8.3.3 LIGHT, OPTICS AND VISION[M20,M53]

8.3.3.1 The Nature of Electromagnetic Waves

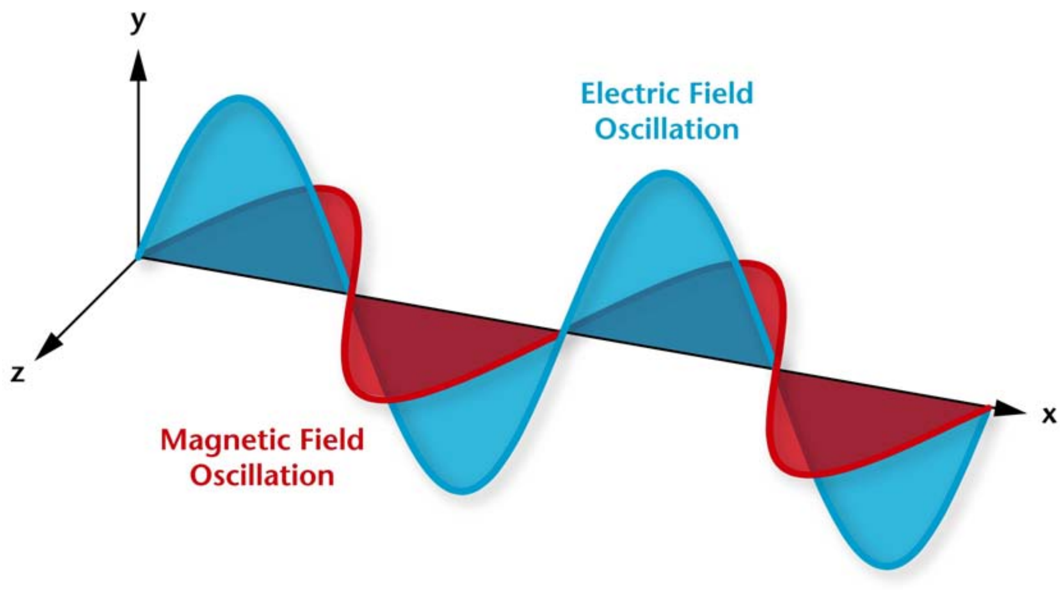
8.3.3.1.1 Electric and Magnetic Fields

We have seen that a wave is the propagation of a disturbance in a medium. All of the waves that we have studied to this point have been mechanical waves, which are only propagated though some physical medium. Electromagnetic waves are different in this respect: they do not require a physical medium for propagation. Instead, as the name suggests, an electromagnetic wave is a disturbance that propagates through an electromagnetic field. While mechanical waves are characterised by the oscillation of particles within the medium through which they propagate, electromagnetic waves are characterised by the oscillation of electric and magnetic field vectors within the electromagnetic field through which they propagate.

Having observed that a changing magnetic field could create an electric field, the Scottish physicist James Clerk Maxwell (1831–1879) showed that a changing electric field also created a magnetic field. A consequence of these observations is that changing electric and magnetic fields trigger each other and these changing fields move, as one, at a speed equal to the speed of light. Maxwell then proposed that light was an electromagnetic wave, the propagation of a disturbance in an electromagnetic field.

8.3.3.1.2 Electromagnetic Waves

Electromagnetic waves are, nonetheless, transverse waves, although individual wave trains comprise both electrical and magnetic components that, like the components of an electromagnetic field, are mutually perpendicular to one another. Any single-frequency electromagnetic wave exhibits a sinusoidal variation in both electric and magnetic fields in space, as illustrated.



3D representation of the oscillating electric and magnetic fields of an electromagnetic wave

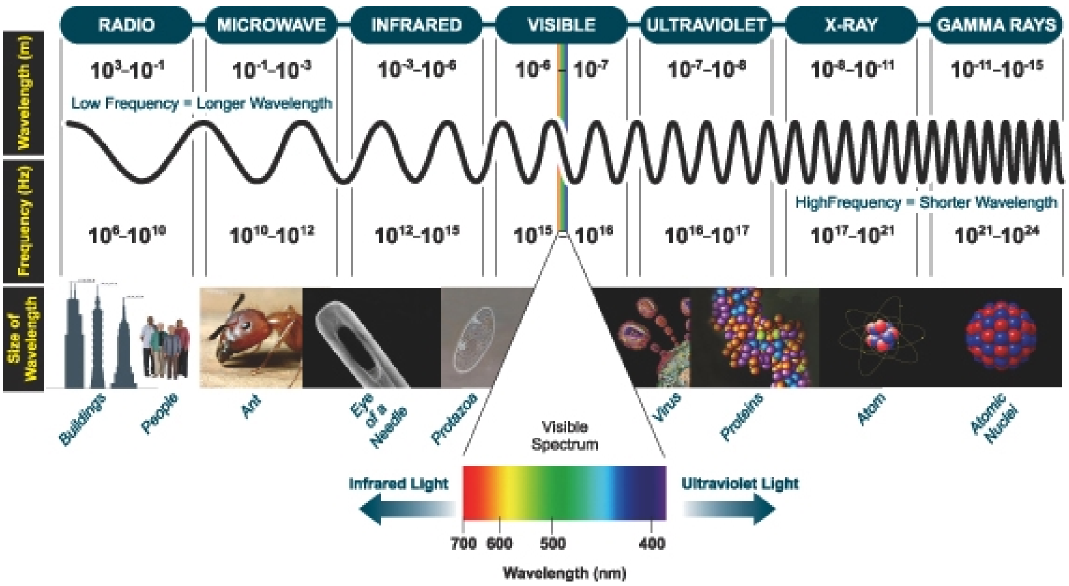
An electromagnetic wave moves or propagates in a direction that is at right angles to the vibrations of both the electric and magnetic oscillating field vectors. By convention, and to simplify illustrations, the vectors representing the oscillating electric and magnetic fields of electromagnetic waves are often omitted (as in the illustration above), although they are understood to still exist.

8.3.3.1.3 Speed of Electromagnetic Waves

All electromagnetic waves travel at the same speed, the speed of light (c). For most practical purposes the speed of light can be taken to be constant—3.00×10^{8} $m{\cdot}s^{-1}$.

8.3.3.1.4 The Electromagnetic Spectrum

Whether transmitted to a radio from the broadcast station, heat radiating from the oven, furnace or fireplace, X-rays of teeth, or the visible and ultra-violet light emanating from the sun, the various forms of electromagnetic radiation all share fundamental wave-like properties. Every form of electromagnetic radiation, including visible light, oscillates in a periodic fashion with crests and troughs, displaying a characteristic **frequency**, **wavelength** and **amplitude** that define the energy, and intensity of the radiation.



The electromagnetic spectrum[1]

The relationship between frequency, wavelength and speed also holds true for all electromagnetic waves:

$$c = f\lambda$$

As a result, and since c is essentially constant, electromagnetic waves can be referred to interchangeably by their respective frequencies and wavelengths, the two being inversely proportional—*i.e.* the greater or longer the wavelength the smaller or lower the frequency, and vice versa. Visible light is usually discussed in terms of its wavelength, while radio waves are usually identified by their frequency.

While much of the discussion hereafter will focus on light, the visible portion of the electromagnetic spectrum, the principles are applicable to all forms of electromagnetic radiation.

8.3.3.1.5 Electromagnetic Wave Intensity

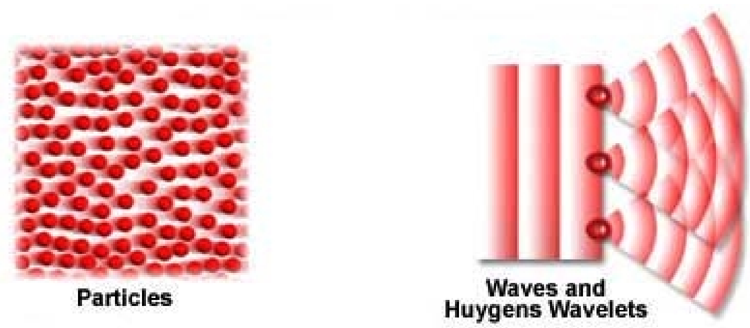
The intensity of electromagnetic waves, and light intensity in particular, is measured in several different ways. Radiant intensity, radiance, irradiance, radiant exitance,

[1] http://www.andor.com/image_lib/introduction/introduction%20(light)/intlight%201%20small.jpg

radiosity, spectral radiance and spectral irradiance are all measurements of the intensity of electromagnetic waves. The two most commonly used, however, are **radiance** ($W \cdot sr^{-1} \cdot m^{-2}$), the measure of projected intensity (coming from a source), and **irradiance** ($W \cdot m^{-2}$), the measure of incident intensity (falling on a surface—*cf.* sound intensity).

8.3.3.2 Wave–Particle Duality of Light[2]

The exact nature of visible light is a mystery that has puzzled man for centuries. Greek scientists from the ancient Pythagorean discipline (*ca.* 500 BC) postulated that every visible object emits a steady stream of particles, while Aristotle (384 BC–322 BC) concluded that light travels in a manner similar to waves in the ocean. Even though these ideas have undergone numerous modifications and a significant degree of evolution over the past 2000 years, the essence of the dispute established by the Greek philosophers remains to this day.



Light as particles and waves

One point of view envisions light as wave-like in nature, producing energy that travels through space in a manner similar to the ripples spreading across the surface of a still pond after being disturbed by a dropped rock. The opposing view holds that light is composed of a steady stream of particles, much like tiny droplets of water sprayed from a garden hose nozzle. During the past few centuries, the consensus of opinion has wavered with one view prevailing for a period of time, only to be overturned by evidence for the other. Only during the first decades of the twentieth century was enough compelling evidence collected to provide a comprehensive answer, and to everyone's surprise, both theories turned out to be correct, at least in part.

In the early eighteenth century, the argument about the nature of light had turned the scientific community into divided camps that fought vigorously over the validity of their favourite theories. One group of scientists, who subscribed to the **wave theory**, centred their arguments on the discoveries of Christiaan Huygens (1629–1695). The opposing camp cited the prism experiments of Sir Isaac Newton (1642–1727) as proof that light travelled as a shower of particles, each proceeding in a straight line until it was refracted, absorbed, reflected, diffracted or disturbed in some other manner. Although Newton, himself, appeared to have some doubt about his **corpuscular theory** on the nature of light, his prestige in the scientific community held so much weight that his advocates ignored all other evidence during their ferocious battles.

Huygens' theory of light refraction, based on the concept of the wave-like nature of light, held that the velocity of light in any substance was inversely proportional to its refractive index. In other words, Huygens postulated that the more light was *bent* or refracted by a substance, the slower it would move while traversing that substance. His

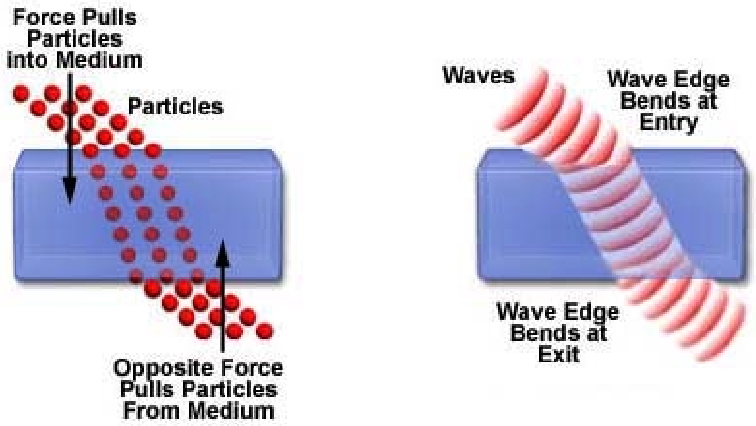
---

2 http://micro.magnet.fsu.edu/primer/lightandcolor/particleorwave.html
http://en.wikipedia.org/wiki/Wave–particle_duality

followers concluded that if light were composed of a stream of particles, then the opposite effect would occur because light entering a denser medium would be attracted by molecules in the medium and experience an increase, rather than a decrease, in speed. Although the perfect solution to this argument would be to measure the speed of light in different substances, air and glass for example, the devices of the period were not up to the task. Light appeared to move at the same speed regardless of the material through which it passed. Over 150 years passed before the speed of light could be measured with a high enough accuracy to prove that the Huygens theory was correct.

Despite the highly regarded reputation of Sir Isaac Newton, a number of prominent scientists in the early 1700s did not agree with his corpuscular theory. Some argued that if light consisted of particles, then when two beams are crossed, some of the particles would collide with each other to produce a deviation in the light beams. Obviously, this is not the case, so they concluded that light must not be composed of individual particles.

Huygens, for all his intuition, had suggested in his 1690 treatise ***Trait de la Lumire*** that light waves travelled through space mediated by the **ether**, a mystical weightless substance, which was supposed to exist as an invisible entity throughout air and space. The search for ether consumed a significant amount of resources during the nineteenth century before finally being laid to rest. In a later volume, ***Huygens' Principle***, he ingeniously described how each point on a wave could produce its own **wavelets**, which then add together to form a wavefront. Huygens employed this idea to produce a detailed theory for the refraction phenomenon, and also to explain why light rays do not crash into each other when they cross paths.



Refraction of particles and waves

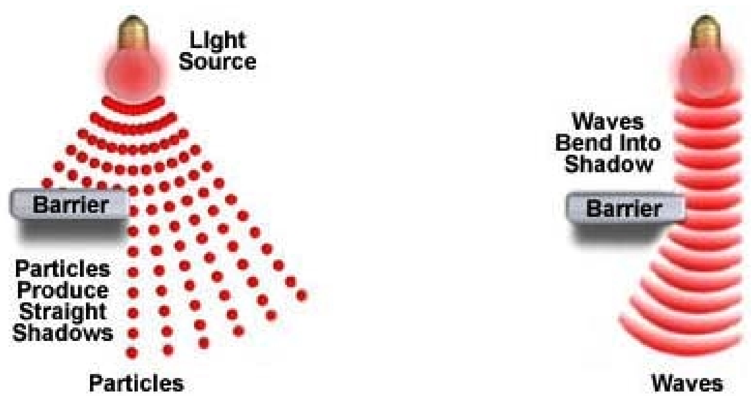
When a beam of light travels between media having different refractive indices, the beam undergoes **refraction**, and changes direction when it passes from one medium to the other. According to Huygens' wave theory, a small portion of each angled wavefront should impact the second medium before the rest of the front reaches the interface. This portion will start to move through the second medium while the rest of the wave is still travelling in the first medium, but will move more slowly due to the higher refractive index of the second medium. Because the wavefront is now travelling at two different speeds, it will bend into the second medium, thus changing the angle of propagation. In contrast, particle theory has a rather difficult time explaining why particles of light should change direction when they pass from one medium into another. Proponents of the theory suggest that a special force, directed perpendicular to the interface, acts to change the speed of the particles as they enter the second medium. The exact nature of this force was left to speculation, and no evidence has ever been collected to prove the theory.

Another excellent comparison of the two theories involves the differences that occur when light is reflected from a smooth, specular surface, such as a mirror. Wave theory speculates that a light source emits light waves that spread in all directions. Upon impacting a mirror, the waves are reflected according to the arrival angles, but with each wave turned back to front to produce a reversed image. The shape of arriving waves is strongly dependent upon how far the light source is from the mirror. Light originating from a close source still maintains a spherical, highly curved wavefront, while light emitted from a distance source will spread more and impact the mirror with wavefronts that are almost planar.

Particles and waves reflected by a mirror

The case for a particle nature for light is far stronger with regard to the reflection phenomenon than it is for refraction. Light emitted by a source, whether near or far, arrives at the mirror surface as a stream of particles, which bounce away or are reflected from the smooth surface. Because the particles are very tiny, a huge number are involved in a propagating light beam, where they travel side by side very close together. Upon impacting the mirror, the particles bounce from different points, so their order in the light beam is reversed upon reflection to produce a reversed image. Both the wave and particle theories adequately explain reflection from a smooth surface. However, the particle theory also suggests that if the surface is very rough, the particles bounce away at a variety of angles, scattering the light. This theory fits very closely with experimental observation.

Particles and waves should also behave differently when they encounter the edge of an object and form a shadow. Newton was quick to point out in his 1704 book **Opticks,** that "Light is never known to follow crooked passages nor to bend into the shadow". This concept is consistent with the particle theory, which proposes that light particles must always travel in straight lines. If the particles encounter the edge of a barrier, then they will cast a shadow because the particles not blocked by the barrier continue on in a straight line and cannot spread out behind the edge. On a macroscopic scale, this observation is almost correct, but it does not agree with the results obtained from light diffraction experiments on a much smaller scale.

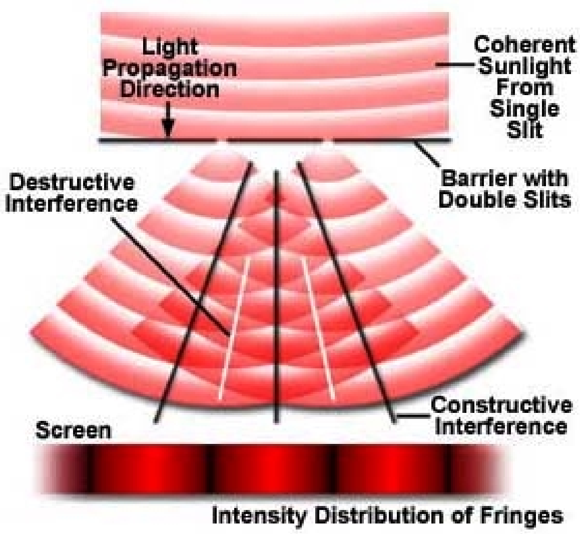


Diffraction of particles and waves

When light is passed through a narrow slit, the beam spreads and becomes wider than expected. This fundamentally important observation lends a significant amount of credibility to the wave theory of light. Like waves in water, light waves encountering the edge of an object appear to bend around the edge and into its geometric shadow, which is a region that is not directly illuminated by the light beam. This behaviour is analogous to water waves that wrap around the end of a raft, instead of reflecting away.

Almost a hundred years after Newton and Huygens proposed their theories, the English physicist Thomas Young (1773–1829) performed an experiment that provided strong evidence for the wave-like nature of light. Because he believed that light was composed of waves, Young reasoned that some type of interaction would occur when two light waves met. In order to test this hypothesis, he used a screen containing a single, narrow slit to produce a coherent light beam (containing waves that propagate in phase) from ordinary sunlight. When the sun's rays encounter the slit, they spread out or **diffract** to produce a single wavefront. If this front is allowed to illuminate a second screen having two closely spaced slits, two additional sources of coherent light, perfectly in step with each other are produced. Light from each slit travelling to a single point halfway between the two slits should arrive perfectly in step. The resulting waves should reinforce each other to produce a much larger wave. However, if a point on either side of the central point is considered, then light from one slit must travel much farther to reach a second point on the opposite side of the central point. Light from the slit closer to this second point would arrive before light from the distant slit, so the two waves would be out of step with each other, and might cancel each other to produce darkness.

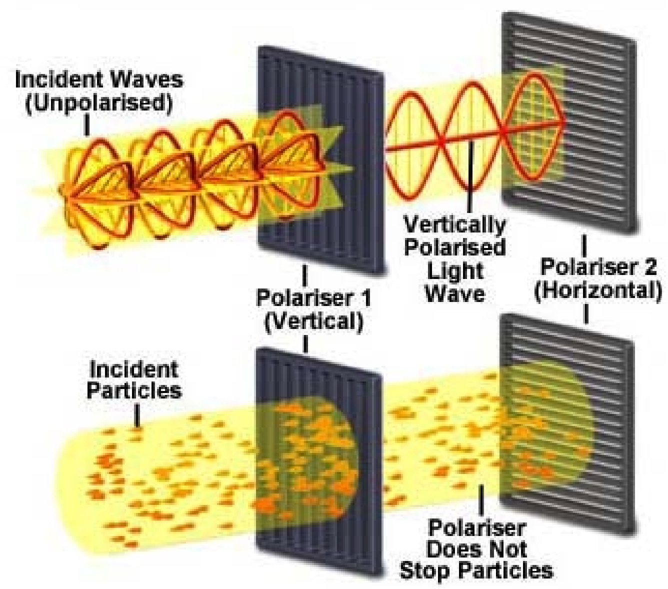
Young discovered that when the light waves from the second set of slits are spread (or diffracted), they meet each other and overlap. In some cases, the overlap combines the two waves exactly in step. However, in other cases, the light waves are combined either slightly or completely out of step with each other. Young found that when the waves met in step, they added together by a process that has come to be termed **constructive interference**. Waves that meet out of step will cancel each other out, a phenomenon known as **destructive** interference. In between these two extremes, various degrees of constructive and destructive interference occur to produce waves having a wide spectrum of amplitudes. Young was able to observe the effects of interference on a screen placed at a set distance behind the two slits. After being diffracted, the light that is recombined by interference produces a series of bright and dark **fringes** along the length of the screen.



Young's Double Slit Experiment

Even more evidence for a wave-like nature of light was uncovered when the behaviour of a light beam between crossed polarisers was carefully examined. Polarising filters have a unique molecular structure that allows only light having a single orientation to pass through. In other words, a polariser can be considered a specialised type of molecular **Venetian blind** having tiny rows of slats that are oriented in a single direction within the polarising material. If a beam of light is allowed to impact a polariser, only light rays oriented parallel to the polarising direction are able to pass through the polariser. If a second polariser is positioned behind the first and oriented in the same direction, then light passing through the first polariser will also pass through the second.

However, if the second polariser is rotated at a small angle, the amount of light passing through will be decreased. When the second polariser is rotated so the orientation is perpendicular to that of the first polariser, then none of the light passing through the first polariser will pass through the second. This effect is easily explained with the wave theory, but no manipulation of the particle theory can explain how light is blocked by the second polariser. In fact, the particle theory did not adequately explain interference and diffraction, effects that were ultimately found to be manifestations of the same phenomenon.

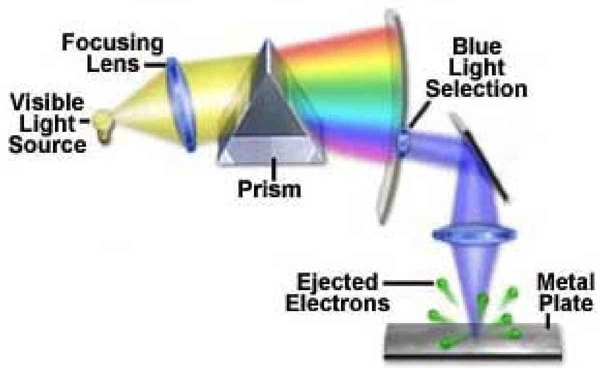


Particles and waves through crossed polarisers

The effects observed with polarised light were critical to the development of the concept that light consists of **transverse** waves having components that are perpendicular to the direction of propagation. Each of the transverse components must have a specific orientation direction that enables it to either pass through or to be blocked by a polariser. Only those waves with a transverse component parallel to the polarising filter will pass through—all others will be blocked.

A major blow to the wave theory, however, occurred behind the scenes in the late 1880s when scientists first discovered that, under certain conditions, light could dislodge electrons from the atoms of several metals. Although at first only a curious and inexplicable phenomenon, it was quickly discovered that ultraviolet light could relieve atoms of electrons in a wide variety of metals to produce a positive electrical charge. German physicist Philipp Lenard (1862–1947) became interested in these observations, which he termed the **photoelectric effect**. Lenard used a prism to split white light into its component colours, and then selectively focused individual bands onto a metal plate. He noticed that, for a specific wavelength of light (blue, for example), the electrons that were ejected produced a constant potential—they each possessed a fixed amount of energy. Decreasing or increasing the amount of light produced a corresponding increase or decrease in the number of electrons liberated, but each still maintained the same energy. In other words, electrons escaping their atomic bonds had energies that were dependent on the wavelength of the light, not its intensity. This is contrary to what would be expected based on the wave theory.

Lenard also noticed a relationship between the wavelength of the incident light, and the energy of the ejected electrons—shorter wavelengths produced electrons with greater amounts of energy.



The Photoelectric Effect

The foundation for a connection between light and atoms was cast in the early 1800s when English chemist and physicist William Hyde Wollaston (1766–1828) discovered that the sun's spectrum was not a continuous band of light, but contained hundreds of missing wavelengths. Over 500 narrow lines, corresponding to missing wavelengths, were mapped by German physicist Joseph von Fraunhofer (1787–1826), who assigned letters to the largest gaps. Later, it was discovered that the gaps were the result of absorption of specific wavelengths by atoms in the sun's outer layer. These observations were some of the first links between atoms and light, although the fundamental impact was not understood at the time.

In 1905, Albert Einstein (1879–1955) postulated that light might actually have some particle characteristics, regardless of the overwhelming evidence for a wave-like nature. In developing his **quantum theory**, Einstein suggested mathematically that electrons attached to atoms in a metal can absorb a specific quantity of light (first termed a **quantum**, but later changed to a **photon**) and thus have the energy to escape. He also speculated that if the energy of a photon were inversely proportional to the wavelength, then shorter wavelengths would produce electrons having higher energies, an hypothesis borne in fact from the results of Lenard's earlier research.

Einstein's theory was solidified in the 1920s by the experiments of American physicist Arthur H. Compton (1892–1962), who demonstrated that photons had momentum, a necessary requisite to support the theory that matter and energy are interchangeable. About the same time, French scientist Louis-Victor de Broglie (1892–1987) proposed that all matter and radiation have properties that resemble both a particle and a wave. Accordingly, de Broglie[3] extrapolated Einstein's famous formula relating mass and energy:

$$E = mc^2 = hf$$

[3] Following a lead from German physicist Max Planck (1858–1947)

where E is the energy of a particle, m the mass, c is the speed of light, h is Planck's constant, and f is the associated frequency. De Broglie's work, which relates the frequency of a wave to the energy and mass of a particle, was fundamental in the development of a new field that would ultimately be utilised to explain both the wave-like and particle-like nature of light. Quantum mechanics was born from the research of Einstein, Planck, de Broglie, Danish physicist Neils Bohr (1885–1962), Austrian-Irish physicist Erwin Schrdinger (1887–1961), and others who attempted to explain how electromagnetic radiation can display what is now termed **duality**, or both particle-like and wave-like behaviour. At times light behaves as a particle, and at other times as a wave. This complementary, or dual, role for the behaviour of light can be employed to describe all of the known characteristics that have been observed experimentally, ranging from refraction, reflection, interference, and diffraction, to the results with polarised light and the photoelectric effect.

8.3.3.3 Light and Colour

When he first observed the dispersion of light through a prism, Isaac Newton identified seven colours in the spectrum that he saw: red, orange, yellow, green, blue, indigo and violet (abbreviated to Roy G. Biv, as a memory aid). The spectrum, however, actually comprises a continuum of colours, from red, with a wavelength of ~700 nm, to violet, with a wavelength of ~400 nm, and the seven that Newton selected were quite arbitrary. In modern usage, indigo is not usually distinguished as a separate colour, and we arbitrarily identify the six remaining points—red, orange, yellow, green, blue and violet (Roy G. Bv)—in the visible spectrum.

When all the wavelengths of the visible spectrum are present together, we perceive the radiation stream as white light, hence the reference to visible light as white light. Technically, white is not a colour at all, but rather the combination of all the colours of the visible spectrum. If none of the wavelengths of the visible spectrum are present in a stream of radiation, we perceive complete darkness, or a totally black environment, although once again black is not actually a colour but merely the absence of any wavelengths in the visible spectrum.

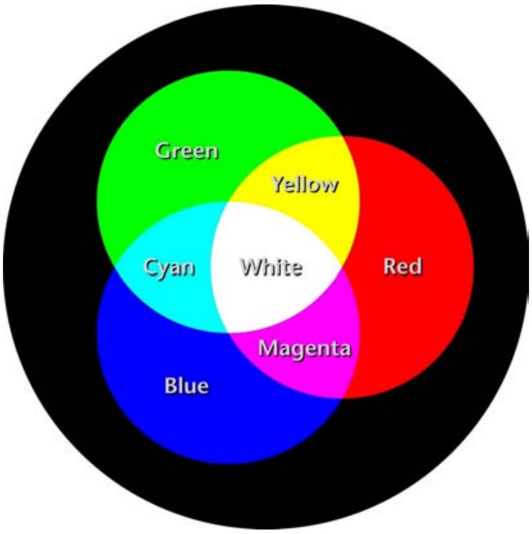
When light of a single wavelength strikes an object, a number of things can happen. The wave could be absorbed by the object, in which case the energy of the wave would be converted to heat; the wave could be reflected by the object; or the wave could be transmitted by and through the object. Rarely however does just a single wavelength of light strike an object. While it does happen, it is more usual that light of many wavelengths or even all visible wavelengths is incident on an object. When this occurs, objects have a tendency to selectively absorb, reflect or transmit light of certain wavelengths. That is, one object might reflect the green wavelengths while absorbing all others. Another object might selectively transmit blue wavelengths while absorbing all others. Whatever the response, the manner in which visible light interacts with an object is dependent upon the wavelength (or frequency) of the light and the nature of the atoms and molecular structure of the object.

The most obvious manifestation of the interaction between light and an object is what we perceive as the colour of the object. The process by which animals perceive colour, however, is not as simple as it might at first seem. It would be quite reasonable to suppose that an object might be perceived to be yellow, for example, if it emitted or reflected light in the yellow part of the visible spectrum, with a wavelength of around 550 nm. But it is not perhaps as obvious that an object that emitted or reflected light from just the red and green parts of the visible spectrum could also appear yellow. Probably even less obvious is the fact that, starting with just three distinct colours, we can produce all the colours of the rainbow.

We will see, when we study the structure of the human eye, that there are very good reasons why we can produce all the colours we can see by mixing different proportions of just three colours, provided these three colours are carefully chosen. For the time being, however, we will consider the two basic mechanisms—interestingly, the mechanism which applies to transmitted light differs from that which applies to reflected light—by which a range of colours can be produced by mixing different amounts of just three colours.

8.3.3.3.1 Additive Colour Mixing

The process by which colours are produced through the transmission of different light wavelengths is known as additive colour mixing. This is the process by which most light sources produce colour, or indeed white light. The more light generated by a light source, the brighter it will appear. If a light source generates all wavelengths in the visible spectrum, the light produced will appear white—adding all the wavelengths together produces white light.



The three primary colours usually used in additive colour mixing are red, green and blue (RGB), and these are known as the **additive primaries**. Note that this is not the only possible set of additive primary colours, it is just the most commonly used. Nonetheless, the various combinations of two of these additive primary colours produces the so-called **secondary colours**, cyan, magenta and yellow, as illustrated, and combinations of various amounts of the three colours produces the range of colours—the **gamut**—available via this process.

It follows that there are combinations of just two colours—red and cyan, for example, a primary colour and its **complimentary** secondary colour—that will produce white light. This is perfectly logical if we remember that cyan is merely the combination of blue and green—the combination of cyan and red is therefore nothing more, or less, than the combination of blue, green and red, which we know to additively produce white. It might be noted from the illustration above that green and magenta, and blue and yellow are also primary/secondary complimentary pairs.

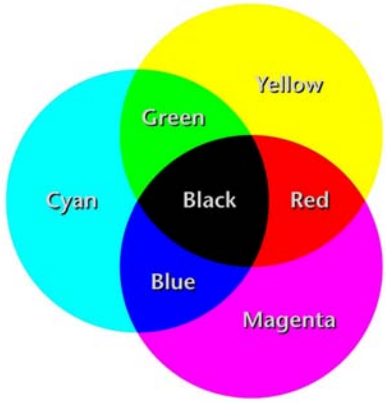
As we will see later, however, these three primary colours, red, green and blue, do not match the colour ranges of the sensors in the human eye perfectly, and are thus not capable of producing the full gamut of colour that we can see.

The most obvious example of additive colour mixing occurs on a television screen or the monitor of a computer. Note that these devices generate light, they do not depend on reflected light. Stage lighting also depends, to some extent, on the principles of additive colour mixing.

8.3.3.3.2 Subtractive Colour Mixing

The process by which colours are produced through the reflection of light is known as subtractive colour mixing. This is the process by which an illuminated object appears a particular colour. The greater the ability of an object to absorb incident light, the less light is reflected. If all light is absorbed, and none reflected, the object will appear black. Ultimately, the colour of an illuminated object will depend on what wavelengths of incident light are absorbed, or *subtracted* from the incident light beam, by the object.

The three primary colours usually used in subtractive colour mixing are cyan, magenta and yellow (CMY), and these are known as the **subtractive primaries**. Note that these are the same colours that comprise the set of secondary additive colours, and like the additive primaries, they are not the only possible set of subtractive primaries. Nonetheless, the various combinations of two of these subtractive primary colours produce the additive primary colours, as illustrated, and combinations of various amounts of the three colours produce the gamut available via the subtractive process. As with the additive colour process described above, since these subtractive primaries are directly related to the additive primaries they do not match the colour ranges of the sensors in the human eye, and are thus not capable of producing the full gamut of colour that we can see.



The most common example of subtractive colour mixing occurs in printed material—the more ink that is deposited on a page, the more incident light is absorbed, and the darker the image. When printing a colour image, the most commonly used process involves the mixing of different amounts of cyan, magenta and yellow ink to create the range of printed colours. Unfortunately, because we cannot produce inks that are absolutely pure in colour, mixing these three inks doesn't actually produce black—the result is more a very dark, muddy brown. To overcome this problem, and because most printed matter is actually text, which is generally black, most printing processes use a fourth ink, black, to achieve a more satisfactory result. This process is commonly referred to as the CMYK process—K for *key*; in this case the key colour is black.

We must also remember that an object can only reflect light that is incident upon it. If the incident light does not contain all visible wavelengths, the 'missing' wavelengths will not be reflected. If a blue object is illuminated with white light, it will effectively absorb the red and green light, and reflect the blue. If the same object is illuminated with red light, it will absorb all the incident light and appear black—under red light, there will be no discernible colour difference between a blue object and a black object.

Similarly, if a red object and a white object are illuminated with white light, the red object will reflect just the red wavelengths (and appear red) while the white object reflects all the incident light (and appear white). If the same two objects are illuminated with just red light, both will reflect all the incident light, and so both will appear to be the same colour—red.

8.3.3.3.3 Blue Skies and Red Sunsets

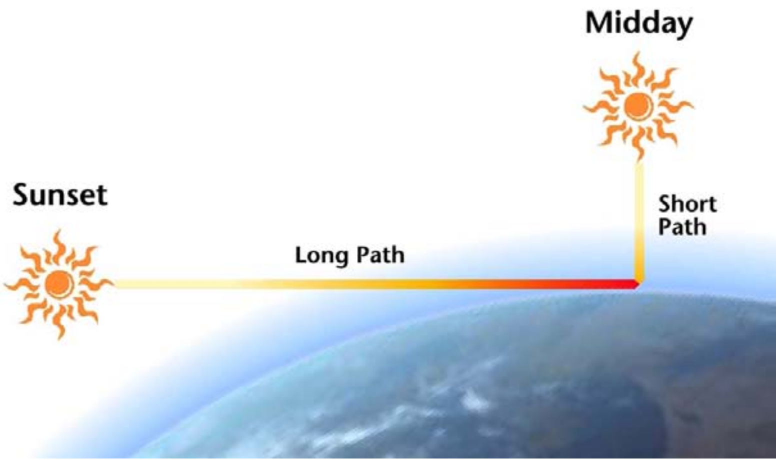
The sun emits waves across the entire electromagnetic spectrum, and in particular across the full visible spectrum. Since all visible wavelengths are present in sunlight, it appears 'white'. This white light illuminates both the Earth and its atmosphere. As discussed above, the interaction of visible light with matter will often result in the absorption of specific wavelengths of light. Those that are not absorbed are either transmitted (by transparent materials) or reflected (by opaque materials). As we look at various objects in our surroundings, the colours that we perceive are dependent upon the colour(s) of light that are reflected or transmitted by those objects. The chlorophyll in plants will absorb red and blue wavelengths so that the plants appear green. The pigments in the petals of flowers will absorb some wavelengths so that they take on their characteristic colours. The black asphalt on the road will absorb most wavelengths and appear black (as well as becoming quite hot!).

Perhaps less obvious is the fact that particles in Earth's atmosphere also absorb, reflect and transmit different parts of the visible spectrum. Of course, a rainbow is one of the more spectacular examples of this phenomenon.

The Earth's atmosphere, however, comprises mainly molecules of nitrogen and oxygen. The size of these molecules makes them most effective in scattering the shorter wavelengths of visible light. Violet light is scattered most, followed by blue light, green light, etc. The longer wavelengths tend to pass through the atmosphere without a significant alteration in their direction. This selective scattering results in the sky's being illuminated with light from the violet/blue end of the visible spectrum and we thus see the sky as being blue in colour.

The light that is not scattered in this way passes through the atmosphere. The longer wavelengths of sunlight thus tend to dominate if we view the sun directly at around midday. Further, while sunlight comprises the entire range of wavelengths of visible light, not all frequencies are equally intense. In fact, sunlight tends to be most rich in the yellow wavelengths. For this reason, with the blue wavelengths scattered, the sun appears yellow during the middle part of the day.

The appearance of the sun, however, changes throughout the day. While it may be yellow during the middle part of the day, it gradually changes as it approaches sunset. As the sun approaches the horizon, sunlight must traverse a greater distance through our atmosphere, as illustrated below.



The midday sun appears yellow, while the setting sun appears red

As the path which sunlight takes through the atmosphere increases in length, the light encounters more and more atmospheric particles. The result is that more and more of the light is scattered, and this scattering even starts to affect the yellow wavelengths. During the sunset hours, the light passing through the atmosphere to our eyes tends to be most concentrated with red and orange wavelengths of light and the sunsets take on a reddish-orange hue. Sunsets can become even more stunning with an increase in smoke or dust in the atmosphere, as the larger size of these particles (compared to that of the atmospheric gases) makes them more effective in scattering the longer yellow and orange wavelengths.

8.3.3.4 Plane Mirrors

The visible portion of the electromagnetic spectrum, light, is of particular importance to most animals because the eyes of animals are sensitive to electromagnetic energy in this region. Optics is the branch of physics that describes the behaviour and properties of light, and the interaction of light with matter.

Light that reaches our eyes invariably comes from one of two categories of object: luminous objects or illuminated objects. Luminous objects are objects that generate their own light. Illuminated objects are objects that are capable of reflecting light to our eyes. The sun is an example of a luminous object, while the moon is an illuminated object. Most of the objects in our physical world are not light-generating objects. They make their presence visibly known by reflecting light, and it is only by reflection that these objects can be seen.

Since the process of reflection is so important in this context, we will begin our study of optics by considering some of the characteristics of mirrors, since image formation by a mirror also depends upon the property of reflection.

The simplest mirror is the plane mirror, a mirror with a flat surface. If an object, such as a pencil, is placed at a distance in front of a plane mirror and illuminated, light rays from the pencil[4] will spread out and reflect off the surface of the mirror. To an observer looking at the mirror, these rays appear to come from a location on the other side of the mirror. By convention, the mirror **image** of an object is said to be at this location behind the mirror.

8.3.3.4.1 Image Characteristics

The image location is the point in space from which the light reflected by an object appears to diverge. In the case of plane mirrors, the image is said to be a **virtual image**. Virtual images are images that are formed in locations that light does not actually reach. Light does not actually pass through the location on the other side of the mirror, it only appears to an observer as though the light is coming from this location. Whenever a mirror creates an image that is virtual, that image will be located behind the mirror. We will see later that some curved mirrors, in addition to forming virtual images, are also capable of forming a **real image**.

If you view an image of yourself in a plane mirror, you will notice that there is an apparent left-right reversal of the image. That is, if you raise your left hand, you will notice that the image raises what would seem to be its right hand. If you raise your right hand, the image raises what would seem to be its left hand. This is often termed **left-right reversal**. This characteristic becomes even more obvious if one views text in a mirror. For example, the word **Physics** is reversed as illustrated:

Object   Mirror image

Not only does the order of letters appear reversed, the actual orientation of the letters themselves appears reversed as well. While there is an apparent left-right reversal of the orientation of the image, however, there is no top-bottom vertical reversal. If you *stand on your feet* in front of a plane mirror, the image does not *stand on its head*. Similarly,

---

[4] This light will actually come from the illumination source, and simply be reflected off the pencil, but for the purpose of the present discussion, we can assume that the light rays originate from the pencil.

the *ceiling* does not become the *floor*. The image is said to be **upright**, as opposed to **inverted**.

This apparent left-right reversal is quite intriguing. Exactly what is happening to cause this word and its letters to be reversed? And why is the reversal observed in the left to right direction and not in the head to toe direction?

To further explore the reason for this apparent left-right reversal, suppose we write the word **Physics** on a transparency and hold it in front of us in front of a plane mirror. If we look at the image of the transparency in the mirror, we would observe the mirror image illustrated above. The letters are written *reversed* when viewed in the mirror. But what if we look at the letters on the transparency? How are those letters oriented? When we face the mirror and look at the letters on the transparency, we observe the unexpected—the mirror image illustrated above. When viewed from the perspective of the person holding the transparency (and facing the mirror), the letters exhibit the same left-right reversal as the mirror image. At least they are reversed when viewed from the perspective of a person who is facing the mirror, a person who *stands behind the transparency*. If you view the transparency from the front, you have a different frame of reference and thus do not see the letters as being reversed. The apparent left-right reversal of an image is simply a frame-of-reference phenomenon.

Another characteristic of mirror images is the relationship between the respective distances of the object and the image from the mirror. For plane mirrors, the object distance is equal to the image distance. That is, the image is the **same distance** behind the mirror as the object is in front of the mirror. If you stand at a distance of 2 meters from a plane mirror, you must focus at a location 2 meters behind the mirror in order to view your image.

Finally, the dimensions of the image are the same as the dimensions of the object. If a 1.8-metre tall person stands in front of a mirror, they will see an image that is 1.8 metres tall. If a coin with a diameter of 18 mm is placed in front of a plane mirror, the image of the coin will also have a diameter of 18 mm. The ratio of the image dimensions to the object dimensions is known as the magnification. Plane mirrors produce images that have a **magnification of 1**.

Thus, plane mirrors produce images that are:

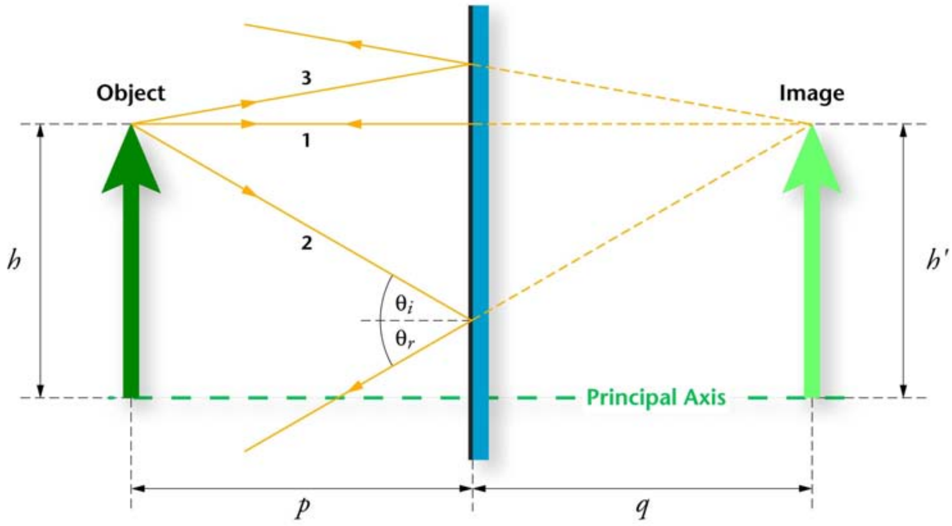
1. **Virtual**;
2. **Upright**; but **left-right reversed**;
3. **The same distance** from the mirror as the object; and
4. **The same size** as the object.

8.3.3.4.2 Ray Diagrams

A ray diagram is a diagram that traces the paths of light rays in order to locate the image of an object that is formed by reflection in a mirror. The construction of a ray diagram is essentially an exercise in geometry.

Typically, when constructing a ray diagram, the object is placed with its base on a Principal Axis, or base line. The base of the image will also be located on this base line, and rays traced from the top of the object are used to locate the top of the image. While they are not generally included in a ray diagram, a similar set of rays from all other points on the object will define the complete image.

Only two rays are actually required to locate any point on the image, but a third is often included to verify the calculation.



Ray diagram for image location with a plane mirror

The first ray is drawn from the object, perpendicular to the surface of the mirror. Because the angle of incidence is 0°, the ray will reflect back on itself. This is ray **1** in the illustration above.

For a plane mirror, the second ray is simply drawn so that its angle of incidence (θ_i) is something other than 0°. In accordance with the Law of Reflection, the angle of reflection (θ_r) is equal to the angle of incidence. Either of the rays **2** or **3** in the illustration above would satisfy this requirement.

Next, the reflected rays are extended until they intersect. If the reflected rays are diverging, as they will always be in the case of a plane mirror, they should be extended backwards, behind the mirror, to the point where they intersect. When rays are extended behind a mirror, they are represented as broken lines, indicating that they are not real rays, but rather *virtual* extensions of the real rays.

A third ray (*cf.* ray **3** or **2** above) can be drawn to confirm the point of intersection.

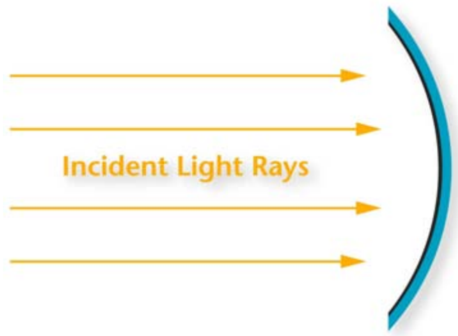
In the illustration above, the point where the extended rays intersect defines the location of the top of the *virtual* image. In the case of a plane mirror, the height of the image (h') is equal to the height of the object (h), and the distance from the image, behind the mirror (q) is equal to the distance from the object to the mirror (p).

8.3.3.5 Spherical Mirrors

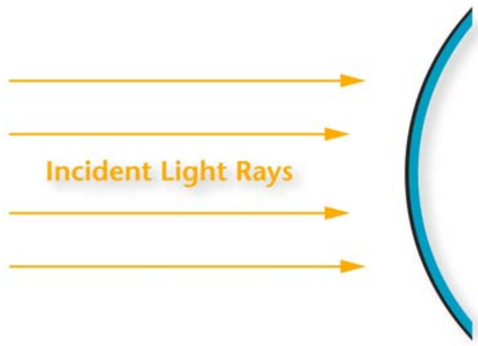
A spherical mirror is a mirror in the form of a slice of a spherical surface. There are thus two kinds of spherical mirror:

- A **concave** mirror, which curves inward like the hollow inside a sphere (it appears *caved-in*);
- A **convex** mirror, which is curved outward, like the outside of a sphere.

The most common examples of concave mirrors are shaving or make-up mirrors—they magnify objects placed close to them. The most common example of a convex mirror is the passenger-side wing mirror on cars—it has a wider field of view than a plane mirror, and reflected objects look smaller and further away than they really are.



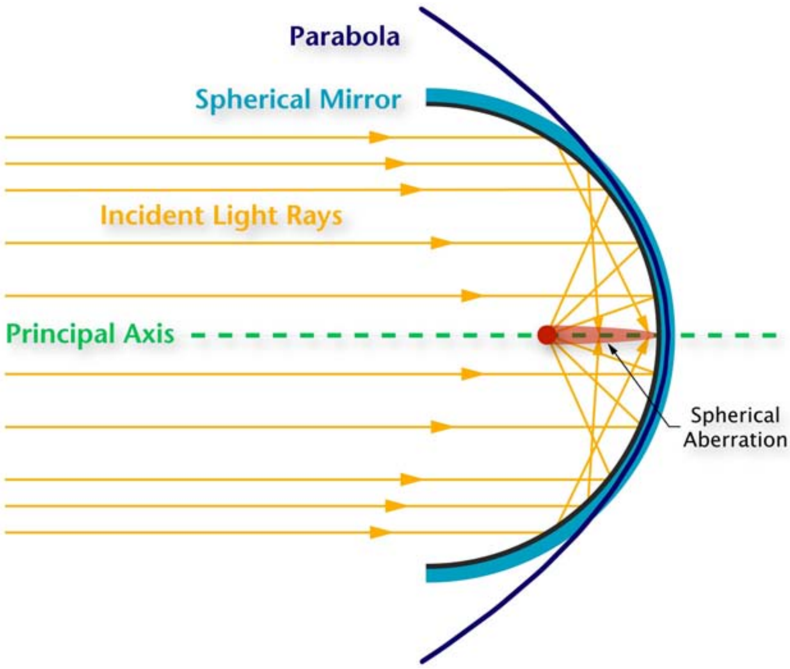
Concave mirror



Convex Mirror

8.3.3.5.1 Properties of Spherical Mirrors[5]

Since the shape of a sphere is very close to that of a parabola in the region around its principal axis, a spherical mirror will behave like a parabolic reflector for light rays that strike the mirror close to its principal axis. As incident light rays move further away from the principal axis, their reflected rays diverge from the focus, producing an effect known as **spherical aberration**.



Close to its principal axis, a sphere approximates the shape of a parabola

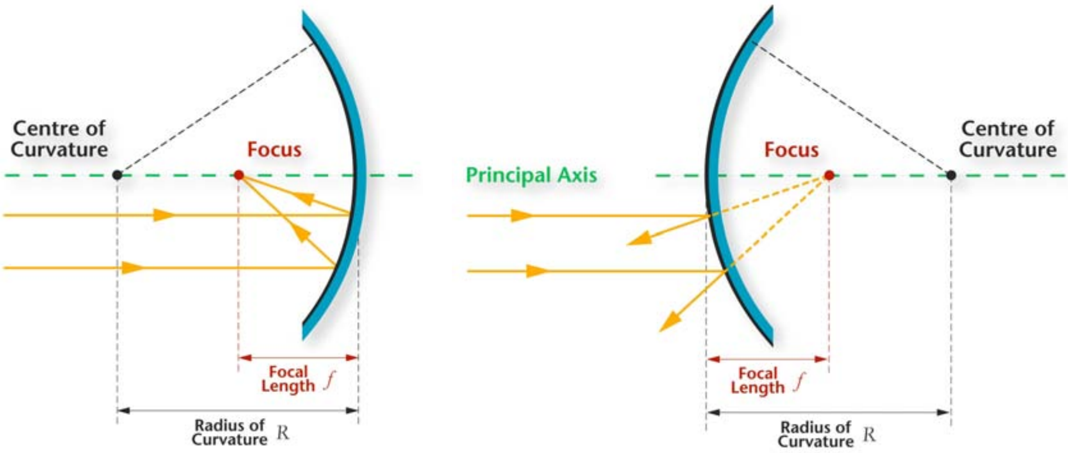
Parabolic reflectors, however, are much more difficult, and thus more expensive, to manufacture than spherical ones. As a result, parabolic reflectors are only used where the spherical aberration of a (more economical) spherical reflector would be a serious problem, as for example, in the receiving dish of a radio telescope or the headlight of a car.

The focal length *f* of a spherical mirror is given by the equation:

$$f = \frac{R}{2}$$

[5] http://farside.ph.utexas.edu/teaching/302l/lectures/node136.html

where R is the radius of curvature of the mirror. This is equally applicable for both concave and convex mirrors, as illustrated below. Note, however, that in the case of a concave mirror, the Focus and Centre of Curvature are in front of the mirror, while in a convex mirror they are behind it.



Concave mirror Convex mirror

8.3.3.5.2 Ray Diagrams

As with plane mirrors, we can determine the positions and sizes of images formed by spherical mirrors geometrically by drawing Ray Diagrams. Ray Diagrams for spherical mirrors are a little more complex than those for plane mirrors, and are constructed as follows:

1. Draw the spherical mirror surface of radius **R**. Draw a line, through **C** (the Centre of Curvature), which crosses the spherical surface near its centre. This is the Principal Axis. Mark **C** and **F** (the focal Point) along the Principal Axis;
2. Draw the object in front of the mirror surface. Draw an incident ray parallel to the Principal Axis from a point on the object to the mirror, and a reflected ray from the mirror through **F**. (The reflected ray, or an extension of the reflected ray must pass through **F**);
3. Draw a second incident ray through **F**, and a reflected ray parallel to the Principal Axis. (The incident ray, or an extension of the incident ray must pass through **F**);
4. The intersection of the two reflected rays, or, for divergent rays, the intersection of their backward extensions, marks the position of the image;
5. To check the accuracy of your drawing, draw a third ray through the Centre of Curvature, **C**. (The ray or its extension must pass through **C**). This ray reflects back onto itself. The intersection of all rays marks the image of the chosen point.

Summary of Rules for Drawing Reference Rays

Ray	Line drawn from Object to Mirror	Line Drawn from Mirror to Image after Reflection
1	Parallel to Principal Axis	Through Focal Point **F**
2	Through Focal Point **F**	Parallel to Principal Axis
3	Through Centre of Curvature **C**	Back along itself through **C**

The relationship between the location of an object, the location of its image, and the position of the focus of the spherical mirror is given by the Spherical Mirror Equation:

$$\frac{1}{\textbf{Object Distance}} + \frac{1}{\textbf{Image Distance}} = \frac{1}{\textbf{Focal Length}}$$

where all distances are measured horizontally from the point where the Principal Axis intersects the mirror. By convention, and in the Ray Diagrams that follow, the pronumerals p, q and f are used to identify the Object Distance, Image Distance and Focal Length respectively, so that the above equation is often written simply as:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

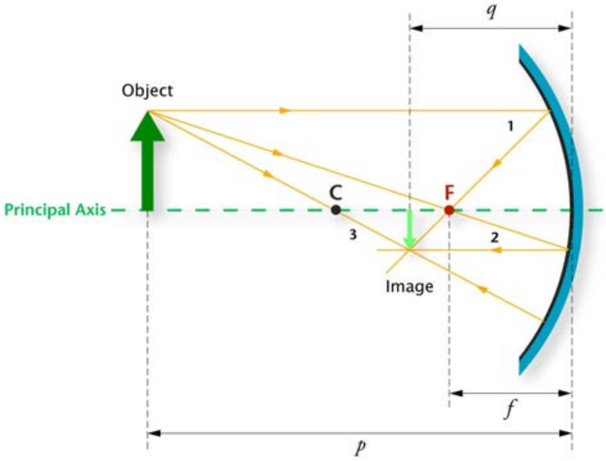
Concave Mirrors

When an object is reflected in a concave mirror, a number of cases can arise.

Case 1a: Object beyond both the Focal Point and Centre of Curvature

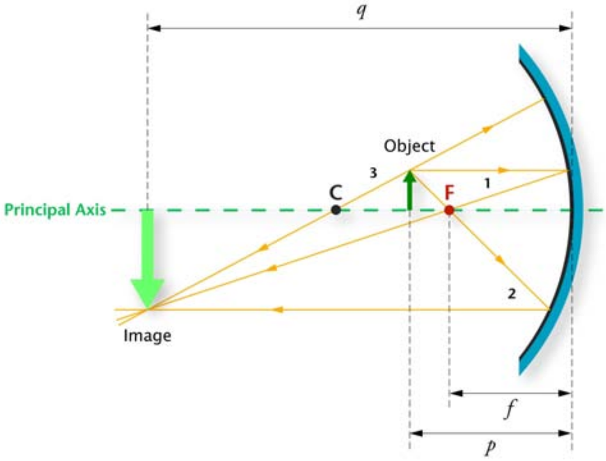
When the object is beyond the Centre of Curvature (**C**), the image is real, inverted, and smaller than the object.

When the object is at the Centre of Curvature (**C**), the image is also at the Centre of Curvature, it is real, inverted, and the same size as the object.



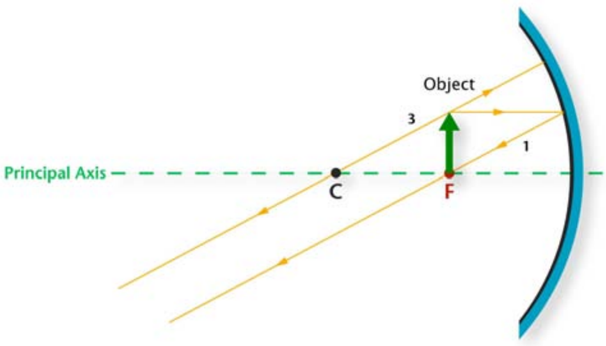
Case 1b: Object between the Focal Point and Centre of Curvature

When the object is between the Focal Point (**F**) and the Centre of Curvature (**C**), the image is real, inverted, and larger than the object.



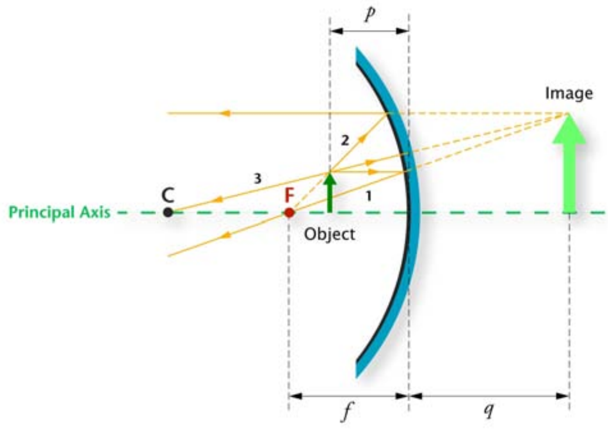
Case 2: Object at the Focal Point

When the object is at the Focal Point (**F**), the image is infinitely far away, and thus is not visible.



Case 3: Object between the Focal Point and the Mirror

When the object is between the Focal Point (F) and the face of the mirror, the image is virtual, upright and larger than the object.

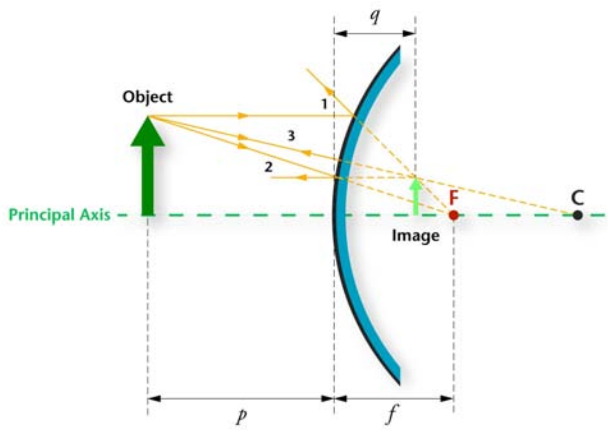


Summary of Image Characteristics for Concave Mirrors

Object Location	**Magnification**	**Attitude**	**Type**	**Position**
Near infinity	<-1	Inverted	Real	At F
Beyond C	<-1	Inverted	Real	Between F & C
At C	-1	Inverted	Real	At C
Between C and F	>-1	Inverted	Real	Beyond C
At F	undefined			
Between F and O	>+1	Upright	Virtual	Behind the mirror

Convex Mirrors

When an object is reflected in a convex mirror, the image is always virtual, upright and smaller than the object.



Magnification

The magnification provided by a spherical mirror is given by the equation:

$$\textbf{Magnification} = \frac{\textbf{Image Height}}{\textbf{Object Height}} = -\frac{\textbf{Image Distance}}{\textbf{Object Distance}}$$

or

$$M = \frac{h'}{h} = -\frac{q}{p}$$

Sign Conventions

By convention, the **focal length** of a **concave mirror is positive**, while that of a **convex mirror is negative**. This and other sign conventions applicable to the mirror equations are listed in the following table.

	f	*p*	*q*	M
Positive	Concave mirror	Object in front of mirror	Image in front of mirror	Image upright
Negative	Convex mirror	*not applicable*	Image behind mirror	Image inverted

8.3.3.5.3 Real and Virtual Images[6]

When light rays are reflected in a mirror, we have noted that, depending on the type of mirror and the location of the object, the image that is formed may be either real or virtual (*cf.* plane mirrors, which only produce virtual images).

A **real image** is an image in which the outgoing rays from a point on an object pass through a single point. It is easiest to observe real images when projected on an opaque screen.

A **virtual image** is an image in which the outgoing rays from a point on an object never actually intersect at a point. A simple example is a plane mirror where your image is perceived to be at twice the distance from you to the mirror. That is, if you are half a metre in front of the mirror, your image will appear at a distance of half a metre behind the mirror.

In drawings of optical systems, virtual rays are conventionally represented by dotted lines. Optical rays represent paths on which light actually travels. A virtual ray (the dotted lines) represent perceived paths as seen by an observer looking into the optical device. The light rays do not travel on these dotted paths.

8.3.3.6 Lenses

A lens is a component of glass or transparent plastic material, usually circular in diameter, which has two primary surfaces that are ground and polished in a specific manner designed to produce either a convergence or divergence of light passing through the material.

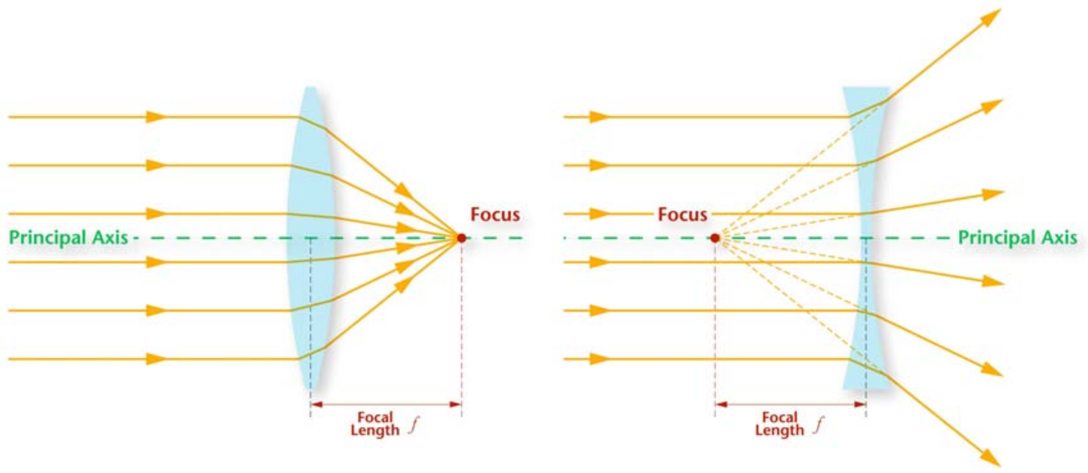
When light rays are reflected from an object and pass through a lens, an image of the object is formed. Unlike a mirror however, where image formation depends upon the property of reflection, image formation by a lens depends upon the property of refraction.

8.3.3.6.1 Types of Lenses

Lenses are classified by the curvature of the two optical surfaces. There are two basic types of lens—converging lenses, also generally known as convex lenses, and diverging lenses, also known as concave lenses. A beam of light travelling parallel to the optical (Principal) axis of a convex lens will be focused to a single point on the axis beyond the lens. This point is known as the **focus** of the lens. The distance from the centre of the lens to the focus is known as the **focal length** of the lens.

The same beam of light will diverge when passing through a concave lens. The divergent beam will nonetheless appear to be emanating from a particular point on the axis on the other side of the lens. This point is also known as the focus, and the distance from the centre of the lens once again the focal length, although in this case the focal length is usually considered to be negative.

6 http://en.wikipedia.org/wiki/Virtual_image



Convex (converging) lens Concave (diverging) lens

8.3.3.6.2 Ray Diagrams

As with mirrors, we can determine the positions and sizes of images formed by lenses geometrically by drawing Ray Diagrams. Ray Diagrams for lenses are constructed as follows:

1. Draw the lens and mark the Focal Point (**F**) and a distance of twice the focal length (**2F**) along the Principal Axis, on each side of the lens;
2. Draw the object to the left of the lens. Draw an incident ray parallel to the Principal Axis from a point on the object (typically the top) to the lens, and a refracted ray from the lens through the point **F**, on the opposite side. If the ray is diverging it should be extended back onto the near side point **F**;
3. Draw a second incident ray, from the same point on the object, through the centre of the lens;
4. Draw a third incident ray, again from the same point on the object, through **F**, and a refracted ray parallel to the Principal Axis. Once again, the incident ray, or an extension of the incident ray must pass through one of the points **F**;
5. The intersection of the three refracted rays, or, for divergent rays, the intersection of their backward extensions, marks the position of the image;

Summary of Rules for Drawing Reference Rays

Ray	Line drawn from Object to Lens	Line Drawn from Lens to Image
1	Parallel to Principal Axis	Through Focal Point **F**
2	Through Centre of Lens	Through Centre of Lens
3	Through Focal Point **F**	Parallel to Principal Axis

Again, in the same way that the relationship between object, image and focal length is given by the mirror equation, a similar relationship exists with thin lenses[7]. This relationship, between the location of an object, the location of its image, and the position of the focal point of a lens is given by the Thin-Lens Equation:

$$\frac{1}{\text{Object Distance}} + \frac{1}{\text{Image Distance}} = \frac{1}{\text{Focal Length}}$$

[7] A thin lens is a lens with a thickness that is negligible compared to its focal length.

where all distances are measured horizontally from the centre of the lens. Again, by convention, and in the Ray Diagrams that follow, the pronumerals p, q and f are used to identify the Object Distance, Image Distance and Focal Length respectively, so that the above equation is often written simply as:

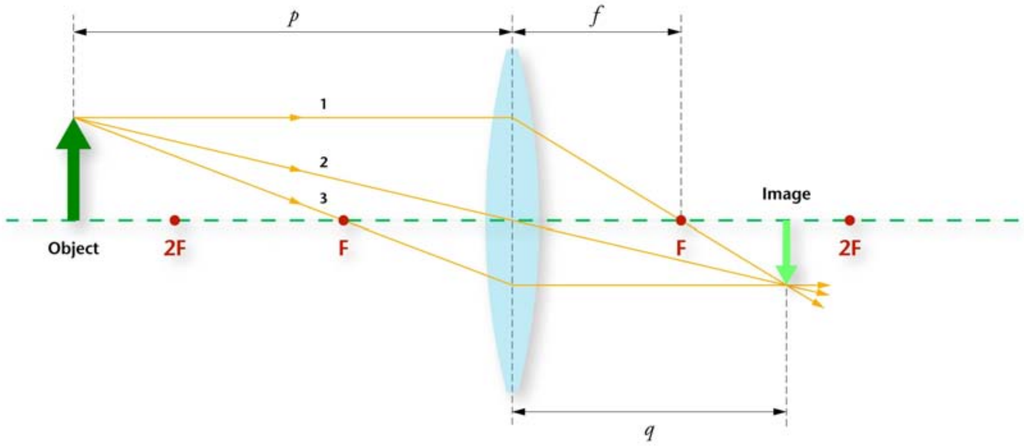
$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

Convex Lenses

When an image of an object is formed through a convex lens, a number of cases can arise.

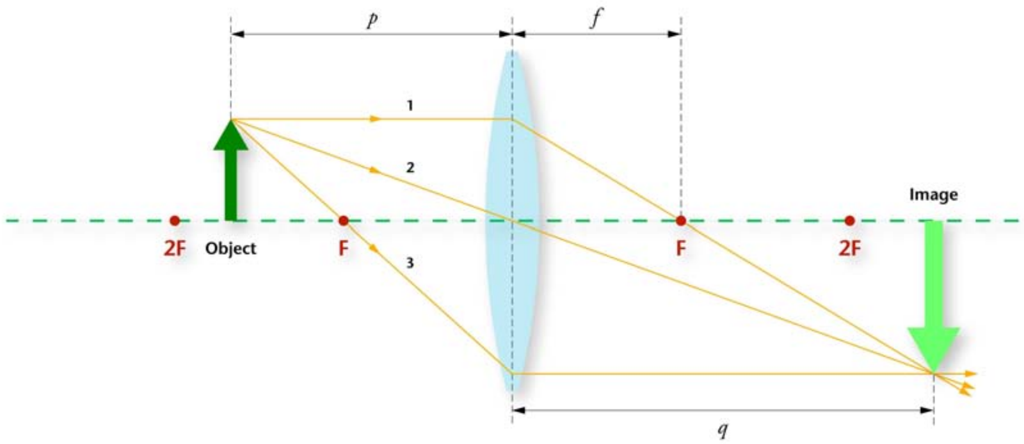
Case 1a: Object beyond twice the Focal Length

When the object is further than twice the Focal Length from the lens, the image is real, inverted, and smaller than the object.



Case 1b: Object between the Focal Point and Centre of Curvature

When the object is between twice the Focal Length (**2F**) and the Focal Point (**F**), the image is real, inverted, and larger than the object.

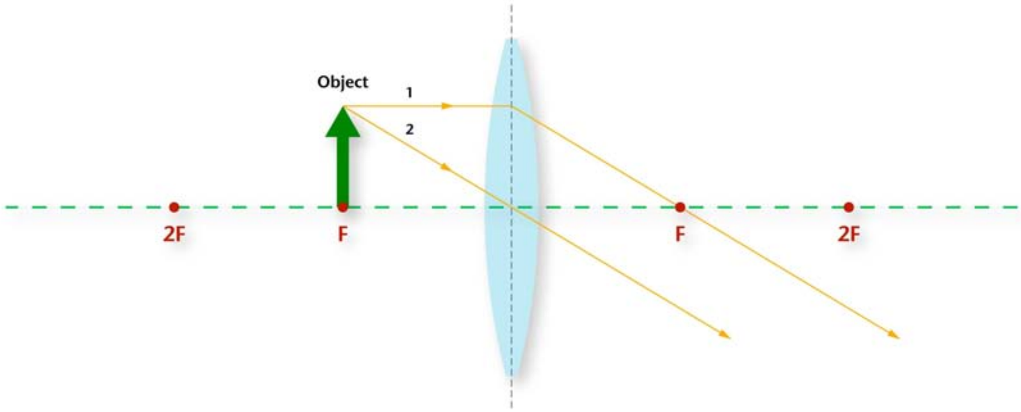


Convex Lens—Object distance greater than, but less than twice, the Focal Length
Image is Real, Inverted and larger than Object

If the object distance is twice the Focal Length (**2F**), the image distance is also twice the Focal Length (**2F**)—the image is real, inverted and the same size as the object.

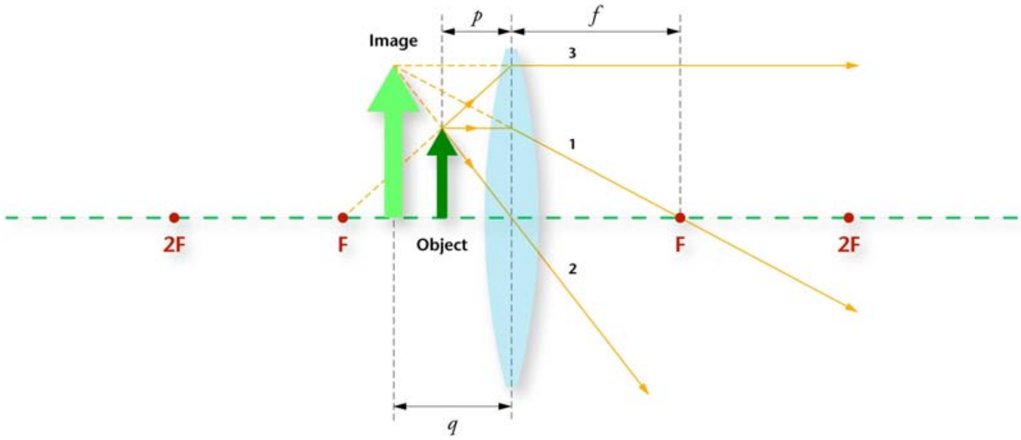
Case 2: Object at the Focal Point

When the object is at the Focal Point (**F**), the image is infinitely far away, and thus is not visible.



Case 3: Object between the Focal Point and the Mirror

When the object is between the Focal Point (**F**) and the face of the lens, the image is virtual, upright and larger than the object.

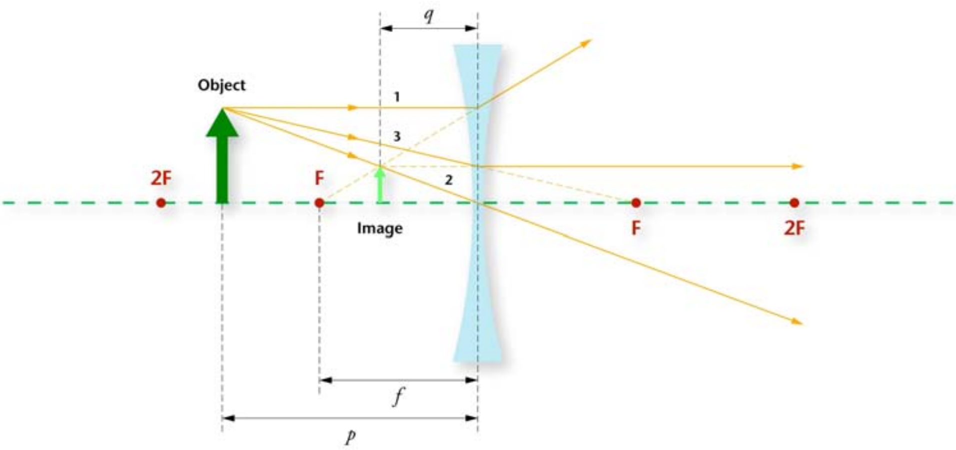


Summary of Image Characteristics for Convex Lenses

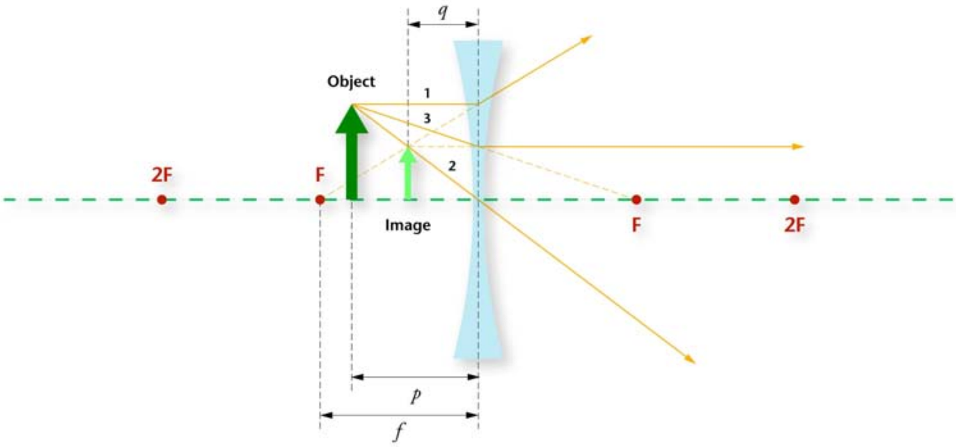
Object Location	**Magnification**	**Attitude**	**Type**	**Position**
Near infinity	<-1	Inverted	Real	At F
Beyond 2F	<-1	Inverted	Real	Between F & 2F
At 2F	-1	Inverted	Real	At 2F
Between 2F and F	>-1	Inverted	Real	Beyond 2F
At F	undefined			
Between F and O	>+1	Upright	Virtual	Same side as object

Concave Lenses

When an image of an object is formed through a concave lens, the image is always virtual, upright and smaller than the object, as illustrated below. The image simply becomes smaller the father away the object is from the lens.



Concave Lens—Object distance greater than the Focal Length
Image is Virtual, Upright and smaller than Object



Concave Lens—Object distance less than the Focal Length
Image is Virtual, Upright and smaller than Object

Magnification

In exactly the same way as for spherical mirrors, the magnification provided by thin lenses is given by the equation:

$$\textbf{Magnification} = \frac{\textbf{Image Height}}{\textbf{Object Height}} = -\frac{\textbf{Image Distance}}{\textbf{Object Distance}}$$

or

$$M = \frac{h'}{h} = -\frac{q}{p}$$

Sign Conventions

By convention, the **focal length** of a **convex lens is positive**, while that of a **concave lens is negative**. This and other sign conventions applicable to the lens equations are listed in the following table.

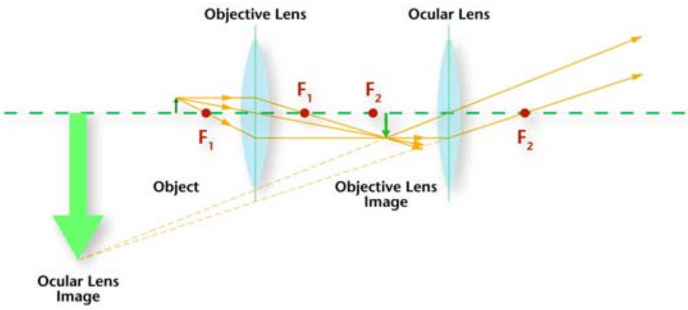
	f	p	q	M
Positive	Convex lens	Object in front of lens	Image behind lens	Image upright
Negative	Concave lens	Object behind lens	Image in front of lens	Image inverted

8.3.3.6.3 Compound Lenses—Microscopes and Telescopes

A compound lens is an array of simple lenses with a common axis. Typically, the use of multiple elements allows for the more effective minimisation of optical aberrations (*e.g.* spherical, chromatic etc.) than is possible with a single element. Compound lens arrangements are used in many optical instruments, although the best known are probably the microscope, telescope and camera. Simple cameras function quite well with a single, simple lens element, while compound lenses are used in more expensive cameras for optimal image quality. Microscopes and telescopes, however, depend on the interaction of at least two lens elements to perform their respective functions.

The Microscope

A simple microscope can be made from two convex lenses. One lens, the **objective lens** is positioned close to the object to be viewed. It is usually a lens with a relatively short focal length (less than 10 mm) that forms an inverted and magnified image within the focal length of the second, **ocular lens**. The ocular lens, or eyepiece lens, acts as a magnifying glass for this real image, producing an enlarged, virtual image from the image provide by the objective lens.

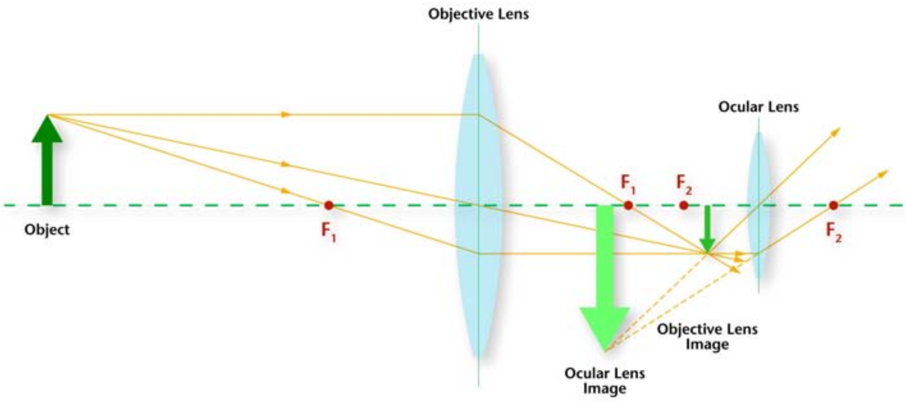


Compound lens arrangement in a microscope
Objective image magnified by Ocular—Shorter focal length Objective

Microscopes are often equipped with two or three interchangeable objective lenses, which provide for different levels of object magnification.

The Telescope

A simple telescope can also be made from two convex lenses. In this case, however, the **objective lens** is usually a lens with a relatively long focal length (more than 100 mm, although astronomical telescopes typically have objective lenses with focal lengths in excess of 1 m). Nonetheless, the objective lens of a telescope also forms an inverted and magnified image within the focal length of the second, **ocular lens**, which performs exactly the same function as the ocular lens in a microscope.



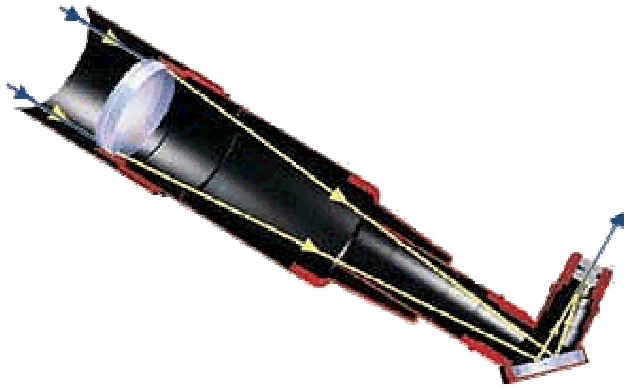
Compound lens arrangement in a telescope
Objective image magnified by Ocular—Longer focal length Objective

Another difference with the telescope is that different levels of magnification, or different powers, are achieved by changing the eyepiece, rather than the objective lens.

The following are illustrations of the three most common types of astronomical telescope in current use.

Classical refracting telescope

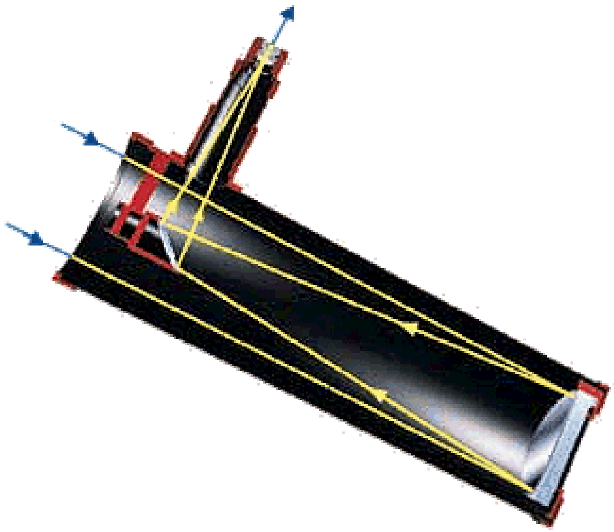
This is the classical telescope design, following the compound lens principles outlined above. The plane mirror at the end of the telescope inverts the image so that it appears erect.



Newtonian reflector telescope

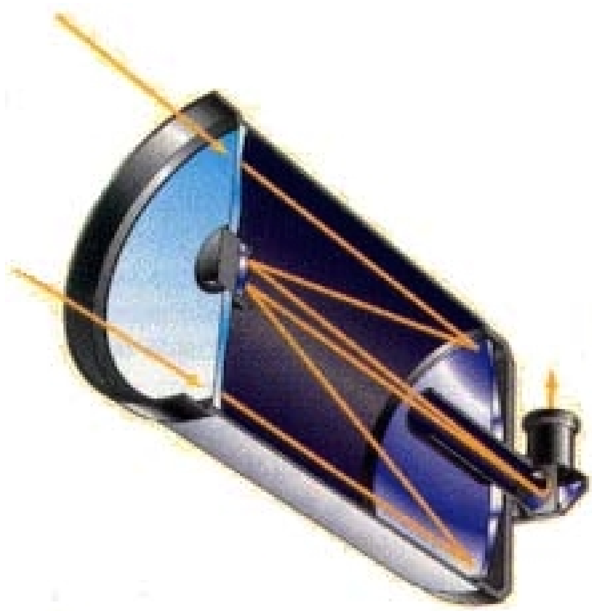
Instead of using a large objective lens, a parabolic mirror is used to bring the image to focus. Unlike lenses, mirrors are not subject to chromatic aberration, and large mirrors are much more economical to manufacture than large precision lenses.

This is an example of a **catadioptric** optical system, one that combines both lenses and mirrors.



Schmidt-Cassegrain reflector telescope

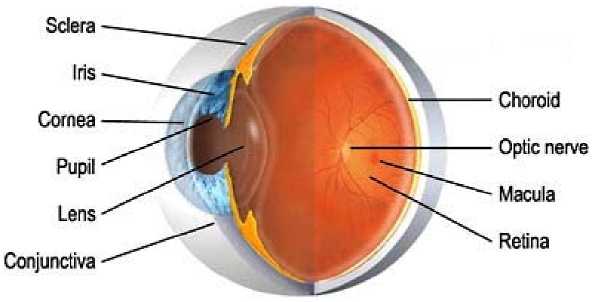
The use of mirrors increases the effective focal length of this instrument. In this case both a primary concave mirror and a secondary convex mirror are employed. The Schmidt design also incorporates a specially ground corrector plate, at the front of the telescope, to correct for spherical aberration in the (relatively low cost, spherical) primary mirror.

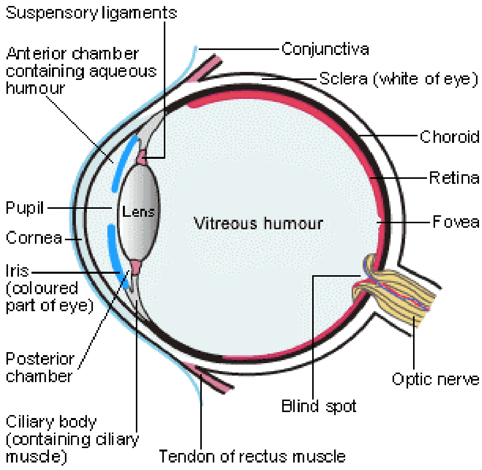


8.3.3.7 The Human Eye

8.3.3.7.1 Eye Anatomy[8]

The primary structural elements of the eye are illustrated below. Generally, these elements can be described in three groups, or layers.





The Outer Layer: The Sclera and Cornea

The **sclera** is the tough, fibrous outer layer of the eyeball that forms the white of the eye. The front of the sclera is covered by the **conjunctiva**, a thin, transparent membrane that is involved in protecting the eye. The conjunctiva also lines the inside of the eyelid.

The **cornea** is a dome-shaped structure at the front of the eye. It is transparent, allowing light to enter the eye, and together with the **lens** helps focus and direct light onto the **retina**.

The Middle Layer: The Uveal Tract (iris, ciliary body and choroid) and Lens

The **iris** is the coloured part of the eye that controls the size of the **pupil**—the black area in the centre of the iris. In bright light, the iris reduces the pupil size to restrict the amount of light entering the eye; in dim light or darkness, the iris opens up the pupil to allow more light in.

The iris sits between the **anterior chamber** and the **posterior chamber** in the front part of the eye. These chambers contain a watery liquid known as **aqueous humour**, which is constantly being produced (by the **ciliary body**) and drained away. Aqueous humour is important for nourishing the lens and cornea.

The lens is a clear, flexible structure that changes shape so that the eye can focus on objects at varying distances. The lens is connected to the ciliary body (which contains the **ciliary muscles**) by **suspensory ligaments**. When the ciliary muscles contract, the normal tension exerted on the lens by the suspensory ligaments is released, and the lens becomes thick and curved, allowing the eye to see close-up objects clearly. When the ciliary muscles relax, the lens becomes thinner, which is necessary for long-range vision.

[8] http://stlukes-eye.com/Anatomy.asp
http://www.mydr.com.au/default.asp?Article=3429

The **vitreous humour** is a jelly-like substance that fills the back portion of the eye behind the lens. As well as helping the eye keep its shape, this clear gel transmits light to the back of the eye.

The **choroid**, a membrane found between the sclera and the retina, lines the back of the eye. It contains many blood vessels that supply oxygen and nutrients to the retina, and is highly pigmented to help absorb light and prevent scattering.

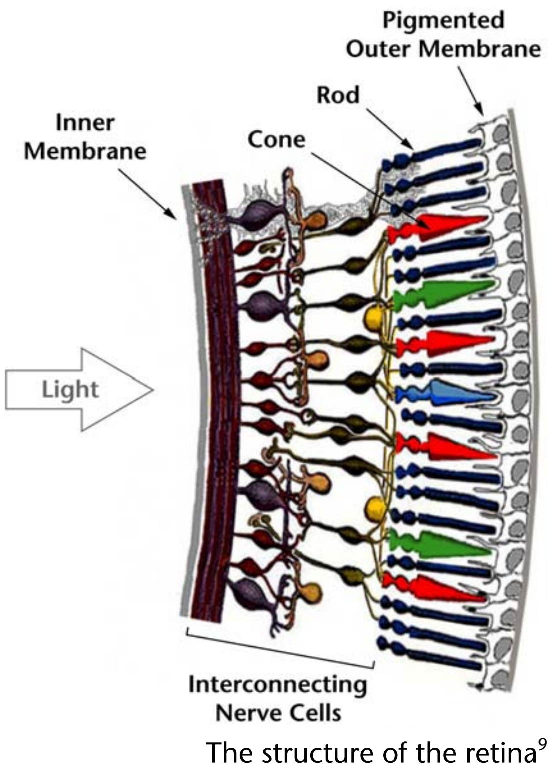
The Inner Layer: The Retina

The retina lines the inside of the back part of the eye, and is the light-sensitive part of the eye. The retina contains millions of cells known as photoreceptors, and each photoreceptor is linked to a nerve fibre. There is a **blind spot**, also known as the optic disc, at the back of each eye where all of these nerve fibres converge to form the **optic nerve**. This blind spot, however, is not usually noticed, because objects that fall on the blind spot of one eye are seen by the other eye.

When an image is detected by the photoreceptors, the relevant information is converted into nerve impulses that are sent to the brain via the optic nerve. The **macula** is a small area of the retina that contains a particularly high concentration of photoreceptors, and is important for sharp central vision. The middle part of the macula, the **fovea**, is the most sensitive area, providing the sharpest vision.

8.3.3.7.2 The Eye's Response to Colour

The retina of the human eye contains two types of photoreceptors: **rods**, which are the most sensitive photoreceptors but are effectively devoid of colour sensitivity, and **cones**, which, although less sensitive than rods, provide us with our ability to distinguish between different colours. In low light conditions, the cones are relatively ineffective and our colour vision is poor. The more sensitive cones, however, still allow us to differentiate between light and dark. As light levels increase, the cones become more effective and we see in vivid colour. Cones are also concentrated in and around the macula, while rods are concentrated more outside this central area of the retina. The result is that, while the macula is important for sharp vision and colour perception, it is our peripheral vision that is most sensitive in low light conditions.



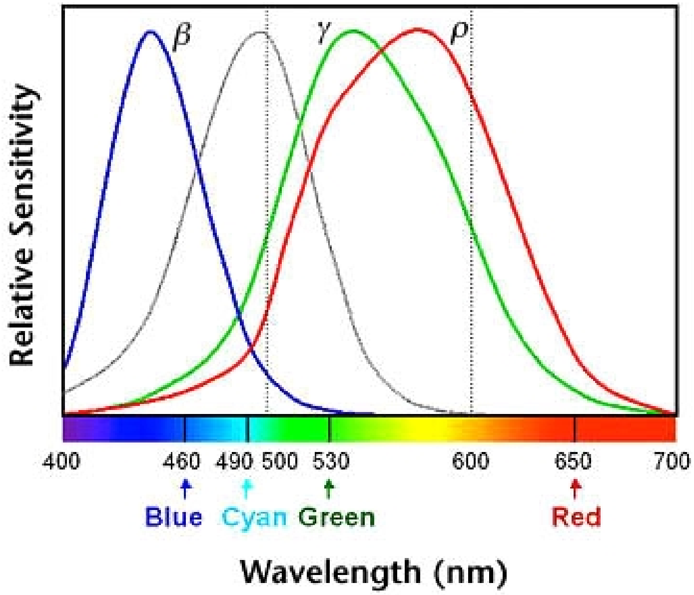
The structure of the retina[9]

The adjacent illustration is a simplified representation of the arrangement of cells in the retina, showing the rods and cones towards the outside. All rods are the same, but the retina contains three different kinds of cone, identified by the Greek letters beta *(β)*, gamma *(γ)* and rho *(ρ)*. The three kinds of cone differ in the wavelengths of light to which they are sensitive, corresponding roughly to blue, green and red wavelengths respectively, and thus provide a means of differentiating

[9] Adapted from: http://webvision.med.utah.edu/sretina.html

between the different wavelengths within the visible spectrum. Our brain interprets different levels of stimulation of the three different kinds of cone as different colours, in the same way that mixing different amounts of the three primary colours produces all the colours of the rainbow.

Although the β and γ sensors do indeed correspond closely to the primary colours blue and green, the ρ sensor is really more sensitive in the yellow-orange part of the spectrum, as illustrated in the spectral sensitivity graph below.



Human spectral sensitivity to colour
The three cone types *(β, γ, ρ)* correspond roughly to Blue, Green, Red[10]
The grey curve is an approximation of the rod sensitivity[11]

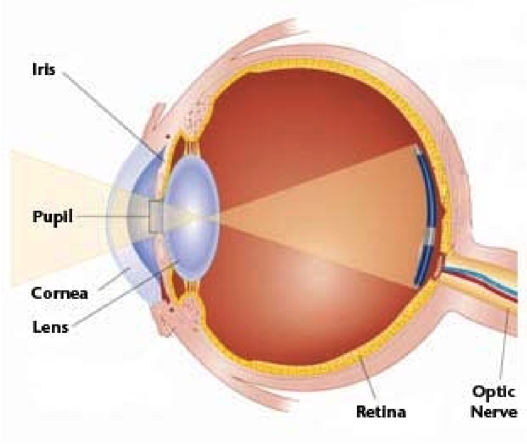
As a result, the full range of human colour vision cannot actually be replicated by mixing the three traditional primary colours, red, green and blue. In fact, because no three pigments or light sources exactly match the sensitivity range of the cones in the retina, *no* combination of three real light sources can duplicate the full **gamut** (colour range) of human vision. Monitors with three [additive] phosphor colours (RGB) have a limited colour gamut; printers with four ink colours (CMYK—the three subtractive colours, cyan, magenta and yellow, and black) have an even smaller gamut. Printers with additional colours—6 to 8 are not uncommon—however, have a larger gamut and are thus better suited to printing, for example, the broad range of colours that are found in photographs.

It is also interesting to note that, while the basic process of colour perception is similar, different animals have different colour sensors. The European starling, for example, has four different cones and the common pigeon five, and none of these are identical to the three human cones. As a result, the colours we see are different from those seen by a starling or pigeon (or any other animal)[12].

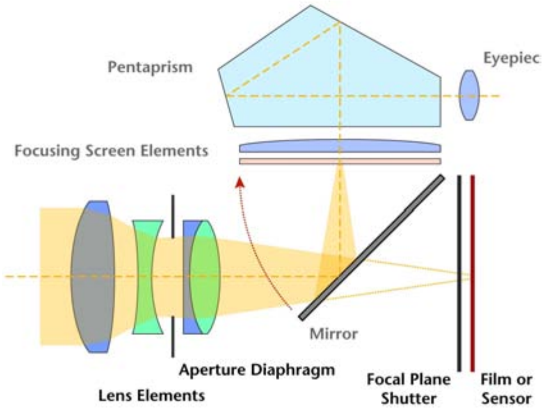
10 http://www.normankoren.com/Human_spectral_sensitivity_small.jpg
11 http://www.yorku.ca/eye/specsens.htm
http://www.marine.maine.edu/~eboss/classes/SMS_491_2003/Week_10.htm
12 Rowe, M.P., Inferring the retinal anatomy and visual capacities of extinct vertebrates, Palaeontologia Electronica 3(1)
http://people.eku.edu/ritchisong/birdbrain2.html

8.3.3.7.3 Eye Function

The function of the eye is often likened to that of a camera and the illustrations on the following page provide a comparison of the essential elements of each.



Cross-section of the human eye[13]



Cross-section of an SLR camera[14]

The cornea and lens of the eye function like the lens elements of a camera. In the eye, the cornea has a geometry that is essentially fixed, and most focusing function is achieved by changing the shape of the lens. In anything other than the most simple cameras, focusing is achieved by moving the individual lens elements back and forth to vary the overall focal length of the lens assembly.

The retina of the eye reacts to the different light intensities and colours that pass through the lens. This is the same function that is performed by the film in a conventional camera, or by the sensor array in a digital camera. The eye, of course, sends its information off to the brain, via the optic nerve, for immediate interpretation. A camera, like the one illustrated below, simply records its information. In this respect, the eye behaves more like a movie or video camera, continually transmitting images to the brain for interpretation.

Although we don't necessarily notice, the brain does not, however, process every piece of information received by the retina. The rods and cones take a certain amount of time to respond to changes, and they *only* respond to changes—if the eye remained perfectly still, and the light falling on the retina did not change at all, the resultant image would quickly fade and disappear. As a result, the brain tends to register changes in images, rather than whole images. If changes occur at a rate of less than about 20 times per second, however, the image will appear to 'flicker' or motion will appear 'jerky', an indication that the brain is capable of processing more information than is being provided. The operation of a camera or video is more obvious in this respect, with the image capture process controlled by the discrete action of the shutter.

The amount of light entering the eye is controlled by the size of the pupil, the hole in the middle of the iris. The aperture diaphragm in a camera performs exactly the same function.

The other camera elements identified in the illustration above simply help the operator to identify the image to be recorded.

Of course, while a camera, even an automatic camera, requires some degree of external operator input for correct operation, the human eye is truly fully automatic.

---

13 http://www.artlex.com/ArtLex/b/images/blindspot.jpg
14 Original artwork from http://en.wikipedia.org/wiki/image:SLR_cross_section.svg

Eye Movement

Movements of the eye is controlled by a set of so-called extraocular muscles. These muscles run from the bones of the orbit to the sides of the eyeball. Each eye has six muscles to help it move in different directions—the lateral, medial, superior and inferior **rectus muscles**, and the superior and inferior **oblique muscles**. People who have a squint (strabismus) may have a problem with one of these muscles or the nerves that control it.

Eye Protection

Your eyes sit in sockets within the bones of your skull (known as the **orbits**), and are surrounded by fat, fibrous tissue and muscle that helps to protect them. Your eyes are also protected by your **eyelids** and **eyelashes**, which block out bright light and help to keep out dust, dirt and other foreign objects.

Tears are produced by the **lacrimal gland**, located in the orbit just above the outside corner of each eye. Tears are swept across the front of your eyes each time you blink, and drain into ducts at the inner corners of the eyes. Tears not only lubricate the eyes, but also work with your eyelids and eyelashes to protect against dirt and infection.

8.3.3.7.4 Abnormalities in the Eye[15]

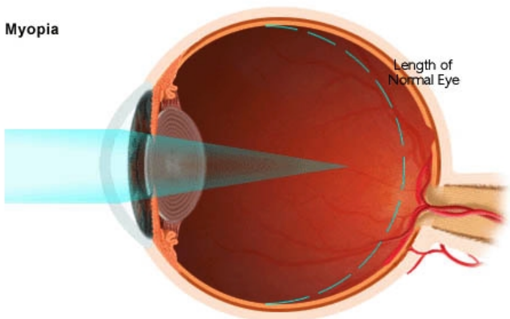
There are three common conditions that can arise as a result of abnormalities in the structure of the eyeball: myopia, hyperopia and astigmatism

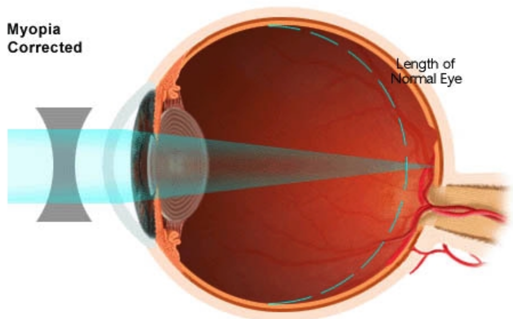
Myopia

Commonly known as nearsightedness, myopia is a condition in which an image of a distant object becomes focused in front of the retina. This can occur either because the eyeball axis is too long or because the refractive power of the cornea or lens is too strong (*i.e.* the cornea or lens is thicker than it needs to be).

Nearsighted people typically see well up close, but have difficulty seeing distant objects clearly. Myopia is the most common refractive error seen in children. It usually becomes progressively worse through adolescence and stabilises in early adulthood. It also appears to be an inherited problem.

Spectacles or contact lenses can help to correct or improve myopia by adjusting the focusing power of the lens as illustrated.





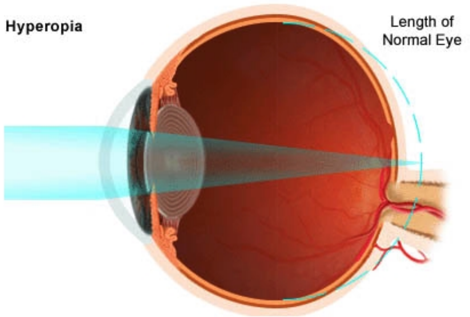
Hyperopia

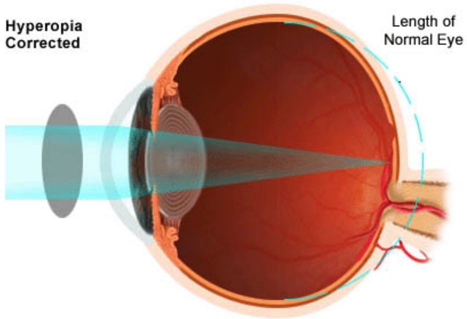
Commonly known as farsightedness, hyperopia is the refractive error in which an image of a distant object becomes focused behind the retina. This can occur either because the eyeball axis is too short, or because the refractive power of the cornea or

[15] http://www.childrenshospital.org/az/Site1517/printerfriendlypageS1517P0.html

lens is too weak (*i.e.* the cornea or lens is thinner than it needs to be). This condition makes close objects appear out of focus.

Spectacles or contact lenses can help to correct or improve hyperopia by adjusting the focusing power of the lens as illustrated.

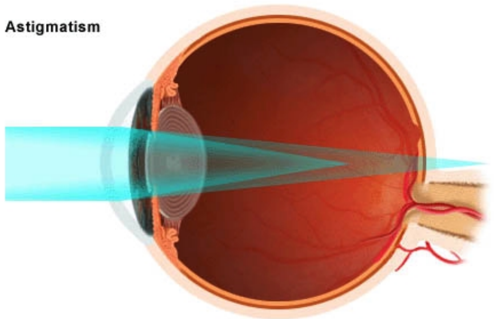




Most people develop hyperopia as the age, usually from around 40 years on, as the eye become less flexible.

Astigmatism

Astigmatism is a condition in which an abnormal curvature of the cornea can cause two focal points to fall in two different locations making objects up close and at a distance appear blurry. Astigmatism often occurs along with nearsightedness or farsightedness, and in some cases can also be corrected with spectacles or contact lenses.



References

The Physics Classroom
(http://www.physicsclassroom.com/Default2.html)

The Wave-Particle Duality of Light
(http://micro.magnet.fsu.edu/primer/lightandcolor/particleorwave.html)

Optics (Light Bench) Experiments
(http://www.physics.isu.edu/~hackmart/opticsintro.pdf
http://www.physics.isu.edu/~hackmart/optics%201.pdf
http://www.physics.isu.edu/~hackmart/optics%202.pdf
http://www.physics.isu.edu/~hackmart/optics%203.pdf)

Interactive Evaluation of Image Formation with Mirrors and Lenses

Image Formation by a Diverging Mirror
(http://www.physics.uoguelph.ca/applets/Intro_physics/kisalev/java/dmirr/index.html)

Image Formation by a Converging Lens
(http://www.physics.uoguelph.ca/applets/Intro_physics/kisalev/java/clens/index.html)

Image Formation by a Diverging Lens
(http://www.physics.uoguelph.ca/applets/Intro_physics/kisalev/java/dlens/index.html)

Imaging with a Lens
(http://www.mtholyoke.edu/~mpeterso/classes/phys301/geomopti/lenses.html)

Lens Combinations: Microscopes
(http://www.mtholyoke.edu/~mpeterso/classes/phys301/geomopti/micro.html))

Lens Combinations: Telescopes
(http://www.mtholyoke.edu/~mpeterso/classes/phys301/geomopti/twolenses.html))