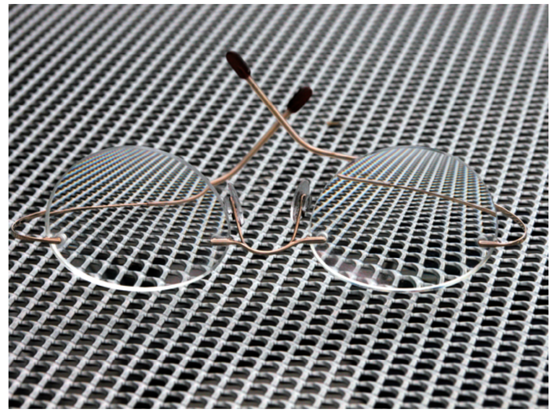
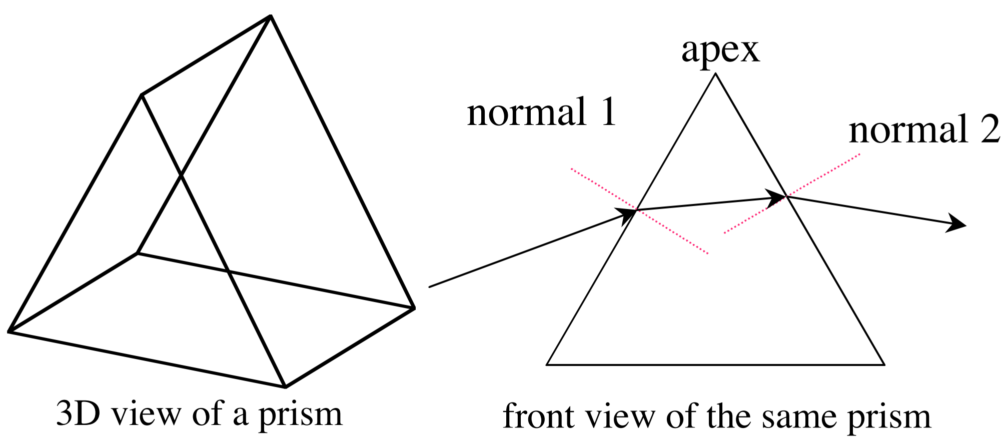
You need to learn the concepts and formulae highlighted in red. The rest of the text is for your intellectual enjoyment, but is not a requirement for homework or exams.

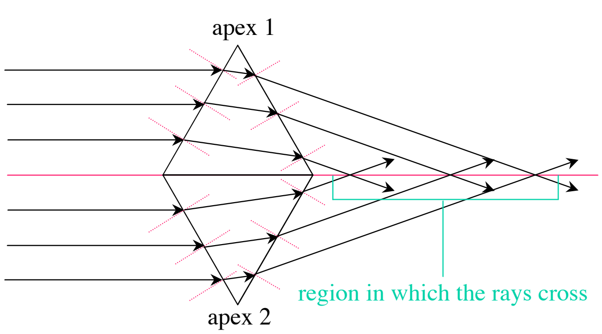
Chapter 2
LENSES



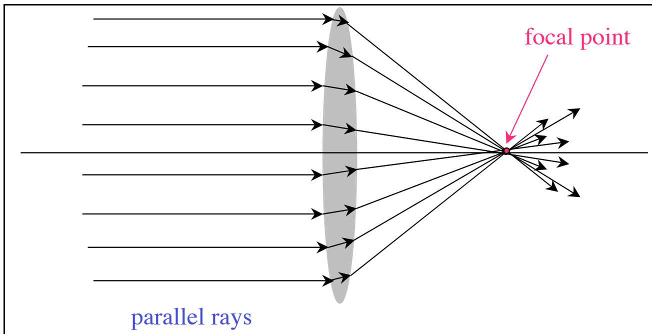
THE PRISM

When the surfaces of a block of glass or plexiglass are not parallel, as in a prism, the incoming and outgoing rays of light are not parallel. In the figure below, the incoming ray is refracted, and kinked towards the normal 1, it is then transmitted through the prism, until it reaches the other surface, and refracted again, away from the normal 2. The end result is ***always*** that the ray is deflected ***away from the apex*** of the prism.





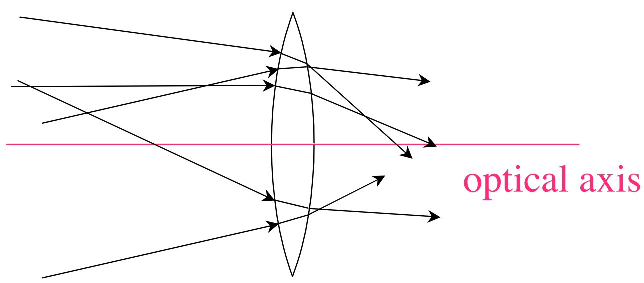
Let us now combine two prisms, and illuminate them with parallel rays of light, as shown on the left, the top prism deflects the rays down, while the bottom prism deflects them up. The rays that travel through equivalent sections of the two prisms cross each other at the center, on the axis. Is there a transparent object that can make ***all*** rays that travel through it cross at the same point?



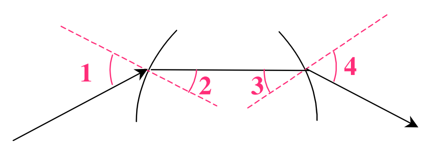
If the two surfaces of the transparent prism are not flat but curved along a spherical surface, the result is a lens. In the diagram here, the lens is convex on both sides, therefore it is a converging lens. All parallel rays coming from the left converge into a single point.

CONVERGING AND DIVERGING LENSES

Converging lenses are ***thicker*** in the middle than near the edges - they ***deflect light rays towards an axis***. You can see this if you apply the law of refraction to the surface where the light beam enters the glass and to the surface where the light beam comes out of the glass. You may think of a little piece of the lens as a prism. The diagram below shows a few possible light rays.



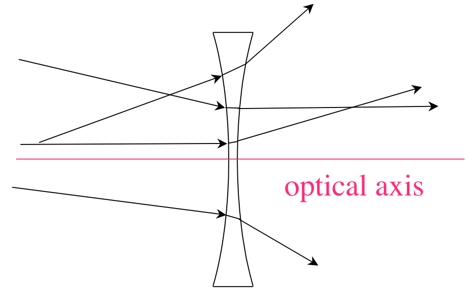
converging lens



enlarged view of part of the lens

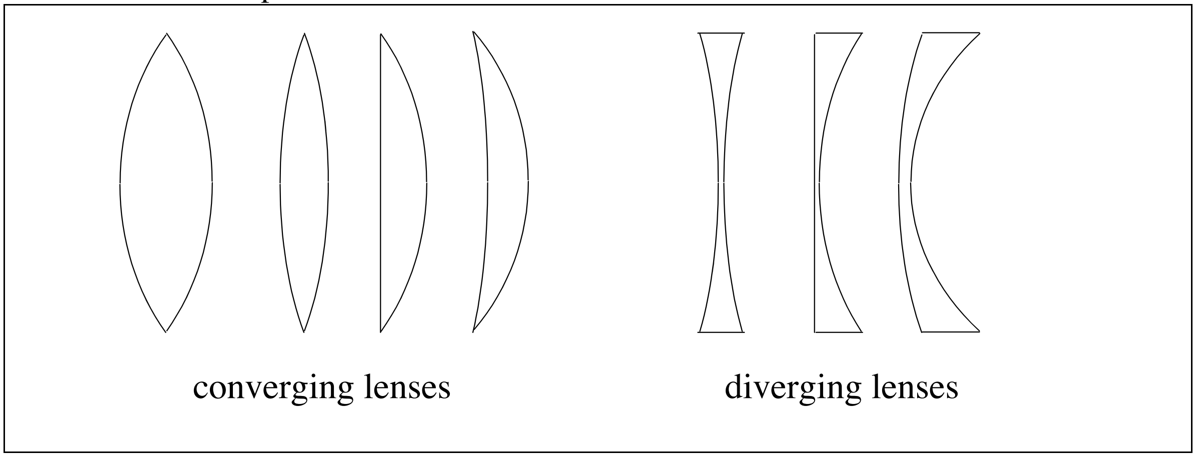
Notice in the enlargement that angle 2 is smaller than angle 1, and angle 4 is larger than angle 3, according to Snell's law.

Diverging lenses are ***thinner in the middle***, and they deflect light rays ***away*** from the axis.



diverging lens

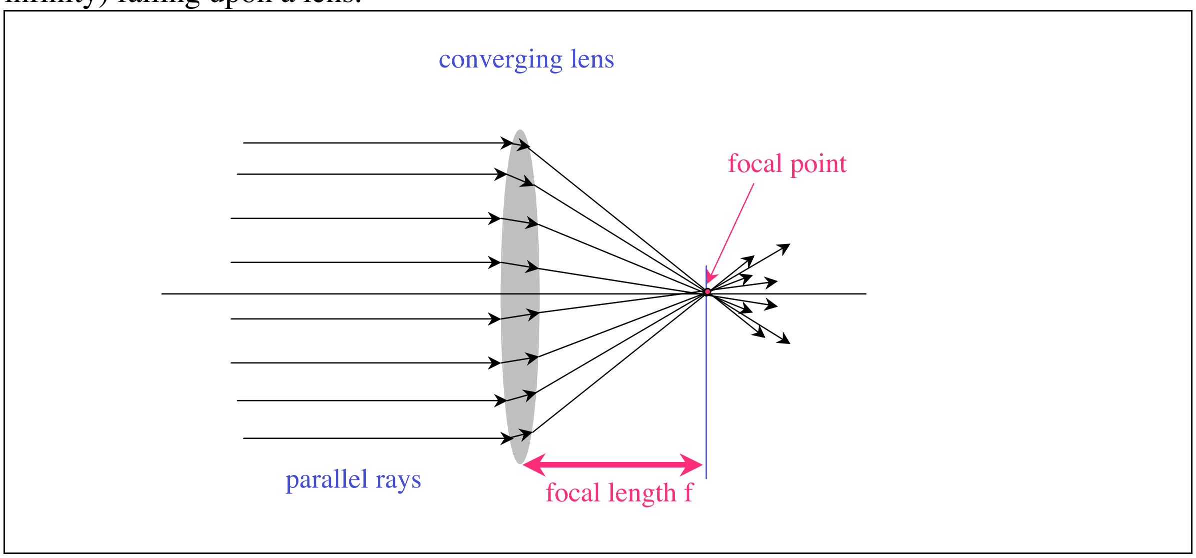
Some other lens shapes:



The more curved the surfaces, the more the rays are deflected, when going through the lens. The higher the refractive index of the lens material, the more the rays are deflected.

FOCAL LENGTH

Consider parallel light rays (for instance light rays from a light source very far away, or at infinity) falling upon a lens.



The rays converge at a point at a certain distance **f** from the lens. This is called ***the focal length*** of the lens. Conversely if you place an object (a source of light) at distance f from the lens, the light rays after passing through the lens will be parallel to each other. A *strong lens* is one that deflects light a lot, meaning it has a *short focal length*. The focal length of camera lenses is always given in millimeters. The most common and inexpensive 35 mm cameras (meaning 35 mm film width, *not* 35 mm focal length!) have f = 50 mm lenses. For eyeglasses, the ***dioptric power*** is usually used instead of the focal length. The relationship between dioptric power (measured in diopters) and f is simple:

$$\text{Dioptric Power (D)} = \frac{1}{f(m)}$$

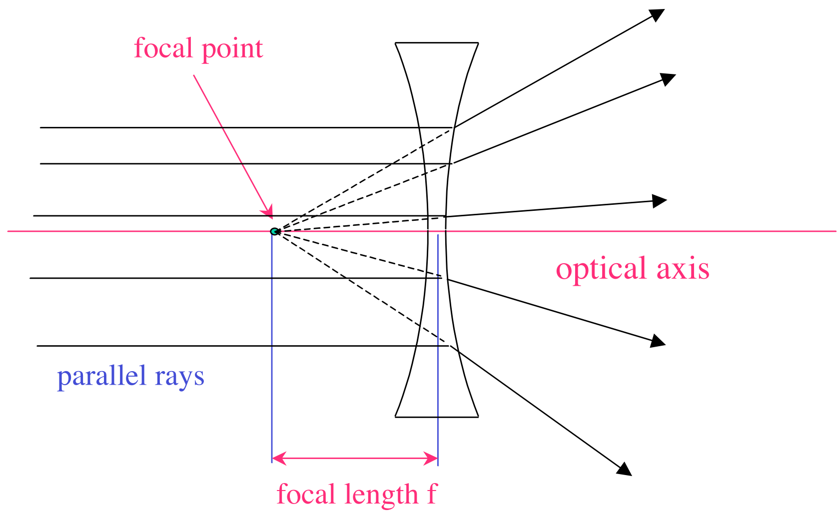
The focal length f in this formula *must* be expressed in *meters*.

Thus a camera lens with a focal length f = 50 mm = 0.05 m has $\frac{1}{0.05}$ = 20 D. The advantage of diopters is that if you combine two lenses one right next to the other, you simply add the diopters. This implies that for two lenses the combined focal length is:

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$$

where f_1 and f_2 are the focal lengths of the two lenses.

Diverging Lens

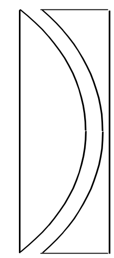


Diverging lens with parallel rays incident. The diverging lens makes the rays diverge, as if they were originating from the focal point. The image of a far away object (at infinity) on the left is formed in the focal point on the left, on the same side of the lens as the object. ***The image is virtual***, that is, there are no actual rays converging in the image point, but only the artificial extension of diverging rays (dashed lines) converge in the focal point.

After the parallel rays go through the diverging lens, they diverge *as if* they came from a point at distance f on the other side of the lens. This distance is also called the ***focal length*** of the lens. To indicate that the (apparent) crossing point of the rays is on the other side of the lens, ***the focal length of a diverging lens is stated as a negative number***. An important advantage of giving the focal length of a diverging lens as a negative number is that it makes the formulas work out automatically. For instance, if you combine a converging lens with 10 cm focal length with a diverging lens of -10 cm focal length the resulting focal length f is calculated from the following equation:

$$\frac{1}{f} = \frac{1}{10} + \left(\frac{1}{-10}\right) = \frac{1}{10} - \frac{1}{10} = 0$$

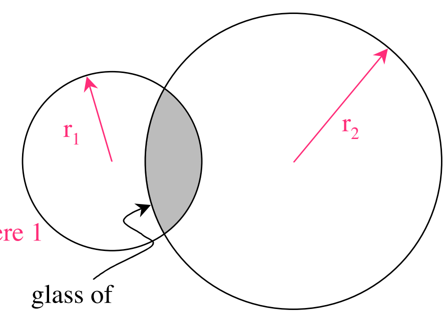
$f = \infty$



In the system of the two lenses, the resulting focal lens is infinite. This means that if the incoming rays are parallel, the outgoing rays are also parallel: the lens combination acts like a flat window pane. Does this make sense? Yes, because when a converging lens and a diverging lens are combined, so that the convex and concave sides are glued together, the resulting object is indeed a parallel slab of glass, like a window pane!

The focal length of lenses depends on the index of refraction of the glass and on the radius of curvature of the two faces of the lens, r_1 and r_2:

$$\frac{1}{f} = (n-1)\left[\frac{1}{r_1} + \frac{1}{r_2}\right]$$



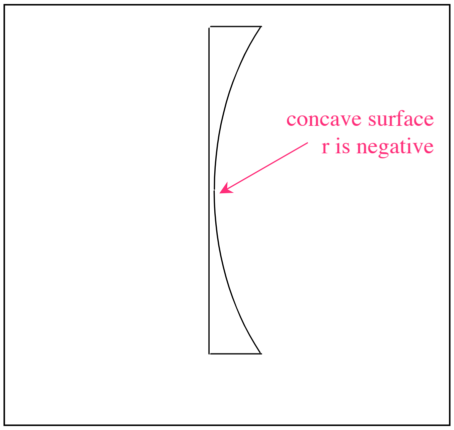
This is called the ***lens-maker's formula***. This formula has important consequences, worth noticing here: there are always two focal points for each lens, and they lie at a distance f on either side of the lens. The lens may have two different radii of curvature, but the focal points are always equidistant. This implies that a lens can be used from one side, or flipped and used from the other side and the image will be identical. If you wear glasses, you can try this right now! Take your glasses off, rotate them 180° around a horizontal axis and look through them (if you rotated them correctly, you should be looking with your right eye through the right lens and with the left eye through the left lens, but inside-out). You can see just as well with the lenses inverted: observing how the lenses are curved, this is really counterintuitive!

If the face of the lens is ***concave***, r is counted ***negative***. The formula makes sense:

if n = 1 (a lens made of air!), $\frac{1}{f}$ is always zero $f \to \infty$

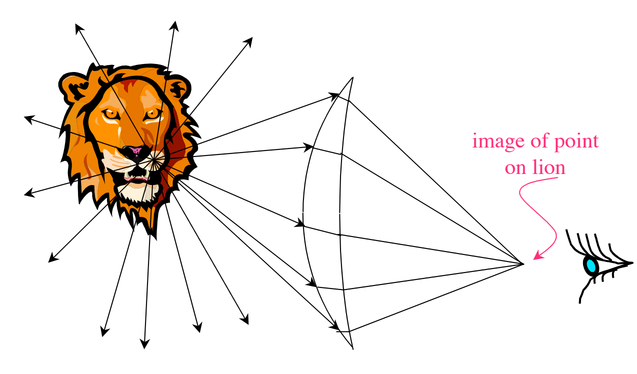
meaning the rays are not deflected. The smaller r_1 and r_2 get, the smaller the focal length. What does the formula tell you if you take a thin sheet of plastic and curve it? Does it still make sense?

Note that, the lens-maker's formula ***works only for thin lenses***, that is, lenses for which the thickness is much less than the focal length.



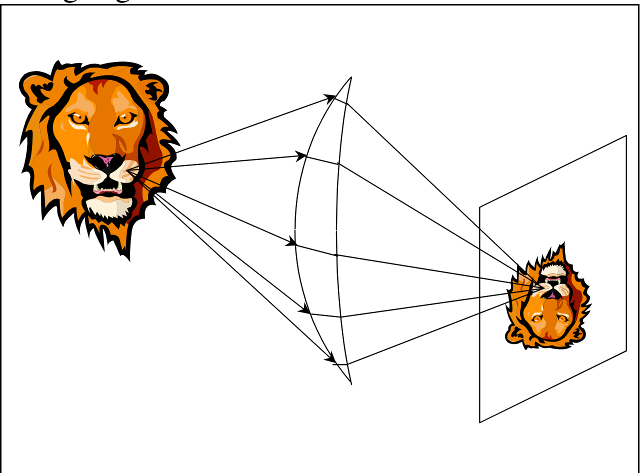
IMAGES - REAL AND VIRTUAL

Sunlight shines on the face of a lion. Think of one tiny point on the lion's cheek. Part of the sunlight hitting that point is diffusely reflected ***in all directions*** - that's why we see the lion with our eyes no matter from what direction we look at it (diffuse reflection of light, as opposed to specular reflection from a mirror which goes in a particular direction only).



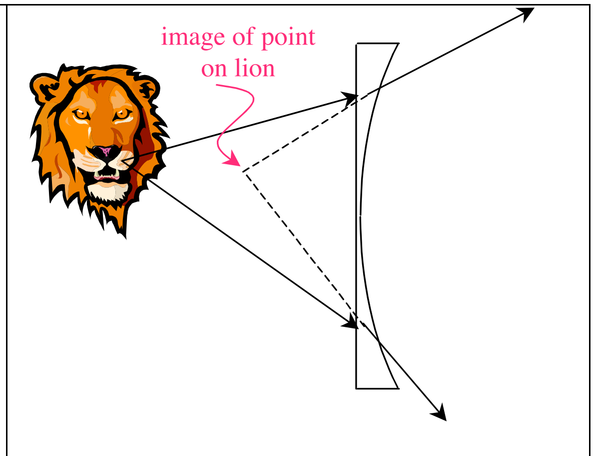
If a lens is placed in such a way that the rays coming from the point of the lion converge to a point on the other side of the lens, you will see light rays, as if you looked at a point of the lion itself, or if you put a screen there, you will see an ***image*** of the point on the lion. This is a real image of the point, because if you put a piece of paper here, the image will form a focused spot on it. In fact you will see an image of the whole lion because point by point the rays going away from the lion converge again on the other side of the lens.

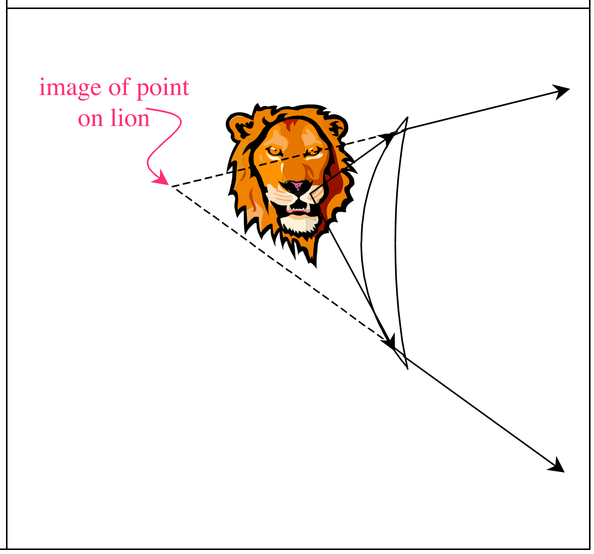
Of course, if there is no screen you will see the image only if you place your eye so that the rays get into your iris. From the side you will not see it unless you place a screen where the image is - the screen will direct some of the light into your eyes, again by diffuse reflection. The image is called ***real*** if (after passing through the lens) ***the rays*** from a given point of the object ***cross each other, or converge, in a point***. A real image can be seen on a screen or on a photographic film.



To form a real image we need a converging lens, but ***converging lenses do not necessarily produce a real image*** as we shall see below. The distance from the object to the lens is called the object distance o. The distance from the lens to the image is called the image distance i. Important note: the distance i is where the rays from the object are focused, but this is *not* the focal length. Focal length refers to parallel incident rays - this requires that the object is very far away (i.e. if o is very large or ∞, the image distance i is equal to the focal length f).

If the object is very near the lens, the lens may not be strong enough to deflect the diverging rays coming from the object sufficiently to make them converge to a point. Or, worse yet, we may have a diverging lens that spreads the rays even more than they were spreading when they got to the lens. In either case, the rays after the lens seem to come from a point at distance i behind the lens (with respect to your eye, see drawing). You can see the image with your eye if you look through the lens, but of course you cannot place a screen or a film where the image is, because no real rays cross there, or anywhere else. This is a ***virtual image***. Your image in a mirror also is ***virtual***: if you place a light bulb near a mirror, there is no way to catch the image on a piece of white paper - you would need an additional lens (as in your eye or in a camera taking a picture of the rays from the mirror) to form a real image. A ***diverging lens***, always produces virtual images. A ***converging lens*** may also form a virtual image, when the object is closer to the lens than the focal point.

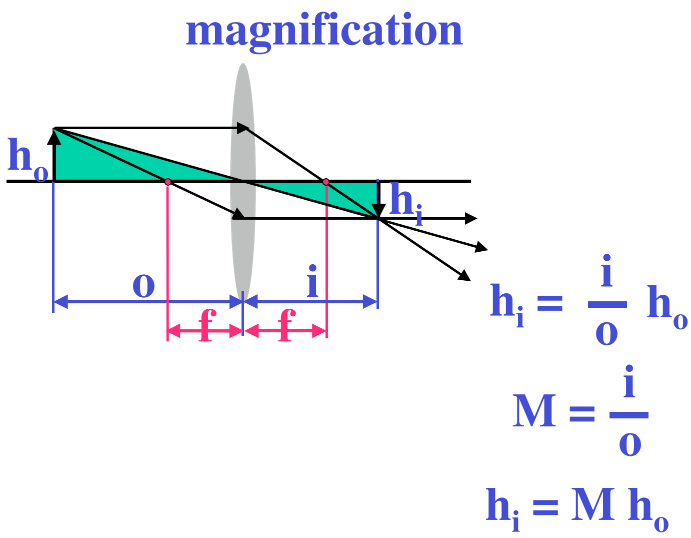




THE LENS FORMULA

The lens formula can be used to answer a variety of questions: where is the image? Is the image real or virtual? How large is the image? The lens formula tells you how ***object distance o*** and ***image distance i*** are related to the focal ***length f:***

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f}$$



The magnification M of a lens can be easily calculated as reported on the left. In the formulas h_o is the height of the object and h_i is the height of the image.

Be careful with the term magnification, since the magnification can be smaller than 1. In this case the lens ***demagnifies***, and the image is smaller than the object. The magnification varies dramatically for converging lenses, which can be magnifying or demagnifying, depending on where the object is positioned with respect to the lens. For diverging lenses the situation is simpler, since they always demagnify.

The magnification is always a positive number: if the image distance i is negative (object and image are on the same side of the lens) take the modulus of i, that is, ***ignore the minus sign***.

Let us now see some examples of how the lens formula and the magnification formula can be used.

1. Take a converging lens of 10 cm focal length and place the object 20 cm from the lens. Where is the image?

$$\frac{1}{i} = \frac{1}{f} - \frac{1}{o} = \frac{1}{10} - \frac{1}{20} = \frac{2}{20} - \frac{1}{20} = \frac{1}{20}$$

i = 20 cm

The image is 20 cm from the lens. How big is the image? The object height is h_o, the image height is h_i.

$$\textit{Magnification} = \frac{h_i}{h_o} = \frac{i}{o}$$

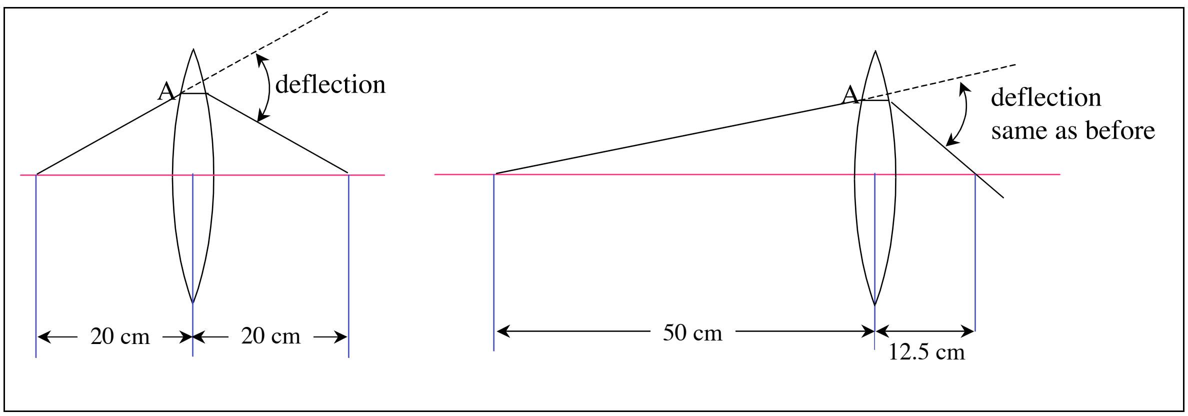
In other words, the image size is to the object size as the image distance is to the object distance. In this example $\frac{h_i}{h_o} = \frac{i}{o} = \frac{20}{20} = 1$ or $h_i = h_o$.
Therefore the image is the same size as the object.

2. Same lens, object at 50 cm:

o = 50 cm ; f = 10 cm

$$\frac{1}{i} = \frac{1}{10} - \frac{1}{50} = \frac{5}{50} - \frac{1}{50} = \frac{4}{50}$$
$$i = \frac{50}{4}\text{cm} = 12.5\text{cm}$$

Image size is now 1/4 [=12.5/50] of object size. Note that as we moved the object from 20 to 50 cm (further away from lens), the image got closer to the lens (from 20 cm to 12.5 cm).



There is a neat and easy way to understand (and thus to remember) this. Look at the diagram above: light hitting point A of the lens is deflected as indicated by the angle labeled “deflection”.

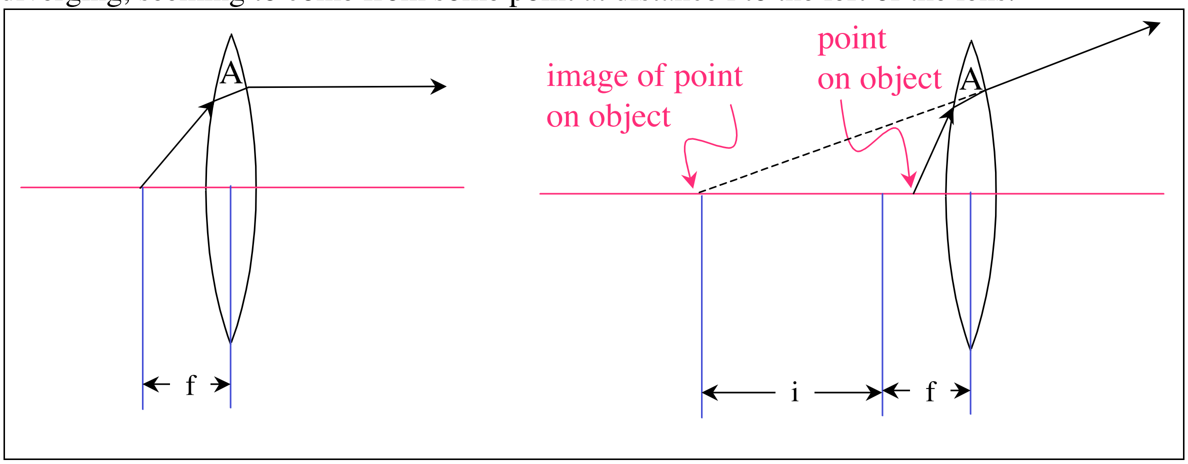
If we move the object further away, the ray coming from an object point on the axis, and hitting point A, will still be deflected by the same angle as before. But the incoming ray hitting point A comes in less steeply when we move the object away, so that the ray after the lens is now steeper and crosses the axis sooner. Eventually, when the object is very far away the ray will cross the axis at a distance f (= 10 cm in our example) from the lens.

3. Same lens, object at 5 cm from lens.

$$\frac{1}{i} = \frac{1}{f} - \frac{1}{o} = \frac{1}{10} - \frac{1}{5} = \frac{1}{10} - \frac{2}{10} = \frac{-1}{10}$$

image distance = -10 cm

The minus sign indicates that the image is on the same side as the object [rather than on the opposite side, as in examples 1 and 2]. This ***negative i*** indicates that the image is ***virtual***. The size of the virtual image is twice the object size [since i/o = 10/5 = 2]. For a converging lens the image is virtual whenever the object distance is less than the focal length of the lens. Again we can use the easy way to understand it, just as we did above: a ray starting on the axis of the lens a distance f from the lens will come out parallel to the axis - that's how focal length was defined. If you move the object closer yet, the ray hitting point A will still be deflected by the same angle as before, but this means it keeps diverging, seeming to come from some point at distance i to the left of the lens.



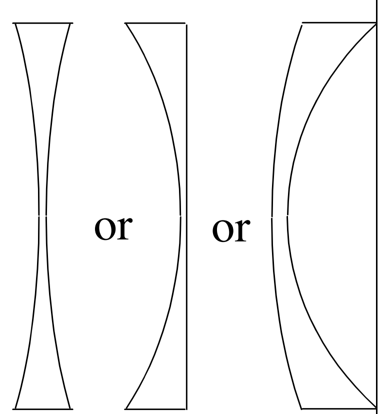
4. ***Diverging lens***, f = -10 cm.
I must use f = -10 cm (negative focal length) in the lens formula.

$$\frac{1}{i} = \frac{1}{-10} - \frac{1}{o} = \frac{1}{-10} - \frac{1}{o}$$

Note that the image distance i will also be negative (virtual image) no matter where I place the object, that is, for any o value. For example if o = 10 cm

$$\frac{1}{i} = \frac{-1}{10} - \frac{1}{10} = \frac{-2}{10} = -\frac{1}{5} \quad i = -5\text{cm}$$

The ***virtual*** image is 5 cm from lens, on the ***same side of the lens*** as the object. The image will be half the size of object, because $M = \frac{i}{o} = \frac{5}{10} = \frac{1}{2}$. Remember to neglect any minus signs when calculating the magnification.



THREE EASY RAYS

Often it is much nicer to draw rays than to use the lens formula to find the image size and position. One can see why the image is where it is. This procedure is called ***ray tracing***, and is performed by very fast computers, to be able to trace very many rays of light, and simulate what happens in nature. Ray tracing is used for scientific applications, for animation movies and for computer graphics. It may seem complicated to do, but in the case of lenses it is very simple. The trick is to use three particularly easy rays.

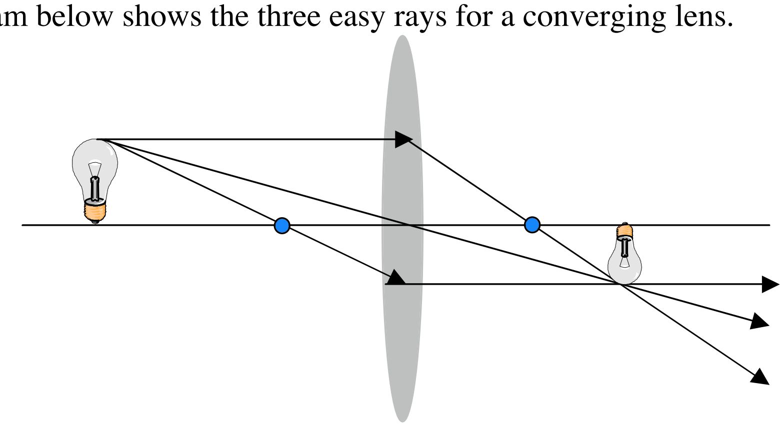
If we select a point on the object, say the top of the light bulb in the diagram below, we do not need to repeat the operation for each point in the light bulb. The ***three easy rays*** to trace ***for a converging lens***, as reported in the diagram below, are:

- A ray from the object ***through the center*** of the thin lens goes ***straight through*** because at the center of the lens the two lens faces are parallel to each other. There is in fact a slight shift, as for a window pane, but it can be neglected for thin lenses.

- A ray from the object to the lens, ***parallel to the optical axis*** goes ***through the focal point on the other side of the lens***. The image point is where these two rays cross. So now we can draw the image as shown in the graph. It is upside down (inverted image).

- A third ray is the ray from the object ***through the focal point*** on the left of the lens. This ray will go out ***parallel to the axis***, and of course will go through the same image point already found.

The diagram below shows the three easy rays for a converging lens.

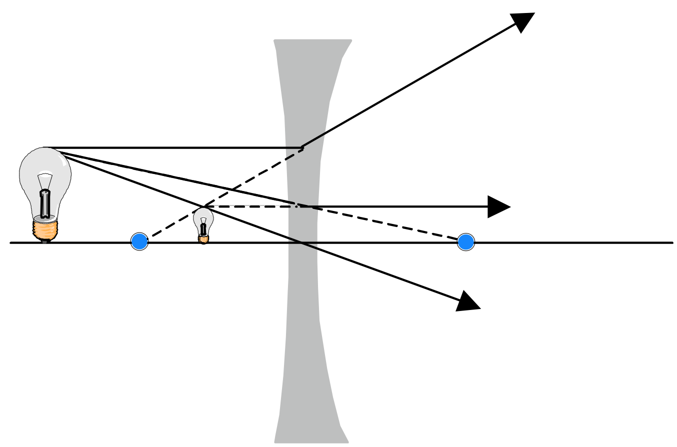


The blue dots mark a distance f from the lens, either way.

Now ***for a diverging lens, the three easy rays***, reported in the diagram below, are:

- As before, the first easy ray is ***through the center of the lens***.
- The second ray is ***parallel to the optical axis***, and is deflected out ***as if it came from the focal point on the left*** of lens. The image point is where the ***extension*** (dashed line) of the first and second rays cross.
- The third ray is the ray from the top of the object ***heading for the focal point on the other side of the lens***. This ray will come out ***parallel to the axis***, seeming to come from the point already identified by the first two rays. This ray may seem more complicated to draw. Remember that all light rays can always be inverted, therefore if this third ray comes from the right, parallel to the optical axis, it will diverge on the left, as if it came from the focal point on the right.

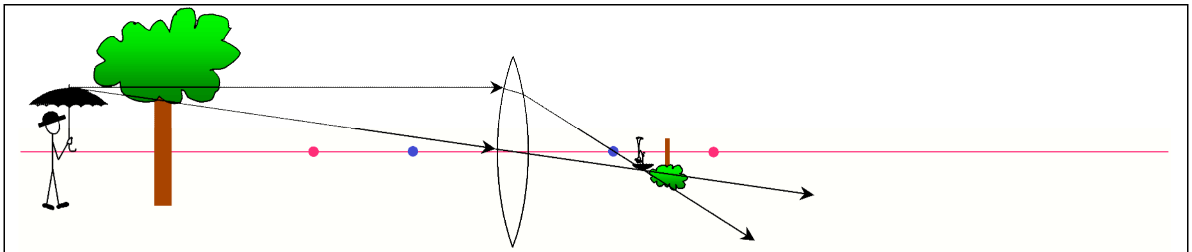
We see, from these three easy rays that the image is virtual, upright and smaller than the object, as always for diverging lenses.



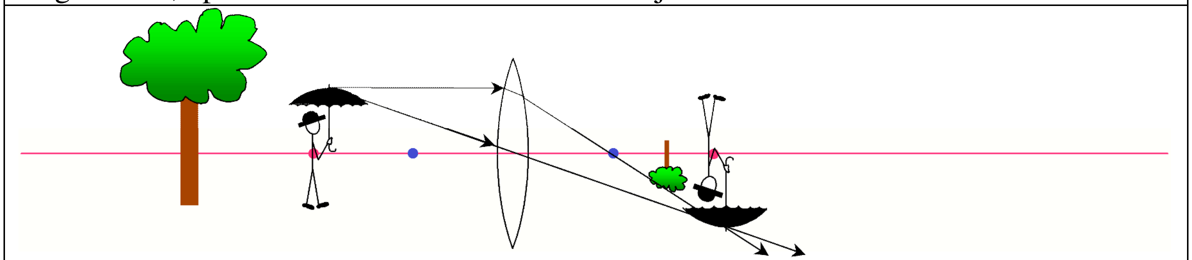
The blue dots mark a distance f from the lens, either way

Note on Magnification

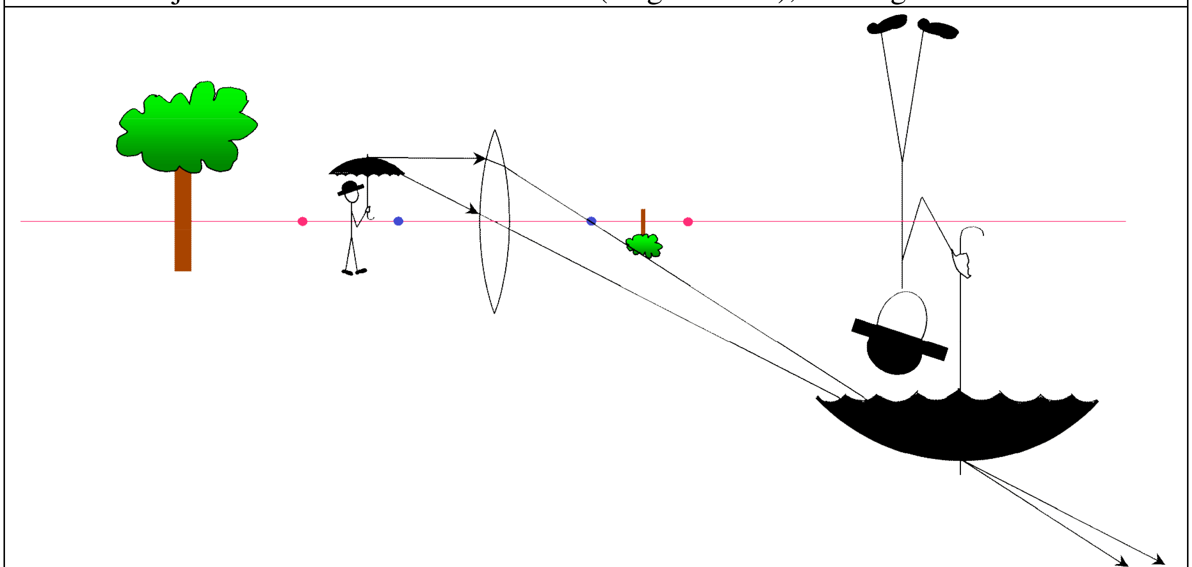
As mentioned, the magnification is always < 1 for diverging lenses, that is, these are always demagnifying lenses. On the other hand, for converging lenses the magnification varies dramatically. Converging lenses can be magnifying or demagnifying, depending on where the object is positioned with respect to the lens. See the figure in the next page for various object positions. Notice that one object (the man with the umbrella) moves, while another object (the tree) is stationary.



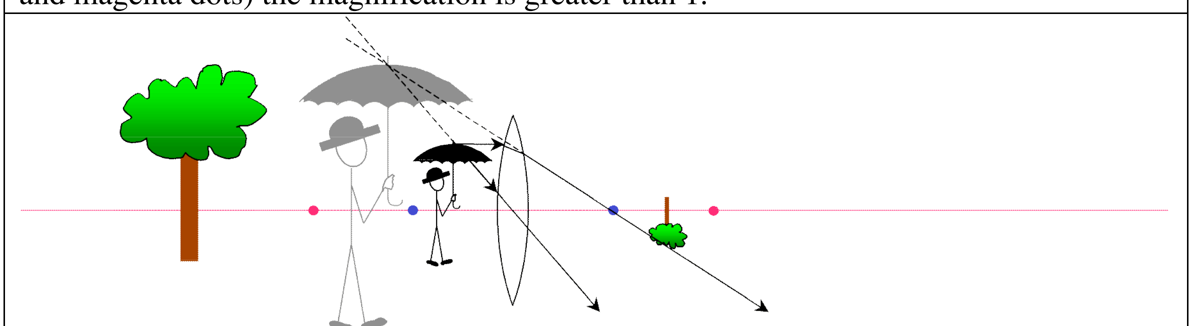
When the object (man with umbrella) is further than the focal distance (blue dots), the image is real, upside down and smaller than the object.



When the object is at twice the focal distance (magenta dots), the magnification is 1.

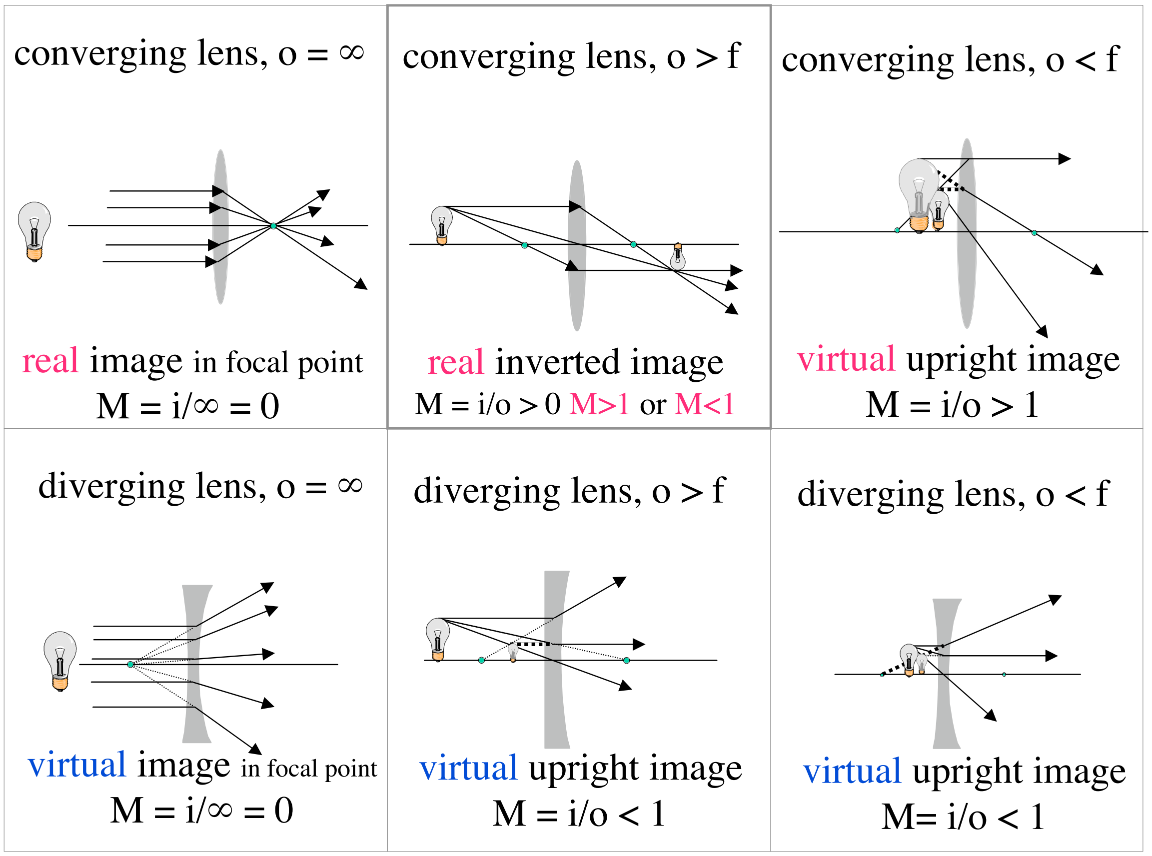


When the object is between the focal distance and twice the focal distance (between blue and magenta dots) the magnification is greater than 1.



When the object is closer than the focal distance, the image is virtual (gray man), upright, and the magnification is greater than 1. (Adapted from Eugene Hecht, *Optics*, Addison-Wesley Publishing Company, Reading MA).

The figure below shows a summary of resulting images and magnifications in all possible cases for converging and diverging lenses: object at infinity, object farther than the focal distance and object closer than the focal distance.

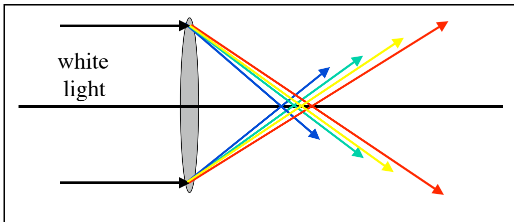


LENS ABERRATIONS

All the formulas and ray tracing we discussed so far work satisfactorily as long as the lenses are thin, and most importantly as long as we are in the ***paraxial approximation***, that is, we only consider ***rays of light close to the optical axis***. Selecting only paraxial rays can be done either using lenses with a very small diameter, or larger lenses with an iris aperture that limits the used diameter of the lens. Decreasing the effective diameter, that is, closing the aperture, limits the ***luminosity***. Consider an object and a lens. Rays from the object to the lens hit all points on the surface of the lens, not only the easy rays we showed before. Each portion of the lens is producing the image. You may therefore cover a portion of the lens, and still obtain the same image. If for example you cover half of the lens, the other half (which now has a D shape), you will obtain the exact same image of the object, with same magnification and position, but this image will have half the intensity. The reason for using larger lenses is to increase the lens luminosity. That is why the best telescopes, which can see the faintest stars, have very large diameters. The same applies to camera lenses: the best cameras have the largest diameter lenses, so they can take photographs in dim conditions of light, with the aperture wide open, that is, using the entire diameter of the lens.

Departing from the paraxial condition, e.g. taking a photograph with the aperture wide open, introduces problems with the image clarity and shape. These problems are called ***lens aberrations***. There are ***chromatic aberrations***, which arise from the fact that the index of refraction n varies for the different wavelengths of light (colors), as well as ***monochromatic aberrations***, which take place even when light of a single wavelength is used. The monochromatic ones are: ***spherical aberration***, coma, astigmatism, field curvature and ***distortion***. We will only discuss the aberrations that mostly affect photography.

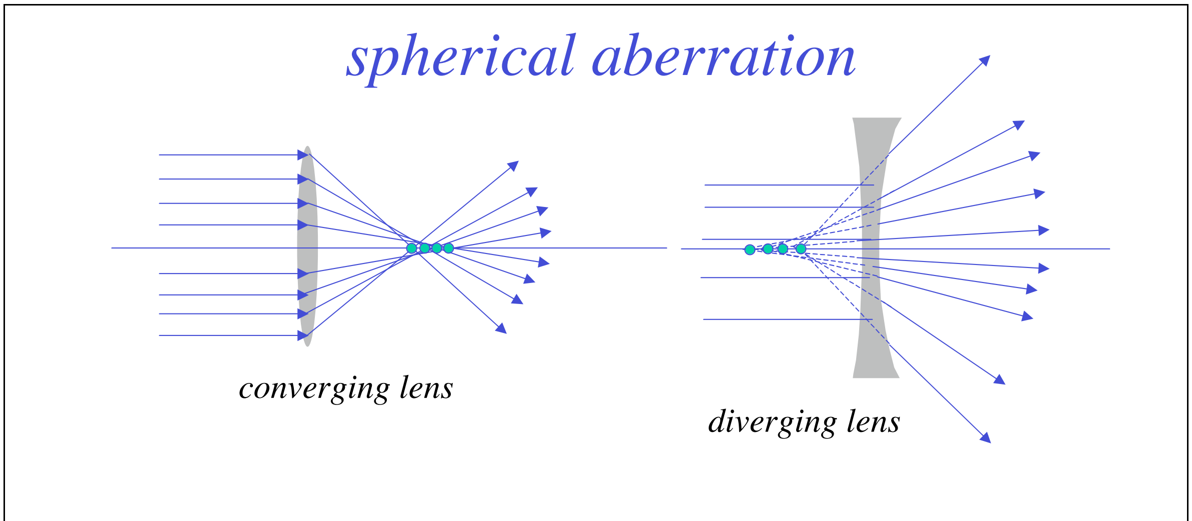
Chromatic aberrations: the index of refraction of glass is a little different for different wavelengths of light (***dispersion***). Rays of different colors are separated by the lens, and blue light rays are deflected more than red light rays. Consequently the focal length of a simple lens is different for blue and for red light. The difference is not very large, but enough to cause trouble: red and blue polka dots on a dress can not be captured in good sharp focus at the same time with a camera using a simple lens. If you focus on the red dots, the blue dots will be a little fuzzy and vice versa. You can see this imaginingto be perfectly in focus for the red dots. The blue dots will then be focused maybe 0.2 mm in front of the film and by the time the rays get to the film they will have already re-diverged a little, and produce a blurry blue dot image. The blur is worse if the lens aperture is large, because the rays diverge more over the 0.2 mm distance. This is because rays of light farther from the axis of the lens must be deflected more to be brought into focus. The more rays of light are deflected by the lens, the greater the effect of chromatic aberration. If you don't need a large lens aperture (i.e. you want to take pictures only outside on sunny days) the solution is simple: make the lens aperture small.



Blue light is deflected by the lens more than red light, the focal point for blue is closer to the lens than the focal point for red. Blue and red objects cannot be in focus at the same time. Closing an iris aperture, i.e. selecting the central rays only, or ***paraxial rays***, reduces chromatic aberration.

Better, more expensive cameras that have ***achromatic lenses*** produce sharper pictures and can be used in dimmer light. Achromatic lenses are built by combining a strongly converging lens, made of low-dispersion glass, with a weakly diverging lens of high-dispersion glass. In this way the dispersion cancels but net focusing remains. You might wonder: why not use a kind of glass with no dispersion? There is no such glass or plastic! Or why not cancel the dispersion by making half of the lens with glass of the opposite characteristic? Again it can't be done, since all transparent materials deflect blue more than red.

Spherical aberration: even monochromatic rays of light, which have a single color, suffer from spherical aberration. This aberration arises from the fact that rays hitting the ***outer diameters*** of a lens are ***deflected more*** than paraxial rays, the rays hitting the lens inner diameters. This is true for both converging and diverging lenses, although the net results cancel, if you use a combination of diverging and converging lenses as shown in the figure below.

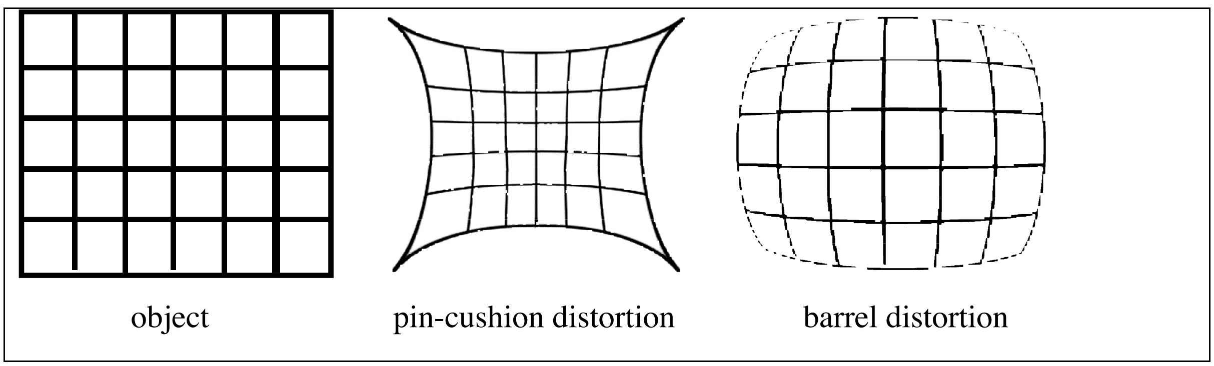


The converging lens moves the focus to the left when using outer diameters of the lens. The diverging lens moves the focus to the right, when going towards outer diameters. Since these two lenses compensate each other, for camera lenses, a combination of diverging and converging lenses is used to minimize spherical aberration.

Another solution to eliminate spherical aberration is to use a small aperture, but again this can only be done in bright light.

Chromatic and spherical aberration, as well as ***coma*** and ***astigmatism***, deteriorate the image clarity. ***Field curvature*** and ***distortion*** deteriorate the image shape, that is, they deform the image.

Distortion: It is not enough that the picture is in sharp focus for all colors and for rays passing through all parts of the lens. Suppose you take a picture of a rectangular building with a rectangular grid of windows - these should be imaged on the film as rectangles rather than as slightly curved lines. Common distortions are the pin-cushion distortion or the barrel distortion:



These distortions arise from the fact that the image ***magnification varies with the image off-axis radius***. In the pin-cushion distortion the center of the image has lower magnification than the edges, while in the barrel distortion the center has higher magnification than the edges.

The design of a good quality camera lens is complicated and expensive, because it has to compensate for all these aberrations for all possible object distances, including objects near the center of the field of view and near the edges. The solution to the problem lies in making ***compound lenses*** by stacking different lenses, each lens curved and spaced from the next lens in the proper way. In practice, computer ray tracing programs do trial and error in simulations of compound lenses, until the best compromise is reached, and only then the lenses are physically mounted in stacks. For most purposes it is not worthwhile to buy the absolutely best lens available since even for an average quality lens the imperfections in the picture are very slight. In particular even cheap lenses can take quite good quality pictures if the lens aperture is kept small enough. However, if your interest is to take photos when you really can't use a tripod or a flash, e.g. for early morning bird shots in the wilderness, you must have a ***fast camera lens***, that is, a lens with a very large diameter (e.g. 500 mm) to increase luminosity and allow a fast exposure with no tripod or flash. Such lenses are always compound, and you may spend more than $10,000 on a single lens. Then you have to carry it around on long hikes, and it may weigh 20 pounds!

Coated lenses: most good-quality camera lenses have coating to reduce the loss of light intensity caused by reflection from the lens surface. This is called ***non-reflective coating***. Some lens coatings may give a noticeable magenta tinge, because a small amount of blue and red lights are reflected, and their sum makes the coating appear magenta. The best quality coatings are perfectly colorless, and they do not let any light be reflected back. Let us see in detail why non-reflective coating is necessary.

Not all the light incident on the glass surface of the lens penetrates the lens and gets refracted. Part of the light is reflected back. The loss in light intensity on the camera film from this reflection is quite serious, because if many lenses are used to avoid aberrations the light goes through many surfaces.

Example:

If the compound lens has 6 lenses, there are 12 surfaces.
If ***only*** ***8%*** of the light intensity is lost to reflection on each surface,

$$0.92^{12} = 0.367$$

only ***36.7%*** of the original light gets to the film!

You would really be better off using a simple lens and avoiding aberrations by always using a small aperture!

This is why, for compound lenses, non-reflective coating should always be used. The coating consists of a thin film of the right index of refraction and just the right thickness, so that reflected light from front and back surface of the coating cancel by destructive interference. So the light cannot be reflected and has to go through the lens.