

CHAPTER FOURTEEN

Analysis of Optical Systems I

Analysis of Optical Systems I

14.1 Introduction

In this chapter we shall look at some very useful, and quite simple, techniques for analyzing optical systems. Most of our attention will be directed towards optical systems that have axial symmetry. So, for example, we shall examine various properties of lenses, flat and spherical mirrors, plane-parallel slabs of transparent material, and polarizers. In our analysis we will be confronted with the problem of how the components of the system affect the passage of light. To describe the behavior of the system in detail we need to know how a wave passing through is modified in terms of intensity, phase shift, direction of propagation, and polarization state. Our discussion will be simplified by assuming that we are dealing with light that is monochromatic, or at least contains only a small range of wavelengths. This allows us to neglect the effects of *dispersion*, the variation of the refractive index with wavelength. Dispersion, defined quantitatively as $dn/d\lambda$, causes the properties of an optical system containing transmissive components such as lenses and optical fibers, to vary with wavelength.

14.2 The Propagation of Rays and Waves Through Isotropic Media

An isotropic optical material is one whose properties are independent of the direction of propagation or polarization state of an electromagnetic

wave passing through the material. Several simple relations hold for such a material:

- (i) The polarization \mathbf{P} induced by the electric field \mathbf{E} of the wave is parallel to \mathbf{E} , so

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \quad (14.1)$$

χ , the electric susceptibility, is a scalar.

- (ii) It follows that the electric displacement vector \mathbf{D} of the wave is parallel to \mathbf{E}

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (14.2)$$

which from Eq. (14.1) gives

$$\mathbf{D} = \epsilon_0(1 + \chi)\mathbf{E} \quad (14.3)$$

or

$$\mathbf{D} = \epsilon_0 \epsilon_r \mathbf{E} \quad (14.4)$$

ϵ_r is called the dielectric constant of the material. For an absorbing (or amplifying material) ϵ_r is complex, $\epsilon_r = \epsilon' - i\epsilon''$. The refractive index of the material is $n = \sqrt{\epsilon_r}$, or in a lossy material $n = \sqrt{\epsilon'}$. The dielectric constant, ϵ_r , and consequently the refractive index, n , depend on the frequency of the wave. This phenomenon is called dispersion and is the reason why a prism separates white light into its constituent colors.

- (iii) The Poynting vector of the wave, $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, which represent the magnitude and direction of energy flow, is parallel to \mathbf{k} , the wavevector. The wavevector \mathbf{k} is perpendicular to the surfaces of constant phase (phasefronts) and is by definition perpendicular to the vectors \mathbf{D} and \mathbf{B} . The time averaged magnitude of \mathbf{S} is called the intensity of the wave

$$I = \langle |\mathbf{S}| \rangle_{AV} \quad (14.5)$$

The local direction of \mathbf{S} is called the *ray* direction. Only for plane waves is the ray direction the same at all points on the phasefront. Many properties of an optical system can be best understood by considering what happens to a ray of light entering the system. How does its direction change as it encounters mirrors and lenses, and what happens to the intensity of the wave to which the ray belongs as it crosses interfaces between different components of the system?

Fig. 14.2.

14.3 Simple Reflection and Refraction Analysis

The phenomena of reflection and refraction are most easily understood in terms of plane electromagnetic waves - those sorts of waves where the direction of energy flow (the ray direction) has a unique direction. Other types of wave, such as spherical waves and Gaussian beams, are also important; however, the part of their wavefront which strikes an optical component can frequently be approximated as a plane wave, so plane wave considerations of reflection and refraction still hold true. The ray direction is the same as the direction of the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. In isotropic media (gases, liquids, glasses, and crystals of cubic symmetry) the ray direction and the direction of \mathbf{k} , the propagation vector, are the same. Except for plane waves the direction of \mathbf{k} varies from point to point on the phase front.

When light strikes a plane mirror, or the planar boundary between two media of different refractive index, the *angle of incidence* is always equal to the *angle of reflection*, as shown in Fig. (14.2). This is the fundamental *law of reflection*.

When a light ray crosses the boundary between two media of different refractive index, the *angle of refraction* θ_2 , shown in Fig. (14.2) is related to the angle of incidence θ_1 , by *Snell's Law*

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} \quad (14.6)$$

This result does not hold true in general unless both media are isotropic.

Since $\sin \theta_2$ cannot be greater than unity, if $n_2 < n_1$ the maximum angle of incidence for which there can be a refracted wave is called the *critical angle* θ_c where

$$\sin \theta_c = n_2/n_1 \quad (14.7)$$

Fig. 14.3a.

as illustrated in Fig. (14.3a).

If θ_1 exceeds θ_c the boundary acts as a very good mirror, as illustrated in Fig. (14.3b). This phenomenon is called *total internal reflection*. Several types of reflecting prisms operate this way. When total internal reflection occurs, there is no transmission of energy through the boundary. However, the fields of the wave do not abruptly go to zero at the boundary. There is said to be an *evanescent wave* on the other side of the boundary, the field amplitudes of which decay exponentially with distance. Because of the existence of this evanescent wave, other optical components should not be brought too close to a totally reflecting surface, or energy will be coupled to them via the evanescent wave and the efficiency of total internal reflection will be reduced. With extreme care this effect can be used to produce a variable-reflectivity, totally-internally-reflecting surface, as shown in Fig. (14.4). If one or both of the media in Fig. (14.2) are anisotropic, like calcite, crystalline quartz, ADP, KDP or tellurium, in general the incident wave will split into two components, one of which obeys Snell's Law (the *ordinary wave*) and one which does not (the *extraordinary wave*). This phenomenon is called *double refraction*, and is discussed in detail in Chapter 18.

If an optical system contains only planar interfaces, the path of a ray of light through the system can be easily calculated using only the law of reflection and Snell's Law. This simple approach neglects diffraction effects, which become significant unless the lateral dimensions (*apertures*) of the system are *all* much larger than the wavelength (say 10^4 times larger). The behavior of light rays in more complex systems containing non-planar components, but where diffraction effects are negligible, is better described with the aid of *paraxial ray* analysis¹. Transmitted and reflected intensities and polarization states cannot be determined

Fig. 14.3b.

Fig. 14.4.

Fig. 14.4a.

by the above methods and are most easily determined by the *method of impedances*.

Fig. 14.5a.

14.4 Paraxial Ray Analysis

A plane wave is characterized by a unique propagation direction given by the wave vector \mathbf{k} . All fields associated with the wave are, at a given time, equal at all points in infinite planes orthogonal to the propagation direction. In real optical systems such plane waves do not exist, as the finite size of the elements of the system restricts the lateral extent of the waves. Non-planar optical components will cause further deviations of the wave from planarity. Consequently, the wave acquires a ray direction which varies from point to point on the phase front. The behavior of the optical system must be characterized in terms of the deviations its elements cause to the bundle of rays which comprise the propagating, laterally-restricted wave. This is most easily done in terms of paraxial rays. In a cylindrically symmetric optical system, for example a coaxial system of spherical lenses or mirrors, *paraxial rays* are those rays whose directions of propagation occur at sufficiently small angles θ to the symmetry axis of the system that it is possible to replace $\sin \theta$ or $\tan \theta$ by θ - in other words paraxial rays obey the small angle approximation.

(a) Matrix Formulation

In an optical system whose symmetry axis is in the z direction, a paraxial ray in a given cross section ($z = \text{constant}$) is characterized by its distance r from the z axis and the angle r' it makes with that axis. If the values of these parameters at 2 planes of the system (an *input* and an *output* plane) are $r_1 r'_1$ and $r_2 r'_2$ respectively, as shown in Fig. (14.5a), in the paraxial ray approximation there is a linear relation between them of the form

$$\begin{aligned} r_2 &= Ar_1 + Br'_1 \\ r'_2 &= Cr_1 + Dr'_1 \end{aligned} \quad (14.8)$$

or, in matrix notation

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r_1 \\ r'_1 \end{pmatrix} \quad (14.9)$$

$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is called the ray transfer matrix, \mathbf{M} ; its determinant is usually unity, i.e. $AD - BC = 1$.

Optical systems made of isotropic material are generally reversible - a ray which travels from right to left with input parameters r_2, r'_2 will leave the system with parameters r_1, r'_1 - thus:

$$\begin{pmatrix} r_1 \\ r'_1 \end{pmatrix} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} \quad (14.10)$$

where the reverse ray transfer matrix $\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}$.

The ray transfer matrix allows the properties of an optical system to be described in general terms by the location of its *focal points* and *principal planes*, whose location is determined from the elements of the matrix. The significance of these features of the system can be illustrated with the aid of Fig. (14.5b). An input ray which passes through the *first focal point*, F_1 , (or would pass through this point if it did not first enter the system) emerges travelling parallel to the axis. The intersection point of the extended input and output rays, point H_1 in Fig. (14.5b), defines the location of the *first principal plane*. Conversely, an input ray travelling parallel to the axis will emerge at the output plane and pass through the second focal point, F_2 , (or appear to have come from this point). The intersection of the extension of these rays, point H_2 , defines the location of the *second principal plane*. Rays 1 and 2 in Fig. (14.5b) are called the *principal rays* of the system. The location of the principal planes allows the corresponding emergent ray paths to be determined as shown in Fig. (14.5b). The dashed lines in this figure, which permit geometric construction of the location of output rays 1 and 2, are called *virtual ray paths*. Both F_1 and F_2 lie on the axis of the system. The axis of the system intersects the principal planes at the *principal points*, P_1 and P_2 , in Fig. (14.5b).

The distance f_1 from the first principal plane to the first focal point is called the *first focal length*; f_2 is the *second focal length*.

In most practical situations, the refractive indices of the media to the

Fig. 14.5b.

left of the input plane (the *object space*) and to the right of the output plane (the *image space*) are equal. In this case we can derive simple relations between the focal lengths f_1 and f_2 and h_1 and h_2 the distances of the input and output planes from the principal planes, measured in the sense shown in Fig. (14.5b).

We can break up the system shown in Fig. (14.56) into 3 parts, the region from the input plane to the first principal plane, the region between the two principal planes, and the region from second principal plane to output plane.

If we write the transfer matrix from the left to the right principal planes as $\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$, then the overall transfer matrix is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & h_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} 1 & h_1 \\ 0 & 1 \end{pmatrix}$$

which gives

$$\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} 1 & -h_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 1 & -h_1 \\ 0 & 1 \end{pmatrix}$$

and therefore

$$\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} A - h_1 C & B - h_1 A - h_2(D - h_1 C) \\ C & D - h_1 C \end{pmatrix}$$

For the second principal rays in Fig. (4.7b), the distance of the ray from the axis, r , does not change, therefore

$$r = (A - h_2 C)r$$

so

$$h_2 = \frac{A - 1}{C}$$

Furthermore, for the first principal ray, whose input angle is r'

$$r = A'r + B'r'$$

so $B' = 0$.

Consequently, if the media to the left of the input plane and the right of the output plane are the same, using $\det(M) = 1$, gives

$$h_1 = \frac{D-1}{C}$$

For the second principal ray in Fig. (14.7b)

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} r_{in} \\ 0 \end{pmatrix}$$

so,

$$r'_2 = C' r_1$$

Furthermore, from Fig. (14.7b), it is easy to see that

$$-r'_2 = \frac{r_{in}}{f_2}$$

Combining Eqs. (14.11f) and (14.11g) gives

$$C = -\frac{1}{f_2}$$

By a similar procedure using the first principal ray it can be shown that

$$C' = -\frac{1}{f_1}$$

where C' is the (21) element of the inverse matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}$.

We have already seen from Eq. (14.) that $C' = C$, therefore,

$$f_1 = f_2 = f$$

Thus, if the elements of the transfer matrix are known, the location of the focal points and principal planes is determined. Graphical construction of ray paths through the system using the methods of *ray tracing* is then straight-forward (see section 14.2.2.(b)).

In using matrix methods for optical analysis, a consistent sign convention must be employed. In the present discussion:

A ray is assumed to travel in the positive z direction from left to right through the system. The distance from the first principal plane to an object is measured positive from right to left - in the negative z direction. The distance from the second principal plane to an image is measured positive from left to right - in the positive z direction. The lateral distance of the ray from the axis is positive in the upward direction, negative in the downward direction. The acute angle between the system axis direction and the ray, say r'_1 in Fig. (14.4a), is positive if a counterclockwise motion is necessary to go from the positive z direction to the ray direction. When the ray crosses a spherical interface

Fig. 14.6a.

the radius of curvature is positive if the interface is convex to the input ray. The use of ray transfer matrices in optical system analysis can be illustrated with some specific examples.

- (i) Uniform optical medium In a uniform optical medium of length d no change in ray angle occurs, as illustrated in Fig. (14.6a), so

$$\begin{aligned} r'_2 &= r'_1 \\ r_2 &= r_1 + dr'_1 \end{aligned}$$

Therefore,

$$\mathbf{M} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \quad (14.12)$$

The focal length of this system is infinite and it has no specific principal planes.

- (ii) Planar interface between two different media At the interface, as shown in Fig. (14.6b), $r_1 = r_2$ and from Snell's Law, using the approximation $\sin \theta = \theta$

$$r'_2 = \frac{n_1}{n_2} r'_1.$$

Therefore

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{pmatrix} \quad (14.13)$$

- (iii) A parallel-sided slab of refractive index n bounded on both sides with media of refractive index 1 (Fig. 14.6c)

In this case,

$$\mathbf{M} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix} \quad (14.14)$$

Fig. 14.6b.

Fig. 14.6c.

Fig. 14.6d.

The principal planes of this system are the boundary faces of the optically dense slab.

- (iv) Thick Lens The ray transfer matrix of the thick lens shown in Fig. (14.7a) is the product of the three transfer matrices:

Fig. 14.6e.

Fig. 14.6f.

Fig. 14.7a.

$$\mathbf{M}_3 \cdot \mathbf{M}_2 \cdot \mathbf{M}_1 = \begin{pmatrix} \text{matrix for} \\ \text{second spherical} \\ \text{interface} \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{medium of} \\ \text{length } d \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{first spherical} \\ \text{of interface} \end{pmatrix}$$

Note the order of these three matrices; \mathbf{M}_1 comes on the right because it operates first on the column vector which describes the input ray.

At the first spherical surface

$$n'(r'_1 + \phi_1) = nr = n(r''_2 + \phi_1)$$

which, since $\phi_1 = r_1/R_1$, can be written as

$$r''_2 = \frac{n'r'_1}{n} + \frac{(n' - n)r_1}{nR_1}$$

The transfer matrix at the first spherical surface is

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 0 \\ \frac{(n'-n)}{nR_1} & \frac{n'}{n} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{D_1}{n} & \frac{n'}{n} \end{pmatrix} \quad (14.15)$$

where $D_1 = (n - n')/R_1$ is called the power of the surface. If R_1 is measured in meters the units of D_1 are *diopters*.

In the paraxial approximation all the rays passing through the lens travel the same distance d in the lens. Thus,

$$\mathbf{M}_2 = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}. \quad (14.16)$$

The ray transfer matrix at the second interface is

$$\mathbf{M}_3 = \begin{pmatrix} 1 & 0 \\ \frac{n-n'}{n'R_2} & \frac{n}{n'} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{D_2}{n'} & \frac{n}{n'} \end{pmatrix} \quad (14.17)$$

which is identical in form to \mathbf{M}_1 . Note that in this case both r'_2 and R_2 are negative. The overall transfer matrix of the thick lens is

$$\mathbf{M} = \mathbf{M}_3\mathbf{M}_2\mathbf{M}_1 = \begin{pmatrix} 1 - \frac{dD_1}{n} & \frac{dn'}{n} \\ \frac{dD_1D_2}{nn'} - \frac{D_1}{n'} - \frac{D_2}{n'} & 1 - \frac{dD_2}{n} \end{pmatrix} \quad (14.18)$$

If Eqs. (14.18) and (14.11) are compared, it is clear that the locations of the principal planes of the thick lens are

$$h_1 = \frac{d}{\frac{n}{n'} \left(1 + \frac{D_1}{D_2} - \frac{dD_1}{n} \right)} \quad (14.19)$$

$$h_2 = \frac{d}{\frac{n}{n'} \left(1 + \frac{D_2}{D_1} - \frac{dD_2}{n} \right)} \quad (14.20)$$

A numerical example will best illustrate the location of the principal planes for a biconvex thick lens.

Suppose:

$$n' = 1(\text{air}), n = 1.5(\text{glass})$$

$$R_1 = -R_2 = 50\text{mm}$$

$$d = 10\text{mm}$$

In this case $D_1 = D_2 = 0.01\text{mm}^{-1}$ and from, Eqs. (14.19) and (14.20),

Fig. 14.7b.

$h_1 = h_2 = 3.448\text{mm}$. These principal planes are symmetrically placed inside the lens. Fig. (14.7b) shows how the principal planes can be used to trace the principal ray paths through a thick lens.

From Eqs. (14.18) and (14.11)

$$r'_2 = \frac{-r_1}{f} + \left(1 - \frac{dD_2}{n}\right) r'_1 \quad (14.21)$$

where the focal length is

$$f = \left(\frac{D_1 + D_2}{n'} - \frac{dD_1D_2}{nn'}\right)^{-1} \quad (14.22)$$

If l_o is the distance from the object O to the first principal plane, and l_i the distance from the second principal plane to the image I in Fig. (14.7b), then from the similar triangles $OO'F_1$, P_1AF_1 and P_2DF_2 , $II'F_2$

$$\begin{aligned} \frac{OO'}{II'} &= \frac{OF_1}{F_1P_1} = \frac{l_o - f_1}{f_1} \\ \frac{OO'}{II'} &= \frac{P_2F_2}{F_2I} = \frac{f_2}{l_i - f_2}. \end{aligned} \quad (14.23)$$

If the media on both sides of the lens are the same, then $f_1 = f_2 = f$ and it immediately follows that

$$\frac{1}{l_o} + \frac{1}{l_i} = \frac{1}{f}. \quad (14.24)$$

This is the fundamental imaging equation.

(v) Thin Lens. If a lens is sufficiently thin that to a good approximation $d = 0$ the transfer matrix is

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{-(D_1+D_2)}{n'} & 1 \end{pmatrix} \quad (14.25)$$

As shown in Fig. (14.6d) the principal planes of such a thin lens are at

the lens. The focal length of the thin lens is f , where

$$1/f = \frac{D_1 + D_2}{n'} = \left(\frac{n}{n'} - 1\right) \left(\frac{1}{R_1} - \frac{1}{R_2}\right) \quad (14.26)$$

so the transfer matrix can be written very simply as

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \quad (14.27)$$

The focal length of the lens depends on the refractive index of the lens material and the refractive index of the medium with which it is immersed. In air

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2}\right). \quad (14.28)$$

For a biconvex lens R_2 is negative and

$$\frac{1}{f} = (n - 1) \left(\frac{1}{|R_1|} + \frac{1}{|R_2|}\right). \quad (14.29)$$

For a biconcave lens

$$\frac{1}{f} = -(n - 1) \left(\frac{1}{|R_1|} + \frac{1}{|R_2|}\right). \quad (14.30)$$

The focal length of any diverging lens is negative. For a thin lens object and image distances are measured to a common point. It is common practice to rename the distances ℓ_o and ℓ_i in this case so that the imaging Eq. (14.24) reduces to its familiar form

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f}, \quad (14.31)$$

u is called the *object distance* and v the *image distance*.

(vi) A length of uniform medium plus a thin lens. (Fig. (14.6e)). This is a combination of systems (i) + (v); its overall transfer matrix is found from Eqs. (14.12) and (14.27) as

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d \\ -1/f & 1 - d/f \end{pmatrix}. \quad (14.32)$$

(vii) Two thin lenses As a final example of the use of ray transfer matrices consider the combination of two thin lenses shown in (Fig. (14.6f).) The transfer matrix of this combination is

$$\mathbf{M} = \begin{pmatrix} \text{matrix for} \\ \text{second lens, } f_2 \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{uniform medium, } d_2 \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{first lens, } f_1 \end{pmatrix} \\ \begin{pmatrix} \text{matrix for} \\ \text{uniform medium, } d_1 \end{pmatrix}$$

which can be shown to be

$$\mathbf{M} = \begin{pmatrix} 1 - \frac{d_2}{f_1} & d_1 + d_2 - \frac{d_1 d_2}{f_1} \\ -\frac{1}{f_1} - \frac{1}{f_2} + \frac{d_2}{f_1 f_2} & 1 - \frac{d_1}{f_1} - \frac{d_1}{f_2} - \frac{d_2}{f_2} + \frac{d_1 d_2}{f_1 f_2} \end{pmatrix} \quad (14.33)$$

The focal length of the combination is

$$f = \frac{f_1 f_2}{d_2 - (f_1 + f_2)}. \quad (14.34)$$

The optical system consisting of two thin lenses is the archetypal system used in analyses of the stability of lens waveguides and optical resonators^{10,12}.

(viii) Spherical mirrors. The object and image distances from a spherical mirror also obey Eq. (14.31) where the focal length of the mirror is $R/2$; f is positive for a concave mirror, negative for a convex mirror. Positive object and image distances for a mirror are measured positive in the normal direction away from its surface. If a negative image distance results from Eq. (14.31) this implies a *virtual image* (behind the mirror).

(b) Ray Tracing

Practical implementation of paraxial ray analysis in optical system design can be very conveniently carried out graphically by *ray tracing*. In ray tracing a few simple rules allow geometrical construction of the principal ray paths from an object point, although these constructions do not take into account the non-ideal behavior, or *aberrations*, of real lenses. The first principal ray from a point on the object passes through (or its projection passes through) the first focal point. From the point where this ray, or its projection, intersects the first principal plane the output ray is drawn parallel to the axis. The actual ray path between input and output planes can be found in simple cases, for example, the path XX' in the thick lens shown in Fig. (14.7b). The second principal ray is directed parallel to the axis; from the intersection of this ray, or its projection, with the second principal plane the output ray passes through (or appears to have come from) the second focal point. The actual ray path between input and output planes can be found in simple cases, for example, the path YY' in the thick lens shown in Fig. (14.7b). The intersection of the two principal rays in the image space produce the image point which corresponds to the same point on the object. If only the back-projections of the output principal rays appear to intersect, this intersection point lies on a *virtual image*.

In the majority of applications of ray tracing a quick analysis of the system is desired. In this case, if all lenses in the system are treated as thin lenses, the position of the principal planes need not be calculated beforehand and ray tracing becomes particularly easy. For a thin lens a third principal ray is useful for determining image location. The ray

Fig. 14.8.

Fig. 14.8a.

from a point on the object which passes through the center of the lens is not deviated by the lens. The use of ray tracing to determine the size and position of the real image produced by a convex lens and the virtual image produced by a concave lens are shown in Fig. (14.8). Fig. (4.8a) shows a converging lens where the input principal ray parallel to the axis actually passes through the focal point. Fig. (14.8b) shows a diverging lens where this ray emerges from the lens appearing to have come from the focal point. More complex systems of lenses can be analyzed the same way. Once the image location has been determined by the use of the principal rays, the path of any group of rays, a ray *pencil*, can be found; see, for example, the cross-hatched ray pencils shown in Fig. (14.8).

The use of ray tracing rules to analyze spherical mirror systems is similar to those described above, although the ray striking the center of the mirror in this case reflects so that the angle of incidence equals the angle of reflection. Figs. (14.8c) and (14.8d) illustrate the use of ray tracing to locate the images produced by concave and convex mirrors.

Fig. 14.8b.

In the case of a convex mirror the ray tracing construction is modified in the same way as it was for a de-focusing lens. A ray parallel to the axis that would pass through the focal point *if it did not first encounter the mirror surface* reflects parallel to the axis. A ray parallel to the axis reflects at the mirror surface so as to appear in projection to have come from the focal point.

(c) Imaging and Magnification.

In Fig. (14.8a) the ratio of the height of the image to the height of the object is called the *magnification*, m . In the case of a thin lens

$$m = \frac{b}{a} = \frac{v}{u} \quad (14.35)$$

For a more general system the magnification can be obtained from the ray transfer matrix equation

$$\begin{pmatrix} b \\ b' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a \\ a' \end{pmatrix} = M \begin{pmatrix} a \\ a' \end{pmatrix} \quad (14.36)$$

where a' is the angle of a ray through a point on the object and b' through the corresponding point on the image. The matrix M in this case includes the entire system from object O to image I . So, in Fig. (14.8a)

$$\mathbf{M} = \begin{pmatrix} \text{matrix for} \\ \text{uniform medium} \\ \text{of length } v \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{lens} \end{pmatrix} \begin{pmatrix} \text{matrix for} \\ \text{uniform medium} \\ \text{of length } u \end{pmatrix} \quad (14.37)$$

for imaging, b must be independent of a' , so $B = 0$.

The magnification is

$$m = \frac{b}{a} = A \quad (14.38)$$

so the ray transfer matrix of the imaging system can be written

$$\mathbf{M} = \begin{pmatrix} m & 0 \\ -\frac{1}{f} & \frac{1}{m} \end{pmatrix}, \quad (14.39)$$

where it should be noted the result, $\det(M) = 1$, has been used.

The *angular magnification* of the system is defined as

$$m' = \left(\frac{b'}{a'}\right)_{a=0}$$

which gives $m' = 1/m$.

Note that $mm' = 1$, a useful general result.

14.4.1 The Use of Impedances in Optics

The method of impedances is the easiest way to calculate the fraction of incident intensity transmitted and reflected in an optical system. It is also a good way to follow the changes in polarization state which result when light passes through an optical system.

The impedance of a plane wave travelling in a medium of relative permeability μ_r and dielectric constant ϵ_r is

$$Z = \sqrt{\frac{\mu_r \mu_0}{\epsilon_r \epsilon_0}} = Z_0 \sqrt{\frac{\mu_r}{\epsilon_r}} \quad (14.40)$$

If $\mu_r = 1$, as is usually the case for optical media, the impedance can be written

$$Z = Z_0/n. \quad (14.41)$$

This impedance relates the transverse \mathbf{E} and \mathbf{H} fields of the wave

$$Z = \frac{E_{tr}}{H_{tr}}. \quad (14.42)$$

When a plane wave crosses a planar boundary between two different media the components of both \mathbf{E} and \mathbf{H} parallel to the boundary have to be continuous across that boundary. Fig. (14.9a) illustrates a plane wave polarized in the plane of incidence striking a planar boundary between two media of refractive index n_1 and n_2 , respectively. In terms of the magnitudes of the vectors involved

$$E_i \cos \theta_1 + E_r \cos \theta_1 = E_t \cos \theta_2 \quad (14.43)$$

$$H_i - H_r = H_t. \quad (14.44)$$

Eq. (14.44) can be written

$$\frac{E_i}{Z_1} - \frac{E_r}{Z_1} = \frac{E_t}{Z_2}. \quad (14.45)$$

Fig. 14.9a.

Fig. 14.9b.

It is easy to eliminate E_t between Eqs. (14.43) and (14.45) to give

$$\rho = \frac{E_r}{E_i} = \frac{Z_2 \cos \theta_2 - Z_1 \cos \theta_1}{Z_2 \cos \theta_2 + Z_1 \cos \theta_1}, \quad (14.46)$$

where ρ is the *reflection coefficient* of the surface. The fraction of the incident energy reflected from the surface is called the reflectance, $R = \rho^2$. Similarly, the *transmission coefficient* of the boundary is

$$\tau = \frac{E_t}{E_i} = \frac{2Z_2 \cos \theta_2}{Z_2 \cos \theta_2 + Z_1 \cos \theta_1}. \quad (14.47)$$

By a similar treatment applied to the geometry shown in Fig. (14.9b) it can be shown that for a plane wave polarized perpendicular to the plane of incidence

$$\rho = \frac{Z_2 \sec \theta_2 - Z_1 \sec \theta_1}{Z_2 \sec \theta_2 + Z_1 \sec \theta_1} \quad (14.48)$$

$$\tau = \frac{2Z_2 \sec \theta_2}{Z_2 \sec \theta_2 + Z_1 \sec \theta_1}. \quad (14.49)$$

If the effective impedance for a plane wave polarized in the plane of

incidence (*P polarization*) and incident on a boundary at angle θ is defined as

$$Z' = Z \cos \theta, \quad (14.50)$$

and for a wave polarized perpendicular to the plane of incidence (*S polarization*) as

$$Z' = Z \sec \theta, \quad (14.51)$$

then a universal pair of formulae for ρ and τ results:

$$\rho = \frac{Z'_2 - Z'_1}{Z'_1 + Z'_2}, \quad (14.52)$$

$$\tau = \frac{2Z'_2}{Z'_1 + Z'_2}. \quad (14.53)$$

It will be apparent from an inspection of Fig. (14.9) that Z' is just the ratio of the electric field component parallel to the boundary and the magnetic field component parallel to the boundary. For reflection from an ideal mirror $Z'_2 = 0$. Note that from Eqs. (14.52) and (14.53)

$$\tau = 1 + \rho \quad (14.53a)$$

In normal incidence Eqs. (14.46) and (14.48) become identical and can be written as

$$\rho = \frac{Z_2 - Z_1}{Z_2 + Z_1} = \frac{n_1 - n_2}{n_1 + n_2} \quad (14.54)$$

Note that there is a change of phase of π in the reflected field relative to the incident field when $n_1 < n_2$.

Since intensity \propto (electric field)², the fraction of the incident energy which is reflected is,

$$R = \rho^2 = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \quad (14.55)$$

R increases with the index mismatch between the two media, as shown in Fig. (14.10).

It is important to note that although the energy reflectance at a boundary is $|\rho|^2$ the energy transmittance is *not* $|\tau|^2$. This is because the impedances are different on the two sides of the boundary. For example, if a wave is incident from medium 1 to medium 2 the intensity is $I_1 = |E_i|^2/2Z_1$, and the transmitted intensity is $I_2 = |E_t|^2/2Z_2$. The transmittance of the boundary is

$$T = \frac{I_2}{I_1} = \tau^2 \frac{Z_1}{Z_2} \quad (14.56a)$$

If the wave were travelling in the reverse direction across the boundary

Fig. 14.10.

then its transmission coefficient, as in Eq. (14.53) would be

$$\tau' = \frac{2Z'_1}{Z'_1 + Z'_2} \quad (14.56b)$$

and

$$\tau\tau' = \frac{4Z'_1Z'_2}{(Z'_1 + Z'_2)^2} = \frac{\tau^2 Z_1}{Z_2} = T \quad (14.56c)$$

If there is no absorption of energy at the boundary the fraction of energy transmitted, called the *transmittance*, is

$$T = 1 - R = \frac{4n_1n_2}{(n_1 + n_2)^2} \quad (14.56)$$

- (a) Reflectance for Waves Incident on an Interface at Oblique Angles
If the wave is not incident normally, it must be decomposed into two linearly polarized components, one polarized in the plane of incidence, and one polarized perpendicular to the plane of incidence.

For example, consider a plane polarized wave incident on an air/glass ($n = 1.5$) interface at an angle of incidence of 30° with a polarization state exactly intermediate between the S and P polarizations.

The angle of refraction at the boundary is found from Snell's Law

$$\sin \theta_2 = \sin 30^\circ / 1.5$$

so $\theta_2 = 19.47^\circ$.

The effective impedance of the P component in the air is

$$Z'_{P1} = 376.7 \cos \theta_1 = 326.23 \text{ ohm}$$

and in the glass

$$Z'_{P2} = \frac{376.7}{1.5} \cos \theta_2 = 236.7 \text{ ohm.}$$

Thus, from Eq. (14.52), the reflection coefficient for the P component is

$$\rho_P = \frac{236.77 - 326.33}{236.77 + 326.23} = -0.159$$

The fraction of the intensity associated with the P component which is reflected is $\rho_P^2 = 0.0253$.

For the S component of the input wave

$$\begin{aligned} Z'_{S1} &= 376.7 / \cos \theta_1 = 434.98 \text{ ohm} \\ Z'_{S2} &= 376.7 / (1.5 \cos \theta_2) = 266.37 \text{ ohm} \\ \rho_S &= \frac{266.37 - 434.98}{266.37 + 434.98} = -0.240 \end{aligned}$$

The fraction of the intensity associated with the S polarization component which is reflected is $\rho_S^2 = 0.0578$. Since the input wave contains equal amounts of S and P polarization, the overall reflectance in this case is

$$R = \langle \rho^2 \rangle_{\text{avg}} = 0.0416 \simeq 4\%$$

Note that the reflected wave now contains more S polarization than P, so the polarization state of the reflected wave has been changed.

- (b) Brewster's Angle Returning to Eq. (14.46) it might be asked whether the reflectance is ever zero. It is clear that ρ will be zero if

$$n_1 \cos \theta_2 = n_2 \cos \theta_1 \quad (14.57)$$

which from Snell's Law, Eq. (14.7), gives

$$\cos \theta_1 = \frac{n_1}{n_2} \sqrt{1 - \left(\frac{n_1^2}{n_2^2}\right) \sin^2 \theta_1}, \quad (14.58)$$

giving the solution

$$\theta_1 = \theta_B = \arcsin \sqrt{\frac{n_2^2}{n_1^2 + n_2^2}} = \arctan \frac{n_2}{n_1} \quad (14.59)$$

The angle θ_B is called *Brewster's angle*, a wave polarized in the plane of incidence and incident on a boundary at this angle is totally transmitted. This fact is put to good use in the design of low-reflection-loss windows in laser systems, as we have seen previously. If Eq. (14.48) is inspected carefully it will be seen that there is no angle of incidence which yields zero reflection for a wave polarized perpendicular to the plane of incidence.

- (c) Transformation of Impedance through Multilayer Optical Systems. The impedance concept allows the reflection and transmission characteristics of multilayer optical systems to be evaluated very sim-

Fig. 14.11a.

ply. If the incident light is incoherent then the overall transmission of a multilayer structure is just the product of the transmittances of its various interfaces. For example, an air/glass interface transmits about 96% of the light in normal incidence. The transmittance of a parallel-sided slab is $0.96 \times 0.96 \equiv 92\%$. This simple result ignores the possibility of interference effects between reflected and transmitted waves at the two faces of the slab. If the faces of the slab are very flat and parallel, and the light is coherent, such effects cannot be ignored. In this case, the method of transformed impedances is useful.

Consider the three-layer structure shown in Fig. (14.11a). The path of a ray of light through the structure is shown. The angles $\theta_1, \theta_2, \theta_3$ can be calculated from Snell's law. As an example consider a wave polarized in the plane of incidence. The effective impedances of media 1, 2 and 3 are:

$$\begin{aligned} Z'_1 &= Z_1 \cos \theta_1 = Z_0 \cos \theta_1 / n_1 \\ Z'_2 &= Z_2 \cos \theta_2 = Z_0 \cos \theta_2 / n_2 \\ Z'_3 &= Z_3 \cos \theta_3 = Z_0 \cos \theta_3 / n_3 \end{aligned} \quad (14.60)$$

It can be shown that the reflection and transmission coefficients of the structure are exactly the same as the equivalent structure in Fig. (14.11b) for normal incidence, where the effective thickness of layer 2 is now

$$d' = d \cos \theta_2. \quad (14.61)$$

the reflection and transmission coefficients of the structure can be cal-

Fig. 14.11b.

culated from its equivalent structure using the transformed impedance concept¹¹.

The transformed impedance of medium 3 at the boundary between media 1 and 2 is

$$Z_3'' = Z_2' \left(\frac{Z_3' \cos k_2 d' + i Z_2' \sin k_2 d'}{Z_2' \cos k_2 d' + i Z_3' \sin k_2 d'} \right) \quad (14.62)$$

where $k_2 = 2\pi/\lambda_2 = 2\pi n_2/\lambda_0$.

The reflection coefficient of the whole structure is now just

$$\rho = \frac{Z_3'' - Z_1'}{Z_3'' + Z_1'} \quad (14.63)$$

and its transmission coefficient

$$\tau = \frac{2Z_3''}{Z_3'' + Z_1'} \quad (14.64)$$

In a structure with more layers, the transformed impedance formula, Eq. (14.62) can be used sequentially, starting at the last optical surface and working back to the first.

Polarization Changes. A change in polarization state can occur in two principal ways: (a) the wave can reflect off an interface for which the reflection coefficient is different for light polarized in the plane of incidence (a *P-wave*), or perpendicular to the plane of incidence (a *S-wave*), and (b) the wave can be transmitted through an anisotropic, or optically active medium that alters the polarization state of the wave. These effects are considered in further detail in Chapter 18 and xxx. Optically active media rotate the plane of polarization of a linearly polarized wave passing through them. This can occur in solids or liquids in which the constituent particles are *chiral* in character, that is they have distinct left or right-handed symmetry. A good example would be an aqueous

sugar solution. A solution of the sugar dextrose, for example, rotates the plane of polarization of a linearly polarized wave in a clockwise direction viewing in the direction of propagation while its *enantiomorph*^{*}, laevulose, rotates the plane of polarization in a counter-clockwise direction. One way to account for the phenomenon of optical activity is for left- and right-hand circularly polarized light to have different phase velocities in the medium. Optical activity can also be induced by the action of an external field. For example, in the liquid nitrobenzene, application of a transverse electric field causes the liquid to become optically active. This is the Kerr effect and is used as we have seen previously to make shutters for Q-switching. Many materials become optically active if they are placed in an axial magnetic field. This is the Faraday effect. The actual amount of rotation is described by the *Verdet* constant α for the material. For a slab of material of length ℓ and an applied axial flux density \mathbf{B}

$$\theta_{rot} = \alpha \ell | \mathbf{B} |$$

If α is positive the direction of rotation is clockwise if the light propagates in the same direction as \mathbf{B} , it is counterclockwise if the direction of propagation is anti-parallel to \mathbf{B} . Thus, a to-and-fro pass of light through the field doubles the angle of rotation.** The Faraday effect is used to good effect in optical isolators; the operation of which is shown schematically in Fig. (14.2). A plane polarized wave is rotated by 45° and cannot then pass the entrance linear polarizer. Thus, the source of radiation is protected from undesirable feedback from optical components further down the chain. The use of isolators is extremely important in high power oscillator/amplifier(s) systems. Without their use a reflected wave of initially weak intensity could be amplified on its return passage through the system with disastrous consequences for the oscillator.

14.5 Appendix 14

14.5.1 Optical Terminology

Light is electromagnetic radiation with a wavelength between $0.1nm$ and $100\mu m$. Of course, this selection of wavelength range is somewhat arbitrary; short wavelength vacuum ultraviolet light between $0.1nm$ and

* The molecule with the same atomic constituents, but mirror image structure.

** This is in contrast to natural optical activity where a to-and-fro pass leads to net rotation.

10nm (so called because these wavelengths - and a range above up to about 200nm - are absorbed by air and most gases) might as easily be called soft X-radiation. At the other end of the wavelength scale, between 100 μ m and 1000 μ m, the sub-millimeter wave region of the spectrum, lies a spectral region where conventional optical methods become difficult, as do the extension of microwave techniques which are easily applied to the centimeter and millimeter wave regions. The use of optical techniques in this region, and even in the millimeter region, is often called quasi-optics.

Table 14.1 summarises the important parameters which are used to characterize light and the media through which it passes. A few comments on the table are appropriate. Although the velocity of light in a medium depends both on the relative magnetic permeability μ_r and dielectric constant ϵ_r of the medium, for all practical optical materials $\mu_r = 1$ so the refractive index and dielectric constant are related by

$$n = \sqrt{\epsilon_r}. \quad (\text{A14.1})$$

When light propagates in an anisotropic medium, such as a crystal of lower than cubic symmetry, n and ϵ_r will, in general, depend on the direction of propagation of the wave and its polarization state. The velocity of light *in vacuo* is currently the most precisely known of all the physical constants - its value is known³ within 4 $cm s^{-1}$. A redefinition of the meter has been proposed⁴ based on the cesium atomic clock frequency standard⁵ and a velocity of light of exactly $2.99792458 \times 10^8 m s^{-1}$. The velocity, wavelength and wavenumber of light travelling in the air have slightly different values than they have *in vacuo*. Tables are available which tabulate corresponding values of $\bar{\nu}$ *in vacuo* and in standard air.

A plane electromagnetic wave travelling in the z direction can, in general, be decomposed into two independent, linearly polarized components *. The electric and magnetic field pairs associated with each of these components are themselves mutually orthogonal and transverse to the direction of propagation. They can be written as (E_x, H_y) and (E_y, H_x) . The ratio of the mutually orthogonal \mathbf{E} and \mathbf{H} components is called the *impedance* Z of the medium.

$$\frac{E_x}{H_y} = \frac{-E_y}{H_x} = Z = \sqrt{\frac{\mu_r \mu_0}{\epsilon_r \epsilon_0}} \quad (\text{A14.2})$$

* The wave could also be decomposed into left- and right-hand circularly polarized components.

Table 14.1 *Fundamental Parameters of Electromagnetic Radiation and Optical Media.*

Parameter	Symbol	Value	Units
Velocity of light in free space <i>in vacuo</i>	$c_0 = (\mu_0 \epsilon_0)^{-1/2}$	2.99792458 $\times 10^8^*$	ms^{-1}
permeability of free space	μ_0	$4\pi \times 10^{-7}$	henry m^{-1}
permittivity of free space	ϵ_0	8.85416×10^{-12}	farad m^{-1}
velocity of light in a medium	$c = (\mu_r \mu_0 \epsilon_r \epsilon_0)^{-1/2}$ $= c_0/n$		ms^{-1}
refractive index	$n = (\mu_r \epsilon_r)^{1/2}$		dimensionless
relative permeability of a medium	μ_r	usually 1	dimensionless
dielectric constant of a medium	ϵ_r		dimensionless
frequency	$\nu = c/\lambda$		Hz $10^9 Hz = 1GHz$ (gigahertz) $10^{12} Hz = 1Thz$ (terahertz)
wavelength <i>in vacuo</i>	$\lambda_0 = c_0/\nu$		m
wavelength in a medium	$\lambda = c/\nu = \lambda_0/n$		m $10^{-3} m = 1mm$ (millimeter) $10^{-6} m = 1\mu m$ (micrometer) $10^{-9} m = 1nm$ (nanometer) $10^{-12} m = 1pm$ (picometer) 1\AA (Angstrom) = $10^{-10} m$ 1μ (micron) = $1\mu m$ $1m\mu$ (millimicron) = $1nm$
wavenumber	$\bar{\nu} = 1/\lambda$		cm^{-1} (Kayser)
wave vector	$\mathbf{k}, \mathbf{k} = 2\pi/\lambda$		m^{-1}
photon energy	$E = h\nu$		J 1 eV (electron volt) = $1.60202 \times 10^{-19} J$
electric field of wave	\mathbf{E}		volt m^{-1}
magnetic field of wave	\mathbf{H}		amp m^{-1}
Poynting vector intensity	$\mathbf{P} = \mathbf{E} \times \mathbf{H}$ $I = \langle \mathbf{P} \rangle_{AV} =$ $ \mathbf{E} ^2/2Z$		watt m^{-2} watt m^{-2}
impedance of medium	$Z = E_x/H_y = -$ $E_y/H_x = (\frac{\mu_r \mu_0}{\epsilon_r \epsilon_0})^{1/2}$		ohm
impedance of free space	$Z_0 = (\mu_0/\epsilon_0)^{1/2}$	376.7	ohm

*See text.

The negative sign in Eq. (2) arises because (y, x, z) is not a right-handed coordinate system.

The Poynting vector, \mathbf{S} , where

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \quad (A14.3)$$

is a vector which points in the direction of energy propagation of the wave. The average magnitude of the Poynting vector is called the intensity I .

$$I = \langle |\mathbf{S}| \rangle_{AV} = \frac{|\mathbf{E}|^2}{2Z} \quad (A14.4)$$

The factor of 2 comes from time-averaging the square of the sinusoidally varying electric field.

The wavevector \mathbf{k} points in the direction perpendicular to the phase front of the wave (the surface of constant phase). In an isotropic medium \mathbf{k} and \mathbf{S} are always parallel.

The photon flux corresponding to an electromagnetic wave of average intensity I and frequency ν is

$$N = I/h\nu = I\lambda/hc, \quad (A14.5)$$

where h is Planck's constant, $6.626 \times 10^{-34} Js$.

For a wave of intensity 1 watt m^{-2} and wavelength $1\mu m$ in *vacuo* $N = 5.04 \times 10^{18}$ photons $s^{-1}m^{-2}$. Photon energy is sometimes measured in electron volts (eV): $1eV = 1.60202 \times 10^{-19} J$. A photon of wavelength $1\mu m$ has an energy of $1.24eV$. It is often important, particularly in the infrared, to know the correspondence between photon and thermal energies. The characteristic thermal energy at absolute temperature is kT . At $300^\circ K$, $kT = 4.14 \times 10^{-21} J = 208.6 cm^{-1} = 0.026eV$.