2018–2019

SCIENCE

Resource Guide

An Introduction to Laser Technology and Its Applications

USAD United States Academic Decathlon®

The vision of the United States Academic Decathlon® is to provide all students the opportunity to excel academically through team competition.

Toll Free: 866-511-USAD (8723) • **Direct:** 712-326-9589 • **Fax:** 651-389-9144 • **Email:** info@usad.org • **Website:** www.usad.org

Table of Contents

Introduction

The laser is among the most important inventions of the twentieth century. Since its introduction in 1960, the laser has made possible a countless number of scientific, medical, industrial, and commercial applications. Theodore Maiman, the inventor of the first working laser, was quoted by *The New York Times* in 1964 as calling the laser "a solution seeking a problem." Today it's nearly impossible to go an entire day without using at least one object that was either manufactured using lasers or that uses lasers in order to function. In this *Science Resource Guide*, we will explore the nature of light and optics, the basic operation of a laser, and the various applications of lasers that have transformed our lives over the course of more than half a century. Along the way, we will consider significant milestones in the development of the laser, as well as the people who contributed to our understanding of both lasers and light itself. You'll also have opportunities to explore light-related phenomena using online simulations and simple demonstrations with everyday materials.

Section I will cover the basic properties of light as an electromagnetic wave. In this section, we will discuss how visible light represents a relatively small segment of the complete electromagnetic spectrum. We will also explore how the colors that we see are caused by interactions between visible light and matter.

In Section II, we will consider the field of optics, which concerns the behavior of light, the formation of images, and the phenomena of reflection, refraction, and diffraction. As we will see, lenses and mirrors can be configured to focus light in particular ways, which will become important in understanding laser operation and applications.

Section III will provide an overview of the laser itself. We will start by considering how light emission can be explained using a quantum model of atomic excitation, and explore various mechanisms of light emission. By the end of this section, you will be familiar with the most important properties of laser light, the main categories of lasers, and the history behind the development of the laser.

Finally, Section IV will provide an overview of major applications of lasers across several categories. Lasers are an important tool in a variety of research fields, and they have also made possible new surgical procedures and manufacturing techniques. Lasers are all around us, in laser pointers, laser printers, and optical disc readers. By the end of this guide, you will be able to explain how lasers work and the significant ways in which they contribute to our daily lives.

NOTE TO STUDENTS: You will notice as you read through the Resource Guide that some key terms and phrases are boldfaced. These terms are included in the glossary at the end of the Resource Guide.

Section I

Light

INTRODUCTION

Light is both a familiar and fascinating phenomenon. Although a **laser** beam and a flashlight beam have very different properties, they are still governed by the same fundamental principles that describe light. In this section, we will set the stage for our discussion of the laser by covering some of the most fundamental properties of light and color.

PROPERTIES OF LIGHT

WAVES

Many phenomena in the known universe are described by the general mathematical concept of *waves*. For example, the beam of light produced by a laser, the sound generated by a thunderclap, and the ripples of water in a pond are all waves. A wave is the motion of a disturbance between two points. You've likely observed how dropping a small object, such as a pebble, into a pond causes ripples of water to spread outward from where the object lands. The pebble disturbs the water immediately surrounding it, which disturbs the water further out, which disturbs the water further out, and so on (Figure 1).

FIGURE 1

A disturbance event, such as a dropped pebble, causes waves on the surface of a still pond.

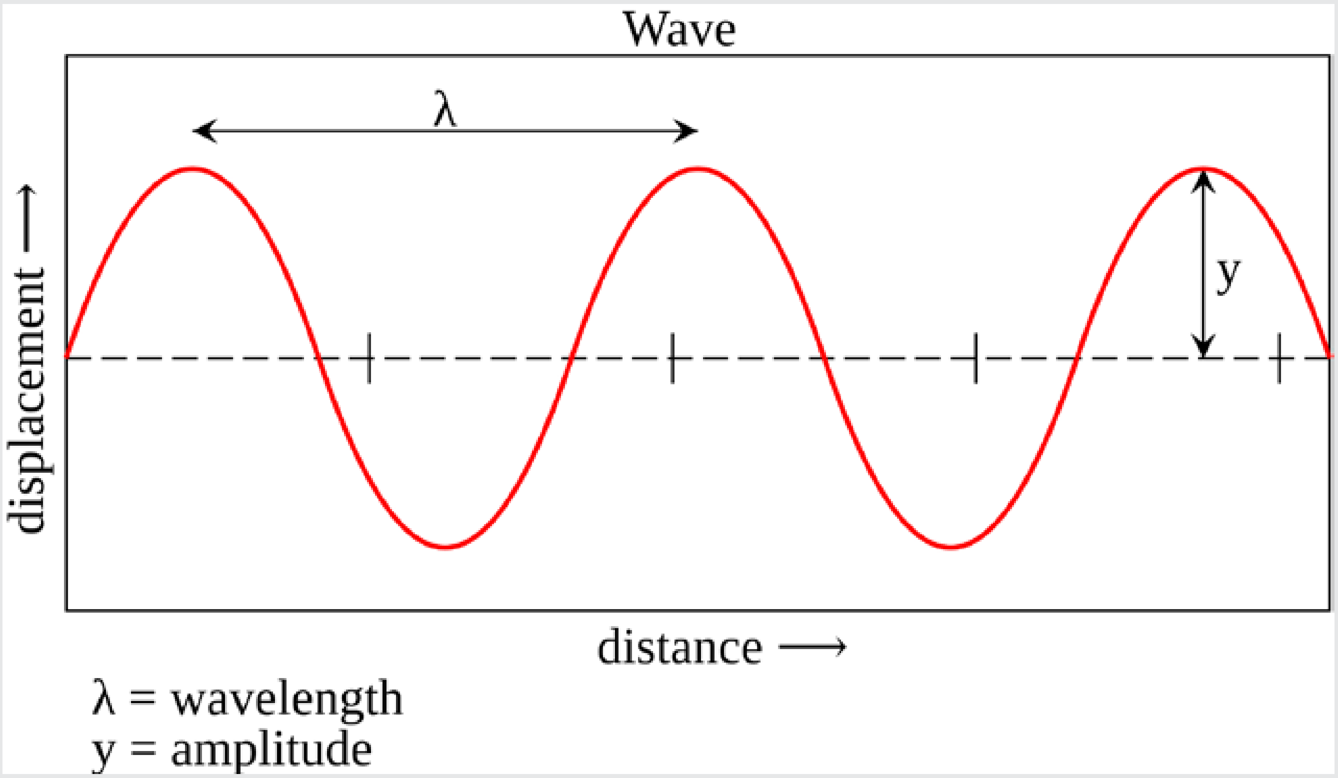
Waves can be classified as transverse or longitudinal. A **transverse wave** is made up of oscillations that are perpendicular to the direction that the wave travels. As we will see, light is an example of a transverse wave. In contrast, for a **longitudinal wave**, the oscillations are parallel to the direction that the wave travels. Sound is an example of a longitudinal wave. Regardless of their classification, all waves are caused by a vibration of some kind. For example, sound waves originate from vibrations of matter, such as a person's vocal chords or a guitar string, which initiate a sequence of collisions between air molecules that eventually reaches your eardrum.

Waves transport *energy* but not *mass*. That is, waves do not carry matter along with them from one point to another as they travel. Think about a floating buoy bobbing up and down in the ocean as waves pass. Although the height of the buoy changes with time, on average its position along the surface of the water does not change. Likewise, individual particles that make up the wave are not moving along with it. Likewise, a sound wave is not carried by individual air molecules moving all the way from the source to your ear. Although the individual molecules vibrate back and forth along the direction of the wave, they exhibit no net motion toward your ear.

WAVE PROPERTIES

Although different types of waves have different physical properties, there are general characteristics that we can use to describe any wave. The **amplitude** of a wave is the maximum displacement from the equilibrium position. For a water wave or a wave on a string, the amplitude is the height of the wave compared to its resting position. For electromagnetic waves, the amplitude refers to the maximum magnitude of the electric and magnetic fields. In general, the amplitude of a wave is related to how much energy it carries. The **wavelength** λ of a wave is defined

FIGURE 2



The amplitude (y) and wavelength (λ) of a general waveform.

as the distance between two consecutive corresponding points on a wave—for example, the distance between two successive peaks (FIGURE 2).

A wave can also be described in terms of its **frequency** f, which is defined as the number of oscillations it makes in a given amount of time. (For instance, think of how many ripples pass by a certain point in a pond every second.) The standard unit of frequency is the hertz (Hz), which equals one oscillation per second. The **period** T of a wave is the amount of time that elapses between oscillations and is related to the frequency by $T = 1/f$. For example, if the frequency of a wave is 10 Hz (i.e., 10 oscillations per second), the corresponding period is 0.1 seconds because each oscillation is 1/10 seconds apart.

Another important characteristic of a wave is the **wave speed** v. The speed of any wave is related to both its frequency and its wavelength. Let's determine the exact relationship by considering the motion of a wave past a fixed point, such as crests of water past the edge of a dock. We can use a ruler to measure how far apart the crests are (i.e., the wavelength) and a stopwatch to measure how much time passes between the arrival of one wave crest and the next (the period). Speed is a distance traveled per unit time, and in the case of a wave the distance is the wavelength, and the time is the period.

$$v = \frac{\text{distance}}{\text{time}} = \frac{\lambda}{T}$$

Recalling that period is the reciprocal of frequency, we can convert this equation to the often more useful form:

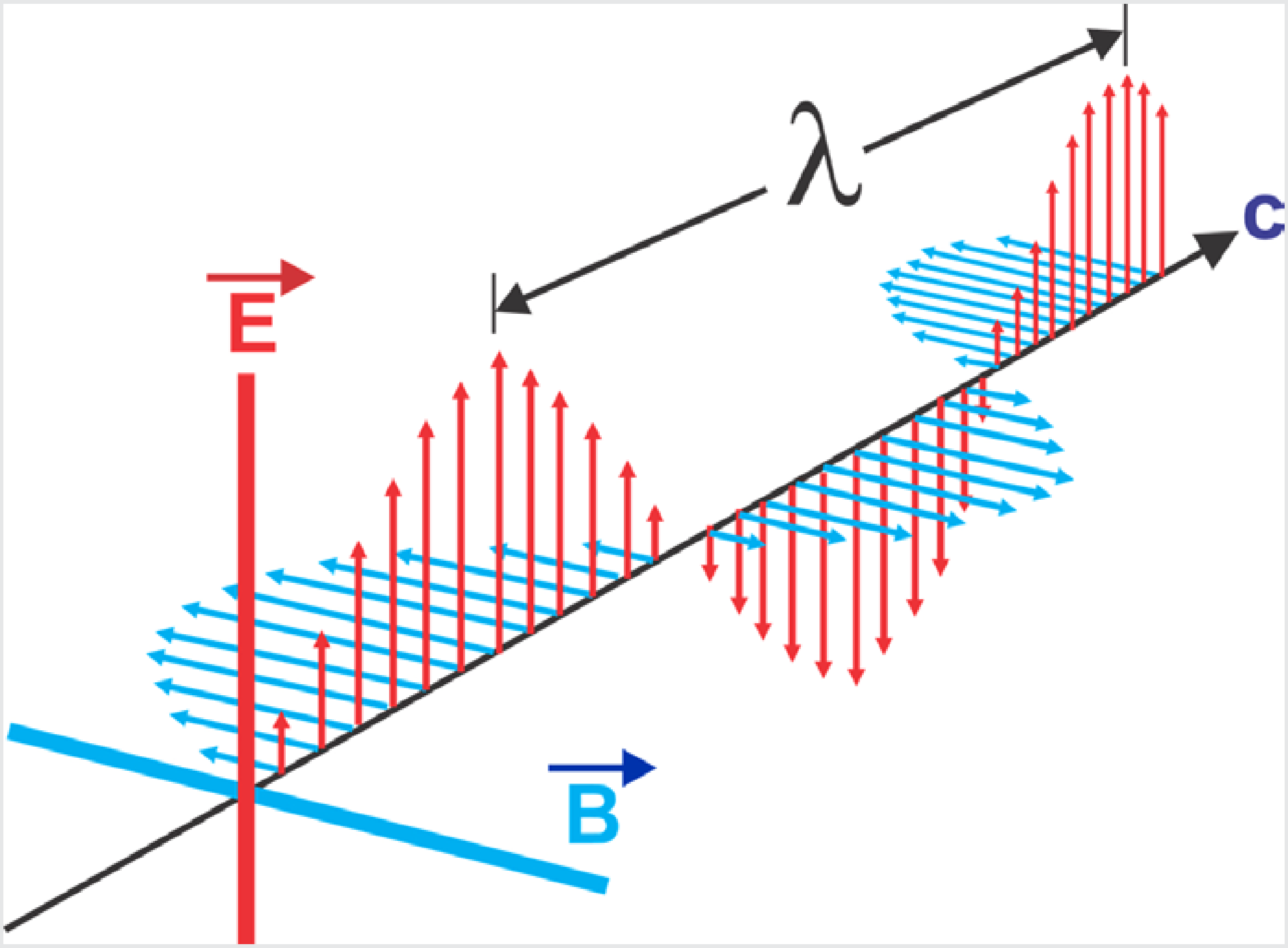
$$v = f\lambda.$$

This relationship applies to any kind of wave, from light to sound to water. Notice that according to this equation, the higher the frequency of a wave, the shorter its wavelength.

ELECTROMAGNETIC WAVES

Light is an **electromagnetic wave**, which is a wave made up of electric and magnetic fields that oscillate as the wave travels (FIGURE 3). Although these electric and magnetic fields are continuously changing in magnitude,

FIGURE 3



An electromagnetic wave is made up of oscillating electric and magnetic fields. The wave moves in the direction perpendicular to both oscillating fields.

they are always directed perpendicular to one another. Furthermore, the fields are always directed perpendicular to the direction of wave propagation. For this reason, electromagnetic waves are classified as transverse. Electromagnetic waves also do not require a physical medium through which to propagate. Unlike **mechanical waves**, such as sound, electromagnetic waves can travel through the empty vacuum of space. This is how light from the Sun and other stars reaches Earth.

We mentioned earlier that waves are caused by some kind of initial vibration. Electromagnetic waves are caused by the vibrations of charged particles, such as electrons and protons. Charged particles are surrounded by electric fields that exert forces on other charged particles. When a charged particle accelerates, the fluctuation in its electric field causes a self-perpetuating "pulse" that spreads out like the ripples in a pond. In the early 1860s, Scottish physicist James Clerk Maxwell derived a system of equations that describes the behavior of electric and magnetic fields, including how a changing electric field can generate a magnetic field and vice versa. In a paper presented to the Royal Society in 1864, Maxwell used these equations to predict the existence of electromagnetic waves. When a sinusoidally-oscillating charged particle creates an oscillating electric field, then an oscillating perpendicular magnetic field is also created. In conjunction, these two mutually perpendicular oscillating fields form an electromagnetic wave.

THE SPEED OF LIGHT

Electromagnetic waves traveling through empty space never speed up or slow down. Indeed, it is a fundamental tenet of physics that *all electromagnetic waves move at the same speed through a vacuum*. Why should this be so? Suppose an electromagnetic wave were to gradually lose speed. According to the laws of electromagnetism, the oscillating electric field would generate a weaker magnetic field, which would in turn generate a weaker electric field, and so on. In short, the wave would die out, and energy would be lost—a violation of the principle of energy conservation. Using the same logic, an electromagnetic wave could not speed up because the electric and magnetic fields would continually strengthen one another, and the energy of the wave would increase without end. As it turns out, there is only one speed at which the electric and magnetic fields perfectly self-perpetuate in a manner consistent with energy conservation. This speed, known as the **speed of light**, c, is equal to 299,792,458 m/s.[1]

Before the 1600s, prominent scientists, including Johannes Kepler and Rene Descartes, believed light to be transmitted instantaneously. Galileo Galilei was among the first to question this viewpoint. To test whether the speed of light could be measured, Galileo proposed an experiment in which two participants would hold covered lanterns a great distance apart. One person would uncover his lantern, and then the second person would uncover his lantern the instant he saw the light from the first lantern. By knowing the distance separating the two lanterns, Galileo could then determine the speed of light. However, limited by both the timekeeping methods of the time and human reaction speeds, Galileo could only estimate that light was at least ten times faster than sound.[2]

In 1676, Danish astronomer Ole Römer was studying the moons of Jupiter when he made an interesting observation regarding their eclipses. Römer noticed that eclipses of Io tended to happen *later* than expected when Earth was farther from Jupiter, and *earlier* than expected when Earth was closer to Jupiter (Figure 4). He attributed the difference to the greater distance the light from Io had to travel to reach Earth. Using the then-accepted value for the diameter of Earth's orbit, Römer calculated the speed of light to be about 220,000,000 m/s, the first quantitative measurement of c.[3]

Around 1850, French physicist Hippolyte Fizeau sought to measure the speed of light via an experiment on Earth—akin to Galileo's lantern experiment centuries earlier. Fizeau constructed an apparatus consisting of a cogwheel and a mirror separated by a distance of eight kilometers. By shining a beam of light through the teeth of the turning cogwheel and observing its reflection in the distant mirror, Fizeau's experiment served as a rapid simulation of the covering and uncovering of Galileo's lanterns (Figure 5). Using this apparatus, Fizeau obtained a value for c equal to 313,000,000 m/s. In 1862, Fellow French physicist Leon Foucault improved upon Fizeau's result by replacing the cogwheel with a rotating mirror, yielding a more accurate

FIGURE 4

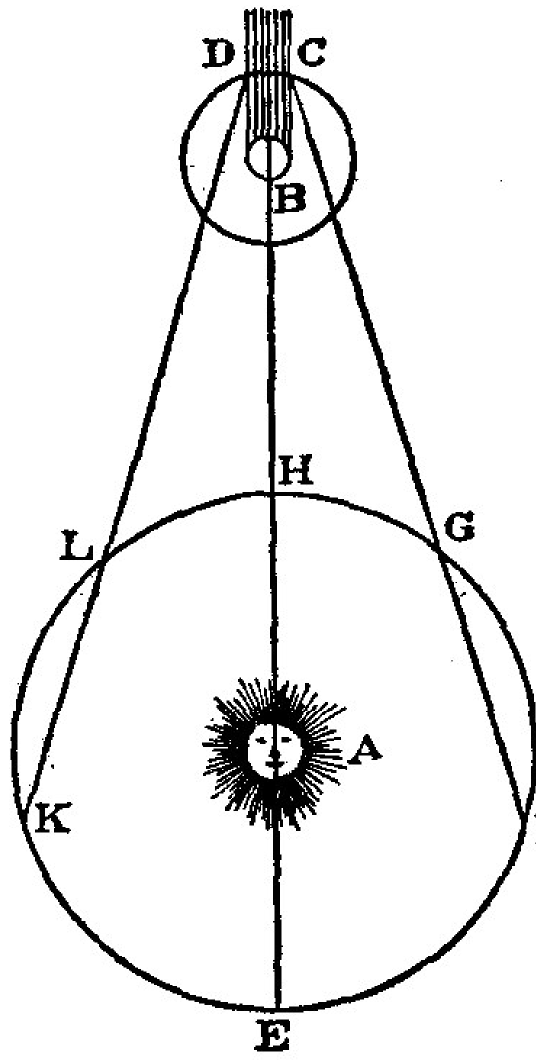


Fig. 70.

Illustration from Ole Römer's original 1676 publication regarding the speed of light. The largest circle is Earth's orbit around the Sun (A). Jupiter is located at (B), eclipsing its moon Io at points (DC).

value of 298,000,000 m/s.[4]

It's important to note that even by the time of Foucault's experiment, light was not yet understood to be an electromagnetic wave. In an 1865 paper, James Maxwell used his system of equations to determine a theoretical value for the speed of electromagnetic waves. He obtained a value of 300,000,000 m/s, in astonishingly close agreement with the speed of light that had been measured only years prior. Maxwell concluded that light was, in fact, an electromagnetic wave: "The agreement of the results seems to show that light and magnetism are affections of the same substance, and that light is an electromagnetic disturbance propagated through the field according to electromagnetic laws."[5]

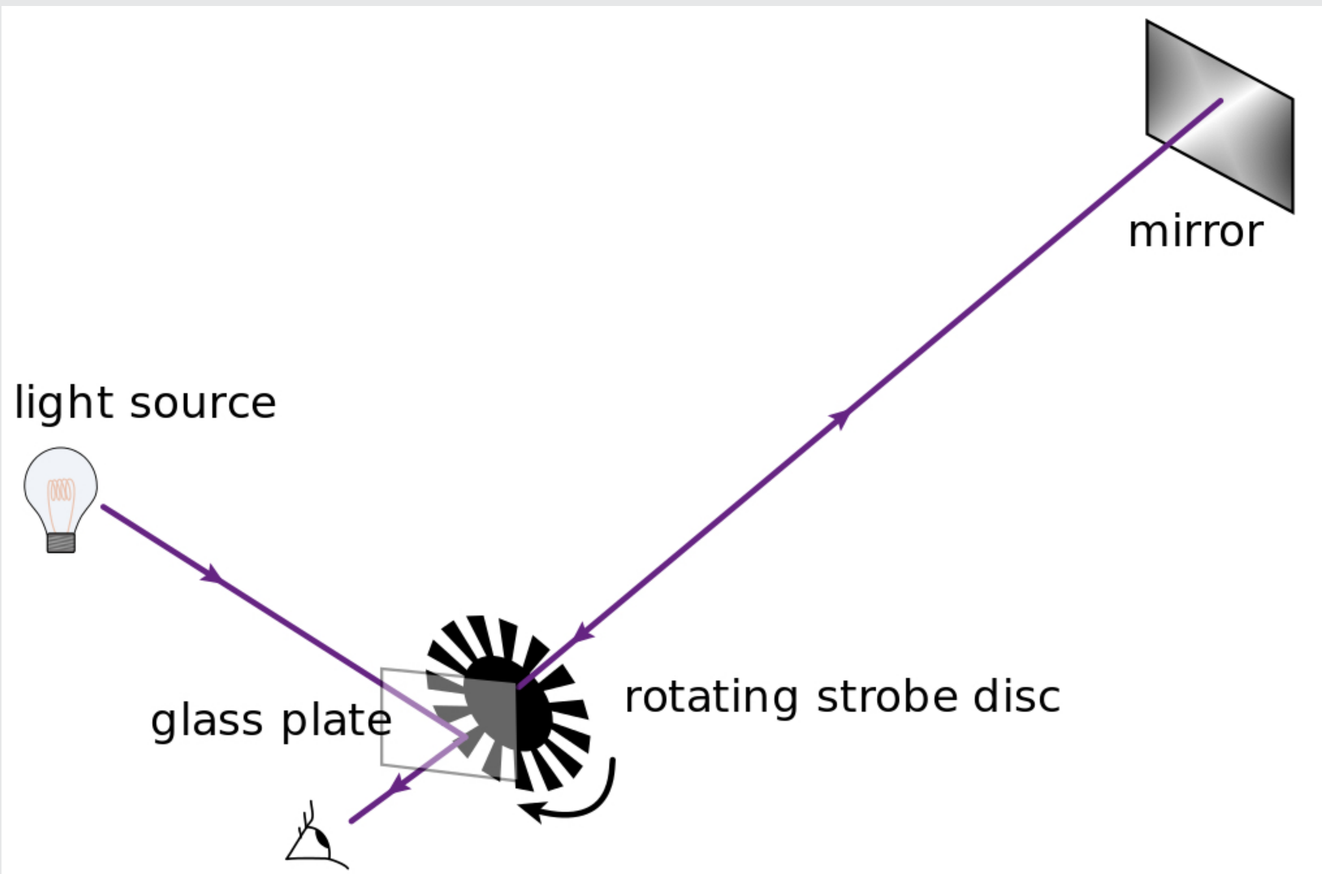
Subsequent measurements of the speed of light were conducted by Albert Michelson over a half-century period beginning in 1879. Michelson's techniques improved upon Foucault's method, and his final measurements in 1931 yielded an even more accurate value of 299,774,000 m/s.[6] Today the speed of light is one of the most well-established values in all of physics. We know its value to such precision that we now define the length of the meter in terms of it.

THE ELECTROMAGNETIC SPECTRUM

We are constantly surrounded by an undulating sea of electromagnetic waves, with each wave characterized by its own amplitude and frequency. In addition to visible light, the space around us is filled with a multitude of waves with different wavelengths, from radio waves to gamma rays. This continuum of electromagnetic waves, arranged by either frequency or wavelength, is called the **electromagnetic spectrum** (Figure 6).

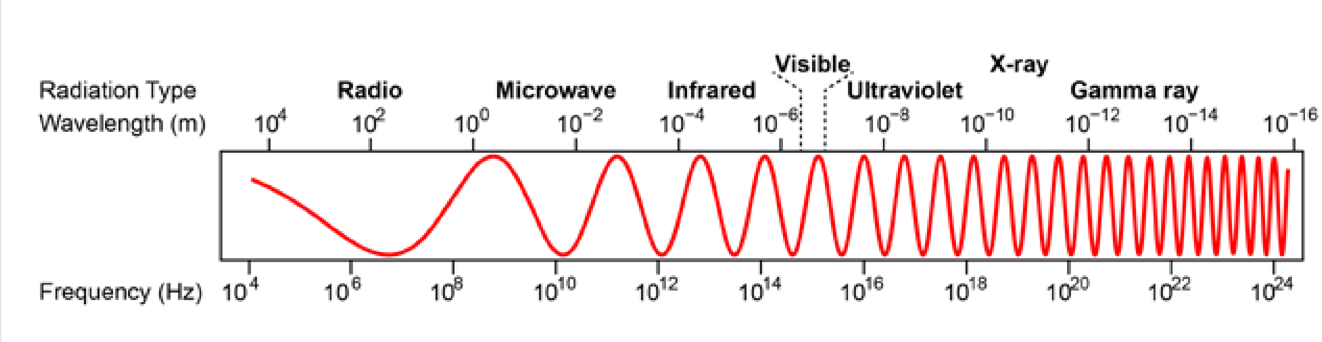
It's worth noting that the labels for each category serve as a historic classification system and do not indicate fundamental differences between each type of wave. Each

FIGURE 5



Fizeau's speed of light measurement apparatus.

FIGURE 6



The electromagnetic spectrum.

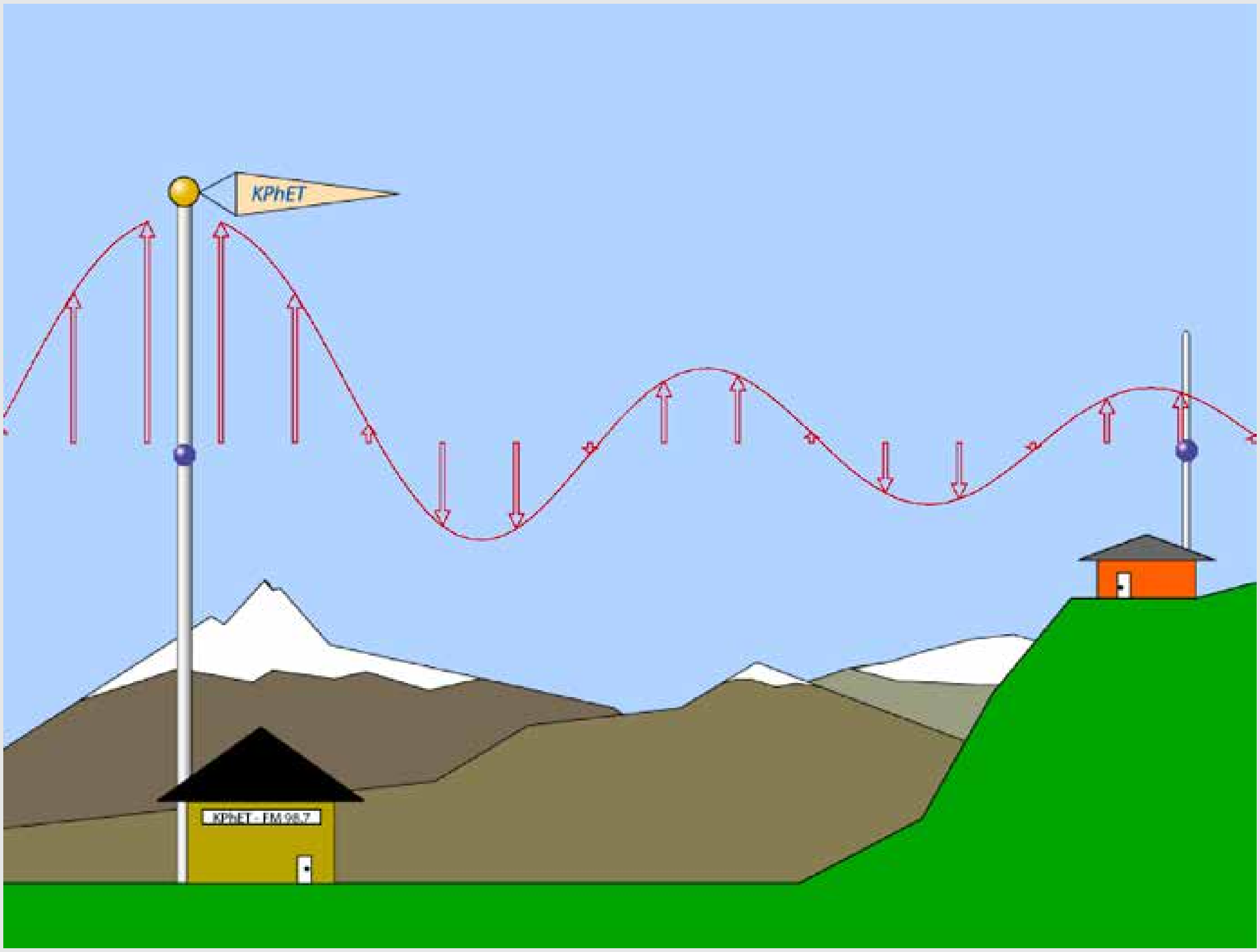
kind of wave in the spectrum is a manifestation of the same electromagnetic wave phenomenon, differing from one another in their frequency of oscillation, which is directly related to the amount of energy they are capable of depositing in matter. There is no sharp boundary between one category of electromagnetic wave and the next.

- Radio waves have the longest wavelengths in the electromagnetic spectrum, typically 30 cm or longer. Some AM radio waves even have wavelengths as long as 1 km. Radio waves are primarily used in radio and television communication, but also in radar systems, satellites, and walkie-talkies.
- Microwaves have wavelengths between 1 mm and 30 cm. Although you are likely most familiar with their application to cooking food in microwave ovens, microwaves are also used for aircraft navigation, radar, and scientific research.
- Infrared (IR) waves have wavelengths longer than visible light, between about 700 nm and 1 mm. IR waves are readily absorbed and emitted by most materials. When absorbed, IR waves cause objects to get hotter by increasing the kinetic energy of atoms and molecules. Infrared has a wide range of applications in meteorology, night-vision devices, communication, and astronomy, to name a few.
- Visible light is the set of wavelengths detectable by the human eye. The colors of the visible spectrum range from violet (≈400 nm) to red (≈700 nm). Human eyes are not equally sensitive to all wavelengths; the peak of sensitivity is around 560 nm (yellow-green light).
- Ultraviolet (UV) light ranges from about 400 nm to 10 nm. The Sun is a significant source of UV light received here on Earth although most is absorbed by the atmosphere. UV light in large doses is harmful to humans, but it also has practical uses in the sterilization of medical instruments.
- X-rays have wavelengths between about 10^{-4} nm and 10 nm. The most familiar application of X-rays is in medicine for the examination of bones, teeth, and organs. X-rays are energetic enough to be harmful to living tissue—you may recall wearing thick shielding to protect your body when having an X-ray performed at the doctor or dentist's office.
- Gamma rays are short-wavelength, highly penetrating waves that are emitted by radioactive nuclei. They have wavelengths between 10^{-14} m and 10^{-10} m. Gamma rays are extremely hazardous to living tissue and can cause irreversible damage if absorbed. However, the destructive capability of gamma rays can also be harnessed to destroy harmful cells such as cancerous tumors.

ANTENNAS AND RADIO WAVES

Antennas can be used to transmit and receive electromagnetic waves. When electrons in an antenna are moved back and forth at a given frequency, an electromagnetic wave is produced with a matching frequency. For example, if the frequency of your favorite radio station is 94.9 MHz, that means the electromagnetic waves produced by the station's antenna (and received by your radio's antenna) are oscillating at a frequency of 94,900,000 cycles per second. By tuning your radio to the same frequency, you can receive a clear signal from that

FIGURE 7



A radio wave produced by a station antenna is received by a home antenna. You can generate your own radio wave pulses using the "Radio Waves & Electromagnetic Fields" PhET applet at phet.colorado.edu.

Photo Credit: PhET applet at phet.colorado.edu.

electromagnetic wave (FIGURE 7).

In 1887, German physicist Heinrich Hertz generated the first radio waves in a laboratory. In a series of subsequent experiments, Hertz demonstrated that even though the waves could not be seen, they exhibited many similar properties to light waves. For instance, Hertz showed that these waves traveled at the speed of light, and he was able to focus and redirect them using reflective surfaces. (Hertz made numerous contributions to experimental physics. You may have noticed that the unit of frequency, the hertz, is named in his honor.)

At the time of his death in 1894, Hertz believed his radio waves served no practical purpose beyond verifying Maxwell's equations. Within a few years, however, the full potential of radio waves was made clear. In the mid-1890s, Italian engineer Guglielmo Marconi revolutionized the world by harnessing radio waves for the purposes of wireless communication. Marconi made the first public demonstration of a radio transmitter and receiver in July 1896.

TRANSPARENT vs. OPAQUE MATERIALS

How is it possible for light waves to pass through a glass window but not a piece of plywood? To answer this question, we must zoom in to examine the atomic building blocks that make up the world around us. Everyday materials are composed of molecules, which are in turn made up of atoms containing electrons. Although electrons are generally bound to a particular atom, the oscillating electric fields of an incident electromagnetic wave can cause an atom's outer electrons to vibrate back and forth.

Many objects have a natural frequency at which they tend to oscillate—think about the specific tone a tuning fork makes when struck, or the rate at which a playground swing moves back and forth when displaced. When these objects encounter a periodic disturbance that matches their natural frequency, their amplitude of oscillation increases strongly. This phenomenon is called **resonance**. You may have encountered resonance when pushing someone in a swing with just the right timing to make the swing go higher, or—more impressively—when observing how a crystal wineglass can shatter when exposed to a musical tone that matches its natural frequency. In the same way, electrons within the molecules of everyday matter have natural frequencies of oscillation; the exact frequency is dependent upon the molecular composition of the material.

When light of a given frequency is incident upon a material with electrons having the same resonant frequency, the light wave is absorbed. The energy of the incident light wave is converted into thermal energy of the atom, which collides and interacts with neighboring atoms, thereby raising the temperature of the material. Materials that absorb particular frequencies of light are said to be **opaque** to those frequencies. Ordinary window glass is opaque to UV light because electrons within glass atoms tend to have natural frequencies in the ultraviolet range. Accordingly, when UV light is incident upon glass, it warms the glass instead of being transmitted.

What happens if incident light does not match the resonant frequency of a material? The electrons are still set into vibration, but at lower amplitudes for shorter periods of time. The vibrating electrons then generate their own electromagnetic waves that propagate outward, and there is less energy converted into heat. If the material is **transparent**, the reemitted waves will be continuously absorbed and emitted by neighboring atoms until the wave exits from the opposite side of the glass. Throughout this process, the frequency of the light does not change as it is "handed off" between atoms. There is, however, a time delay on the order of nanoseconds between when each atom absorbs and reemits a wave. Consequently, light passes more slowly through transparent materials than it does in a vacuum.

In contrast, if the material is opaque, the wave will not be passed through the inside of the material. Instead, the electrons closest to the surface of the object will vibrate briefly in response to the light and then re-emit a reflected light wave back outward. Most objects around us are opaque, and we only see the outside of them due to the light that is re-emitted back from the electrons on the surface (or, in other words, reflected).

Metals tend to be opaque to most electromagnetic waves. The outermost electrons in metals are not bound to particular atoms and are free to migrate throughout the material, which explains why metals tend to be particularly good conductors of both electricity and heat. When light is incident upon the surface of a metal, these free electrons respond more to the electric fields in the wave than they would if they were bound to a particular atom. As a result, the electrons redistribute themselves in such a way that they set up an electric field that exactly cancels the field of the incident wave. (The ability of electrons to redistribute in this way is a defining characteristic of conductors.) This redistribution produces an outgoing wave that matches the frequency of the incoming wave, which is the reflected light that we see from a shiny metal surface. If there is resistance to the motion of the electrons, the result will be less reflected light and a duller appearance to the metal.

VISIBLE LIGHT AND COLOR

THE VISIBLE SPECTRUM

The colors of light we see depend on the frequencies of light that reach our eyes. The lowest-frequency light in the visible spectrum appears to most people as *red* and the highest-frequency light as *violet*. Although there are an infinite number of frequencies within the visible spectrum, conventionally we often group them into the seven colors Red, Orange, Yellow, Green, Blue, Indigo, and Violet. You can use the memory device "ROYGBIV" (often pronounced like "Roy G. Biv") to remember the colors of the visible spectrum in order from lowest to highest frequency.

It's important to note that although we often describe light as "blue" or "red" based on its frequency, color is not an inherent property of any electromagnetic wave. Nor is the color of an object on or within the object itself. Color is a *physiological experience*, involving our eyes and brain, that differs from person to person. For instance, even if the same light waves reached their eyes, a person with color blindness might perceive those waves differently from a person with normal color vision.

Light that is made up of a mixture of all colors in the visible spectrum appears white. For example, white light from the Sun is a composite of all visible wavelengths of light. In the late 1660s, Sir Isaac Newton first demonstrated that white light is a composite of colored light by using a prism to separate sunlight into a spectrum of colors and then recombining them with a second prism (Figure 8). (Later we will discuss how a prism separates out individual colors from white light.)

The Sun emits a range of electromagnetic wave frequencies, the vast majority (over 99 percent) belonging to ultraviolet, visible, and infrared portions of the spectrum. The peak of solar intensity resides squarely within the visible spectrum for humans. That's certainly no coincidence—evolutionary processes over millions of years have caused human eyes to be most sensitive to the frequencies of light that are emitted most strongly by the Sun. Some animals and insects are sensitive to electromagnetic waves that are beyond the visible spectrum for humans. Bees, for example, are sensitive to a portion of the solar spectrum that extends into the ultraviolet, which enables them to detect UV patterns that direct them toward pollination sites on certain species of flowers.

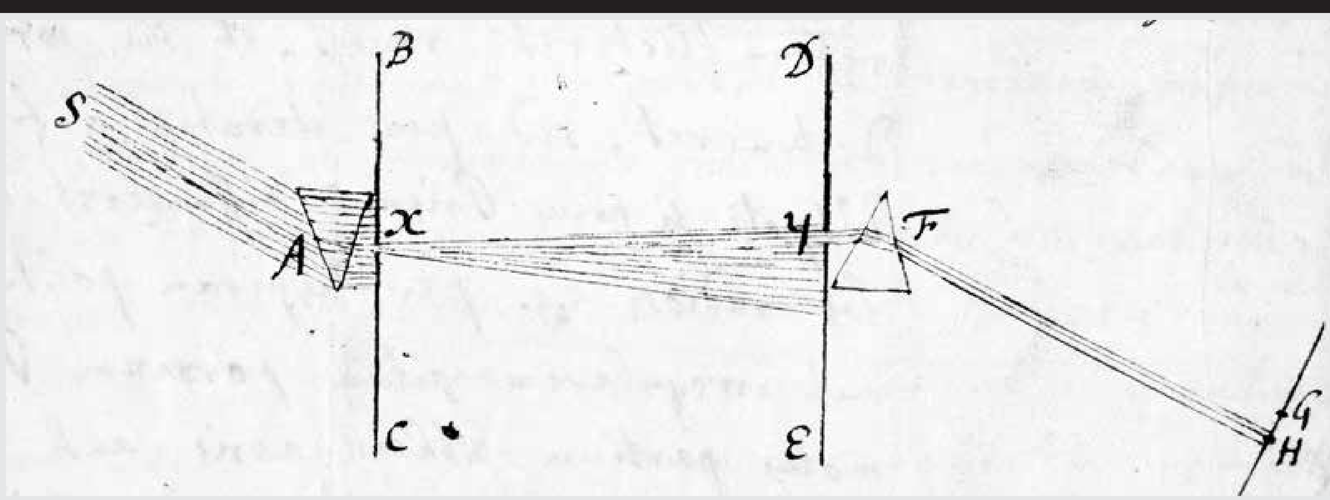
SELECTIVE ABSORPTION AND REFLECTION

What does it mean for an object to appear red or yellow or violet? Some objects we see, such as the Sun, the stars, campfires, lamps, and lasers, are *sources* of visible light. We see them because they emit electromagnetic waves in the visible spectrum as the result of electrons being rearranged within atoms and molecules. However, we also see objects that aren't sources of light, but only because they reflect light from other sources to our eyes.

Objects absorb light of certain frequencies and reflect the rest. The frequency of the reflected light determines the apparent color of the object. A red apple, for instance, appears red because it absorbs all frequencies of light incident upon it except red, which it reflects. Materials that reflect light of all visible frequencies appear to be the same color as the light that is incident upon them. Conversely, objects that absorb all incident light reflect no visible frequencies and appear black.

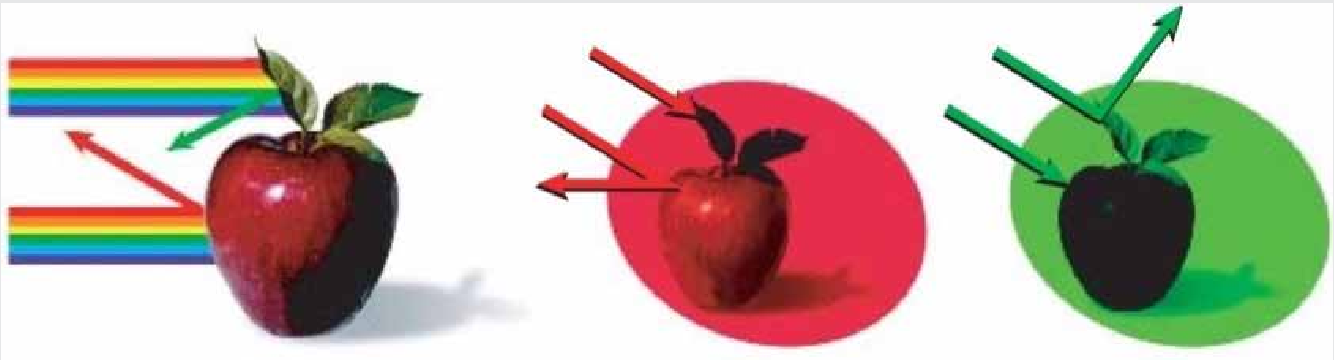
It's important to note that an object can only reflect frequencies that exist within the light that is incident upon it. In other words, the apparent color of an object is dependent upon what kind of light illuminates it. Consider an apple that appears red when illuminated by white light. Under red light, the apple will still appear red because red wavelengths are clearly still able to be reflected. However, under green light, the apple will appear black. Why? The green wavelengths of light are completely absorbed by the apple, and none are reflected to the observer (Figure 9). Incandescent bulbs are often described as producing a "warmer" light than fluorescent bulbs because the light waves they emit tend to be mostly from the low-frequency (i.e., red) side of the visual spectrum. As a result, parts

FIGURE 8



An illustration of Newton's double prism experiment, taken from a 1671 letter to the Royal Society.

FIGURE 9



A red apple reflects red light and absorbs the rest. Under red light, the leaves appear black. Under green light, the apple appears black.

of objects that reflect red light tend to be more visually prominent when illuminated by an incandescent bulb.

SELECTIVE TRANSMISSION

Ordinary window glass appears colorless because it transmits all frequencies of visible light equally well. It's also possible for materials to absorb certain frequencies of visible light but allow others to be transmitted. Rubies, for example, are a variety of the mineral corundum, which consists of aluminum oxide in a crystalline structure. In its purest form, corundum is colorless. The red color of rubies is due to chromium impurities scattered throughout the crystal structure, which absorb frequencies of light within the visible spectrum except red. Interestingly, sapphires also belong to the corundum family, but instead contain titanium and iron impurities that lend them their stunning blue color.

It's also possible to artificially produce materials that are transparent only to certain frequencies. For example, colored glass is produced by adding metals and metal oxides during the manufacturing process. Cobalt blue glass is known for its deep blue coloring, which is due to the cobalt oxide it contains (Figure 10).

BLUE SKIES AND RED SUNSETS

Using the ideas we've covered so far, you have enough information to answer an age-old question: *Why is the sky blue?* As we discussed earlier, the colors we see are caused by objects selectively reflecting certain frequencies of light to our eyes and absorbing the rest. But how can air reflect light at all? Strictly speaking, the light is not exactly reflected, but instead is *scattered* off the molecules that make up the atmosphere. Scattering in this context refers to light being absorbed and re-emitted by the atmospheric molecules.

Earth's atmosphere is made up of molecular gases—primarily nitrogen (N_2) and oxygen (O_2), but also argon (Ar), carbon dioxide (CO_2), and other gases in trace amounts. The N_2 and O_2 molecules have resonant frequencies at the high-frequency end of the visible spectrum. As a result, these molecules tend to scatter blue light from the Sun more strongly than they scatter red light. The scattered blue light is sent in all directions, some of

FIGURE 10

Cobalt blue glassware.

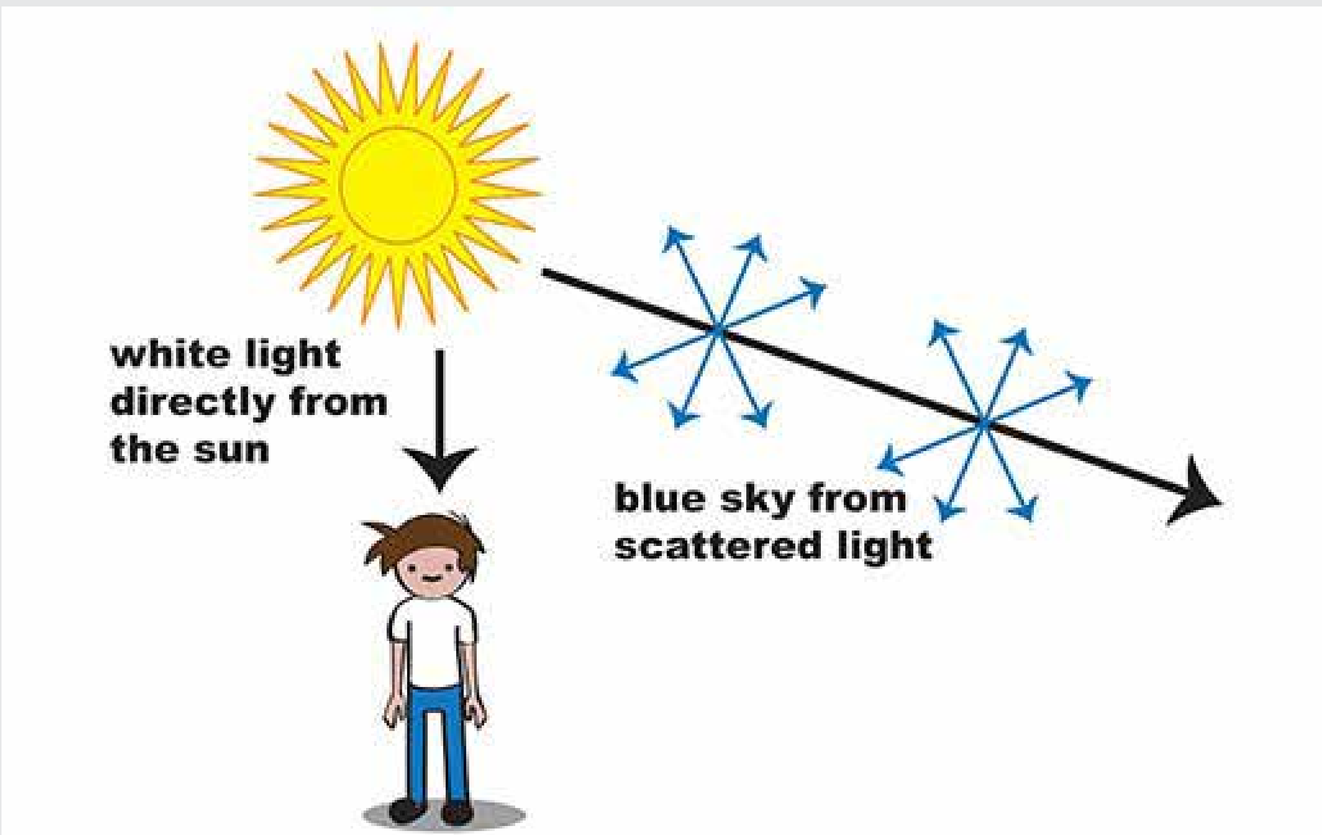
which reaches our eyes (Figure 11). The result is a vibrant blue sky. This effect is known as **Rayleigh scattering** after its discoverer, British physicist Lord Rayleigh. Violet light is actually scattered more strongly by atmospheric particles than blue light, but our eyes are more sensitive to blue than violet.

Of course, the sky does not always appear blue. In your lifetime, you've likely witnessed many breathtaking sunsets, made up of deep reds and oranges, as the Sun dips close to the horizon. At sunset, light from the Sun has to travel further through the atmosphere in order to reach the observer. By the time the light from the setting Sun reaches your eyes, the higher frequencies of light have scattered away, whereas the lower frequencies (corresponding to the red end of the visible spectrum) have not (Figure 12). What's left are the brilliant, warm colors of the sunset.

Try It! Sunset in a Glass

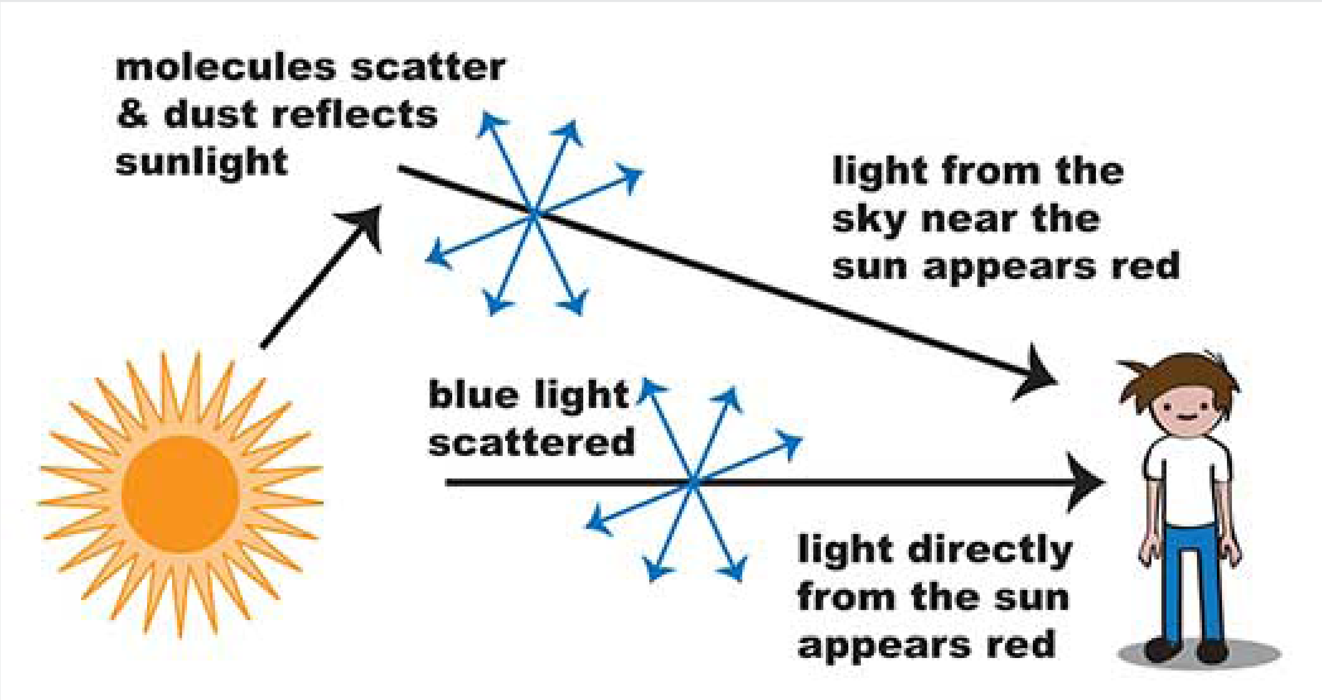
Obtain a clear drinking glass, water, milk, and a bright flashlight (LED light works best). Fill the glass with water and then stir in about ½ tablespoon of milk. Point the lit flashlight directly at the glass from the side, and move your head around the glass to view the scattered light from all angles. The light should appear *blue* when viewed from the sides of the glass at right angles to the flashlight, and *red* when viewed straight on from the side opposite the flashlight. In this experiment, the light from the flashlight is scattering off the milk particles suspended in the water in a manner similar to how light scatters off particles in the atmosphere.

FIGURE 11



The sky appears blue because blue light is scattered the most toward our eyes due to Rayleigh scattering.

FIGURE 12



The red colors of sunsets are what remain after the blue frequencies have scattered away.

SECTION I SUMMARY

- A wave is the motion of a disturbance between two points. All waves transport energy but not mass. Light is an electromagnetic wave, made up of oscillating electric and magnetic fields. Electromagnetic waves are produced when charged particles accelerate.
- All electromagnetic waves travel at the same speed, the speed of light (c), in empty space. The speed of light is equal to 299,792,458 m/s.
- The electromagnetic spectrum is a classification of all electromagnetic waves by frequency. Each type of wave in the spectrum represents the same electromagnetic phenomenon, differing from one another only in their frequency of oscillation.
- An antenna is a means of producing electromagnetic waves of a particular frequency by moving electrons back and forth. Heinrich Hertz produced the first radio waves in 1887, and Guglielmo Marconi demonstrated their application to communication in 1896.
- Transparent objects allow incoming light waves to be transmitted through them. Opaque objects absorb incoming light and reflect it back at their surface.
- The visible spectrum represents a small fraction of the entire electromagnetic spectrum. The colors we see depend on the frequencies of light that reach our eyes. White light is a composite of all visible wavelengths of light and can be separated into a full spectrum using a prism.
- Objects absorb certain frequencies of light and reflect the rest. The frequency of the reflected light determines the apparent color of the object. Objects that absorb all incident frequencies reflect no light and appear black.
- The sky appears blue due to Rayleigh scattering of sunlight from atmospheric molecules. Higher frequencies within the visible spectrum (e.g., blue) are scattered more strongly toward observers.

Section II

Optics

INTRODUCTION

Light can be made more useful when it is shaped, redirected, and focused by mirrors and lenses. As we will see, lasers use mirrors and lenses for their basic operation as well as to make them more useful for practical purposes. **Optics** is a branch of physics involving the properties of light and how light interacts with matter through reflection, refraction, and diffraction. In addition to lasers, optics plays an important role in the function of cameras, magnifiers, telescopes, microscopes, and many other devices we use on a daily basis.

GEOMETRIC OPTICS

THE RAY APPROXIMATION

Centuries of observations and experiments have demonstrated that light travels in a straight line until it encounters a boundary with another medium. This statement generally holds true provided that the light is interacting with objects much larger than its wavelength. At these length scales, we can represent beams of light as rays, or straight lines with a direction. This simplification is called the **ray approximation**. The part of optics in which the ray approximation is valid is known as **geometric optics**, or ray optics. Throughout most of this section, we will apply the ray approximation when analyzing the path that light waves follow through various materials. Later in this section, we will consider the behavior of waves when the ray approximation is no longer valid.

REFLECTION

As we described in Section I, most objects we see do not emit light of their own, but rather they re-emit light that is incident upon their surface from a source of light. When light reaches the surface of a material it is either absorbed, transmitted through the material, or reflected. **Reflection** occurs when light is not absorbed or fully transmitted, and some portion of the wave changes direction at the surface of the material without actually entering the material itself. All substances absorb at least some incoming light and reflect the rest.

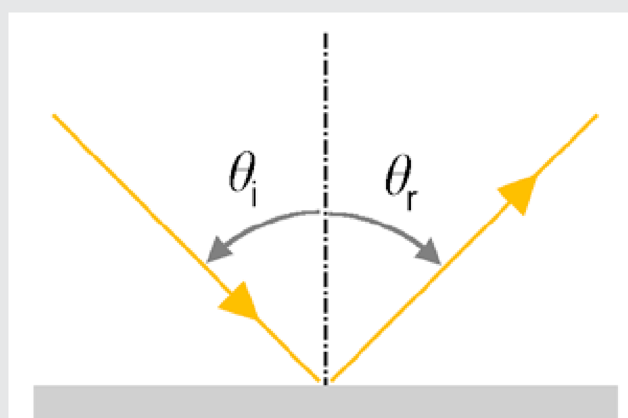
The Law of Reflection

Consider a light ray that is incident upon a smooth, reflective surface (Figure 13). The incoming ray makes an angle θ_i, known as the *angle of incidence*, with an imaginary line perpendicular to the reflecting surface. This line is called the **normal**. (In physics, we often use the word *normal* as another word for *perpendicular*.) The outgoing ray makes an angle θ_r, called the *angle of reflection*, with the normal. According to the **law of reflection**, the angle of incidence is equal to the angle of reflection. Symbolically, this can be written as follows:

$$\theta_i = \theta_r ,$$

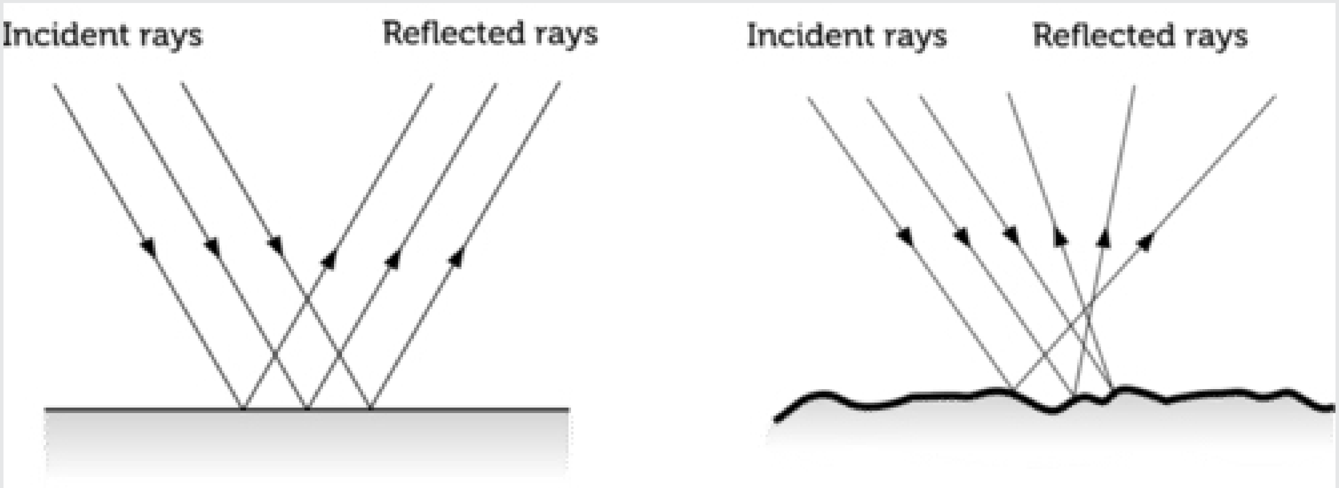
where θ_i is the angle of incidence and θ_r is the angle of reflection. To put this in words, the angle that a light ray makes as it reflects from a surface is always equal to the angle the incoming light ray makes with the surface.

FIGURE 13



The law of reflection states that the angle of incidence (θ_i) equals the angle of reflection (θ_r).

FIGURE 14



Specular reflection (left) and diffuse reflection (right).

Specular and Diffuse Reflection

Reflection of light from a smooth surface is called **specular reflection**. During specular reflection, incident light rays remain parallel to one another, resulting in a mirror-like reflection of incoming light. A mirror is, of course, an excellent example of a specular reflector, but specular reflection can also be observed from window glass or still water.

Surfaces do not need to be smooth in order to cause reflection. Reflection from a rough surface is called **diffuse reflection**. As you can see in Figure 14, the law of reflection still applies to each incoming light ray; however, the irregularity of the surface causes each individual ray to be reflected in a slightly different direction. The reflection of light from most walls is an example of diffuse reflection; walls tend to reflect light back into a room, but they are (generally) not shiny and do not form images as mirrors do. Most objects around us are visible as a result of diffuse reflection.

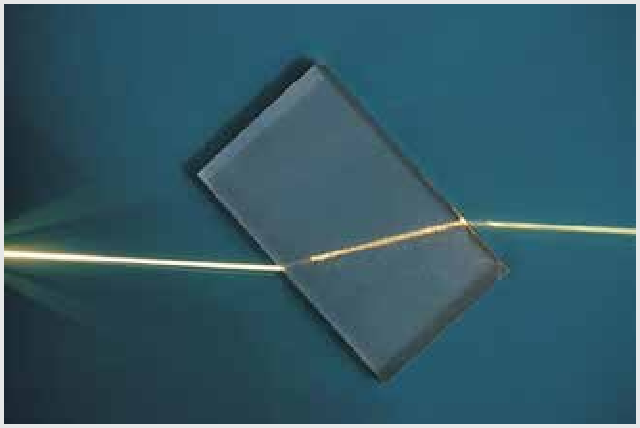
The "smoothness" of a surface is relative to the wavelength of the wave that is being reflected. If the height of the irregularities on a surface are small compared to the wavelength of the incident light, the surface will appear smooth to the wave and it will undergo specular reflection. Therefore, it's possible for the same surface to appear rough to one wave and to appear smooth to another wave with a longer wavelength. For example, the dish surfaces of large radio telescopes, some more than fifty meters in diameter, cause visible light to be diffusely reflected, but they also act as specular reflectors for long-wavelength radio waves.

REFRACTION

When a light ray encounters a boundary between two different transparent materials, the transmitted ray changes direction in a process called **refraction** (Figure 15). Refraction occurs because light travels at different speeds on each side of the boundary. The refracted path taken by light entering a different medium can be described using **Fermat's principle of least time**. In 1662, French mathematician Pierre de Fermat proposed that the path that light takes between two points is the path that minimizes the travel time. For two points within the same material, this is obviously the straight-line path connecting them. But what about two points in two different materials?

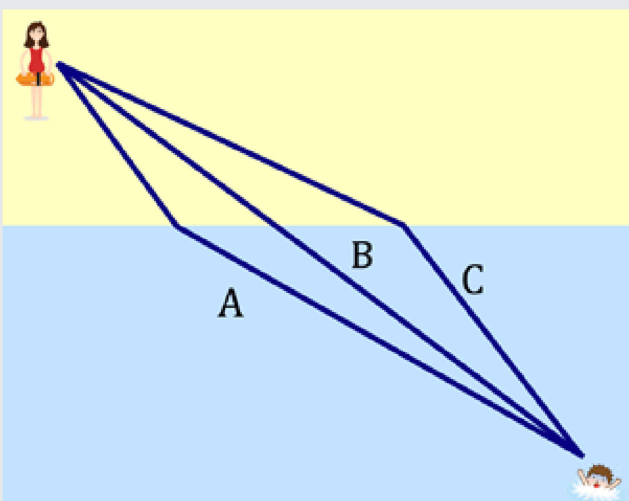
We can use an analogy to think about how Fermat's principle applies to refraction. Imagine you are serving as a lifeguard on a beach when you suddenly notice a swimmer in distress (Figure 16). Your running speed on the beach is faster than your swimming speed in the water. Which path will allow you to reach the swimmer in the least amount of time? Path A requires a greater distance traveled through water than on the beach, so it can't be the fastest route. Although path B is a straight line, you could reduce the overall travel time by covering more distance over land and less through water. Without going into a formal derivation, path C is the choice that minimizes the time to reach the

FIGURE 15



A light ray refracting as it enters a plastic block, and again as it leaves.

FIGURE 16



Three paths a lifeguard could take to reach a swimmer in distress. The least-time path, C, resembles the path of a light ray entering a medium in which it moves more slowly.

swimmer. Likewise, a light ray entering a material in which it moves more slowly will follow a bent path similar to path C. Just as the optimal path for a lifeguard would depend on his or her swimming and running speeds, the exact angle the ray follows will depend on the speed at which light travels in each material.

TABLE 1

Medium	Index of Refraction
Vacuum	1 (exact)
Air	1.0003
Water	1.33
Crown Glass	1.52
Sapphire	1.77
Diamond	2.42

Indices of refraction for various media.

The Law of Refraction

We stated previously that in a vacuum devoid of any matter, light always travels at the speed of light, c. When light passes through a transparent medium such as air, water, or glass, it slows down by a factor characteristic to that medium called the index of refraction. The **index of refraction**, n, of a material is defined as the speed of light in a vacuum, c, divided by the speed of light in the material, v:

$$n = \frac{\text{speed of light in a vacuum}}{\text{speed of light in the material}} = \frac{c}{v}$$

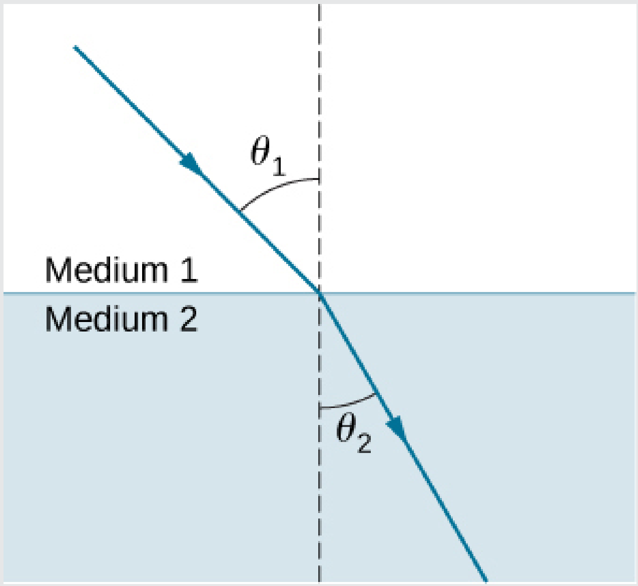
Rearranging, we can see that the speed of light within a material with index of refraction n is given by $v = c/n$. A few points are worth noting:

- n is a dimensionless number (i.e., it has no units).
- n is always greater than or equal to 1.
- n is defined to be 1 for a vacuum.

Table 1 shows the index of refraction for some common materials. In general, light travels more slowly in materials with higher indices of refraction. Correspondingly, a light ray will bend more when passing from a vacuum into a material with a higher index of refraction than it would when passing into a material with a lower index of refraction.

When a light wave passes between two materials, its frequency does not change. Without going into a great deal of detail, this is because a light wave must be "continuous" as it passes between regions of differing

FIGURE 17



The law of refraction.

refractive indices. That is, in a given period of time, the number of wave crests passing into one side of an interface must equal the number of wave crests leaving the other side. Recall from Section I that for a wave, the speed, frequency, and wavelength are related by $v = f\lambda$. Accordingly, when the speed of a light wave is reduced within a transparent medium, its wavelength must also be reduced by the same factor. If a wave has a wavelength λ in a vacuum, its wavelength in a medium with index of refraction n will be λ/n.

The equation relating the angle of incidence and the angle of refraction is known as **Snell's law** after seventeenth-century Dutch astronomer and mathematician W. Snell. According to Snell's law, when a light ray travels from a region with index of refraction n_1 to a region with an index of refraction n_2, the angle of incidence Θ_1 is related to the refracted angle Θ_2 by the following equation:

$$n_1 \sin \Theta_1 = n_2 \sin \Theta_2.$$

Although we won't provide a full mathematical treatment of Snell's law, we will describe some general rules of refraction. When a light ray enters a medium with a higher index of refraction, it bends *toward* the normal. Furthermore, the greater the refractive index, the more the light ray will bend toward the normal. When a light ray enters a medium with a lower index of refraction, it bends *away from* the normal.

Mirages

Have you ever been traveling in a car on a hot day and noticed what appear to be pools of water covering the road in the distance? As you continue to drive, you never reach the water despite the very convincing image you saw only moments before. This is a familiar example of a **mirage**, which is a naturally occurring optical illusion due to the refraction of light through the atmosphere. The air just above the surface of the roadway is very hot compared to air higher up, and light travels faster through the hot air because it is less dense. As a result, light that is incident upon the road ahead of you is refracted back upward toward your eyes (Figure 18). The fact that the incoming light is behaving in a manner consistent with reflection—although the effect is entirely due to refraction—causes your brain to interpret the region above the road as a reflective surface such as water.

We often associate mirages with the desert, as the hot temperatures cause varying air densities that refract light through various angles. This causes a wavy, shimmering effect that can also be seen above a hot grill or stove. Although it is common to think of mirages as being hallucinations or "tricks of the mind," they are the result of true refractive behavior of light rays and can therefore be photographed.

A similar refractive effect causes the Sun to be visible at the horizon even after it has "set." Earth's atmosphere is

FIGURE 18

A mirage of distant pools of water on a hot roadway, caused by refraction.

less dense at higher altitudes; accordingly, the least-time pathway for sunlight is a curved arc that favors higher altitudes where it travels more quickly. The curved path is due to the fact that the index of refraction changes gradually, layer by layer, in the atmosphere. To an observer on Earth's surface, the image of the Sun persists at the horizon even after the Sun itself should technically be out of view (FIGURE 19). The image of the Sun may also appear "squished" vertically just as it disappears from sight due to the same effect of refraction.

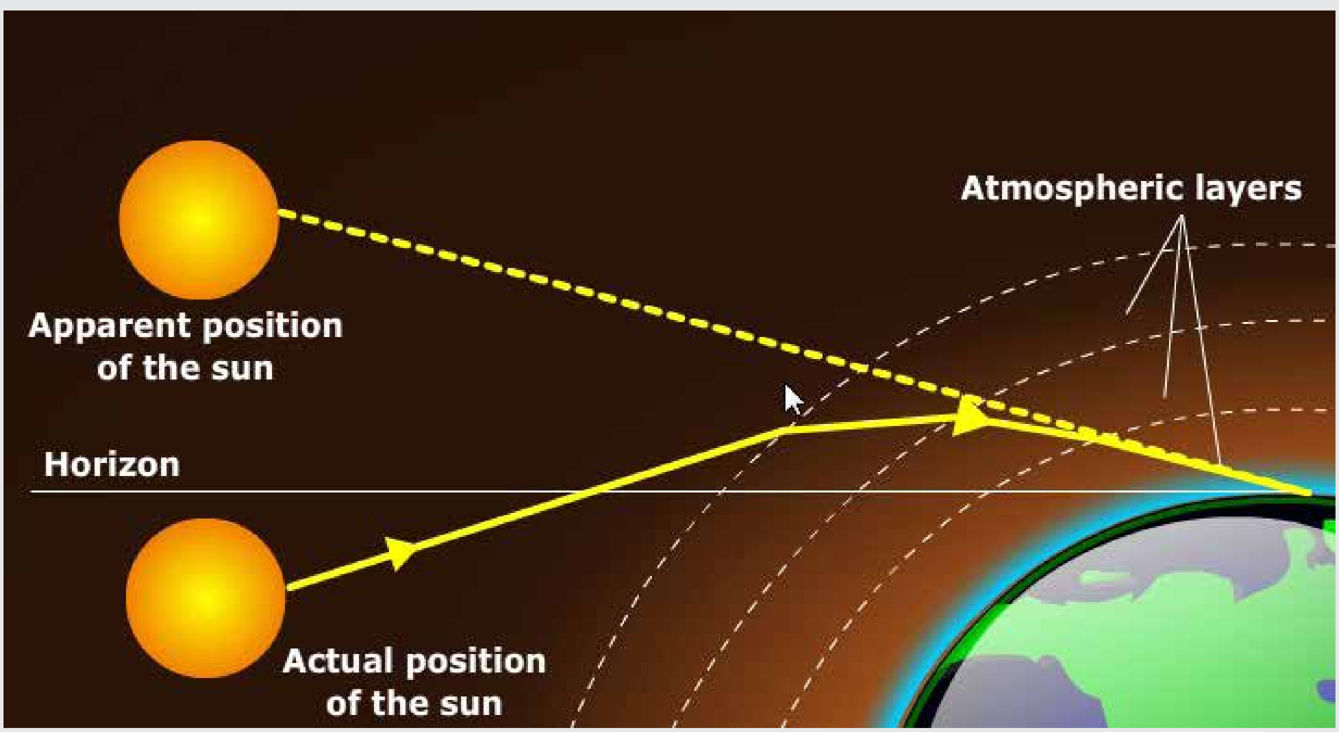
DISPERSION

Section I described how light behaves within transparent materials by introducing the principle of resonance. Light is mediated by electron oscillators within materials. Electron oscillators in most transparent materials, such as glass, resonate most strongly with frequencies of light in the ultraviolet range. Therefore, light waves with frequencies closer to the violet end of the visible spectrum are more likely to interact with oscillators in these materials and travel more slowly on average as a result. Indeed, violet light travels about 1 percent more slowly in ordinary glass than red light. The index of refraction is slightly greater for violet light than red light, and violet light will undergo a greater amount of refraction as a result. **Dispersion** is the separation of visible light into a spectrum of colors due to this dependence of refractive index on wavelength. Dispersion is commonly illustrated using a **prism**, which is a piece of glass or plastic typically cut into a triangular block. When white light is passed through one of the angled faces of a prism, a spectrum of colors is produced through dispersion.

Rainbows

Rainbows are caused by millions of raindrops acting as tiny prisms dispersing sunlight. A light ray from the Sun enters the droplet and is refracted. At the far side of the droplet, the ray is reflected and returns to the front side, where it once again is refracted as it passes back into the surrounding air (FIGURE 20a). The slight difference in angle between each wavelength of light causes the vivid spectrum of colors that we see. Although each droplet produces a full spectrum of colors, an observer's eye is only positioned to see one of those colors. The droplets at the very top of the rainbow direct red light toward the observer; violet light dispersed from this droplet passes above the observer's eye. Likewise, the droplets at the bottom edge of the rainbow direct violet light toward the observer (FIGURE 20b).

FIGURE 19



The Sun appears to linger at the horizon due to refraction through Earth's atmosphere.

In order for a rainbow to be visible, water droplets (in a cloud or falling rain) must be present in one part of the sky, and the Sun must be shining in the opposite part of the sky. (This is the reason rainbows are often associated with the passing of storms.) On a clear day, you can create your own rainbow by viewing a cascade of droplets from a garden hose or lawn sprinkler. An observer positioned between the Sun and the water droplets, facing away from the Sun, will see a spectrum of colors in a circular arc. If it weren't for the ground getting in the way, you would actually see the rainbow trace out a full circle—in fact, under the right conditions, it's possible to view a circular rainbow from an airplane or very tall building.[7]

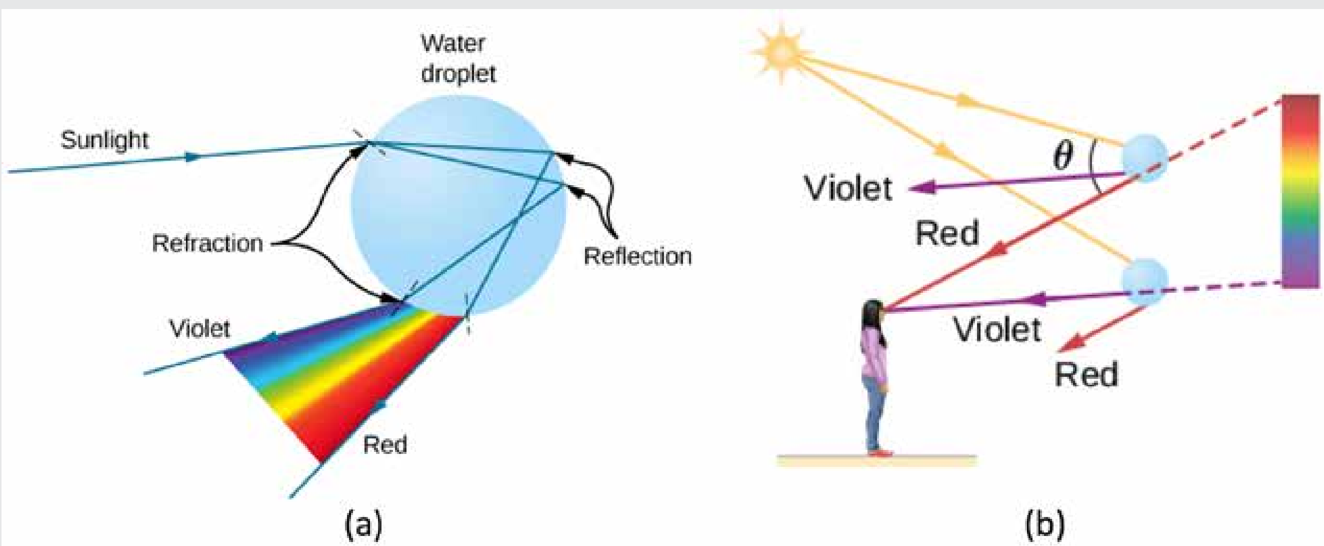
A secondary, or "double," rainbow is a rare sight requiring just the right conditions to view. The secondary rainbow appears above the primary and results when light undergoes an additional reflection before exiting the droplet. This second reflection causes the order of colors in the secondary rainbow to be reversed. The secondary rainbow is also much dimmer than the primary due to light intensity lost from the additional reflection. The next time you see a rainbow, take a moment to appreciate how refraction, reflection, and dispersion each contribute to this splendid visual phenomenon.

TOTAL INTERNAL REFLECTION

In general, when light moves between two different transparent materials, some light is refracted and some is reflected. **Total internal reflection** is an optical phenomenon in which a light ray traveling from a region of higher refractive index to a region of lower refractive index is completely reflected at the boundary. Consider what happens as the angle of incidence is gradually increased (FIGURE 21): the angle of refraction also gradually increases until, at the **critical angle**, the refracted ray will be oriented at 90° relative to the normal. For all angles of incidence greater than the critical angle, no refracted ray will be produced, and the ray will be entirely reflected back into the medium. The reflected light ray obeys the law of reflection.

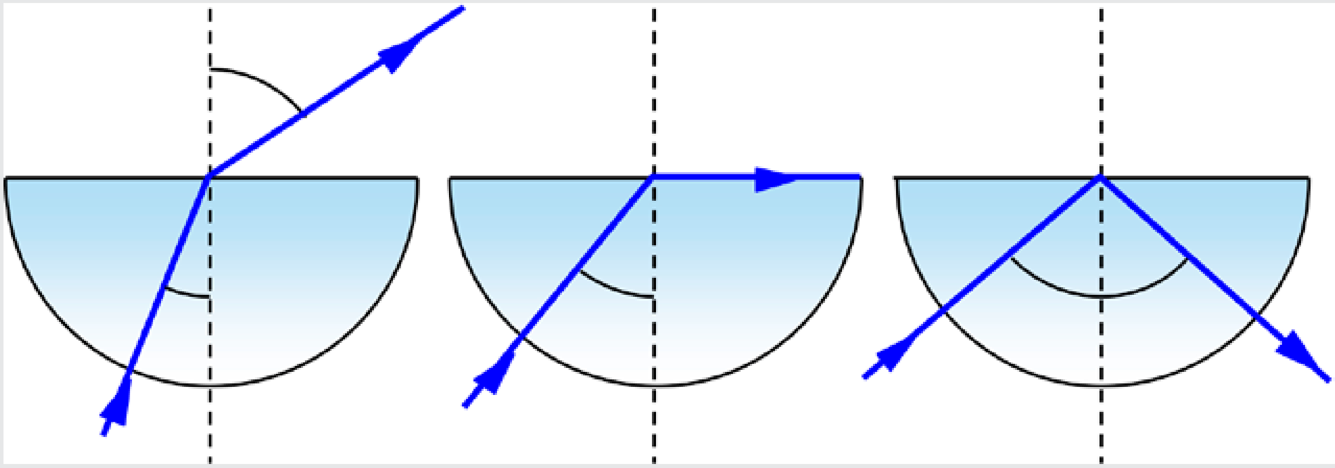
The effect of total internal reflection explains why diamonds sparkle more than other materials. Because diamond has a relatively high index of refraction (typically 2.42), the critical angle is small (about 24.4°), and therefore light rays typically undergo many internal reflections before striking at a steep enough angle to leave the diamond. Diamonds are typically shaped by jewelers in such a way that the light is concentrated at only a few places along the surface of the diamond that it can exit, thereby contributing to the sparkle. Dispersion of light also contributes to the brilliant array of colors that are often visible from the surfaces of well-shaped diamonds.

FIGURE 20



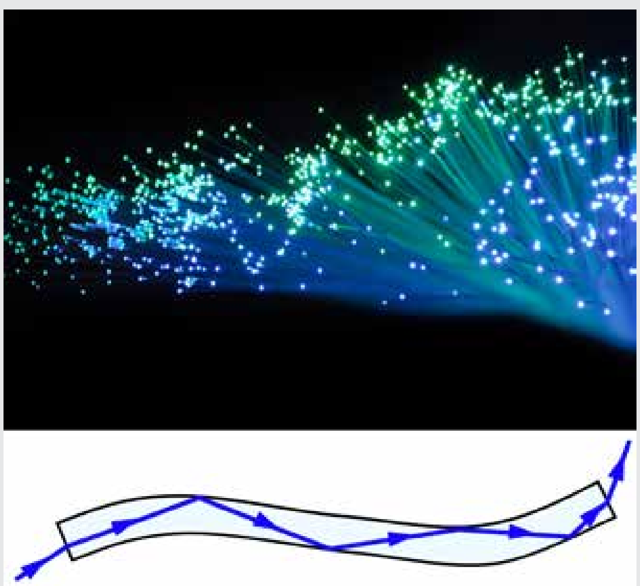
(a) Dispersion of light in a spherical water droplet, which causes a rainbow to be visible. (b) The position of colors in the rainbow is caused by colors emerging in different directions.

FIGURE 21



The critical angle is one for which the angle of refraction is 90°. Above this angle of incidence, no light is refracted, and all light is reflected back into the material.

FIGURE 22



Fiber-optic cables are an important application of total internal reflection.

Fiber-Optics

Another important application of total internal reflection is optical fiber technology (FIGURE 22). Optical fibers act as "pipes" through which light rays undergo a series of total internal reflections. In this way, the light is transmitted along the length of the fiber even as it bends and twists. Fiber-optic cables are flexible, meaning they can transmit light to—and collect images from—otherwise difficult-to-access areas such as a patient's stomach or the inside of a motor. Optical fibers can even be manufactured smaller than the diameter of a blood vessel.

Fiber-optic cables have several advantages over copper wires and cables in communication. Optical fibers are light, flexible, and can be manufactured at a lower cost compared to conventional communication wiring. Optical fibers also have a greater bandwidth than copper wires, meaning more information can be transmitted per unit time. Unlike electrical signals, light is also not affected by fluctuations in temperature or nearby magnetic fields, so interference and "noise" are less of a concern. Finally, optical fibers are more secure than conventional wiring because they are more difficult to tap into by possible eavesdroppers.

MIRRORS AND LENSES

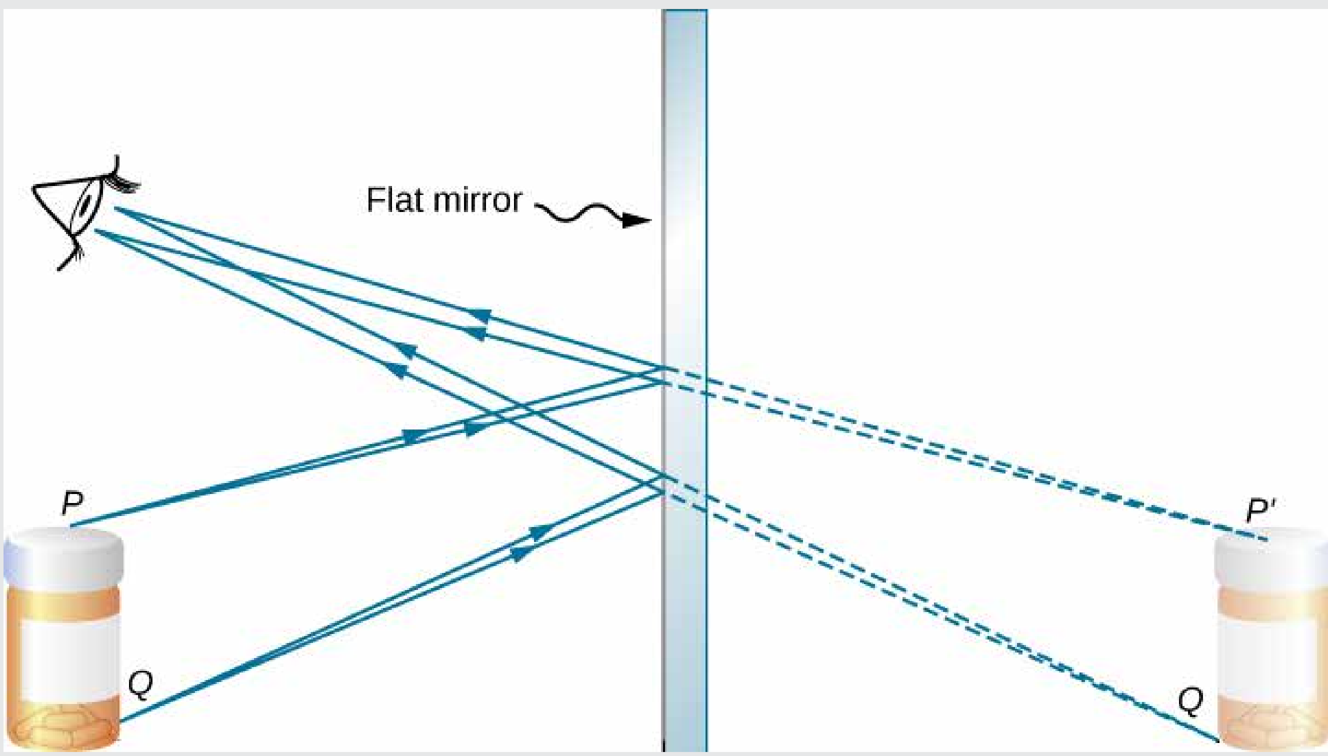
PLANE MIRRORS

A plane mirror, which has a flat surface, is the simplest type of mirror. Consider a point P on the lid of a pill bottle placed in front of a plane mirror (FIGURE 23). Light rays propagate outward in all directions from point P, and many of these rays reach the mirror. When the rays encounter the surface of the mirror, they are reflected back according to the law of reflection. If we look at the rays diverging from the mirror's surface, they *appear* to be emerging from a single point (P'), which is the point at which an observer sees an **image** of the pill bottle lid. We can apply the same reasoning to point Q at the base of the bottle. A little geometry (and our own experience) shows that the image formed by a plane mirror appears as far behind the mirror as the object is in front of the mirror. Furthermore, the size of the image is equal to that of the object. In other words, plane mirrors do not magnify, or enlarge, an image relative to the original object.

Images can be classified as *real* or *virtual*. A **real image** is formed by light rays passing through the location of the image. If a surface such as a screen is placed at the location of a real image, the image will be visible. For example, images projected onto a movie theater screen are real images, as are images focused by a camera onto film or a photosensitive surface. In contrast, light rays appear to diverge from a **virtual image**, but they do not actually pass through where the image is formed. For example, the image of the pill bottle in Figure 23 is a virtual image. Our system of visual perception, consisting of our eyes and brain working together to interpret visual stimuli, cannot ordinarily distinguish between an object and the image of that object reflected in a mirror. The illusion of an object existing behind a mirror is due to the fact that light from a reflected object enters the eye in exactly the same manner as if it had been emitted by a physical object at the location of the image.

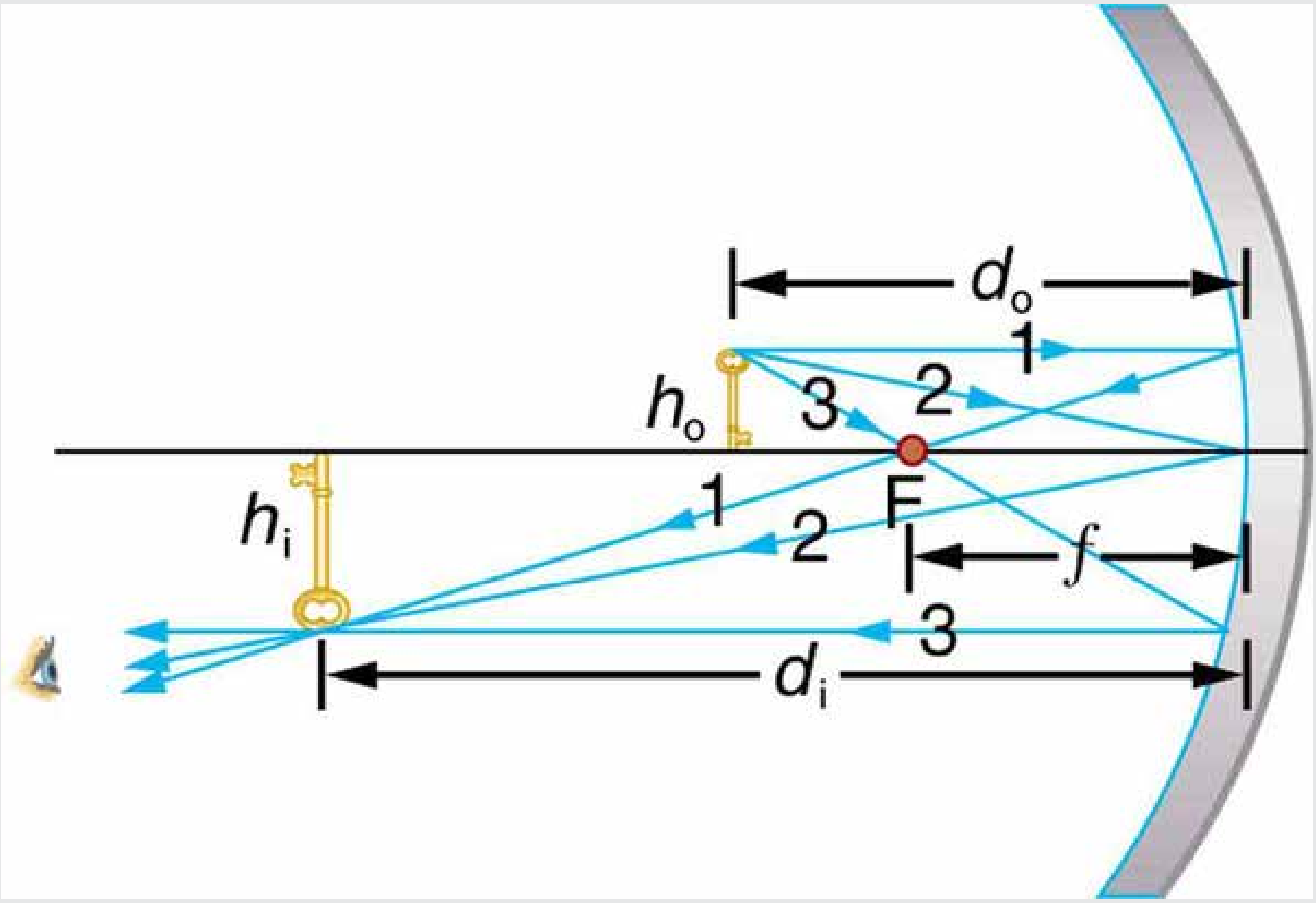
Ray tracing is the technique of determining or following (tracing) the paths that light rays take in order to determine the position and size of an image. A diagram (like the one in Figure 23) showing the paths that these rays take is called a **ray diagram**. The laws of reflection and refraction, along with some simple approximations, allow us to predict the paths the light rays will take. Ray diagrams for curved mirrors and lenses include a **principal axis**, which is an imaginary line passing at a right angle to the surface of the mirror or lens, through its center. As we will see, ray diagrams can help us understand how mirrors and lenses form images in a wide variety of situations.

FIGURE 23



Reflection of an object from a plane mirror. The image is the same size as the object and appears to be the same distance from the mirror as the object.

FIGURE 24



A ray diagram for an image formed by a concave mirror when the object is outside the focal point (F). Ray 1 enters parallel to the principal axis and reflects through the focal point; ray 2 reflects at equal angles at the center of the mirror; and ray 3 enters through point F and leaves parallel to the principal axis.

CONCAVE AND CONVEX MIRRORS

Not all mirrors we use are flat. Cosmetic mirrors, security mirrors in stores, and passenger side mirrors in cars all have curved surfaces that cause images to be formed in different ways compared to plane mirrors. For curved mirrors, the sizes and distances of objects and images are no longer equal, which can cause images to appear warped or distorted. If you've ever looked into the reflective surface of a spherical holiday ornament or a curved "funhouse" mirror, you may recall seeing a distorted version of your face or body. We can think of a curved mirror as many small plane mirrors connected together, each oriented at a slightly different angle from the one adjacent to it. We can still apply the law of reflection to this situation, with the caveat that the normal to the surface will be different at every point along the surface.

Concave Mirrors

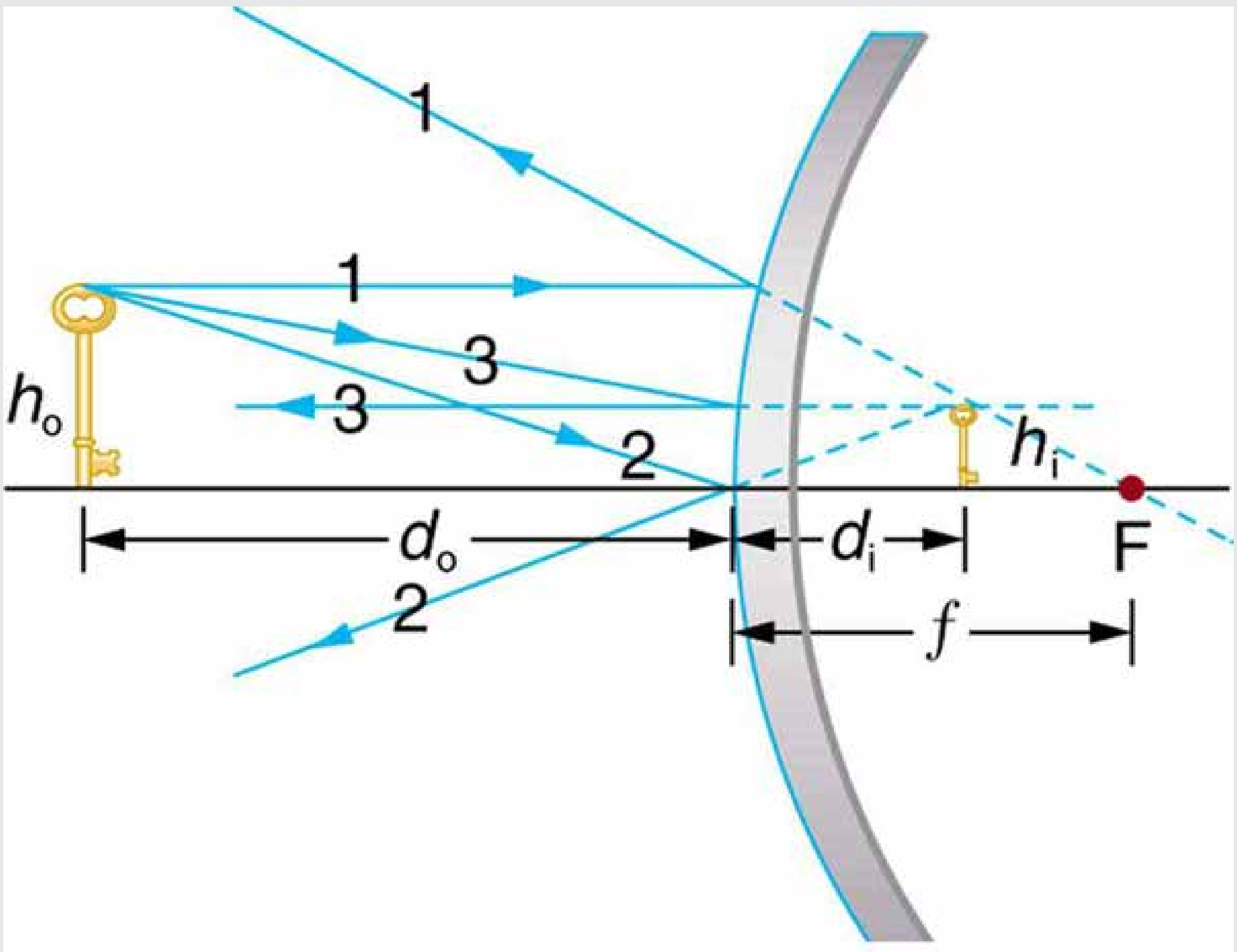
A **concave mirror** is shaped such that its reflective surface curves inward, away from incident light. Concave mirrors reflect incoming parallel light rays inward toward a single point, called the **focal point**, which lies along the principal axis of the mirror. The distance from the surface of the mirror to the focal point is called the **focal length**, often abbreviated *f*. Concave mirrors produce different types of images depending on the position of the object relative to the mirror. For objects between the mirror and the focal point, the image is virtual, upright, and magnified. When the object is located outside the focal point, the resulting image is real and inverted (Figure 24). As the object moves farther and farther away, the light rays arrive at the mirror very nearly parallel to one another, and the resulting image is formed very close to the focal point of the mirror. We won't show ray diagrams for each of these cases here, but you can find them in most introductory physics texts.

Since concave mirrors can form upright, magnified images of objects that are within their focal point, they are particularly useful for viewing an enlarged reflection of one's face when shaving or applying makeup. Solar cookers also use concave mirror to focus sunlight at a specific point to cook food. Concave mirrors are also useful for projecting rays straight outward from a light source, as in flashlights and automobile headlights. As we will discuss in a later section, concave mirrors are also used in the construction of optical cavities, which are important in laser operation.

Convex Mirrors

A **convex mirror** is shaped such that its reflective surface curves outward toward the light source. Convex mirrors

FIGURE 25



A ray diagram for an object reflecting from a convex mirror. The rays follow the same rules as in Figure 24, but must be extended through the surface of the mirror. The image is virtual, upright, and smaller than the object.

also have a focal point, but it is located on the opposite side of the mirror from the object. Although convex mirrors reflect light outward and cannot be used to focus light, they can create images. Images produced by convex mirrors are always virtual, upright, and reduced (i.e., smaller in size than the object), as seen in Figure 25. As the object approaches the mirror, the image becomes larger until it is approximately the same size as the object (when the object is in contact with the mirror).

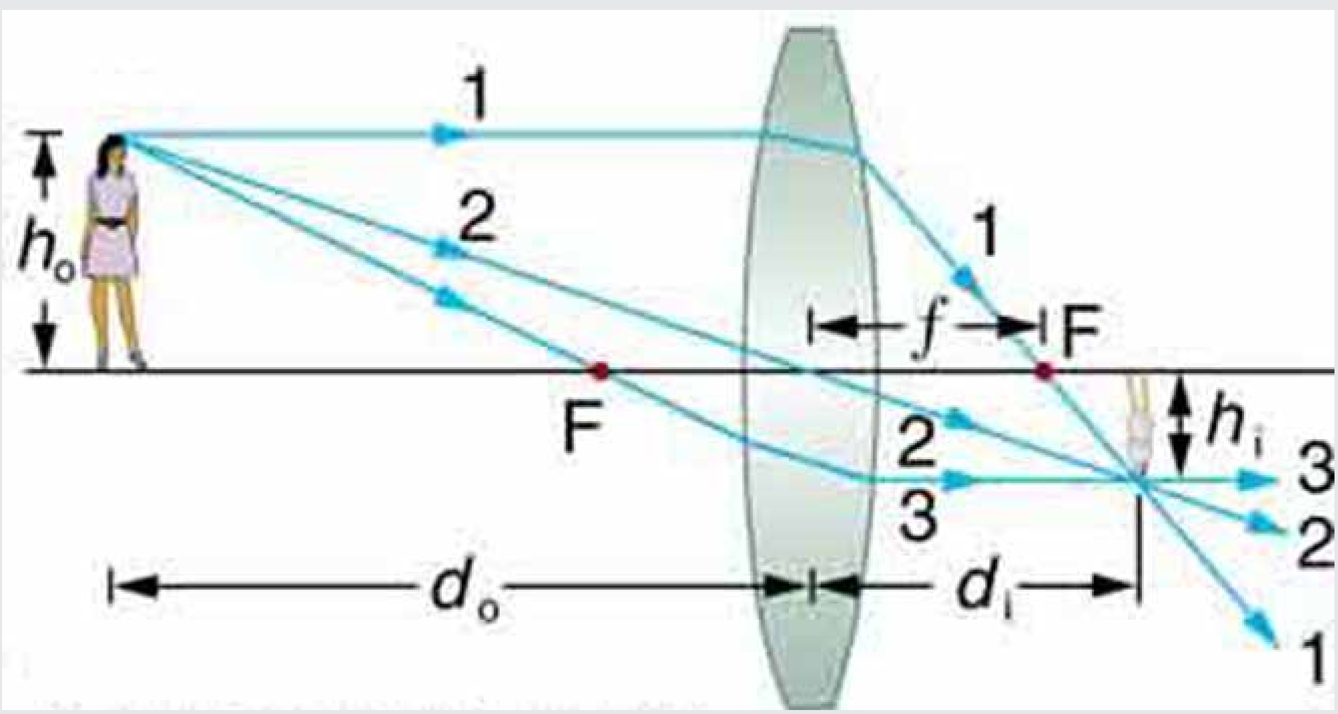
You may be familiar with the safety warning "OBJECTS IN MIRROR ARE CLOSER THAN THEY APPEAR" that is required by law to be etched on passenger side mirrors of motor vehicles in the United States. Convex mirrors are preferable for side mirrors are because they always produce an upright image while providing the driver a wider field of view than a plane mirror would. However, the reduced size of the image may lead the driver to believe that the corresponding objects are farther away than they actually are, hence the warning message.

Convex mirrors are also often mounted in large buildings such as hospitals, stores, or schools to allow people to see around the corners of hallways or aisles. Likewise, they are often seen mounted along roadways where there is reduced visibility or the potential for "blind spots" around turns. Convex mirrors made from blown glass were likely used as early as the fifteenth century, as they are depicted in paintings from that period. Convex mirrors were likely more common than plane mirrors during that time because glass-blowing technology of the time made curved mirrors easier to produce.[8]

LENSES

We have seen how materials can cause rays of light to bend through refraction. A **lens** is a carefully shaped piece of transparent material, such as plastic or glass, that refracts light rays to produce an image. Lenses are essential to many optical instruments that we use on a regular basis, including eyeglasses and cameras. We can think of a lens as a series of connected prisms, each with a slightly different surface angle, that work together to form a bright, focused image of an object. Like curved mirrors, lenses also

FIGURE 26



A ray diagram for an image formed by a converging lens. The three labeled rays correspond to the three principal rays defined in the text.

have focal points at which they concentrate light. However, lenses focus light rays through refraction, not reflection. Lenses also have two focal points, one on each side, because light can enter from either face of the lens.

Converging Lenses

A **converging lens**, also known as a convex lens, refracts incoming parallel light rays so that they meet, or converge, at a focal point. Converging lenses can be identified by their shape, as they have convex outer surfaces. In other words, they are thicker toward their center than at their edges. An object infinitely far away from a converging lens will create a point image at the focal point. The image is real, which means that it can be projected on a screen. This provides a simple method for determining the focal length of a converging lens such as a magnifying glass—use the lens to create an image of a distant light source, such as the Sun or a high ceiling light, and measure the distance from the lens to the image.

When constructing ray diagrams for lenses, there are three light rays that are particularly useful because their behavior is easy to predict using the refracting properties of lenses. These rays are known as **principal rays**:

1. A light ray traveling parallel to the principal axis (the *parallel ray*) will be focused at the focal point. This comes from the definition of focal point.
2. A ray passing through the center of the lens (the *central ray*) will continue in a straight line with no net refraction. This occurs because both sides of a lens are parallel to one another along any path through the center of the lens. By approximating the lens as having a negligible thickness, we can assume that the ray does not deviate from a straight line.
3. A light ray passing through the focal point (the *focal ray*) will leave the lens parallel to the principal axis on the opposite side of the lens.

Figure 26 provides an example of how these three rays can be used to predict the location and size of an image for a converging lens.

A converging lens can produce a real or a virtual image of an object, depending on where the object is located relative to the focal point. As a distant object approaches the focal point of a converging lens, the image becomes larger and farther away. When the object is between a converging lens and its focal point, the light rays from the object diverge when they pass through the lens, and an enlarged virtual image is formed on the same side of the lens as the object. This case describes a simple magnifier, which we will discuss later in this section.

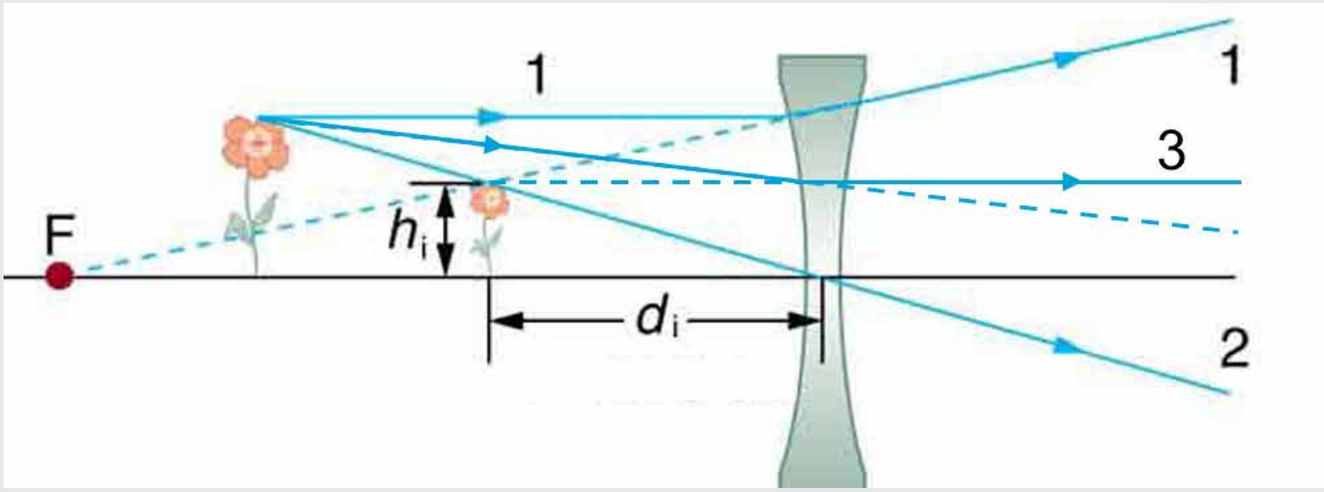
Try It! Make a Pinhole Camera

Instead of lenses, early cameras used small apertures the size of pinholes to produce an image. You can make your own pinhole camera by removing one side of a cardboard box and covering the open side with a semitransparent sheet of tracing or tissue paper. Using a very thin, sharp instrument such as a pin or needle, poke a hole in the side of the box opposite the paper, and point the hole toward an illuminated object. An inverted image of the object will appear on the sheet of paper opposite the pinhole. The smaller you make the pinhole, the sharper the image will be.

You can even make an actual "camera" by using unexposed photographic film in place of the tracing paper, closing the back of the box so that no extra light can enter, and adding a removable flap to cover the pinhole. When you're ready to take a picture, open the flap to allow light from the object to expose the film, producing a photograph. You'll need to test various exposure times depending on the type of film and the amount of available light.

Pinhole cameras are excellent tools for viewing solar eclipses. When the camera is pointed at the Sun, a crisp image of the eclipse (that's safe to view) will appear on the screen. Even without constructing a camera, you can observe a similar "pinhole" effect the next time you are walking underneath a leafy tree on a sunny day. When the spaces between the leaves are small compared to the height of the tree, they act as "pinholes" that project circular spots of light on the ground. The many circles are each individual images of the Sun cast by different openings between leaves.

FIGURE 27



A ray diagram for an image formed by a diverging lens. The three principal rays are labeled.

Diverging Lenses

A **diverging lens**, or concave lens, causes incoming light rays to spread out, or diverge. The focal point of a diverging lens is defined as the point from which the diverged rays appear to originate. Diverging lenses can only produce virtual images. When a diverging lens is used alone, the image is always virtual, upright, and reduced in size compared to the object. Diverging lenses have concave surfaces; they are thicker toward their edges than at their center.

Ray diagrams for diverging lenses use the same principal rays as converging lenses. However, these rays may need to be extended backward in order to determine the point from which they appear to have originated. Figure 27 provides an example of a ray diagram for a diverging lens. The first ray, parallel to the axis, appears to come from the focal point on the same side of the lens as the object. The second ray passes through the center of the lens and is not refracted. The third ray is drawn as if it were going to the focal point on the opposite site of the lens. As this ray passes through the lens, it is refracted parallel to the principal axis and must be extended backward, as shown by the dashed line. The location of the tip of the image is the point at which the three rays appear to have originated.

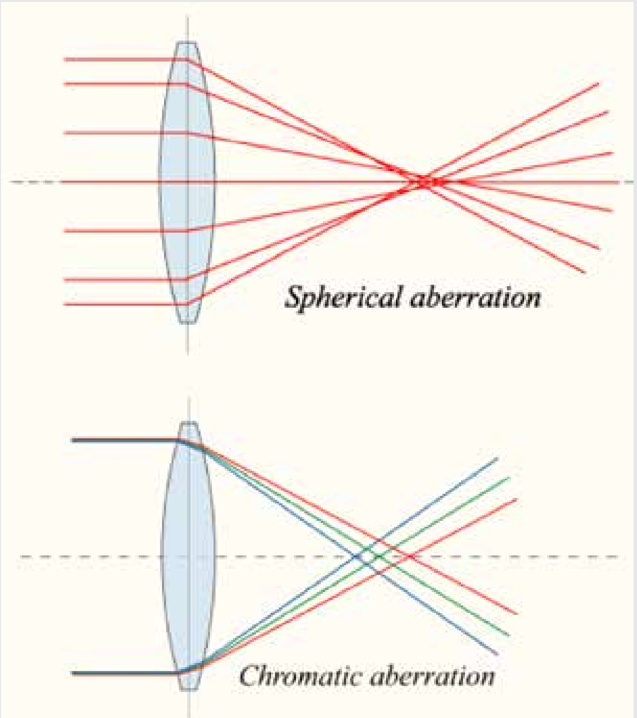
It's important to note that although we have been using three rays in order to locate the position of an image through ray tracing, many more rays pass through the lens and are also refracted to pass through the same image point as the others. Light rays also leave every point on the object (either through emission or reflection) and pass through every part of the lens. What would you expect to happen to the image of the person in Figure 26 if the top half of the lens were covered up? It's natural to think that an image of half the person would be formed, but this is incorrect—you would still see a complete image of the person, although it would be fainter because less light is being focused to produce it.

Spherical and Chromatic Aberration

No lens or mirror produces a perfect image. Due to refractive and reflective effects, real mirrors and lenses often behave differently than we might expect according to a simple ray analysis. A distortion of an image caused by unintended lens or mirror behavior is called an **aberration**. There are two primary types of aberrations: *spherical* and *chromatic*.

Spherical aberration occurs when light passing through the edges of a lens or reflecting from the edges of a mirror is focused at a slightly different point than light that is incident near the center (Figure 28), resulting in a blurred image. In a camera, spherical aberration can be reduced by using an adjustable aperture to control the amount of light passing through each part of the lens. As the aperture size is made

FIGURE 28



Spherical and chromatic aberration.

smaller, only the center of the lens is used to focus light. However, this also reduces the amount of light being used to produce an image, so a longer exposure time may be necessary to compensate. Spherical aberration in mirrors can be corrected by using a parabolic rather than spherical mirror.

Chromatic aberration is caused by light of different colors (i.e., wavelengths) being refracted by different amounts by a lens and, therefore, converging at slightly different points (FIGURE 28). Recall that violet light undergoes a greater degree of refraction than red light, so violet light accordingly has a shorter focal length than red light for the same lens. Chromatic aberration can be corrected by combining lenses of different types of glass. Although curved mirrors can cause spherical aberration, they do not introduce chromatic aberration because no refraction takes place.

These defects also affect image formation by the lenses of your eyes. Your pupils change size, or dilate, in order to regulate the amount of light entering your eyes. When your pupils are at their most narrow, light passes mostly through the center of the lens, where spherical and chromatic aberrations are not as pronounced. This is one reason why our vision tends to be much sharper in bright light.

WAVE OPTICS

HUYGENS' PRINCIPLE

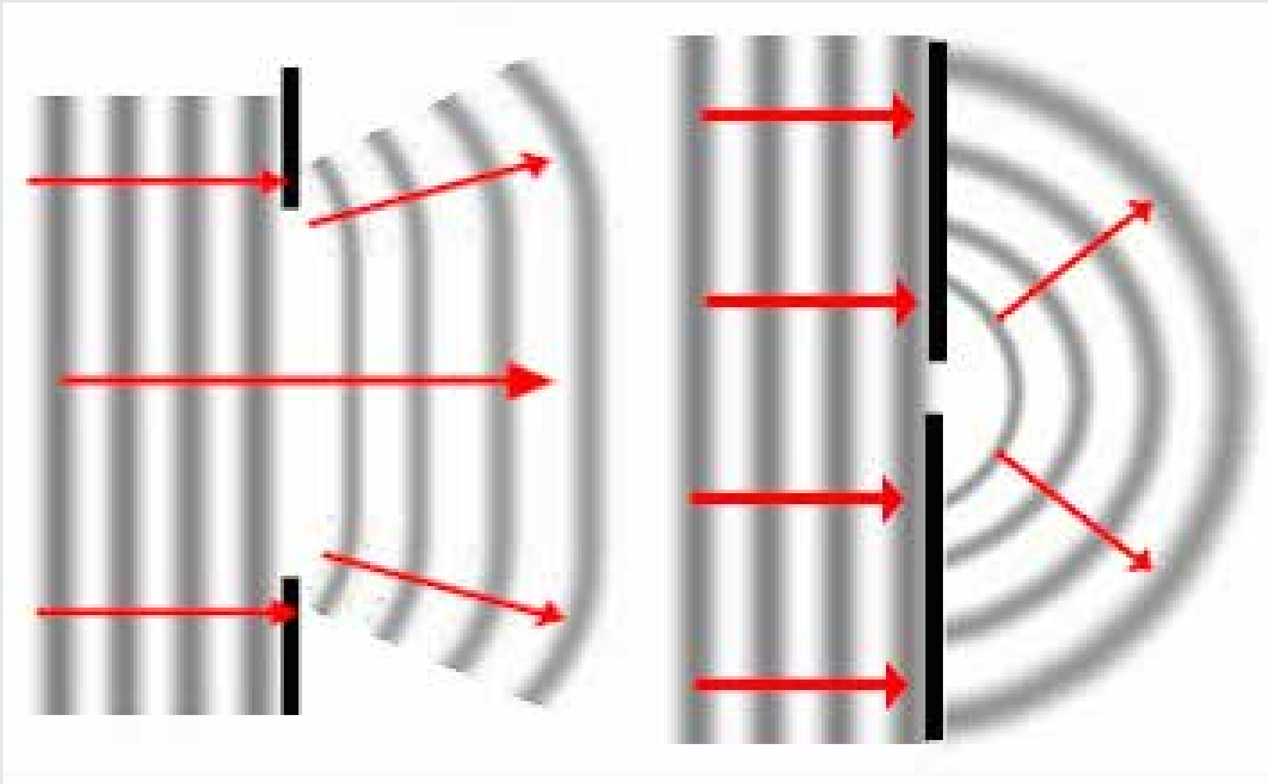
In the mid-seventeenth century, long before Maxwell proposed the existence of the electromagnetic wave, scientists disagreed about the exact nature of light. Sir Isaac Newton, based on numerous experiments involving prisms and mirrors, argued that light must be made up of weightless particles, called corpuscles, that travel in straight lines and rebound elastically from surfaces. Newton's theory could account for the basic geometric phenomena of reflection and refraction, but other optical effects could not be adequately explained. For example, light passing through a narrow slit had been observed to spread outward instead of casting a sharp image of the slit, which could not be easily explained with a corpuscle model.

In 1690, Christiaan Huygens, a Dutch physicist and contemporary of Newton, published an alternative wave theory of light. Huygens envisioned light as a wave that propagates radially outward from a source, just as ripples spread outward from a pebble dropped in water. **Huygens' principle** states that every point of a wavefront can be treated as itself a source of secondary wavelets that spread out in all directions. Although it took several centuries before it became widely accepted, Huygens' wave theory is now fundamental to our understanding of light. The effects of diffraction, interference, and polarization are all direct consequences of the wave nature of light.

DIFFRACTION

We have already described two ways in which light can change its direction of propagation: reflection and refraction. We will cover one more, but it can only be accounted for by considering the wave behavior of light. You may have noticed how you can hear sounds even when the source is around a corner or on the opposite side of a barrier. This is due to **diffraction**, which is the bending of a wave around an edge of a surface or through an aperture (FIGURE 29). Huygens' principle helps explain how light spreads through openings by considering every point along the front of the wave to be itself a point source of light.

FIGURE 29



Diffraction of waves through an aperture.

When the size of an opening is large compared with the wavelength of light, it casts a shadow with sharp, well-defined edges that resembles the shape of the opening. As we narrow the opening to be closer to the wavelength of light, the edges of the shadow become fuzzier and the light emerging from the opening spreads out. Diffraction occurs for all waves, not just for light or sound. For each kind of wave, the same general rule applies: diffraction is most apparent when the wavelength of the wave is comparable to the opening or obstruction in the path of the wave.

A good example comes from radio waves. The radio waves used in AM radio have wavelengths approximately 100 times longer than those used in FM radio (about 300 meters compared with about 3 meters). Accordingly, AM radio waves have an easier time diffracting around buildings; AM radio reception is often better than FM radio reception in densely populated areas as a result.

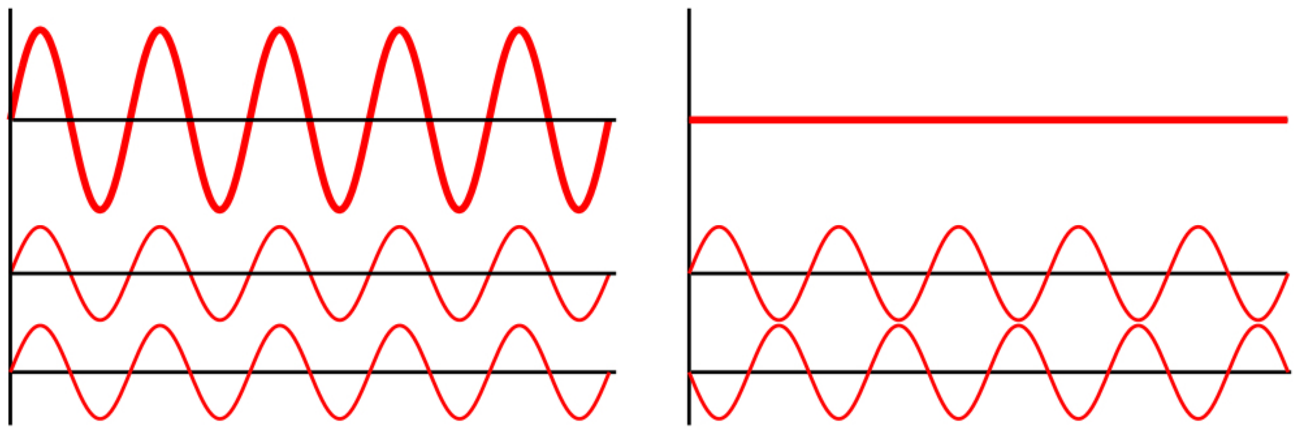
Diffraction effects also place a limitation on how useful conventional microscopes can be for viewing very small objects. For instance, most viruses are about 20–300 nm in diameter, which is smaller than the wavelengths of visible light. When viewed using a conventional microscope, the edges of the virus become indistinct due to diffraction effects. Electron microscopes, which rely on the fact that all matter exhibits wave-like properties, use fast-moving electrons to probe objects at this scale. The wavelength of electrons in such a microscope is thousands of times smaller than the wavelength of visible light, so diffraction effects from the object under observation are minimized.

INTERFERENCE

Superposition

What happens when two waves in the same medium interact? Traveling waves can meet and even pass through one another without either individual wave being altered or destroyed. When two waves are in the same place

FIGURE 30



Totally constructive (left) and totally destructive (right) interference of two waves.

at the same time, we can add their amplitudes together at each individual point to find the amplitude of the resulting wave (FIGURE 30). This is known as the principle of **superposition**. In general, the phenomenon of two waves interacting to form a resultant wave is known as **interference**. Interference effects are possible with any kind of wave. However, these effects are easiest to observe and understand when we consider interference between waves of the same frequency.

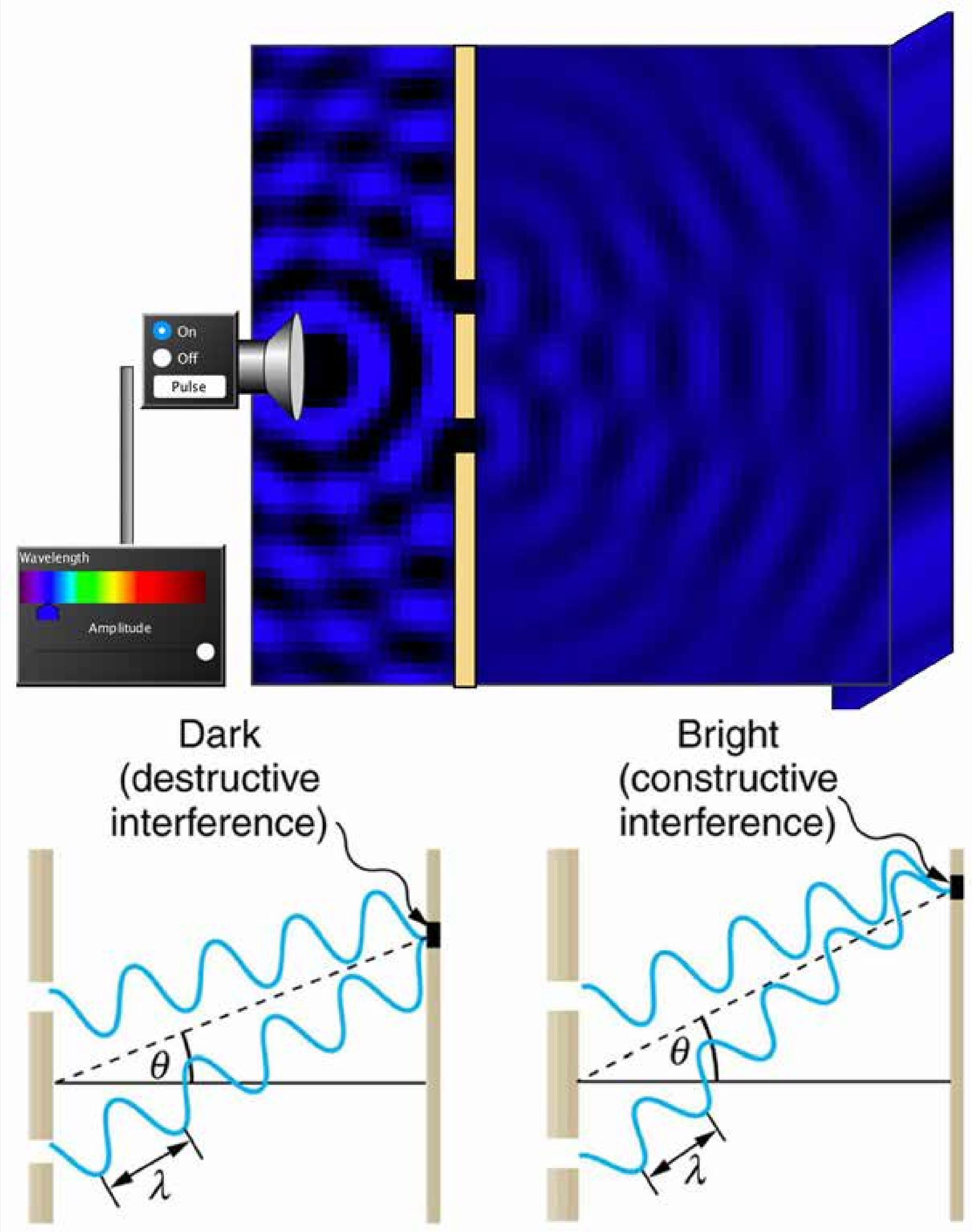
When two waves overlap such that their crests and troughs align perfectly, we say that they are "in phase," and the resultant wave has a larger amplitude than either of the individual waves. This is called **constructive interference** because the waves build constructively on one another. What happens if we shift one of the waves so that the crests of one wave align with the troughs of another? In this case, adding the amplitudes yields zero at every point along the wave. In other words, the waves cancel each other out in what is known as **destructive interference**. Destructive interference between reflected sound waves can result in acoustical "dead spots" in theaters, where no sound is audible. For this reason, acoustical engineers must design performance spaces to minimize the effects of destructive interference. Destructive interference can be put to practical use in noise-cancelling headphones, which measure ambient sound and then generate an out-of-phase sound wave with the same amplitude to cancel out the ambient sound.

Double Slit Interference

In 1801, Thomas Young investigated a phenomenon known as **double slit interference**, in which light passing through two thin parallel slits produces a pattern of bright and dark fringes on a screen (FIGURE 31). This effect is due to alternating constructive and destructive interference, which can only be explained by treating light as a wave. Light waves exiting each slit have the same phase (i.e., they have crests lined up with crests) but take paths of different lengths to reach the same point on the screen. Bright fringes occur where a crest from one slit reaches the screen at the same time as a crest from the other slit. Dark fringes occur where a crest from one slit reaches the screen at the same time as a trough from the other slit, thereby canceling each other out. Young's discovery provided convincing evidence in favor of Huygens' wave model of light.

We can use a similar analysis to consider diffraction effects from a single slit with appreciable width. Recall from Huygens' principle that we can treat every point along a wavefront as a point source of light. By applying Huygens' principle to the slit, we can consider individual points along the slit as "in-phase" sources of light. The light "emitted" by each tiny part of the slit travels a different distance to reach a corresponding point on the viewing screen, and therefore they can arrive in or out of phase. Constructive and destructive interference between these individual wavelets results in a pattern of bright and dark regions

FIGURE 31

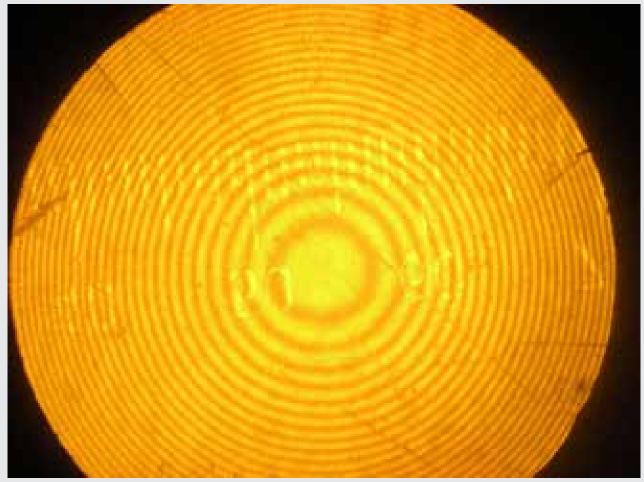


Top: A simulation of double slit interference. You can explore diffraction and interference effects using the "Wave Interference" [PhET applet](phet.colorado.edu) at phet.colorado.edu. Bottom: The bright and dark fringes are a result of constructive and destructive interference, respectively.

FIGURE 32

A layer of very evenly spaced bacteria floating on the surface of a fish tank acts as a natural diffraction grating.

FIGURE 33



Newton's rings, as observed through a microscope.

along the screen. Effectively, the presence of the slit causes the emerging wave to interfere with itself, producing a diffraction pattern.

Diffraction Gratings

Interference patterns can also be created using more than two slits. A **diffraction grating** uses a large number (often thousands per centimeter) of equally spaced, parallel slits to produce interference effects. Like prisms, diffraction gratings spread light into a spectrum of colors; however, the effect is due to interference and not dispersion. Compared to double slit interference patterns, the interference patterns produced by diffraction gratings tend to be more spread out, and the dark fringes tend to be broader because the greater number of slits provides more opportunities for destructive interference to occur. Diffraction gratings can be made by etching very fine lines on glass using a diamond point; the narrow spaces between the scratches then act as slits.

You've probably noticed a spectrum of colors when viewing the underside of a CD or DVD. A CD contains evenly spaced grooves separated by about 1 micrometer (one millionth of a meter). These grooves act as a diffraction grating that separates different wavelengths of light toward your eye. Because the surface is reflective, the resulting interference pattern can be viewed simply by looking at the CD. Diffraction effects are also observed in nature; small repeating formations such as striated muscle tissue, crystals, or bacterial layers can act as diffraction gratings, which is often useful to scientists in determining their dimensions and structure (Figure 32).

Thin-Film Interference

Long before Young's double slit experiment, Sir Isaac Newton made observations of interference effects, but he did not fully grasp their significance. Newton placed a spherical lens on top of a flat piece of glass, leaving a small air gap between the edges of the two pieces of glass. When this combination is illuminated by light, a series of concentric circular fringes appear when viewed from above. These fringes, known as **Newton's rings**, are similar to the bright and dark fringes observed in double slit interference (Figure 33). Newton's rings are caused by incident light moving through different thicknesses of glass at every point along the lens, and then interfering with light that is reflected from the top of the lens when it re-emerges at the top surface. Newton's rings are useful in testing the curvature of precisely ground lenses; asymmetric ring patterns indicate that the lens is not perfectly spherical and must be reshaped.

Interference effects are also responsible for the iridescent colors seen in a soap bubble or an oil slick. These colors appear when the thickness of the transparent layer is close to the wavelength of light used to illuminate it. Consider multiple rays of light incident on a thin transparent material such as a layer of oil floating on water (Figure 34). Some light waves are reflected from the top surface of the film, whereas others are transmitted through the film and reflect

FIGURE 34

Thin-film interference effects seen in a spill of diesel fuel. The colors magenta, yellow, and cyan are prominent.

back up from the bottom layer. Because each of these waves travels a different distance, they interfere with one another when they emerge from the top surface of the film.

When a thin film is illuminated by white light, the thickness of the film determines which wavelength of light will undergo destructive interference. The light emerging from the film at that point will appear as the **complementary color** of the light that was canceled. (A complementary color is the color that remains when another color is subtracted from white light.) For example, if red light is canceled out due to destructive interference, the light that emerges will appear as the complementary color of red, which is cyan. Magenta (the complementary color of green) and yellow (the complementary color of blue) also tend to be prominent in thin soap films and layers of oil. The thickness of the film may vary slightly from point to point, causing other colors to destructively interfere and other complementary colors to appear as a result. This gives the film a rainbow-like appearance (FIGURE 34).

By viewing an oil or gasoline slick from different angles, you will see the colors shift even though the film itself has not changed thickness. This is because the light reaching your eyes is incident at a different angle, and therefore the transmitted wave travels a different distance through the film. The colors on the surface of a soap bubble also change as the bubble becomes thinner and thinner due to evaporation. Thicker walls tend to cancel out longer wavelength light (i.e., red), leaving cyan. As the wall gets thinner, shorter wavelengths such as blue undergo destructive interference. Eventually the bubble is too thin to cause interference of any visible wavelengths of light, and it appears colorless just before popping.

Try It!
Thin-Film Interference

The next time you are washing dishes in the sink, you can try making observations of interference effects with thin soap films. Try inverting a cup or mug in soapy water to create a thin film across the opening. What colors do you see? If you turn the cup sideways, the film will flow downward to become thicker toward the bottom in a narrow "wedge" shape. You should see layers of color corresponding to the varying thickness of the film. At the thinnest region near the top of the film, the film appears colorless because no interference in the visible range is occurring.

POLARIZATION

Polarization is a property of light that plays an important role in how it interacts with materials. Later we will see that polarization is important to the operation of devices such as lasers. The polarization of an electromagnetic wave refers to the plane in which the electric field oscillates as the wave propagates. To help visualize the polarization of a wave, imagine you take a rope attached to a wall at arm height and pull it taut. Then, imagine shaking your arm up and down rhythmically to produce a transverse wave that travels along the length of the rope. We can think of this wave as an analog for a light wave, with the displacement of the rope at any point corresponding to the magnitude and direction of the electric field. (The magnetic field, although an essential component of the electromagnetic wave, is not present in this analogy.) In this case, the wave is polarized along the vertical axis in which you are moving your arm.

We described in Section I how electromagnetic waves are produced when a charged particle, such as an electron, undergoes acceleration. An electron that accelerates along a vertical axis will produce vertically polarized light, whereas an electron that accelerates along a horizontal axis will produce horizontally polarized light. **Linearly polarized** light, also known as plane polarized light, is made up of waves that are polarized along a single axis. At every point in a linearly polarized wave, the electric field is always directed parallel to the axis of polarization.

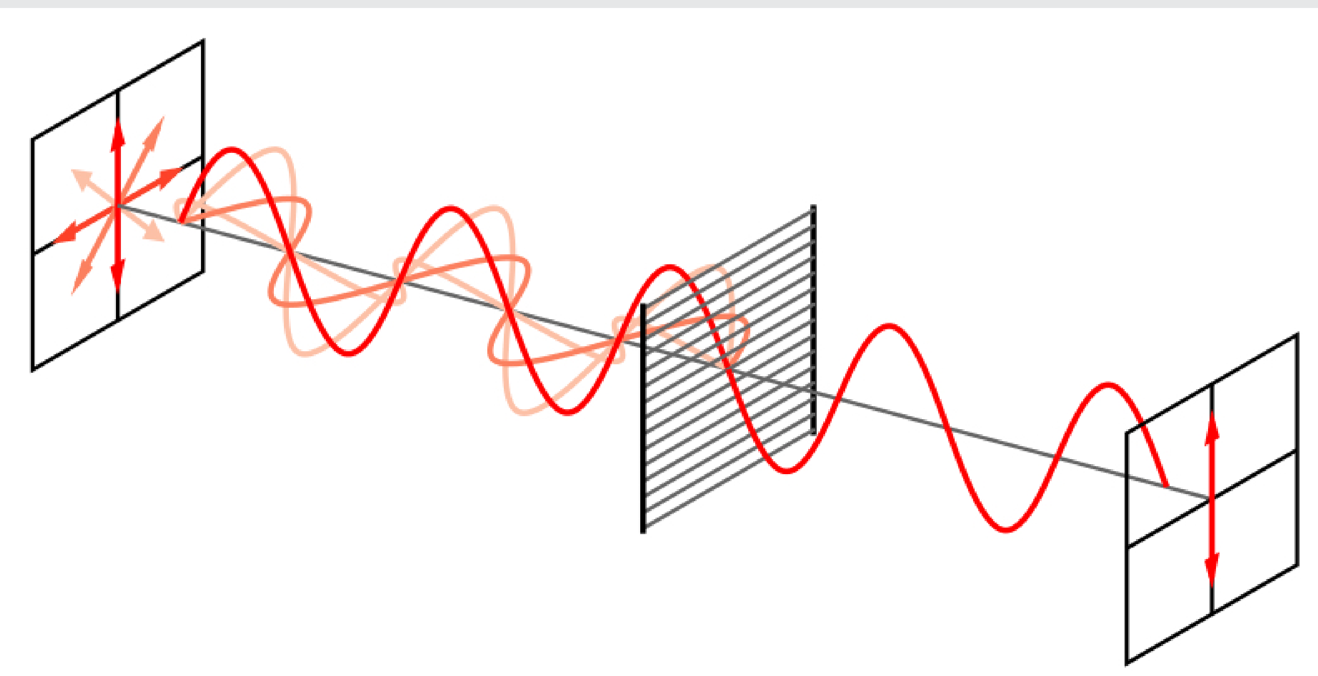
Unpolarized light is made up of waves that are polarized along many different axes. Most everyday light sources, such as light bulbs, candles, and the Sun, emit unpolarized light. Light from these sources is produced by a multitude of electrons that are vibrating in all different directions, so there is no preferred direction of polarization. How, then, can we obtain polarized light? The most common method is to use a specially prepared material to separate out light of a specific polarization from unpolarized light.

In 1938, E. H. Land discovered a material that could polarize light, which he called the **Polaroid**. A common Polaroid filter is made up of many molecules called *hydrocarbons* arranged in long parallel strands. The material is prepared such that the constituent molecules are especially good conductors, and they tend to absorb light that is polarized along their axis of orientation while allowing light that is polarized perpendicular to this axis to pass through (Figure 35). What would you expect to see if you stacked two polarizing filters, oriented them perpendicular to one another, and then looked through them? No light would pass through both filters, so they would appear completely dark.

Light reflected from nonmetallic surfaces is partially polarized by the surface. The amount of polarization depends on the angle of incidence. Most of the reflected light is polarized parallel to the surface from which it is reflected. There is one angle of incidence, called the **polarizing angle**, at which reflected light will be completely polarized parallel to the surface of the interface. The polarizing angle depends on the index of refraction of both materials on either side of the reflective interface. The polarizing angle is also known as **Brewster's angle** after Sir David Brewster, a Scottish physicist who first discovered it.

The polarization of reflected light is a common effect that plays a role in our everyday lives. Light that is reflected from surfaces such as water, snow, or glass undergoes some degree of polarization. Polarized sunglasses tend to have vertically-oriented polarization axes because most

FIGURE 35



A polarizer takes incident unpolarized light and transmits linearly polarized light.

of the light they are designed to block is reflected from horizontal surfaces and therefore tends to be horizontally polarized. Likewise, photographers often use polarizing filters to reduce unwanted glare from reflective surfaces such as water or glass.

OPTICAL INSTRUMENTS

Lenses can be combined in various ways to form images that we would not be able to see with our eyes alone. In the remainder of this section, we will consider three familiar examples of optical instruments: simple magnifiers, microscopes, and telescopes. Each of these instruments can be understood by applying similar optical principles.

SIMPLE MAGNIFIER

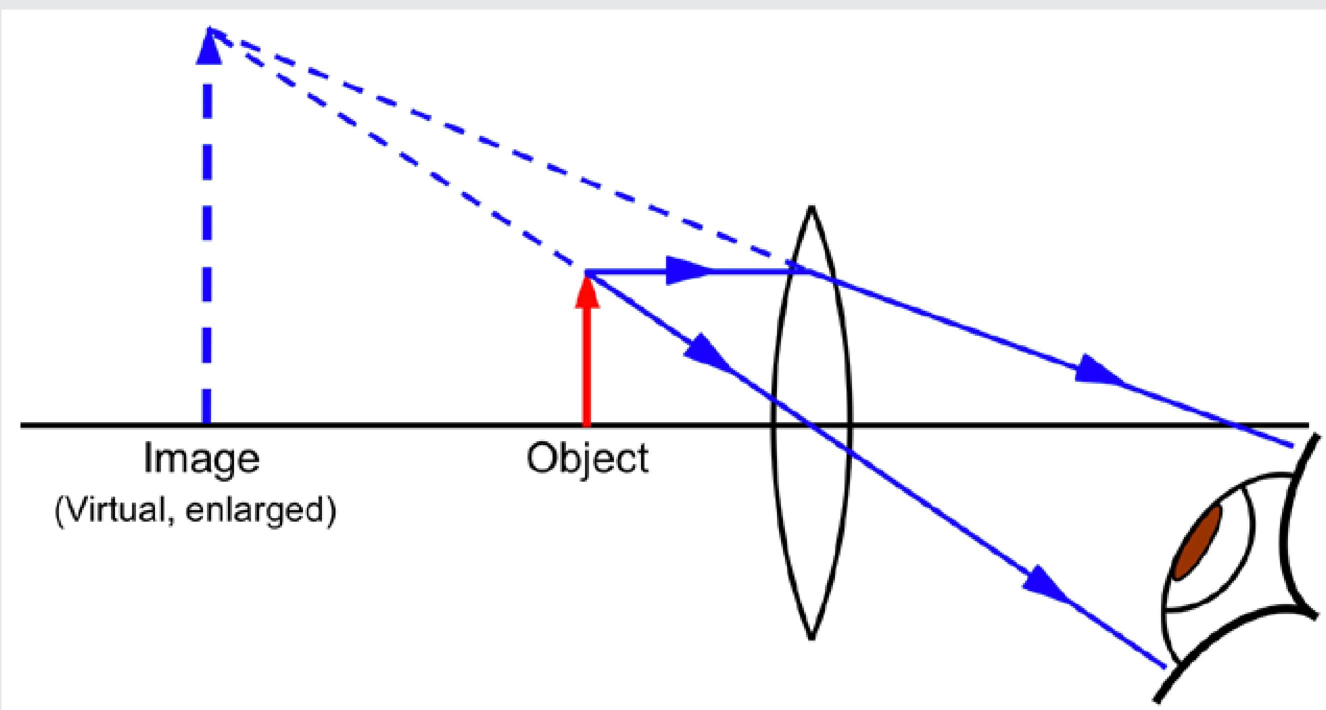
The simple magnifier is among the most basic optical instruments, as it consists of just a single converging lens (Figure 36). If you've ever used a magnifying glass, you are familiar with the ability of a simple magnifier to produce enlarged images of small objects such as newsprint or an insect. In order for magnification to occur, the object must be placed within the focal length of the lens and viewed from the opposite side. The resulting image is virtual, magnified, upright, and appears farther away than the object actually is.

MICROSCOPES

A single lens can only provide so much magnification of small objects. In order to probe finer details at small scales, lenses can be combined in order to increase their overall magnifying power, as in a microscope. In general, when considering the behavior of light in multi-lens systems, we can treat the image produced by one lens as the object for the other lens. A **compound microscope** consists of two converging lenses; the lens closer to the object is called the *objective lens* and the lens closer to the observer is called the *eyepiece* (Figure 37). The object to be observed is placed just outside the focal length of the objective lens. The purpose of the objective lens is to form a real, inverted image at a point within the focal length of the eyepiece. The eyepiece then acts as a simple magnifier for this intermediate image. The final image is virtual, magnified, and inverted.

The inventor of the compound microscope is unknown,

FIGURE 36



A simple magnifier consists of a single converging lens. The solid vertical arrow is the object, and the dotted vertical arrow is the resulting virtual image.

although the first models were demonstrated around 1621. Many seventeenth-century scientists, including Galileo, experimented with multi-lens systems for the purposes of magnification. Although our ability to precisely manufacture lenses has steadily improved the magnification power of compound microscopes over several centuries, there is a fundamental limit to how powerful such microscopes can be. This limit is governed not by the strength of the lenses but by the nature of light itself; in order to be seen, the object being viewed under a microscope must be at least as large as a wavelength of light (i.e., 400–700 nm). For smaller length scales, electron microscopes are a more appropriate instrument because they operate with smaller wavelengths.

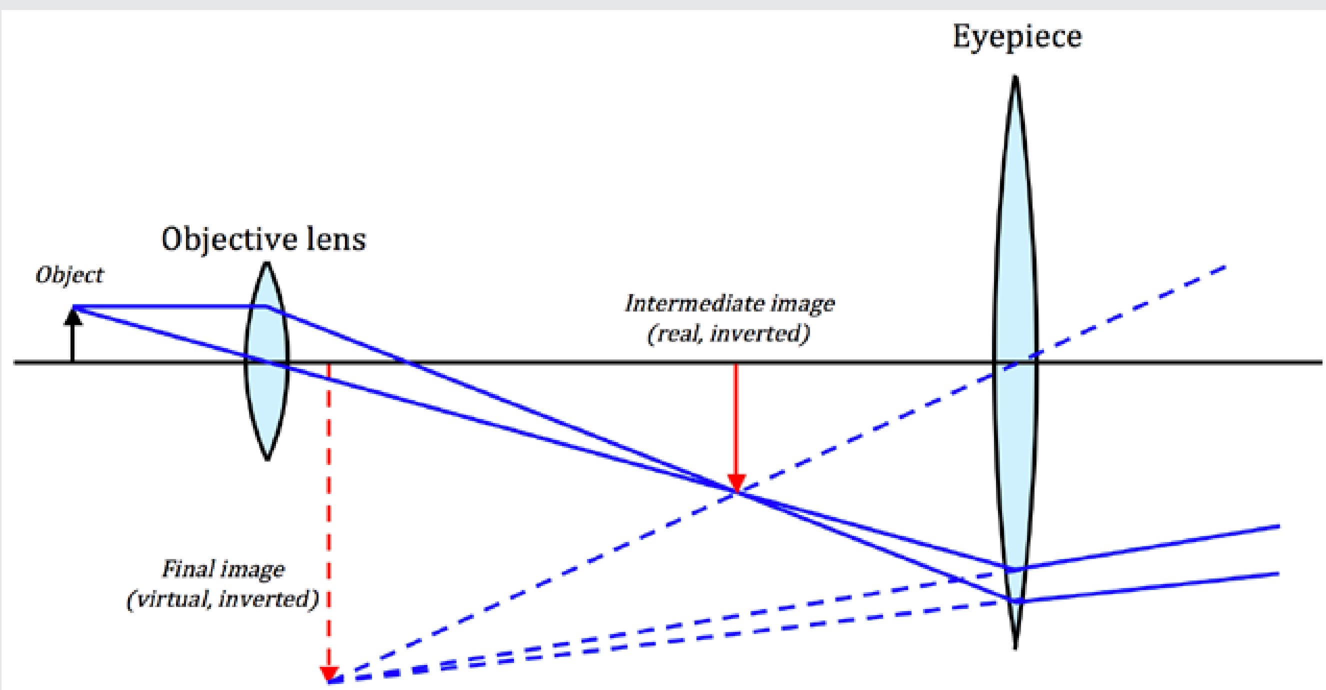
TELESCOPES

Telescopes are used to view very distant objects such as planets, stars, nebulae, and galaxies. A **reflecting telescope** uses a concave mirror to collect and focus light from very distant objects such as planets, moons, and stars. Isaac Newton is credited with assembling the first reflecting telescope in 1688. Since Newton's time, reflecting telescopes have been instrumental in countless astronomical discoveries, including the discovery of the planet Uranus by William Herschel in 1781. The Hubble Space Telescope, which employs a 2.4-meter concave mirror, was deployed in 1990 from Space Shuttle Discovery and remains in operation today. The Keck Observatory on Mauna Kea in Hawaii features two reflecting telescopes with mirror diameters of 10 meters, placing them among the largest optical telescopes in the world today.

A **refracting telescope** uses a combination of lenses to form an image of the distant object. Like a microscope, refracting telescopes use a converging objective lens that is nearer to the object and an eyepiece through which the observer views the final image. Because the object to be viewed is so far away, light rays reach the objective lens nearly parallel to one another and are focused to form a real, inverted image essentially at the focal point of the objective lens (Figure 38). By positioning the eyepiece so that this intermediate image is just inside its focal point, it then acts as a simple magnifier to produce the final virtual image. The length of a refracting telescope is essentially the sum of the focal lengths of the objective lens and the eyepiece.

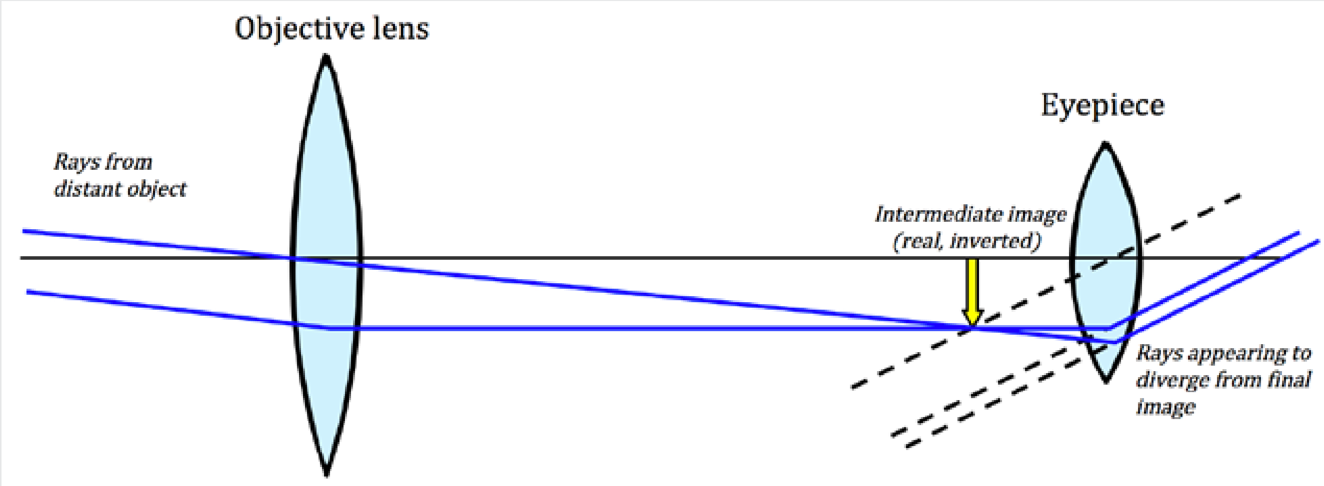
Refracting telescopes are useful for viewing larger, brighter objects in the night sky such as the moon and planets. Some astronomical objects, such as galaxies, are so distant that a large diameter telescope is necessary to collect enough

FIGURE 37



A ray diagram of a compound microscope.

FIGURE 38



A ray diagram of a refracting telescope.

light for proper viewing. Manufacturing the large lenses that would be necessary for a refracting telescope is difficult and costly; large-diameter lenses are also heavy and prone to sagging. Furthermore, as we have discussed previously, refracting telescopes are more susceptible to chromatic aberration than reflecting telescopes that use concave mirrors to focus light. For these reasons, reflecting telescopes tend to be more useful for the purposes of astronomical research.

SECTION II SUMMARY

- Light travels in a straight line until it encounters a boundary with another medium. The ray approximation is a simplification that allows us to represent light as rays, or straight lines with a direction.
- The law of reflection states that the angle of incidence equals the angle of reflection. Reflection from a smooth surface is called specular reflection; reflection from a rough surface is called diffuse reflection.
- Refraction is the bending of light as it moves between transparent materials with different indices of refraction. The index of refraction of a material is equal to the speed of light in a vacuum divided by the speed of light in that material.
- The speed of light in a transparent material is also slightly dependent on wavelength. Consequently, different colors of light undergo different amounts of refraction, an effect known as dispersion. Dispersion causes the spectrum of colors produced by prisms and observed in rainbows.
- Total internal reflection is an optical phenomenon in which a light ray traveling from a region of higher refractive index to a region of lower refractive index is completely reflected at the boundary. Total internal reflection is the basis for fiber-optic technology.
- A mirage is a naturally occurring optical illusion caused by light refracting through air of different densities. A commonly observed mirage is the illusion of pools of water on a hot roadway.
- Real images are formed by light rays passing through the location of the image. Virtual images are not formed by light rays meeting; they are images from which light rays appear to be diverging.
- Curved mirrors produce images with different characteristics than plane mirrors. Convex mirrors are curved outward toward incoming light and produce only virtual images. Concave mirrors are curved inward away from incoming light and produce real or virtual images depending on the location of the object.
- A lens is a shaped piece of transparent material that

refracts light rays to produce an image. Converging lenses refract parallel light rays so that they meet at a focal point. Diverging lenses refract parallel light rays so that they spread apart from a focal point.

- We can use a ray diagram to predict where a lens will produce an image of an object. By tracing rays that emerge from a point object and using the properties of lenses, we can find the image point where the rays converge (or in the case of a virtual image, where the rays appear to be diverging from).
- No lens or mirror produces a perfect image. Spherical aberration occurs when light passing through the edges of a lens is focused at a slightly different point than light passing near the center. Chromatic aberration is a consequence of dispersion.
- In the seventeenth century, Christiaan Huygens was an early proponent of the wave theory of light. According to Huygens' principle, light waves spread outward from a source like ripples in water.
- Traveling waves can meet and pass through one another without affecting or destroying either wave. According to the principle of superposition, we can add the amplitude of each wave together point-by-point to find the amplitude of the resulting wave.
- Constructive interference occurs when two overlapping waves are in phase; their amplitudes combine to produce a wave with a greater amplitude. Destructive interference occurs when two out of phase waves combine to cancel each other out.
- The iridescent colors seen in a soap bubble or oil slick are caused by an effect called thin-film interference. Light that is transmitted through a thin transparent film may be reflected back up and interfere with light that was reflected from the top of the film.
- Linearly polarized light is made up of waves that are polarized along a single axis. Unpolarized light can be polarized after passing through a polarizing material, such as Polaroid, or after reflecting from a surface such as the ocean or a roadway. Brewster's angle is a special angle of incidence at which light reflected from a surface will be completely polarized parallel to the surface.
- Multiple lenses can be combined, as in a microscope or telescope, to form images that we would not be able see with our eyes alone. A compound microscope uses two converging lenses to a produce a magnified image of a small nearby object. A refracting telescope uses two converging lenses to produce an image of a distant object.

Section III
Lasers

INTRODUCTION

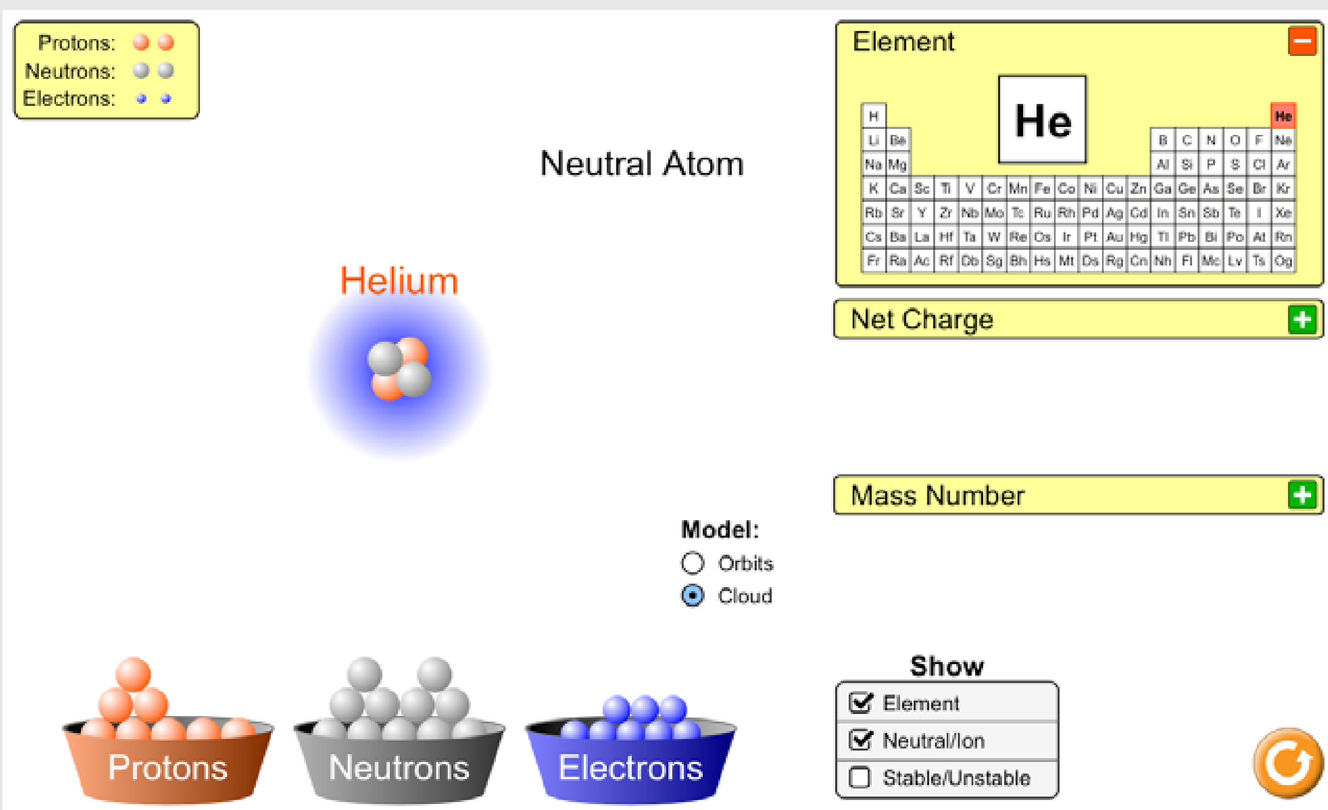
The first part of this section will cover the basic building blocks of laser operation, beginning with how light is emitted by atoms. From there, we will describe the processes of incandescence, fluorescence, and phosphorescence, all of which are relevant to laser operation. The section will conclude with an overview of how lasers work and a history of the development of the laser.

LIGHT EMISSION AND ABSORPTION

THE ATOM

In order to understand how light is emitted, we must first consider the structure and properties of the atom. **Atoms** are the building blocks of the world around us. Atoms contain a **nucleus**, which is a dense core that comprises most of the atom's mass, and an outer region inhabited by bound

FIGURE 39



A neutral atom contains an equal number of protons and electrons. You can combine protons and electrons to form atoms using the "Build an Atom" PhET app at phet.colorado.edu.

electrons. The nucleus is made up of positively charged particles called protons as well as neutrons, which have no electric charge. The nucleus is surrounded by negatively charged electrons. These electrons occupy "shells" that have different spatial distributions. Some are spherical, some are barbell-shaped, and others have more complex distributions. Due to the attractive force between oppositely charged particles, the farther an electron is from the nucleus, the more potential energy it has with respect to the nucleus.

A substance composed of only one kind of atom is called an **element**. Each element in the periodic table has a characteristic **atomic number**, which is equal to the number of protons in the nucleus of an atom of that element. Electrons and protons have the same magnitude, but opposite sign, of electric charge. Because atoms are electrically neutral, they must have the same number of electrons as protons (Figure 39). (Atoms can gain or lose electrons to become **ions**, which have an overall negative or positive charge.) Hydrogen, which has an atomic number of 1, contains a single proton in its nucleus, orbited by a single electron. Likewise, all carbon atoms (atomic number 6) contain six protons and six electrons.

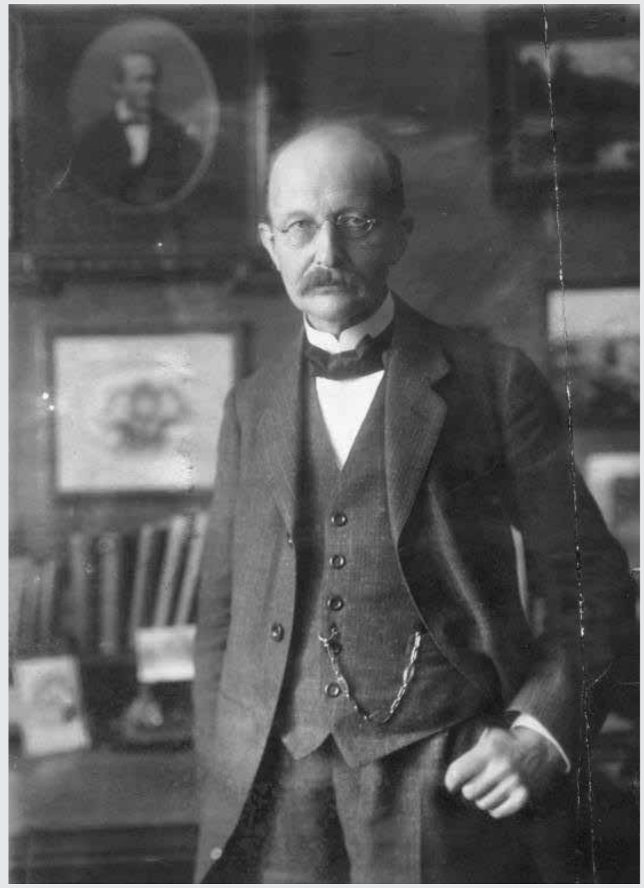
QUANTIZATION

Lasers are fundamentally quantum mechanical devices. That is, the operation of a laser is fully dependent on the quantum nature of light and matter. Let's explain what this means. Before 1900, scientists assumed energy could vary continuously and be endlessly subdivided. No experiment had provided evidence to contradict this viewpoint. In 1900, a German physicist named Max Planck (Figure 40) proposed that the energy could only be emitted or absorbed in discrete bundles that are multiples of a fundamental unit, or **quantum**, of energy. This relationship can be expressed as

$$E = nhf,$$

where $h = 6.63 \times 10^{-34}$ J·s is called Planck's constant; f is the frequency of oscillation; and n is a positive integer. That is, energy could be exchanged in amounts of hf, $2hf$, $3hf$, etc., but not anything in between. Because the energy can only take certain specific values, we say that it is "quantized." As an analogy, imagine dunes of sand that appear smooth from a distance, but are in fact coarse and grainy when observed close up. Planck's constant is small enough that we do not notice the quantization of energy in our everyday experience—the quantized packets of energy are too small to be individually perceived by our senses.

FIGURE 40

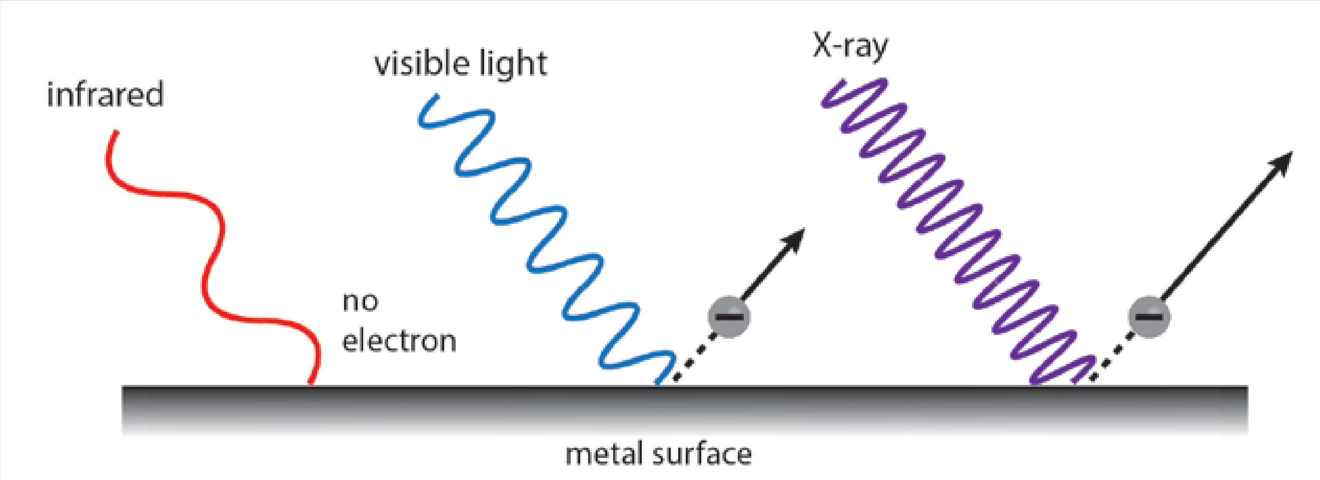


Max Planck proposed the quantum hypothesis that energy is made up of discrete units.

PHOTONS

Experiments conducted around the same time as Planck's hypothesis revealed that light itself exhibits similar quantum behavior. Heinrich Hertz had discovered in 1887 that metals illuminated with ultraviolet light tended to produce sparks. J. J. Thomson later determined that these sparks were actually electrons being ejected from the surface of the metal. This phenomenon became known as the **photoelectric effect**. Physicists initially attempted to use Maxwell's electromagnetic wave model to explain the photoelectric effect. They reasoned that since electromagnetic fields exert forces on charged particles, the oscillating fields in a light wave could theoretically push an electron within an atom back and forth as one might push a child on a swing. Eventually, the electron could have enough energy to escape the atom, and the metal, altogether.

FIGURE 41



The photoelectric effect. Light below a certain frequency does not eject electrons, whereas higher-frequency light ejects electrons with more energy. These observations cannot be explained by treating light as a wave.

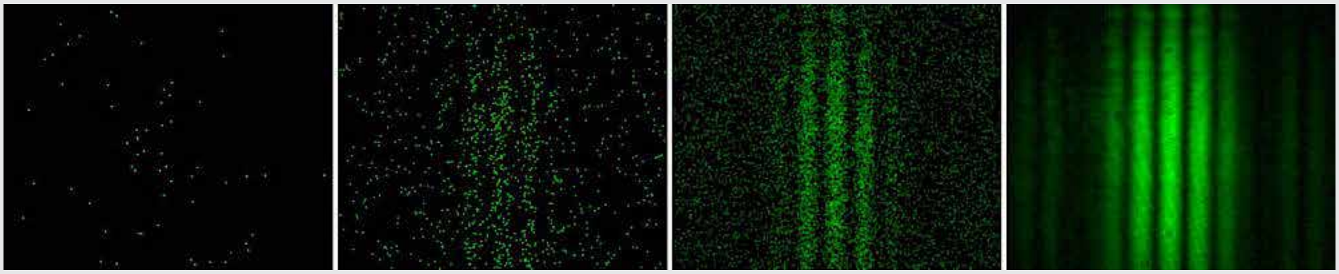
However, further investigation of the photoelectric effect in 1902 resulted in several observations that contradicted physicists' expectations under the wave model. For instance, electrons were emitted almost instantaneously (less than a nanosecond) after the metal was illuminated, whereas the wave model predicted a time delay. Furthermore, light below a certain frequency was not observed to eject any electrons, regardless of the intensity of the light. According to the existing theory, any frequency of light would be expected to eject electrons if the incident light was sufficiently intense. In addition, increasing the frequency of the incident light resulted in higher-energy electrons being emitted—however, there should have been no relationship between light frequency and electron energy when treating light as a wave (FIGURE 41).

Drawing from Planck's theory, in 1905 Albert Einstein proposed that light itself is quantized into discrete energy packets that exhibit properties of particles. Einstein argued that the photoelectric effect could only be explained as packets of light colliding with electrons, thereby transferring their energy and ejecting the electrons from the metal. These packets of light energy eventually became known as **photons**. Furthermore, Einstein argued that the energy of these photons obeys Planck's relation $E = hf$, which was consistent with experimental results from the photoelectric effect. Thus, an individual photon from a higher-frequency beam of light carries more energy than a photon from a lower-frequency beam of light. Einstein was awarded the Nobel Prize in 1921, in part for his explanation of the photoelectric effect.

Photons are packets of energy transferred from the electromagnetic field. Although they have zero mass, photons have both energy and momentum. It takes an extremely large number of photons to make up most of the radiation we experience on a daily basis because the energy of any individual photon in the visible spectrum is so small. For example, a quick estimate can be made that a single 100 watt light bulb will release on the order of 10^{21} photons every *second*. Normally we do not "feel" an individual photon any more than we would feel an individual droplet of water while swimming through the ocean.

Before we move on, we should address an important point. Although they seem contradictory, the photon, or particle, nature of light does <u>not</u> invalidate the wave nature of light we discussed in the previous section. Light exhibits properties of <u>both</u> waves and particles, depending on the experiment (FIGURE 42). This seemingly contradictory behavior is known as **wave-particle duality**. The photoelectric effect is a demonstration of the particle nature of light, whereas double-slit interference is a wave phenomenon. Wave-particle duality is a puzzling aspect of the physical world that took the world's leading physicists many years to accept.

FIGURE 42



A double slit illuminated with extremely weak laser light, so that photons effectively pass through one by one. Each tiny dot is a single photon detection. Notice how an interference pattern is gradually built up over time, illustrating the dual nature of light.

From Weis and Dimitrova.[9]

THE BOHR MODEL

In 1913, Niels Bohr proposed a model of the hydrogen atom that incorporated the quantum aspects of Planck's hypothesis. Bohr adapted the existing planetary model of the atom, in which electrons revolved around a central nucleus, but added a key assumption: Bohr suggested that the electron's energy was only allowed to take on particular values and could not be anything in between those values. Furthermore, these energy levels correspond to specific fixed orbits around the nucleus (Figure 43). A staircase serves as a useful analogy for the energy levels in an atom. Just as it's not possible for an object to reside in between two steps, an electron cannot occupy a state between two energy levels.

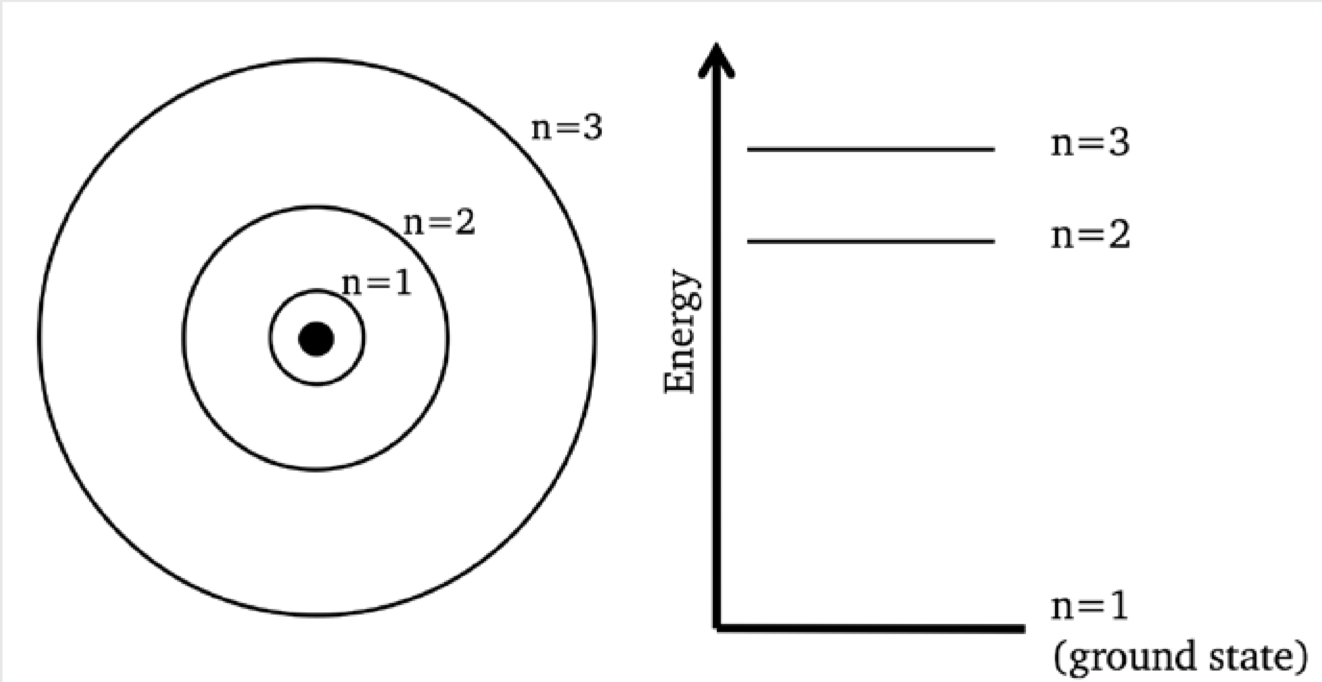
Bohr indexed the energy levels of the hydrogen atom according to a **quantum number** that could take the whole number values $n = 1, 2, 3$, and so on. The lowest allowable energy level, for which $n = 1$, is known as the **ground state** of the atom. In this state, the electron is also closest to the nucleus. Energy levels that are higher than the ground state are called **excited states**. Electrons must accept energy, for example by absorbing a photon, in order to move from one energy level to a higher one. Only certain frequencies of photons can be absorbed by an individual atomic system, and therefore each energy level diagram is quantized by the allowable transitions between states, which is unique for each type of atom.

Hydrogen is the simplest atom in the universe due to its single orbital electron. As a result, physicists often use it as the standard model to explain the basic rules of quantum mechanics and atomic energy levels. In atoms with more than one electron, the energy levels follow more complicated rules, and additional quantum numbers are required to specify the energy state of an electron. Physicists associate these quantum numbers with quantities describing specific characteristics of the electron, including the shell, the subshell, position in a shell, and spin. Fortunately, the details of these quantum states are not necessary to obtain a basic understanding of lasers, so we won't cover them here.

According to an important rule called the *Pauli exclusion principle*, each electron in an atom must occupy a unique quantum state. Equivalently, each electron in an atom must have a unique set of quantum numbers. Shells and subshells usually are identified by a number and letter, such as *1s* or *2p*. The number refers to the principal quantum number, the same as Bohr's original quantum number n. The letter identifies sub-shells formed under complex quantum-mechanical rules. Each shell can contain two or more electrons, but each electron must have a unique set of quantum numbers. In the simplest case of two-electron shells, the two electrons spin in opposite directions.

If atoms are in their ground state, electrons fill in energy levels from the lowest-energy or innermost shell until the number of electrons equals the number of protons in the nucleus. The primary quantum number of the last ground-state electron corresponds to the row the element occupies on the periodic table. This "stacking" property of electrons helps explain some of the chemical properties observed within families of elements in the periodic table. For instance, the noble gases helium, argon, neon, krypton, and xenon all have filled electron shells and therefore do

FIGURE 43



The Bohr model of the hydrogen atom includes fixed orbitals (left) with corresponding energy levels (right). Only the first three levels are shown.

not react or combine easily with other elements. In contrast, alkali metals such as lithium and sodium have an unpaired electron and tend to be highly reactive.

There are many more types of energy levels in addition to those found in solitary atoms. Molecules have electronic energy levels of their own, which become more complex as the number of electrons and atoms in the molecule increases. Molecules also have energy levels that depend on vibrations of atoms within them and on the rotation of the entire molecule. All of these energy levels are quantized as well.

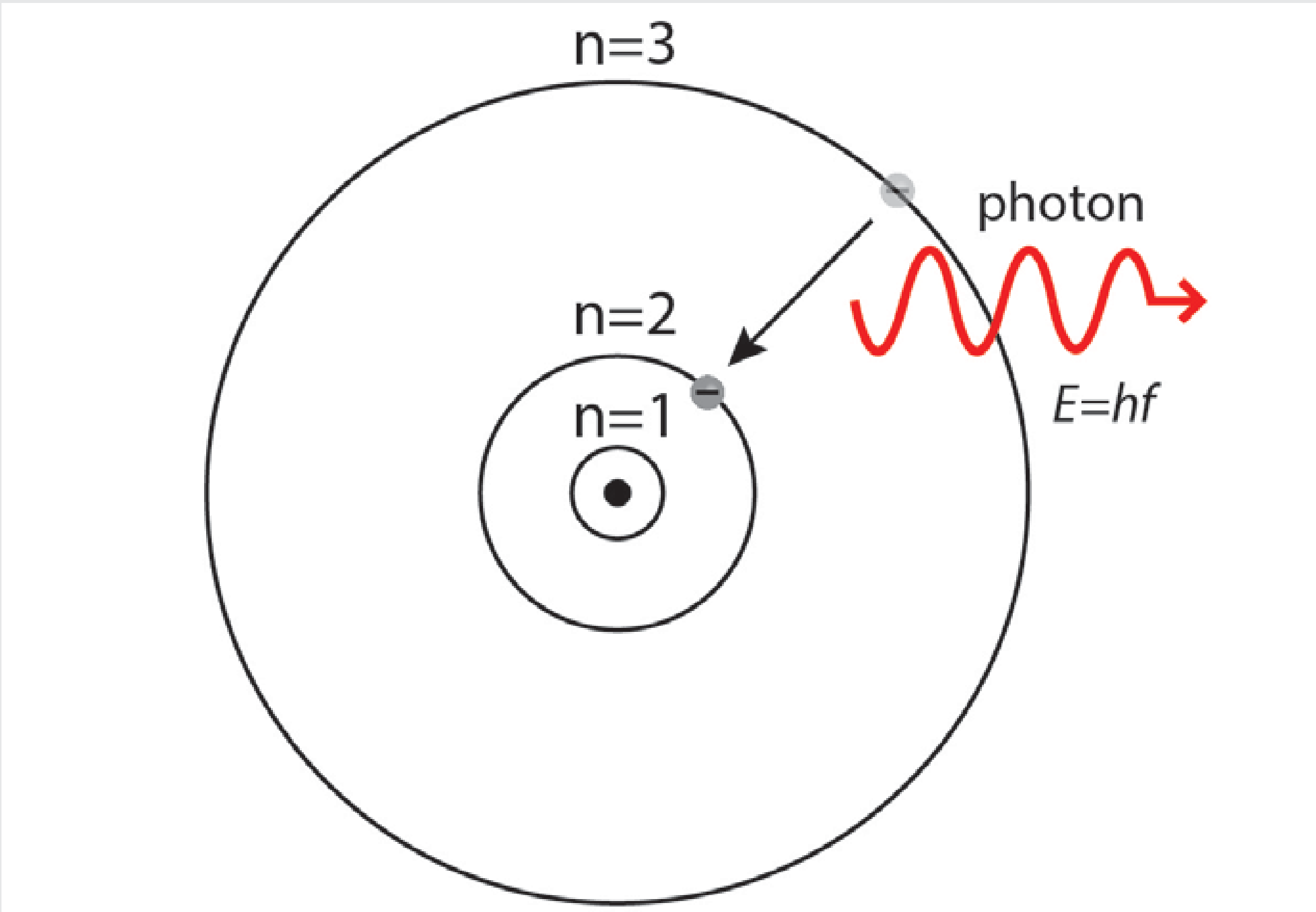
ATOMIC EXCITATION AND EMISSION

Each element has a characteristic set of energy levels that are common to every atom of that element. The process of an electron moving from a lower to a higher energy level is called **excitation**, which occurs when the atom absorbs energy corresponding to a particular energy level transition. This energy input can result from absorbing a photon, but can also come from kinetic energy due to collisions with other particles or thermal energy from being heated. Once they are in an excited state, electrons can drop to lower energy levels by emitting a photon with energy equal to the difference in energy between the two levels. The frequency of the emitted light is given by the Planck relation $E = hf$ (FIGURE 44).

A neon light is a familiar example of a **gas-discharge lamp**, which produces light as a result of atomic de-excitation (FIGURE 45). Gas-discharge lamps consist of glass tubes filled with a noble gas such as argon, neon, or xenon. The color of the lamp depends on which gas is contained within the tube. True neon lights are reddish-orange in color; other colors, such as green or blue, are produced by different gases. At each end of the tube are electrodes that are connected to a voltage source. The voltage causes electrons to leave one of the electrodes and vibrate back and forth within the tube, colliding with millions of atoms as they do so. The collisions with the electrons transfer energy to the atoms, exciting their orbital electrons to higher energy levels. The atoms emit a characteristic photon as they de-excite, and the process continues.

City street lights commonly use a form of gas-discharge lamp to produce light. Street lights containing mercury vapor produce light with intense blue and violet components, giving them a different appearance than white light from an incandescent light bulb. Many street lights

FIGURE 44



An atom emits a photon when an excited electron drops to a lower energy level. The energy of the photon is equal to the difference in energy between the two levels.

now use low-pressure sodium vapor that consumes less energy; these lights are primarily yellow in color. While walking along a city street at night, you may be able to tell what type of gas is contained within the street light based on the color of light it produces. Using a prism or diffraction grating, you can actually separate and identify the individual frequencies that each lamp emits.

Certain elements produce brilliant colors of light when ignited, which is another result of atomic de-excitation. For example, table salt produces a bright yellow flame, because a common transition between energy levels in sodium corresponds to yellow light. Likewise, copper produces a bluish-green flame, and lithium burns bright magenta red. These characteristic colors served as a basis for chemical analysis and the discovery of new elements in the mid-nineteenth century.

Aurora borealis and aurora australis, also known as the Northern and Southern Lights, are spectacular visual phenomena that can be explained by atomic excitation. The Sun emits a continual stream of charged particles that become trapped in Earth's magnetic field and spiral down through the atmosphere. These particles collide with atoms and molecules in the atmosphere, thereby exciting them to higher energy levels. Oxygen atoms emit a pale green color, nitrogen molecules produce red-violet light, and nitrogen ions emit blue-violet light. In addition to visible light, auroras also emit infrared, ultraviolet, and X-ray radiation.

FIGURE 45

Neon lights are a form of gas-discharge lamp, which produces light as a result of atomic de-excitation.

EMISSION SPECTRA

The characteristic set of wavelengths emitted by a collection of excited atoms is called an **emission spectrum**. Emission spectra can be viewed with a measurement device called a **spectroscope**, which uses a prism or grating to separate light emitted by a collection of excited atoms into component wavelengths. FIGURE 46 shows a common configuration for a spectroscope, in which the emitted light is passed through a thin slit and projected onto a viewing screen. Each component color is refracted to a definite position on the screen, forming a distinct image of the slit as a narrow line. These differently colored lines are often called spectral lines.

FIGURE 47 shows emission spectra for an incandescent source, a hydrogen lamp, and a collection of excited iron atoms. The emission spectrum for hydrogen contains four distinct lines in the visible spectrum, which occur as a result of a hydrogen atom transitioning to the $n = 2$ state from a higher energy level. This set of emission lines is known as the Balmer series. Other sets of hydrogen emission lines outside the visible spectrum include the Lyman series (which involves transitions to $n = 1$) and the Paschen series ($n = 3$).

The more electrons an atom contains, the more complicated the energy level structure becomes. The hydrogen spectral lines are relatively simple, containing only a few visible lines; this is because hydrogen contains only a single electron. For atoms with larger atomic numbers, the electrons interact with one another and with the nucleus, creating many more energy levels and thus many more possible transitions. Iron, which has an atomic number of 26, contains 26 electrons. Notice that iron has many more visible spectral lines than hydrogen due to its more complex electron configuration.

Since each element has a characteristic set of energy levels, the emission spectrum is also unique to each element. We can therefore think of these spectral lines as a sort of "fingerprint" that indicates the presence of a specific element. The process of analyzing spectral lines to identify the chemical makeup of an excited sample is called **atomic spectroscopy**. Atomic spectroscopy came into its own as a chemical analysis technique in the late 1800s.

FIGURE 46

A basic spectroscope apparatus using a prism and a narrow slit in cardboard (top). Notice the difference in spectral lines from an incandescent bulb (middle) and a fluorescent bulb (bottom).

The probability that a particular transition will occur depends on the population of the two states and their quantum characteristics. Certain types of quantum transitions are far more likely than others, so each transition to a lower energy state has its own characteristic lifetime, or average length of time for an electron to undergo a transition between those two states. As we will see, these lifetimes are an important factor in laser operation. Complex interactions between adjacent atoms in liquids or solids blur energy levels so that simple absorption and emission lines are not visible. This creates significant differences between gas and solid-state lasers.

Try It!
Make a Spectroscope

You don't need special equipment in order to observe spectral lines for yourself. You can use a spare CD and a cardboard paper towel tube to assemble your own spectroscope. Detailed instructions and pictures can be found at the Exploratorium website.[10] The CD acts as a diffraction grating to spread out the individual wavelengths of visible light received by a radiant object. Test out your spectroscope by viewing an incandescent bulb and a fluorescent light. Try observing a variety of light sources of various colors, including city street lights, neon lights, LEDs, and incandescent sources. (*Caution: You should not use the spectroscope to view light sources that would be unsafe to view with your eye alone, such as the Sun or laser light.*) What differences do you observe between the emission spectra for these light sources, and what does it tell you about how they emit light?

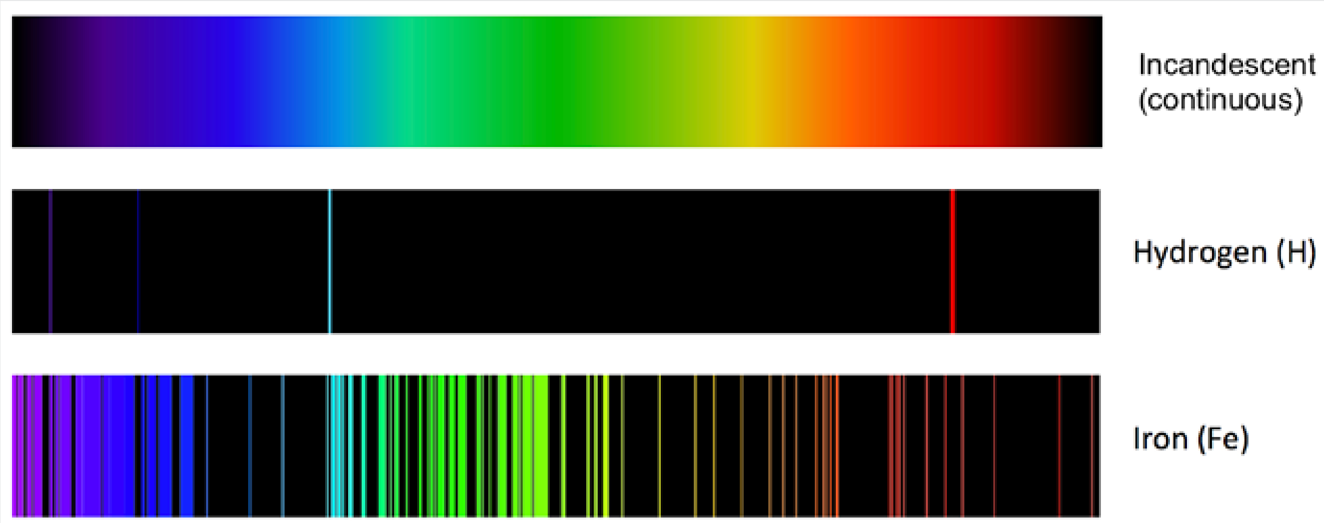
INCANDESCENCE

All matter that has a temperature above 0 K ("absolute zero") radiates electromagnetic waves, which is known as thermal radiation. At room temperature, objects radiate mostly infrared waves, which cannot be detected by our eyes but can be captured by an infrared-sensitive camera. When objects become sufficiently hot (around 800 K), they begin to glow with a visible color dependent on their temperature. Familiar examples include a red-hot stove burner or a yellow-white light bulb filament (Figure 48). **Incandescence** is the production of light due to an object being at a high temperature.

Unlike light from a neon tube, thermal radiation is emitted as a continuous range of wavelengths instead of a discrete set of spectral lines. Does this mean the tungsten atoms in a light bulb filament contain an infinite number of energy levels? No, it doesn't—if we were to vaporize the tungsten atoms and excite them, they too would emit a characteristic set of discrete spectral lines. So, what's going on?

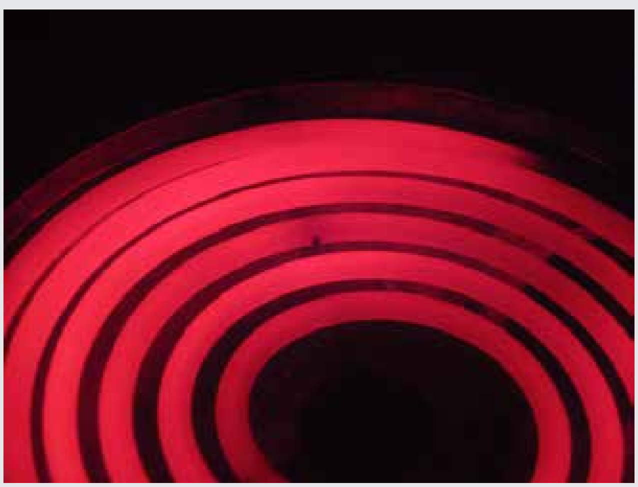
When in the gas phase, atoms are separated by large distances relative to their size and thus have a low probability of interacting with one another. The process of electron excitation and de-excitation within one atom will not be affected by the presence of other atoms within the gas.

FIGURE 47



The emission spectra for an incandescent source, hydrogen, and iron.

FIGURE 48



A hot stove coil glows red due to incandescence. A large fraction of energy is also being radiated as infrared waves, which we perceive as heat.

In contrast, atoms in the solid phase are bunched together and may even share electrons in the case of metals. Energy levels of closely-packed atoms effectively overlap, allowing electrons to undergo a much greater number of possible transitions. Consequently, atoms in the solid phase tend to emit photons with a continuous range of frequencies.

The frequencies of light emitted through incandescence are dependent on the temperature of the object. Figure 49 shows several intensity distribution curves for objects at various temperatures. Notice that the intensity of radiation is not evenly distributed across wavelengths; for any given temperature, there is a "peak" wavelength that is emitted with the greatest intensity. Furthermore, experimental evidence has shown that the peak wavelength is inversely proportional to the temperature of the object.

$$\lambda_{peak} \propto \frac{1}{T}$$

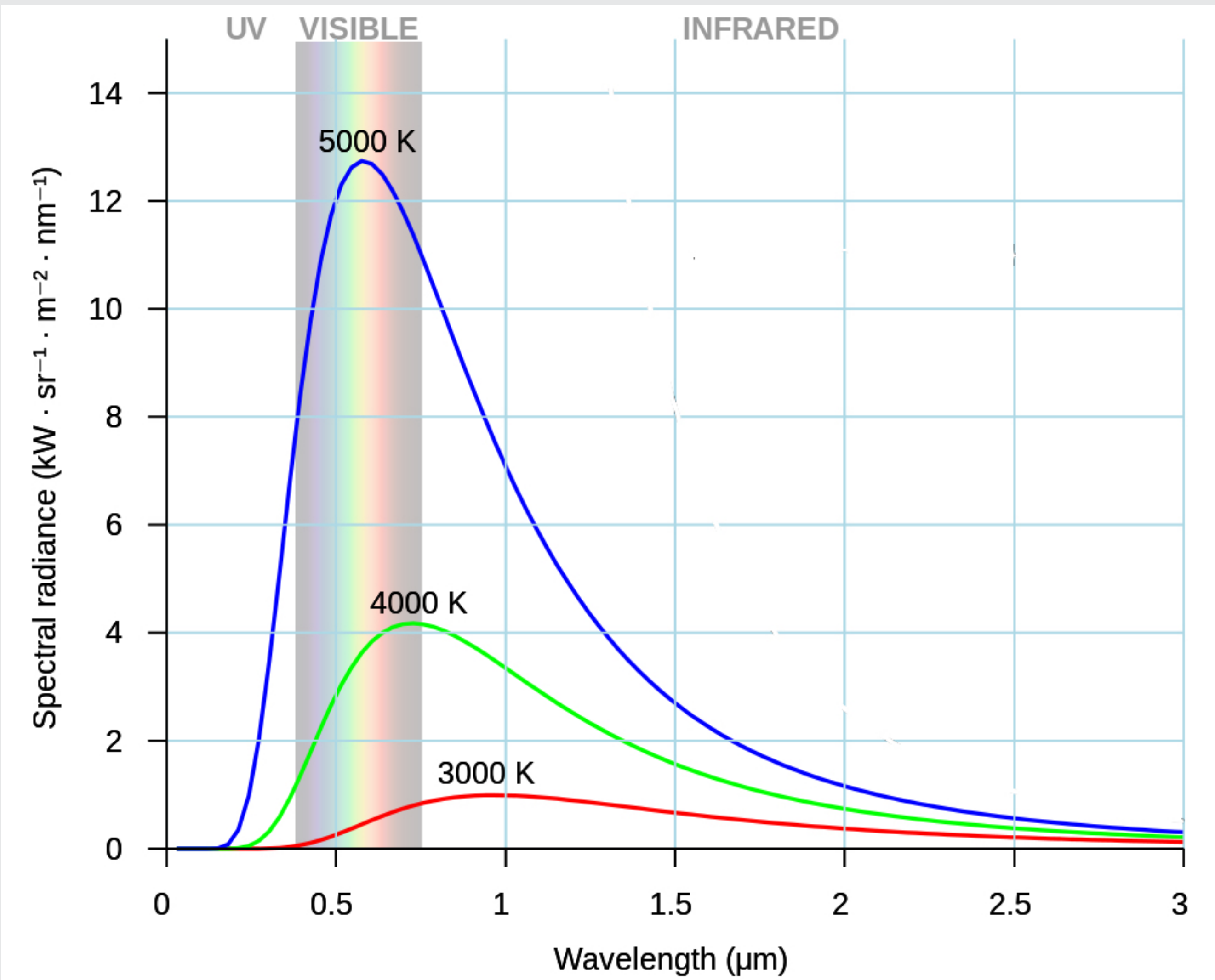
Accordingly, if the kelvin temperature of an incandescent object is doubled, the peak wavelength of emitted electromagnetic waves will be halved. This relationship allows us to easily compare the surface temperature of distant stars based on their distributions of emitted light. For instance, since the wavelength of violet light is just about half that of red light, we can estimate that a violet-hot star has approximately twice the surface temperature of a red-hot star. As an object is heated from room temperature, it first glows "red-hot" as the peak wavelength shifts into the visible spectrum from the infrared. Eventually, as the distribution peak overlaps the entire visible spectrum, the object will appear "white hot."

ABSORPTION SPECTRA

Electrons can move to higher energy levels by absorbing electromagnetic radiation. An **absorption spectrum** is created by passing a continuous spectrum of electromagnetic waves through a collection of atoms. What emerges from the atoms is a continuous spectrum except for black lines, called absorption lines, where specific wavelengths of light have been absorbed. FIGURE 50 shows a segment of the emission and absorption spectra for hydrogen. Notice that the positions of absorption lines correspond with the emission lines for the same element.

The Sun is an incandescent source of light, yet if we zoom in to the emission spectrum of sunlight we find that it is not perfectly continuous—instead, there are thin lines where certain frequencies of light are missing. These absorption lines are known as *Fraunhofer lines* after Joseph von Fraunhofer who first discovered them and determined their wavelengths. Fraunhofer lines occur because the light emitted from the body of a star is absorbed by the atmosphere of cooler gases that surround it. In 1868, scientists deduced that a set of then-unknown spectral lines from the Sun belonged to a yet undiscovered element, which they called helium. Thus, solar atomic spectroscopy revealed helium to be a new element almost three decades before it was first found on Earth.

FIGURE 49



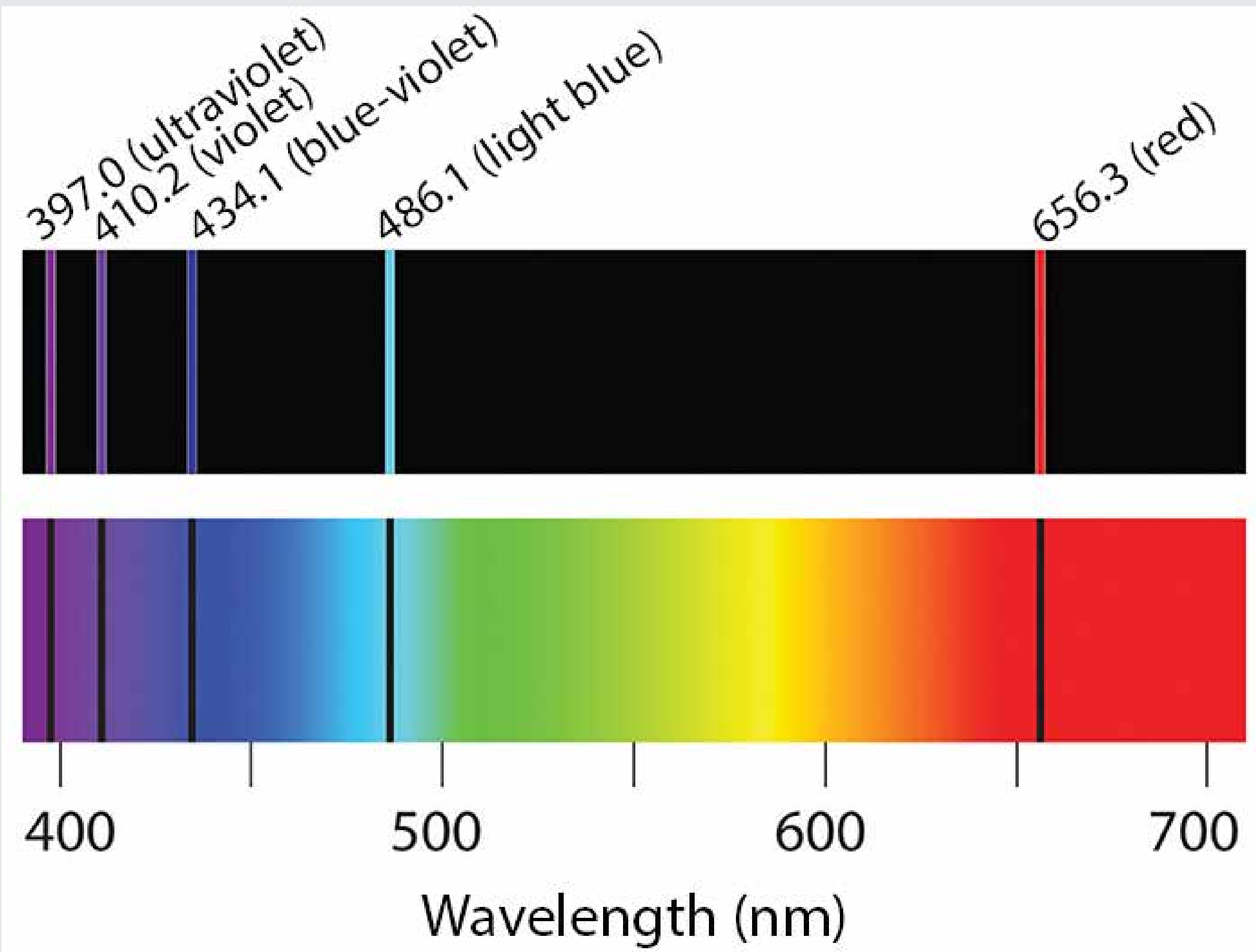
Intensity distributions of emitted electromagnetic waves for objects at 3000 K, 4000 K, and 5000 K. Notice how the peak wavelength becomes shorter as the temperature of the object increases.

Spectral lines from stars also provide evidence of how fast they are moving relative to us. If you've ever stood near a train as it passes, you may have experienced how the sound of the whistle seems to become lower in pitch at the moment the train moves past your position. Although the frequency of the whistle has not changed, the sound waves that reach your ears are shifted lower in frequency. This frequency shift due to a moving source, known as the **Doppler effect**, also applies to electromagnetic waves. Because we know the exact frequencies of spectral lines due to measurements here on Earth, we can compare them with spectral lines from distant objects to determine the relative shift. If the spectral lines are shifted toward the red side of the spectrum (i.e., lower in frequency), we know the star or galaxy is moving away from Earth. Nearly all galaxies we observe exhibit a red shift in their spectral lines, which serves as evidence that the universe is expanding.

FLUORESCENCE

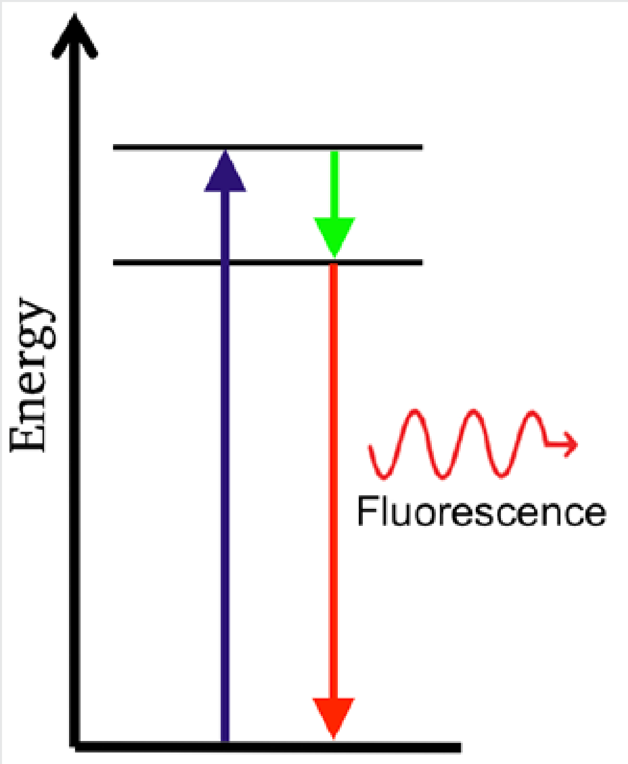
Different de-excitation pathways can cause atoms to emit different combination of photons. **Fluorescence** occurs when a material absorbs a photon and then de-excites by emitting a photon with a lower frequency (i.e., energy). Returning to the staircase analogy we explored earlier, imagine leaping to the top of a small staircase in a single jump. You could get back to the bottom by stepping down each lower step one by one, or by skipping over a step and then taking a smaller step. Likewise, an atom excited by high-frequency (and thus higher-energy) ultraviolet light can take smaller steps to a lower energy state by emitting

FIGURE 50



Emission spectrum (top) and absorption spectrum (bottom) for hydrogen.

FIGURE 51



An energy level diagram of fluorescence. After absorbing a photon, an atom emits a photon with a different frequency by dropping down two energy levels in a row.

lower-frequency (and thus less-energetic) visible light as it returns to the ground state (Figure 51). This property will become important later in explaining laser operation.

You may be familiar with fluorescent posters that glow vibrant colors when exposed to an artificial source of ultraviolet light (sometimes known as a "blacklight"). Fluorescent dyes are used in paints and fabrics to make them glow when they are bombarded with ultraviolet photons in sunlight. Fluorescence has also been introduced as an anti-counterfeiting security feature on paper currency around the world, including in the United States. If you shine ultraviolet light on a $10, $20, $50, or $100 bill printed within the past decade, you will reveal a fluorescent thread that cannot otherwise be seen with visible light alone.

An example of natural fluorescence is the brilliant response of various minerals to ultraviolet light (Figure 52). Many natural science museums have dedicated exhibits to showcase the wide array of colors produced by minerals

FIGURE 52

A collection of fluorescent minerals illuminated with ultraviolet light.

of different compositions. Each mineral has a different crystal structure made up of a particular combination of elements, giving them each a distinct set of energy levels. Like the fluorescent ink on a black light poster, the atoms in the mineral structures are excited by high-energy ultraviolet photons. As the excited electrons cascade down to lower energy levels, they emit photons with frequencies corresponding to the tiny energy-level spacings between each step. Every excited atom emits a characteristic set of frequencies, and no two different minerals emit light of exactly the same color.

PHOSPHORESCENCE

Atoms in excited states normally have lifetimes on the order of nanoseconds; that is, they undergo de-excitation within about 10^{-9} seconds of being excited. However, due to various quantum mechanical effects, certain excited states can have lifetimes on the order of microseconds or even milliseconds. Energy states that exhibit this degree of relative stability are said to be **metastable**. Metastable states exist because certain electron transitions are less probable than others under the rules of quantum mechanics. Electrons that are excited to metastable states effectively become "stuck" because the de-excitation process occurs at a relatively low rate. Materials with this property are said to exhibit **phosphorescence**.

You're already familiar with the phenomenon of phosphorescence if you've ever observed "glow in the dark" paint, toys, stickers, or light switches. Many such

FIGURE 53

This bird figurine glows in the dark due to phosphorescence.

materials can be made to emit light by first exposing them to daylight. The visible light excites the atoms within the material to a higher energy level, where they remain in a metastable state while gradually undergoing de-excitation. The persistent glow is caused by atoms emitting photons during the de-excitation process. Unlike fluorescent materials, which stop glowing immediately after the external source of radiation is turned off, phosphorescent materials can glow for seconds, minutes, or even hours after no longer being illuminated (Figure 53).

LAMPS

Incandescent Lamps

An incandescent lamp, or incandescent light bulb, consists of a glass globe enclosing a thin tungsten filament. An electric current heats the filament to a temperature between 2000 K and 3300 K, at which point it emits a continuous spectrum of light with an intensity peak in the infrared. Argon gas is commonly used to fill incandescent light bulbs; if oxygen reached the hot tungsten filament, it would undergo rapid oxidation and be destroyed. Over time, the filament gradually evaporates until it breaks and no longer conducts electricity—at this point, the bulb has "burned out." A halogen lamp is an incandescent lamp that includes a small amount of a halogen gas, such as iodine, that slows the rate of evaporation of the tungsten filament and extends the lifetime of the bulb. Halogen lamps operate at a higher temperature and are slightly more energy efficient than standard incandescent bulbs.

Although they have been ubiquitous for more than a century, incandescent bulbs are not particularly efficient sources of light. A standard incandescent light bulb converts less than 5 percent of the energy it consumes into visible light—the rest is emitted as invisible frequencies of electromagnetic waves, which we may perceive as heat. Accordingly, incandescent bulbs are in the process of being phased out around the world in favor of more energy-efficient light sources. Some countries, including Mexico, have banned them entirely.

Fluorescent Lamps and CFLs

A fluorescent lamp is a long, cylindrical glass tube filled with a low-pressure mercury vapor. Fluorescent lamps operate similarly to gas-discharge lamps. Electrodes connected to a voltage source discharge energetic electrons into the tube, where they collide with mercury atoms and excite them to a higher energy level. The excited mercury atoms emit photons primarily with ultraviolet frequencies. The emitted ultraviolet light interacts with *phosphors*, powdery materials coating the inside of the tube. Upon absorbing an ultraviolet photon, the phosphors fluoresce, thereby emitting an assortment of low-frequency photons that collectively we perceive as white light. A compact fluorescent lamp (CFL) is a small fluorescent tube lamp that is twisted into a helical shape and attached to a screw-in base like that of an incandescent bulb. CFLs are more energy efficient than incandescent bulbs and have average lifetimes around ten times longer. Different phosphors can be used to coat the inside of the tube in order to produce different shades of white light (Figure 54).

The primary disadvantage of fluorescent lamps, both compact and regular, is that they pose a significant environmental and health hazard due to their mercury content. Fluorescent bulbs must be disposed of properly so as not to release poisonous mercury into the surrounding air or water. While CFL bulbs were promoted for a few years as an energy-efficient alternative to incandescent bulbs, the lowering cost of LED (light-emitting diode) bulbs in recent years has made them a preferred candidate for safe energy-efficient lighting. In early 2016, General Electric announced that they would be phasing out CFL production in the United States in favor of LED bulbs.[11]

FIGURE 54

Comparison of an incandescent light bulb (left) and five CFL bulbs. They differ slightly in color due to different types of phosphor coating inside the tubes.

Light-Emitting Diodes

A **diode** is an electronic device that allows electric charge to pass through it in only one direction. Diodes serve many different functions in a variety of electronic devices we rely on every day. For example, diodes are commonly used in power supplies to convert alternating current (AC) to direct current (DC). A photodiode is another type of diode that produces an electric current when light is incident upon it, which is essential to the operation of solar cells. **Light-emitting diodes** (LEDs) work like photodiodes in reverse; they emit light when an electric current is passed through them.

An LED consists of a junction between two semiconductor layers. A **semiconductor** is a material that can be made to conduct electricity under some conditions but not others. One semiconductor layer (the "n-type") contains excess electrons, whereas the other layer (the "p-type") contains "holes" where electrons could be accepted. A barrier prevents the electrons from moving across the boundary. If a voltage (from a battery, for instance) is placed across the two semiconductor layers, the electrons will be energetic enough to cross the barrier and fill the holes in the adjoining layer. In doing so, the electrons lose energy in the same manner as atomic electrons dropping to a lower energy level. This energy is released in the form of a photon of visible light (Figure 55).

The color of the emitted light is determined by the depth of the energy "holes," which is a property of the elements used to create the semiconductor layers. A larger drop in energy will correspond to a higher energy photon (i.e., more blue) whereas a lower drop will correspond to a lower energy photon (i.e., more red). The first light-emitting diodes (LEDs) were developed in the 1960s and were only capable of emitting red light. In the 1990s, LEDs of many different colors became widely available.

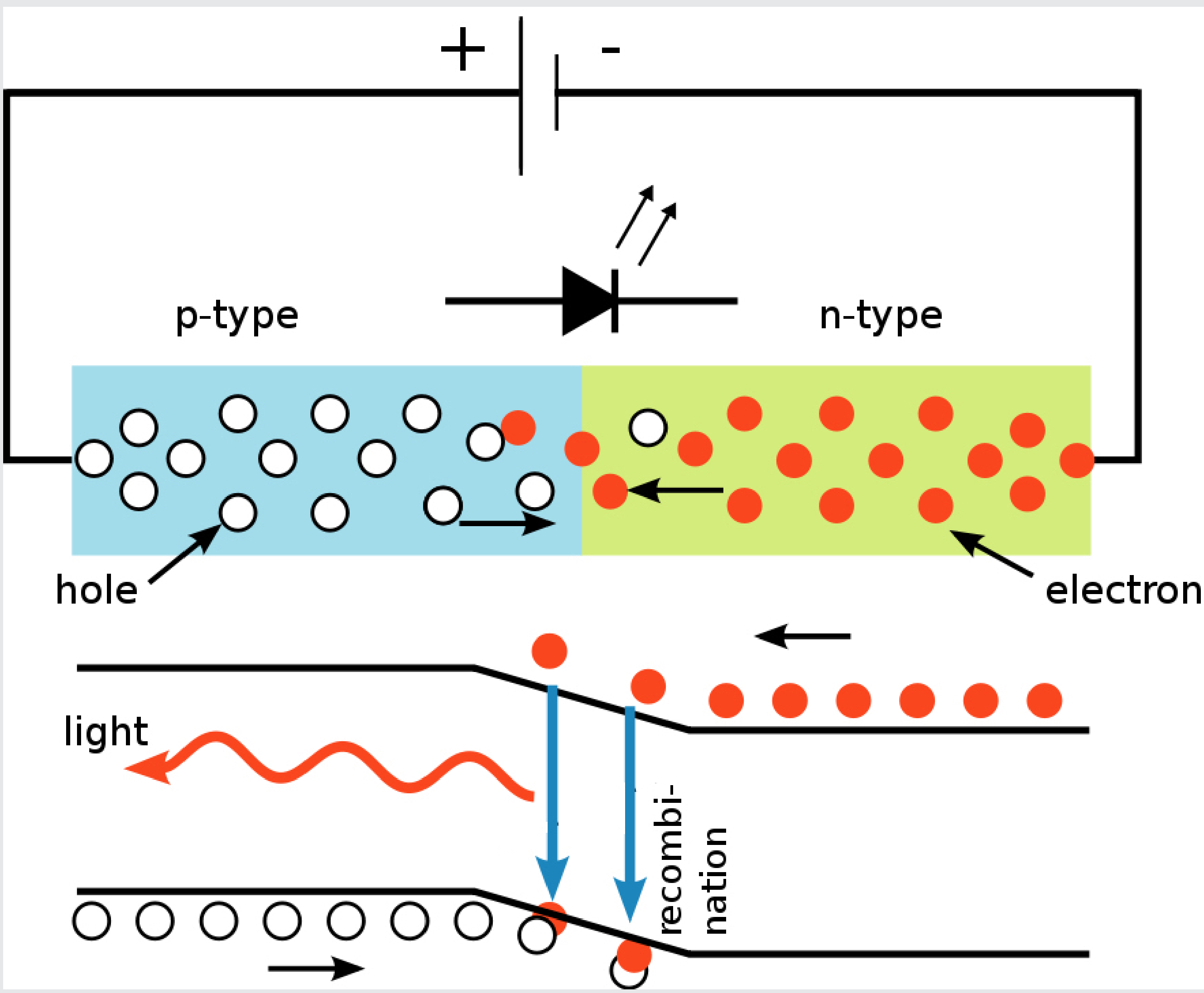
Today, LEDs are used in a wide variety of lighting applications in commercial, industrial, and residential settings. LED arrays are commonly found in traffic lights, automobile brake lights, and electronic billboards and displays. Within the last few years, LED light bulbs have supplanted CFLs as the consensus solution for energy-efficient lighting in most applications. LED bulbs have long lifetimes—approximately five times that of a CFL bulb, and forty times that of an incandescent bulb. Furthermore, LEDs consume less energy than other bulb designs and do not have disposal risks. LEDs also do not undergo abrupt failure as incandescent bulbs do; instead they very gradually decrease in brightness over many hours of use.

LASERS

LASER OPERATION

A **laser** is a device that emits a narrow beam of single-wavelength, coherent light as a result of stimulated emission. The term *laser* began as an acronym for "Light Amplification by Stimulated Emission of Radiation." The output power of laser light can vary from a few thousandths of a watt (in the case of laser pointers) to several thousand watts (industrial laser cutters). Every laser contains an

FIGURE 55



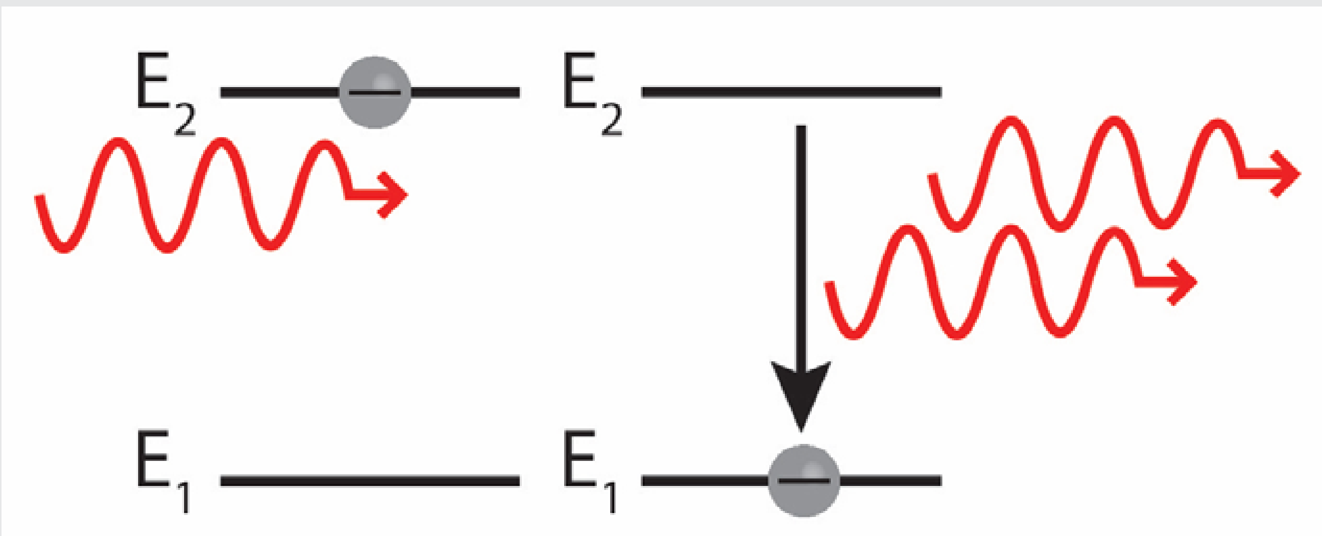
A light-emitting diode emits photons when excess electrons from an n-type semiconductor layer fill holes in a p-type semiconductor layer.

active medium, which is the source of atoms that will undergo cycles of excitation and de-excitation to release photons that will form the laser beam. Depending on the type of laser, the active medium can be a solid, liquid, or gas. Let's walk through the necessary conditions for laser operation, all of which build upon the principles of energy levels, excitation, fluorescence, and phosphorescence that we covered earlier in this section.

Stimulated Emission

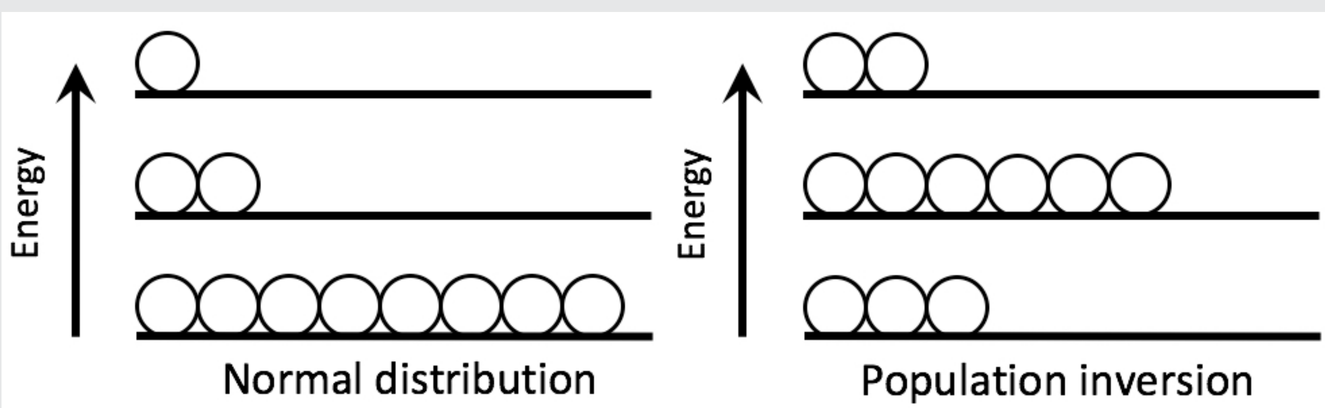
We have previously discussed how an atom can become excited from a lower to a higher energy level by absorbing a photon with energy equal to the difference between the energies of two levels. Eventually, the atom will return to a lower energy state by emitting a photon in a process known as **spontaneous emission**. Although we can predict on average how long it will take for de-excitation to occur (i.e.,

FIGURE 56



Stimulated emission occurs when a photon triggers a de-excitation event.

FIGURE 57



A population inversion exists when more atoms are in an excited state than in the ground state.

the lifetime of the state), it's impossible to predict exactly when a specific excited atom will spontaneously emit a photon. Furthermore, the direction of the emitted photon is also random.

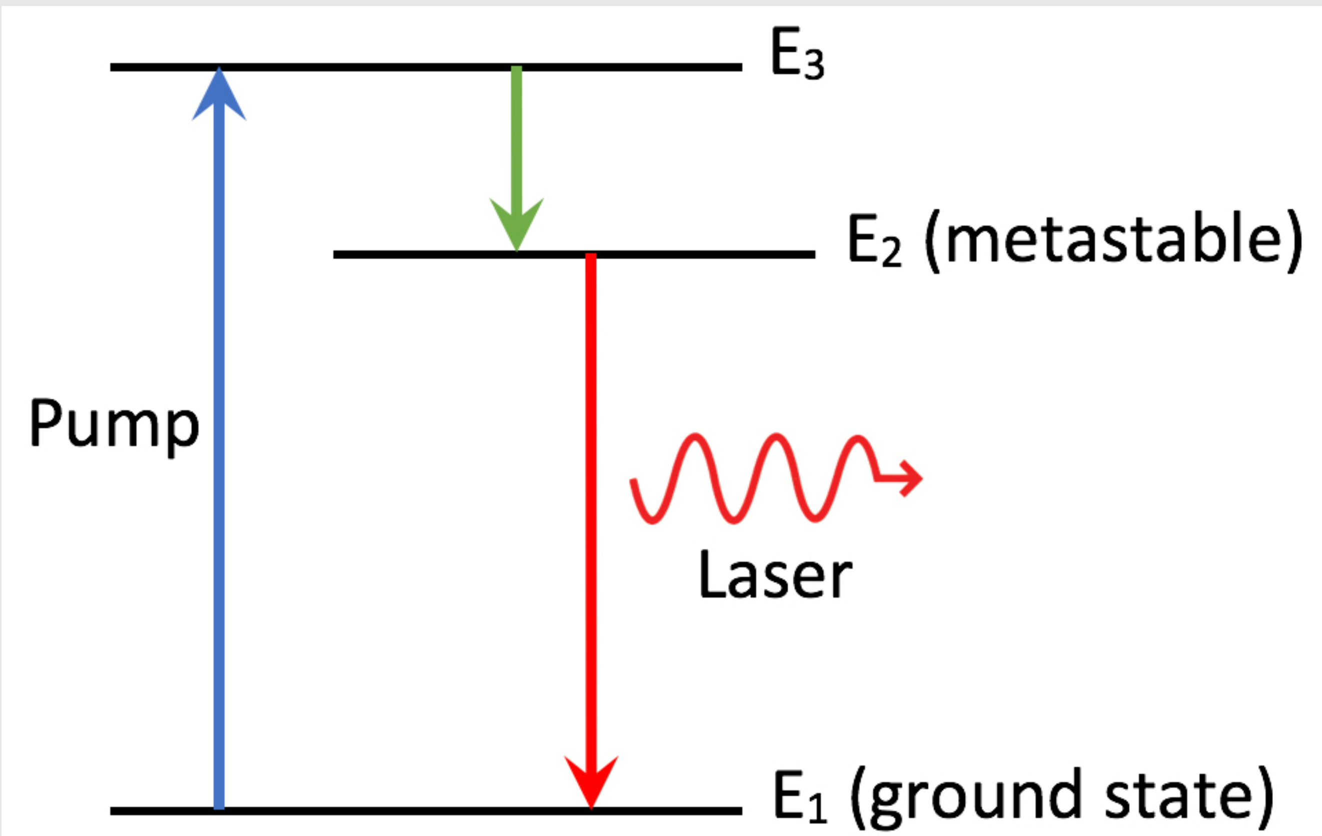
Atoms can also transition to a lower energy level through a process called **stimulated emission**. While an atom is in an excited state, the oscillating electric field from a passing photon (at the same frequency as—or very close to—the transition frequency of the excited electron) can cause the atom to emit a second photon that oscillates in precise synchrony with the first (FIGURE 56). Most significantly, the emitted photon is identical in frequency, phase, and direction to the first photon. These identical photons are what make up a laser beam.

Population Inversion

A functional laser depends on not just a handful of stimulated emission events, but rather a full cascade that will continually generate a stream of stimulated emission photons. When a collection of atoms is in thermodynamic equilibrium (that is, it is not exchanging energy with its surroundings), the vast majority of the atoms tend to be in the lowest possible energy state. This poses a problem because an emitted photon is almost certain to be absorbed by an atom in the ground state instead of stimulating emission in an excited atom.

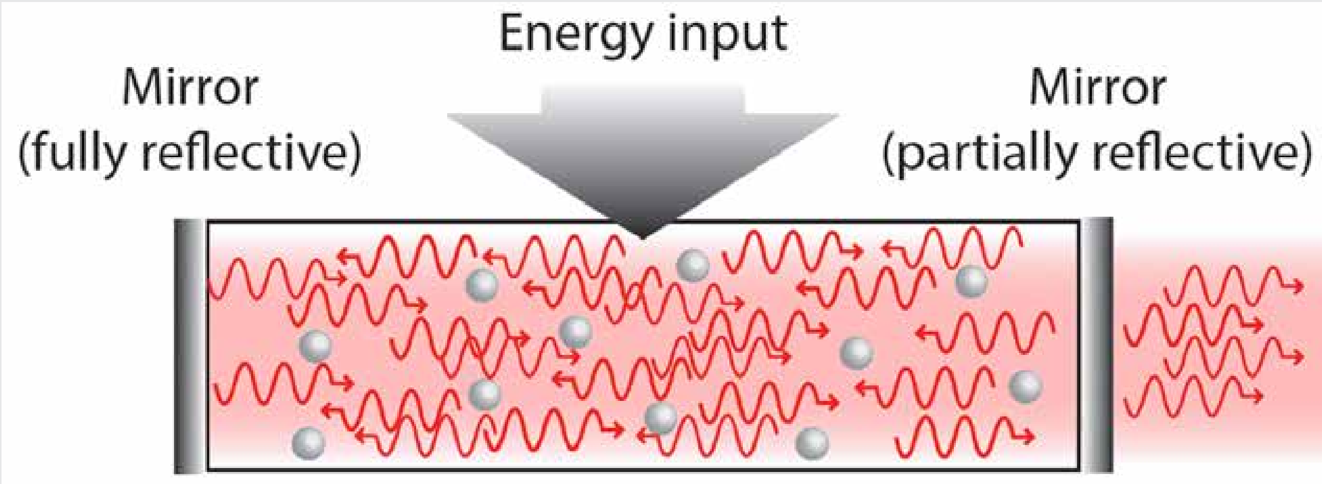
Clearly, in order to sustain the chain reaction of stimulated emission events, the atoms must be prepared such that more are in an excited state than are in the ground state. This

FIGURE 58



A simplified energy diagram of a three-level laser. The pump source excites the atoms to the highest energy level, where they are likely to drop down to the metastable state. Because the lifetime of the metastable state is relatively long, stimulated emission is likely to occur before the electrons undergo spontaneous emission.

FIGURE 59



An optical resonator is essential to amplify emitted light waves while allowing a fraction of the photons to escape, forming the output beam.

condition is known as **population inversion** (Figure 57). Stimulated emission will continue as long as population inversion exists within the active medium, but it will slow down and eventually stop if a majority of the atoms are no longer in the higher-energy state.

Energy Levels

Achieving population inversion and sustaining a chain of stimulated emission events requires a configuration of energy levels with specific characteristics. If the excited atoms undergo spontaneous emission before stimulated emission can take place, there will not be a sustained beam of identical photons traveling in the same direction. For this reason, the excited state must be metastable to ensure that atoms will remain in that state long enough to sustain a population inversion. Some amount of spontaneous emission is unavoidable, but the longer the lifetime of the state, the easier it is to ensure that stimulated emission will dominate.

Although two-level lasers are theoretically possible under certain conditions, it is generally unfeasible to excite atoms directly into a metastable state. A more practical approach involves three energy levels: a ground state, an excited state with a short lifetime (relative to the other two transitions), and a metastable state with a slightly lower energy (Figure 58). Atoms are excited, or "pumped," into the higher energy level, where they quickly de-excite to the metastable level. The metastable level is chosen to have a lifetime generally a thousand times longer than that of the higher energy level. This process accumulates a large population of atoms in the metastable state, thereby establishing a population inversion between it and the ground state.

A laser is not itself a source of energy; rather, it needs energy from an external source in order to continually maintain population inversion. An external source of energy, called the **pump source**, excites the atoms in the active medium to the metastable state. Although a variety of techniques have been used in laser pumping, the most common methods are optical pumping and electrical pumping. Optical pumping is the use of an intense light source, such as a flashlamp, arc lamp, or external laser to excite atoms through photon absorption. Electrical pumping involves the use of an electric discharge or current to cause atomic excitation.

Three-level lasers are not an optimal solution for laser operation; a more efficient method involves four energy levels. A four-level laser adds an additional energy level above the ground state, which becomes the lower level for the laser emission transition. In other words, the lower level of the laser transition is not the ground state, which makes it easier to maintain a population inversion between the laser transition states. In a four-level laser, the lower level of the laser transition starts out nearly empty because it has a higher energy than the ground state. Thus, only exciting a small fraction of the ground state atoms to the metastable state is sufficient to establish a population inversion.

Optical Cavity

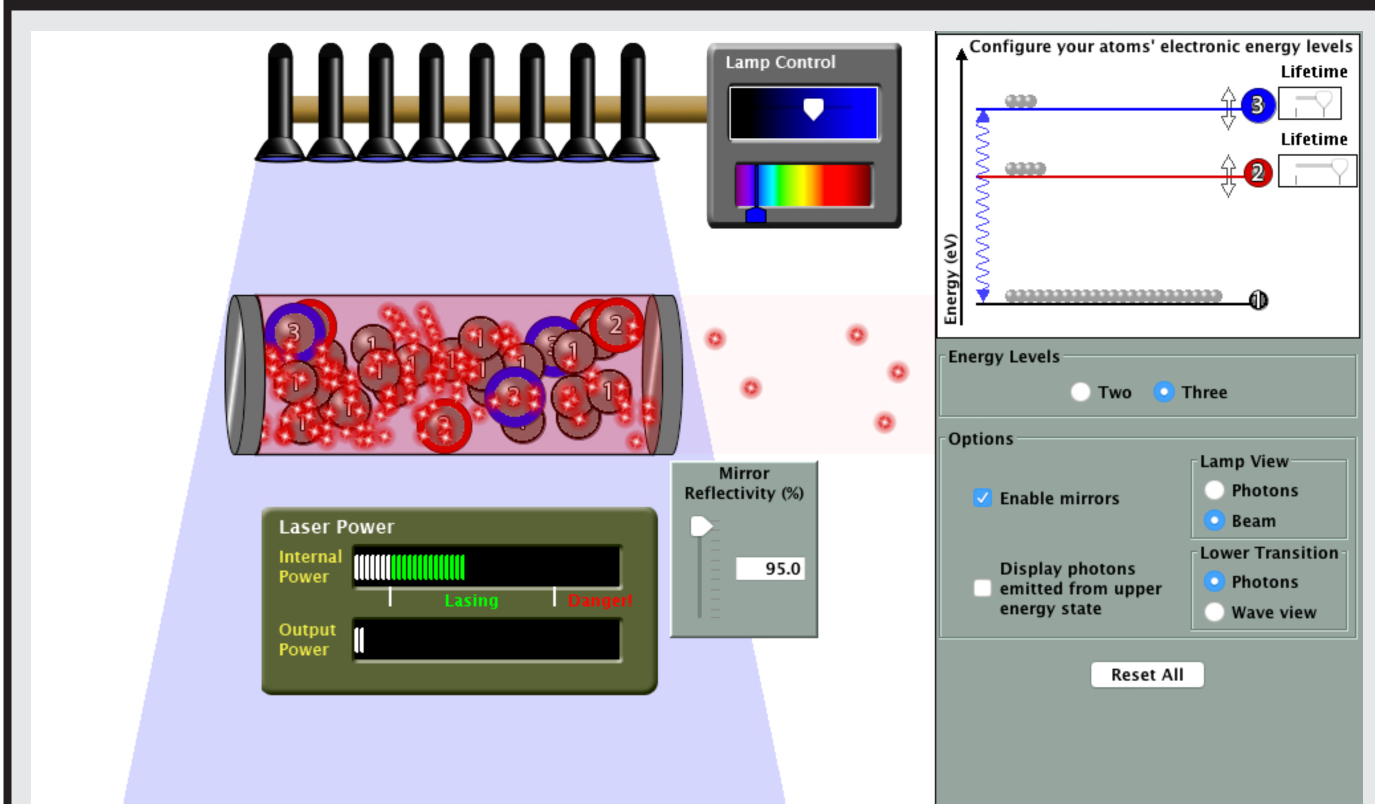
Once population inversion is achieved (i.e., most of the atoms are in an excited state), a single atom undergoing de-excitation emits a photon that causes a chain reaction of stimulated emission events: that photon causes stimulated emission in a nearby atom, which emits an identical photon, which stimulates the emission of another photon, and so on. In order to sustain the chain of stimulated emission events and amplify the laser beam, the active medium must be surrounded by an **optical cavity** (also known as a laser cavity or optical resonator).

The simplest optical cavity consists of two parallel mirrors, one of which is slightly transmitting to allow the output of the laser beam. One mirror is coated to be completely reflective (100 percent chance of reflection) whereas the other is coated to be partially reflective (~95 percent chance of reflection). Photons that "leak" from the partially reflective mirror form the laser beam. It's possible for light to reflect back and forth several hundred times before exiting the resonator (FIGURE 59).

PROPERTIES OF LASER LIGHT

Laser light has several properties that make it useful for many practical applications. Laser light is monochromatic, directional, and coherent. By comparison, ordinary white light is a combination of many wavelengths of light, emits isotropically (in all directions), and is a mixture of many out-of-phase wavelengths. These three properties of laser light are what can make it more hazardous than ordinary light—laser light is capable of depositing a lot of energy within a small area. We will discuss each of these properties in greater detail.

FIGURE 60



Now that you've got the basics of laser operation down, try building your own laser with the "Lasers" PhET app at phet.colorado.edu. You can configure the energy levels, cavity reflectivity, and wavelength. How powerful can you make your laser?

Monochromatic

Because the photons emitted by a laser all correspond to the same energy transition, they all have the same frequency. Single-frequency light such as this is often described as **monochromatic**. In comparison, recall that thermal radiation (such as that produced by an incandescent light source) produces a continuous spectrum of frequencies with different intensities.

Laser light is not perfectly monochromatic; there is some "spread" due to Doppler shifts from the motion of atoms or molecules within the active medium. The line width or bandwidth of a laser describes the spread of its spectrum of emitted frequencies. However, this spread is extremely narrow compared to the spectrum of frequencies emitted by, say, an incandescent light bulb.

Directional

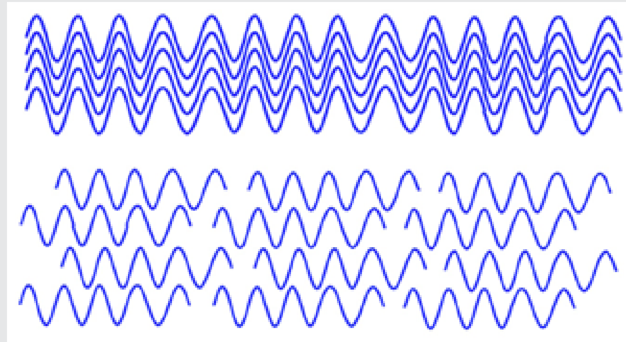
The light from a laser emerges as a very narrow beam with very little divergence, or spread. We often refer to a beam with this property as being **collimated**. If you've used a laser pointer while giving a presentation or distracting a feline companion, you are familiar with the ability of lasers to project a point of light, even from a relatively large distance. A laser's high degree of collimation is a direct result of the precise alignment of the parallel mirrors that form the optical cavity. As the light waves reflect back and forth many times within the cavity, the mirrors constrain the waves to an axis perpendicular to the surfaces of both mirrors. Any light that is slightly "off-axis" will be lost from the cavity and thus will never form part of the final beam.

The highly collimated nature of a laser beam makes it highly useful but also highly dangerous. You should never look directly into a laser beam because the highly parallel rays can focus to a nearly microscopic dot on the retina of your eye, causing almost instant damage to the retina. On the other hand, the ability of lasers to focus so precisely contributes to their wide range of both medical and industrial applications. In medicine, lasers can be used as sharp scalpels; in industry, they can serve as fast, powerful, and computer-controllable cutting tools.

Coherent

Light that is made up of waves that are "in-phase" relative to one another is said to be **coherent**. In other words, the peaks and troughs of the waves exactly align (FIGURE 61). An ordinary light source, such as an incandescent light bulb, produces light that is incoherent, meaning the

FIGURE 61



Coherent waves (top) and incoherent waves (bottom). Each individual wave has the same frequency.

waves have random phases. Even a collection of waves with identical frequencies can be incoherent if they are not in-phase relative to one another. LEDs, for example, emit light that is single-frequency but not coherent. Coherence in laser light is a direct consequence of stimulated emission.

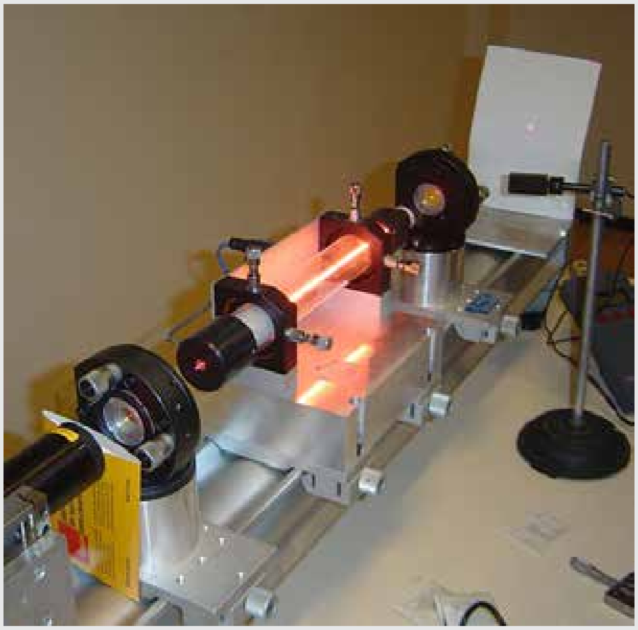
Coherence in laser light is important for observing interference effects, which has important applications in precision measurement. **Interferometry** is the use of superimposed waves to make extremely fine measurements of small displacements, surface irregularities, or changes in refractive index. A basic interferometer uses a beam splitter and mirrors to overlap beams of light from a coherent source such as a laser. By slightly adjusting the path length difference between the two beams, shifts in the observed interference pattern will occur as the relative phases of the two beams change. This technique allows measurements to be made on the length scale of the wavelength of light being used.

TYPES OF LASERS

Gas Lasers

Gas lasers use a low-pressure gas mixture as an active medium. Most gas lasers are excited by passing an electric current through the gas, delivered by electrodes placed at opposite ends of the tube. The helium-neon laser (or HeNe laser) is a common gas laser that produces light in the visible spectrum at a wavelength of 632.8 nm. HeNe laser cavities contain a mixture of helium and neon gas. The helium atoms are excited by an applied current and then collide with neon atoms to excite them to the state that causes the 632.8 nm radiation. The bright red output and relatively low cost of HeNe lasers make them well suited for

FIGURE 62



A helium-neon laser.

many low-power applications in educational and research laboratories (FIGURE 62).

Other examples of gas lasers include the carbon dioxide (CO_2) laser, which is a high-efficiency laser that operates in the infrared band. Carbon dioxide lasers are commonly used for high-power industrial applications such as welding and cutting.

Excimer lasers are a type of gas laser that rely on the excitation of "dimer" molecules, such as argon fluoride, that are stable only in the excited state. Excimer lasers were first demonstrated in the mid-1970s and are capable of removing extremely fine layers of surface material by breaking molecular bonds without burning or heating the surrounding area. For this reason, excimer lasers are well-suited for precision etching of plastics or semiconductor circuits, as well as delicate eye surgery such as LASIK.

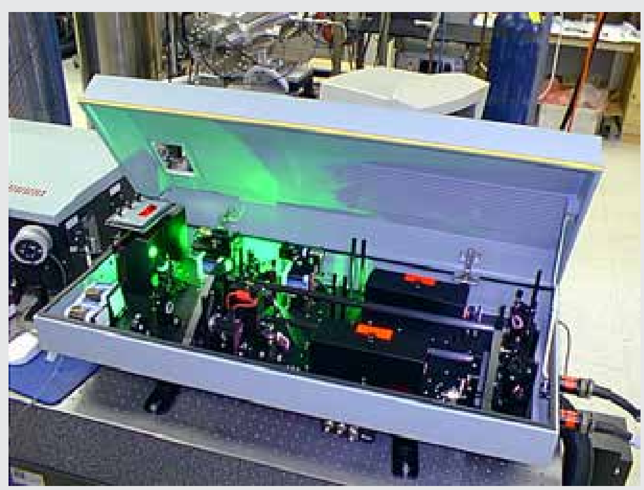
An advantage of gas lasers over other laser types is that the gas medium tends to be both relatively inexpensive and largely resistant to damage. However, gas lasers are also typically larger than other types due to the low density of the medium. In recent years, gas lasers have seen a decline in sales as they have gradually been replaced by solid-state and semiconductor lasers for many commercial applications. For example, HeNe lasers were originally used in grocery store checkout scanners, but have largely been replaced by laser diodes for this purpose.

Solid-State Lasers

Solid-state lasers use an active medium consisting of a solid crystalline or glass rod (known as the *host*) containing light-emitting atoms (the *active species*). The first laser ever built was a solid-state laser using synthetic ruby, which is corundum (aluminum oxide crystal) with chromium as an active species. In most solid-state laser materials, the active species is identified first (typically by its chemical symbol), followed by the host material. For example, the Ti:sapphire laser consists of titanium atoms in sapphire crystal. Both the active species and the host are important in solid-state lasers. The active species determines the laser transition, but its interactions with the host may shift the wavelength slightly.

Neodymium (Nd) is commonly used as an active species in solid-state laser crystals. For example, the Nd:YAG (neodymium-yttrium aluminum garnet) is one of the most common types of laser, with applications in research, medicine, manufacturing, and other fields (FIGURE 63). Nd:YAG lasers typically emit infrared light with a wavelength of 1064 nm although other wavelengths are possible. Other host materials for neodymium include YLF (yttrium lithium fluoride) and glass. The host material is selected based on its optical, thermal, and mechanical properties.

FIGURE 63



A frequency-doubled 1064 nm Nd:YAG laser.

Semiconductor Diode Lasers

Semiconductor diode lasers, commonly known as laser diodes, operate using the same basic principles as light-emitting diodes, but with some important differences. Like an LED, a laser diode consists of two semiconductor layers, an n-type with an excess of electrons and a p-type with electron "holes" to be filled. The semiconductor layers are separated by a microscopic region called the *active layer* that serves as the optical resonator. Laser diodes operate at much higher currents than LEDs, typically around ten times greater. Whereas an LED emits photons in all directions from its junction layer, a laser diode is configured with reflective ends to form an optical resonator in the region between the semiconductor layers. In a laser diode, stimulated emission occurs when a photon emitted by one electron transition triggers another electron to fill a hole, and so on, resulting in a coherent beam of light that emerges from one side of the diode.

Laser diodes are compact and easy to mass-produce. In terms of sheer numbers, they are the most common type of laser. Their small size makes them well suited for use in low-power applications such as laser pointers, laser printers, and CD/DVD players (Figure 64). Laser diodes can also be operated at lower voltages than other types of lasers. While gas and solid-state lasers require input voltages on the order of kilovolts, laser diodes can be operated at only a few volts.

Laser diodes are typically not as collimated as beams from other types of lasers. In many cases, an external lens is used to correct the shape of the beam, which contributes to the overall fragility of the laser since damage to the lens could render it non-functional. Furthermore, the delicate nature of the semiconductors makes laser diodes more sensitive to static discharges and currents. Excess electrical current can cause the diode to become inoperable. Laser diodes can also degrade in power efficiency over time, gradually requiring more power to output the same beam intensity.

FIGURE 64

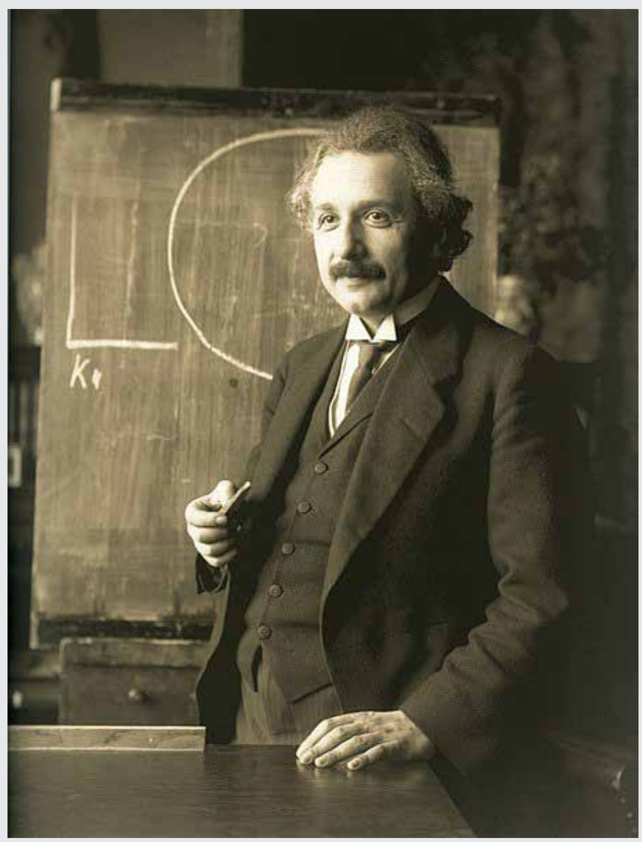
A laser diode, with a penny for scale.

HISTORY OF THE LASER

1900–50: EARLY FOUNDATIONS

The origins of the laser can be traced back to an idea from Albert Einstein in the formative years of quantum theory. Having emerged as an early visionary in twentieth-century physics with his 1905 explanation of the photoelectric effect, Einstein continued to make contributions to the theory of light emission (Figure 65). In 1917, Einstein published a paper entitled *Zur Quantentheorie der Strahlung* ("On the Quantum Theory of Radiation") in which he proposed the concept of stimulated emission.

FIGURE 65



Albert Einstein originally proposed the concept of stimulated emission.

In 1947, Willis Lamb and R.C. Retherford were the first to experimentally demonstrate stimulated emission. In 1950, French physicist Alfred Kastler proposed the technique of optical pumping, whereby a collection of atoms could be excited to a higher energy state through absorption of incident light. This idea would later have tremendous implications for laser operation, and Kastler was awarded the 1966 Nobel Prize in Physics for this contribution.

EARLY 1950s: CHARLES TOWNES AND THE MASER

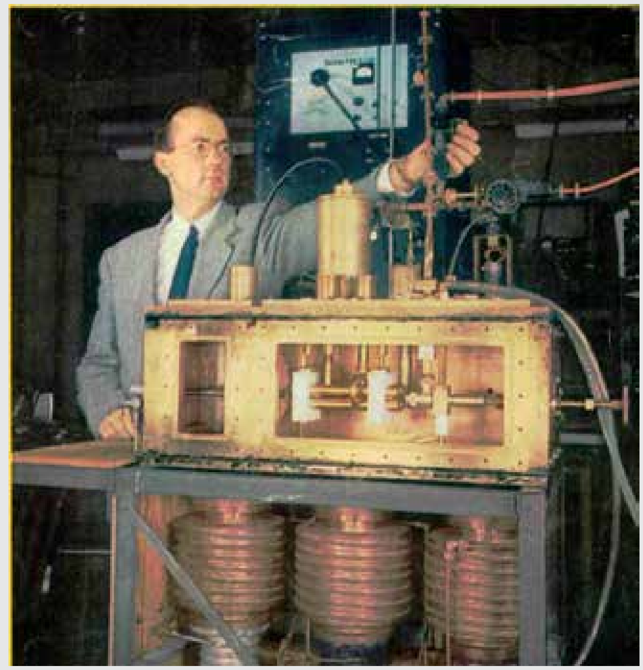
By 1950, stimulated emission seemed to be an interesting phenomenon with few practical applications due to the difficulty of achieving the necessary population inversion. That would change following a breakthrough from Charles H. Townes. During World War II, Townes had worked to develop radar systems that operated at microwave frequencies. After the war, Townes applied his expertise in microwaves to the field of molecular spectroscopy, using microwaves to study the structure of molecules and atoms. Townes joined the faculty at Columbia University in 1948 and became a professor in 1950.

In 1951, Townes was sitting on a park bench in Washington, D.C., when he devised a method for producing stimulated emission at microwave frequencies. Townes's idea was to a pass a beam of excited ammonia molecules through a reflective cavity, where they would emit microwaves. The cavity would contain and amplify the waves as they were emitted, thereby establishing a stimulated emission condition. A small hole in one side of the cavity would permit the coherent beam of microwaves to escape. Townes dubbed his device the "maser," which stood for "Microwave Amplification by Stimulated Emission of Radiation."

Eminent physicists of the time, including Niels Bohr and Isidor Rabi, expressed skepticism that Townes' maser idea was practical or even possible.[12] Nevertheless, Townes began work on the maser with two graduate students at Columbia named James P. Gordon and Herbert J. Zeiger. In 1954, Townes and his team demonstrated the first working maser using excited ammonia molecules in accordance with Townes' original vision. Townes' maser produced about 10 nanowatts of power at a wavelength around 1 cm (Figure 66).

While Townes was developing the maser, Soviet physicists Nikolay Basov and Aleksandr Prokhorov were trying to model stimulated emission in quantum systems with more than two energy levels. In 1955, Basov and Prokhorov suggested that optical pumping could be used to create a population inversion in a three-level quantum system. Basov and Prokhorov's three-level model correctly anticipated the configuration that would ultimately be used in the first laser.

FIGURE 66



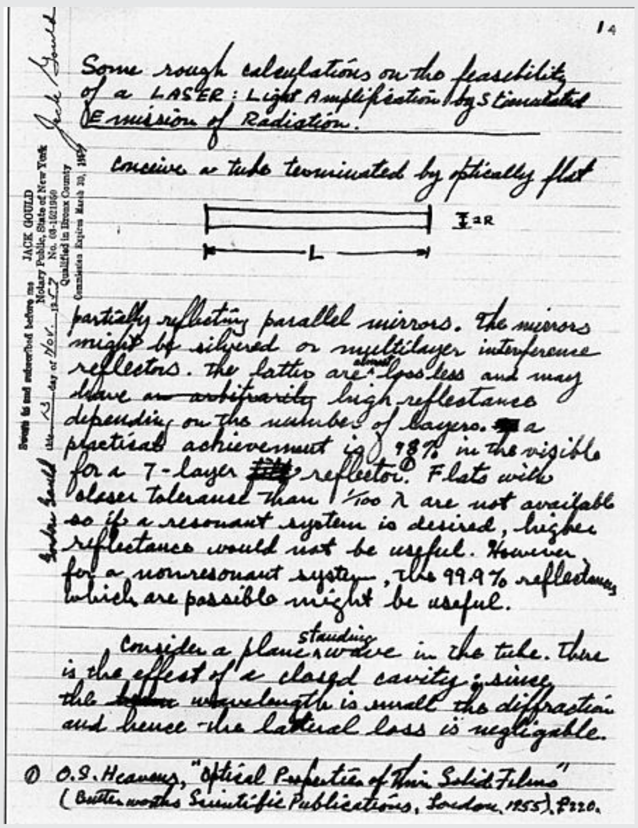
Charles Townes standing next to the ammonia maser, the precursor to the laser.

LATE 1950s: LASER CONCEIVED

Once the maser had been firmly established, a natural question emerged: could the same principle of stimulated emission be extended to visible wavelengths of light? In 1957, Charles Townes and Arthur Schawlow explored the possibility of infrared light amplifiers but eventually moved on to visible wavelengths, which they called "optical masers." Townes partnered with a Columbia graduate student named Gordon Gould who was interested in optical pumping. On November 13, 1957, Gould wrote an entry in his lab notebook wherein he used the term "laser," which is believed to be the first usage of the term (Figure 67).

As the possibility of developing a working laser seemed closer than ever, many scientists around the world recognized the groundbreaking possibilities presented by the device. Gordon Gould filed a patent application in April 1959, but it was rejected by the U.S. Patent Office. Instead, the patent was awarded to Schawlow and Townes on behalf of Bell Labs, who had applied nine

FIGURE 67


14

Some rough calculations on the feasibility of a LASER: Light Amplification by Stimulated Emission of Radiation.

Conceive a tube terminated by optically flat

L

partially reflecting parallel mirrors. The mirrors might be silvered or multilayer interference reflectors. The latter are almost lossless and may have an arbitrarily high reflectance depending on the number of layers. A practical achievement is 98% in the visible for a 7-layer reflector. Flats with closer tolerance than 1/100 λ are not available so if a resonant system is desired, higher reflectance would not be useful. However for a nonresonant system, the 99.9% reflectance which are possible might be useful.

Consider a plane standing wave in the tube. There is the effect of a closed cavity; since the wavelength is small the diffraction and hence the lateral loss is negligible.

① O.S. Heavens, "Optical Properties of Thin Solid Films" (Butterworths Scientific Publications, London, 1955) $220.

Gordon Gould's 1957 notebook entry in which he coined the term "LASER."

months earlier. Gould sued, and after thirty years of legal proceedings he was awarded forty-eight patents, including optical pumping and a variety of laser applications. The development of the laser, which was a product of many theoretical and experimental breakthroughs over the course of several decades, serves as an example of how difficult it can be to assign credit for scientific innovations.

1960s: THE RACE TO BUILD THE LASER

Following a 1958 publication by Townes and Schawlow confirming that masers operating at visible wavelengths were a theoretical possibility, all that remained was for a research team to demonstrate the first working version. Several teams across the United States joined the race to construct the first laser. At Columbia, Townes began work on a potassium gas laser, since its energy levels were well understood. Arthur Schawlow explored the use of ruby crystals at Bell Laboratories, since they were in ample supply. Irwin Wieder at Westinghouse Research Laboratories also tried using ruby, but faced challenges with generating enough energy to achieve population inversion. Ali Javan, a former doctoral student of Townes at Columbia, attempted a gas laser using a helium-neon mixture.

On May 16, 1960, Theodore Maiman succeeded in demonstrating the first functional laser at Hughes Research Laboratories in California. Maiman's laser used a synthetic ruby crystal the size of a fingertip to produce 694 nm red laser light. Although other research teams had also been experimenting with ruby as an active medium, Maiman's innovation was to use a flashlamp to deliver an intense optical pulse that would excite a sufficient number of atoms (FIGURE 68). Although Maiman's laser could only operate in brief, intense pulses, it was the first demonstration of laser operation.

Maiman's discovery set off a flurry of activity among the other research teams. Careful analysis of the technique led to the development of several other laser types less than a year later. Shortly after Maiman's demonstration, Arthur Schawlow succeeded in producing laser light using a different kind of ruby crystal. In December 1960, Ali Javan, and William R. Bennett, Jr., and Donald R. Herriott succeeded in demonstrating the first gas laser, the HeNe, at Bell Labs. In 1962, Robert Hall demonstrated the first diode laser using gallium arsenide (GaAs) as the semiconductor material. Other laser types would debut in the 1960s, including the CO_2 and Nd:YAG lasers in 1964; and the tunable dye laser in 1967.

The 1964 Nobel Prize in Physics was awarded jointly to Charles Townes, Nikolay Basov, and Aleksandr Prokhorov "for fundamental work in the field of quantum electronics, which has led to the construction of oscillators and amplifiers based on the maser-laser principle."[13] Half was awarded to Townes; the remaining half was shared by Basov and Prokhorov.

1970s–PRESENT

During the 1970s, laser technology began to enter our daily lives with a steady stream of practical applications. On June 26, 1974, a pack of chewing gum became the first product ever scanned by a laser barcode reader. In 1978, LaserDisc technology, which used HeNe lasers to read optical discs, became commercially available for the first time. Although LaserDiscs did not succeed commercially in the long-run, the same technology led to the introduction of audio CD players in 1982. The first commercial laser printer was released by IBM in 1976, and in 1984 laser printing entered mainstream

FIGURE 68

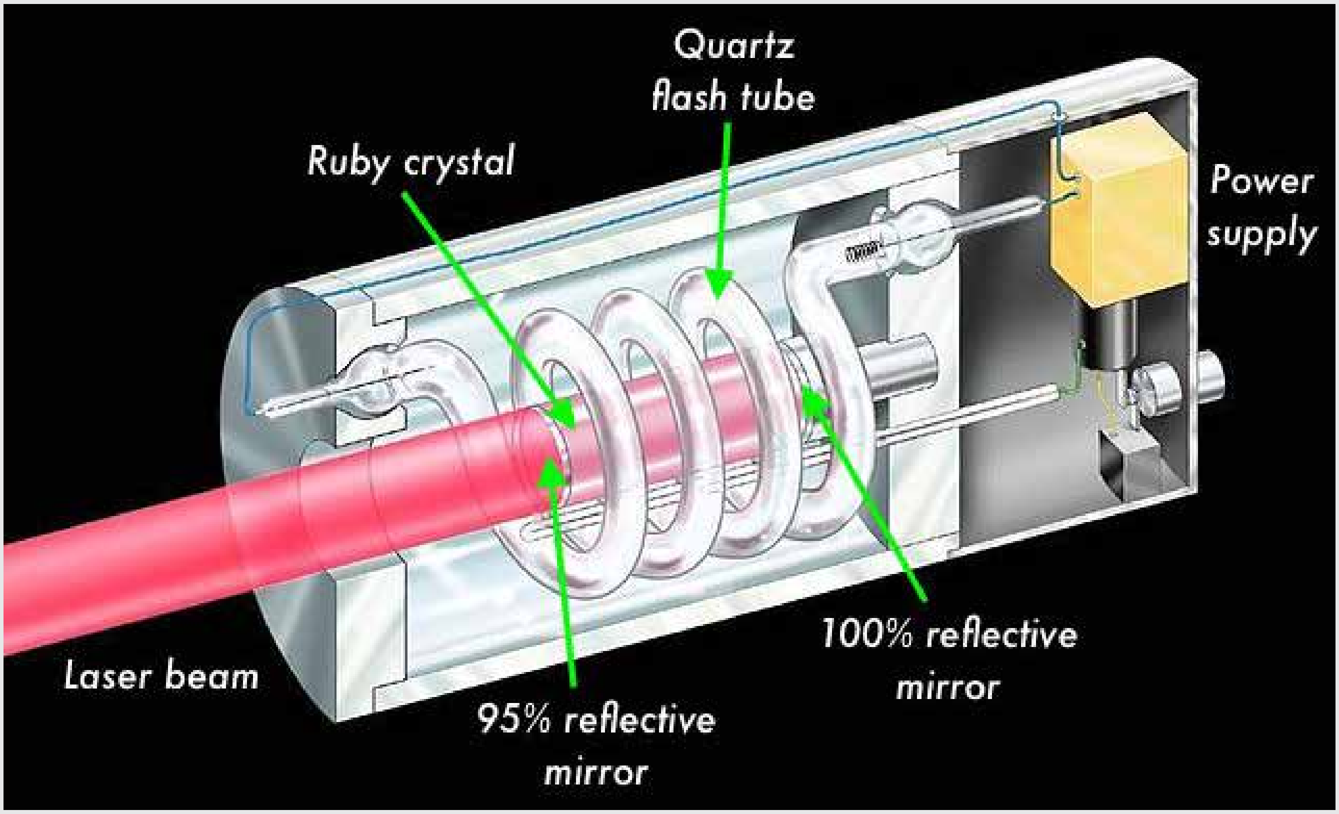


Diagram of a ruby laser.

availability with Hewlett-Packard's introduction of the desktop-size HP Laserjet. In the next section, we will cover a much wider assortment of applications of laser technology.

SECTION III SUMMARY

- Atoms are the building blocks of the world around us. Each element has a characteristic atomic number, which is equal to the number of protons in the nucleus of an atom of that element. Neutral atoms contain equal numbers of protons and electrons.
- Lasers rely on the quantum nature of light and matter. In 1900, Max Planck proposed that energy can only be exchanged in discrete bundles that are multiples of a fundamental unit, or quantum, of energy.
- In 1905, Albert Einstein proposed that light itself is quantized into discrete energy packets that exhibit properties of particles. These packets eventually became known as photons. The energy of a photon is directly proportional to its frequency.
- Electrons occupy quantum states within atoms, corresponding to energy levels that are quantized. The lowest energy level is known as the ground state; higher energy levels are called excited states.
- When atoms become excited, electrons move to higher energy levels. An electron can transition to a lower energy level by emitting a photon with energy equal to the difference in energy between the two levels.
- The characteristic set of wavelengths emitted by a collection of excited atoms is called an emission spectrum. Since every element has a unique electron configuration, the collection of wavelengths emitted by an element is also unique.
- Incandescence is the production of light due to an object being at a high temperature. Incandescent objects emit a continuous spectrum of wavelengths.

The peak wavelength depends on the object's temperature.

- Fluorescence occurs when a material absorbs a photon and then de-excites by emitting a photon with a lower frequency. Phosphorescence is a related process that takes place over a longer time scale because the de-excitation occurs from a metastable state, which has a longer lifetime.
- Incandescent light bulbs, fluorescent lights, and light-emitting diodes all rely on different mechanisms to produce light. As a result, they have different energy efficiencies and emit light with different properties.
- Lasers use stimulated emission of atoms to produce light that is monochromatic, directional, and coherent. In order to sustain a chain of stimulated emission events, a population inversion must be maintained between a metastable state and a lower-energy state.
- Although they all rely on stimulated emission and optical resonators, gas lasers, solid-state lasers, and laser diodes all use different types of materials as active media. Consequently, each type of laser has advantages and disadvantages that make it suitable for a different set of applications.
- In 1954, Charles Townes invented the maser, the precursor to the laser that operated at microwave frequencies. Theodore Maiman developed the first working laser, a solid-state ruby laser, in 1960.

Applications of Lasers

INTRODUCTION

Lasers are used for a variety of applications in scientific, medical, industrial, and commercial fields. In this section, we will describe some of the most significant applications of lasers in these categories, as well as some of the scientific figures who developed these applications. As we will see, the output power and wavelength of a laser determines the kind of applications for which it is most appropriate.

SCIENTIFIC APPLICATIONS

LASER SPECTROSCOPY

In the previous section, we described how elements can be identified by their characteristic absorption and emission spectra; this analysis technique is known as spectroscopy. **Laser spectroscopy** is a branch of spectroscopy in which a laser is used to illuminate the sample being studied in order to determine a precise absorption spectrum. The 1981 Nobel Prize in Physics was awarded in part to Nicolaas Bloembergen and Arthur Schawlow for the development of laser spectroscopy. Due to their high degree of monochromaticity, lasers output a much narrower band of wavelengths than conventional light sources, allowing researchers greater precision in identifying the precise wavelengths that have been absorbed. For example, an ordinary light source could show that absorption occurs around 750 nm for a given sample. A laser, on the other hand, could show that light was in fact absorbed at two wavelengths in that range: 749.87 and 750.13 nm. Furthermore, lasers can also identify absorption lines too weak to see using other techniques.

Laser absorption spectroscopy is most often performed using tunable lasers, which can be "swept" across a particular range of emitted wavelengths. For example, an absorption experiment might involve a titanium-sapphire laser that sweeps from 740 to 770 nm once every thirty seconds (Figure 69). If the wavelength varies uniformly with time, the precise absorption wavelength can be determined by timing when the sensor measured absorption. In the described configuration involving a Ti:Sapphire laser with a 30-second scan at a uniform rate, a detection of absorption 20.37 seconds after the start of the sweep would indicate an absorption wavelength of 760.37 nm. Tunable semiconductor lasers can be swept across that wavelength range even faster.

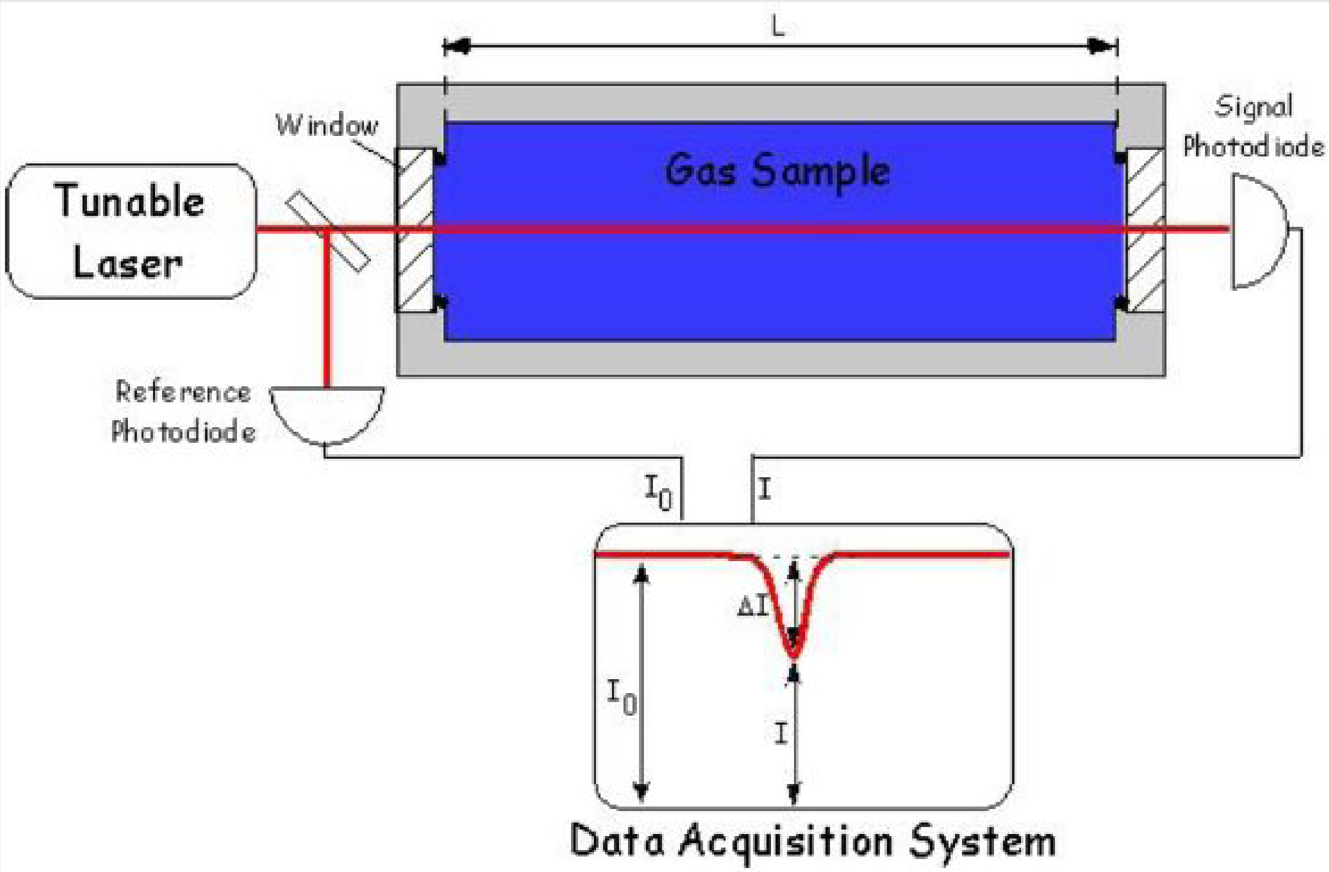
Infrared absorption spectroscopy, involving wavelengths of a few micrometers, is a reliable technique for identifying organic compounds. This is because molecules are commonly identified by their patterns of vibration and rotation, and the frequencies of infrared light match the vibrational and rotational frequencies of most molecules, resulting in absorption. Laser sources are available in the infrared region, but non-laser sources are easier and less costly to use when the extremely high resolution of laser sources is not essential.

CONFOCAL LASER SCANNING MICROSCOPY

Lasers can also be used as part of microscopes, although the manner in which an image is formed is very different from a conventional microscope. **Confocal laser scanning microscopy**, or CLSM, is a technique that reconstructs a three-dimensional rendition of the specimen. First, a laser beam is focused onto a single point on a two-dimensional "slice" of the sample to be imaged. Mirrors are then used to scan the laser back and forth everywhere along that two-dimensional plane. As the laser scans, molecules illuminated by the laser light undergo fluorescence and emit light that is refocused through a pinhole aperture to be collected by a sensor. An amplifier is then used to convert the intensity of the emitted light into an electrical signal (Figure 70). The information collected about each section of the object is then used to create the three-dimensional model of the specimen.

The basic principle of confocal scanning microscopy was first developed by Marvin Minsky, a Harvard postdoctoral

FIGURE 69



A schematic illustrating laser absorption spectroscopy with a tunable laser. I_0 is the reference intensity of the laser, and I is the signal intensity after passing through the sample. The signal intensity is reduced by an amount ΔI relative to the reference, indicating absorption at that wavelength.

student, in the mid-1950s. However, Minsky's technique initially had limited practical utility because it required a high-intensity light source and substantial information processing power. The development of the laser in the next decade brought renewed interest in confocal scanning microscopy. A functioning confocal laser scanning microscope was first demonstrated in 1969 by M. David Egger and Paul Davidovits at Yale. Egger would go on to publish the first recognizable images of cells in 1973. Advances in both computer and laser technology in the following decades, coupled with new algorithms for digital manipulation of images, led to a growing interest in confocal microscopy. The first commercially available instruments utilizing this technique began to appear in 1987.

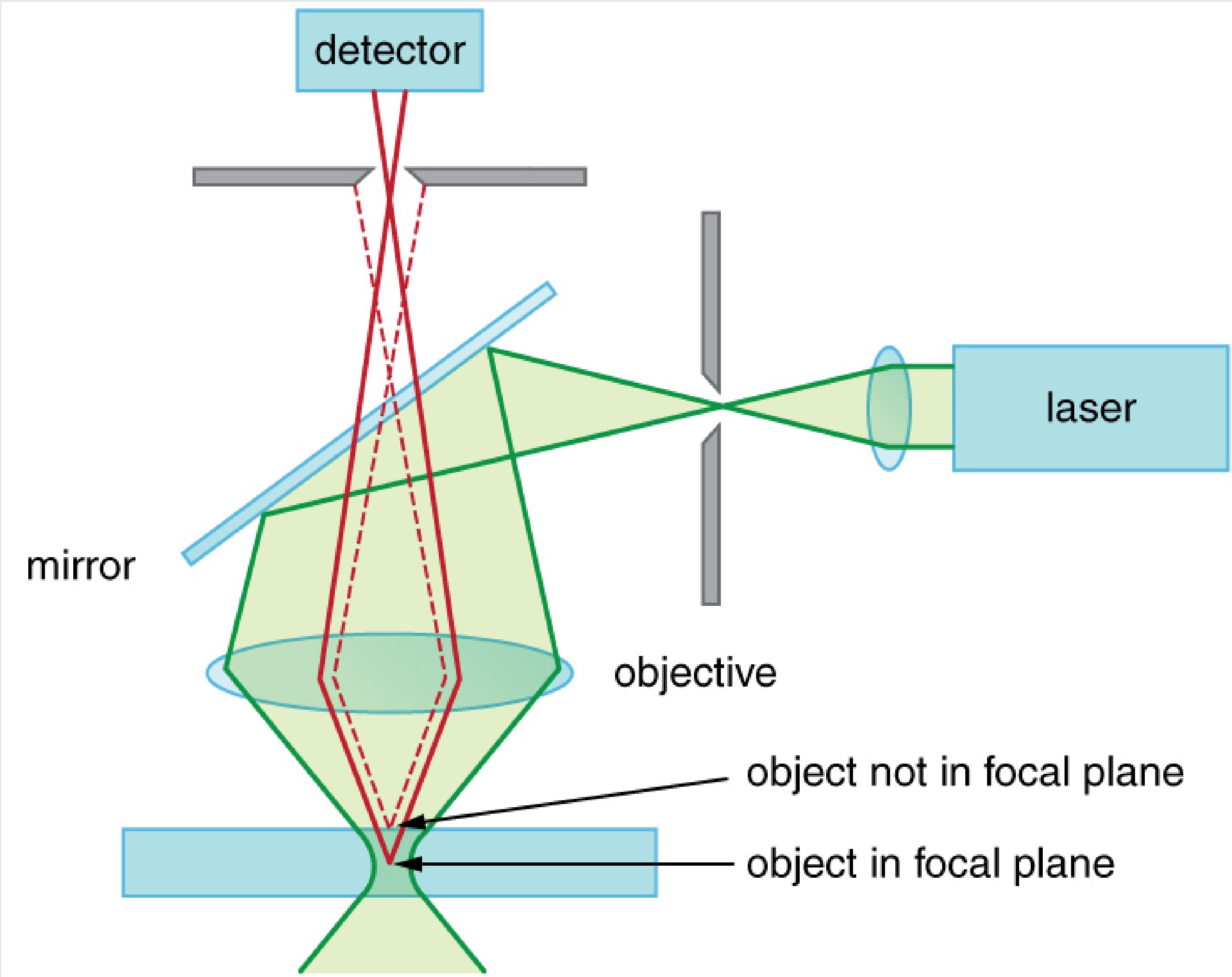
Since the 1990s, improvements in computer processing speeds, advances in digital storage technology, and the development of more stable and powerful lasers have significantly widened the applicability of CLSM as a research tool. Today, CLSM is used to conduct routine investigations on molecules, cells, and living tissues that were not possible even a few years ago. You can learn more about CLSM using the scanning microscope simulator at the Olympus Microscopy Resource Center.[14]

MANIPULATING TINY OBJECTS

Optical Tweezers

Laser light is not only useful for probing matter at microscopic length scales, it can also be used to confine matter for experimental purposes. In 1970, Arthur Ashkin at Bell Laboratories first reported on the ability of laser light to exert

FIGURE 70



A confocal laser scanning microscope.

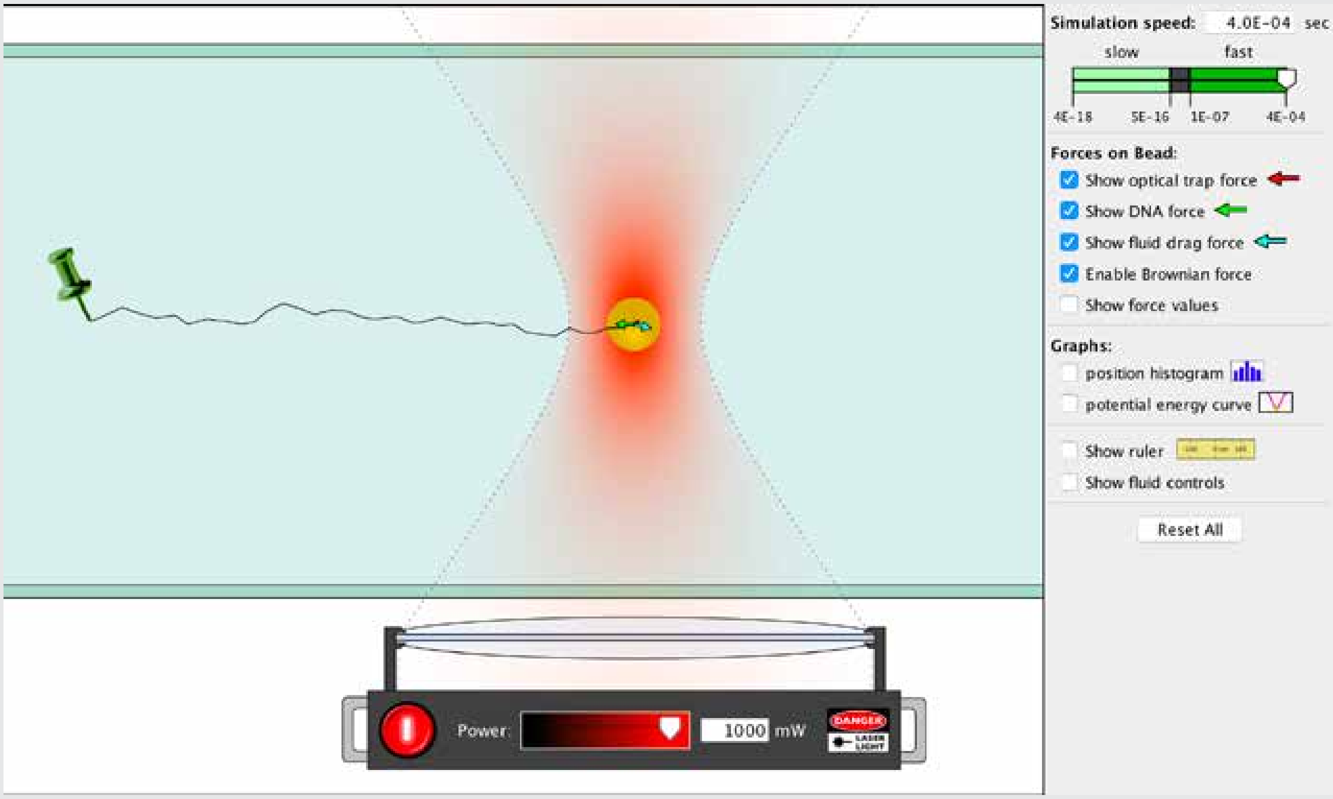
forces on microscopic particles. This discovery led to the development of **optical tweezers**, which are scientific instruments that use a tightly focused beam of laser light to hold microscopic particles stable in three dimensions. The beam can then be used to move these particles around, similar to the way a pair of tweezers can be used to manipulate small objects. The optical tweezing effect was demonstrated by Ashkin and other colleagues in 1986.

Optical tweezers rely on the fact that photons have momentum. As we discussed earlier in the guide, when light is transmitted through an object, it refracts and changes direction. This change in direction corresponds to a change in momentum of the photon. According to momentum conservation, the object acquires a change in momentum that is equal and opposite to that of the photon. By precisely focusing the beam of photons on the object to be constrained, these continual momentum "kicks" effectively push the object inward toward the center of the beam.

Optical tweezers have been used to sort cells, track the movement of bacteria, and measure piconewton-scale forces on microscopic particles.

A primary application of optical tweezers is studying the properties of DNA. A strand of DNA can be attached to a glass or polystyrene bead that is then "gripped" by the

FIGURE 71



Use optical tweezers to virtually manipulate an individual DNA strand using the "Optical Tweezers and Applications" PhET app at phet.colorado.edu.

optical tweezers. By pulling on the bead, scientists have measured the elasticity of DNA as well as the amount of force needed to break bonds within a DNA molecule. DNA molecules within cells are wrapped around proteins but are unwound during the process of replication. In 2001, researchers at Cornell used optical tweezers to unwind a strand of DNA from a protein and discovered that the molecule spontaneously rewrapped to its original state once released. The researchers concluded that enzymes within the cell must act as "molecular motors" that are capable of exerting forces on the DNA strand.

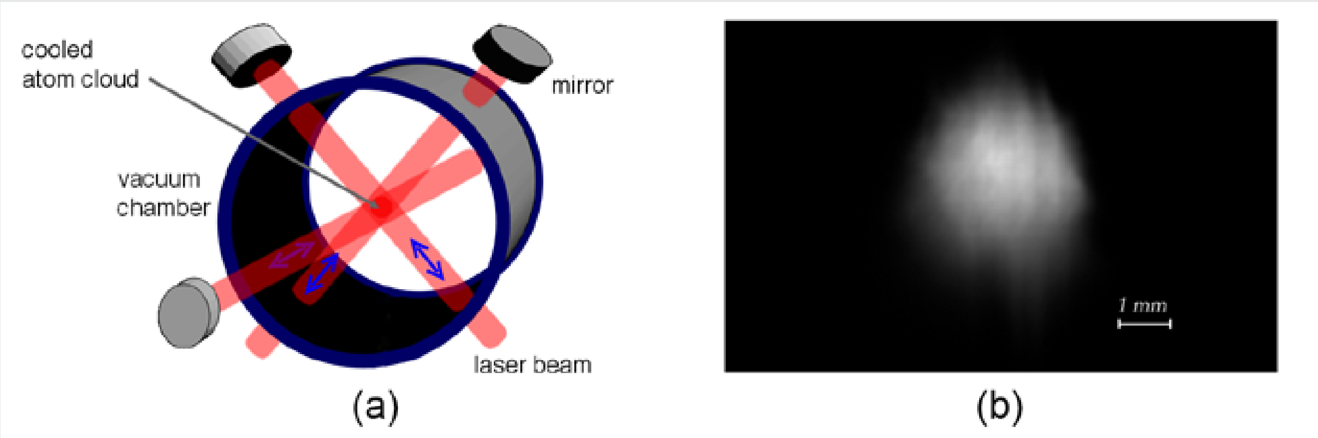
Laser Cooling

Lasers can also be used to control and slow down a collection of atoms in a technique known as **laser cooling**, for which Claude Cohen-Tannoudji, Bill Phillips, and Steven Chu were awarded the 1997 Nobel Prize in Physics. The temperature of any substance is a measurement of how much kinetic energy, on average, its constituent particles have. At room temperature, atoms and molecules in the air move at an average velocity of around 500 meters per second. Naturally, it is difficult to conduct experiments—and measure effects due to quantum mechanics—on individual atoms moving at this speed.

The basics of laser cooling can be understood through a combination of quantum and classical principles. When an atom absorbs a photon, it "recoils" in the direction the photon was moving due to conservation of momentum. The excited atom will then undergo spontaneous emission, releasing a photon in a random direction. After many thousands of these absorption and emission processes, the atoms will tend to be pushed away from the laser beam. We can therefore think of the laser as exerting an overall "pressure" on the gas.

In order to "cool" the atoms, physicists use a technique called Doppler cooling to selectively slow atoms moving

FIGURE 72



(a) Schematic of a laser cooling apparatus (b) A fluorescence image of laser-cooled cesium atoms. Based on the intensity of fluorescence, an estimated 1.5-2 billion atoms are contained within the cloud.

toward the laser. According to the Doppler effect, the apparent frequency of any wave is shifted due to relative motion between the source and receiver. If you are moving toward a source of light, the frequency you observe will be slightly higher in frequency than what you would observe at rest. Therefore, by setting the cooling laser frequency to slightly below an energy transition, only atoms moving toward the laser will be excited by the light and thus slowed down. A conventional atom trap uses six separate laser beams directed inward toward a vacuum chamber in order to cool a collection of atoms (Figure 72).

How cold can we make a collection of atoms? As laser cooling techniques have improved, physicists have been able to cool atoms to temperatures progressively closer to absolute zero. In 2015, a Stanford research lab reported cooling a rubidium gas to 50 trillionths of a Kelvin.[15] The average speed of a rubidium atom in this gas would be less than 70 thousandths of a millimeter per second.

Atomic Clocks

Our ability to control atoms using methods such as laser cooling has led to substantial advances in precision timekeeping and standardization. Early clocks measured time through primarily mechanical means, such as the swinging of a pendulum. These devices tend to be sensitive to small environmental variations or energy dissipation through friction or air resistance, making them unsuitable as a standard for precision timekeeping. In contrast, the frequencies of atomic transitions are extremely stable and are identical across all atoms of the same element.

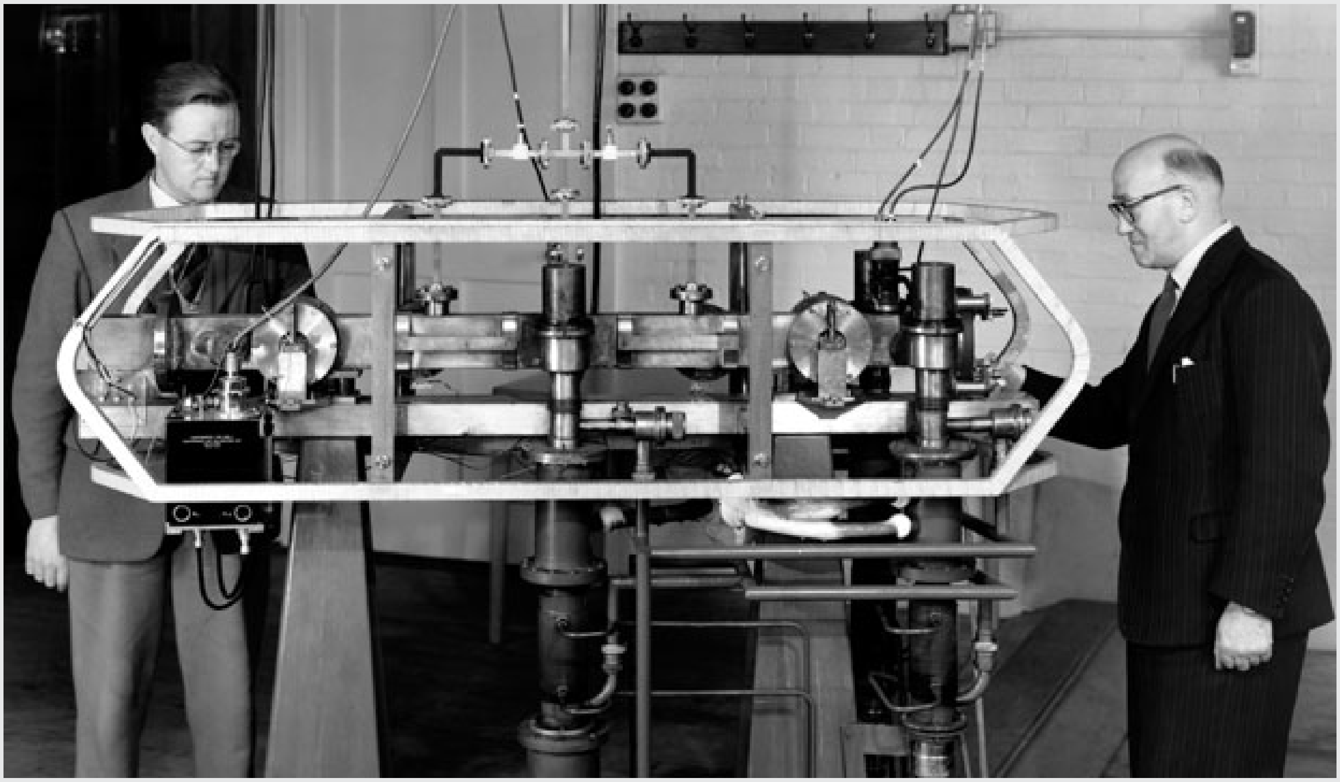
Since 1967, the second has been defined as the unit of time that elapses during exactly 9,192,631,770 oscillations of the radiation produced by transitions between the two hyperfine levels of the cesium-133 ground state. (Hyperfine splitting results from an extremely small shift in ground state energy due to magnetic interaction between the electron spin and nuclear spin.) Atomic clocks have been created using various elements including hydrogen, rubidium, mercury, strontium, and aluminum. Cesium is the current standard used by the National Institute of Standards and Technology (NIST), which regulates standards of measurement in the United States.

The second is the most accurately known unit in the SI system of measurement. The first cesium atomic clock, created in 1955, was accurate to within one second in 300 years (Figure 73). Today, cesium clocks have achieved accuracy to within one second every 300 million years.[16] The substantial increase in accuracy can be attributed to ongoing advances in laser cooling and spectroscopy techniques, which have improved our ability to accurately measure frequencies and reduce experimental noise in atomic systems.

NUCLEAR FUSION

Extremely high-power lasers are being used to initiate nuclear fusion reactions as part of research efforts into fusion

FIGURE 73



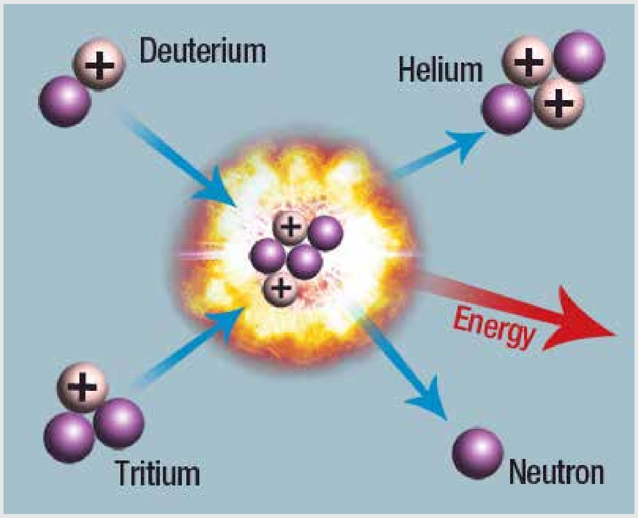
The first cesium atomic clock.

as a viable power source. **Nuclear fusion** is a type of nuclear reaction in which two smaller atomic nuclei combine to form a larger nucleus. (Nuclear fusion can be thought of as the opposite of nuclear *fission*, in which a heavy nucleus splits into smaller nuclei.) Fusion reactions involving light nuclei, such as hydrogen, release energy. Chains of these fusion reactions serve as the power source for stars, including the Sun. An important fusion reaction is the combination of *deuterium* (a hydrogen nucleus consisting of a proton and a neutron) and *tritium* (a hydrogen nucleus consisting of a proton and two neutrons) (Figure 74).

Nuclear fusion is often considered the ultimate power source because it relies on an abundant fuel source (hydrogen), does not produce radioactive waste, and does not present a risk of the catastrophic "meltdown" associated with nuclear fission reactors. However, many challenges remain before fusion can become a viable power source. The greatest challenge to controlled nuclear fusion is the extremely high temperatures necessary to ensure that nuclei are energetic enough to fuse. Fusion of hydrogen nuclei in the Sun requires temperatures of at least ten million Kelvin, and temperatures of that magnitude are difficult to achieve and maintain in a laboratory or power plant. The possibility of laser-induced fusion was first suggested around 1962, not long after the invention of the laser itself. The idea behind laser-induced fusion is to heat and compress hydrogen isotopes to the high temperatures and pressures needed for their nuclei to undergo fusion as they do in stars. This is done by focusing very short and very high-energy laser pulses onto a target made of fusion fuel, thereby heating and compressing the target to very high pressures and temperatures for a fleeting instant. Inertial forces implode the fusion targets, so this approach is called *inertial confinement fusion*.

Inertial confinement fusion relies on high-intensity laser beams to heat and compress a tiny pellet that contains a mixture of deuterium and tritium (Figure 75). The goal of inertial confinement is to cause the fusion reaction to proceed fast enough that there is not enough time for nuclei to move apart. Laser energy is concentrated in short, intense pulses, and the light is distributed uniformly

FIGURE 74



Nuclear fusion.

across the surface of the target, producing a symmetrical explosion that forces the target material inward while heating it to high temperatures. The deuterium-tritium pellet is quickly ionized into a plasma and heated to temperatures above 10^8 K. Fusion is achieved in under 10^{-9} s, or less than one billionth of a second.

Government scientists have built a series of increasingly larger and more powerful lasers, which have demonstrated that laser fusion is possible but not easy. Each step paved the way for a bigger and more powerful laser. The U.S. Department of Energy began construction of the latest in the sequence, the National Ignition Facility (NIF), in 1997 at the Lawrence Livermore National Laboratory in California. (The name comes from the goal of crossing a threshold called **fusion ignition**, at which the fusion reaction becomes self-sustaining.) The NIF was completed in March 2009 and is currently among the largest and highest-energy inertial confinement fusion devices in the world.

The NIF occupies three connected buildings that together are 704 feet long, 403 feet wide, and 85 feet tall. The NIF uses the most powerful laser assembly ever constructed, driven by a system of amplified neodymium-glass lasers, to ignite fusion in a hydrogen pellet 2 mm in diameter. In 2012, the NIF conducted a laser pulse with a power of 500 trillion watts—equal to 1,000 times the power used by the entire United States at any instant in time. In 2013, NIF achieved an important milestone—more energy was

FIGURE 75

The deuterium-tritium target pellet at the National Ignition Facility.

extracted from the fusion reaction than was used to trigger the reaction itself.[17] Although progress is being made, fusion ignition remains an elusive target that may take at least another decade of further research to achieve.

MEDICAL APPLICATIONS

LASER SURGERY

Lasers have had a tremendous impact on medicine. Laser surgery techniques are based upon an understanding of how laser light interacts with biological tissue, which entirely depends on the wavelength of light being used and the nature of the tissue. The most important component of biological tissue is water, which absorbs strongly in the infrared range of the electromagnetic spectrum. In fact, water is such a strong absorber of infrared and accounts for so much of the bulk of soft tissue that it is a fair approximation to consider tissue as absorbing light as if it was water. Water absorbs about 80 percent of the incident 10.6-µm wavelength of a CO_2 laser in the first 20 µm, corresponding to the surface of exposed skin or tissue. Absorption is even stronger at shorter wavelengths of 3–6 µm, although it is not uniform through the infrared band.

Carbon dioxide lasers are a common choice for laser surgery because they are readily available and emit 10.6-µm wavelength light that is easily absorbed by tissue. A CO_2 laser beam focused on tissue at a sufficiently high intensity will cause cells to be vaporized due to energy absorption.

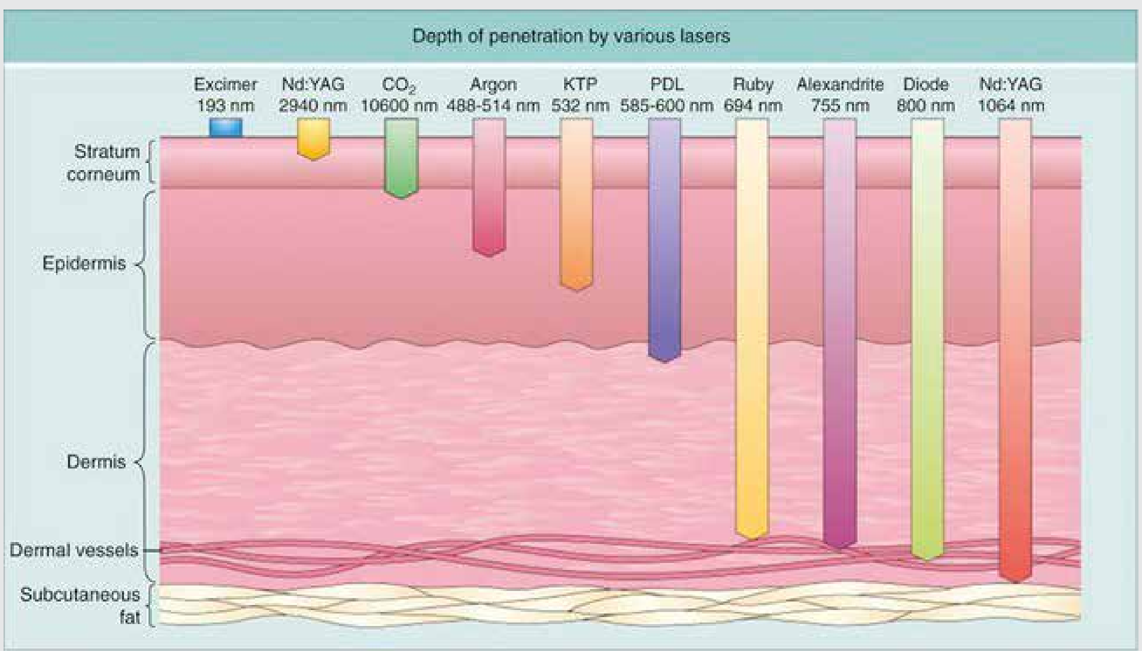
The absorption is so strong that only the upper layer of cells is vaporized; these cells absorb virtually all of the light, so the lower level of cells survives with little damage (Figure 10). The lower layer suffers even less damage from a laser emitting at a wavelength of 3 μm because water absorbs that wavelength even more strongly.

Another advantage of the 10.6-μm CO_2 laser wavelength is that it penetrates just deep enough into tissue to seal small blood vessels and stop bleeding. This cauterization effect makes the CO_2 laser especially valuable for surgery in regions rich in blood vessels, such as the gums and the female reproductive tract, by giving the surgeon a tool to remove thin layers of blood-rich tissue. This is particularly useful in the treatment of gum disease and of endometriosis, a condition affecting several million American women. Carbon dioxide lasers also are used in a type of heart surgery that creates new paths for blood vessels in the heart, called transmyocardial revascularization. In fact, certain types of surgeries were only made possible with the introduction of the laser because they involved incisions that resulted in a high degree of bleeding with a conventional scalpel.

Carbon dioxide lasers at 10.6 μm cannot be used to cut bone because bone contains less water than other tissue. However, lasers emitting at wavelengths of 3 μm can cut bone because bone contains some water that strongly absorbs that wavelength. Surgeons generally use other tools for cutting bone and for many other types of surgery, particularly general surgery on internal organs, where lasers have no particular advantages over conventional scalpels.

Other laser wavelengths in the ultraviolet, visible, and near infrared have other advantages for certain types of medical treatment. For example, lasers can be used to treat "port-wine stains," which are a type of dark red birthmark resulting from defective blood vessels near the surface of the skin. In this case, the laser is selected to match the peak absorption of the blood vessels causing the blemish. As another example, the 193-nm wavelength of the argon-fluoride excimer laser is strongly absorbed by the lens of the human eye, making it ideal for refractive surgery. Recent advances in laser medicine have been made possible by matching the laser wavelength and power level to specific treatment needs.

FIGURE 76



Penetration depths of laser light for various laser types. CO_2 lasers are capable of destroying only the outermost layer of skin cells while leaving the deeper tissue unaffected.

Refractive Eye Surgery

The most commonly advertised type of laser surgery today uses an argon-fluoride excimer laser to reshape the lens of the eye to correct refractive defects that otherwise would require people to wear corrective eyeglasses or contact lenses. Refractive surgery procedures are used to correct both myopia (near-sightedness) and hyperopia (far-sightedness) as well as astigmatism, which is blurred vision due to uneven focusing of light by the eye. The development of the excimer laser in 1970 paved the way for the use of lasers in refractive eye surgery. Researchers in the early 1980s discovered that 193-nm pulses from argon-fluoride excimer lasers could ablate, or dissolve, tissue very efficiently from the cornea, the transparent front layer of the eye. As the cornea provides much of the eye's refractive power, the idea of using surgical lasers to correct vision by reshaping the cornea followed almost immediately (FIGURE 77).

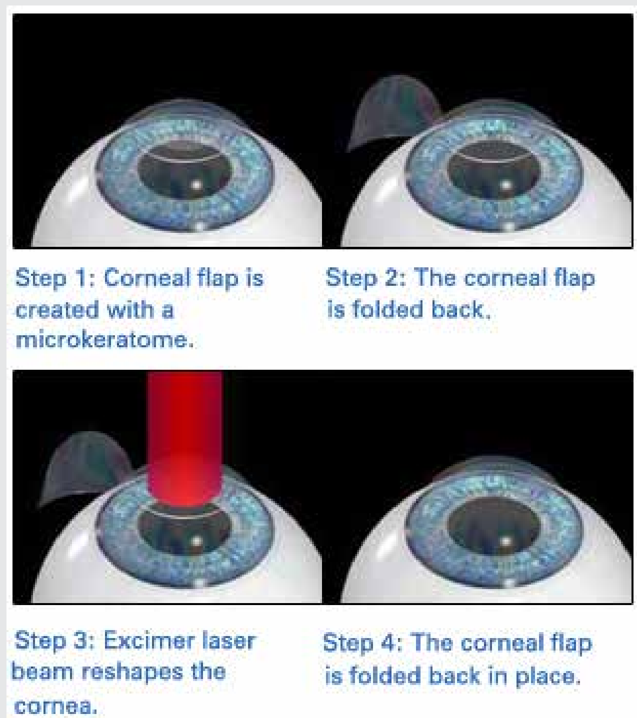
FIGURE 77

LASIK refractive surgery procedure being performed at National Naval Medical Center Bethesda in 2007.

The first type of laser refractive surgery, known as *photorefractive keratectomy* (PRK), uses an excimer laser to ablate material from the cornea after the top layer of the cornea, known as the *epithelium*, has been chemically wiped away. Concern over the effects of removing the epithelium led to the development of **LASIK** (*laser-assisted in-situ keratomileusis*), in which the epithelium is peeled back before the excimer laser ablates tissue from the underlying layer, then put back in place afterward (FIGURE 78). The LASIK procedure was approved in the United States in 1999; by 2002 LASIK had already become the world's most common elective procedure.[18] Although LASIK is currently the most common form of refractive surgery, PRK is still in use and new variations are under development.

Refractive surgery techniques have become steadily more efficient and successful as our laser and computational technology have improved. Current systems can measure the refractive profile of the eye to compute the precise corrections

FIGURE 78



The LASIK surgical procedure involves peeling back the epithelium and then reshaping the cornea using light from an excimer laser.

required to compensate for the eye's defects. However, some complications remain, and accuracy of the correction often depends on the healing response of the eye. Some patients require a second surgery to complete correction.

LASER DERMATOLOGY

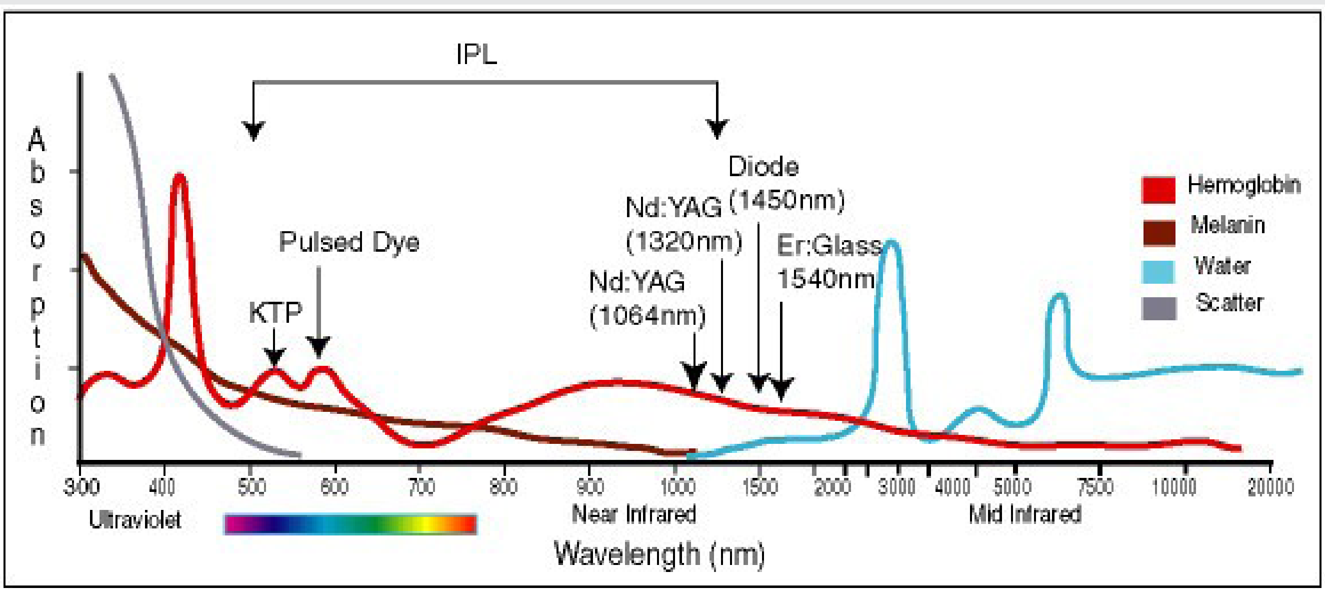
Laser dermatology procedures achieved early success in treating dark-red birthmarks called "port wine stains," which often appear on the face or neck. These discolorations result from blood in networks of abnormal blood vessels just under the surface of the skin. Because port wine stains tend to be spread over the surface of the skin, they cannot be treated by conventional surgery, and concealment with cosmetics was often the only solution. Dermatologists made early attempts to treat port wine stains using blue-green argon laser light, but these procedures were painful, worked only with dark blemishes, and could cause scarring. These problems were reduced by switching to pulsed dye lasers tuned to emit near the peak absorption of hemoglobin in blood at about 580 nm (Figure 79). At this wavelength, light absorption is concentrated in blood vessels, thereby avoiding skin blisters, and can be used with lighter birthmarks or on children. Lasers also are used for treating other skin blemishes, such as acne scars.

Tattoo removal is another familiar medical application of lasers. Tattoos consist of colored ink that has been inserted into the dermis layer of the skin. Laser tattoo removal works by illuminating tattoos at a wavelength absorbed by the tattoo pigment, thereby breaking down the pigment molecules to bleach the tattoo. The effectiveness of the procedure depends on the ink; some tattoos almost disappear, but others fade only slightly. A successful laser treatment will generally leave a slightly discolored spot where a traditional tattoo used to be. Laser tattoo removal may be made easier due to a new family of encapsulated tattoo inks, which are designed for people who might have second thoughts. The ink is contained within tiny transparent plastic spheres, which protect them from degrading in the body. A single laser treatment splits open the spheres, dumping the dye into cells, which break down the molecules to bleach the color.

Laser skin resurfacing is another cosmetic procedure in which a carbon dioxide laser is scanned across aged areas, particularly the face, neck, and hands, to remove the surface layer. This process removes surface wrinkles and blemishes, thereby exposing a fresh layer of skin. Laser skin resurfacing requires a few weeks for the skin to heal and return to normal color.

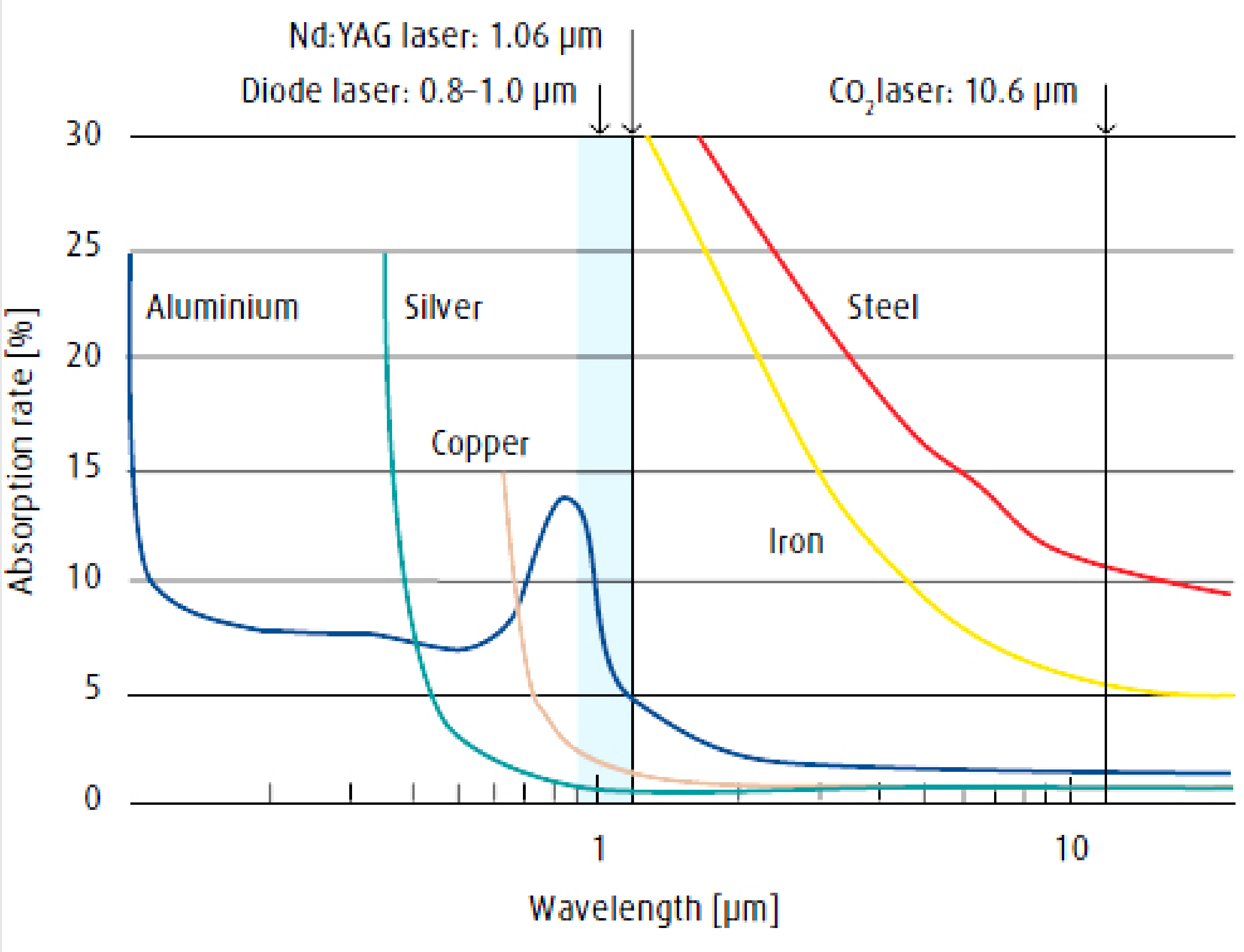
Lasers can also be used for the purpose of hair removal, which is accomplished by directing near-infrared laser light at hair follicles. A pigment called melanin in the hair and

FIGURE 79



Relative absorption versus wavelength of laser light by hemoglobin, melanin, and water.

FIGURE 80



Absorption rates of various metals by wavelength. Higher values indicate that light of that wavelength is more readily absorbed by the metal.

follicle absorbs the light, heating and killing the follicle, and thus stopping hair growth in areas where hair is undesired. The effectiveness of laser hair removal treatment depends on the contrast between skin and hair color; the darker the hair and the lighter the skin, the more effective the treatment. Different lasers may be used for different skin and hair combinations. The 755-nm output of a pulsed alexandrite laser works well, but only on light skin. Arrays of diode lasers emitting pulses at 810 nm are used for light to medium skin, and neodymium lasers emitting at 1064 nm are used for darker skin. It is more difficult to remove light-colored hair than dark-colored hair with laser light due to the lack of melanin pigment.

LASER DENTISTRY

There are three primary applications of laser dentistry. Lasers can be used to "drill" teeth, using solid-state erbium lasers that emit 3-µm wavelength light. Although these lasers cannot remove solid enamel well, they can remove decayed areas and prepare cavities for repair. More important to both dentist and patient, laser pulses are less threatening than the traditional whirring mechanical drill and can be less painful as well. As mentioned earlier, carbon dioxide lasers can also be used to treat gum disease by removing swollen tissue. Laser teeth whitening is another popular cosmetic dental treatment that uses visible argon and neodymium lasers to activate a special gel solution that has been applied to the patient's teeth.

INDUSTRIAL APPLICATIONS

Many industrial and manufacturing applications rely on high-power lasers to cut, shape, and join materials. Energy transfer to the object by the laser depends on absorption at the laser wavelength, which varies considerably among materials (Figure 80). In general, metals tend to reflect more light at longer wavelengths and absorb a larger fraction of laser energy at shorter wavelengths. Accordingly, most metals are easier to cut with shorter wavelengths of laser light. Titanium is a notable exception to this pattern because it absorbs a relatively high 8 percent of light from a CO_2 laser and is almost impossible to cut with conventional saws. Other factors also affect whether lasers or conventional tools are more appropriate for a machining task. For example, aluminum is rarely cut with lasers because it reflects strongly across the spectrum and is soft enough to be cut easily with ordinary saws or blades.

Absorption at the solid surface does not tell the whole story, however. Once the material absorbs some energy, the material melts, then vaporizes. The absorption of most materials increases with temperature, so the first bit of heating is the hardest, and liquid metals often absorb more energy than the metals in the solid phase. Vaporization can create problems by forming a layer of vapor or plasma between the laser and the object. This plasma layer absorbs some laser energy, blocking it from reaching the object. Sometimes, turbulence produced by heating the object can remove some or all of this plasma, but in other cases the plasma must be removed or blown away to improve energy absorption by the material.

The optics that focus the laser beam control how deeply the laser can penetrate the material. A lens with a short focal length focuses light into a tighter spot, but the beam spreads out faster beyond the point where the focal spot diameter is smallest, reducing the concentration of light deeper in the hole. A lens with a longer focal length does not produce as tight a focal spot, but the beam does not spread out as fast, so it can drill deeper holes and cut thicker sheets of material. Machinists must take depth of focus into account when selecting the appropriate lens for the drilling or cutting task.

LASER DRILLING

Laser drilling is accomplished with one or more short laser pulses, each of which removes some material. The drilling process requires high peak power and works well for a sequence of short repetitive pulses. (If the energy is delivered over too long an interval in a long pulse, the peak power may not reach high enough levels to vaporize and remove material from the hole.) The number of pulses required depends on factors including wavelength, peak power, the nature of the material, and repetition rate. Laser pulses can even be made powerful enough to drill holes in diamond, which is among the hardest materials on Earth. For most efficient drilling, it may be necessary to change the optical focus during drilling in order to reach deeper into the drilled hole. The laser beam and/or the object can also be moved to adjust the spot being drilled. It also is important to make sure that the vaporized material is removed from the drilling site.

LASER CUTTING

Laser cutting can be thought of as drilling a series of overlapping holes in a material. In cutting, the laser beam and/or the object move continuously, and the beam itself normally operates continuously rather than being pulsed (Figure 81). Laser cutting is typically accomplished with the assistance of a jet of air, oxygen, or dry nitrogen. For nonmetals, the role of the jet is to blow debris away from the cutting zone and improve the quality of the cut. Lasers can cut readily through sheets of wood, paper, and plastic, but thick materials are more difficult. Foods also tend to char when cut, leaving an unappetizing black layer. In other words, lasers aren't particularly good tools for slicing bread—unless you prefer your toast burnt to a crisp.

Laser cutting of metals works slightly differently and uses a different type of jet. In this case, the laser beam heats the metal to a temperature hot enough that it burns as the

FIGURE 81

A four-kilowatt industrial laser cutting a sheet of metal.

oxygen from the jet passes over it. This process is properly called "laser-assisted cutting" because the oxygen in the jet is actually responsible for the combustion process that causes the cut.

Carbon dioxide lasers are commonly employed for laser cutting applications. The primary advantages of laser cutting over mechanical cutting, such as with a saw, blade, or drill, are speed and accuracy. Laser cutters can be precisely controlled, and in many cases, automated. The main disadvantage of laser cutting is the relatively high power consumption of industrial lasers. In recent years, laser cutting technology has become more accessible and affordable, making it more widely available to hobbyists.

LASER BEAM WELDING

Welding is effectively the opposite of cutting; cutting separates one object into two or more pieces, whereas welding joins two or more pieces together into a single whole (FIGURE 82). Both processes require the entire thickness of the material to be heated, but the specific conditions differ.

In order to form a solid bond, the objects to be welded must be fitted together precisely, leaving very little room between them. The laser beam then heats the edges of the two plates to their melting points, causing them to fuse together at the points where they are in contact. If the pieces are not in contact along the entire junction, the weld will be flawed. The laser beam must also penetrate the entire depth of the weld (i.e., the thickness of the material) to form a proper joint. The heat of welding transforms the metal through the entire depth of the weld and in a zone extending on both sides of the actual junction. The width of the weld zone depends on the composition of the materials being welded, the power delivered by the laser, and the speed with which the laser beam moves along the joint. The farther from the mid-point of the weld, the less the heat affects the metal.

Machinists must consider several factors before welding, including the composition of the materials and their thicknesses. The materials do not have to be identical, but metallurgists have learned that some types of materials do not bond together well, no matter how thoroughly they are heated or what technique is used for welding. Like cutting and drilling, laser welding is also limited by the thickness of material it can handle. However, it has gained wide acceptance in manufacturing a wide variety of products in a range of sizes, including razor blades and heart pacemaker cases. Laser beam welding is commonly employed in high-volume applications that involve automation, such as in the automotive industry. Both CO_2 and Nd:YAG lasers are commonly used for welding purposes.

FIGURE 82

Laser beam welding process. The laser beam is invisible and coming directly from above.

LASER MARKING, ENGRAVING, AND ETCHING

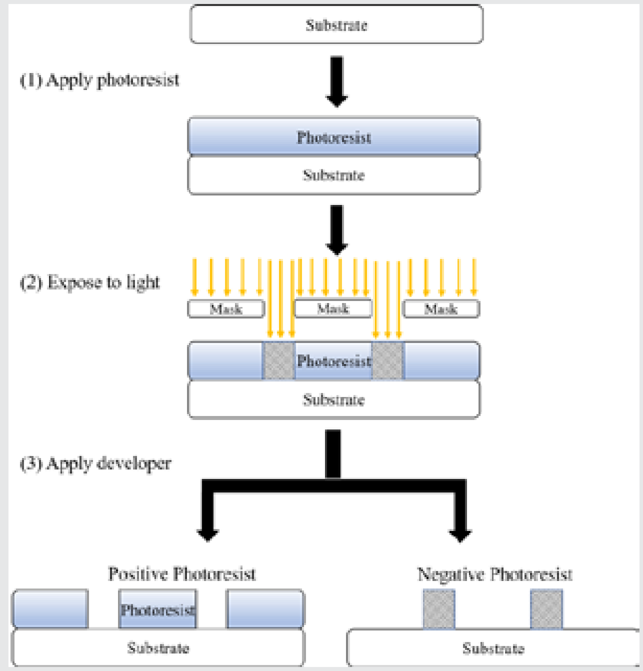
In addition to drilling or cutting completely through materials, lasers can be used to leave a permanent mark on a material, either by oxidizing the surface, carving out a shallow indentation, or melting the outer layer. Laser marking is achieved by using a low-power laser beam to heat and oxidize under the surface of the material, leaving a dark discoloration. This process leaves the surface intact but creates permanent high-contrast marks.

Laser engraving uses a higher-power laser to effectively "drill" a shallow indentation in the outer layer of the material by vaporizing a tiny amount of surface material. As with the laser drilling process, several laser passes may be used to create deeper marks. Laser engraving is commonly used to permanently mark serial numbers or logos on products or electronic components (FIGURE 83). *Laser etching* is a subset of laser engraving in which heat from the laser beam melts the surface of the material, causing it to expand slightly and leaving a raised mark. Wood, plastics, metals, stone, and glass can all be etched using lasers.

FIGURE 83

An electronic component that has been marked by a laser.

FIGURE 84



A basic illustration of the photolithography process.

PHOTOLITHOGRAPHY

Digital devices we use every day, including smart phones, computers, and tablets, all rely on integrated circuits, or microchips, in order to function. These integrated circuits are made up of patterned layers of semiconductor material and are fabricated in a laser process known as **photolithography**. The first step in the photolithography process is to coat the surface of a semiconductor wafer with a light-sensitive material called a *photoresist*. Next, an ultraviolet light source illuminates the photoresist through a pattern called a *mask*, thereby exposing patterns in the photoresist. Depending on the type of photoresist, either the exposed or unexposed areas are etched away chemically. If additional layers of circuit are needed, another layer of photoresist is deposited, and the process is repeated to create another layer of patterns on the wafer (Figure 84).

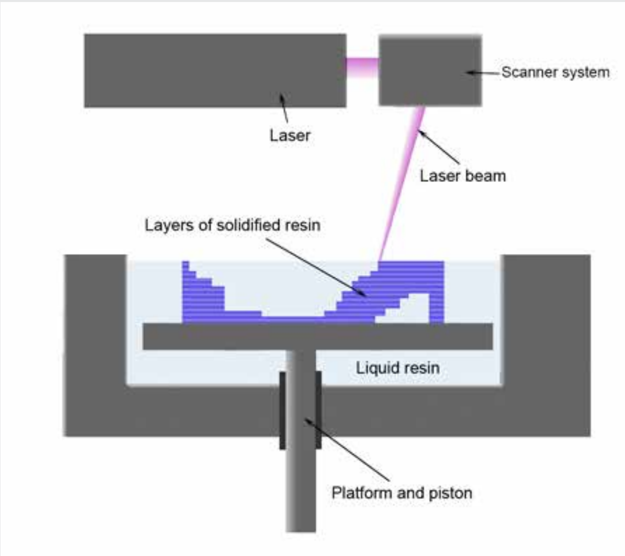
The driving factor in increasing the power and complexity of integrated circuits is the number of circuit elements that can be integrated on a single chip, which in turn depends on the minimum feature size that can be fabricated. In order to shrink the feature size, photolithography has moved to shorter and shorter wavelengths. As such, ultraviolet excimer lasers have become the preferred light sources for fabricating microchips. The 248-nm wavelength of krypton-fluoride excimer lasers was used in the past, but the current standard is the 193-nm wavelength of argon-fluoride (ArF) excimer lasers.

The resolution of photolithography can also be enhanced by using optical systems that focus the light tightly, so it can expose features only a fraction of a wavelength. Photolithography techniques have been improved by immersing the optics in liquids with high refractive index, in a process known as immersion lithography. This has the effect of increasing the focusing power of the lens to make tighter focal spots. Using a 193 nm ArF excimer laser and liquid immersion techniques, fabricators have succeeded at printing features smaller than 50 nm. Other efforts are focused on improving the optical systems to stretch ArF laser technology to produce finer details. Semiconductor fabrication is by far the largest single application for excimer lasers, so improving ArF technology is an important goal within the laser industry.

STEREOLITHOGRAPHY

Lasers can also be used to "build" three-dimensional objects, using techniques very different from other types of laser materials working. These techniques are among several competing technologies for *rapid prototyping*, which produces three-dimensional objects in shapes specified by computers. Although the parts typically are used in visualization or testing, eventually they may be made on demand for general use.

FIGURE 85



Stereolithography apparatus.

One laser technique, called **stereolithography**, relies on the fact that some liquids solidify when exposed to ultraviolet light. In a stereolithography system, an ultraviolet laser beam is focused onto a platform slightly below the surface in a bath of liquid that solidifies when exposed to ultraviolet light. The computer controls scanning of the beam across the surface of the liquid, forming a thin layer of solid on the platform. Then the platform is lowered further into the liquid, and the laser beam scans the surface again, following another pattern to build the next layer of the model. After that layer is formed, the platform is lowered again, and the cycle repeats layer by layer to build a three-dimensional model (Figure 85).

A related rapid prototyping technique is *selective laser sintering*, in which a high-power laser beam is scanned to melt small particles in a powder that covers a small platform. Then the platform is lowered, and the process is repeated to solidify another layer. Other than its use of a powder and longer-wavelength lasers, the technique is similar to stereolithography. Although stereolithography is fast and versatile, it is relatively expensive compared to other rapid prototyping techniques. Nevertheless, stereolithography systems are commercially available to hobbyists and machinists.

OTHER APPLICATIONS

BARCODE READERS

Barcode readers, also known as barcode scanners, are a familiar application of laser technology. Barcode readers work by scanning light across a pattern of dark lines of varying thickness. The dark bars absorb the light, whereas white spaces between the bars reflect the light. A photodiode measures the intensity of reflected light and interprets it as a string of information—for example, a product code in the case of a grocery store scanner or a book identification code at a library. The first laser barcode scanner was used in 1974. Although HeNe lasers were originally employed for this purpose, they have gradually been replaced with cheaper and more compact diode lasers.

LASER POINTERS

Laser pointers are small, battery-powered handheld devices capable of emitting a low-power beam of laser light in the visible spectrum (Figure 86). Laser pointers are typically used to highlight certain items when giving a presentation, but they have other uses as well. For example, construction companies often employ laser pointers as part of surveying equipment or leveling apparatuses. Robotics systems also use laser pointers as a guidance system to orient and direct a robot toward a target. Laser light shows at entertainment venues use laser pointers to provide an impressive spectacle, often coupling them

FIGURE 86

Three laser pointers emitting light of different wavelengths in the visible spectrum.

withsmoke machines to make the full laser beam visible due to Rayleigh scattering.

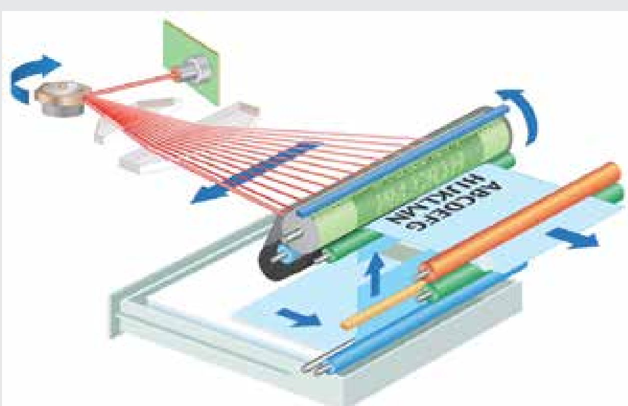
Because they are commercially available, the output power of laser pointers is carefully regulated to prevent accidental eye damage resulting from misuse. According to the United States Food and Drug Administration, any laser with an output power exceeding 5 mW cannot be sold or marketed as a laser pointer. It's important to note that even "safe" low-power laser light can cause eye damage if viewed directly for extended periods. You should always exercise caution when using a laser device of any output power.

LASER PRINTERS

Laser printers use a laser beam to guide ink to be placed on the paper. When you click "Print," your computer sends a stream of data to the laser printer, typically made up of millions of *bytes*, or single characters that encode the information you want printed. Inside the printer, an electronic circuit interprets this data. A rotating drum accumulates a positive electric charge. The circuit activates a laser diode that reflects a beam off a rotating hexagonal mirror that scans it back and forth across the rotating drum. Wherever the laser beam strikes the drum, the positive charge is replaced by a region of negative charge instead. The position of the laser beam is precisely guided by the circuit to match the printing instructions: negative charge where the page should be black, and positive charge where the page should be white (FIGURE 87).

An ink roller comes in contact with the rotating drum,

FIGURE 87



A laser printer.

FIGURE 88

Publicity photo of the Xerox 9700, the first commercial laser printer in 1977.

passing along positively charged particles of ink called toner. The positively charged toner particles are attracted to the negatively charged regions on the drum, but repelled by the positively charged regions. Thus, an image of the page builds up in ink on the drum. A sheet of paper is then fed from a hopper within the printer, being given a positive charge as it moves toward the drum. As the paper passes by the drum, the negatively charged toner particles are transferred over to it, and a final heated roller fuses the toner to the page.

In 1969, Gary Starkweather at Xerox developed the idea for laser printing. In 1977, Xerox introduced the first commercial laser printer, the Xerox 9700. Early laser printers tended to be expensive and bulky—the Xerox 9700 was roughly the size of four modern-day photocopiers joined together (FIGURE 88). By the late 1970s, several major computer companies, including IBM, Hewlett-Packard, and Canon were competing to develop affordable mass-market laser printers. In May 1984, Hewlett-Packard introduced the HP LaserJet, the first desktop laser printer, at a retail price of $3,495. The Apple LaserWriter printer was introduced the following year. These printers are credited with bringing desktop publishing to the mainstream.

In the decades since their introduction, laser printers have become faster, more affordable, and are now capable of printing higher-quality images than earlier models. They are commonly employed for high-volume printing needs, such as in offices and schools. Although laser printers were once only capable of black-and-white printing, color versions

FIGURE 89

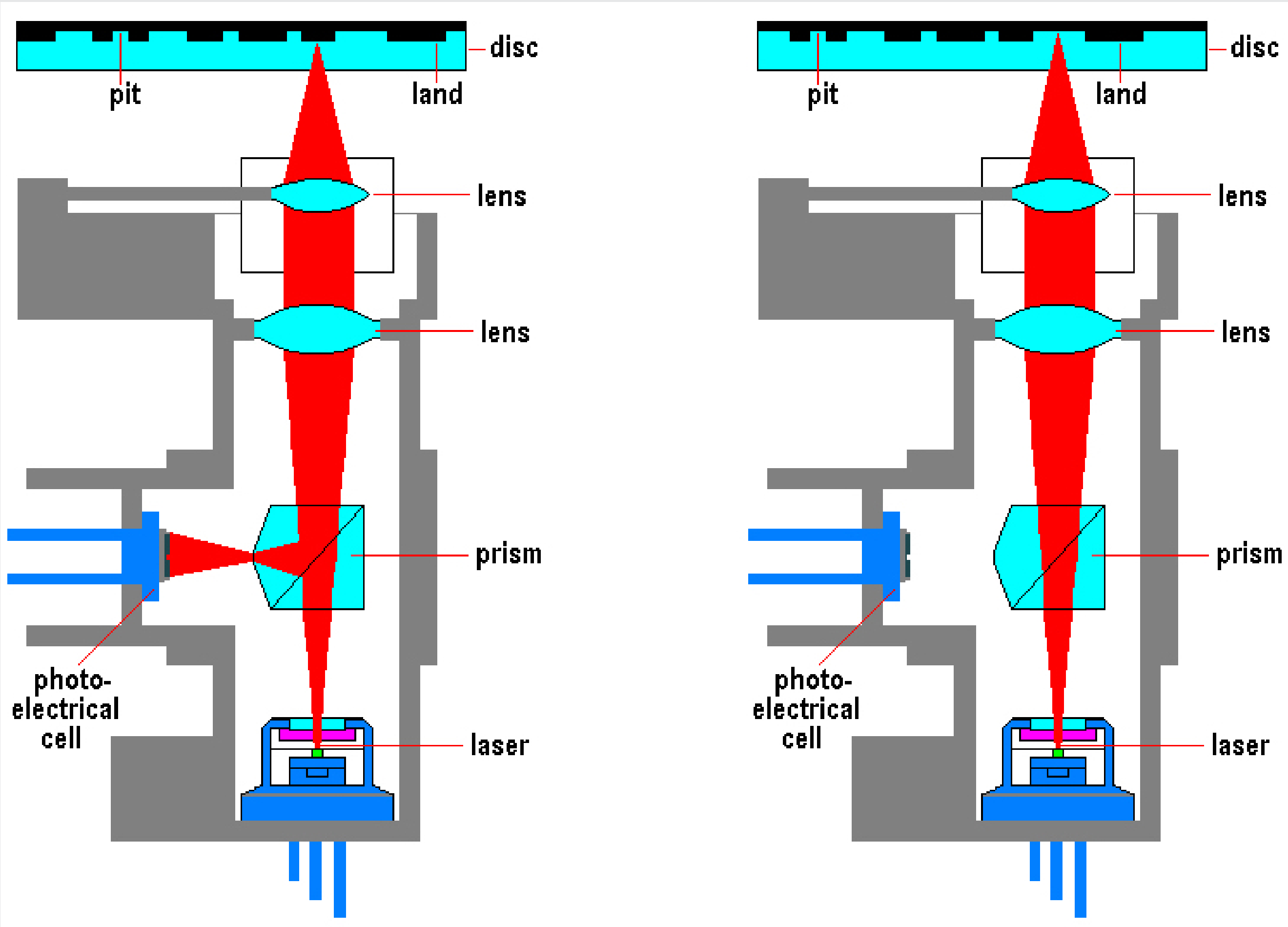


Diagram showing how a laser is used to read information from a CD.

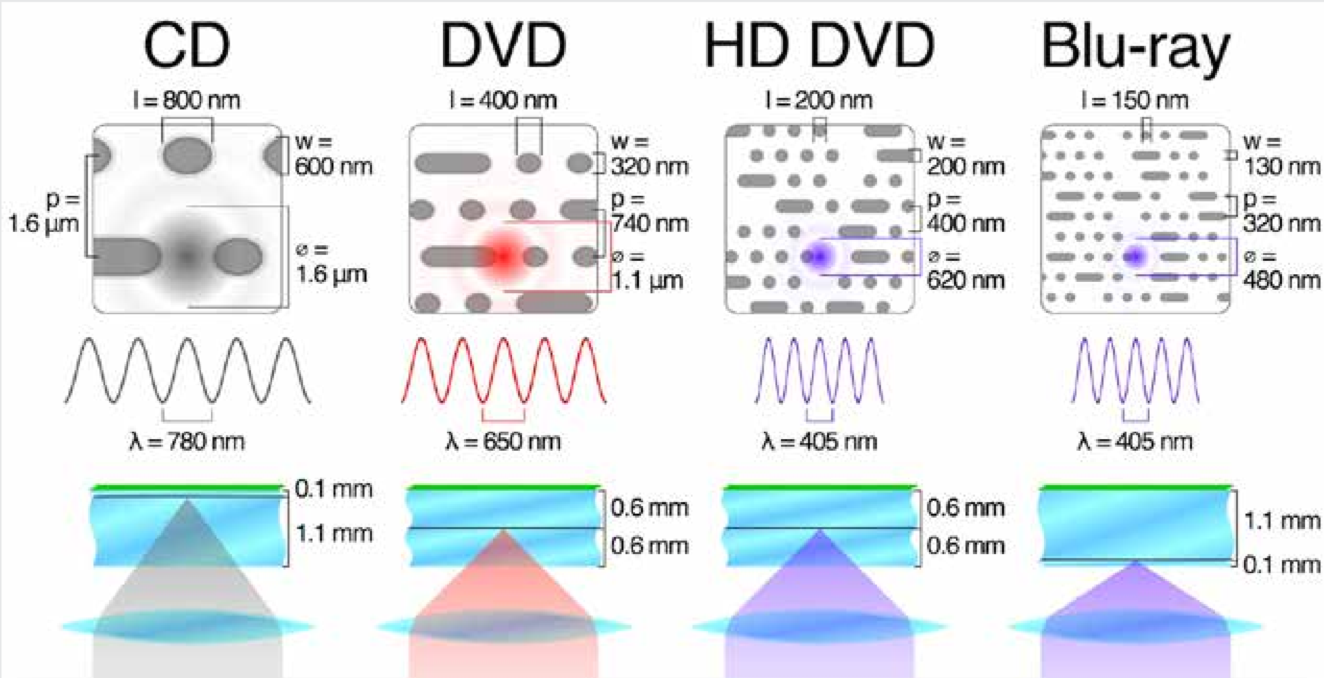
are now readily available. Color laser printers work by making three passes using different colored toners, one for each of the three colors needed to generate the full range of colors for reflected images.

OPTICAL DISCS: CDs, DVDs, AND OTHERS

In terms of sheer numbers, the widest application of laser technology is the recording and playing of information on optical discs such as compact discs (CDs) or digital versatile discs (DVDs). Let's take a closer look at how this works.

A standard CD is a disc of metal and plastic with a 4.7-inch diameter. A CD consists of three layers: a plastic base containing the track information, a reflective aluminum layer, and a transparent polycarbonate layer that protects the track information. Information on a CD or DVD is encoded as a string of ones and zeros, which can be interpreted by a disc player as a song or movie. A bump, known as a "pit," corresponds to a zero byte of information. The depth of these pits is configured to be one-quarter of a wavelength of the laser light in the plastic (recall that the wavelength of light changes in a transparent material compared to air). Therefore, when the beam reflects back from the surface of the pit, it will be shifted a half-wavelength compared to the rest of the wave and undergo destructive interference. The photodetector then interprets the corresponding low light intensity as a zero (Figure 89).

FIGURE 90



A comparison of the laser wavelength and storage capacity of CDs, DVDs, HD DVDs, and Blu-ray. Dimensions indicated are track separation (p), pit width (w) and minimum length (l), and laser spot size (ø) and wavelength (λ).

The precursor to the CD, the LaserDisc, was introduced in 1978. A standard LaserDisc was a 12-inch-diameter metal and plastic disc that had the appearance of an enlarged CD. The LaserDisc format was introduced as a higher quality alternative to VHS videotape recordings. Ultimately, the high costs of players and discs, combined with their inability to record, led to the commercial failure of the technology. In 1995, the DVD format was introduced by Panasonic, Philips, Sony, and Toshiba. With the advent of digital media storage and internet file distribution in recent years, CDs have declined in popularity and sales. It may not be long until CDs join LaserDisc as an obsolete media format.

A standard CD player uses an infrared laser with a wavelength of 780 nm. DVD players use shorter wavelength lasers around 650 nm, which allows the pit depth, track separation, and pit length to be scaled down accordingly (Figure 90). For this reason, a standard DVD can store approximately thirty times more information than a CD. More recent technology, such as HD DVD and Blu-ray discs, use even shorter wavelength lasers (405 nm, at the violet end of the visible spectrum). Accordingly, these formats have about ten times the storage capacity of a DVD, and 300 times that of a CD.

HOLOGRAPHY

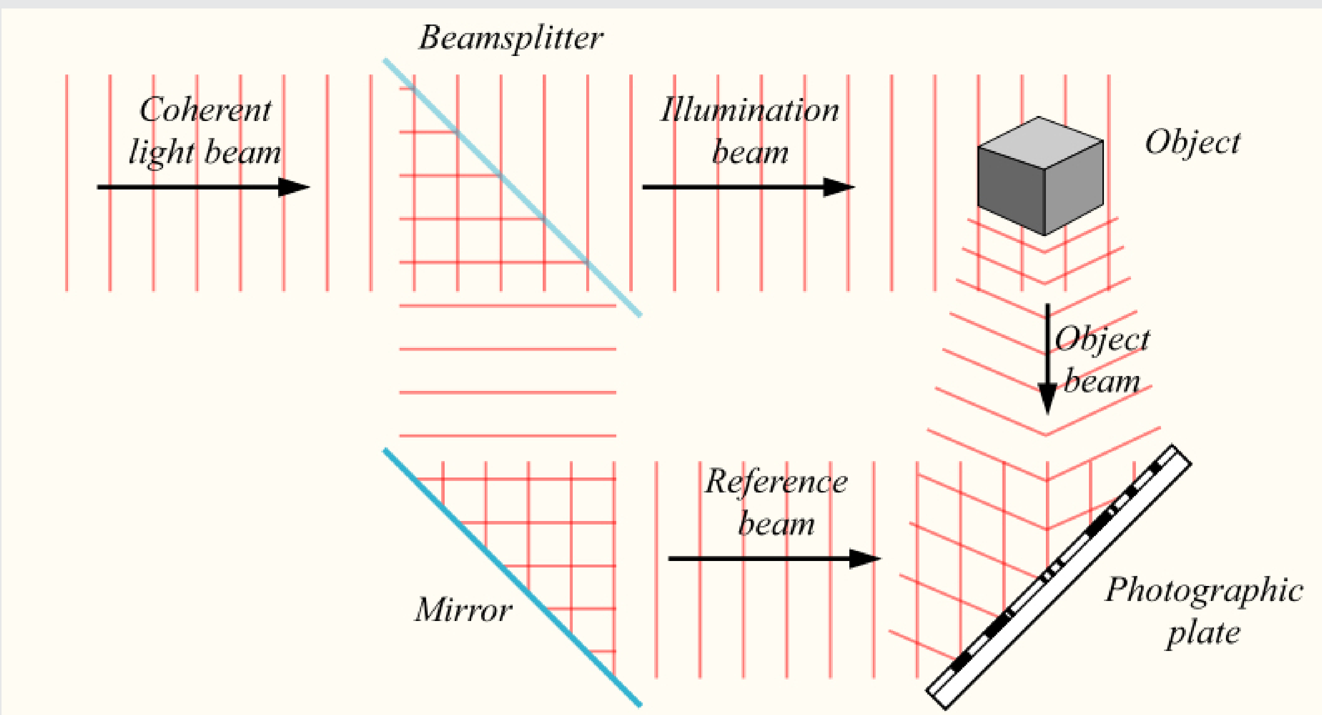
Holography is a method of producing a three-dimensional recording of an object. The hologram itself is not an image, and in many cases, it cannot be viewed under normal lighting conditions. Holograms require laser light for both recording and viewing the hologram. The coherent nature of laser light is essential to the formation of holograms. Holograms record light scattered from multiple directions rather than a single direction, as in a photograph. As such, the hologram can be viewed from multiple angles and has apparent depth (Figure 91). Holography should not be confused with lenticular 3D imagery, which interlaces images from multiple angles to create an illusion of depth, or stereoscopic imagery as seen in 3D films.

A hologram is formed by splitting a laser beam into two separate beams, one of which illuminates and scatters off the object to be recorded (Figure 92). A mirror is used to

FIGURE 91

Two photographs of a hologram of a mouse, taken from two different angles.

FIGURE 92



Hologram recording process.

direct the other beam to illuminate a photographic plate that will record the hologram. When the two laser beams reach the photographic plate, the resulting interference pattern is encoded on the plate. This pattern appears random. One cannot determine the image encoded on a hologram simply by looking at the plate, just as one cannot identify what music is recorded on a CD by looking at its surface. The hologram can be viewed by shining a laser beam—identical to the one used to record it—onto the photographic film and viewing the diffraction pattern.

Recent advances in holography methods have resulted in holograms that can be viewed under normal white light without the aid of a laser. The most common version of these holograms is called a *rainbow* hologram because it reflects a rainbow-like spectrum of colors that shifts depending on the viewing angle. Rainbow holograms are often included as a security feature on credit cards, passports, and even international currency. Hungarian-British physicist Dennis Gabor was awarded the 1971 Nobel Prize in Physics for the invention of holography. Gabor devised the method of holography in the late 1940s, years before lasers had been invented.

SECTION IV SUMMARY

- Lasers are used for a variety of applications in scientific, medical, industrial, and commercial fields. The output power and wavelength of a laser determines the kind of applications for which it is most appropriate.
- Laser spectroscopy is an analysis technique in which a laser is used to illuminate the sample being studied in order to determine a precise absorption spectrum. Laser absorption spectroscopy is most often performed using tunable lasers, which can be "swept" across a particular range of emitted wavelengths.
- Optical tweezers are scientific instruments that use a tightly focused beam of laser light to hold microscopic particles stable in three dimensions. Optical tweezers have been used to sort cells and measure the elasticity of DNA molecules.
- A related confinement technique, laser cooling, uses specially configured lasers to selectively slow down a collection of atoms. In 2015, a Stanford research lab reported cooling a rubidium gas to 50 trillionths of a Kelvin.
- Lasers have been applied to nuclear fusion research in a process known as inertial confinement fusion. At the National Ignition Facility, tremendously high-power lasers compress a deuterium-tritium pellet to very high pressures and temperatures for a fleeting instant.
- The most commonly advertised type of laser surgery today, LASIK, uses an argon-fluoride excimer laser to reshape the lens of the eye to correct refractive defects. Lasers have also been widely adopted for cosmetic skin treatments to remove wrinkles, improve skin tone, and remove unwanted hair.
- Many industrial and manufacturing applications rely on high-power lasers to cut, shape, and join materials. The most appropriate type of laser depends on the particular machining task as well as the composition and thickness of the material.
- Photolithography is a fabrication technique for integrated circuits, which are found in most digital devices. Ultraviolet excimer lasers are the current standard for photolithography.
- Laser printers rely on a precisely controlled laser to "draw" places on a rotating drum for electrically charged toner to land. The toner is then transferred to sheets of paper that are fed past the drum.
- A familiar application of laser light is the playing of CDs and DVDs. A laser beam of a specific frequency is used to illuminate "bumps" on an internal layer of the disc, which encodes the information on the disc as a string of bytes.
- Holography is a method of producing a three-dimensional recording of an object using laser light. A hologram is formed by creating an interference pattern from scattered light on a photographic plate, which can be read by a laser that is identical to the one used to record it.

Conclusion

Over the course of this guide, we have worked to develop an understanding of how lasers operate and what makes them such powerful tools across a wide range of applications. We saw how science's understanding of light and other electromagnetic waves has evolved over the course of several centuries, to the point that we are now able to harness and amplify light emitted by individual atoms into powerful beams of radiation for our own purposes. Although there are many different categories of lasers—gas, solid-state, semiconductor diode, and excimer—they all rely on the principle of stimulated emission. The wavelength and output power of a laser determine the types of applications for which that laser is most suitable. Our understanding of how lasers interact with matter has continuously grown in the half century since they were originally invented, leading to new applications with each passing year.

In Section I of the resource guide, we introduced light as a type of electromagnetic wave. We discussed the properties of light and considered how visible light represents the same electromagnetic phenomenon as radio waves and X-rays.

Section II covered ray and wave optics. We discussed how a ray model of light can be used to understand how light changes direction due to reflection and refraction, as well as to predict how images will be formed by lenses and mirrors. We also found that other phenomena such as diffraction, polarization, and interference can only be explained by considering the wave nature of light.

In Section III, we discussed the properties and history of the laser. We started by considering how light is emitted by excited atoms, and then we described various forms of light emission, including incandescence, fluorescence, and phosphorescence. We found that lasers can be modeled using a combination of these emission mechanisms, and we discussed the essential properties of laser light: monochromaticity, directionality, and coherence.

The final section of the resource guide gave an overview of some of the most significant applications of lasers. While lasers have served as an important research tool since their invention, they have also come into their own as powerful medical tools and versatile instruments for precision machining. Now that you've reached the end of the guide, we hope you'll have a deeper appreciation for the physics at work the next time you listen to a CD, watch a Blu-ray disc, or use a laser pointer.

TIMELINE

1676 –	Ole Römer makes the first qualitative measurement of the speed of light.
1690 –	Christiaan Huygens publishes his wave theory of light.
1862 –	Leon Foucault makes an improved measurement of the speed of light.
1865 –	James Maxwell publishes a set of equations describing electric and magnetic fields.
1868 –	Helium is discovered as a new element using evidence from solar spectral lines.
1887 –	Heinrich Hertz produces radio waves in a laboratory.
1887 –	Heinrich Hertz discovers the photoelectric effect.
1896 –	Guglielmo Marconi demonstrates the use of radio waves for communication.
1900 –	Max Planck suggests energy is quantized in discrete units, marking the beginning of quantum theory.
1902 –	Philipp Lenard makes observations of the photoelectric effect that contradict classical theory.
1905 –	Albert Einstein explains the photoelectric effect by theorizing that light is quantized.
1913 –	Niels Bohr proposes a quantized model of the atom.
1917 –	Albert Einstein proposes the concept of stimulated emission.
1938 –	E. H. Land invents the Polaroid.
1950 –	Alfred Kastler proposes the technique of optical pumping.
1951 –	Charles Townes devises the idea for the ammonia maser.
1955 –	The first cesium atomic clock is constructed.
1957 –	Gordon Gould coins the term "laser" in a scientific notebook entry.

1960 –	Theodore Maiman demonstrates the first functional laser, using ruby crystal.
1960 –	The first gas laser, the helium-neon, is demonstrated.
1962 –	Robert Hall demonstrates the first semiconductor diode laser.
1964 –	The Nobel Prize in Physics is awarded to Charles Townes, Nikolay Basov, and Aleksandr Prokhorov.
1969 –	Gary Starkweather develops the idea for laser printing at Xerox.
1969 –	Confocal laser scanning microscopy is first demonstrated.
1970 –	Arthur Ashkin first reports on the ability of laser light to exert forces on microscopic particles.
1974 –	The first product is scanned using a laser barcode reader.
1977 –	Xerox introduces the first commercial laser printer, the Xerox 9700.
1978 –	LaserDisc media format is introduced.
1982 –	Compact discs (CDs) become commercially available.
1984 –	Hewlett-Packard introduces the first desktop laser printer, the HP LaserJet.
1986 –	Arthur Ashkin and colleagues publish an article on the optical tweezing effect.
1995 –	DVD (Digital video disc) format is introduced.
1997 –	Construction begins on the National Ignition Facility (NIF).
1999 –	The LASIK refractive surgery procedure is approved in the United States.
2009 –	The National Ignition Facility is completed.
2013 –	The National Ignition Facility achieves the milestone of extracting more energy from a fusion reaction than was used to trigger it.

Glossary

aberration – a distortion of an image caused by unintended lens or mirror behavior

absorption spectrum – a continuous spectrum interrupted by dark bands, observed when light passing through a substance is absorbed at specific frequencies

active medium – the source of atoms that undergo excitation in a laser; can be a solid, liquid, or gas

amplitude – the maximum displacement from the equilibrium position of a wave

atomic number – a characteristic number for each element, equal to the number of protons in the nucleus of an atom of that element

atomic spectroscopy – the process of analyzing spectral lines to identify the chemical makeup of an excited sample

atoms – the building blocks of the world around us, made up of a positively charged nucleus surrounded by electrons

Brewster's angle – see *polarizing angle*

coherent – describes waves that are in-phase; that is, the peaks of each wave are aligned with one another

collimated – describes a beam of light with very little divergence, or spread

complementary color – the color that remains when another color is subtracted from white light; for example, the complementary color of red is cyan

compound microscope – an optical instrument that uses a combination of two lenses to form an enlarged virtual image of a small object

concave mirror – a mirror shaped such that its reflective surface curves inward, away from incident light; concave mirrors reflect incoming parallel light rays inward toward a focal point

confocal laser scanning microscopy (CLSM) – a technique that uses a laser to reconstruct a three-dimensional rendition of a specimen by scanning back and forth across two-dimensional slices and detecting light emitted through fluorescence

constructive interference – interference between two waves that are in phase, meaning that their crests and troughs align perfectly, resulting in a wave with an overall larger amplitude

converging lens – a lens that refracts incoming parallel light rays so that they meet, or converge, at a single point

convex mirror – a mirror shaped such that its reflective surface curves outward toward the light source; convex mirrors cannot focus light but can create images

critical angle – an angle of incidence above which no light is refracted and all light is reflected back into the material

destructive interference – interference between two waves that are completely out of phase, resulting in the waves canceling each other out

diffraction – the bending of a wave around an edge of a surface or through an aperture

diffraction grating – an optical component consisting of a large number (often thousands per centimeter) of equally spaced parallel slits, which is used to produce interference effects

diffuse reflection – reflection of light from a rough surface, causing reflected rays to be scattered in different directions

diode – an electronic device that allows electric charge to pass through it in only one direction

dispersion – the separation of light into colors due to the dependence of refractive index on wavelength

diverging lens – a lens that causes light rays to spread out, or diverge

Doppler effect – a perceived change in frequency of a wave due to relative motion between the source and receiver

double slit interference – a phenomenon in which light passing through two thin parallel slits produces a pattern of bright and dark fringes on a screen

electromagnetic spectrum – a classification of electromagnetic waves by frequency or wavelength

electromagnetic wave – a wave made up of electric and magnetic fields that oscillate as the wave travels

element – a substance composed of only one kind of atom

emission spectrum – bands of light at specific frequencies, emitted by a source of electromagnetic radiation

excitation – the process of an electron moving from a lower to a higher energy level

excited state – any state of an atom that is higher in energy than its lowest (ground) state

Fermat's principle of least time – optical principle stating that the path that light takes between two points is the path that minimizes the travel time

fluorescence – an emission process in which an atom absorbs a photon and then de-excites by emitting a photon with a lower frequency (i.e., energy)

focal length – the distance from a lens to its focal point

focal point – the point at which parallel light rays meet after passing through a converging lens or being reflected by a concave mirror

frequency – the number of oscillations per unit time of a wave or other repeating event

fusion ignition – the point at which a nuclear fusion reaction becomes self-sustaining

gas-discharge lamp – a light source consisting of a glass tube filled with a noble gas that is excited by an electric discharge between two electrodes

geometric optics – the part of optics in which the ray approximation is valid

ground state – the lowest possible energy state of an atom

holography – a method of producing a three-dimensional recording of an object

Huygens' principle – states that every point of a wavefront can be treated as itself a source of secondary wavelets that spread out in all directions

image – the likeness of an object, formed by light rays that leave that object and are reflected by a mirror or refracted by a lens

incandescence – the production of light due to an object being at a high temperature

index of refraction – a characteristic value of a transparent material, equal to the speed of light in a vacuum divided by the speed of light in the material

interference – the phenomenon of two waves interacting to form a resultant wave

interferometry – the use of superimposed waves to make extremely fine measurements of small displacements, surface irregularities, or changes in refractive index

ion – an atom that has gained or lost electrons; ions have an overall electric charge

laser – a device that emits a beam of monochromatic, coherent light; *laser* is an acronym for "light amplification by stimulated emission of radiation"

laser cooling – the technique of slowing atoms using specially tuned laser light

laser spectroscopy – a branch of spectroscopy in which a laser is used to illuminate the sample being studied

LASIK – a refractive surgery technique wherein the outer layer of the eye is peeled back and the underlying layer reshaped using a laser; *LASIK* is an acronym for "laser-assisted in-situ keratomileusis"

law of reflection – an optical principle stating that for a reflected light ray, the angle of incidence is equal to the angle of reflection

lens – a carefully shaped piece of transparent material, such as plastic or glass, that refracts light rays to produce an image

light-emitting diode (LED) – a type of diode, consisting of a junction between two semiconductor layers, that emits light when an electric current is passed through it

linearly polarized – describes light made up of waves that are polarized along a single axis

longitudinal wave – a wave that is made up of oscillations that are parallel to the direction that the wave travels

mechanical wave – a wave that requires a physical medium through which to propagate

metastable – describes an energy transition with a relatively long lifetime

mirage – a naturally occurring optical illusion due to the refraction of light through the atmosphere

monochromatic – describes light that is made up of a single frequency

Newton's rings – an optical effect made up of concentric bright and dark rings, caused by alternating constructive and destructive interference of light as it reflects from a spherical lens and a flat piece of glass

normal – an imaginary line that is perpendicular to the surface of a material

nuclear fusion – a nuclear reaction in which two light nuclei combine to form a heavier nucleus

nucleus – the dense core of an atom, made up of protons and neutrons

opaque – describes a material that absorbs a particular frequency of light rather than transmitting it

optical cavity – an arrangement of precisely aligned parallel mirrors that sustains stimulated emission, amplifies the laser light, and constrains the photons emitted during laser operation to a narrow beam

optical tweezers – scientific instruments that use a tightly focused beam of laser light to hold microscopic particles stable in three dimensions

optics – a branch of physics involving the properties of light and how light interacts with matter through reflection, refraction, and diffraction

period – the amount of time that elapses between consecutive oscillations of a wave or periodic disturbance

phosphorescence – the emission of light from a metastable state; the emission process takes place over a longer time scale than fluorescence due to the lower probability of de-excitation

photoelectric effect – the emission of electrons from the surface of a material when it is illuminated with light of certain frequencies

photolithography – the process of using high-intensity ultraviolet light, such as that from a laser, to create integrated circuits layer by layer; each layer is made up of a light-sensitive material called a photoresist that is exposed to laser light in a specific pattern

photon – a quantum of electromagnetic radiation that has energy proportional to its frequency

polarization – the plane in which the electric field of an electromagnetic wave oscillates as the wave propagates

polarizing angle – the angle at which reflected light will be completely polarized parallel to the surface of the interface.

Polaroid – a material invented by E. H. Land that can polarize light along a single axis

population inversion – a necessary condition for laser operation; occurs when more atoms within a system occupy a higher energy state than occupy a lower energy state

principal axis – an imaginary line passing at a right angle to the surface of the mirror or lens, through its center

principal rays – three light rays that are particularly useful for constructing ray diagrams because their behavior is easy to predict using the refracting properties of lenses

prism – a piece of glass or plastic typically cut into a triangular block, which can be used to separate white light into a spectrum of colors through dispersion

pump source – the energy source for a laser; the pump source excites atoms in the active medium to a higher energy level

quantum – a discrete unit of energy

quantum number – a number that indexes a particular quantum state of an atom

ray approximation – a simplification of light beams as rays

ray diagram – a diagram showing the paths that light rays take when being reflected or refracted to form an image

ray tracing – the technique of determining or following (tracing) the paths that light rays take in order to determine the position and size of an image

Rayleigh scattering – describes the scattering of light by particles in the atmosphere, which results in high-frequency visible light, such as blue and violet, being scattered most strongly toward our eyes

real image – an image formed by the convergence of actual light rays; a real image can be projected onto a surface

reflecting telescope – an optical instrument that uses a combination of mirrors to collect and focus light, thereby forming an image of a distant object

reflection – the change in direction of a light ray at the boundary between two materials such that it returns into the medium in which it originated

refracting telescope – an optical instrument that uses a combination of lenses to collect and focus light, thereby forming an image of a distant object

refraction – the change in direction of a light ray as it is transmitted between transparent materials with different indices of refraction

resonance – a phenomenon in which a periodic disturbance matches the natural frequency of a resonator, causing its amplitude of oscillation to increase strongly

semiconductor – a material that can be made to conduct electricity under some conditions but not others

Snell's law – an equation relating the angle of incidence and angle of refraction for a refracted ray

spectroscope – a measurement device that separates light emitted by a collection of excited atoms into component wavelengths, used for viewing emission spectra

specular reflection – reflection of light from a smooth surface, resulting in reflected light rays remaining parallel to one another

speed of light – the speed at which all electromagnetic waves propagate in a vacuum, equal to 299,792,458 m/s

spontaneous emission – the random emission of a photon due to an atomic electron moving to a lower energy state

stereolithography – a rapid prototyping technique in which a laser is scanned back and forth across the surface of a liquid as it is gradually lowered on a platform, thereby solidifying the model layer by layer

stimulated emission – photon emission that is triggered by the oscillating electric field from a passing photon with the same frequency; the emitted photon is identical in frequency, phase, and direction to the first photon

superposition – principle stating that when two waves overlap, we can add their amplitudes together at each individual point to find the amplitude of the resulting wave

total internal reflection – an optical phenomenon in which a light ray traveling from a region of higher refractive index to a region of lower refractive index is completely reflected at the boundary

transparent – describes a material that transmits a particular frequency of light rather than absorbing it

transverse wave – a wave that is made up of oscillations that are perpendicular to the direction that the wave travels

virtual image – an image formed by light rays that do not actually converge at the image's location

wave speed – the speed at which a wave propagates, equal to the frequency times the wavelength

wave-particle duality – the observation that light exhibits properties of both particles and waves

wavelength – the distance between successive points on a wave (e.g., the distance between consecutive peaks

Notes

1. This value is exact because the length of the meter is defined using *c*.
2. Dennis F. Vanderwerf, *The Story of Light Science: From Early Theories to Today's Extraordinary Applications*. (Springer International Publishing, 2017) 4.
3. Vanderwerf 6.
4. Vanderwerf 10.
5. Maxwell, J. C. "A Dynamical Theory of the Electromagnetic Field." 1865.
6. Vanderwerf 11.
7. http://earthsky.org/earth/can-you-ever-see-the-whole-circle-of-a-rainbow.
8. S.J. Schechner (2005). "Between knowing and doing: Mirrors and their imperfections in the Renaissance." *Early Science and Medicine* 10(2), 137–162.
9. A. Weis and T. L. Dimitrova. "Wave-particle duality of light for the classroom." http://www.sps.ch/en/articles/progresses/wave-particle-duality-of-light-for-the-classroom-13/.
10. https://www.exploratorium.edu/snacks/cd-spectroscope.
11. M. Egan, "Say Goodbye, Say Hello: GE Stops Making CFLs, Says Go, Go, Go To LEDs." Feb. 1, 2016. http://ge.com.
12. C. H. Townes, *How the Laser Happened* (1999), p. 69.
13. Nobel Prize in Physics 1964, http://nobelprize.org.
14. http://olympus.magnet.fsu.edu/primer/java/confocalsimulator/index.html.
15. T. Kovachy, et al., "Matter Wave Lensing to Picokelvin Temperatures," Phys. Rev. Lett. 114, 143004 (2015).
16. "NIST Launches a New U.S. Time Standard: NIST-F2 Atomic Clock" April 3, 2014. http://www.nist.gov/pml/div688/nist-f2-atomic-clock-040314.cfm.
17. Philip Ball, "Laser fusion experiment extracts net energy from fuel" Feburary 12, 2014. http://www.nature.com/news/laser-fusion-experiment-extracts-net-energy-from-fuel-1.14710.
18. "Looking Back: The History of Laser Vision Correction" July 11, 2013. https://www.lasik.com/articles/laser-vision-correction-history/.

Bibliography

"Bright Idea: The First Lasers." American Institute of Physics, 2017. https://history.aip.org/exhibits/laser/.

Bromberg, Joan Lisa. *The Laser in America, 1950–1970.* The MIT Press, 1993.

Hecht, Jeff. *Beam: The Race to Make the Laser.* Oxford University Press, 2005.

Hecht, Jeff. *Understanding Lasers: An Entry-Level Guide.* 3rd ed., John Wiley & Sons, 2008.

Hecht, Jeff and Teresi, Dick. *Laser: Light of a Million Uses,* Dover Publications, 1998.

Hewitt, Paul G. *Conceptual Physics,* 12th ed. Pearson, 2015.

Hitz, C. Breck, et al. *Introduction to Laser Technology.* 4th ed. John Wiley & Sons, 2012.

Rose, Melinda. "A History of the Laser: A Trip Through the Light Fantastic." Photonics Spectra (online), May 2010. https://www.photonics.com/Article.aspx?AID=42279

Serway, Raymond A. and Faughn, Jerry S. *Holt Physics.* Holt, Rinehart, and Winston, 2006.

Serway, Raymond, et al. *College Physics.* 8th ed. Cengage Learning, 2009.

Townes, Charles H. *How the Laser Happened.* Oxford University Press, 1999.

Vanderwerf, Dennis F. *The Story of Light Science: From Early Theories to Today's Extraordinary Applications.* Springer International Publishing, 2017.