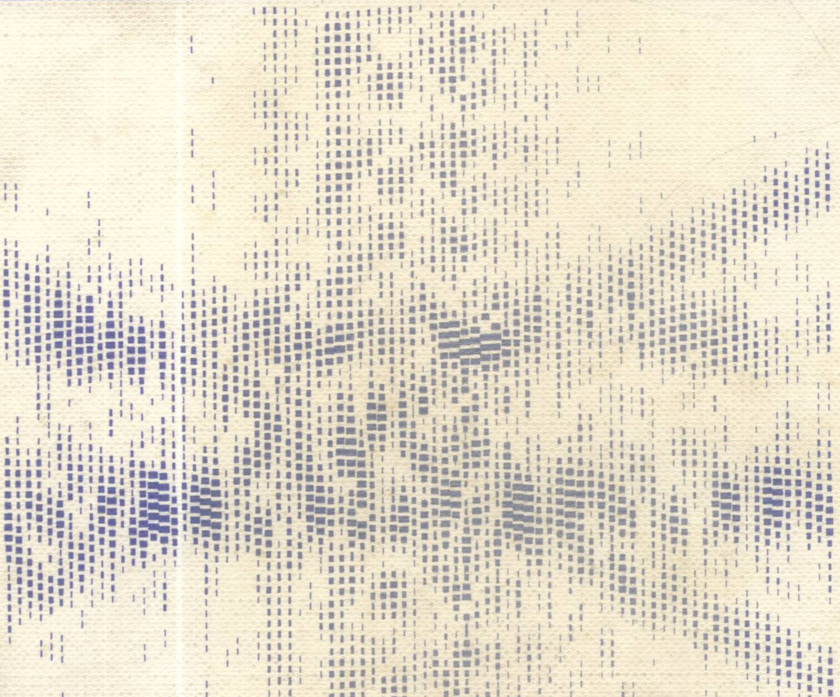
PHYSICS PROGRAMS

1. Optics



Edited by

A. D. BOARDMAN

Physics Programs

Optics

Edited by
A. D. Boardman
Department of Pure and Applied Physics, University of Salford

A Wiley–Interscience Publication

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto

British Library Cataloguing in Publication Data:

Physics programs.
Optics
1. Physics—Programmed instruction
I. Boardman, A. D.
530′.07′7 QC21.2 80–40123

ISBN 0 471 27729 0

Typeset in Northern Ireland, at The Universities Press (Belfast) Ltd. and printed by Pitman Press, Bath

Physics Programs

Optics

Preface

This small book is the optics section of the larger textbook called *Physics Programs* that covers the four areas, optics, magnetism, solid state/quantum physics, and applied physics. Each chapter given here is self-contained, with enough theory given for the topic discussed, and the associated computer programs, to be well understood. The programs are guaranteed in the sense that they are copied directly from fully working source texts on the computer. They can be used, possibly with minor adjustments, on any computing system. If what is required is a classroom demonstration, or the engagement of a class in a simple sequence of exercises, then the programs may be used without understanding the coding. The programs are, however, liberally strewn with comments so that they can be used for more advanced projects in which an understanding of the program is required.

The material given here covers ray tracing and lens aberrations, computer-generated holograms, and a technique currently used in the study of surface polaritons. It is hoped that this small set of programs will be of great interest to the many students and teachers of optics. All of these chapters are suitable for undergraduates at some stage in their studies, although, of the three, the chapter on holograms is the most sophisticated.

Salford A. D. BOARDMAN

Contents

Preface vii

Ray Tracing and Lens Aberrations 1
P. A. Young

Attenuated Total Reflection Analysis of Surface Polaritons 47
G. C. Aers and A. D. Boardman

Computer-generated Holograms 79
A. D. Boardman and M. E. S. Chapman

Physics Programs
Edited by A. D. Boardman

CHAPTER 1

Ray Tracing and Lens Aberrations

P. A. Young

1. INTRODUCTION

A great deal of our knowledge of the physical and biological world comes from our use of microscopes, telescopes, cameras, and other optical devices that use light waves to form images of greater brightness or detail than we can obtain from our eyes alone. A basic part of the design of such optical systems is the tracing of rays through them and the determination of their deviations from perfect imagery, the so-called aberrations.

The two concepts that are important in discussing the propagation of light and the formation of images are the wavefront and the ray. A wavefront is defined as the locus of points which the light has taken the same time to reach, and the ray as the direction in which the light energy is travelling. In isotropic materials, such as glass, the rays are perpendicular to the wave-fronts.

Because light is a wave-motion it can be diffracted, and the concepts of ray and wavefront breakdown in situations in which diffraction is important. These include points at which light waves converge to form images. Diffraction, in fact, makes it impossible to realize a ray physically, by, for example, passing light through smaller and smaller pinholes set so as to define a direction of energy travel; nevertheless, it remains a useful idealization and it is the propagation of these ideal rays, and associated waves, that is the province of geometrical optics.

1.1 Fermat's principle

A fundamental link between the wavefront and the ray is provided by Fermat's principle of least time that, in words, is

the path taken by a light ray is such that the time of travel is a minimum.

Subsequent work has shown that, although a minimum is often involved, a better statement is that the time should be stationary (maximum, minimum, or inflexion point). In the notation of the calculus of variations this is written as

$$\delta \int dt = 0, \tag{1}$$

where t is the time and the integral is taken between suitable limits. Now if the speed v of the wave, i.e. ds/dt, and the refractive index $n = c/v$ is introduced then equation (1) has the form

$$\delta \int n \cdot ds = 0 \tag{2}$$

where c, the constant velocity of light *in vacuo*, has been deleted. The quantity $\int n \cdot ds$ is called the optical path along the ray. For regions of constant refractive index $\int n\, ds$ is ns which is the familiar rule that

Optical path = refractive index × geometrical path

It follows from Fermat's principle that light rays obey the observed laws of geometrical optics,[1] and in particular for refraction, if a ray (called the incident ray) in a medium of refractive index n strikes a surface, at an angle I to the normal, that separates it from a medium of refractive index n' then it continues as a refracted ray at an angle I' to the normal such that

(a) the incident ray, the refracted ray, and the normal lie in one plane;
(b) the angles and refractive indices obey Snell's law, viz.:

$$n \sin I = n' \sin I'. \tag{3}$$

2. RAY TRACING IN THE PARAXIAL APPROXIMATION

It is a consequence of the wave nature of light that just as rays are a physical impossibility so also is the ideal optical system defined as one in which all rays leaving a single object point converge on (or appear to diverge from) a unique image point, and even within the realm of geometrical optics the quasi-ideal system (ignoring diffraction) can only be realized in a few cases, of which the plane mirror is the simplest example.[2] The situation, however, is not as bad as it seems because sufficiently close approximations to ideal systems can be obtained as to be practically useful, and it is the closeness to ideal imagery that is specified by the aberrations. Furthermore, these may be determined by ray tracing using the laws of geometrical optics.

However, before any detailed design is carried out to find the exact form and nature of the deviations from perfect imagery, as revealed by the actual paths of the rays, it is useful to use what is known as the paraxial or Gaussian approximation in which all rays are assumed to be close to the axis of a system and all angles are assumed to be small. These assumptions lead to the paraxial equations which can be used, for instance, to determine: (1) the system focal length; (2) the position of the ideal, or Gaussian, image from which the deviations can be measured; (3) an estimate of the size of the aberration of the image as measured by the difference in optical path between the actual ray and the Gaussian image ray.

Assume that the system has a unique axis of symmetry and that a paraxial ray is described by the two parameters u and y, as shown in Figure 1a; u is the angle the ray makes with the axis, and y is the distance from the axis of a point on the ray, usually at one of the optical surfaces. (Note that for small u and y, no distinction is made between the position of a ray intercept on a surface or on its tangent plane.)

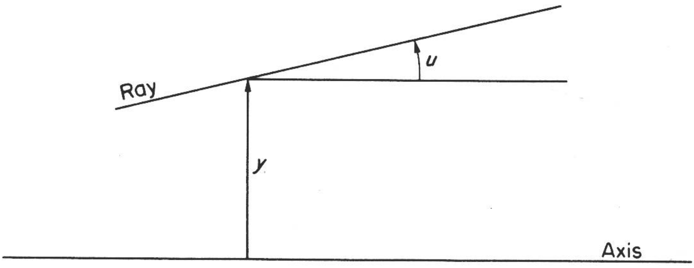


Figure 1a. Ray parameters

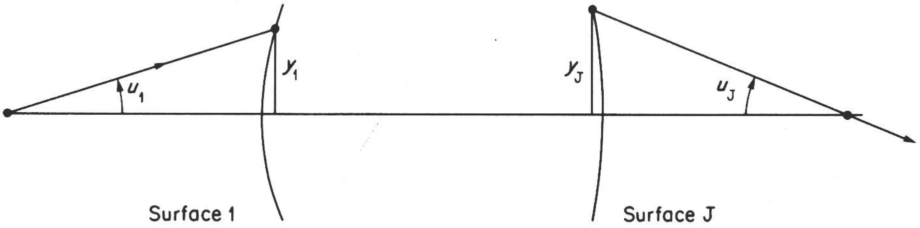


Figure 1b. Initial and final ray parameters

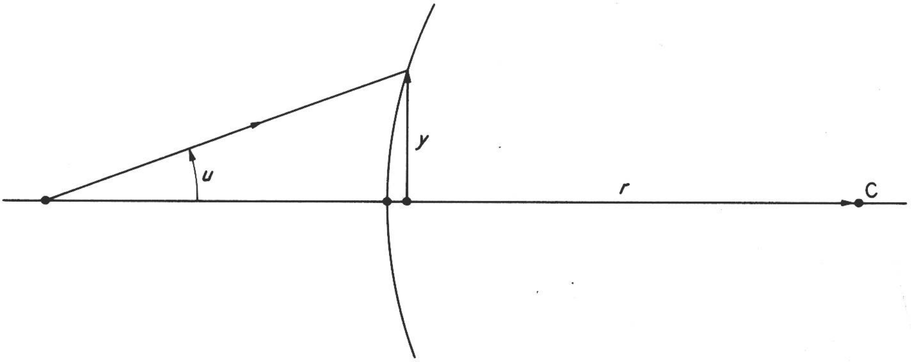


Figure 1c. Positive parameters

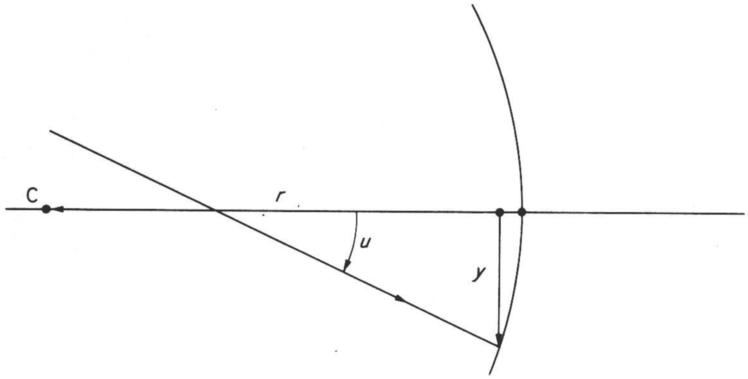


Figure 1d. Negative parameters

The tracing of a ray means, as shown in Figure 1b, the determination of the angle u_J and height y_J at which the ray leaves the final (Jth) surface of the system, given the angle u_1 and height y_1 at which it enters the first surface.

The rays are assumed to traverse the system from left to right and the normal Cartesian conventions on the signs of distances and angles apply. The radii r of surfaces are positive if they are convex to the left. Figures 1c and 1d, in which C is the centre of curvature, show situations in which all the quantities are positive or negative respectively. The procedure for ray tracing is broken down into two parts, these are refraction at a surface and transfer from one surface to the next.

2.1 Refraction and transfer

The sth surface AB of radius r_s dividing two regions of refractive indices n_s and n'_s is shown in Figure 2a. Suppose that a ray makes axial angles u_s and

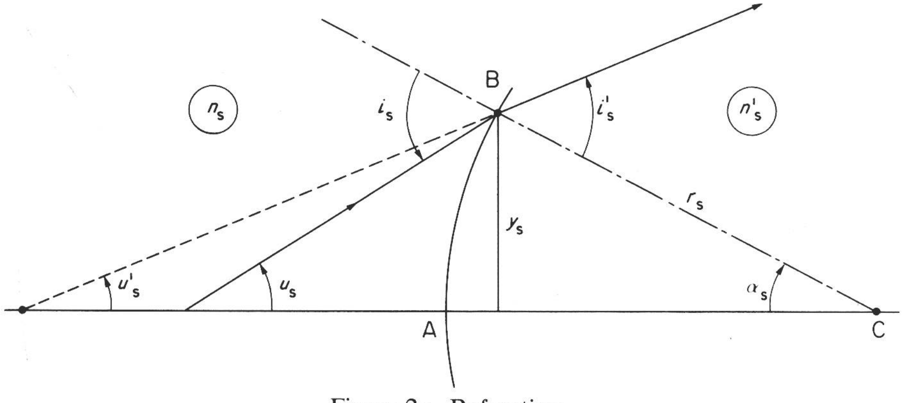


Figure 2a. Refraction

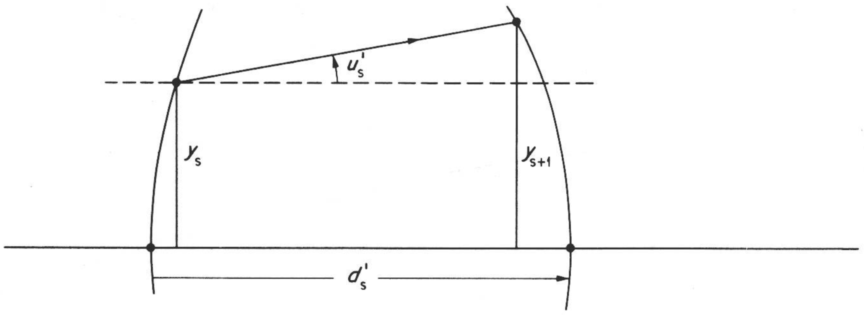


Figure 2b. Transfer

u'_s before and after refraction respectively, and have corresponding angles of incidence and refraction of i_s and i'_s. Furthermore, suppose that C is the centre of curvature of the surface and that the incident height y_s subtends an angle of α_s at the centre. (It is a consequence of the sign convention on distances that α_s is *positive*.) Then, from the diagram,

$$\begin{aligned} i_s &= \alpha_s + u_s, \\ i'_s &= \alpha_s + u'_s. \end{aligned} \tag{4}$$

Now for small angles, Snell's law, equation (3), becomes $n'i' = ni$ so that

$$n'_s(\alpha_s + u'_s) = n_s(\alpha_s + u_s). \tag{5}$$

Also, for small angles, $\alpha_s = y_s c_s$ where $c_s = 1/r_s$ is the curvature of the surface. (Note that $c_s \to 0$ when $r_s \to \infty$ and can thus be used numerically for plane surfaces.) Hence, on substituting for α_s we find

$$n'_s u'_s = n_s u_s - y_s K_s. \tag{6}$$

where K_s, the power of the surface, is

$$K_s = (n'_s - n_s)c_s. \tag{7}$$

After a ray leaves the surfaces at angle u'_s and height y_s, it proceeds to the $s+1$ surface, a distance d'_s away along the axis, and intercepts it at height y_{s+1}. In Figure 2b, it is seen, for small angles and heights, that

$$y_{s+1} = y_s + d'_s u'_s. \tag{8}$$

2.2 Ray tracing procedure

A given ray is traced through a system by successive use of equations (6) and (8), noting that at each surface $n_{s+1} = n'_s$ and $u_{s+1} = u'_s$. The trace is started in one of two ways:

(1) An initial axial point O on the object, a distance l_1 from the first

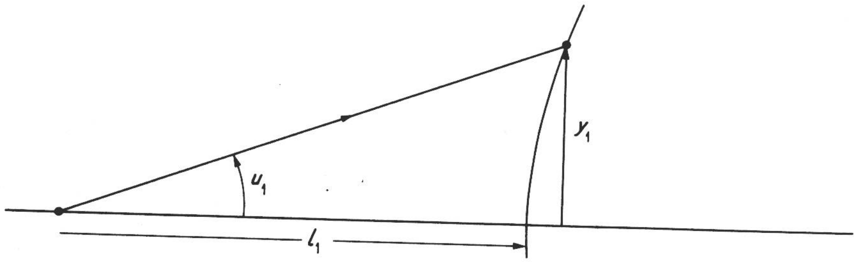


Figure 2c. Starting with a ray at a given incidence height

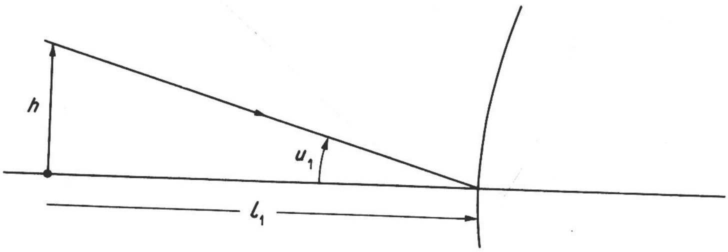


Figure 2d. Starting with a ray from a known object height

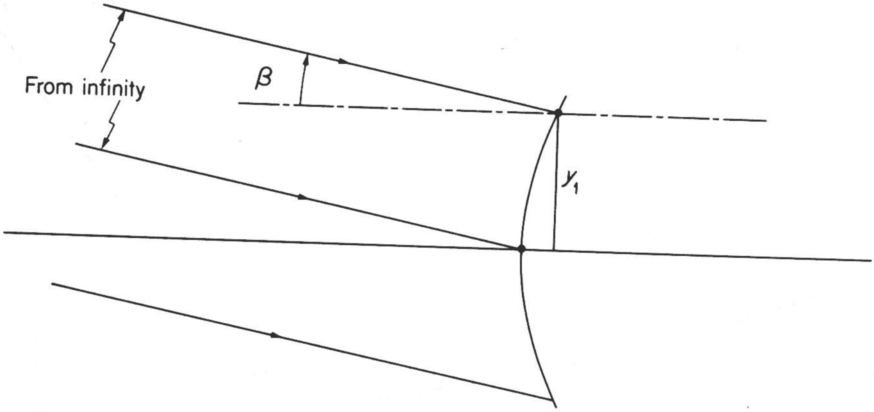


Figure 2e. Rays from infinity, field angle β

surface, is chosen together with an incidence height of y_1 as shown in Figure 2c. Then

$$u_1 = y_1/l_1, \tag{9}$$

and the trace starts with a refraction.

(2) An object of height h is selected at a distance l_1 (Figure 2d) and a ray through the centre of the first surface is chosen, then

$$\begin{aligned} y_1 &= h, \\ u_1 &= -h/l_1, \end{aligned} \tag{10}$$

and the trace starts with a transfer.

Note that for objects at infinity h and l_1 simultaneously tend to infinity but the field angle $\beta = -h_1/l_1$ remains constant, as shown in Figure 2e. If a ray at height y_1 on the first surface is chosen, then with $u_1 = \beta$, the trace starts with a refraction.

3. STOPS AND PUPILS

It should be pointed out that the 'surface' referred to above can be simply a circular hole for which $n' = n$ and $c = 0$. Such apertures, or stops, are often placed in optical systems to limit the extent of the beams passing through, and also to control the aberrations. Amongst the various stops and lens apertures in an optical system there will be one, or its image, which seen from the object side subtends the smallest angle at the axial object point: this is known as the entrance pupil and it limits the maximum angle that the rays can make with the axis and still pass through the system.

If the pupil is an image of a stop the corresponding stop is called the aperture stop, if the pupil is real it is itself the aperture stop. The image of the entrance pupil as seen from the image side is the exit pupil. For an off-axis object point the pupil will still, to a large extent, control the angular

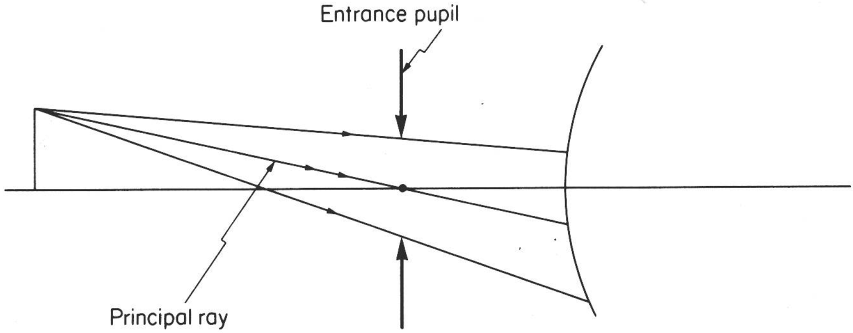


Figure 3a. Entrance pupil and principal ray for a real pupil

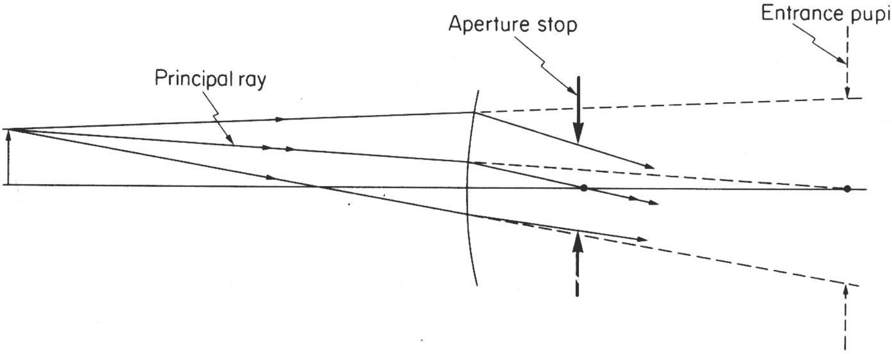


Figure 3b. Entrance pupil and principal ray for a virtual pupil

aperture of the rays that pass through the system. The ray through the centre of the entrance pupil, that defines a cone of rays that can pass through the system, is an important ray and is known as the principal ray. Figures 3a and 3b show stops, pupils, and principal rays in an optical system; the principal ray is indicated by double arrows.

4. FOCAL LENGTH

If the initial angle u_1 is zero and the initial height y_1 is finite then one can determine the focal length. After passing through the system the ray will (except in what are known as telescopic systems) leave the final surface at a finite angle, u'_J, and height y_J and pass through the focal point F'. The point on the axis directly below the intersection point of the initial ray and the final ray is the principal point P', distance p' from the last surface, whilst the distance from the last surface to F' is known as the back focal length, bfl'.

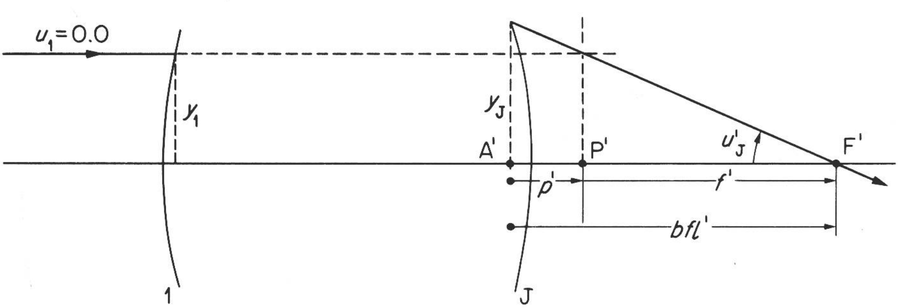


Figure 3c. Focal parameters, P', F', p', f', bfl'

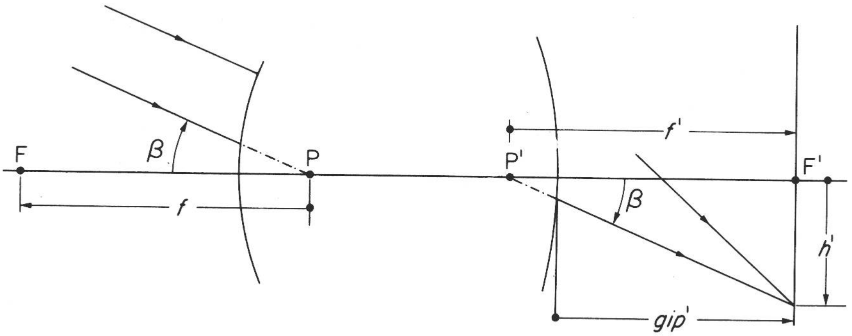


Figure 3d. Gaussian image plane and size for an infinite object

In Figure 3c, it is seen that

focal length:	$f' = -y_1/u'_J$,	(11)
back focal length:	$bfl' = -y_J/u'_J$,	(12)
position of principal point:	$p' = bfl' - f'$.	(13)

If a ray be traced from infinity back through the lens ($u'_J = 0$) then similar points F and P and distances f, p, and ffl (front focal length) are defined on the object side. In systems in which, as is usually the case, $n_1 = n'_J$ the focal lengths f and f' are numerically equal and the principal points, P and P' are also the so-called nodal points such that a ray directed towards P on the object side leaves the system on the image side as if directed away from P'. This is shown in Figure 3d.

4.1 Gaussian image

If the position of the Gaussian, or paraxial, image with respect to the last, Jth, surface of the system, is gip' and the size of the image is h' then

(1) for an object effectively at infinity, as shown in Figure 3d, the image distance is

$$gip' = bfl' \tag{14}$$

and its size is

$$h' = f'\beta; \tag{15}$$

(2) for an object at a finite distance l, and of size h, the image position is found by tracing a paraxial ray from the axial object at any non-zero angle u_1 and compatible $y_1 = lu_1$. The image position is then given, as shown in Figure 3e, by

$$gip' = y_J/u'_J \tag{16}$$

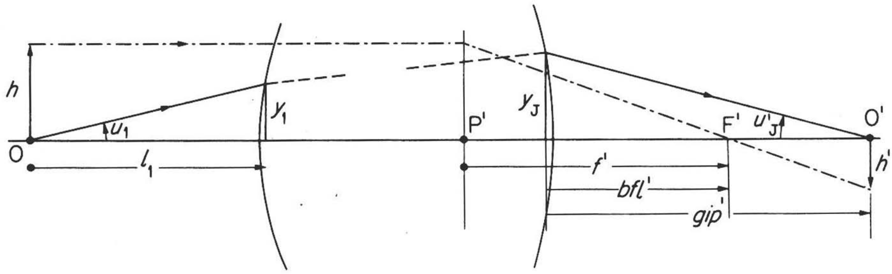


Figure 3e. Gaussian image plane and size for a finite object

The size of the image can be seen, from Figure 3e, to be given by

$$\frac{-h'}{F'O'}=\frac{h}{P'F'}$$

so that

$$h'=-(gip'-bfl')h/f'. \tag{17}$$

5. LENS ABERRATIONS

In a perfect optical system all rays leaving a point object, O, converge on (or diverge from) a point image, O'. If we apply Fermat's principle to the rays travelling from the object to the image via the system then any ray takes a minimum time so that all rays from O to O' must take the same time; this fact is more conveniently stated as: the optical path along all the rays from object to image are equal. A suitable way to measure the defects of any real optical system is therefore in terms of the differences in the optical paths of

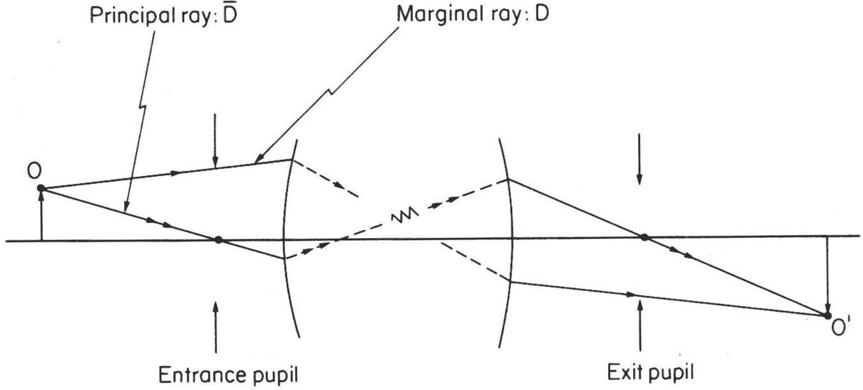


Figure 4a. Principal ray and marginal ray

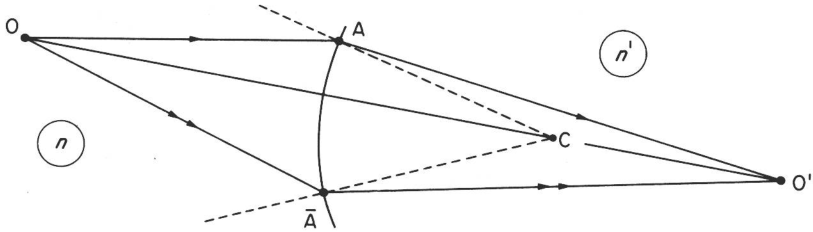


Figure 4b. Wave aberration due to refraction

the rays, which is done by comparing the optical path of any given ray with that of a reference ray, the latter being chosen as the principal ray. This is shown in Figure 4a and the conventional choice of a measure W, of the aberration, is

$$W = \bar{D} - D, \tag{18}$$

where $\bar{D}$ is the total optical path of the principal ray and D is the total optical path of the given ray.

5.1 Aberration due to one refraction

In Figure 4b O is an object and O' is its image in a surface separating media of refractive indices n and n'. $O\bar{A}O'$ is a principal ray and OAO' is a given ray that, since it is usually taken to be a ray at the edge or margin of the pupil, is called a marginal ray. Then

$$\begin{aligned} \bar{D} &= n \cdot O\bar{A} + n' \cdot \bar{A}O', \\ D &= n \cdot OA + n' \cdot AO', \end{aligned} \tag{19}$$

and the difference in optical path, defined by equation (18), is

$$\begin{aligned} W &= n'(\bar{A}O' - AO') + n(O\bar{A} - OA) \\ &= n'(\bar{A}O' - AO') - n(\bar{A}O - AO) \\ &= \Delta\{n(\bar{A}O - AO)\}, \end{aligned} \tag{20}$$

where Δ means take the difference of the value in the expression after and before refraction.

5.2 Spherical aberration

The most important aberration, which is present even for axial objects, is spherical aberration. It is also the one that is most readily calculated from equation (20). Spherical aberration takes the general form of producing differing focusing positions for different incident heights, as shown in Figures 5a and 5b. The focus F_G is the Gaussian focus (F') and is the one that is