Imaging

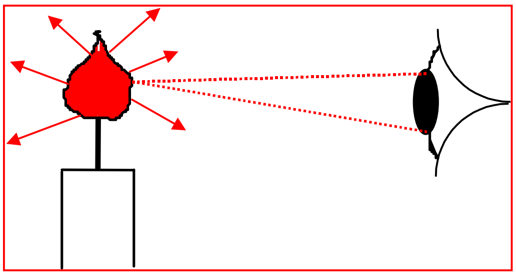
More optical devices are based at least in part on an understanding of imaging than on all other optical phenomena put together. Imaging devices have driven the development of optics for some 350 years. People have been trying to understand imaging for very much longer. Unfortunately, early enquirers of more than 2000 years ago all tackled the problem of trying to understand sight. In one respect this was certainly an unfortunate choice of subject because sight is a complicated phenomenon. Even when the image formation of the eyeball was finally appreciated, which it was in outline by the end of the 16th century, no-one was much nearer to explaining the complexities of the visual experience we all have. An upside-down image of the world in front of us appears at the back of our eyeballs. So what? How does this explain vision?

One fact that reveals a lot about the pace of development of knowledge in general, and science in particular, is how long it has taken to develop the optical instruments we now take for granted. It took some 250 years from the early days of microscopes and telescopes before a really good understanding of the optics involved was obtained. It's not as if there was hardly any interest in the subject, either. When optical instruments improved, so did science in many areas. Optical instruments formed a significant manufacturing and commercial trade. There was implicit pressure from science, from manufacturing and from commerce to improve performance. Yet still it took 250 years, and the best part of the 20th century in addition before manufacturing industry could turn out the high quality of equipment we now take for granted.

My first point is that imaging as it's now presented is a much more modern subject than you might have guessed from the comparative antiquity of optical instruments. My second point is that when introductory textbooks present imaging as a straightforward application of the basic principles of straight-line propagation, refraction and reflection, they may mislead you into believing that the subject is pretty simple and, by implication, without too much of interest to detain you. You may well finish this section with the same impression. To do justice to the depth of modern optical instrument design needs a course by itself, and quite a long course at that. These pages are just the opening chapter in a long story. The pursuit of that story still earns a good living for a good many people around the world.

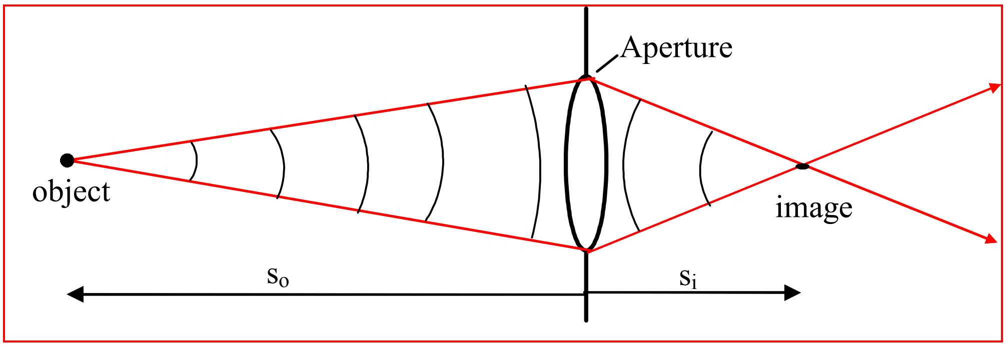
Forming an image

Each point on an object sends out light in many directions. To see an object point we intercept a (narrow) pencil of light coming from the point. The object appears at the apex of the observed cone of light. This is the key to understanding image formation.



Vergence

The concept of convergence or divergence of a pencil of rays is included in the general idea of *vergence*. The following sketch shows a spherical wavefront from an object point converging to a new centre that is the image at a distance s_i from an aperture. The **vergence** of the wavefront in the medium n_2 is defined as:



$vergence = \frac{n_2}{s_i}$, +ve for a converging wave; -ve for a diverging wave. The units are **dioptres**, equivalent to m^{-1}. The incident wave has: $vergence = -\frac{n_1}{s_o}$.

The effect of a lens

Look at the sketch above again. A refracting surface at the aperture **alters the vergence** of an incident wave by the **power** D of the surface. i.e.

$$\frac{n_2}{s_i} = -\frac{n_1}{s_o} + D$$

or

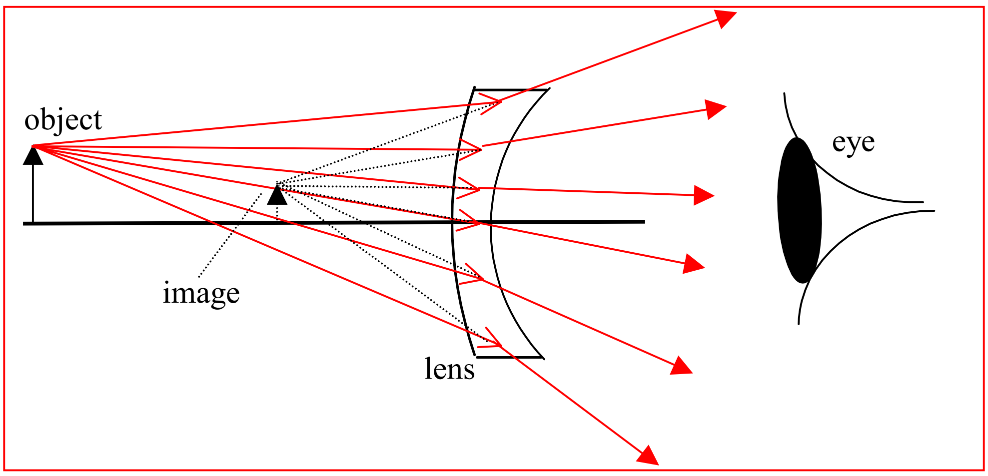
$$\frac{n_2}{s_i} + \frac{n_1}{s_o} = D \quad ,$$

which is the fundamental imaging equation of a refracting or reflecting surface.

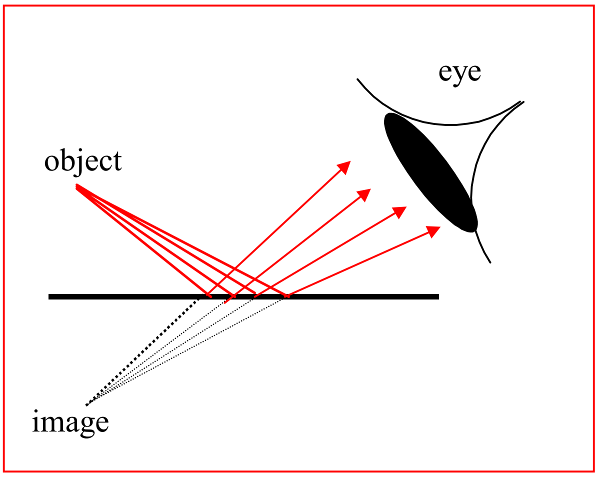
The point of the concept of *vergence* is that it is a useful idea to accompany the wave picture of light and it gives a physical interpretation of the imaging equation. This equation links the distances from the surface that you'll find object and image.

Graphic examples of imaging

- Image formation requires the divergence or convergence of a bundle of neighbouring rays from each object point.



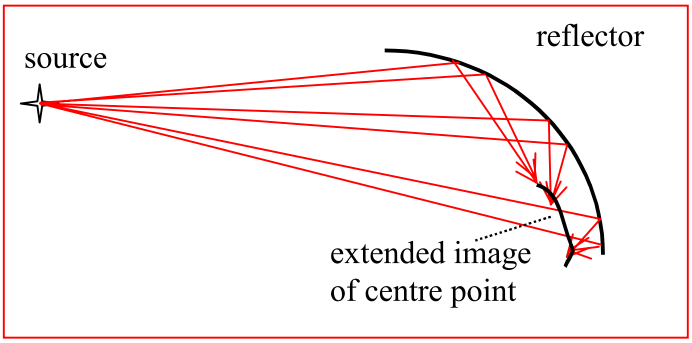
- The apex of the pencil is the position of the image.
- The image of an extended object is the composite image of all its points.
- Simple image forming surfaces alter pencils of light in 2 ways:
  - they alter the vergence of the pencil
  - they may bend the axis of the pencil



Imperfect and perfect imaging

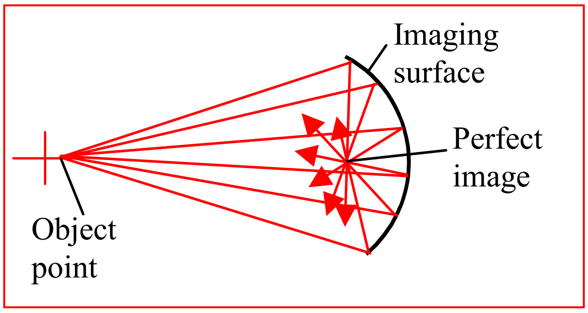
- Images are formed independently by each separate part of an image-forming surface. [Contrast the earlier viewpoint based on the concept of eidola (ειδολα)].

- For a single object point, in general there is no necessity for all the image points created by different parts of a surface to coincide. The most usual situation is for the image to be blurred in some way.



- As a smaller and smaller part of an image-forming surface is used, any blur will become less. Finer detail of a composite image will be visible. For example, the basic reason for poor eyesight is that different parts of the eyelens are creating pencils of light from each object point that don't coincide on the retina. By looking through a very small hole, without glasses, detail can be seen that the whole eye cannot resolve. *Demo*

- The **perfect image-forming surface** is one where images from all parts of the surface exactly coincide. Hence, there is one image point for each object point. All possible object points are said to make up **object space**; all possible image points make up **image space**. In mathematical terms, a perfect image-forming surface performs a **one-one mapping** from object space to image space. Each pair of points is said to be **conjugate**.



- A second feature of perfect imaging (not emphasised in *Hecht*) is that **perfect image mapping must be linear**. What this means is illustrated by the following idea. Consider a ruler or scale as an object standing upright parallel to the image forming surface. Equal object distances in the object plane, like the distances between successive markings on the ruler, must become equal image distances. The following table tries to show what is happening:

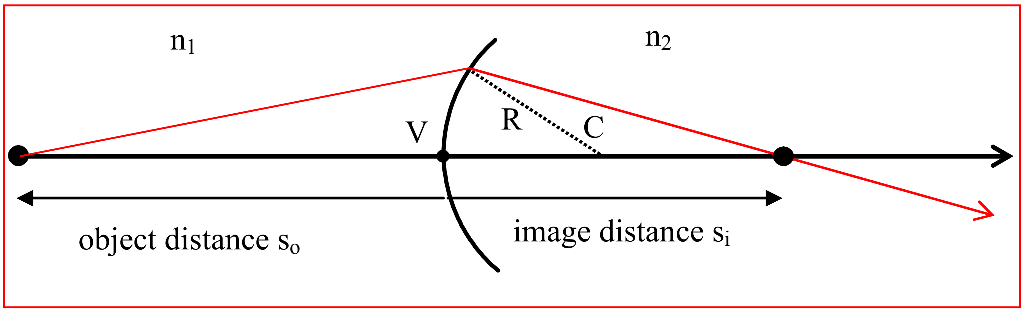
e.g.

object scale:	1	2	3	4	5	6	7	8	9
image scale:	3	6	9	12	15	18	21	24	27
non-linear map:	1	4	9	16	25	36	49	64	81

The perfect image-forming surface cannot be realised in practice, except for a plane mirror. Perfect imaging is an ideal that can be approached from several directions and the success of modern optics shows it can be quite closely approached for many practical purposes. *Hecht* describes imaging by elliptical mirrors, parabolic mirrors, hyperbolic refracting surfaces and other *aspheric surfaces* in special circumstances.

Spherical surfaces and the paraxial approximations

Spherical surfaces form perfect images in the paraxial approximations. This is a remarkable result with great practical implications, because spherical surfaces are by far the easiest curved surfaces to make.



Perfect images are the limit of *paraxial* optics and spherical image-forming surfaces. Paraxial optics means:

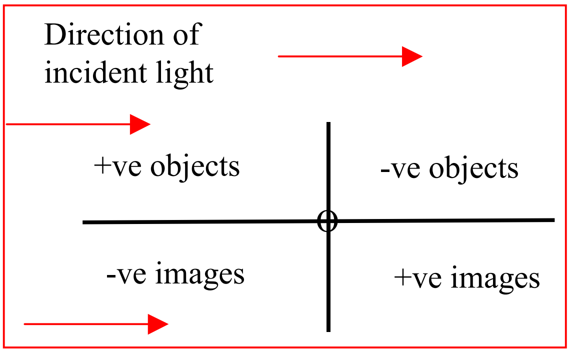
(a) Angles of incidence and refraction are small. The implication of this is that if θ is such an angle, then the 'small angle approximations' of mathematics can be used, namely $\sin\theta \approx \tan\theta \approx \theta$ (in radians). For example, Snell's law you would expect as $n_1\sin\theta_i = n_2\sin\theta_t$ but in the paraxial approximations it simplifies to $n_1\times\theta_i = n_2\times\theta_t$.

(b) Off-axis distances are small compared with the curvatures of surfaces and with object and image distances. The implication of this is that the distances rays of light travel between object, image and surface don't differ much between from the ray coming from an axial object point going along the axis to the vertex. Object and image distances are taken as measured along the principal axis from a fixed origin.

Measurements are made from the **vertex** of the surface, namely the point where the surface intersects the axis of the optical system (marked as V in the diagram above). The result of following through the geometry of the figure above using the paraxial approximations is the comparatively simple relationship between **object distance** (s_o, measured from the vertex) and the **image distance** (s_i, measured from V) given by *Hecht* (fig 5.8)

$$\frac{n_2}{s_i} + \frac{n_1}{s_o} = \frac{n_2 - n_1}{R} = \mathrm{D} \text{ , the power of the surface (in dioptres).}$$

Sign convention



For an algebraic equation to give the right answer when images and objects can be on either side of a lens you need a sign convention to determine which quantities are positive and which are labelled as negative numbers. The sign convention used by *Hecht* is one of two different ones you will find in textbooks. It is known as the **real is positive sign convention**.

First, you have to choose the points from where measurements are made. This depends on what kind of optical element you are considering.

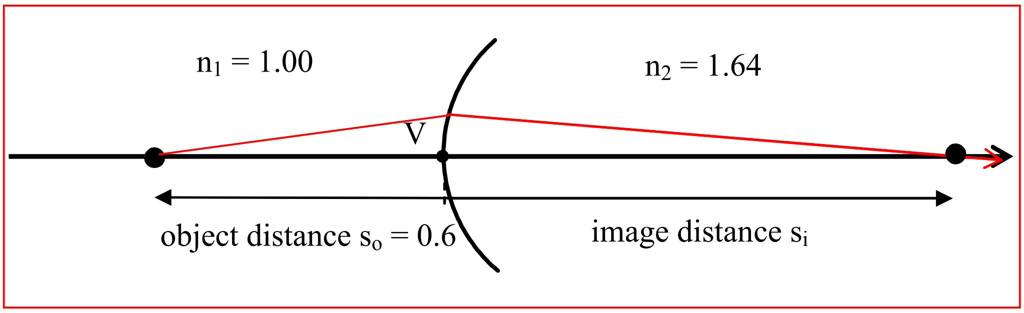
An **element** can be either
a) a single surface, the most basic element. The origin for measurements is the vertex, V
b) A thin lens, the usual element in introductory texts. The origin is the centre of the lens.
c) A compound system. Compound systems won't be dealt with in this course but are introduced in PX2505. There is a choice of origin. One possible selection is to use the focal points, namely F_o for object points and F_i for image points.

Seondly, you must decide which direction is positive. Suppose that light is incident from left to right. For lenses, objects to the left of the origin have +ve s_o (i.e. real objects); images to the right of the origin have +ve s_i (i.e. real images).

Worked example

Some optics courses spend quite an effort in practising various calculations of image location. We will only look at a few examples using the algebraic imaging equations above.

An object is placed on the axis of a refracting surface of radius of curvature 200 mm; the object is 600 mm from the surface in air (refractive index 1.00) and the surface has refractive index 1.64. *Where is the image?*



- Given: $n_1 = 1.00$; $n_2 = 1.64$; R = 200 mm; s_o = 600 mm ≡ 0.6 m . Hence the surface power D = $(n_2 - n_1)/R = (1.64 - 1.00)/0.2 = 3.2$ dioptres.

  Find s_i from $\frac{n_2}{s_i} + \frac{n_1}{s_o} = D$.

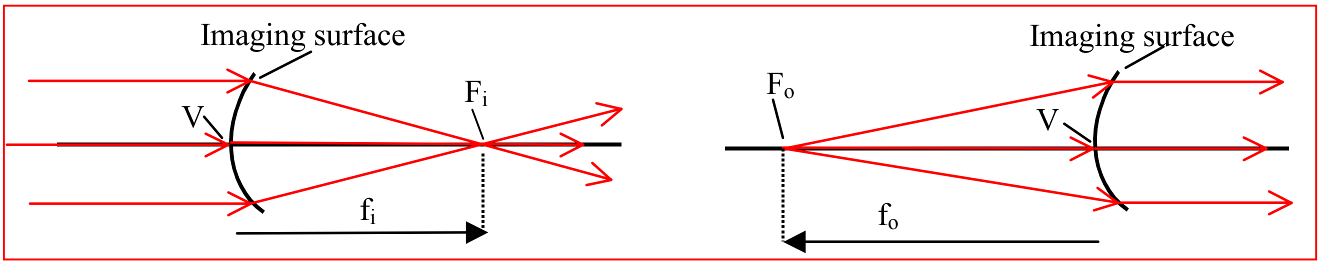
Therefore $\dfrac{1.64}{s_i} + \dfrac{1.00}{0.6} = 3.2$ and hence $s_i = 1.07$ m .

Focal points and focal lengths

Every camera lens has its focal length stamped on its mount; so do enlarger lenses, microscope objectives, telescope eyepieces and virtually every separate lens you can buy. Look up a lens catalogue and the focal length of each component is specified. What is this focal length? It is a distance from a measurement origin to a special point called the focal point. Actually, each image-forming surface has two focal points and it's time to introduce them. The **focal points** of a surface (or lens or complete optical system) are two points used in conjunction with locating the image. F_i is in image space, F_o is in object space.

- The **image (or second) focal point F_i** is the image point for an axial object at ∞.
- The **object (or first) focal point F_o** is the axial object point whose image is at ∞.
- The **image focal length** is the distance $f_i = VF_i$.
- The **object focal length** is the distance $f_o = VF_o$.

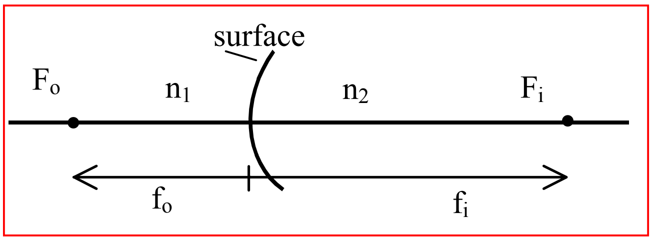
It helps to draw a picture corresponding to the imaging of these descriptions.



Using the imaging equation on page 4, you will easily see that the power of a single imaging surface is related to its focal lengths:

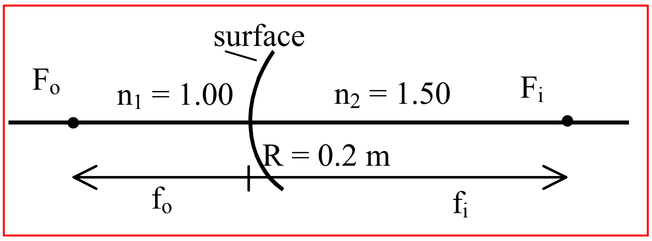
$$\frac{n_2}{f_i} = D = \frac{n_1}{f_o} \quad .$$

This tells us where the focal points are, namely on either side of the vertex a distance away that depends on the media refractive indices and the surface power.



A worked example should make this clear.

An imaging surface between media of refractive indices 1.00 and 1.50 has radius of curvature 200 mm. *What is the power of the surface? What are its focal lengths?*



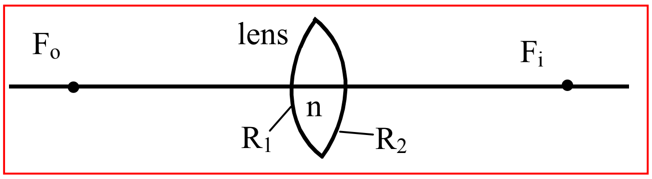
➢ $D = (1.50 - 1.00)/0.2 = 2.5$ dioptres; $f_i = 1.50/D = 0.6$ m; $f_o = 1.00/D = 0.4$ m .

One final point worth remembering here is that the second focal point is NOT THE SAME as the image point in most circumstances. If you look back at the definition of this focal point you'll see that they are only the same for an axial object at infinity.

The usefulness of focal points

- If you know the focal lengths for an image forming surface, and one other fact about it, e.g. its radius of curvature, or the refractive index of the medium in which the object is, you can calculate all its imaging.
- Focal points are two out of six **cardinal points** that characterise an imaging component.
- A lens is simply two imaging surfaces, one after the other. A lens has 2 focal points:

  - Focal points are usually equi-spaced from the centre.
  - A lens is specified by its rear focal length, i.e. its focal length in image space.



Imaging equation for a thin lens

With two surfaces, one after the other but with no significant distance between them, we have the well-known **lensmaker's formula**

$$\frac{1}{s_o}+\frac{1}{s_i}=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right)=\frac{1}{f}=\mathrm{D} \quad .$$

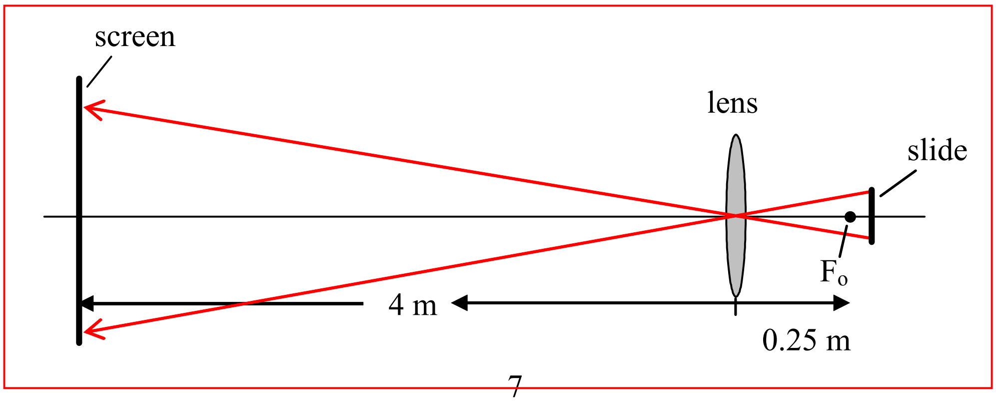
This is simply obtained by using the image produced by the first surface as the object for the second surface. You should be able to show it yourself from the basic single surface imaging equation on page 4.

Lens imaging example

An image of a slide is to be formed 4 m in front of a lens of focal length 250 mm. *How far from the lens must the object be placed?*

$$\frac{1}{s_o}+\frac{1}{s_i}=\frac{1}{f}\ , \ \textit{therefore}\ \ \frac{1}{s_o}+\frac{1}{4}=\frac{1}{0.25}\ , \ \textit{giving}\ \ s_o=\frac{4}{15}\ m=0.267\ m\ .$$

Imaging by a photographic enlarger lens



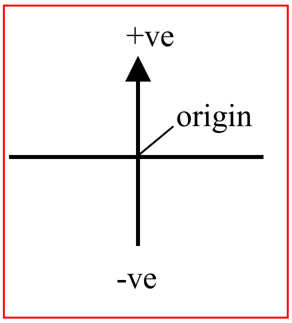
An enlarger lens has a focal length of 50 mm and is positioned 60 mm beneath the negative being enlarged. *How far below the lens must the printing paper be placed so that the image is in focus?* In terms of symbols: $f = 50$ mm; $s_o = 60$ mm; s_i ?

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}\ ,\ \ therefore\ \ \frac{1}{60} + \frac{1}{s_i} = \frac{1}{50}\ ,\ giving\ s_i = 300\ mm\ .$$

Notice that so long as the same units are used for all the known distances, the result is in these units.

Magnification

The transverse magnification is given by the ratio of image height to object height. A sign convention is used to measure distances perpendicular to the optic axis. Using this, inverted images are assigned a –ve magnification.



$$transverse\ magnification\,(M_T) = \frac{image\ height}{object\ height} = -\frac{s_i}{s_o}\ .$$

Magnification examples

- *What magnification did the slide projector in the example above give the projected picture?*
  - $s_i = 4$ m; $s_o = 4/15\ m$, hence,

    Magnification = $-s_i/s_o = \times -15$

  What magnification did the enlarger produce of the negative?
  - $s_i = 300\ mm;\ s_o = 60\ mm$, hence

    Magnification = $-s_i/s_o = -300/60 = \times -5$

The negative signs in these examples show that the images are upside down.

More Examples

Try these examples yourself to see if you've understood the workings of the imaging equations.

Hecht

(5.13) *Determine the focal length of a planar-concave lens (n=1.5) having a radius of curvature 100 mm. What is its power in dioptres*?
Answer: in the definition of power of a lens (p7), $R_1 = \infty$, $R_2 = +0.1$ m. Hence $D = -5$ dioptres and $f = -200$ mm.

(5.15) *What focal length lens is needed to produce an image on a screen 900 mm behind a lens when the object is 450 mm in front of the lens*?

Answer: from the thin lensmaker's equation, f = 300 mm.

(5.18) *Two +ve lenses of focal lengths 0.3 m and 0.5 m respectively are separated by a distance 0.2 m. A miniature frog rests on the principal axis 0.5 m in front of the first lens. Where is the final image with respect to the final lens*?

Answer: image from the first lens is at a distance 1.5/2 m = 0.75 m; object for second lens has $s_o = -0.75 + .2$ m $= -0.55$ m; hence image from second lens is at distance $s_i = 0.262$ m.

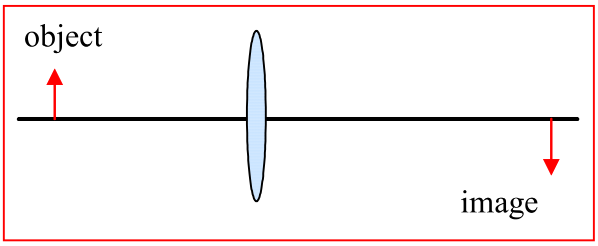
Summary of object and image sign convention and notation

- Object space covers the full range of coordinates from $+\infty$ to $-\infty$; likewise image space.
- Object distances s_o are measured in **object space** from the first principal point H_1; image distances s_i are measured in **image space** from the second principal point H_2. These principle points are defined in optics texts. All we need to know here is that for a simple surface, both points are located at the vertex. For a thin lens, both points are located at the centre of the lens. Those taking PX2505 will meet the full definition of 'principal points'.
- Optical elements (e.g. lenses) are considered one at a time.
- The object for an optical element is either a real physical object or the image created by the previous element.
- In Hecht's *real-is-positive* sign convention, **real objects** have s_o +ve. They are located *before* H_1, as measured in the direction that light is incident on the optical element. See page 4.
- **Virtual objects** have -ve s_i. They are located *after* the optical element, measured in the direction that light is incident on the optical element.
- **Real images** will appear on a screen placed at the image position. They occur *after* the optical element, as measured in the direction light leaves the element. They have +ve s_i.
- **Virtual images** occur *before* the optical element, as measured in the direction that light leaves the element.

1)

Real object +ve s_o

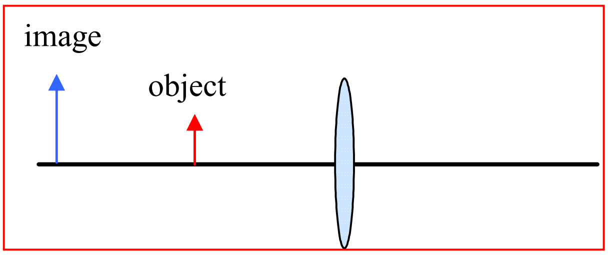
Real image +ve s_i



2)

Real object +ve s_o

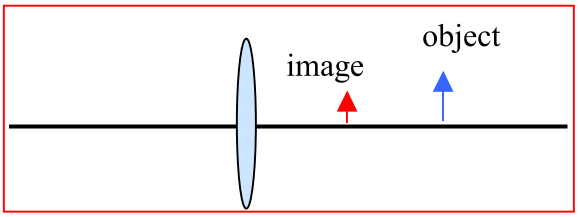
Virtual image -ve s_i



3)

Virtual object -ve s_o

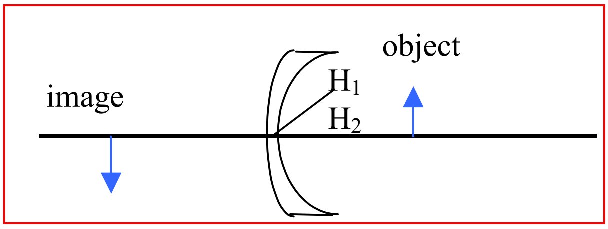
Real image +ve s_i



4)

Virtual object -ve s_o

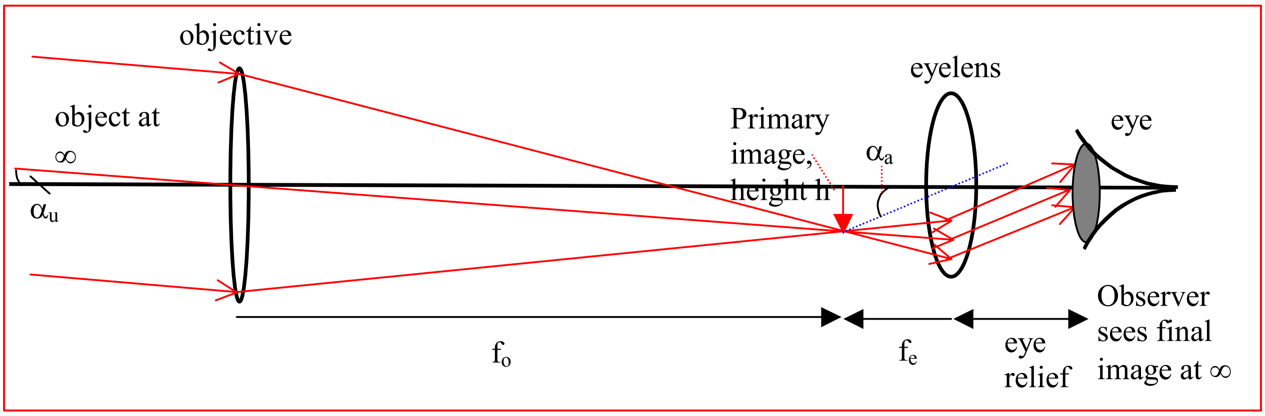
Virtual image -ve s_i



Telescopes

Telescopes are not only used by astronomers, gamekeepers, birdwatchers, Friends of the Earth, etc. but are also present on a diverse range of instruments such as theodolites, spectrometers, autocollimators, cathetometers, etc. Binoculars are essentially a pair of folded telescopes, the folding being achieved by prisms.

Most telescopes can be considered as developments of the Keplerian design, introduced in our first year practical course. This has two lenses, called the **objective** and the **eyelens**.



See Hecht fig 5.101. The diagram has been drawn with both object and final image at ∞.

Each lens is doing a special job. The first lens, called the *objective*, is there to form a real image of the object a long way off. The objective image in a telescope is called the *primary image*. Because light that goes through the centre of the lens is undeviated, you should be able to see that the further off this primary image is the larger it will be. The primary image is formed in the rear focal plane of the objective, because the object is at infinity. Hence, the longer the focal length of the objective (written f_o), the larger is the primary image. Other things being equal, you would expect longer telescopes to have a larger magnifying power and we'll soon see that this is so.

The second lens, called the *eyelens*, is just there to enable you to examine very closely the image formed by the objective. The eyelens functions exactly the same way as a "magnifying glass". When you look through the eyelens, you are focusing on something that is only a few tens of mm in front of your eye. The eyelens, though, is arranged to produce its image at infinity so that your eye is optically relaxed when focusing on it. This is simply achieved by placing the primary image in the front focal plane of the eyelens. The distance of the eyelens from the primary image is, therefore, f_e the focal length of the eyelens. The optics of the eyelens will be sketched out on the next page. Actually, more than one distance is important, not just the distance in front of the eyelens that what you are looking. When you look through an eyepiece your eye should be at a comfortable distance behind the eyelens and not jammed up against it. This comfortable distance is called the *eye relief*. It is part of the optical design to get the eye relief right for the average observer.

In summary, you'll see that both object and final image are normally at infinity for a telescope. The separation of the lenses is just the sum of their focal lengths, $f_o + f_e$. One of the most important features of a telescope as an optical instrument is its magnifying power. This has to be defined with a care. What does a telescope do? It produces an image in which any feature subtends a bigger angle than it would do if you looked at it unaided. The key idea is that the telescope produces *angular magnification*. The linear size of the image is not important. What is important is the angle it subtends, compared with what the object subtends when looked at directly.

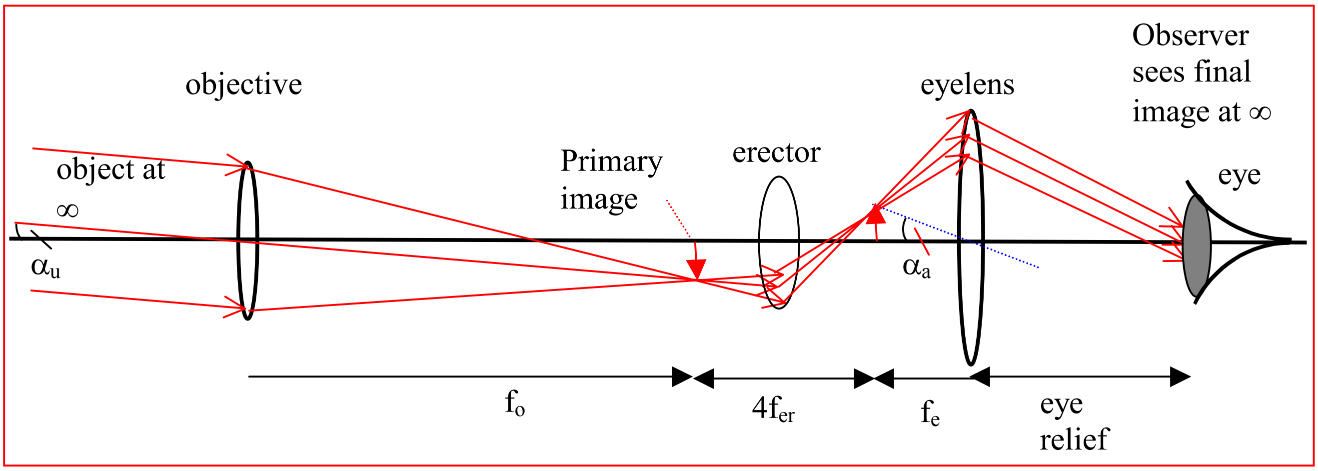
$$angular\ magnifying\ power\,(MP) = \frac{angle\ subtended\ by\ image\ at\ eye\,(\alpha_a)}{angle\ subtended\ by\ object\ at\ eye\,(\alpha_u)}$$

$$\alpha_a \approx \tan\alpha_a = \frac{h}{f_e}\text{, in the paraxial approximation .}$$

$$\alpha_u \approx \tan\alpha_u = -\frac{h}{f_o}\ .$$

$$\therefore MP = -\frac{f_o}{f_e}\text{, the } - \text{ve sign indicating an inverted image .}$$

Erecting telescope

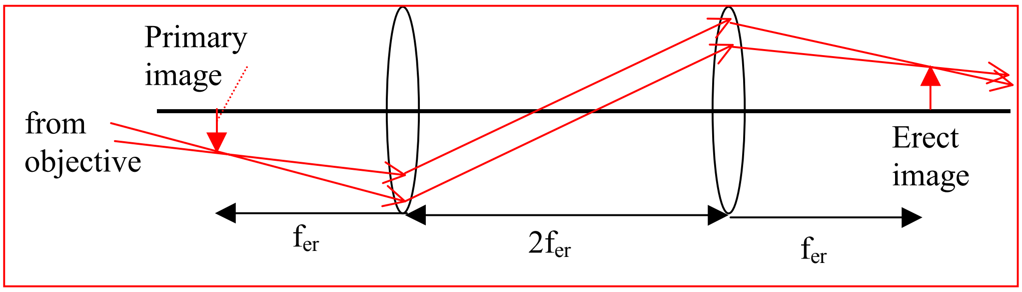


The 2-lens telescope is not really suitable for anything. The basic Keplerian telescope is good for astronomical work if the eyepiece includes a field lens, which we'll introduce soon. The terrestrial telescope needs an optical image erecting system. In principle, this can be provided by a single lens. Note that:

- The eye relief has got larger
- The erector limits the field of view
- A larger eyepiece diameter is needed

The 3-lens telescope does work, as you may remember from the first year lab. However, both erector lens and eyelens need a field lens, which we shall come to shortly.

The most common form of erector is a configuration sometimes known as **Rheita's erector**. It is essentially a telescope within a telescope, a simple pair of equal lenses of magnification ×-1. See Hecht fig. 5.103.



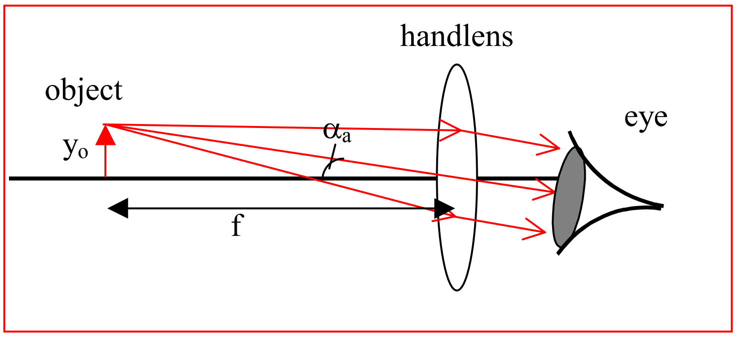
Even more common is the use of a dual prism for erecting. This folds the optical path length making the whole device so short that two can be placed side-by-side to make a pair of binoculars. The arrangement was illustrated in the 'fundamentals' notes that discussed reflection.

Eyelens - handlens - loupe - simple microscope

The eyelens in a telescope works exactly like a simple handlens. This lens owes its magnifying power to the fact that there is a closest distance that we can see objects clearly, called **the nearest distance of distinct vision**, d_o. If this were not the case, we could get a closer and closer look at objects simply by bringing them nearer our eye. For a *standard observer*, the nearest distance of distinct vision is taken as 0.25 m.

What the simple handlens does is give us an in-focus image of an object placed near the front focal plane of the handlens. If the front focal plane is quite close to your eye, then effective magnification is achieved.
The magnifying power MP is given

by $$MP = \frac{\alpha_{a(\text{aided})}}{\alpha_{u(\text{unaided})}}$$.



Since $\alpha_a = \frac{y_o}{f}$;

and $\alpha_u = \frac{y_o}{d_o}$,

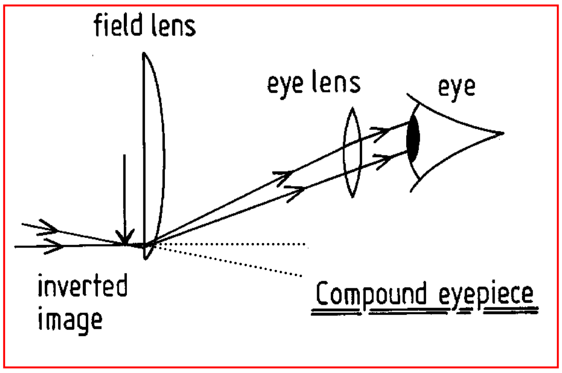
$$MP = \frac{d_o}{f} = 0.25D$$

Notice that by putting the object in the front focal plane, the final image is at ∞. Your eye is then relaxed when looking at it and it should be no strain to look through the lens for while. Hecht discusses the case when the image is at d_o, in which case the magnifying power is greater by +1. i.e. $MP = 0.25D + 1$.

e.g. *What is the power of a "×10" eyepiece?* $10 = 0.25\,D$ and hence $D = 40$ dioptres.

Compound eyepiece with field lens

The **field lens** deviates light that would have missed the following lens into it, as in this illustration from the first-year lab manual.



Field lens advantages:

- Allows a smaller diameter following lens to be used, thereby increasing image quality.
- Increases the unvignetted field of view.
- Can be used to control the eye-relief.
- The edge of the field lens acts as a field stop.

A **field stop** is an aperture in the object field or any image field that limits the field.

Real eyepieces have two or more lenses. One kind that is a popular basis of eyepieces found in telescopes is the Ramsden eyepiece, which is essentially that illustrated. It was invented by the great instrument maker Jesse Ramsden, over 200 years ago. A subtlety about the Ramsden eyepiece is that the focal lengths of the two components are chosen equal. This turns out to make the overall power of the lens minimally dependent on dispersion of the refractive index of the glass used and hence produces the best chromatic performance. Because the object being viewed through the eyepiece is close to the front of the first lens, a calibration scale can be placed at exactly the same place and seen to be simultaneously in focus with the image.

Merely being able to form in principle the correct image is not really the issue with real optical components. The real issues concern such things as fields of view and whether they are uniformly illuminated, and the quality of the image as determined by various measures of departure from the ideal. You can pay good money for eyepieces with names like Kelner, Plössl and Erfle that have been developed from the Ramsden eyepiece to reduce the defects inherent in a simple lens arrangement and to improve the image quality.

Summary

The mention of imaging defects is a natural lead-in to the subject of what can go wrong with images and why imaging isn't as simple as these pages so far suggest. All the image location relationships and concepts like focal lengths are based upon the assumption of the *paraxial approximations* that lead to spherical surfaces forming perfect images. The paraxial approximations are never exactly true and in consequence real images are never as nice as the ones calculated using the paraxial approximations. Any imperfection of an image is called an **image aberration**. We've met **chromatic aberration**, which is not connected with the paraxial approximations. Five recognised aberrations are connected with the breakdown of the paraxial aberrations. You may recognise some of their names: **spherical aberration**, **coma**, **astigmatism**, **field curvature** and **distortion**. Recognising their existence and learning what to do to reduce those that are important in any particular application has been a major effort in optical instrument design over the past 150 years. It is an effort that is still continuing.

Our course is intended as an overview of *Light Science* and unfortunately will have to leave out any discussion of this topic that is at the very heart of optical instrumentation. There are now computer programmes available that allow the relatively unskilled to look at aspects of real imaging that previously only the specialists could tackle. One of these programs is called *Winlens*. It is produced by a lens design company called 'Spindler and Hoyer', who have given us permission to put it on the undergraduate classroom servers and make it available. Anyone wishing to try it can find it alongside *Raytrace*, and you are welcome to explore many of the inbuilt examples that come with it.

There is another raft of questions about instrument design that this course isn't going to cover. These include topics like 'What field of view can be seen through a particular instrument?'; 'Is the field of view uniformly illuminated?' (it generally isn't unless you take some trouble); 'Is there an optimum place to put your eye behind an observing instrument?' (yes); 'How much detail can be seen in the final image produced?'; 'Are the diameters of the lenses wide enough to create a bright image?' and so on. Optical instruments aren't just a succession of lenses or lenses and mirrors. They are instruments with requirements of mechanical stability, ergonomic utility, integrity of materials used under a wide range of environmental conditions and requirements of contemporary design. It is part of the intention of the web-page exercise that you explore some of these facets.

What has this section tried to do? I have tried to give an idea of the basic principle of imaging, the concept of imperfect and perfect imaging and what the paraxial approximations are that define the perfect imaging of spherical surfaces. We have looked at the simplest case of a single imaging surface and seen how object and image positions are linked, how the focal points and focal lengths are defined and how magnification is produced. This led to the corresponding ideas for thin lenses and how it is comparatively simple to work out the image position for a lens of given power. Finally, we have looked at how the telescope has been developed from a simple 2-lens concept to a device using 5 lenses (for a terrestrial telescope). It is a short beginning to the subject of optical instruments, but I hope one you can build on in the future when you want to, or need to.

JSR