CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

1

Introduction: centred optical systems

1.1 The common properties of optical instruments

Optical instruments come in all shapes and sizes, from fly-on-the-wall surveillance cameras to 10 metre segmented astronomical reflecting telescopes, and in shapes from microscopes to sextants to periscopes to spectrographs to cine-projectors.

Whatever their purpose, they all have two things in common.

(1) They are *image-forming* devices, intended to make a picture, to form an image of a luminous source. The image may be on a cinema screen, on a photographic emulsion, on a CCD surface or on the retina of an eye.
(2) They are, with one important exception, *centred systems*. That is to say they comprise a series of curved surfaces of transparent materials or reflecting materials or both. The centres of curvature of the various elements all lie on a straight line called the *optic axis*. Light passes from the object, through successive elements until it emerges to form an image.

This is a slight over-simplification of course. There are occasional plane reflectors along the path, as in a periscope for example, but these are for convenience rather than for any peculiar optical properties they possess.

1.2 Optical elements

There are four basic optical elements: the lens, the mirror, the diffraction grating and the prism. What follows now concerns the first two of these. The others have chapters of their own.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

1.2.1 The lens

Lenses have been known since antiquity. The word lens is the Roman word for a lentil. A lens may be simple as in a magnifying glass or composed of several *elements* as in a telescope or camera, and the word is used indiscriminately. We are concerned for the moment with the simple element, which is a disc of glass with two polished surfaces, one or both of which may be curved. The curved surfaces are usually spherical, although non-spherical surfaces are becoming commonplace at the time of writing.

Its chief properties are (a) its *aperture* and (b) its focal length. The aperture is the diameter of the transparent area. The focal length is the distance from the *vertex* [1] of the lens of the image formed of an object at $-\infty$. If the curvature is concave the focal length is negative and no real image can be formed. Negative lenses are almost invariably part of a more complex lens system which does form a real image.

Lenses may be described as *biconvex*, *plano-convex* or *meniscus*, depending on the direction of curvature of the two surfaces. Two lenses may have the same aperture and focal length, but focal length depends only on the *difference* in the two curvatures, so that two lenses of the same focal length but different front surface curvature may have different image-forming properties. As a simple guide to these image-forming properties, remember that as a general rule refraction should be equally distributed among the various surfaces for a good sharp image. For example, the image quality of a plano-convex lens will depend markedly on which surface faces the object side.

Lenses are traditionally made of glass and much research has gone into making high quality optical glass, free from bubbles, striae [2] and stones (which cause scattering), free from stresses (which cause birefringence) and of high homogeneity (to avoid aberrations). Rare earths such as thorium and lanthanum have been incorporated to achieve high index and low dispersion (both desirable properties), and refractive indices between 1.4 and 1.9 are available (at a price); but polymer materials such as polystyrene and polycarbonates, with refractive indices $\sim$1.6, are also available for ordinary optical instruments.

Instruments which consist only of lenses, such as refracting telescopes and microscopes, are called 'dioptric' systems; those comprising only mirrors are 'catoptric' and when both are used they are 'catadioptric'. Reflecting telescopes and Cassegrain long-focus camera lenses come into this latter category.

[1] The word is used loosely here. Strictly *surfaces*, not lenses, have vertices, and the vertex is the point where the optic axis meets the surface.
[2] Known as sleaks in the trade.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

1.2.2 The mirror

Early mirrors were of bronze. In the seventeenth century alloys were found which were hard enough to be ground and polished like glass, and eventually *speculum metal* was perfected, an alloy of approximately two-thirds copper and one-third tin. With a high polish it reflected about 70% of the incident light. In 1835 Justus von Liebig (1801–1873) of the University of Giessen described a method for precipitating colloidal silver from an ammoniacal solution of silver nitrate [3] and it then became possible to deposit highly reflecting silver surfaces on to polished glass. These surfaces, when newly deposited, had a reflectivity of about 92% but as the industrial revolution expanded they tended to tarnish in the sulphurous atmosphere which developed around coal-burning cities.

With the development of vacuum techniques, evaporation of metals became possible and reflecting surfaces of aluminium were found to be chemically durable and with a reflectivity of about 88%. These were improved further by *overcoating* them with evaporated films of silicon monoxide to give a hard and washable surface finish. Reflectivity improves in the infra-red, and gold surfaces, as well as being chemically durable, show a reflectivity in the region of 99%. In the far ultra-violet aluminium may be overcoated with magnesium fluoride to improve its reflectivity which otherwise falls rapidly as the wavelength shortens.

Mirror surfaces may be plane, spherical, paraboloidal or elliptical, and some current mirror-forming techniques allow various, even more elaborate surfaces to be formed by computer-controlled grinding and polishing, chiefly for large astronomical telecopes.

1.3 Concepts in optical instrument design

Let us consider first some features common to all optical instruments, which we need in order to understand how they work. We begin with the description of light.

1.3.1 Rays of light

First of all is the assumption that light travels in straight lines, *rays*, emitted by an object, passing through various lenses and mirrors and converging finally to form an image on a focal surface: a screen, a photographic plate, a CCD or a retina.

[3] An early version of Tollens' reaction.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

This assumption is one of the 'convenient fictions' of physics, useful for describing what happens and for predicting the behaviour of light in instrument design; but the model fails in the end to take account of diffraction and interference in the *minutiae* of image formation. Modern optical design requires the *wave theory* of light propagation to account for the fine detail in an image.

Wave theory supposes that light is an electromagnetic field which propagates outwards from a point source in *wavefronts*, spherical surfaces on which the electric field is a maximum, and which expand outwards at the speed of light. The perpendicular distance between successive wavefronts is the wavelength, measured either in angstrom units (1 Å $= 10^{-10}$ m, roughly the diameter of an atom) or in nanometres (10^{-9} m), the standard SI unit.[4]

A lens or a concave mirror can be used to convert part of these diverging wavefronts into spherically converging wavefronts. In a lens, the process is governed by *refraction*, the speed of light being less in a dense transparent medium than in air.

However, it is easier to describe the process of refraction by connecting an object and its image by a ray, the direction changing at each refracting surface.

1.3.2 Optical ray diagrams

It is a convention, not always observed, that diagrams showing rays passing through optical systems show the optic axis horizontal, and show rays coming from the left, the object side or input side, and leaving the system on the right, the image or output side. The usual Cartesian formalities are observed and lengths measured to the left of the lens – the input side – are considered to be negative and those measured to the right are considered positive.

1.3.3 Refraction and the law of sines

The earliest known exposition of the law governing the refraction of light at a surface was by the mathematician Ibn Sahl (940–1000) of Baghdad in the tenth century. It was rediscovered several times in Europe following the Renaissance and today is known in English-speaking countries as *Snell's law* after Willebrord Snel van Royen (1580–1626) of Leiden. It is also more simply known as the 'law of sines'.

In 1657 Pierre de Fermat (1601–1665) published his *Principle of Least Time*, which asserts that light travels from object to image by that path which takes

[4] Physicists tend to use nanometres, while astronomers cleave to the old angstrom unit. Both units will be found in this book.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

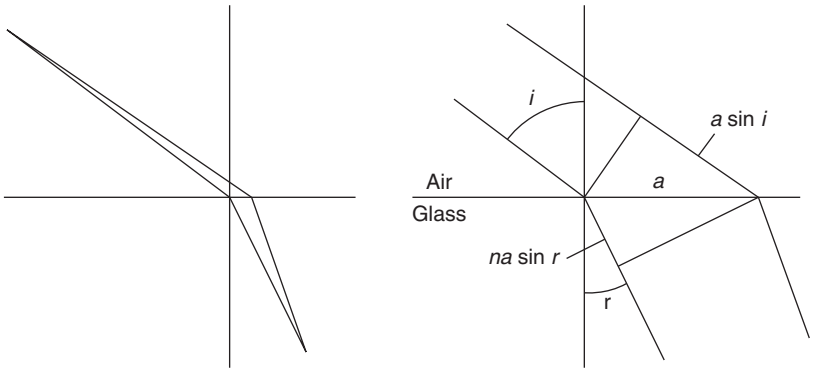

Figure 1.1 The law of sines (Snell's law). When the optical path length is near a minimum, a small change in direction makes no change in the total path length. The small extra path in air is exactly matched by the small diminished path in glass. When the total change is zero, the law of sines follows immediately.

the shortest time. From Fermat's principle it is a simple exercise in differential calculus [5] to derive Snell's law, that the sine of the angle of incidence divided by the sine of the angle of refraction is a constant, the *refractive index*, usually [6] denoted by the letter n:

$$n = \frac{\sin i}{\sin r}$$

The derivation of the 'law of sines' is simple. The small change of *time*, Δt, is given by:

$$\Delta t = a \sin i - n.a \sin r$$

since the speed of light is less in glass by a factor n. At a minimum the change in time is zero, which gives the 'law of sines' immediately.

1.3.4 Optical paths

We occasionally need three different path lengths to describe the passage of light through refracting media.

(1) The *geometrical path*, the path actually laid out on the drawing board.
(2) The *optical path*, the geometrical path multiplied by the refractive index of the medium through which each section of the path the ray is travelling.

[5] And is evidence that Fermat had the idea of the *differential* calculus before Newton and Leibnitz, even though he never made the connection with integration.
[6] But in many older books by the Greek letter μ.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

Alternatively (and more accurately), it is the number of wavelengths of light between object and image, multiplied by the vacuum wavelength of the light. All the rays between an object point and its image point must have the same optical path length.

(3) The *reduced path*, the geometrical path *divided* by the refractive index. This is the path length you would measure if you were using a rangefinder to find the apparent distance to an object at the bottom of a swimming pool.

1.3.5 Apertures and stops

The *aperture* of a lens is self-evident. It is the diameter of the transparent part of the lens. The *physical* diameter is a millimetre or two greater than this to allow for 'turn-down' at the circumference in the manufacturing process, and to allow for the 'cell' or retaining device which holds the lens in place.

Stops are opaque screens normal to the optic axis, with circular holes to limit the size of the ray bundle passing through the instrument. They may fulfil several purposes. They may be positioned to control the aberrations of the system, they may be 'anti-glare' stops, used to reduce scattered light from various internal surfaces which would otherwise reduce the contrast of the image and prevent accurate photometry, and if they are variable in diameter as in a camera iris diaphragm, they control the intensity of light forming the final image.

1.3.6 Pupils

There is always one defining stop in an optical instrument, the *iris*. It may simply be the rim of a lens or it may be somewhere inside the sytem, in which case its image, seen through all the elements on the input side, is the *entry pupil* of the system and the corresponding image on the output side is the *exit pupil*. Pupils are not necessarily inside the system. In a telescope for instance, the exit pupil is several millimetres to the right of the last element of the eyepiece: it is where you put your eye-pupil when inspecting a distant object. All the rays of a ray bundle must pass through the pupils of a system.

1.3.7 Relative aperture

This is popularly referred to as 'focal ratio' or 'F-number' and is denoted by F. It is the ratio of the lens focal length to its aperture. As such it is a measure of the intensity of light at an image point and consequently is useful in photography and photometry. It is usually expressed as a fraction of the focal

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

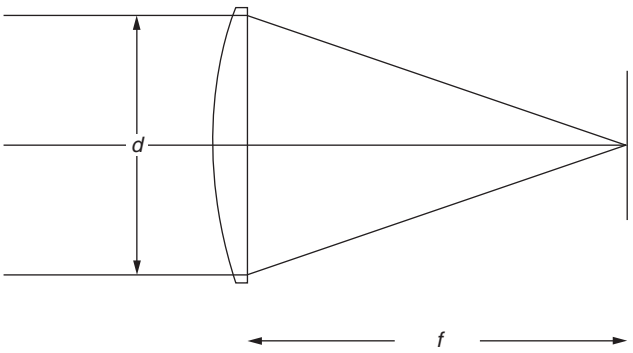

Figure 1.2 The focal ratio or *relative aperture* or F-number. This is the ratio of the focal length of the lens to its clear aperture. In general it is defined as the ratio of the focal length to the pupil diameter and is expressed as, for example, f/8 or f/6.3.

length so that, for example, a lens with relative aperture of f/6 has a clear aperture of one-sixth of its focal length. In complex lenses with many elements it is the exit *pupil* diameter as a fraction of the focal length which measures the focal ratio.

1.3.8 Numerical aperture

In a microscope the object is extremely tiny. In order to see a great enlargement, the light from the object must be collected from the greatest possible solid angle and turned into a large image at a much greater focal ratio. For microscope objective lenses we refer to its *numerical aperture*, defined as:

$$\mathrm{NA} = n \sin \theta_m$$

where θ_m is the angle which the ray from the object to the edge ('the margin') of the objective lens makes with the optic axis *at the object*, and n as usual is the refractive index in the object space, i.e. *before* the first surface of the lens. Numerical aperture is considered at greater length in the chapter on microscopes. So-called *immersion objectives* in microscopes take advantage of the refractive index of oil to increase the NA by a factor n. For most practical purposes, the numerical aperture is related to the relative aperture or F-number by

$$\mathrm{NA} = 1/2F$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

1.3.9 Étendue

This indicates the radiated power which can be transmitted through the system from a distant extended source of light. It refers mostly to photometers and radiation detectors and is a measure of the flux-collecting capability of an optical instrument. It can be defined in two ways.

(1) It is the exit pupil area multiplied by the solid angle subtended there by a detector.
(2) It is the sensitive area of a detector multiplied by the solid angle subtended there by the exit pupil.

One must be careful here: each pixel of a CCD is a separate detector. The rate at which information is gathered in a properly designed experiment will depend on the *total* étendue of the device, i.e. on the total sensitive area of the CCD multiplied by the solid angle subtended by the exit pupil.

The concept of étendue as the product of area × solid angle is one of the most profound and important concepts in optics and radiometry, and one which determines the design of many types of optical instrument. Étendue is an invariant of an optical system, and is *conserved* throughout various stages of re-imaging.

1.3.10 The Helmholtz–Lagrange invariant

This is best described with a diagram (see Figure 1.3). If objects are imaged and then re-imaged throughout a system, then the height H_i of the ith image, multiplied by the angle from the image base to the rim of the next aperture A_{i+1}, and multiplied by the refractive index of the medium between, is a constant throughout the system. There are in fact two such invariants in any instrument, defined by two orthogonal planes containing the optic axis. If there is

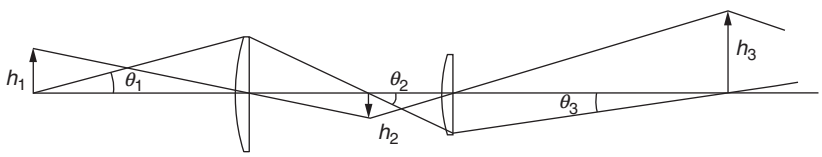

Figure 1.3 The Helmholtz–Lagrange invariants of a centred system. The product of the image size, h_n and the angle θ_n which the next, or previous, aperture subtends at the image position, is a constant of the system. There are two of these invariants for any system, but if there is circular symmetry about the optic axis, they are identical. In an asymmetrical system such as a spectrograph, they are not.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

rotational symmetry they are the same but if the instrument is *anamorphic* they will be different, each being separately conserved though the various stages. Taken together they represent the étendue.

A simplified and practical version is to say that the image size divided by the focal ratio is a constant. People who photograph through telescopes will understand this one: if you want a brighter image, i.e. a smaller focal ratio, you have to accept a smaller image. The amount of light coming through is always the same. Étendue is conserved.

1.3.11 The vignette

This is a term which derives from the word 'vine', which was used classically, suitably entwined, as a tasteful decorative border to a fresco, and subsequently adopted by Victorian photographers to describe the border of a portrait which softened the hard outline by allowing the image to fade gradually at the edges. In optics a vignette is something highly undesirable, and usually indicates that an instrument has not been designed properly or is being used beyond its design limits.

An optical system is said to vignette when there is no single controlling stop to regulate the passage of light through it, so that the intensity of an image varies from point to point from the centre outwards.

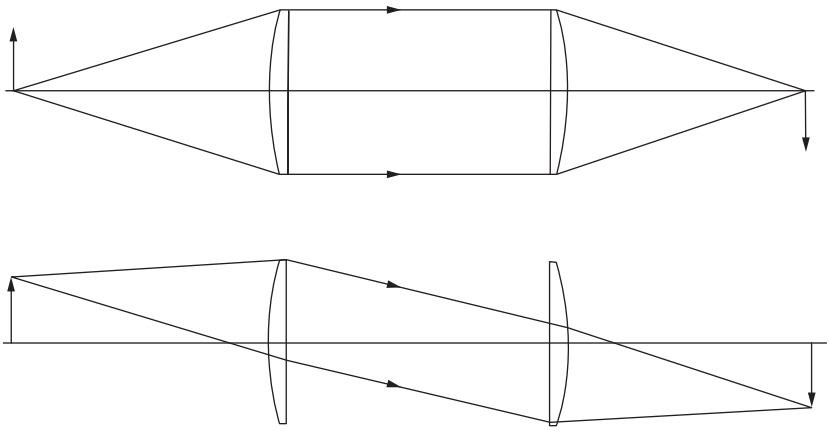

Figure 1.4 The vignette. The flux of light through this system decreases as the obliquity of the ray bundle increases so that the second lens cannot be completely filled with light. A stop half-way between the two lenses would ensure that the intensity is constant over a wide field, and a field stop may be placed on the image plane to mark the limit of the field of uniform illumination.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-05054-9 - An Introduction to Practical Laboratory Optics
J. F. James
Excerpt
More information

1.3.12 Conjugate points

These are defined as points connected by ray bundles. A point on an object is *conjugate* to the corresponding point on its image. In this sense an extended image is, in the mathematical sense, a *map* of the object because there is a one-to-one relation between points of each. The idea is not constrained to object and image points. Points on the entry and exit pupils of a system are conjugate. Light which has passed one point must also necessarily pass through its conjugate. When an instrument has several optical stages it is important that the exit pupil of one be the entry pupil of the next to ensure that all light passes through the system without vignetting.

1.3.13 Ray bundles

Rays from an object point to its conjugate image point spread out and pass through different parts of the optical system. All have the same optical path length [7] and collectively form a ray bundle. An important member of this bundle is the *chief-ray*, generally the central ray of the bundle, which is distinguished by the property that *the chief-ray of a bundle crosses the optic axis at the pupils of the system*.

The chief-ray is of particular importance when interference filters are used to do spectrometry. These filters have a narrow spectral pass-band and the peak wavelength transmitted depends on the angle of incidence. The pass-band of the filter moves to shorter wavelength as the angle of incidence increases. The effect is serious, so that it is important that the chief-ray of every ray bundle be everywhere normal to the surface as it passes through the filter.

1.3.14 Telecentricity

A system is said to be *telecentric* when all the chief-rays entering or leaving the system are parallel to the optic axis. It can be ensured for rays entering the sytem by putting a stop at the rear focus of the objective lens. It is of particular practical importance in instruments such as profile projectors to ensure that a true silhouette of an object is produced at the image position. Such systems are said to be 'telecentric on the object side'. In Figure 1.5 it is necessary for the chief-rays of all bundles to pass through the filter on the right parallel to the optic axis. It is achieved here by placing a lens near the focal plane, of the same focal length as the objective on the left.

[7] Give or take a few wavelengths, owing to aberrations.

© in this web service Cambridge University Press www.cambridge.org