3\ \hbar)(0.27) \approx 1.7\ \hbar$$

The answer is greater than $\frac{\hbar}{2}$, in accordance with Heisenberg's uncertainty principle.

Chapter 14 Problems

1. A particle travels in one dimension with the simple harmonic oscillator potential, $V(x) = \frac{1}{2}Cx^2$. The particle is known to be in the ground state.

(A) Determine the coefficient A_0.

(B) Determine the expectation value for the particle's position. Interpret your answer.

(C) Determine the expectation value for the particle's momentum. Interpret your answer.

(D) Determine the variance in the particle's position.

(E) Determine the variance in the particle's momentum.

(F) Show that your answers to Parts D and E satisfy Heisenberg's uncertainty principle.

Want help? Check the solution at the end of the chapter.

Answers: 1. (A) $A_0 = \frac{C^{1/8}m^{1/8}}{\pi^{1/4}\hbar^{1/4}}$ (B) 0 (C) 0 (D) $\frac{\hbar^{1/2}}{2^{1/2}C^{1/4}m^{1/4}}$ (E) $\frac{\hbar^{1/2}C^{1/4}m^{1/4}}{2^{1/2}}$ (F) $\Delta x \Delta p_x = \frac{\hbar}{2}$

Solutions to Chapter 14

1. According to page 205, for the 1D simple harmonic oscillator, the wave function for the ground state is $\psi_0 = A_0 e^{-\xi^2/2}$, where $\xi = \frac{C^{1/4} m^{1/4}}{\hbar^{1/2}} x$.

(A) We can solve for the coefficient A_0 by normalizing the wave function.

$$\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = \int_{x=-\infty}^{\infty} A_0 e^{-\xi^2/2} A_0 e^{-\xi^2/2}\,dx = A_0^2 \int_{x=-\infty}^{\infty} e^{-\xi^2}\,dx = 1$$

In the last step, we used the identity $e^{-y}e^{-y} = e^{-2y}$. Since $\xi = \frac{C^{1/4} m^{1/4}}{\hbar^{1/2}} x$, it follows that $\xi^2 = \frac{C^{1/2} m^{1/2}}{\hbar} x^2$. Note that $\left(\hbar^{1/2}\right)^2 = \hbar^{1/2}\hbar^{1/2} = \hbar^{1/2+1/2} = h^1 = \hbar$ and that $\left(C^{1/4}\right)^2 = C^{1/2}$. We can write this in the form $\xi^2 = ax^2$, where $a = \frac{C^{1/2} m^{1/2}}{\hbar}$.

$$A_0^2 \int_{x=-\infty}^{\infty} e^{-ax^2}\,dx = 1$$

According to page 203, $\int_{x=-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$.

$$A_0^2 \sqrt{\frac{\pi}{a}} = 1 \quad \to \quad A_0^2 = \sqrt{\frac{a}{\pi}} = \left(\frac{a}{\pi}\right)^{1/2} \quad \to \quad A_0 = \left(\frac{a}{\pi}\right)^{1/4}$$

$$A_0 = \left(\frac{C^{1/2} m^{1/2}}{\pi\hbar}\right)^{1/4} = \boxed{\frac{C^{1/8} m^{1/8}}{\pi^{1/4} \hbar^{1/4}}}$$

We applied the rule $\left(y^b\right)^c = y^{bc}$. For example, $\left(m^{1/2}\right)^{1/4} = m^{1/8}$. Note that there are two ways to go about normalizing this wave function.

- We thought of ψ as a function of x, $\psi(x)$, expressed the normalization condition in the form $\int_{x=-\infty}^{\infty} \psi^*(x)\psi(x)\, dx = 1$, and found that $A_0 = \frac{C^{1/8} m^{1/8}}{\pi^{1/4}\hbar^{1/4}}$, where $\psi_0(x) = A_0 e^{-ax^2/2}$.
- We could think of ψ as a function of ξ, $\psi(\xi)$, express the normalization condition in the form $\int_{\xi=-\infty}^{\infty} \psi^*(\xi)\psi(\xi)\, d\xi = 1$, and find that $B_0 = \frac{1}{\pi^{1/4}}$, where $\psi_0(\xi) = B_0 e^{-\xi^2/2}$.

The difference is a simple change of coordinates. It's just a matter of perspective.

(B) Perform the integral for $\bar{x}$. Recall from Chapter 13 that $\Psi(x,t) = \psi(x)e^{-iEt/\hbar}$: Thus $\Psi^* x \Psi = \psi^* e^{iEt/\hbar} x \psi e^{-iEt/\hbar} = \psi^* x \psi$. Recall from Part A that $\psi_0 = A_0 e^{-\xi^2/2}$, where $\xi = \frac{C^{1/4} m^{1/4}}{\hbar^{1/2}} x$.

$$\bar{x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x\, \Psi(x,t)\, dx = \int_{x=-\infty}^{\infty} \psi^*(x)\, x\, \psi(x)\, dx$$

$$\bar{x} = \int_{x=-\infty}^{\infty} A_0 e^{-\xi^2/2}\, x\, A_0 e^{-\xi^2/2}\, dx = A_0^2 \int_{x=-\infty}^{\infty} x\, e^{-\xi^2}\, dx$$

In the last step, we used the identity $e^{-y} e^{-y} = e^{-2y}$. Just like we did in Part A, we will write $\xi^2 = ax^2$, where $a = \frac{C^{1/2} m^{1/2}}{\hbar}$.

$$\bar{x} = A_0^2 \int_{x=-\infty}^{\infty} x\, e^{-ax^2}\, dx$$

According to page 203, $\int_{x=-\infty}^{\infty} x e^{-ax^2}\, dx = 0$. This is because the integrand is an odd function (since x is an odd function while e^{-ax^2} is an even function) while the limits of integration are symmetric. That is, by asymmetry, the negative area integrated over the interval $-\infty < x \le 0$ cancels the positive area integrated over the interval $0 \le x < \infty$.

$$\boxed{\bar{x} = 0}$$

Physically, for the 1D simple harmonic oscillator, the particle is oscillating back and forth about the equilibrium position. Thus, it should make sense that the particle is at the equilibrium position on *average*, where $x_e = 0$ since $V = \frac{1}{2}Cx^2$. Recall from Hooke's law problems in first-year physics that $V = \frac{1}{2}C(x - x_e)^2$ and $F = -C(x - x_e)$ are the potential energy and force for a spring oscillating about the equilibrium position x_e (where the spring constant has been written as C instead of the usual k in order to avoid confusion with the wave number).

(C) Perform the integral for $\overline{p_x}$. Recall from Chapter 13 that $\Psi(x,t) = \psi(x)e^{-iEt/\hbar}$: Thus $\Psi^* \hat{p}_x \Psi = -i\hbar\Psi^* \frac{\partial}{\partial x}\Psi = -i\hbar\, \psi^* e^{iEt/\hbar} \frac{\partial}{\partial x}\psi e^{-iEt/\hbar} = -i\hbar\psi^* \frac{\partial}{\partial x}\psi$. Recall from Chapter 13 that the momentum operator is $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ in one dimension.

$$\overline{p_x} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \hat{p}_x\, \Psi(x,t)\, dx = -i\hbar \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, \frac{\partial}{\partial x}\, \Psi(x,t)\, dx$$

$$\overline{p_x} = -i\hbar \int_{x=-\infty}^{\infty} \psi^*(x)\, \frac{\partial}{\partial x}\, \psi(x)\, dx = -i\hbar \int_{x=-\infty}^{\infty} A_0 e^{-\xi^2/2}\, \frac{\partial}{\partial x}\, A_0 e^{-\xi^2/2}\, dx$$

$$\overline{p_x} = -i\hbar A_0^2 \int\limits_{x=-\infty}^{\infty} e^{-\xi^2/2} \frac{\partial}{\partial x} e^{-\xi^2/2}\, dx$$

Note that the derivative is with respect to x. Recall that $\xi = \frac{C^{1/4}m^{1/4}}{\hbar^{1/2}} x$. We either need to make this substitution now or apply the chain rule: We will make the substitution. Just like we did in Part A, we will write $\xi^2 = ax^2$, where $a = \frac{C^{1/2}m^{1/2}}{\hbar}$.

$$\overline{p_x} = -i\hbar A_0^2 \int\limits_{x=-\infty}^{\infty} e^{-ax^2/2} \frac{\partial}{\partial x} e^{-ax^2/2}\, dx = -i\hbar A_0^2 \int\limits_{x=-\infty}^{\infty} e^{-ax^2/2} (-ax)\, e^{-ax^2/2}\, dx$$

Use the identity $e^{-y}e^{-y} = e^{-2y}$. Note that the two minus signs cancel.

$$\overline{p_x} = i\hbar A_0^2 a \int\limits_{x=-\infty}^{\infty} x\, e^{-ax^2}\, dx = i\hbar A_0^2 \frac{C^{1/2}m^{1/2}}{\hbar} \int\limits_{x=-\infty}^{\infty} x\, e^{-ax^2}\, dx$$

$$\overline{p_x} = iA_0^2 C^{1/2} m^{1/2} \int\limits_{x=-\infty}^{\infty} x\, e^{-ax^2}\, dx$$

According to page 203, $\int_{x=-\infty}^{\infty} xe^{-ax^2}\,dx = 0$. This is the same integral that we encountered in Part B.

$$\boxed{\overline{p_x} = 0}$$

Physically, for the 1D simple harmonic oscillator, the particle is oscillating back and forth about the equilibrium position. Thus, it should make sense that the particle has zero momentum on *average*, since the forward momentum of its oscillation will cancel the backward momentum.

(D) Since the variance in the particle's position is defined as $\Delta x = \sqrt{\overline{x^2} - \bar{x}^2}$, we will first need to perform the integral for $\overline{x^2}$ (which turns out to be different from $\bar{x}^2$).

$$\overline{x^2} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\, x^2\, \Psi(x,t)\, dx = \int_{x=-\infty}^{\infty} \psi^*(x)\, x^2\, \psi(x)\, dx$$

$$\overline{x^2} = \int_{x=-\infty}^{\infty} A_0 e^{-\xi^2/2}\, x^2\, A_0 e^{-\xi^2/2}\, dx = A_0^2 \int_{x=-\infty}^{\infty} x^2\, e^{-\xi^2}\, dx$$

In the last step, we used the identity $e^{-y}e^{-y} = e^{-2y}$. Just like we did in Part A, we will write $\xi^2 = ax^2$, where $a = \frac{C^{1/2} m^{1/2}}{\hbar}$.

$$\overline{x^2} = A_0^2 \int_{x=-\infty}^{\infty} x^2\, e^{-ax^2}\, dx$$

According to page 203, $\int_{x=-\infty}^{\infty} x^2 e^{-ax^2}\, dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$.

$$\overline{x^2} = \frac{A_0^2}{2a}\sqrt{\frac{\pi}{a}} = \frac{A_0^2 \pi^{1/2}}{2aa^{1/2}} = \frac{A_0^2 \pi^{1/2}}{2a^{3/2}}$$

Note that $\sqrt{a} = a^{1/2}$ and $aa^{1/2} = a^1 a^{1/2} = a^{1+1/2} = a^{3/2}$. Recall from Part A that $A_0 = \left(\frac{a}{\pi}\right)^{1/4}$.

$$\overline{x^2} = A_0^2 \frac{\pi^{1/2}}{2a^{3/2}} = \left[\left(\frac{a}{\pi}\right)^{1/4}\right]^2 \frac{\pi^{1/2}}{2a^{3/2}} = \frac{a^{1/2}}{\pi^{1/2}} \frac{\pi^{1/2}}{2a^{3/2}} = \frac{1}{2a}$$

Note that $\left[\left(\frac{a}{\pi}\right)^{1/4}\right]^2 = \frac{a^{1/2}}{\pi^{1/2}}$ because $(y^b)^c = y^{bc}$ and $\left(\frac{1}{4}\right)(2) = \frac{2}{4} = \frac{1}{2}$. Also note that $\frac{a^{1/2}}{a^{3/2}} = \frac{1}{a}$ because $\frac{y^m}{y^n} = y^{m-n}$, $\frac{1}{2} - \frac{3}{2} = -\frac{2}{2} = -1$, and $a^{-1} = \frac{1}{a}$. Recall from Part A that $a = \frac{C^{1/2} m^{1/2}}{\hbar}$, such that $\frac{1}{a} = \frac{\hbar}{C^{1/2} m^{1/2}}$.

$$\overline{x^2} = \frac{1}{2a} = \frac{\hbar}{2C^{1/2} m^{1/2}} = \frac{\hbar}{2\sqrt{Cm}}$$

We're not finished yet. We still need to find the variance in the particle's position. Recall from Part B that $\bar{x} = 0$. Note, for example, that $\sqrt{C^{1/2}} = \left(C^{1/2}\right)^{1/2} = C^{1/4}$ because $(y^b)^c = y^{bc}$.

$$\Delta x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{\frac{\hbar}{2C^{1/2}m^{1/2}} - (0)^2} = \sqrt{\frac{\hbar}{2C^{1/2}m^{1/2}}} = \boxed{\frac{\hbar^{1/2}}{2^{1/2}C^{1/4}m^{1/4}}}$$

(E) Since the variance in the particle's momentum is defined as $\Delta p_x = \sqrt{\overline{p_x^2} - \overline{p_x}^2}$, we will first need to perform the integral for $\overline{p_x^2}$ (which turns out to be different from $\overline{p_x}^2$).

$$\overline{p_x^2} = \int_{x=-\infty}^{\infty} \Psi^*(x,t)\,\hat{p}_x^2\,\Psi(x,t)\,dx = (-i\hbar)^2 \int_{x=-\infty}^{\infty} \Psi^*(x,t)\,\frac{\partial^2}{\partial x^2}\,\Psi(x,t)\,dx$$

$$\overline{p_x^2} = (-1)^2 i^2 \hbar^2 \int_{x=-\infty}^{\infty} \psi^*(x)\,\frac{\partial^2}{\partial x^2}\,\psi(x)\,dx = (1)(-1)\hbar^2 \int_{x=-\infty}^{\infty} A_0 e^{-\xi^2/2}\,\frac{\partial^2}{\partial x^2}\,A_0 e^{-\xi^2/2}\,dx$$

$$\overline{p_x^2} = -\hbar^2 A_0^2 \int_{x=-\infty}^{\infty} e^{-\xi^2/2}\,\frac{\partial^2}{\partial x^2}\,e^{-\xi^2/2}\,dx$$

Note that the derivative is with respect to x. Recall that $\xi = \frac{C^{1/4}m^{1/4}}{\hbar^{1/2}}x$. We either need to make this substitution now or apply the chain rule: We will make the substitution. Just like we did in Part A, we will write $\xi^2 = ax^2$, where $a = \frac{C^{1/2}m^{1/2}}{\hbar}$.

$$\overline{p_x^2} = -\hbar^2 A_0^2 \int_{x=-\infty}^{\infty} e^{-ax^2/2} \frac{\partial^2}{\partial x^2} e^{-ax^2/2}\, dx$$

Let's work out this second derivative one step at a time. (Students who don't realize that the second derivative will involve the product rule often wind up with a critical sign mistake.)

$$\frac{\partial^2}{\partial x^2} e^{-ax^2/2} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} e^{-ax^2/2}\right) = \frac{\partial}{\partial x}\left[\left(-\frac{2ax}{2}\right) e^{-ax^2/2}\right] = -a\frac{\partial}{\partial x}\left(x\, e^{-ax^2/2}\right)$$

Apply the product rule with $f = x$ and $g = e^{-ax^2/2}$.

$$\frac{\partial^2}{\partial x^2} e^{-ax^2/2} = -a\frac{\partial}{\partial x}(fg) = -a\left(g\frac{\partial f}{\partial x} + f\frac{\partial g}{\partial x}\right) = -a\left(e^{-ax^2/2}\frac{\partial}{\partial x}x + x\frac{\partial}{\partial x}e^{-ax^2/2}\right)$$

$$\frac{\partial^2}{\partial x^2} e^{-ax^2/2} = -a\left[e^{-ax^2/2}(1) + x\left(-\frac{2ax}{2}\right)e^{-ax^2/2}\right] = -a\left(e^{-ax^2/2} - ax^2 e^{-ax^2/2}\right)$$

$$\frac{\partial^2}{\partial x^2} e^{-ax^2/2} = -ae^{-ax^2/2} + a^2x^2e^{-ax^2/2}$$

Substitute this second derivative into the previous integral.

$$\overline{p_x^2} = -\hbar^2 A_0^2 \int_{x=-\infty}^{\infty} e^{-ax^2/2}\left(-ae^{-ax^2/2} + a^2x^2e^{-ax^2/2}\right) dx$$

$$\overline{p_x^2} = -\hbar^2 A_0^2 \int_{x=-\infty}^{\infty} -ae^{-ax^2}\, dx - \hbar^2 A_0^2 \int_{x=-\infty}^{\infty} a^2x^2e^{-ax^2}\, dx$$

In the last step, we used the identity $e^{-y}e^{-y} = e^{-2y}$. The first two minus signs cancel.

$$\overline{p_x^2} = \hbar^2 A_0^2 a \int_{x=-\infty}^{\infty} e^{-ax^2}\, dx - \hbar^2 A_0^2 a^2 \int_{x=-\infty}^{\infty} x^2e^{-ax^2}\, dx$$

According to page 203, $\int_{x=-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$ and $\int_{x=-\infty}^{\infty} x^2e^{-ax^2}\, dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$.

$$\overline{p_x^2} = \hbar^2 A_0^2 a \sqrt{\frac{\pi}{a}} - \hbar^2 A_0^2 a^2 \frac{1}{2a}\sqrt{\frac{\pi}{a}} = \hbar^2 A_0^2 a \sqrt{\frac{\pi}{a}} - \frac{\hbar^2 A_0^2 a}{2}\sqrt{\frac{\pi}{a}} = \hbar^2 A_0^2 a \left(1 - \frac{1}{2}\right)\sqrt{\frac{\pi}{a}}$$

In the last step, we factored out $\hbar^2 A_0^2 a \sqrt{\frac{\pi}{a}}$. Note that $1 - \frac{1}{2} = \frac{1}{2}$.

$$\overline{p_x^2} = \hbar^2 A_0^2 a \left(\frac{1}{2}\right) \sqrt{\frac{\pi}{a}} = \frac{\hbar^2 A_0^2 a}{2} \sqrt{\frac{\pi}{a}} = \frac{\hbar^2 A_0^2 a \pi^{1/2}}{2a^{1/2}} = \frac{\hbar^2 A_0^2 a^{1/2} \pi^{1/2}}{2}$$

Note, for example, that $\sqrt{\pi} = \pi^{1/2}$. Recall from Part A that $A_0 = \left(\frac{a}{\pi}\right)^{1/4}$.

$$\overline{p_x^2} = A_0^2 \frac{\hbar^2 a^{1/2} \pi^{1/2}}{2} = \left[\left(\frac{a}{\pi}\right)^{1/4}\right]^2 \frac{\hbar^2 a^{1/2} \pi^{1/2}}{2} = \frac{a^{1/2}}{\pi^{1/2}} \frac{\hbar^2 a^{1/2} \pi^{1/2}}{2} = \frac{\hbar^2 a}{2}$$

Note that $\left[\left(\frac{a}{\pi}\right)^{1/4}\right]^2 = \frac{a^{1/2}}{\pi^{1/2}}$ because $(y^b)^c = y^{bc}$ and $\left(\frac{1}{4}\right)(2) = \frac{2}{4} = \frac{1}{2}$. Recall from Part A that $a = \frac{C^{1/2} m^{1/2}}{\hbar}$.

$$\overline{p_x^2} = \frac{\hbar^2 C^{1/2} m^{1/2}}{2\hbar} = \frac{\hbar C^{1/2} m^{1/2}}{2}$$

We're not finished yet. We still need to find the variance in the particle's momentum. Recall from Part C that $\overline{p_x} = 0$.

$$\Delta p_x = \sqrt{\overline{p_x^2} - \overline{p_x}^2} = \sqrt{\frac{\hbar C^{1/2} m^{1/2}}{2} - (0)^2} = \sqrt{\frac{\hbar C^{1/2} m^{1/2}}{2}} = \boxed{\frac{\hbar^{1/2} C^{1/4} m^{1/4}}{2^{1/2}}}$$

Note: Some books express their answers in terms of the angular frequency (ω) rather than the spring constant (C). Recall from first-year physics that $\omega = \sqrt{\frac{C}{m}}$. (Recall that we are using C for the spring constant instead of the usual k in order to avoid possible confusion with the wave number.) Square both sides to see that $C = m\omega^2$. Multiply by m to see that $Cm = m^2\omega^2$. Raise this to various powers to get the following equations.

$$C^{1/2}m^{1/2} = m\omega \quad , \quad C^{1/4}m^{1/4} = m^{1/2}\omega^{1/2} \quad , \quad C^{1/8}m^{1/8} = m^{1/4}\omega^{1/4}$$

(F) Recall Heisenberg's uncertainty principle from Chapter 11: $\Delta x \Delta p_x \geq \frac{\hbar}{2}$. Calculate $\Delta x \Delta p_x$ and compare it with $\frac{\hbar}{2}$. Recall from Parts D and E that $\Delta x = \frac{\hbar^{1/2}}{2^{1/2}C^{1/4}m^{1/4}}$ and $\Delta p_x = \frac{\hbar^{1/2}C^{1/4}m^{1/4}}{2^{1/2}}$.

$$\Delta x \Delta p_x = \left(\frac{\hbar^{1/2}}{2^{1/2}C^{1/4}m^{1/4}}\right)\left(\frac{\hbar^{1/2}C^{1/4}m^{1/4}}{2^{1/2}}\right) = \frac{\hbar^{1/2}\hbar^{1/2}C^{1/4}m^{1/4}}{2^{1/2}C^{1/4}m^{1/4}2^{1/2}} = \frac{\hbar^{1/2}\hbar^{1/2}}{2^{1/2}2^{1/2}} = \boxed{\frac{\hbar}{2}}$$

Recall the rule that $\hbar^{1/2}\hbar^{1/2} = \hbar^{1/2+1/2} = \hbar^1 = \hbar$. The answer is exactly equal to $\frac{\hbar}{2}$, which agrees with the greater than or equal sign ($\geq$) in Heisenberg's uncertainty principle. (This happens to be a rare case where $\Delta x \Delta p_x$ is exactly equal to $\frac{\hbar}{2}$.)

WAS THIS BOOK HELPFUL?

A great deal of effort and thought was put into this book, such as:

- Breaking down the solutions to help make physics easier to understand.
- Careful selection of problems for their instructional value.
- Multiple stages of proofreading, editing, and formatting.
- Physics instructors and students provided valuable feedback.

If you appreciate the effort that went into making this book possible, there is a simple way that you could show it:

Please take a moment to post an honest review.

For example, you can review this book at Amazon.com or BN.com (for Barnes & Noble).

Even a short review can be helpful and will be much appreciated. If you're not sure what to write, following are a few ideas, though it's best to describe what's important to you.

- Were you able to understand the explanations?
- Did you appreciate the list of symbols and units?
- Was it easy to find the information you were looking for?
- How much did you learn from reading through the examples?
- Would you recommend this book to others? If so, why?

GET A DIFFERENT ANSWER?

If you get a different answer and can't find your mistake even after consulting the hints and explanations, what should you do?

Please contact the author, Dr. McMullen.

How? Visit one of the author's blogs (see below). Either use the Contact Me option, or click on one of the author's articles and post a comment on the article.

monkeyphysicsblog.wordpress.com
improveyourmathfluency.com
chrismcmullen.com

Why?

- If there happens to be a mistake (although much effort was put into perfecting the answer key), the correction will benefit other students like yourself in the future.
- If it turns out not to be a mistake, **you may learn something** from Dr. McMullen's reply to your message.

99.99% of students who walk into Dr. McMullen's office believing that they found a mistake with an answer discover one of two things:

- They made a mistake that they didn't realize they were making and learned from it.
- They discovered that their answer was actually the same. This is actually fairly common. For example, the answer key might say $t = \frac{\sqrt{3}}{3}$ s. A student solves the problem and gets $t = \frac{1}{\sqrt{3}}$ s. These are actually the same: Try it on your calculator and you will see that both equal about 0.57735. Here's why: $\frac{1}{\sqrt{3}} = \frac{1}{\sqrt{3}}\frac{\sqrt{3}}{\sqrt{3}} = \frac{\sqrt{3}}{3}$.

Every effort was made to ensure that the final answer given to every problem is correct. But all humans, even those who are experts in their fields and who routinely aced exams back when they were students, make an occasional mistake. So if you believe you found a mistake, you should report it just in case. Dr. McMullen will appreciate your time.

ABOUT THE AUTHOR

D Dr. Chris McMullen has over 20 years of experience teaching university physics in California,

Oklahoma, Pennsylvania, and Louisiana. Dr. McMullen is also an author of math and science workbooks. Whether in the classroom or as a writer, Dr. McMullen loves sharing knowledge and the art of motivating and engaging students.

The author earned his Ph.D. in phenomenological high-energy physics (particle physics) from Oklahoma State University in 2002. Originally from California, Chris McMullen earned his Master's degree from California State University, Northridge, where his thesis was in the field of electron spin resonance.

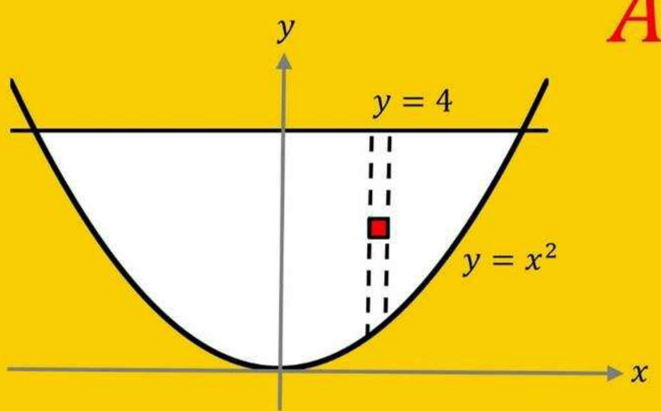
As a physics teacher, Dr. McMullen observed that many students lack fluency in fundamental math skills. In an effort to help students of all ages and levels master basic math skills, he published a series of math workbooks on arithmetic, fractions, long division, algebra, trigonometry, and calculus entitled *Improve Your Math Fluency*. Dr. McMullen has also published a variety of science books, including introductions to basic astronomy and chemistry concepts in addition to physics workbooks.

www.amazon.com/author/chrismcmullen

Author, Chris McMullen, Ph.D.

50 CHALLENGING CALCULUS PROBLEMS

$$A = \int_{x=-2}^{2} (4 - x^2)\,dx$$



FULLY SOLVED

Chris McMullen, Ph.D.

SCIENCE

Dr. McMullen has published a variety of science books, including:

- Basic astronomy concepts
- Basic chemistry concepts
- Balancing chemical reactions
- Calculus-based physics textbook
- Calculus-based physics workbooks
- Calculus-based physics examples
- Trig-based physics workbooks
- Trig-based physics examples
- Creative physics problems
- Modern physics

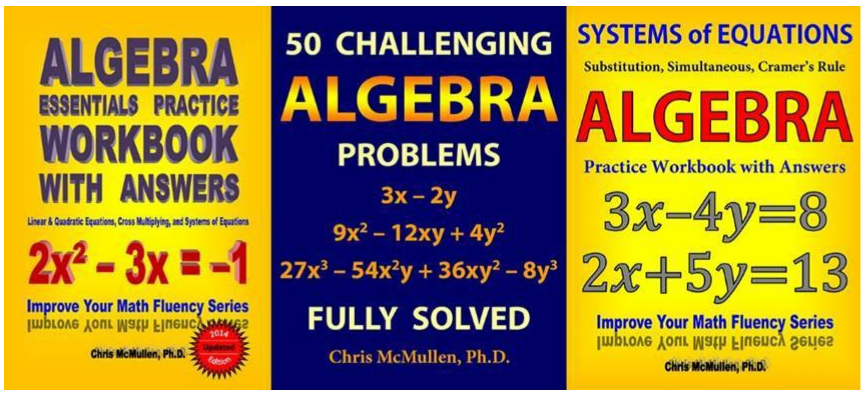
www.monkeyphysicsblog.wordpress.com

ALGEBRA

For students who need to improve their algebra skills:

- Isolating the unknown
- Quadratic equations
- Factoring
- Cross multiplying
- Systems of equations
- Straight line graphs

www.improveyourmathfluency.com



PUZZLES

The author of this book, Chris McMullen, enjoys solving puzzles. His favorite puzzle is Kakuro (kind of like a cross between crossword puzzles and Sudoku). He once taught a three-week summer course on puzzles. If you enjoy mathematical pattern puzzles, you might appreciate:

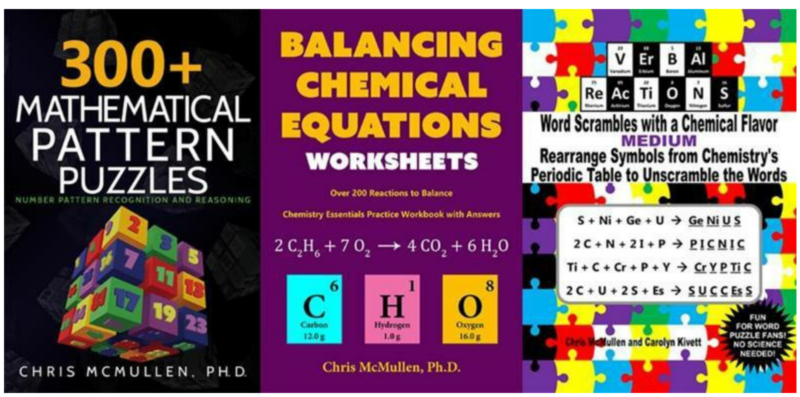
300 + Mathematical Pattern Puzzles

Number Pattern Recognition & Reasoning:

- Pattern recognition
- Visual discrimination
- Analytical skills
- Logic and reasoning
- Analogies
- Mathematics

Dr. McMullen's newest math puzzle book is:

Pyramid Math Puzzle Challenge



THE FOURTH DIMENSION

Are you curious about a possible fourth dimension of space?

- Explore the world of hypercubes and hyperspheres.
- Imagine living in a two-dimensional world.
- Try to understand the fourth dimension by analogy.
- Several illustrations help to try to visualize a fourth dimension of space.
- Investigate hypercube patterns.
- What would it be like to be a four-dimensional being living in a four-dimensional world?
- Learn about the physics of a possible four-dimensional universe.