14 Optics

14.1 INTRODUCTION

Maxwell's theory demonstrated that light is an electromagnetic phenomenon, as was explained in Chapter 3. We have consequently used optical examples extensively to this point. But there are enough special considerations in the generation of coherent light by means of lasers, and in the use of this radiation, that a chapter for specific discussion of these is justified. We consequently concentrate on optical waveguides, optical resonators, and some related optical components in this chapter. In referring to "optical frequencies," we generally mean the visible range, the near infrared, and the near ultraviolet (wavelengths, say from around 10 μm to 100 nm), but the key point is the relation of size and curvature of wavefronts to wavelength. Thus most of the concepts also apply to the far infrared (wavelengths $\sim$ 100 μm) and many of the analyses to millimeter waves and microwave radio frequencies when relation of size to wavelength is appropriate. The term *quasi-optical* is often applied to such situations.

For a number of practical applications, the propagation of light may be described by the behavior of the *rays*, which are the normals to the wavefronts. These are straight lines in a homogeneous medium, change direction according to Snell's law at the boundary between dielectric media, and are in general curved for inhomogeneous media. The basis for these rules and some important applications will be presented in the first part of this chapter.

The guiding of optical waves is generally accomplished by dielectric waveguides of the type introduced in Sec. 9.2. Hollow-pipe guides with metal boundaries are not only difficult to fabricate in the small sizes one would find at these wavelengths, but the metal boundaries would produce much greater losses at the optical frequencies than can be obtained with good dielectric guides. The principles of dielectric guiding have been presented in Chapter 9, but there are a number of special considerations worth developing in more detail than was done in that introduction. In particular, dispersion of signals in such guides is of importance in telecommunication uses. Also, a class of inhomogeneous dielectric guides, known as *graded index guides*, has unique and useful properties and leads to waves with gaussian form in the transverse plane. Similar but diverging *gaussian beams* are found in space (or other homogeneous media) and their transformation by lenses and other optical components is especially important. The spreading of such beams is a special case of diffraction, studied in Chapter 12.

Optical resonators are likewise not different in principle from other resonators we have studied in Chapter 10. But as with the guides, the questions of size in relation to wavelength and the properties of materials at optical frequencies produce some practical differences. In particular, there is much use of resonant systems which are open to space in the transverse direction, with the waves reflected longitudinally between plane or spherical mirrors placed normally to an axis. Gaussian beams again play a role in such resonators.

Several materials properties important to optics, especially anisotropic and nonlinear properties of materials, have been covered in Chapter 13. In this chapter we consider an important aspect of nonlinearity—that leading to *solitons* in which dispersive and nonlinear effects tend to cancel. In soliton propagation short pulses may propagate large distances without much change in shape. Finally, we give an introduction to the information processing applications that arise from diffraction theory and the Fourier transforming properties of lenses. Additional details on these subjects are covered in many excellent texts.[1–4] Quantum effects, such as the laser principle and interactions between light and semiconductors, are covered in books on quantum electronics.[5]

Ray or Geometrical Optics

14.2 Geometrical Optics Through Applications of Laws of Reflection and Refraction

We have already applied the results of the plane-wave analysis to many optical problems. The results of that analysis (Chapter 6) are strictly applicable only to plane waves which are uniform over an infinite wavefront and which fall upon infinite plane boundaries between media. Nevertheless, one would expect the results to be useful whenever the wavefront extends and is uniform over many wavelengths, and when the boundaries are large in comparison with wavelength. Both boundaries and wavefront may be nonplanar so long as radii of curvature are also large in comparison with wavelength. In such cases one obtains a great deal of information about the waves by tracing the *rays* that represent the normals to the wavefronts, using locally the law of reflection and Snell's law of refraction. These were developed in Chapter 6 for uniform plane waves

[1] *A. Yariv,* Optical Electronics, *4th ed., Saunders, Philadelphia, 1991.*
[2] *H. A. Haus,* Waves and Fields in Optoelectronics, *Prentice Hall, Englewood Cliffs, NJ, 1984.*
[3] *B. E. A. Saleh and M. C. Teich,* Fundamentals of Photonics, *Wiley, New York, 1991.*
[4] *A. E. Siegman,* Lasers, *University Science Books, Mill Valley, CA, 1986.*
[5] *For example, A. Yariv,* Quantum Electronics, *3rd ed., Wiley, New York, 1989.*

by considering phase requirements on continuity conditions at a boundary. These laws were actually first developed by observing the behavior of rays of light falling upon various materials before the wave nature of light was clearly established, and formed the basis for an extensive art and science of optics well before Maxwell. It is called *geometrical optics* or *ray optics*. We shall show the formal development in the next section; here we use only the physical justification discussed above and give some useful examples. In this section we consider only homogeneous media, except for discontinuities at boundaries.

Example 14.2a
SPHERICAL MIRROR

Consider first a reflecting surface of spherical form with rays (normals to the wavefronts of the plane waves) parallel to the axis and falling on the mirror. As shown in Fig. 14.2*a*, a ray at radius r from the axis makes angle θ with the radius vector from P to the center of the sphere C. Then if the law of reflection applies locally to the region around P, the reflected ray PF also makes angle θ with radius PC and crosses the axis at F. We tentatively call this the *focus* F and its distance from the sphere, *focal length* f. The reflected ray makes angle 2θ with respect to the axis. Then, from trigonometric relationships,

$$\frac{r}{R - z} = \tan \theta \tag{1}$$

$$\frac{r}{f - z} = \tan 2\theta \equiv \frac{2 \tan \theta}{1 - \tan^2 \theta} \tag{2}$$

The equation for the sphere is

$$r^2 + (R - z)^2 = R^2 \tag{3}$$

By substituting (1) in (2) and utilizing (3) we can solve for f:

$$f = R\left\{1 - \frac{1}{2}\left[1 - \left(\frac{r}{R}\right)^2\right]^{-1/2}\right\} \tag{4}$$

We see that f is a function of radius r at which rays enter, so not all rays are focused to the same point. There is spherical *aberration* as shown in Fig. 14.2*b*. However, for rays very near the axis, with r negligible in comparison with R, the expression (4) reduces to

$$f \approx \frac{R}{2} \tag{5}$$

so that in this approximation the focal length is independent of r and is just half the radius of curvature of the mirror.

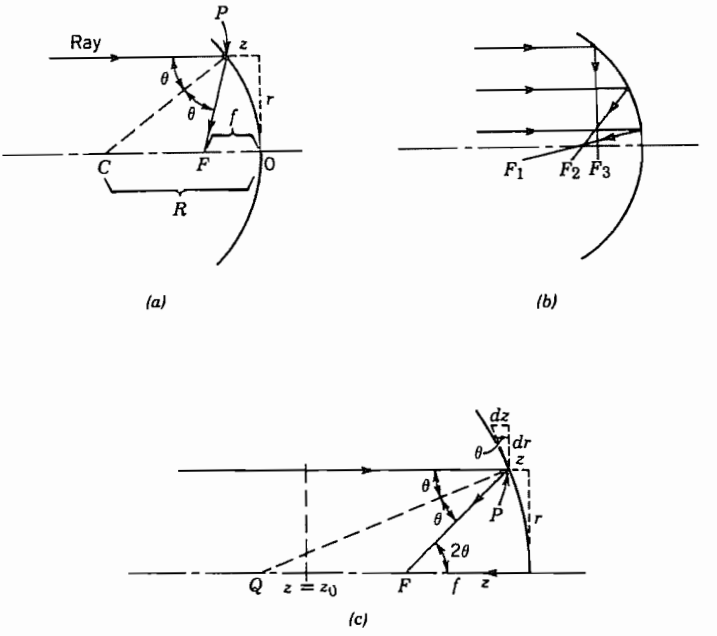


FIG. 14.2 (*a*) Ray entering spherical mirror parallel to axis. (*b*) Rays entering spherical mirror at various radii. (*c*) Parabolic mirror and ray reflected to focus F.

Spherical mirrors are used in forming the resonant systems for various lasers. They often have radii of curvature of a few meters, with rays entering a few millimeters from the axis so that the spherical aberration is small and the approximation (5) for focal length well justified.

Example 14.2b
PARABOLIC MIRROR

To eliminate the spherical aberration just shown, the reflecting surface should be a paraboloid. The paraboloid (Fig. 14.2*c*) may be defined by

$$z = \frac{r^2}{4g} \tag{6}$$

where g is a constant shown below to be focal length. Slope of the mirror at radius r is then

$$\frac{dr}{dz} = \left(\frac{dz}{dr}\right)^{-1} = \frac{2g}{r} \tag{7}$$

The slope of the normal to the mirror, PQ, is just the negative reciprocal of (7):

$$\left(\frac{dr}{dz}\right)_{PQ} = -\frac{r}{2g} = -\tan\theta \tag{8}$$

From the trigonometry of the figure,

$$\tan 2\theta = \frac{r}{f - z} \tag{9}$$

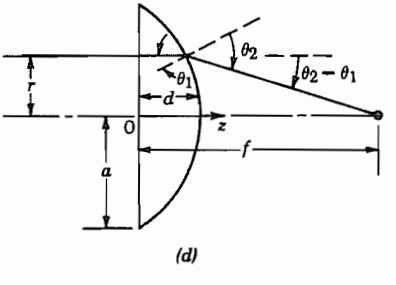
Using the identity in (2), with (8) it can be shown that f is independent of the incident radius r and equals the constant g in (6).

We have here used only the law of reflection but it is interesting to use this example to show the relation to wave optics. In the latter, one can show that all rays entering parallel to the axis and reflecting to the focus have the same phase delay from some reference plane $z = z_0$ to the focus F. Since the medium is homogeneous, this means that all such paths are of equal length, as is known to be the case for a paraboloid.

Example 14.2c
THIN LENS

We next consider the thin lens, the requirement of thinness implying that lens thickness is small in comparison with focal length. For simplicity we make one surface a plane as illustrated in Fig. 14.2*d*. The lens material is considered transparent, with index of refraction n, the surrounding medium being air or space with refractive index unity. Consider first the derivation of the law for the curved surface from Snell's law of refraction. A ray entering parallel to the axis at radius r from the axis experiences no refraction at the plane surface, but does refract at the right-hand curved surface, as shown in Fig. 14.2*d*. If θ_1 and θ_2 are the angles on the two sides from the normal to the curved surface, Snell's law gives

$$\frac{\sin\theta_2}{\sin\theta_1} = \frac{n_1}{n_2} = n \tag{10}$$



(d)

Fig. 14.2 (*d*) Thin lens showing focusing of entering parallel rays to a point.

If rays are paraxial (i.e., are nearly parallel to the axis) angles will be small so that sines may be approximated by angles themselves:

$$\theta_2 \approx n\theta_1 \tag{11}$$

Also assume that r/f is small so that

$$\frac{r}{f} = \tan(\theta_2 - \theta_1) \approx \theta_2 - \theta_1 \approx (n - 1)\theta_1 \tag{12}$$

The angle θ_1 can be related to the slope of the curved surface:

$$\frac{dz}{dr} = -\tan\theta_1 \approx -\theta_1 \approx -\frac{r}{(n - 1)f} \tag{13}$$

Upon integrating we find

$$z \approx -\frac{r^2}{2(n - 1)f} + d \tag{14}$$

where the constant d is the value of z at $r = 0$, the maximum thickness of the lens. This result is a parabolic surface, but since it was derived with several approximations, it is usual to approximate this by a spherical surface. If R is the radius of curvature of the sphere,

$$z = \sqrt{R^2 - r^2} - (R - d) \approx d - \frac{r^2}{2R} \tag{15}$$

where we assume $r \ll R$. In comparing with (14),

$$\frac{1}{f} = \frac{(n - 1)}{R} \tag{16}$$

From another point of view, the lens may be considered a phase transformer with rays at different radii from the axis arriving at the focus with the same phase. Let us illustrate this with a doubly convex lens (Fig. 14.2*e*). With approximations as in (15), z_1 and z_2 are

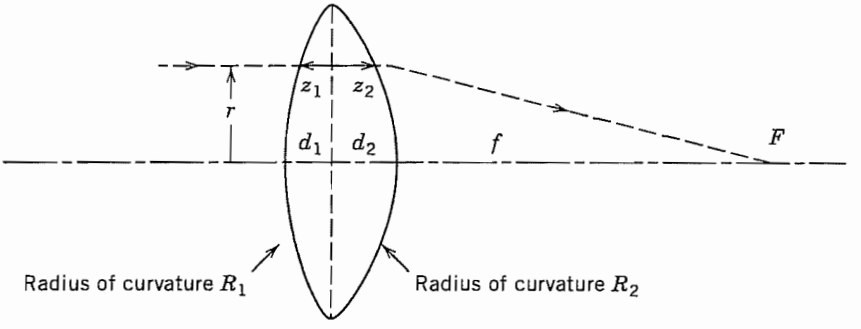
$$z_1 = d_1 - \frac{r^2}{2R_1}, \qquad z_2 = d_2 - \frac{r^2}{2R_2} \tag{17}$$

where all quantities are positive in the convention of Fig. 14.2*e*. The extra phase introduced by the lens at radius r, over that of the path in air, $n = 1$, is

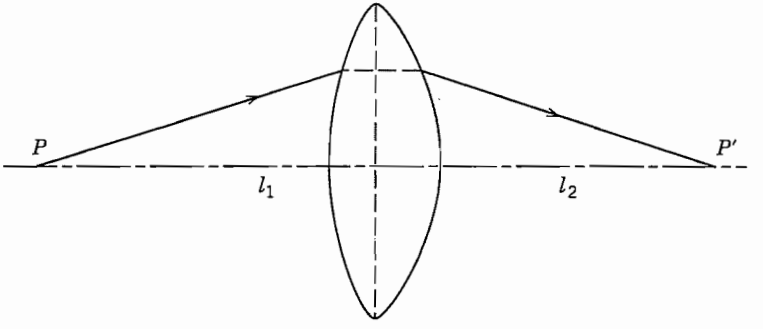
$$\phi(r) = k_0(n - 1)(z_1 + z_2) = k_0(n - 1)\left[(d_1 + d_2) - \frac{r^2}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right)\right] \tag{18}$$

This may be written in terms of the maximum radius of the lens a where z_1 and z_2 are zero:

$$\phi(r) = k_0(n - 1)(d_1 + d_2)[1 - r^2/a^2] \tag{19}$$



(e)



(f)

FIG. 14.2 (*e*) Thin lens; focal length found from phase condition. (*f*) Imaging of point P to P' by thin lens.

Now if we add the phase in going from radius r of the lens to the focus, total phase is

$$\phi = \phi(r) + k_0\sqrt{r^2 + f^2} \approx \phi(r) + k_0\left(f + \frac{r^2}{2f}\right), \qquad r \ll f \tag{20}$$

Using (18), the condition for ϕ to be independent of r is

$$\frac{k_0(n-1)r^2}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{k_0 r^2}{2f}$$

or

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{21}$$

Example 14.2d
RELATION OF OBJECT AND IMAGE DISTANCES IN THIN LENS

Either Snell's law or the requirement on constancy of phase for different paths may be used to derive the relationship for object and image distances in a thin lens. Let us use the latter method for the simple lens analyzed above. This is redrawn in Fig. 14.2*f* to show a ray path going from object point P, distance d_1 from the lens, to image point P', distance d_2 on the other side. The phase delay between the two points for a ray passing through the lens at radius r is

$$\phi_{21} = \phi(r) + k_0\left[\sqrt{l_1^2 + r^2} + \sqrt{l_2^2 + r^2}\right] \tag{22}$$

Using (18), (21), and approximations based on $r \ll l_1, l_2$,

$$\phi_{21} = k_0\left[(n-1)(d_1 + d_2) - \frac{r^2}{2f} + l_1 + \frac{r^2}{2l_1} + l_2 + \frac{r^2}{2l_1}\right] \tag{23}$$

For this to be independent of r,

$$\frac{1}{l_1} + \frac{1}{l_2} = \frac{1}{f} \tag{24}$$

which is the well-known relationship between object and image distances for a thin lens.

14.3 GEOMETRICAL OPTICS AS LIMITING CASE OF WAVE OPTICS

In the preceding section we argued the case for geometrical optics on physical grounds. Now we show formally the terms that are neglected in this approximation. Let us assume fields of the form

$$\mathbf{E} = \mathbf{e}(x, y, z)e^{-jk_0S(x,y,z)} \tag{1}$$

$$\mathbf{H} = \mathbf{h}(x, y, z)e^{-jk_0S(x,y,z)} \tag{2}$$

where $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$. We next substitute (1) and (2) in Maxwell's equations for a linear, source-free, isotropic medium, making use of the expansion of the curl of a vector multiplied by a scalar:

$$\nabla \times \mathbf{E} = \left[\nabla \times \mathbf{e} - jk_0\,\nabla S \times \mathbf{e}\right]e^{-jk_0S} = -j\omega\mu\mathbf{h}e^{-jk_0S} \tag{3}$$

$$\nabla \times \mathbf{H} = \left[\nabla \times \mathbf{h} - jk_0\,\nabla S \times \mathbf{h}\right]e^{-jk_0S} = j\omega\varepsilon\mathbf{e}e^{-jk_0S} \tag{4}$$

Rearranging, we have

$$\nabla S \times \mathbf{e} - \frac{\omega\mu}{k_0}\mathbf{h} = \frac{1}{jk_0}\nabla \times \mathbf{e} \tag{5}$$

$$\nabla S \times \mathbf{h} + \frac{\omega\varepsilon}{k_0}\mathbf{e} = \frac{1}{jk_0}\nabla \times \mathbf{h} \tag{6}$$

At this point we make the approximation that leads to the geometrical optics formulation, assuming that the multipliers **e** and **h** vary little in wavelength. The curl of **e** represents a derivative with distance, and $k_0 = 2\pi/\lambda_0$, so the right side of (5) is small if **e** varies only a small amount in distance λ_0. Similarly, the right side of (6) can be neglected if **h** likewise varies a small amount in a wavelength. Equations (5) and (6) then lead to

$$\mathbf{h} \approx \frac{1}{\mu_r\eta_0}\nabla S \times \mathbf{e} \tag{7}$$

$$\mathbf{e} \approx -\frac{\eta_0}{\varepsilon_r}\nabla S \times \mathbf{h} \tag{8}$$

Let us now interpret the expressions. If S is real, the function k_0S may be considered a phase function with surfaces S = constant being surfaces of constant phase and the gradient of S giving the local "direction of propagation." We see from (7) and (8) that **e** and **h** are then in phase and are normal to each other and to the local direction of propagation, as would be expected if the local behavior is that of a plane wave. If S is complex, (1) and (2) show that there is attenuation of the wave, and then (7) and (8) yield a phase shift between **e** and **h**, as is expected for attenuating media.

We next obtain a rather important equation relating the gradient of S to the local refractive index. To do this, we substitute (7) in (8) and expand the triple vector product to obtain

$$\mathbf{e} = -\frac{1}{\mu_r\varepsilon_r}[\nabla S(\mathbf{e}\cdot\nabla S) - \mathbf{e}(\nabla S\cdot\nabla S)] \tag{9}$$

Since **e** and ∇S are at right angles, the first term on the right is zero and

$$|\nabla S|^2 = n^2(x, y, z) \tag{10}$$

where n is refractive index $= (\mu_r\varepsilon_r)^{1/2}$. This equation is known as the *eikonal* equation of geometrical optics.[6] In rectangular coordinates,

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2(x, y, z) \tag{11}$$

Note that for a homogeneous medium with n a constant this is satisfied by

$$S = n(x\cos\alpha + y\cos\beta + z\cos\gamma) \tag{12}$$

[6] *M. Born and E. Wolf,* Principles of Optics, *6th ed., p. 112, Pergamon Press, New York, 1980.*

with

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1 \tag{13}$$

which is the phase variation one would expect for a uniform plane wave propagating in an oblique direction defined by direction cosines $\cos \alpha$, $\cos \beta$, and $\cos \gamma$. Substitution in (7) and (8) would lead to the expected relative orientation of $\mathbf{e}$ and $\mathbf{h}$ for such a wave (Prob. 14.3c).

Let us next look at power and energy relations for the rays in this formulation. The average Poynting vector is

$$\mathbf{P}_{av} = \tfrac{1}{2} \operatorname{Re}[\mathbf{E} \times \mathbf{H}^*] = \tfrac{1}{2} \operatorname{Re}[\mathbf{e} \times \mathbf{h}^*] \tag{14}$$

Substitution of (7) and expansion of the triple vector product gives

$$\mathbf{P}_{av} = \frac{1}{2\mu_r \eta_0} [\nabla S(\mathbf{e} \cdot \mathbf{e}^*) - \mathbf{e}^*(\mathbf{e} \cdot \nabla S)] \tag{15}$$

The last term is zero since $\mathbf{e}$ is normal to ∇S. The first term may be written in terms of average energy density in electric fields, $u_e = \varepsilon(\mathbf{e} \cdot \mathbf{e}^*)/4$:

$$\mathbf{P}_{av} = \frac{4u_e}{2\mu_r \eta_0 \varepsilon_r \varepsilon_0} \nabla S = \frac{c}{n} (2u_e)\hat{\mathbf{s}} \tag{16}$$

where $\hat{\mathbf{s}} = \nabla S/n$ is the unit vector in the local direction of the ray. Note that the average stored energy densities in electric and magnetic fields are equal, as can be shown from (7) and (8), using the vector identity $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$:

$$\frac{\varepsilon}{4} (\mathbf{e} \cdot \mathbf{e}^*) = -\frac{1}{4c} [\mathbf{e} \cdot (\nabla S \times \mathbf{h}^*)] = \frac{\mu}{4} (\mathbf{h} \cdot \mathbf{h}^*) \tag{17}$$

Thus the term $2u_e$ is total energy density and we see from (16) that the Poynting vector is equal in magnitude to energy density multiplied by c/n, as would be expected if local behavior is as in plane waves. It also has the ray direction given by ∇S so that a tube defined by ray directions (Fig. 14.3) has, for a lossless medium, the same power traversing each cross section. This is analogous to the continuity of flux in a flux tube as met in static fields.

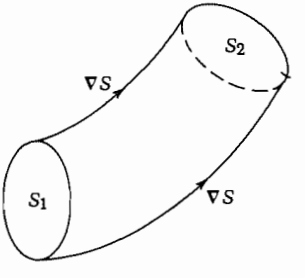


FIG. 14.3 Tube of constant power flow.

Corrections to the geometrical optics formulation have been made[7] by expanding **e** and **h** in asymptotic series in terms of wavelength and including contributions of various orders from the right-hand sides of (5) and (6).

14.4 Rays in Inhomogeneous Media

The general formulation for ray propagation given in the preceding section gave the expected but almost trivial result for homogeneous media, but is of much more use with materials for which the refractive index varies with position. In such media the rays are curved and the geometrical optics approach can often give a great deal of information about the nature of the waves. Important practical examples of inhomogeneous media include the graded index fibers used for guiding of optical waves, the surface guides of integrated optics made by diffusion or ion implantation, and the lower-frequency example of radio waves in the ionosphere.

Before applying the general formulation, we can obtain an expression for curvature of the ray and a physical understanding of the bending process, by assuming that the change of index occurs in a series of steps of differential size. If Snell's law is applied to the surface separating a differential region with index n from one with $n + dn$ as in Fig. 14.4*a*, there results

$$\frac{n}{n + dn} = \frac{\sin(\theta + d\theta)}{\sin \theta} = \frac{\sin \theta + d\theta \cos \theta}{\sin \theta} \tag{1}$$

from which we find

$$\cot \theta \, d\theta = -\frac{dn}{n} \tag{2}$$

The radius of curvature is found from the intersection of the lines erected perpendicular to the ray, and can be related to the length ds of the ray segment in the layer with index $n + dn$ and the angle $d\theta$ or, alternatively, in terms of the ray segment dR along the radius of curvature:

$$R = \frac{ds}{d\theta} = \frac{dR}{d\theta \cot \theta} \tag{3}$$

Using (2), this may be written as

$$\frac{1}{R} = -\frac{1}{n}\frac{dn}{dR} = \hat{\mathbf{R}} \cdot \nabla(\ln n) \tag{4}$$

In graphical ray tracing, the radius of curvature can be calculated from point to point as the index changes, and the ray traced by joining the corresponding arcs at suitably small intervals.

[7] *M. Kline and I. W. Kay,* Electromagnetic Theory and Geometric Optics, *Wiley Interscience, New York, 1965.*

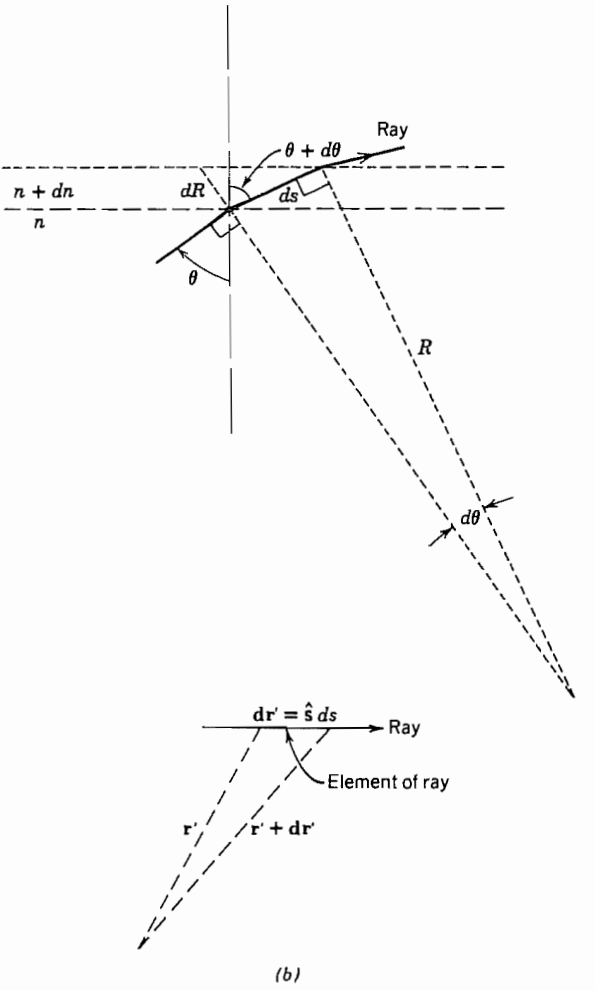


FIG. 14.4 (*a*) Construction showing radius of curvature of ray in inhomogeneous medium. (*b*) Element of ray and definitions used in general formulation.

With this physical understanding of the basis for ray curvature, let us apply the formal development of Sec. 14.3. As shown in Fig. 14.4*b*, let $\mathbf{r}'$ be the vector position of a point on the ray from an arbitrary reference. Its differential change is

$$\mathbf{dr}' = \hat{\mathbf{s}}ds \tag{5}$$

where s is distance along the ray and $\hat{\mathbf{s}}$ is a unit vector in the ray direction. But from Eq. 14.3(10) and discussion after Eq. 14.3(8), the gradient of S has a magnitude equal to refractive index and the direction of $\hat{\mathbf{s}}$:

$$\nabla S = n\hat{\mathbf{s}} \tag{6}$$

Thus

$$n\frac{d\mathbf{r}'}{ds} = \nabla S \tag{7}$$

Taking an additional derivative,

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}'}{ds}\right) = \frac{d}{ds}(\nabla S) \tag{8}$$

But because of (6)

$$\frac{d}{ds}(\nabla S) = \nabla n \tag{9}$$

so that (8) becomes

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}'}{ds}\right) = \nabla n \tag{10}$$

Still other useful forms, including the relation (4), may be derived from this (Prob. 14.4c). We now illustrate the use of the above through two practical examples.

Example 14.4a
THIN CELL WITH QUADRATIC INDEX VARIATION

As a simple but useful example, consider a thin cell (Fig. 14.4*c*) with refractive index varying quadratically with radius r from the axis:

$$n(r) = n_0\left[1 - \Delta\left(\frac{r}{a}\right)^2\right] \tag{11}$$

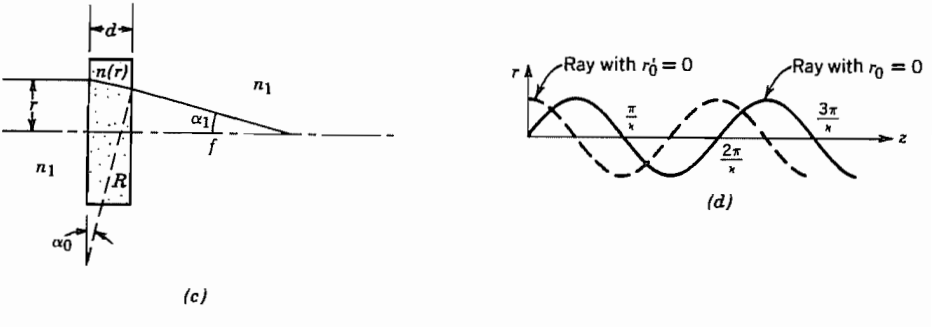


FIG. 14.4 (*c*) A cell filled with an inhomogeneous material, acting as a lens. (*d*) Ray paths through a long quadratic index medium.

Such a variation is approximated when a laser beam passes through certain materials, either because of heating or other nonlinear effects. Self-focusing or self-defocusing of the laser beam may result. A ray entering from the left, parallel to the axis and at radius r from it, is curved, with radius of curvature given by (4):

$$\frac{1}{R} = -\frac{1}{n(r)}\frac{dn(r)}{dr} = \frac{2n_0 r\Delta}{a^2 n(r)} \tag{12}$$

If cell thickness d is small, this curvature applies throughout the cell and the angle at exit, measured inside the cell, is approximately

$$\alpha_0 \approx \frac{d}{R} \tag{13}$$

There is a further deflection by Snell's law at exit from the cell, and if angles are small and the external medium has index n_1,

$$\alpha_1 \approx \frac{n(r)\alpha_0}{n_1} = \frac{2n_0 r\ \Delta\ d}{a^2 n_1} \tag{14}$$

The focal length of this lens is then

$$f \approx \frac{r}{\alpha_1} = \frac{a^2 n_1}{2n_0\ \Delta\ d} \tag{15}$$

We thus see that the cell with quadratic index variation acts as a thin lens, converging if Δ is positive (index decreasing with increasing radius) and diverging for the opposite variation.

Example 14.4b

RAYS IN GRADED INDEX FIBER WITH QUADRATIC VARIATION

Suppose the dielectric with quadratic index variation with radius is not a thin slice, as in the first example, but a long cylinder, as in graded index fibers used for optical guiding. Assume that all rays have small slopes so that $ds \approx dz$. Then (10) becomes

$$\frac{d}{dz}\left[n(r)\frac{d\mathbf{r}'}{dz}\right] = \hat{\mathbf{r}}\frac{dn(r)}{dr} \tag{16}$$

The radius vector $\mathbf{r}'$ from origin 0 to a general point on the ray is

$$\mathbf{r}' = \hat{\mathbf{r}}r + \hat{\mathbf{z}}z \tag{17}$$

Substitution in (16) leads to the simpler relation

$$\frac{d}{dz}\left[n(r)\frac{dr}{dz}\right] = \frac{dn(r)}{dr} \tag{18}$$

If $n(r)$ varies quadratically with r, as in (11),

$$\frac{d^2r}{dz^2} = \frac{-2n_0r\Delta}{a^2n_0[1 - \Delta(r/a)^2]} \approx -\left(\frac{2\Delta}{a^2}\right)r \tag{19}$$

The last approximation assumes Δ small. The solution to (19) is

$$r(z) = r_0 \cos \kappa z + \frac{r_0'}{\kappa} \sin \kappa z \tag{20}$$

where

$$\kappa = \frac{\sqrt{2\Delta}}{a} \tag{21}$$

and r_0 and r_0' are, respectively, radius and slope of the ray at $z = 0$. Two special ray paths are shown in Fig. 14.4*d*.

14.5 Ray Matrices for Paraxial Ray Optics

The rays in many optical instruments and laser resonators or guiding systems have only small slopes with respect to an axis of symmetry, as in the example of the quadratic index rod or fiber analyzed in the preceding section. In such cases the rays are called *paraxial rays* and the optics deriving from this condition is called *gaussian optics*.[8] It is convenient for this class of problems to describe the effect of various optical components by ray matrices that relate output radius and slope of a ray to the input radius and slope. The advantage of the matrix formulation is mainly in its ease of handling combinations of elements.

$$\begin{bmatrix} r_{\text{out}} \\ r'_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} r_{\text{in}} \\ r'_{\text{in}} \end{bmatrix} \tag{1}$$

We illustrate with some simple examples.

Example 14.5a
Homogeneous Medium

This is the simplest case, and although almost trivial, nevertheless an important one. As shown in Fig. 14.5*a*, the ray in a homogeneous medium is a straight line so that for

[8] *See Born and Wolf*[6] *for use with lenses; see Yariv*[1] *for use with laser resonators.*

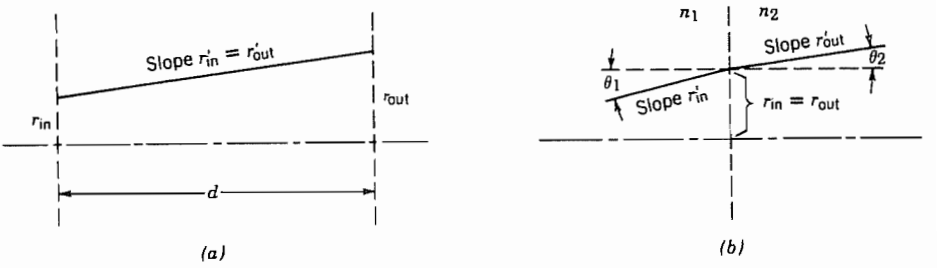


FIG. 14.5 (*a*) Ray in homogeneous medium. (*b*) Ray passing between media of different refractive index at small angle.

a distance d,

$$\begin{aligned} r_{out} &= r_{in} + dr'_{in} \\ r'_{out} &= r'_{in} \end{aligned} \quad \text{or} \quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \tag{2}$$

Example 14.5b
THIN LENS

Interpretation of Eq. 14.2(24) shows that a lens acts to change the slope of an incoming ray, but if it is thin enough, the change in radius is negligible. Thus

$$\begin{aligned} r_{out} &= r_{in} \\ r'_{out} &= r'_{in} - \frac{r_{in}}{f} \end{aligned} \quad \text{or} \quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\dfrac{1}{f} & 1 \end{bmatrix} \tag{3}$$

Example 14.5c
PARABOLIC OR SPHERICAL MIRROR

A curved mirror is similar to the lens except that the reflected ray returns rather than passing through to the other side. For a fixed coordinate system, this would result in opposite signs for the slopes of the two cases, and thus reverse the sign of the $1/f$ term in (3). However, because of the use of these in cascaded systems to be described below, it is more convenient to take the direction of the ray as determining the positive coordinate. That is, for the mirror, positive z is toward the mirror for the incident wave and away from the mirror for the reflected wave. Thus for a parabolic mirror of focal length f, the matrix is the same as in (3). For a spherical mirror, spherical aberration is negligible for paraxial rays so that $f = R/2$.

Example 14.5d
PLANE DISCONTINUITY BETWEEN DIELECTRICS

If a plane perpendicular to the axis separates a medium with refractive index n_1 from a second medium with n_2 (Fig. 14.5*b*) the ray is not changed in radius at the boundary, but slope is changed by Snell's law. For paraxial rays, sines of the angles may be replaced by slopes.

$$\begin{aligned} r_{\text{out}} &= r_{\text{in}} \\ r'_{\text{out}} &= \frac{n_1}{n_2} r'_{\text{in}} \end{aligned} \qquad \text{or} \qquad \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{n_1}{n_2} \end{bmatrix} \tag{4}$$

Example 14.5e
SPHERICAL SURFACE SEPARATING DIELECTRICS

For a spherical boundary between dielectrics with refractive indices n_1 and n_2 (Fig. 14.5*c*), use of Snell's law and the approximation of sines and tangents by their angles gives the ray matrix

$$\begin{bmatrix} 1 & 0 \\ \left(\dfrac{n_2 - n_1}{n_2}\right)\dfrac{1}{R_1} & \dfrac{n_1}{n_2} \end{bmatrix} \tag{5}$$

where R is positive in the sense shown in Fig. 14.5*c*.

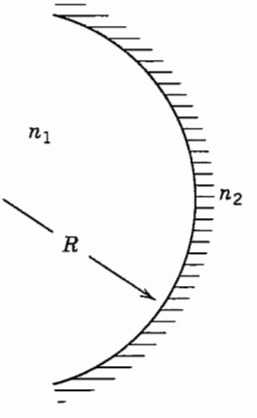


FIG. 14.5c Spherical boundary between dielectrics.

Example 14.5f
ROD WITH QUADRATIC INDEX VARIATION

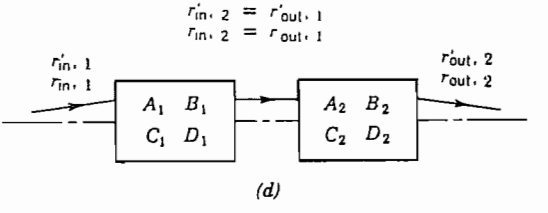
For a rod or fiber of length d with refractive index varying quadratically with radius, Eqs. 14.4(20) and (21) give the ray matrix

$$\begin{bmatrix} r_{\text{out}} \\ r'_{\text{out}} \end{bmatrix} = \begin{bmatrix} \cos \kappa d & \frac{1}{\kappa} \sin \kappa d \\ -\kappa \sin \kappa d & \cos \kappa d \end{bmatrix} \begin{bmatrix} r_{\text{in}} \\ r'_{\text{in}} \end{bmatrix}, \qquad \kappa = \frac{\sqrt{2\Delta}}{a} \tag{6}$$

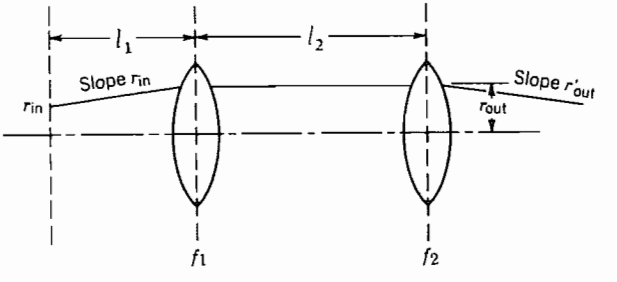
Example 14.5g
ELEMENTS IN TANDEM

We will see in this example the advantage of the matrix formulation in handling combinations of elements. Thus if one element is followed by another, as sketched in Fig. 14.5*d*, $r_{\text{out, 1}} = r_{\text{in, 2}}$ and $r'_{\text{out, 1}} = r'_{\text{in, 2}}$ so that

$$\begin{bmatrix} r_{\text{out, 2}} \\ r'_{\text{out, 2}} \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} r_{\text{in, 2}} \\ r'_{\text{in, 2}} \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} r_{\text{in, 1}} \\ r'_{\text{in, 1}} \end{bmatrix} \tag{7}$$



(d)



(e)

FIG. 14.5 (*d*) Cascade arrangement of two general optical elements. (*e*) Cascade arrangement of two thin lenses with homogeneous regions between and on each side.

The ray matrix of the combination is thus the product of the individual ray matrices starting from the one nearest the output. The procedure is extended in a straightforward manner to any number of elements in cascade, with the final element first in the product chain. As an example, consider the combination of two lenses and two homogeneous regions sketched in Fig. 14.5e. The ray matrix for the combination is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f_2} & 1 \end{bmatrix} \begin{bmatrix} 1 & l_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f_1} & 1 \end{bmatrix} \begin{bmatrix} 1 & l_1 \\ 0 & 1 \end{bmatrix} \tag{8}$$

from which

$$\begin{aligned} A &= 1 - \frac{l_2}{f_1} \\ B &= l_1 + l_2\left(1 - \frac{l_1}{f_1}\right) \\ C &= -\frac{1}{f_1} - \frac{1}{f_2}\left(1 - \frac{l_2}{f_1}\right) \\ D &= -\frac{1}{f_2}\left[l_1 + l_2\left(1 - \frac{l_1}{f_1}\right)\right] + \left(1 - \frac{l_1}{f_1}\right) \end{aligned} \tag{9}$$

We shall interpret these expressions in the following section to show that this combination is useful in forming optical guiding systems or optical resonators.

14.6 Guiding of Rays by a Periodic Lens System or in Spherical Mirror Resonators

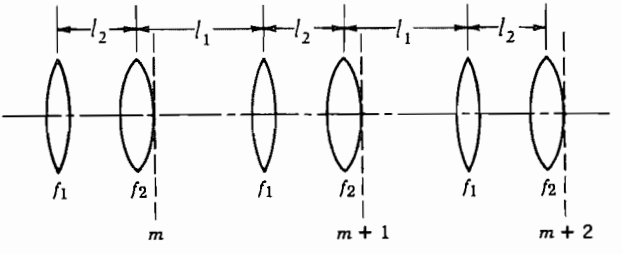
If the lens pair of Fig. 14.5*d* is repeated, as in Fig. 14.6*a*, a periodic system results. It will be seen that for such systems, rays may be guided or confined for certain combinations of parameters, but may diverge for other combinations. Let us apply the ray matrix between reference planes m and $m + 1$ with A, B, C, D defined for one unit of the periodic system:

$$r_{m+1} = Ar_m + Br'_m \tag{1}$$

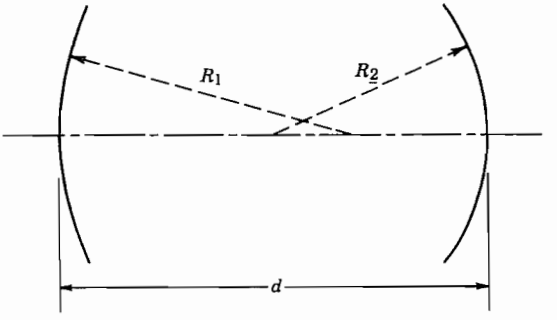
$$r'_{m+1} = Cr_m + Dr'_m \tag{2}$$

Equation (1) may be solved for r'_m.

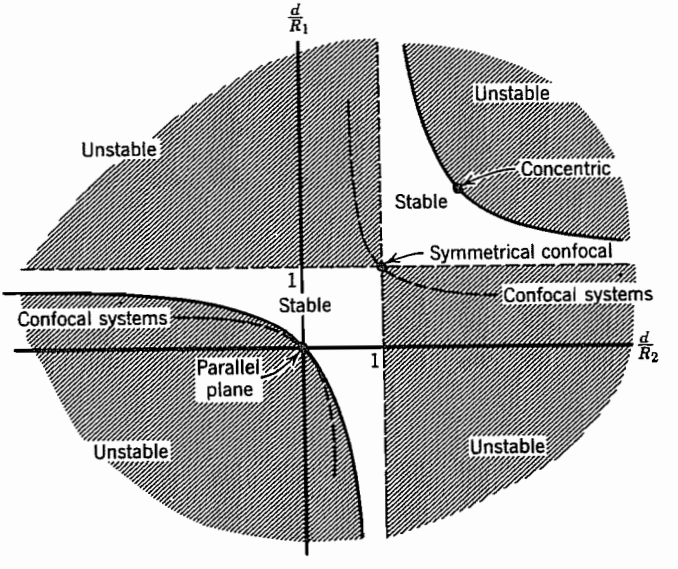
$$r'_m = \frac{1}{B}(r_{m+1} - Ar_m) \tag{3}$$



(a)



(b)



(c)

FIG. 14.6 (*a*) A periodic lens system. (*b*) Spherical mirror resonator showing two coaxial mirrors separated by distance *d*. (*c*) Diagram showing stable (confined) conditions for spherical mirror resonator of (*b*).

For this periodic system, a similar equation applies between references $m + 1$ and $m + 2$:

$$r'_{m+1} = \frac{1}{B}(r_{m+2} - Ar_{m+1}) \tag{4}$$

If (3) and (4) are substituted in (2) there results

$$r_{m+2} - (A + D)r_{m+1} + (AD - BC)r_m = 0 \tag{5}$$

For optical systems of concern to us, it can be shown that $AD - BC = 1$. (This can be checked from Eq. 14.5(9) for the system of Fig. 14.6*a*.) Equation (5) then simplifies to

$$r_{m+2} - (A + D)r_{m+1} + r_m = 0 \tag{6}$$

This linear difference equation may be solved by assuming solutions of exponential form, much as with linear differential equations with constant coefficients. Let

$$r_m = r_0 e^{\pm jm\theta} \tag{7}$$

Substitution of (7) in (6) yields

$$e^{\pm 2j\theta} - (A + D)e^{\pm j\theta} + 1 = 0 \tag{8}$$

This is a quadratic in $e^{\pm j\theta}$ so that the solution is

$$e^{\pm j\theta} = \left(\frac{A + D}{2}\right) \pm \left[\left(\frac{A + D}{2}\right)^2 - 1\right]^{1/2} \tag{9}$$

But since $e^{j\theta} = \cos\theta + j \sin\theta$, this is consistent with

$$\cos\theta = \left(\frac{A + D}{2}\right) \tag{10}$$

Thus there are real solutions for θ with r_m bounded if

$$-1 \leq \left(\frac{A + D}{2}\right) \leq 1 \tag{11}$$

But if (11) is not satisfied, θ is imaginary and one of the solutions (7) is a growing exponential as m increases so that the ray is not bounded. The two cases are frequently called *stable* and *unstable*, respectively.

For the system of Fig. 14.6*a*, Eqs. 14.5(9) apply and the condition for a stable or confined solution is

$$-1 \leq 1 - (l_1 + l_2)\left(\frac{1}{2f_1} + \frac{1}{2f_2}\right) + \frac{l_1 l_2}{2f_1 f_2} \leq 1 \tag{12}$$

If all spacings are the same, $l_1 = l_2 = d$, (12) may be rearranged so that the stability condition is

$$0 \leq \left(1 - \frac{d}{2f_1}\right)\left(1 - \frac{d}{2f_2}\right) \leq 1 \tag{13}$$

The above discussion for the periodic lens system may be applied directly to the spherical mirror resonator by the relations between mirrors and lenses discussed in Sec. 14.5. A typical optical resonator used in laser systems consists of two spherical mirrors with radii of curvature R_1 and R_2, aligned with a common axis (Fig. 14.6*b*). Rays bounce back and forth between the two mirrors (in the geometrical optics picture), and if the system is stable, they are confined within some maximum radius. From Sec. 14.5, this system is equivalent to the periodic lens systems of Fig. 14.6*a* with $l_1 = l_2 = d$ so that (13) applies directly, with $f_1 = R_1/2$ and $f_2 = R_2/2$. The stability condition for Fig. 14.6*b* is thus

$$0 \leq \left(1 - \frac{d}{R_1}\right)\left(1 - \frac{d}{R_2}\right) \leq 1 \tag{14}$$

It will be shown later that this condition is confirmed by a wave analysis, so it is a very important relation for such resonators. Figure 14.6*c* shows the stable (unshaded) and unstable (shaded) regions in a plane with d/R_1 as ordinate and d/R_2 as abscissa. The special cases shown are the parallel plane ($R_1 = R_2 = \infty$), the concentric ($R_1 = R_2 = d/2$), and the confocal ($R_1/2 + R_2/2 = d$). These will be discussed more in Sec. 14.15 using the wave analysis.

Dielectric Optical Waveguides

14.7 Dielectric Guides of Planar Form

The principle of guiding electromagnetic waves by dielectric guides was shown in Sec. 9.2. Such guides have become very useful for optical communication devices. We will consider the optical fibers used for transmission of optical communication signals in later sections. Here we want to consider the simpler planar forms, which have been utilized in the thin-film devices of integrated optics. Figure 14.7*a* shows such a guide with three dielectrics separated by parallel-plane interfaces. The central medium 2 is often called the *film*, the lower region 3 the *substrate*, and the upper region 1 the *superstrate* or *cover*. If the refractive index of the film is higher than that of the materials above and below, there is the possibility of having guided waves through the phenomenon of total reflection as explained in Sec. 9.2. Here, however, we wish to analyze by returning to the basic field equations.

We assume propagation as $e^{-j\beta z}$ in the z direction of Fig. 14.7*a* and neglect variations with y. In this case the wave solutions divide into TM and TE types. We first consider the latter, with the components E_y, H_x, and H_z. The Helmholtz equation for E_y in each

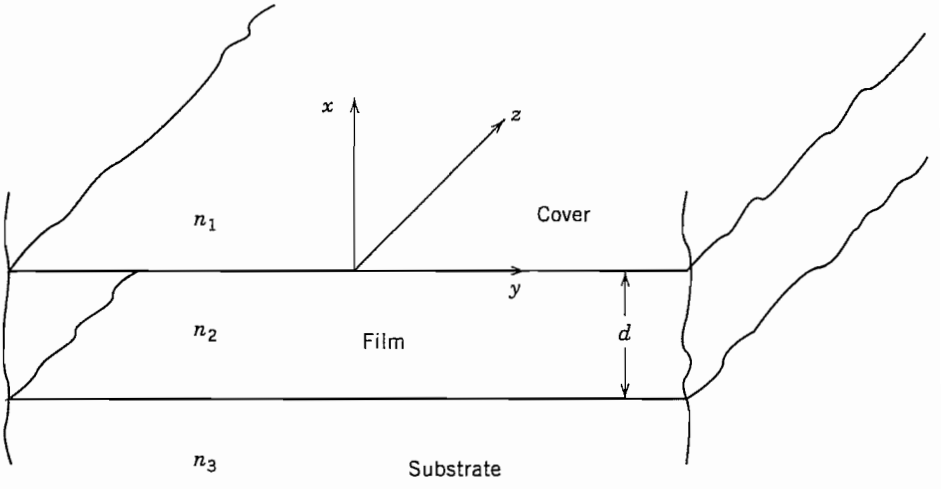


FIG. 14.7a Section of parallel-plane dielectric guide. Propagation is in the z direction.

region is then

$$\frac{d^2E_{yi}}{dx^2} = (\beta^2 - k_i^2)\, E_{yi}, \qquad i = 1,2,3 \tag{1}$$

Solutions of this are either exponentials or sinusoids. We choose exponentials in regions 1 and 3 so that fields may decay with increasing distance from the film. From continuity conditions, it can be shown (Prob. 14.7a) that the solution in the film must then be of sinusoidal form in x. Components H_x and H_z are found, respectively, from the x and z components of the Maxwell equation

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \tag{2}$$

The field components of the TE waves in the three regions are then

$$\left.\begin{aligned} E_{y1} &= Ae^{-qx} = -\left(\frac{\omega\mu}{\beta}\right)H_{x1} \\ H_{z1} &= -\frac{1}{j\omega\mu}\frac{\partial E_y}{\partial x} = \frac{q}{j\omega\mu}Ae^{-qx} \end{aligned}\right\} x > 0 \tag{3}$$

$$\left.\begin{aligned} E_{y2} &= B\cos hx + C\sin hx = -\left(\frac{\omega\mu}{\beta}\right)H_{x2} \\ H_{z2} &= \frac{h}{j\omega\mu}\,[B\sin hx - C\cos hx] \end{aligned}\right\} -d \le x \le 0 \tag{4}$$

$$\left.\begin{aligned} E_{y3} &= De^{px} = -\left(\frac{\omega\mu}{\beta}\right)H_{x3} \\ H_{z3} &= -\frac{p}{j\omega\mu}De^{px} \end{aligned}\right\} x < -d \tag{5}$$

where

$$q^2 = \beta^2 - k_1^2, \qquad h^2 = k_2^2 - \beta^2, \qquad p^2 = \beta^2 - k_3^2 \tag{6}$$

Continuity of tangential field components requires that E_y and H_z be continuous at $x = 0$, and again at $x = -d$. The four resulting equations allow reduction from the four arbitrary constants A, B, C, and D to a single one and development of the determinantal equation

$$\tan hd = \frac{h(q + p)}{h^2 - pq} \tag{7}$$

Although this can be solved graphically[9] for the symmetric case and for the useful case in which $n_3 - n_1 \gg n_2 - n_3$ (Prob. 14.7e), it has been solved numerically[10] and the important results are shown in Fig. 14.7*b* with these definitions:

$$v = \frac{2\pi d}{\lambda_0}\sqrt{n_2^2 - n_3^2} \tag{8}$$

$$b = \frac{n_{\text{eff}}^2 - n_3^2}{n_2^2 - n_3^2} \tag{9}$$

$$a = \frac{n_3^2 - n_1^2}{n_2^2 - n_3^2} \tag{10}$$

where

$$n_{\text{eff}} = \frac{\beta}{k_0} \tag{11}$$

It is seen that the parameter v is proportional to the ratio of thickness to wavelength, but also depends upon the differences in refractive index between guiding region and substrate. Parameter b determines the value of β in terms of an effective refractive index n_{eff}. Note that when $b = 0$, the wave travels with the velocity of light in the substrate material, and when b is unity, it travels with the velocity in the film material. The parameter a describes the degree of asymmetry. For $a = 0$, $n_1 = n_3$, and for $a = \infty$, $n_3 - n_1 \gg n_2 - n_3$. As an example of use of the figure, take a glass film on silica substrate with air above, $n_1 = 1$, $n_2 = 1.55$, $n_3 = 1.50$. Thickness $d = 1.46$ μm and $\lambda_0 = 0.6$ μm. Then $v = 6.0$ from (8) and $a = 8.2$ from (10). From Fig. 14.7*b* we find that two modes, the $m = 0$ and $m = 1$ modes, are guided. Their b values are read as 0.82 and 0.31, respectively. Using (9), this gives $n_{\text{eff}} = 1.541$ and 1.516, respectively. Both have velocities between light velocities of film and substrate, but the higher-order wave is nearer cutoff and so has more energy in the substrate and travels with a velocity near c/n_3. The lower-order mode is further from cutoff, has more of its energy in the film, and so has velocity nearer c/n_2.

[9] *R. E. Collin,* Field Theory of Guided Waves, *2nd ed., IEEE Press, Piscataway, NJ, 1991.*
[10] *H. Kogelnik and V. Ramaswamy,* Appl. Opt. ***13****, 1857 (1974).*

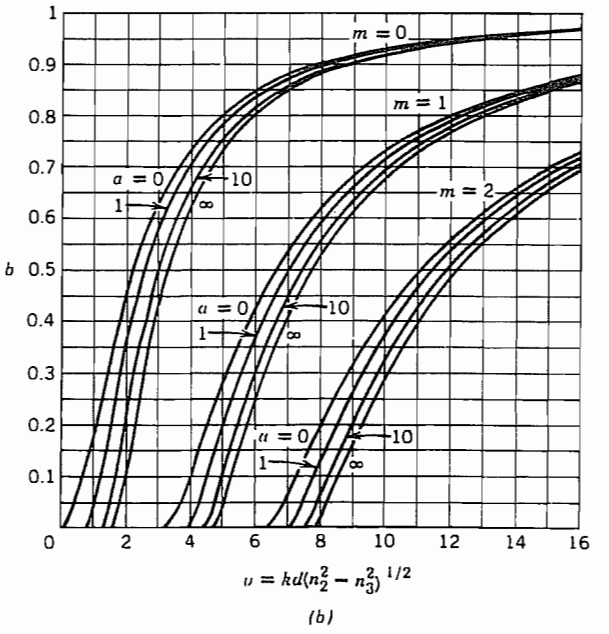


FIG. 14.7*b* Diagram showing normalized phase constant for first few modes of planar slab dielectric waveguides versus normalized frequency, with a range of parameters. (Taken from Kogelnik and Ramaswamy.[10])

For TM waves the analysis is essentially the dual of the above except that we must take into account the different ε's (μ was assumed the same for all regions). The resulting determinantal equation is

$$\tan hd = \frac{h[pn_2^2/n_3^2 + qn_2^2/n_1^2]}{h^2 - pqn_2^4/n_1^2n_3^2} \tag{12}$$

Figure 14.7*b* may also be used for this case when $n_2 \approx n_3$ if a is taken as

$$a_{\text{TM}} = \frac{n_2^4\,(n_3^2 - n_1^2)}{n_1^4\,(n_2^2 - n_3^2)} \tag{13}$$

Note from the figure that for all modes except the $m = 0$ mode for the symmetric case ($n_1 = n_3$), there is a cutoff condition. For lower frequencies or lesser thicknesses, the modes change from guided modes to radiating modes as the condition for total reflection from the interfaces can no longer be satisfied. The ranges are (assuming $n_2 > n_3 > n_1$)

Guided waves	$k_0n_3 < \beta < k_0n_2$
Substrate radiation modes	$k_0n_1 < \beta < k_0n_3$
Modes with radiation above and below	$0 < \beta < k_0n_1$
Physically unrealizable (growing in both regions)	$\beta > k_0n_2$

Although there is only a discrete set of guided modes, there is a continuous spectrum of radiation modes. It can be shown that the modes (including the radiation modes) are orthogonal over the interval $-\infty < x < \infty$ and that the totality of guided and radiation modes form a complete set. However, the set is awkward to use for expansion of arbitrary functions because of the infinite extent of the radiation modes, and their continuous spectrum.[11–13]

Finally we note that although we have developed the determinantal equation from a direct solution of Maxwell's equations, it can also be developed rigorously from the picture of wave reflections at an angle, introduced in Sec. 9.2.

14.8 Dielectric Guides of Rectangular Form

The study of planar guides in the preceding section established the basic principles of dielectric guiding but for most practical purposes it is desirable to confine the wave laterally as well as in depth. One useful configuration is that of a dielectric of rectangular cross section embedded in a substrate of lower index, as indicated in Fig. 14.8*a*. Often the higher dielectric region is formed by diffusion or ion implantation to make an inhomogeneous guiding region, as illustrated in Fig. 14.8*b*. The latter case can be analyzed only numerically once the distribution of index is known (although analytic solutions for certain index variations have been given[13]), but analysis of the rectangular approximation to Fig. 14.8*b* may still be useful. Even the rectangular configuration is hard to analyze, but approximate solutions are available which give a good idea of the behavior.

A numerical analysis of a dielectric guide of rectangular cross section surrounded by a dielectric of lower index was given by Goell.[14] This utilized expansions of the wave in circular harmonics. Most often, one is concerned with operation well above cutoff so that most of the energy is concentrated in the guiding region, and Marcatili has supplied a very useful approximate method for this range.[15] Consider the rectangular dielectric 1 of Fig. 14.8*c*. In the general case, different dielectrics 2, 3, 4, and 5 surround the guiding region. Except near cutoff, the evanescent fields in the external regions die off rapidly so that one need not worry about the difficult problem of matching fields in the shaded corner regions. Moreover, if index differences are not too great, the waves are nearly transverse electromagnetic and are found to break into two classes: E^x_{pq} with

[11] *D. Marcuse,* Light Transmission Optics, *2nd ed., Krieger, Melbourne, FL, 1982.*

[12] *D. Marcuse,* Theory of Dielectric Optical Waveguides, *2nd ed., Academic Press, San Diego, CA, 1991.*

[13] *H. Kogelnik, in* Guided Wave Optoelectronics *(T. Tamir, Ed.), 2nd ed., Springer-Verlag, New York, 1990.*

[14] *J. E. Goell,* Bell. Syst. Tech. J. ***48****, 2133 (1969).*

[15] *E. A. J. Marcatili,* Bell Syst. Tech. J. ***48****, 2071 (1969).*

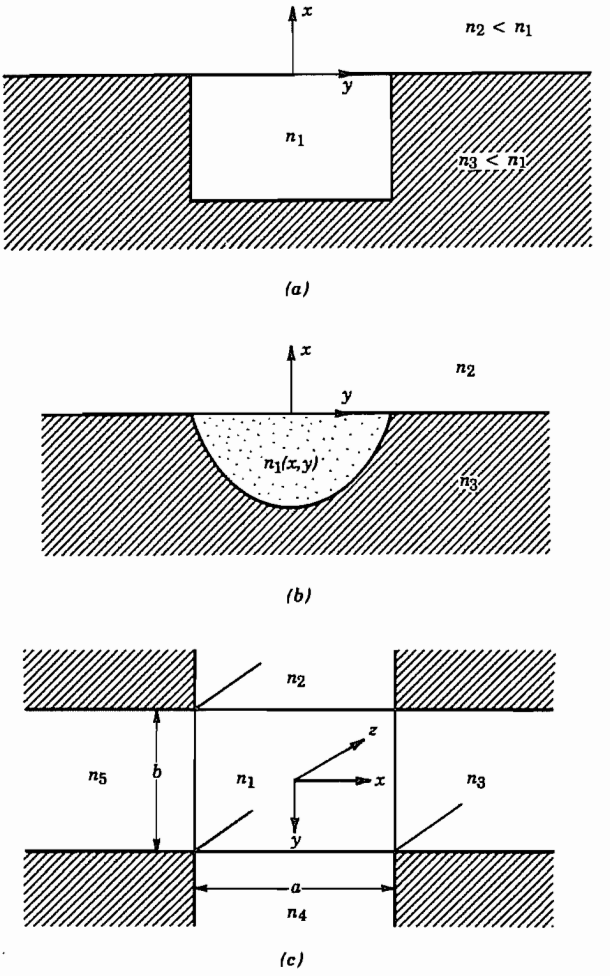


FIG. 14.8 (*a*) Dielectric guide with guiding confined in the *y* direction. (*b*) Similar inhomogeneous guide made by diffusion or ion implantation techniques. (*c*) Model for Marcatili analysis.

negligible E_y, and E_{pq}^y with negligible E_x. For the latter class, with propagation as $e^{-j\beta z}$,

$$E_{y1} = C_1 \cos(k_x x + \phi_1) \cos(k_y y + \phi_2) \tag{1}$$

$$E_{y2} = C_2 \cos(k_x x + \phi_1) e^{\alpha_2 y} \tag{2}$$

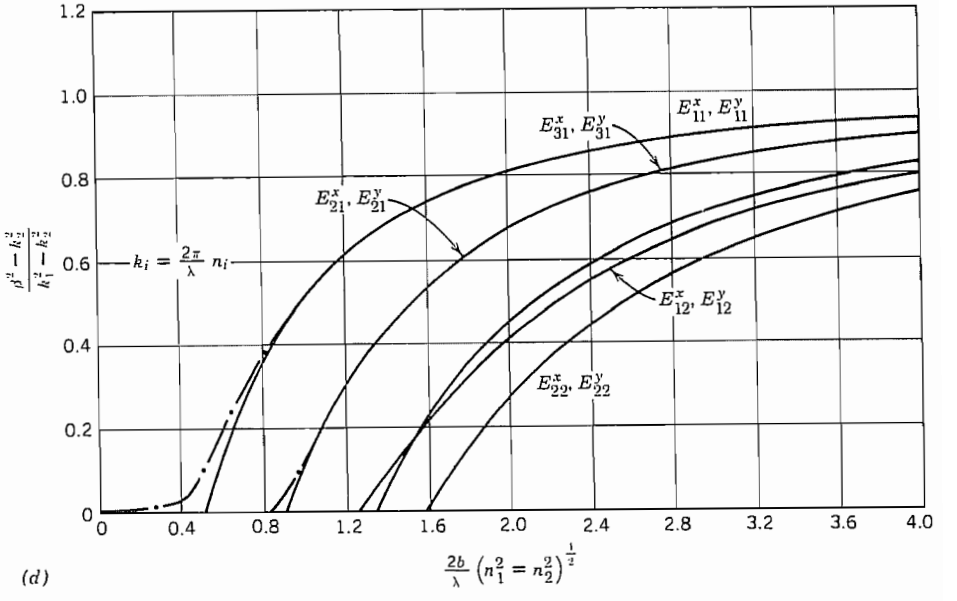
$$E_{y3} = C_3 e^{-\alpha_3 x} \cos(k_y y + \phi_2) \tag{3}$$

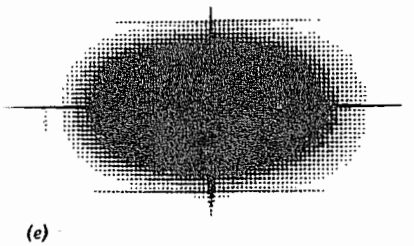
where

$$\alpha_2 = [\beta^2 + k_x^2 - k_2^2]^{1/2} \tag{4}$$

$$\alpha_3 = [\beta^2 + k_y^2 - k_3^2]^{1/2} \tag{5}$$

The solution for region 4 is similar to that for 2, and the solution for region 5 is similar to that for 3, but with exponentials of opposite sign. Remaining field components are obtained from Maxwell's equations. Components E_x and H_y are negligible for this class; continuity of other tangential components at the four boundaries $x = \pm a/2, y = \pm b/2$ relate all constants to C_1 and give the determinantal equation for β. Figure 14.8*d* gives results when all surrounding dielectrics are the same, $n_5 = n_4 = n_3 = n_2$, with $n_1/1.05$




(e)

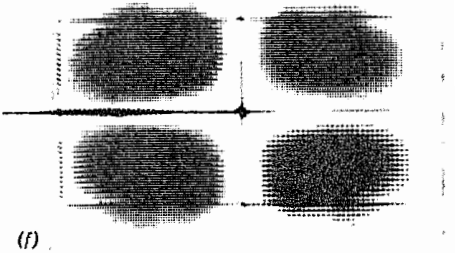

(f)

FIG. 14.8 (*d*) Curves giving normalized phase constant versus normalized frequency of several modes in rectangular dielectric guide surrounded by a common dielectric. Solid curves are by Marcatili's approximate method,[15] and dot–dash curves by Goell's computer solutions.[14] (*e*) Intensity picture of E_{11}^y mode for $a/b = 2$ from Goell.[14] (*f*) Similar picture for E_{22}^y mode.[14] Figs. *d*, *e*, and *f* reprinted with permission of the *Bell System Technical Journal*, Copyright 1969, AT&T.

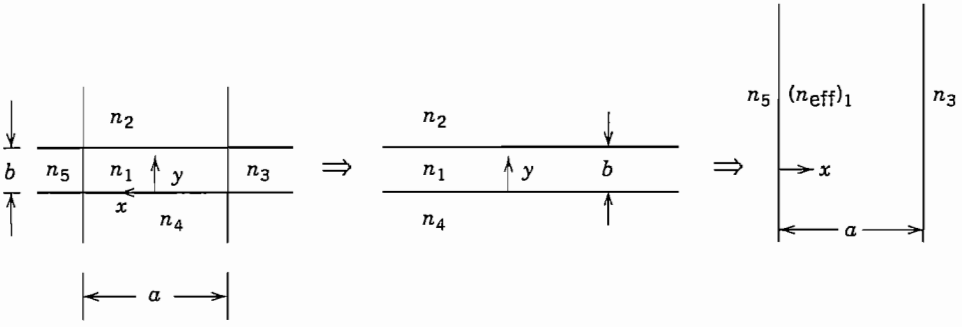


FIG. 14.8*g* Two-step method for obtaining effective index for a rectangular dielectric guide.

$< n_2 < n_1$ and $b/a = \frac{1}{2}$. Also plotted for comparison are results from Goell's computer calculations for these parameters showing that there is good agreement except near cutoff. Figures 14.8*e* and *f* show a picture of two E^y_{mn} modes in a dielectric with $b/a = \frac{1}{2}$, $n_1 = 1.02$, $n_2 = 1$, and $(2b/\lambda_0)(n_1^2 - n_2^2)^{1/2} = 2$.

Effective Index Method Many important guiding systems used with semiconductor lasers or integrated optics have lateral dimensions of the guiding regions appreciably larger than the depths. That is, $a \gg b$ in Fig. 14.8*c*. In such cases a simple method called the *effective index method* has been used to give useful approximate results.[16] The basis for this approach is in the recognition that for such cases the variations in the vertical direction are dominant, assuming comparable mode orders in x and y. Thus, in the Helmholtz equation with the $e^{-j\beta z}$ propagation factor,

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} = (\beta^2 - k^2)E \tag{6}$$

the second derivative in x can be neglected in the first approximation. The solution is then that of the planar guide of Sec. 14.7, with y as the variable instead of x. The curves of Fig. 14.7*b* may then be used to give a first approximation to effective index, $(n_{\text{eff}})_1$. Confinement in the x direction is then accounted for by using this effective index between the side dielectrics, n_3 and n_5, and again using the planar analysis of Fig. 14.7*b* with only x variations, leading to an improved $n_{\text{eff}} = \beta/k_0$. The two steps are indicated in Fig. 14.8*g*.

As with the Marcatili method, the effect of the corner materials is omitted, and as with that method, results are best when operation is well above cutoff.

[16] *J. Bous,* IEEE J. Quant. Electronics **QE-18**, *1083 (1982).*

14.9 Dielectric Guides of Circular Cross Section

Many practical dielectric guides, including the fibers used for optical communications, are of circular cross section. We consider in this section a dielectric with uniform permittivity ε_1 extending to $r = a$, with a second dielectric of lower permittivity ε_2 surrounding it, and permeability of both materials taken as μ_0 (Fig. 14.9*a*). (This is called a *step index fiber* in optical communications; the useful *graded index fiber* is discussed in the next section.)

Any rectangular component of field, such as E_z, satisfies the scalar Helmholtz equation, which for a wave propagating as $e^{\mp j\beta z}$, in circular cylindrical coordinates, is

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r}\frac{\partial E_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_z}{\partial \phi^2} + (k^2 - \beta^2)E_z = 0$$

For guided modes, $k_1^2 - \beta^2 > 0$ in the core, so ordinary Bessel functions result there. Only the first kind is used, to maintain fields finite on the axis. For the outer material or cladding, $\beta^2 - k_2^2 > 0$ for guided modes so that modified Bessel functions are utilized. Only the second solution is retained, so that fields die off properly at infinity. Axial field H_z satisfies a similar equation and we may write

$$\begin{aligned}
&r < a && r > a \\
E_{z1} &= AJ_l\left(\frac{ur}{a}\right)\begin{Bmatrix}\cos l\phi \\ \sin l\phi\end{Bmatrix} && E_{z2} = CK_l\left(\frac{wr}{a}\right)\begin{Bmatrix}\cos l\phi \\ \sin l\phi\end{Bmatrix} \\
H_{z1} &= BJ_l\left(\frac{ur}{a}\right)\begin{Bmatrix}\sin l\phi \\ \cos l\phi\end{Bmatrix} && H_{z2} = DK_l\left(\frac{wr}{a}\right)\begin{Bmatrix}\sin l\phi \\ \cos l\phi\end{Bmatrix} \\
u^2 &= (k_1^2 - \beta^2)a^2 && w^2 = (\beta^2 - k_2^2)a^2
\end{aligned} \tag{1}$$

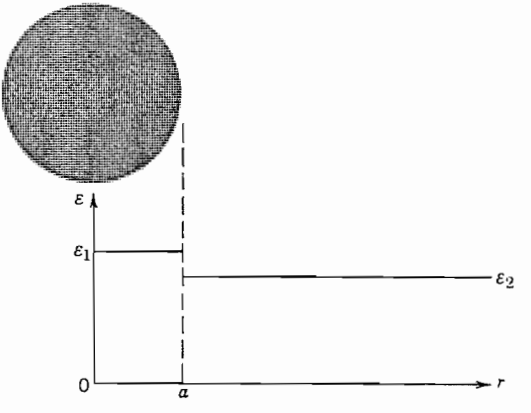


Fig. 14.9*a* Step index optical fiber.

The transverse field components may be obtained from these through Eqs. 8.9(1)–(4), where $k_{c1}^2 = u^2/a^2$ for region 1 and $k_{c2}^2 = -w^2/a^2$ for region 2. The results for E_ϕ and H_ϕ (needed in applying continuity) are

$$E_{\phi 1} = \left[\pm \frac{j\beta a^2 l}{ru^2} AJ_l\left(\frac{ur}{a}\right) + \frac{j\omega\mu a}{u} BJ_l'\left(\frac{ur}{a}\right) \right] \begin{Bmatrix} \sin l\phi \\ \cos l\phi \end{Bmatrix} \tag{2}$$

$$E_{\phi 2} = \left[\mp \frac{j\beta a^2 l}{rw^2} CK_l\left(\frac{wr}{a}\right) - \frac{j\omega\mu a}{w} DK_l'\left(\frac{wr}{a}\right) \right] \begin{Bmatrix} \sin l\phi \\ \cos l\phi \end{Bmatrix} \tag{3}$$

$$H_{\phi 1} = \left[-\frac{j\omega\varepsilon_1 a}{u} AJ_l'\left(\frac{ur}{a}\right) \mp \frac{j\beta a^2 l}{u^2 r} BJ_l\left(\frac{ur}{a}\right) \right] \begin{Bmatrix} \cos l\phi \\ \sin l\phi \end{Bmatrix} \tag{4}$$

$$H_{\phi 2} = \left[\frac{j\omega\varepsilon_2 a}{w} CK_1'\left(\frac{wr}{a}\right) \pm \frac{j\beta a^2 l}{w^2 r} DK_l\left(\frac{wr}{a}\right) \right] \begin{Bmatrix} \cos l\phi \\ \sin l\phi \end{Bmatrix} \tag{5}$$

For the symmetric case with $l = 0$, the solutions break into separate TM and TE sets, the former with E_z, H_ϕ, and E_r (expression for the last component not shown) and the TE set with H_z, E_ϕ, and H_r. The continuity condition of $E_{z1} = E_{z2}$ and $H_{\phi 1} = H_{\phi 2}$ at $r = a$ gives for the TM set

$$\frac{J_1(u)}{J_0(u)} = -\frac{\varepsilon_2 u}{\varepsilon_1 w} \frac{K_1(w)}{K_0(w)} \tag{6}$$

The continuity condition of $H_{z1} = H_{z2}$, $E_{\phi 1} = E_{\phi 2}$ at $r = a$ for the TE set gives the same equation with the factor $\varepsilon_2/\varepsilon_1$ missing. Cutoff for the dielectric guide may be considered to be the condition for which fields in the outer guide extend to infinity, which happens if $w = 0$. For $w = 0$, $K_1/K_0 = \infty$, so that $J_0(u_c) = 0$. So at cutoff,

$$u_c = a(k_1^2 - k_2^2)^{1/2} = 2.405, 5.52, \ldots \tag{7}$$

If $l \neq 0$, the fields do not separate into TM and TE types, but all fields become coupled through the continuity conditions. Applying continuity of E_z, H_ϕ, H_z, and E_ϕ at $r = a$, the four arbitrary constants of (1)–(5) reduce to a single constant and there results the determinantal equation

$$\left[\frac{k_1 J_l'(u)}{uJ_l(u)}\right]^2 + \left[\frac{k_2 K_l'(w)}{wK_l(w)}\right]^2 + \frac{(k_1^2 + k_2^2)}{uw}\left[\frac{J_l'(u)K_l'(w)}{J_l(u)K_l(w)}\right] = \frac{\beta^2 l^2 v^4}{u^4 w^4} \tag{8}$$

where

$$v = \sqrt{u^2 + w^2} = \frac{2\pi a}{\lambda}\sqrt{n_1^2 - n_2^2} \tag{9}$$

Solutions of this transcendental equation leads to *hybrid modes.* Although not purely TM or TE, H_z is dominant in one set of solutions, designated HE_{lp} modes, whereas E_z is dominant in a set designated EH_{lp} modes. Curves of β/k_0 as a function of v are shown for several of the modes in Fig. 14.9*b*.[17] The HE_{11} mode is special in that it has

[17] *D. B. Keck, in* Fundamentals of Optical Fiber Communications *(M. L. Barnoski, Ed.), Academic Press, San Diego, CA, 1976.*

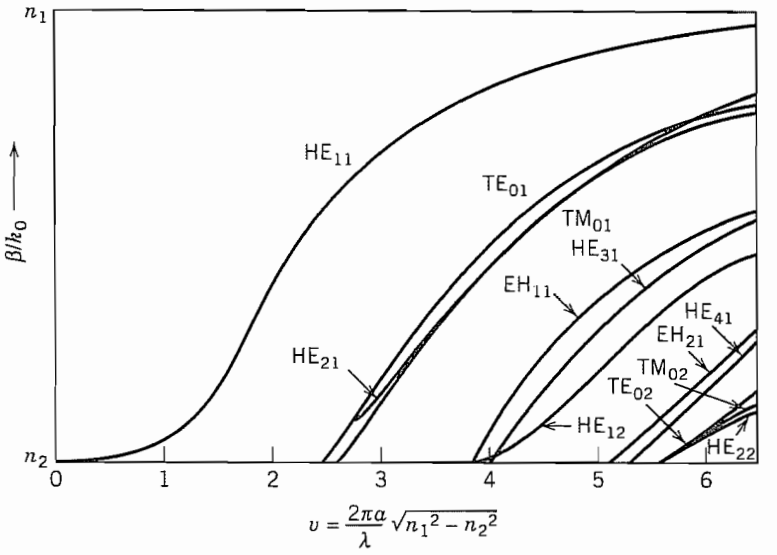


FIG. 14.9*b* Normalized propagation constant as a function of v parameter for a few of the lowest-order modes of a step waveguide.[17]

no cutoff frequency and so is often called the *dominant mode*. Although it has no strict cutoff, energy is primarily in the guiding core only when the core size is appreciable in comparison with wavelength. This mode has been used in dielectric radiators[18] and is also important in optical fibers.[19] Some photographs of the light distribution in various modes or combinations of modes are shown in Fig. 14.9*c*. For high-data-rate fiber communications it is desirable to have only one propagating mode to avoid intermode distortion, as will be further discussed in Sec. 14.11.

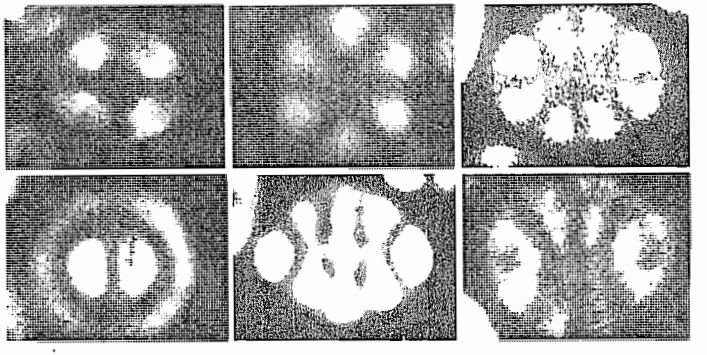


FIG. 14.9*c* Photographs of modes in the step-index optical fiber. (From Ref. 19.)

[18] *G. E. Mueller and W. A. Tyrrell,* Bell Syst. Tech. J. ***26****, 837 (1947).*
[19] *N. S. Kapany,* J. Opt. Soc. Am. ***51****, 1067 (1961).*

where

$$k^2(0) = \omega^2\mu_0\varepsilon(0) = \frac{\omega^2}{c^2}n^2(0) = \left(\frac{2\pi}{\lambda_0}\right)^2 n^2(0) \tag{10}$$

Let us first consider circular cylindrical coordinates with no ϕ variations. Equation (9) then becomes

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r}\frac{\partial E}{\partial r} + \frac{\partial^2 E}{\partial z^2} + k^2(0)\left[1 - 2\Delta\left(\frac{r}{a}\right)^2\right]E = 0 \tag{11}$$

It can be shown by substitution that this is solved by

$$E(r, z) = Ae^{-(r/w)^2}e^{-j\beta z} \tag{12}$$

where

$$w = \left(\frac{2a^2}{\Delta k^2(0)}\right)^{1/4} = \left[\frac{a\lambda_0}{\pi n(0)}\right]^{1/2}\left(\frac{1}{2\Delta}\right)^{1/4} \tag{13}$$

and

$$\beta = k(0) - \frac{2}{w^2 k(0)} = k(0) - \frac{2}{a}\left(\frac{\Delta}{2}\right)^{1/2} \tag{14}$$

so that the variation of E with radius is of gaussian form with a radius w to the $1/e$ value of field dependent upon a, λ_0, $n(0)$, and Δ. This radius is usually called *beam radius* although fields do extend beyond. As a numerical example, if $\Delta = 0.01$, $a = 50\ \mu$m, $\lambda_0 = 1\ \mu$m, $n(0) = 1.5$, w is found to be 8.67 μm.

When w is large compared with wavelength, transverse variation of E is small in a wavelength, and the mode is nearly transverse electromagnetic with axial components of **E** and **H** negligible and the transverse components normal to each other and related by $[\mu_0/\varepsilon(0)]^{1/2}$. This fundamental mode is consequently designated TEM_{00}.

Higher-order modes may be found by returning to (9). First by expansion of ∇^2 in rectangular coordinates,

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} + k^2(0)\left[1 - 2\Delta\left(\frac{r}{a}\right)^2\right]E = 0 \tag{15}$$

the solution may be shown to be

$$E = A_{mp}H_m\left(\frac{\sqrt{2}x}{w}\right)H_p\left(\frac{\sqrt{2}y}{w}\right)e^{-(x^2+y^2)/w^2}e^{-j\beta_{mp}z} \tag{16}$$

where $H_m(\zeta)$ are Hermite polynomials of order m satisfying the differential equation[22]

[22] *M. R. Spiegel,* Mathematical Handbook, *Schaum's Outline Series, McGraw-Hill, New York, 1968.*

$$\frac{d^2H_m(\zeta)}{d\zeta^2} - 2\zeta\frac{dH_m(\zeta)}{d\zeta} + 2mH_m(\zeta) = 0 \tag{17}$$

and defined by

$$H_m(\zeta) = (-1)^m e^{\zeta^2}\frac{d^m}{d\zeta^m}e^{-\zeta^2} \tag{18}$$

In these higher-order modes, w is the same as in (13) but β_{mp} is given by

$$\beta_{mp}^2 = k^2(0) - 2\frac{\sqrt{2\Delta}}{a}k(0)(m + p + 1) \tag{19}$$

Or with Δ small,

$$\beta_{mp} \approx k(0) - \frac{\sqrt{2\Delta}}{a}(m + p + 1) \tag{20}$$

Similarly if ∇^2 is expanded in circular cylindrical coordinates, (9) is

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r}\frac{\partial E}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E}{\partial \phi^2} + \frac{\partial^2 E}{\partial z^2} + k^2(0)\left[1 - 2\Delta\left(\frac{r}{a}\right)^2\right]E = 0 \tag{21}$$

The solution may be shown to be

$$E = B_{mp}\left(\frac{\sqrt{2}r}{w}\right)^m L_p^m\left(\frac{2r^2}{w^2}\right)e^{-r^2/w^2}e^{\pm jm\phi}e^{-j\beta_{mp}z} \tag{22}$$

where $L_p^m(\xi)$ are associated Laguerre polynomials[22] of order p and degree m, satisfying the differential equation

$$\xi\frac{d^2L_p^m(\xi)}{d\xi^2} + (m + 1 - \xi)\frac{dL_p^m(\xi)}{d\xi} + (p - m)L_p^m(\xi) = 0 \tag{23}$$

and defined by

$$L_p^m(\xi) = \frac{d^m}{d\xi^m}\left[e^{\xi}\frac{d^p}{d\xi^p}(\xi^p e^{-\xi})\right] \tag{24}$$

Here also w is as in (13) but phase constant β is

$$\beta \approx k(0) - \frac{\sqrt{2\Delta}}{a}(2m + p + 1) \tag{25}$$

Since the fiber is cylindrical, it might seem that we would only be interested in the latter set, but the Hermite forms are not only simpler, but are often generated by asymmetric excitations. Each of the sets is complete, so an arbitrary distribution can be expanded in either of the two sets.

14.11 Intermode Delay and Group Velocity Dispersion

In an information transmission system, the optical wave is modulated (often digitally) and by Fourier analysis there must be a band of frequencies transmitted to represent the modulated wave. If group velocity varies over this frequency band, the envelope is distorted as the signal propagates down the fiber, as shown in Sec. 8.16. Such *group dispersion* limits the useful transmission distance for a given information rate. We shall see below that group dispersion may arise either from material properties or from the characteristics of waveguide modes themselves. In addition, signal distortion may arise in multimode guides because of the different velocities of the various modes, even at the same frequency; this effect will be considered first.

Intermode Delay In a fiber with many propagating modes, Figs. 14.9*b* and *d* show that some modes will be near cutoff with most of the energy in the cladding, some will be far from cutoff with most of the energy in the core, and others will be in between. Thus a single pulse at the input, if it excites multiple modes, may end as a multiple pulse at the output and an estimate of intermode group delay for length L of the multimode fiber is

$$\Delta T_g = \frac{L}{c}(n_1 - n_2) \tag{1}$$

For a 1% difference in n_1 and n_2, with n about 1.5, the initial pulse would yield multiple pulses spread over about 50 ns for each kilometer of propagation, severely limiting for high-data-rate, long-distance systems.

An advantage of the graded index fiber is that intermode delay is less limiting than in the above. If we calculate group velocity from the approximate expression for β of a quadratic index fiber, Eq. 14.10(20), group delay for mode (m, p) is

$$T_g = L\frac{d\beta_{mp}}{d\omega} = \frac{L}{c}\left[n(0) + \frac{\omega dn(0)}{d\omega}\right] \tag{2}$$

Thus, to this degree of approximation, group delay is the same for all modes and we do not receive multiple pulses at the output corresponding to a single pulse at the input. The more accurate expression 14.10(19) would show some intermode delay and practical differences between the real graded fiber and the ideal model also add some delay, but graded index fibers are capable of appreciably higher data rates for a given distance than the step index fibers for which (1) was an estimate.

Group Velocity Dispersion For a single mode, group delay over length L is

$$T_g = \frac{L}{v_g} = L\frac{d\beta}{d\omega} \tag{3}$$

So for a band of frequencies $\Delta\omega$ carrying the desired information, the variation of delay over this band is approximately

$$\Delta T_g \approx L \frac{d^2\beta}{d\omega^2} \Delta\omega \tag{4}$$

Thus a pulse will spread at a rate proportional to the second derivative of β with frequency. One source of such group velocity dispersion is *waveguide dispersion*, arising from the frequency dependence of β for a given guided mode, with refractive index considered independent of frequency. For step index fibers, values may be estimated from the curves plotted in Fig. 14.9*b* or calculated numerically from the implicit forms of Sec. 14.9. This contribution to dispersion is generally less important than the contributions arising from *material dispersion.*

Material dispersion arises from the variation of refractive index with frequency. From (4) with $\beta \approx \omega n/c$, this is

$$\Delta T_g \approx \frac{L}{c} \frac{d^2(n\omega)}{d\omega^2} \Delta\omega \tag{5}$$

In fiber technology it is common to express dispersion in terms of wavelength spread rather than frequency spread, with (4) written

$$\Delta T_g = LD\Delta\lambda \tag{6}$$

where D is related to $d^2\beta/d\omega^2$ by

$$D = -\frac{2\pi c}{\lambda^2} \frac{d^2\beta}{d\omega^2} \tag{7}$$

This is commonly expressed in picoseconds per kilometer of fiber length and nanometer of wavelength spread. Some representative curves of D versus wavelength[23] are shown in Fig. 14.11. Material dispersion is seen to be larger than waveguide dispersion except near the wavelength of zero dispersion, which for silica is around 1.3 μm. The zero-dispersion wavelength can be shifted by doping the material or by using multiple dielectrics in the cladding, or both. Operation near a zero-dispersion wavelength may be important in minimizing envelope distortion.

Normal dispersion ($d^2\beta/d\omega^2 > 0$ or $D < 0$) occurs for wavelengths shorter than that for zero dispersion, and anomalous dispersion ($d^2\beta/d\omega^2 < 0$ or $D > 0$) for the longer wavelengths.

For a highly coherent source such as a good laser, $\Delta\omega$ of (4) or $\Delta\lambda$ of (6) comes from the frequency spectrum of the modulated signal. For example a gaussian pulse of width τ would spread to τ' in distance L, as in Eq. 8.16(14):

$$\tau' = \tau\left[1 + \left(\frac{8L}{\tau^2} \frac{d^2\beta}{d\omega^2}\right)^2\right]^{1/2} \tag{8}$$

[23] *B. J. Ainslie and C. R. Day,* J. Light Wave Tech. ***LT-4****, 967 (1986).*

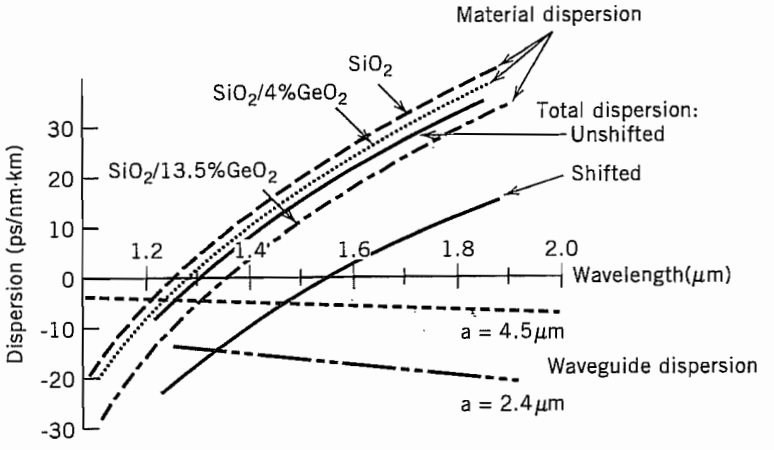


FIG. 14.11 Dispersion of single-mode fibers as function of material composition and core radius a. (After Ainslie et al.[23])

For an incoherent source such as a light-emitting diode (LED) or imperfect laser, $\Delta\omega$ or $\Delta\lambda$ may arise from the frequency variations of the source, exceeding the spectral width of the signal. In that case, the spectral width of the source is used in (4) or (6).

14.12 NONLINEAR EFFECTS IN FIBERS: SOLITONS

Silica and the related materials used in optical fibers have such a high degree of linearity that nonlinear effects might not be expected. Nevertheless, the small fiber cross sections lead to intensities high enough to produce small nonlinear interactions even with modest powers, and the low loss allows these interactions to occur over long lengths of the fiber. We shall concentrate here on the self-phase modulation effect in which refractive index changes with the intensity of the wave, which in turn changes the phase of the wave. Other important effects, some useful and some undesirable, include intermodulation products among several signals propagating in the same fiber, harmonic generation, and several parametric processes.[24] Stimulated Raman scattering arises from interaction of the optical wave with vibrational modes of the silica molecules and stimulated Brillouin scattering from interaction with acoustic waves in the fiber. Both may cause undesirable frequency shifts, but have also proven useful as optical amplifiers or tunable optical oscillators. All of these are well treated in several texts.[25]

[24] *G. P. Agrawal,* Nonlinear Fiber Optics, *Academic Press, San Diego, CA, 1989.*

[25] *See, for example, B. E. A. Saleh and M. C. Teich,* Fundamentals of Photonics, *Chap. 19, Wiley, New York, 1991; or A. Yariv,* Quantum Electronics, *3rd ed., Chap. 18, Wiley, New York, 1989.*

The important nonlinear effect in fibers, in the language of Sec. 13.7, is a $\chi^{(3)}$ effect. It is more common to express it in terms of a change in the refractive index:

$$n(E) = n_0 + \Delta n = n_0 + n_2|E|^2 \tag{1}$$

This Δn produces phase change in length L:

$$\Delta\phi = L\Delta\beta \approx \frac{L\omega}{c}\Delta n \tag{2}$$

For silica, $n_2 \approx 2.3 \times 10^{-22}$ m^2/V^2 so 10 W in a fiber of 10-μm core diameter would result in a field of about 8×10^6 V/m, producing an index change by (1) of only 1.5×10^{-8}. From (2), however, this would produce phase change of π in about 43 m for $\lambda = 1.3$ μm.

To illustrate the importance of dispersion and nonlinear index working together, we will neglect transverse variations and losses, assume small nonlinear effects, neglect dispersion terms higher than $d^2\beta/d\omega^2$, and consider only one polarization so that electric field may be treated as a scalar. Expressing the field $E(z, t)$ as a Fourier integral,

$$E(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} E(z, \omega)e^{j\omega t}\, d\omega \tag{3}$$

Each Fourier component propagates with its phase constant $\beta(\omega)$,

$$E(z, \omega) = E(0, \omega)e^{-j\beta z} \tag{4}$$

so that

$$\frac{\partial E(z, \omega)}{\partial z} = -j\beta E(z, \omega) \tag{5}$$

For the dispersive effect, we expand β in a Taylor series up to second-order terms, as in Sec. 8.16, and add the nonlinear perturbation. Since it is not rapidly varying with frequency, it is evaluated at ω_0:

$$\beta(\omega) = \beta(\omega_0) + \frac{\partial\beta}{\partial\omega}(\omega - \omega_0) + \frac{1}{2}\frac{\partial^2\beta}{\partial\omega^2}(\omega - \omega_0)^2 + \frac{\omega_0}{c}\Delta n \tag{6}$$

Now if we consider a pulse with envelope $A(z, t)$ modulating the optical carrier of angular frequency ω_0,

$$E(z, t) = A(z, t)e^{j(\omega_0 t - \beta_0 z)} \tag{7}$$

The pulse envelope $A(z, t)$ may be expressed in a Fourier integral in its base frequency, $\omega_m = \omega - \omega_0$:

$$A(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} A(z, \omega_m)e^{j\omega_m t}d\omega_m \tag{8}$$

From (7) and (8) we may show

$$E(z, \omega) = A(z, \omega_m)e^{-j\beta_0 z} \tag{9}$$

So substitution in (5) gives

$$\frac{\partial A(z,\,\omega_m)}{\partial z} - j\beta_0 A(z,\,\omega_m) = -j\beta A(z,\,\omega_m) \tag{10}$$

Now using Eq. (6) and using $d\beta/d\omega = 1/v_g$, $\beta'' = d^2\beta/d\omega^2$, and Δn from (1),

$$\frac{\partial A(z,\,\omega_m)}{\partial z} = -j\left[\frac{\omega_m}{v_g} + \frac{\beta''}{2}\,\omega_m^2 + \frac{\omega_0 n_2}{c}\,|E|^2\right] A(z,\,\omega_m) \tag{11}$$

We next make an inverse Fourier transform of (11), noting that the inverse Fourier transform of $(j\omega_m)^n A(z,\,\omega_m)$ is $\partial^n A(z,\,t)/\partial t^n$. The result is

$$\frac{\partial A(z,\,t)}{\partial z} + \frac{1}{v_g}\frac{\partial A(z,\,t)}{\partial t} = \frac{j}{2}\,\beta''\,\frac{\partial^2 A(z,\,t)}{\partial t^2} - \frac{j\omega_0}{c}\,n_2|A|^2 A(z,\,t) \tag{12}$$

Equation (12) is a nonlinear equation giving the effect of both dispersion and nonlinearity on wave propagation. (It is often normalized and transformed to moving coordinates, leading to a standard form known as the *nonlinear Schroedinger equation*, but the present form is adequate for our purposes.) If terms on the right were zero, the pulse amplitude $A(z,\,t)$ would travel without change at group velocity v_g as expected. The first term on the right represents group velocity dispersion and leads to pulse broadening as explained in Secs. 8.16 and 14.11. The second term on the right gives the effect of the nonlinearity, and the self phase modulation described qualitatively above.

Solitons An important solution of (12) is the fundamental soliton, or solitary wave, expressed by

$$A(z,\,t) = A_0\,\operatorname{sech}\left(\frac{t - z/v_g}{\tau_0}\right) e^{jz/4z_0} \tag{13}$$

where $z_0 = \tau_0^2/2\beta''$ and τ_0 is a measure of the width of a propagating pulse. Equation (13) is a solution for a particular amplitude satisfying the condition

$$A_0 = \frac{1}{\tau_0}\left(\frac{-\beta'' c}{\omega n_2}\right)^{1/2} \tag{14}$$

We see first that β'' must be negative for there to be a real solution of (12). That is, operation must be in the region of anomalous dispersion, which for silica occurs for wavelengths longer than about 1.3 μm. Then, for a particular amplitude related to pulse width and the characteristics of the fiber, the envelope (13) propagates at group velocity v_g without change of shape. The group velocity dispersion, which tends to make the pulse broaden, is compensated by the nonlinear effect, which tends to compress the pulse.

Although the preceding analysis has involved several idealizations, it has proven useful in predicting the observed solitons in actual fibers. Note that the true solitons exist only for specific power levels. Thus when propagating with attenuation there is

some pulse broadening as power is lost. But it is a graceful degradation, and experience has shown that additional power to restore the balance need not be added for many kilometers.

There are other solutions to (12), called higher-order solitons. These alternately broaden and contract, returning to the original shape in distance z_0.

Pulse Compression Another important application of (12) is in the compression of short optical pulses. In this technique, a fiber with normal dispersion ($\beta'' > 0$) may propagate a pulse, the nonlinear effect producing a frequency shift or "chirp" and the group velocity dispersion causing the pulse to broaden. With the right parameters, the frequency variation may be an almost linear function of time and the pulse may then be compressed by introducing separate anomalous dispersion, for example, by a combination of prisms and/or gratings. This is related to soliton propagation, except that there the anomalous dispersion is in the fiber, continuously compensating for the dispersion, and in the fiber pulse compressor the two functions are separate.

Extensive study of (12) requires numerical solution. Curves for optimum design of the pulse compressors have been so obtained.[26]

Gaussian Beams In Space and In Optical Resonators

14.13 Propagation of Gaussian Beams in a Homogeneous Medium

The modes of a graded index fiber, studied in Sec. 14.10, propagate with a constant pattern in the fiber if the modes are properly started. Any tendency to diffract is just countered by the distributed focusing effect of the lens-like medium. Let us imagine that the modes come to the end of the fiber, as in Fig. 14.13, and excite similar modes in space or other homogeneous material. It seems clear that they will spread by diffraction in the external region. Gaussian modes and their higher-order extensions thus become fundamental forms in homogeneous regions and may be excited in a variety of ways by lasers or other coherent sources. It is thus important to understand their properties.

The main propagation variation in the homogeneous region is still expected as e^{-jkz}, but there are other variations with z so that we may write

$$E(r, \phi, z) = \psi(r, \phi, z)e^{-jkz} \tag{1}$$

[26] *W. J. Tomlinson, R. H. Stolen, and C. V. Shank,* J. Opt. Soc. Am. B **1**, *139 (1984). Also see G. P. Agrawal,* Nonlinear Fiber Optics, *Chap. 6, Academic Press, San Diego, CA, 1989.*

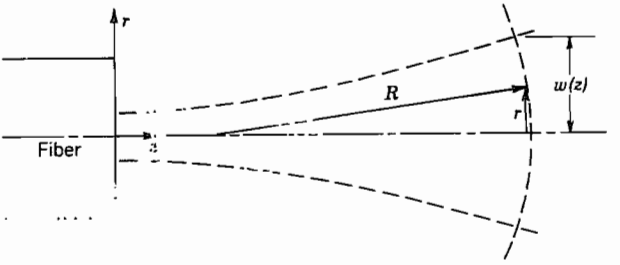


FIG. 14.1 ... si... beam, after leaving the graded index-fiber in which it is guided, will spread by di... ...effects.

We make use of Eq. 14.10(9) with $\Delta = 0$ and $k(0) = k$ and assume that ψ is slowly varying in a wavelength so that ψ'' may be neglected in comparison with ψ', where primes denote derivatives with z. There results

$$\nabla_t^2\psi - 2jk\psi' = 0 \tag{2}$$

Considering again the fundamental mode with $\partial/\partial\phi = 0$, it has been found that a useful form[27] for ψ is

$$\psi(r, z) = A \exp\left\{-j\left[P(z) + \frac{kr^2}{2q(z)}\right]\right\} \tag{3}$$

where $P(z)$ and $q(z)$ are complex. In particular,

$$\frac{1}{q(z)} = \frac{1}{R(z)} - \frac{2j}{kw^2(z)} \tag{4}$$

where $R(z)$ is a radius of curvature of the approximately spherical wavefronts (Fig. 4.13), and $w(z)$ is a measure of the gaussian beam radius at position z. Thus the imaginary part of $1/q$ gives the gaussian variation with radius and the real part gives the phase related to the curvature of wavefronts. The latter may be seen for paraxial rays with $r \ll z$ since, by Fig. 14.13,

$$kz = k\sqrt{R^2 - r^2} \approx kR\left(1 - \frac{r^2}{2R^2}\right) = k\left(R - \frac{r^2}{2R}\right) \tag{5}$$

If (3) is substituted in (2) with ∇_t^2 in cylindrical coordinates with $\partial/\partial\phi = 0$,

$$-\left(\frac{k}{q}\right)^2 r^2 - 2j\left(\frac{k}{q}\right) + \frac{k^2r^2q'}{q^2} - 2kP' = 0 \tag{6}$$

[27] *H. Kogelnik,* Appl.Opt. **4**, *1562 (1965).*

Since the equation is to hold for all r, coefficients of like powers of r must separately be zero:

$$\left(\frac{k}{q}\right)^2 = \left(\frac{k}{q}\right)^2 q' \tag{7}$$

$$P' = \frac{-j}{q} \tag{8}$$

Equation (7) has a solution of the form

$$q(z) = q_0 + z \tag{9}$$

which, substituted in (8), leads to

$$P(z) = -j \ln\left(1 + \frac{z}{q_0}\right) \tag{10}$$

At $z = 0$, curvature of the wavefronts is zero (Fig. 14.13), so from (4)

$$q_0 = \frac{jkw_0^2}{2} \tag{11}$$

with w_0 being the minimum beam radius. Substitution of (9) and (10) in (3), and the result in (1), after some manipulation with the complex quantities q and P, leads to the following expression for E:

$$E(r, z) = A \frac{w_0}{w(z)} e^{-r^2/w^2(z)} e^{-j[kz + kr^2/2R(z) - \eta(z)]} \tag{12}$$

where

$$w(z) = w_0\left[1 + \frac{z^2}{z_0^2}\right]^{1/2} \tag{13}$$

$$R(z) = z + \frac{z_0^2}{z} \tag{14}$$

$$\eta(z) = \tan^{-1}\left(\frac{z}{z_0}\right) \tag{15}$$

$$z_0 = \frac{kw_0^2}{2} \tag{16}$$

As expected, the gaussian beam spreads out while maintaining its gaussian form for each cross section. For large z, the angle of divergence is

$$\theta \approx \frac{w(z)}{z} \approx \frac{w_0}{z_0} = \frac{2w_0}{kw_0^2} = \frac{2}{kw_0} \tag{17}$$

This is consistent with a diffraction analysis for the distribution at the waist, $z = 0$, as seen in Eq. 12.14(18).

Higher-order modes exist, as with the graded index fiber. In rectangular coordinates, these are

$$E_{m,p}(x, y, z) = A_{mp} \frac{w_0}{w(z)} H_m\left(\frac{\sqrt{2}x}{w(z)}\right) H_p\left(\frac{\sqrt{2}y}{w(z)}\right) \times e^{-(x^2+y^2)/w^2(z)} e^{-j[kz + k(x^2+y^2)/2R(z) - (m+p+1)\eta(z)]} \tag{18}$$

where the Hermite polynomials H_m and H_p are as defined in Sec. 14.10, and $w(z)$, $R(z)$, and $\eta(z)$ are as in (13), (14), and (15), respectively. For circular cylindrical coordinates, it can be shown that

$$E_{mp}(r, \phi, z) = A_{mp}\left[\frac{\sqrt{2}r}{w(z)}\right] L_p^m\left(\frac{2r^2}{w^2(z)}\right) e^{-r^2/w^2(z)} e^{\pm jm\phi} e^{-j[kz + kr^2/2R(z) + (2m+p+1)\eta(z)]} \tag{19}$$

where $L_p^m(\xi)$ is the associated Laguerre polynomial defined in Sec. 14.10 and other quantities are as defined above.

Other uses and properties of the gaussian beams will be studied in the next two sections.

14.14 Transformation of Gaussian Beams by Ray Matrix

The transformation of gaussian beams passing through a combination of optical elements, including lenses, homogeneous regions, and dielectric discontinuities, has been shown[27] to follow from the q parameter of Eq. 14.13(4) and the A, B, C, D parameters of the ray matrix defined in Sec. 14.5. The gaussian beams are assumed to be coaxial with the optical elements, and the paraxial (small slope) approximations satisfied. Thus, within these approximations, the value of q_2 at the output of a region with defined parameters A, B, C, D in terms of input q_1 is

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D} \tag{1}$$

where

$$\frac{1}{q_i} = \frac{1}{R_i} - \frac{2j}{kw_i^2}, \qquad i = 1, 2 \tag{2}$$

R_i is the radius of curvature of the wavefront and w_i is the beam radius to e^{-1} value of field, as in Sec. 14.10. We will first show this for special elements and then argue the more general case.

Homogeneous Region For a homogeneous region of length z, from Eq. 14.5(2), $A = 1, B = z, C = 0$, and $D = 1$ so (1) gives

$$q_2 = q_1 + z \tag{3}$$

which is consistent with the solution found in Eq. 14.13(9). In particular, if z is measured from the waist so that $q_1 = q_0 = jkw_0^2/2$, values of R and w at plane z are

$$\frac{1}{R(z)} - \frac{2j}{kw^2(z)} = \frac{1}{q_2} = \frac{1}{z + jkw_0^2/2} \tag{4}$$

Rationalization of this gives

$$R(z) = z + \frac{k^2w_0^4}{4z} \tag{5}$$

$$w^2(z) = w_0^2\left[1 + \frac{4z^2}{k^2w_0^4}\right] \tag{6}$$

which are the forms found in Eqs. 14.13(14) and 14.13(13).

Thin Lens A thin lens would be expected to leave beam radius w unchanged, but modify radius of curvature R by $-1/f$ (Fig. 14.14*a*). Thus, the transformation expected is

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f} \tag{7}$$

Since $A = 1, B = 0, C = -1/f$, and $D = 1$ from Eq. 14.5(3), (7) is also consistent with (1).

Combination of Elements If one element follows another, as a lens following a homogeneous region (Fig. 14.14*b*), then one can extend the form (1) to following elements, so that

$$q_3 = \frac{A_2q_2 + B_2}{C_2q_2 + D_2} \tag{8}$$

But this bilinear form is consistent with matrix multiplication so that we can define total A_t, B_t, C_t, D_t parameters as in Eq. 14.5(7):

$$\begin{bmatrix} A_t & B_t \\ C_t & D_t \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \tag{9}$$

Then

$$q_3 = \frac{A_tq_1 + B_t}{C_tq_1 + D_t} \tag{10}$$

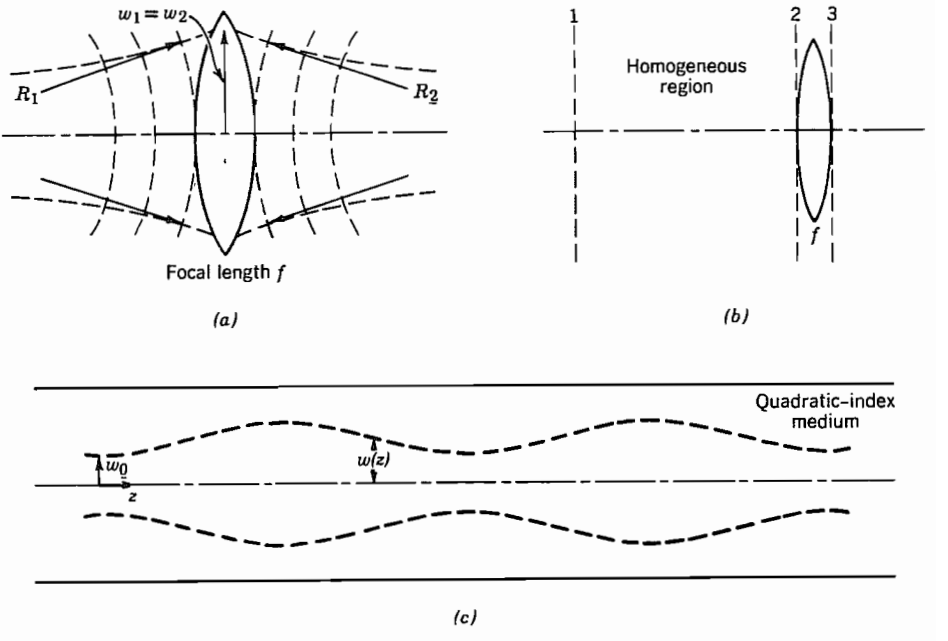


FIG. 14.14 (*a*) Thin lens of focal length f leaves gaussian beam radius unchanged but modifies radius of curvature by Eq. 14.14(7). (*b*) Combination of regions transforms gaussian beam according to overall ray matrix for the regions. (*c*) Gaussian beam in a graded-index medium has variable radius $w(z)$ if not started with zero slope at the equilibrium radius.

In addition to the lens and homogeneous region, the transformation can be shown to apply to a plane dielectric discontinuity with the gaussian beam at normal incidence (Prob. 14.14a) and, of course, to a spherical mirror since the relationship of this to a lens has already been shown. Since most useful optical systems are composed of combinations of these basic elements, the *ABCD* law for gaussian beams is very general and most useful.

Example 14.14
ROD WITH QUADRATIC INDEX VARIATION

For the rod with quadratic index variation, which has been shown to act as a continuous lens to rays and which can support gaussian modes of constant radius if excited properly, we may now utilize the transformations of this section to see what happens if the gaussian beam is not properly started. We use (1) with *ABCD* values from Eq. 14.5(5). Let $q_1 = q_0$ and $q_2 = q(z)$.

$$\frac{1}{q(z)} = \frac{C + D/q_0}{A + B/q_0} = \frac{-\kappa \sin \kappa z + (\cos \kappa z)/q_0}{\cos \kappa z + (\sin \kappa z)/\kappa q_0} \tag{11}$$

This equation tells how $R(z)$ and $w(z)$ vary with distance for a general incident R_0 and w_0. To interpret, let us assume plane wavefronts, $R_0 = \infty$ at the input. Then (11) gives

$$\frac{1}{R(z)} - \frac{2j}{kw^2(z)} = \frac{-\kappa \sin \kappa z - (2j/kw_0^2) \cos \kappa z}{\cos \kappa z - (2j/kw_0^2\kappa) \sin \kappa z} \tag{12}$$

from which beam radius is

$$w^2(z) = w_0^2\left[\cos^2 \kappa z + \left(\frac{2}{kw_0^2\kappa}\right)^2 \sin^2 \kappa z\right] \tag{13}$$

or, using trigonometric identities,

$$w^2(z) = \frac{w_0^2}{2}\left[\left(1 + \frac{4}{k^2w_0^4\kappa^2}\right) + \left(1 - \frac{4}{k^2w_0^4\kappa^2}\right) \cos 2\kappa z\right] \tag{14}$$

so that beam radius is constant at w_0 if

$$\kappa^2 = \frac{4}{k^2w_0^4} \tag{15}$$

which is consistent with the condition Eq. 14.10(13); but if this condition is not satisfied, the beam varies periodically with z, as sketched in Fig. 14.14*c*. A similar variation will occur if the beam is introduced with nonzero slope.

14.15 Gaussian Modes in Optical Resonators

It was noted in Sec. 14.6 that a typical laser resonator consists of the region between two coaxial spherical mirrors, as pictured in Fig. 14.15*a*. The active laser material may fill all or part of the internal region. Such arrangements (including the limiting case of plane mirrors) have long been used as interferometers. Schawlow and Townes[28] suggested that these would also be useful as the resonant structures for laser action. In contrast to the closed-cavity resonators studied in Chapter 10, radiation from the open sides will damp out all but the modes with primarily axial propagation. It is not entirely obvious that any low-loss modes will remain, but it seems that if the diffraction pattern from mirror M_1 is largely contained within the diameter of mirror M_2 and vice versa, diffraction losses should be small. From diffraction theory (Sec. 12.13) the condition requires that the Fresnel number N be larger than unity,

$$N = \frac{a_1a_2}{\lambda d} > 1 \tag{1}$$

where a_1 and a_2 are radii of the mirrors and d is the spacing.

The above general conclusions were verified through an important computer analysis

[28] *A. L. Schawlow and C. H. Townes,* Phys. Rev. **12**, *1940 (1958).*

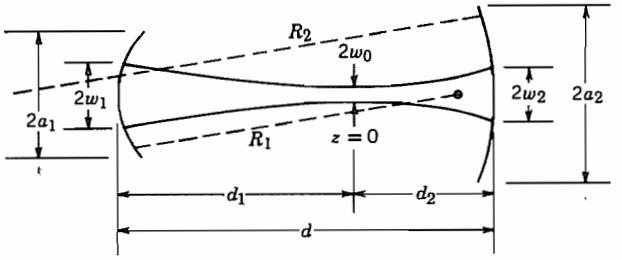


FIG. 14.15*a* Typical spherical mirror resonator with mirrors matching the curved wavefronts of the gaussian beam.

of Fox and Li.[29] They assumed an initial field distribution over one of the mirrors and, from diffraction theory, calculated the resulting pattern at the second mirror. Using this, diffraction theory was used to give the new distribution over mirror 1 and so on through many iterations. There was eventual convergence to stable field patterns showing different modal forms, and diffraction loss was low when condition (1) was satisfied.

Once it is known that stable, low-loss modes exist in the open structure, it is reasonable to look for approximate solutions of Maxwell's equation to represent these. Through the work of Pierce,[30] Boyd and Gordon,[31] Goubau and Schwering,[32] and Kogelnik,[33] there developed the understanding of the gaussian, Hermite–gaussian and Laguerre–gaussian modes discussed in the preceding sections. The remainder of this discussion will consequently apply these to the resonator of Fig. 14.15*a*. See Fig. 14.15*b* for photographs of some modal patterns.

In considering the gaussian modes within spherical mirror resonators, let us consider first the fundamental gaussian mode. It has been shown in Sec. 14.13 that these modes have nearly spherical wavefronts, so the approach is to match the curvature of these to the mirrors. Using the forms for radius of curvature given by Eq. 14.13(14), we may then write[34]

$$R_1 = d_1 + \frac{z_0^2}{d_1} \tag{2}$$

$$R_2 = d_2 + \frac{z_0^2}{d_2} \tag{3}$$

[29] *A. G. Fox and T. Li,* Bell Syst. Tech. J. ***40**, 453 (1961).*

[30] *J. R. Pierce,* Proc. Natl. Acad. Sci. USA ***47**, 1808 (1961).*

[31] *G. D. Boyd and J. P. Gordon,* Bell Syst. Tech. J. ***40**, 489 (1961).*

[32] *G. Goubau and F. Schwering,* IRE Trans. Antennas Propagation ***AP-9**, 248 (1961).*

[33] *H. Kogelnik,* Appl. Opt. ***5**, 1550 (1966).*

[34] *It is more logical from a geometric point of view to consider radii of curvature negative when center of curvature is to the right of the mirror, but from a practical point of view, the sign for a given mirror should not depend upon its orientation, so we make the practical choice and consider radii of curvature positive when the reflecting surface is concave.*

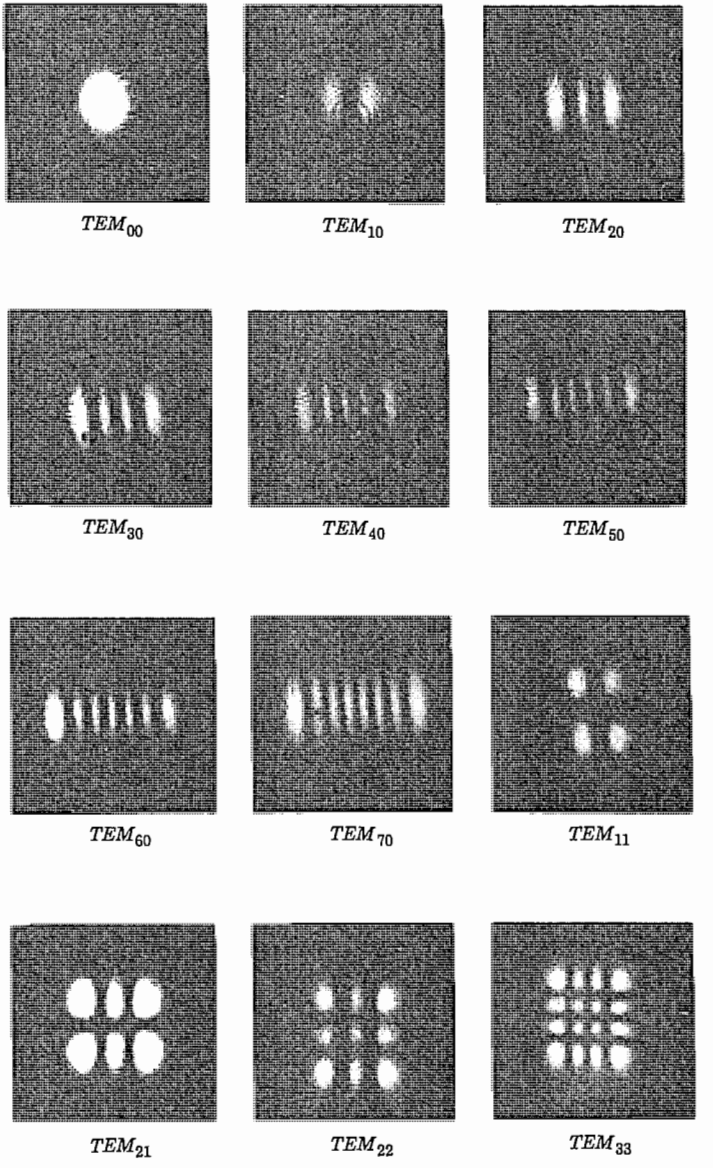


FIG. 14.15*b* Photographs of modal patterns in optical resonators. [From H. Kogelnik and W. W. Rigrod, *Proc. IRE* **50**, 220 (1962).]

where

$$z_0 = \frac{kw_0^2}{2} = \frac{\pi w_0^2 n}{\lambda_0} \tag{4}$$

and n is the refractive index. One kind of design problem is that in which wavelength and position of the beam waist (position of minimum radius) are specified, and w_0 selected to fill as much of the active region as possible. Then mirror curvatures are immediately calculable from (2)–(4). As an example, for a CO_2 laser with $\lambda_0 = 10.6$ μm, $n = 1$, $d_1 = d_2 = 0.2$ m, suppose w_0 is selected as 3 mm. Then $R_1 = R_2 = 35.8$ m.

A second type of design problem for laser resonators is one in which R_1 and R_2 are given together with wavelength and spacing d and it is desired to find position of the waist, its radius, and the beam radii at the two mirrors. Equations (2) and (3) are first solved for d_1 and d_2:

$$d_1 = \frac{R_1}{2} \pm \sqrt{\left(\frac{R_1}{2}\right)^2 - z_0^2} \tag{5}$$

$$d_2 = \frac{R_2}{2} \pm \sqrt{\left(\frac{R_2}{2}\right)^2 - z_0^2} \tag{6}$$

and

$$d_1 + d_2 = d \tag{7}$$

Substitution of (5) and (6) in (7), after some algebra, gives

$$z_0^2 = \frac{d(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{(R_1 + R_2 - 2d)^2} \tag{8}$$

With z_0 known, minimum spot size is determined from (4) and spot size at the mirrors, by using Eq. 14.13(13), to give

$$w_i = w_0\left[1 + \left(\frac{d_i}{z_0}\right)^2\right]^{1/2}, \qquad i = 1, 2 \tag{9}$$

As an example, if $d = 1$ m, $R_1 = 2$, $R_2 = 4$, $z_0^2 = 0.938$ from (8). Then $d_1 = 1.5$ and $d_2 = 0.5$ from (2) and (3). If $\lambda_0 = 1$ μm, minimum spot size w_0 is 0.554 mm from (4) and $w_1 = 0.70$ mm, $w_2 = 0.57$ mm from (9). Several special cases are worth considering individually.

Plane Mirrors If $R_1 = R_2 = \infty$, (8) shows that z_0 is infinite, and from (4) and (9), $w_0 = w_1 = \infty$. From a practical point of view, the mode spreads out until limited by the edges of the mirrors. Although there are then appreciable diffraction effects, the numerical analysis of Fox and Li demonstrated that under certain conditions such systems are still useful. Plane reflectors are most often met in solid-state lasers such as

ruby or neodymium–YAG for which the dielectric discontinuity at $r = a$ provides additional confinement.

The Concentric Resonator Consider a symmetric resonator with $R_1 = R_2 = d/2$ so that centers of curvature coincide in the midplane. From (8), $z_0 = 0$, which would make minimum spot size zero from (4) and spot size at the mirrors infinite from (9). This unrealistic situation is also limited by diffraction of the finite mirrors, yielding a finite minimum spot size at the center in a practical case. Like the plane example, diffraction losses are high, but usable in some cases.

The Symmetric Confocal Resonator If $R_1 = R_2 = d$ so that the foci of the two mirrors coincide at the midpoint of the resonator, expression (8) becomes indeterminant. We consequently look at the gaussian beam transformation of the preceding section. The ray matrix elements can be obtained from the equivalent periodic lens system and Eq. 14.5(9) with $d_1 = d_2 = d = 2f_1 = 2f_2$. Then $A = -1$, $B = 0$, $C = 0$, and $D = -1$ for a round trip. Thus a beam introduced with any beam radius and radius of curvature reproduces itself after a round trip by the transformation of Eq. 14.14(1):

$$q_1 = \frac{Aq_2 + B}{Cq_2 + D} = q_2 \tag{10}$$

This interesting case is highly degenerate. However, if we select the symmetric mode with waist at the midplane, then the proper solution of (8) gives $z_0 = d/2$ and

$$w_0 = \left(\frac{d\lambda_0}{2n}\right)^{1/2} \tag{11}$$

Since z_0 is half the cavity length in this case, it is often interpreted as the confocal parameter. Spot sizes at the mirrors are then $\sqrt{2}w_0$ for this symmetric mode.

Note that the above three special cases are on the edge of the stability diagram derived from ray concepts in Sec. 14.6. The applicability of the diagram for gaussian modes and the consequences of operation outside the "stable" range will be discussed in the following section.

14.16 Stability and Resonant Frequencies of Optical-Resonator Modes

From the ray optics point of view, it was found in Sec. 14.6 that rays are confined within some maximum radius in a periodic system if

$$\left|\frac{A + D}{2}\right| \leq 1 \tag{1}$$

where A and D are ray matrix elements for the system. For the resonator with spacing d and coaxial mirrors having curvature R_1 and R_2, this led to

$$0 \leq \left(1 - \frac{d}{R_1}\right)\left(1 - \frac{d}{R_2}\right) \leq 1 \tag{2}$$

Resonators satisfying this condition are said to be *stable*, and the diagram of Fig. 14.6*c* showing regions for which the condition is satisfied is called a *stability diagram*. To show that the condition applies to gaussian modes of the type studied in the preceding section, let A, B, C, D be the ray matrix parameters for a round trip in the resonator, and q the gaussian beam parameter defined by Eq. 14.14(2). Then q_2 at the end of a round trip is related to q_1 at the beginning by Eq. 14.14(1). But if the mode remains of the same form, $q_1 = q_2$ and

$$\frac{1}{q_1} = \frac{Cq_1 + D}{Aq_1 + B} \tag{3}$$

This leads to a quadratic equation in $1/q$ and has the solution

$$\frac{1}{q_1} = \left(\frac{D - A}{2B}\right) \pm \left[\left(\frac{D - A}{2B}\right)^2 + \frac{C}{B}\right]^{1/2} \tag{4}$$

As noted in Sec. 14.5, the ray parameters are real for the elements we have considered, and also $AD - BC = 1$, so (4) reduces to

$$\frac{1}{q_1} = \frac{1}{R_1} - \frac{2j}{kw_1^2} = \left(\frac{D - A}{2B}\right) \pm \left[\left(\frac{D + A}{2}\right)^2 - 1\right]^{1/2} \tag{5}$$

Thus for a real value of beam radius w_1, $(A + D)/2$ must not have magnitude greater than unity, so that (1) applies in general and (2) for the specific system of two mirrors separated by a homogeneous region of length d.

As has been noted, stability of an optical resonator means, in a ray optics point of view, that rays are confined within some maximum radius; from the wave point of view, it means that the field energy is essentially all within this maximum radius. Unstable resonators have the opposite property, but Siegman has shown that they may nevertheless be useful.[35] The rays, or diffraction field, escaping outside the mirror may be utilized as the output, and this has proved to be the preferred means of output coupling for many high-power lasers.

The resonant frequencies are determined by the requirement that the phase shift after a round trip is a multiple of 2π, or of π for one way. Thus from Eq 14.13(18),

$$kz_2 - kz_1 + (m + p + 1)[\eta(z_2) - \eta(z_1)] = l\pi \tag{6}$$

where m, p, and l are integers. Using Eq. 14.13(15) and the sign conventions of the preceding section, $z_2 = d_2$ and $z_1 = -d_1$,

$$k(d_1 + d_2) + (m + p + 1)\left[\tan^{-1}\left(\frac{d_2}{z_0}\right) + \tan^{-1}\left(\frac{d_1}{z_0}\right)\right] = l\pi \tag{7}$$

Longitudinal modes are defined by the integer l. Thus, the frequency spacing between longitudinal modes is determined by maintaining m and p constant and changing l by

[35] *A. E. Siegman, Proc. IEEE* ***53****, 277(1965); also see A. E. Siegman, Lasers, Chap. 22, University Science Books, Mill Valley, CA, 1986.*

unity:

$$\frac{(\omega_{l+1} - \omega_l)n}{c}(d_1 + d_2) = \frac{2\pi\,\Delta f_l n d}{c} = \pi \tag{8}$$

or

$$\Delta f_l = \frac{c}{2nd} \tag{9}$$

As example, if $d = 0.10$ m and $n = 1.5$, $\Delta f_l \approx 1$ GHz.

Transverse modes are defined by the integers m and p. Frequency spacing between these is determined by maintaining l constant and changing either m or p by unity:

$$\frac{(\omega_{m+1} - \omega_m)n(d_1 + d_2)}{c} + \tan^{-1}\left(\frac{d_2}{z_0}\right) + \tan^{-1}\left(\frac{d_1}{z_0}\right) = 0 \tag{10}$$

or

$$\Delta f_t = \frac{(\omega_{m+1} - \omega_m)}{2\pi} = -\frac{c}{2\pi nd}\left[\tan^{-1}\left(\frac{d_2}{z_0}\right) + \tan^{-1}\left(\frac{d_1}{z_0}\right)\right] \tag{11}$$

Note that as the mirrors approach planes, $z_0 \to \infty$ as shown in the preceding section, and transverse modes become very close together. For a concentric resonator, $z_0 \to 0$ and $\Delta f_t \to c/2nd$. That is, transverse mode spacing is the same as longitudinal mode spacing for concentric resonators. For a symmetric confocal resonator, $d_1 = d_2 = z_0$, $\Delta f_t = c/4nd$ showing that transverse mode spacing is half the longitudinal mode spacing in this case.

Basis for Optical Information Processing

14.17 Fourier Transforming Properties of Lenses

Combinations of lenses have been found useful for information processing by optical means because of the Fourier transforming properties of these systems[36] which follow from diffraction theory (Sec. 12.13). Here we will use the Fresnel form in the paraxial approximation, Eq. 12.13(10),

$$\psi(x, y, z) = \frac{je^{-jkL}}{\lambda L}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi(x', y', z')\exp\left\{-\frac{jk}{2L}[(x - x')^2 + (y - y')^2]\right\}dx'\,dy' \tag{1}$$

[36] *J. W. Goodman,* Introduction to Fourier Optics, *McGraw-Hill, New York, 1968; also see B. E. A. Saleh and M. C. Teich,* Fundamentals of Photonics, *Chap. 4, Wiley, New York, 1991.*

where $\psi(x', y', z')$ is the given distribution of some scalar field component (such as E_x) in the plane z', and $\psi(x, y, z)$ is the resulting distribution of that component in a plane distance L away, where $L = |z - z'|$ (Fig. 14.17*a*).

First note that for L large enough, terms kx'^2/L and ky'^2/L in (1) are negligible,

$$\psi(x, y, z) = \frac{je^{-jk[L+(x^2+y^2)/2L]}}{\lambda L} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(x', y', z')\exp\left[\frac{jk}{L}(xx' + yy')\right] dx'\, dy' \tag{2}$$

which is the Fraunhofer diffraction studied earlier. The integral in (2) is a two-dimensional Fourier transform (Sec. 7.11). The phase factor in front is often unimportant in optical problems, but note that it can be maintained constant if ψ is observed on

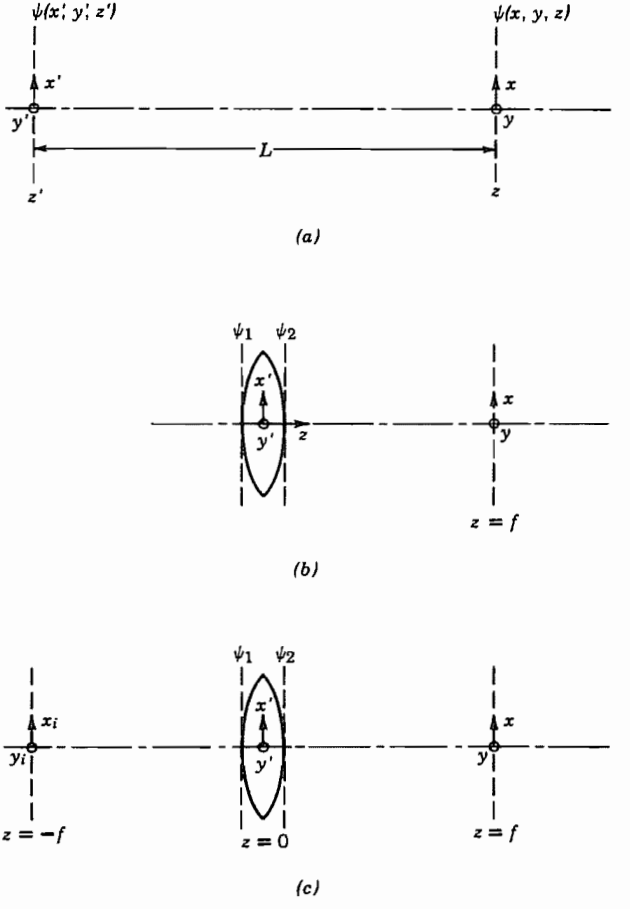


FIG. 14.17 (*a*) Diffraction of a scalar field component between planes z' and z. (*b*) Diffraction with a lens introduced to correct spherical wavefronts. (*c*) Diffraction from input focal plane of lens to output focal plane, yielding Fourier transforming properties.

a spherical surface, $L + (x^2 + y^2)/L = \text{constant}$. Thus with these understandings of the phase term, the diffracted distribution in the Fraunhofer approximation can be considered the Fourier transform of the given distribution.

The next point to be made is that the restriction to large distances can be eliminated if a lens is placed just beyond the input plane to correct for the phase factors involving x'^2 and y'^2. An ideal thin lens does not change amplitude, but modifies phase quadratically with distance from the axis, as in Eq. 14.2(19). Thus referring to Fig. 14.17*b*,

$$\psi_2(x', y', 0+) = \psi_1(x', y', 0-)e^{-(jk/2f)[a^2 - x'^2 - y'^2]} \tag{3}$$

where a is the outer radius of the lens and f is focal distance. We assume that a is large enough to intercept all of the important optical field so that we can continue to integrate over infinite limits. If (3) is then substituted in (1), the phase terms in x'^2 and y'^2 are canceled, provided $L = f$. Thus at the focal plane of the lens, within the Fresnel approximation,

$$\begin{aligned}\psi(x, y, f) = {} & \frac{je^{-jk[f + (x^2 + y^2 + a^2)/2f]}}{\lambda f} \\ & \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_1(x', y', 0-)\exp\left[\frac{jk}{f}(xx' + yy')\right] dx'\, dy'\end{aligned} \tag{4}$$

So Fourier transforming properties are again seen, with the multiplying phase factor again a function of x and y. As above, this phase factor can be made constant by observing ψ on a nearly spherical surface passing through the focus.

Finally, the variable phase term of (4) can be eliminated if one expresses the field $\psi_1(x', y', 0)$ just before the lens as a transform of its value at the focal plane before the lens at $z = -f$ (see Fig. 14.17*c*). Transformation from the input at $z = -f$ to the plane just in front of the lens is by the Fresnel diffraction integral (1):

$$\begin{aligned}\psi_1(x', y', 0-) = {} & \frac{je^{-jkf}}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i(x_i, y_i, -f) \\ & \times \exp\left\{-\frac{jk}{2f}[(x' - x_i)^2 + (y' - y_i)^2]\right\} dx_i\, dy_i\end{aligned} \tag{5}$$

When (5) is substituted in (4) there results

$$\begin{aligned}\psi_0(x, y, f) = {} & \frac{je^{-jk(2f + a^2/2f)}}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i(x_i, y_i, -f) \\ & \times \exp\left[\frac{jk}{f}(xx_i + yy_i)\right] dx_i\, dy_i\end{aligned} \tag{6}$$

Equation (6) made use of the integral

$$\int_{-\infty}^{\infty} e^{j(\alpha x^2 + \beta x)}\, dx = \sqrt{\frac{j\pi}{\alpha}}\, e^{-j(\beta^2/4\alpha)} \tag{7}$$

Here the phase factor before the integral represents a constant phase delay so that the distribution in the output focal plane is directly proportional to the Fourier transform of field distribution in the input focal plane. This may be seen more clearly by defining

$$p = x_i\sqrt{\frac{k}{f}}, \qquad q = y_i\sqrt{\frac{k}{f}}, \qquad u = x\sqrt{\frac{k}{f}}, \qquad v = y\sqrt{\frac{k}{f}}$$
$$f(u, v) = j\psi_0(x, y, f)e^{jk(2f+a^2/f)}, \qquad g(p, q) = \psi_i(x_i, y_i, -f) \tag{8}$$

in which case (6) is in the exact form of the two-dimensional Fourier integral:

$$f(u, v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(p, q)e^{j(up+vq)}\,dp\,dq \tag{9}$$

This important property has made possible Fourier transforming, convolution, correlation and filtering of spatial functions by lenses and combinations of lenses.[36,37]

The Fourier transforming properties noted above also give us some physical insight as to why gaussian modal forms are found for spherical mirror resonators or periodic lens systems in preceding sections. A gaussian function is known to Fourier transform to a gaussian, so this form would be expected to persist as a mode in the periodic lens system. The resonator with two spherical mirrors has already been shown to be equivalent to the periodic lens system. Hermite–gaussian and Laguerre–gaussian forms likewise transform into themselves, so these would be expected to be the higher-order mode forms. In fact an alternate way of deriving such modes is by setting up the Fresnel diffraction integrals and requiring that the function repeat in a period. Solution of the resulting integral equation leads to modal solutions of the form already found.[4]

14.18 Spatial Filtering

A two-dimensional information function may be impressed upon an optical wave by passing the wave through a screen of the desired transmission property, $T(x, y)$. The arrangement might be as in Fig. 14.18*a*: laser light is passed through a pinhole at the

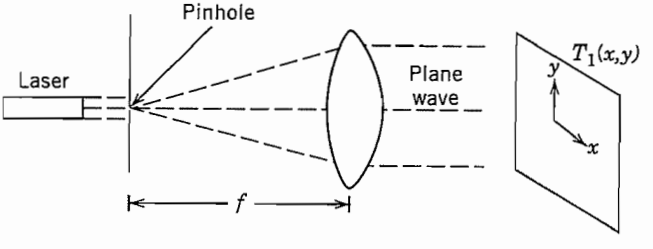


Fig. 14.18*a* Point source at focal point of lens produces a plane wave; information is put on this wave by screen with transmission $T_1(x, y)$.

[37] *J. L. Horner (Ed.), Optical Signal Processing, Academic Press, San Diego, CA, 1987.*

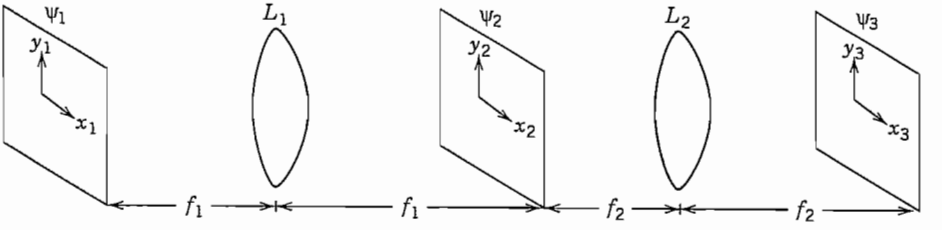


FIG. 14.18*b* Lens L_1 Fourier transforms ψ_1 to ψ_2. T_2 performs spatial filtering. The result is inverse transformed to new function ψ_3.

focus of a lens, transforming it into a plane wave exiting the lens, followed by the transmission screen. T may be complex so that phase as well as amplitude modulation may be impressed upon the wave. If desired, this two-dimensional information function, could be modified by directly passing it through other transmission screens. This would be analogous to operating on a signal which is a function of time directly in the time domain. But as with such signals in time, it may be preferable to filter in the frequency domain. So with our spatial information function, filtering in the Fourier plane to modify the spatial frequencies provides an alternate, and sometimes preferable, way of transforming the information.

A simple configuration which illustrates spatial filtering is the two-lens arrangement of Fig. 14.18*b*. A uniform plane wave, obtained as in Fig. 14.18*a*, has the desired input function $\psi_1(x_1, y_1)$ impressed upon it at the input focal plane of lens L_1. Its Fourier transform then appears at the output focal plane, as explained in the preceding section. A screen with transmission properties $T_2(x_2, y_2)$ may be placed at this plane to modify the desired spatial frequencies. Lens L_2 then produces the inverse Fourier transform at the exit focal plane of L_2, resulting in a modified two-dimensional function.

The spatial filters may modify amplitude only, phase only, or a combination. The simplest to make are binary filters for which holes are cut in an opaque screen so that there is either full or zero transmission for any given portion of the screen. For many purposes it is desirable to have the spatial filter change with time. *Spatial light modulators* are used for such purposes.

Note that the arrangement of Fig. 14.18*b*, with no spatial filter in the Fourier plane, reproduces the input image, inverted. It is also changed in size by the ratio f_2/f_1.

Example 14.18
SIMPLE BINARY FILTER

As a specific example, let us consider the input optical function as the rectangular function of Fig. 14.18*c*, resulting from a grating placed in the path of the plane wave. For simplicity, consider only variations in x. (Cylindrical lenses are required.) Since

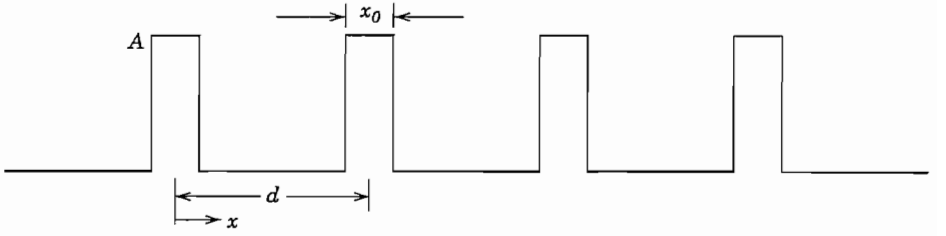


FIG. 14.18c Function ψ_1 taken as a periodic pattern of dark and light.

this is a periodic function, it may be written as the Fourier series:

$$\psi_1(x_1) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi n x_1/d} \tag{1}$$

$$C_n = \frac{A}{\pi n} \sin \frac{n\pi x_0}{d} \tag{2}$$

Use of Eq. 14.17(6) to give ψ_2 in the Fourier plane yields

$$\psi_2(x_2) = K_1 \int_{-\infty}^{\infty} \sum_n C_n e^{jx_1(2\pi n/d + kx_2/f_1)}\, dx_1 \tag{3}$$

where K_1 is the complex constant before the integrand in Eq. 14.17(6).

Since the impulse or delta function $\delta(\zeta)$ is given by the integral,

$$\delta(\zeta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\zeta x}\, dx \tag{4}$$

eq. (3) yields

$$\psi_2(x_2) = 2\pi K_1 \sum_{n=-\infty}^{\infty} C_n \delta\left(\frac{kx_2}{f_1} + \frac{2\pi n}{d}\right) \tag{5}$$

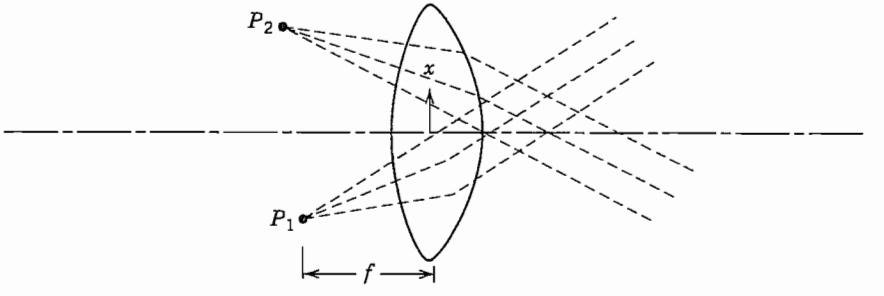


FIG. 14.18d Lens focuses rays from lines P_1 and P_2 to plane waves at an angle. Interference then produces sinusoidal variation with x.

The pattern in the Fourier plane thus consists of a series of lines, spaced along the axis with spacing $2\pi f_1/kd$ or $(\lambda/d)f_1$. Amplitude in each line decreases with increasing distance from the axis. It is then straightforward in principle to design a spatial filter to modify each line (spatial harmonic) to produce a desired output.

Suppose first that a binary filter is used, blocking all lines except that of zero order on the axis. This clearly transforms to a plane wave at the output of L_2, eliminating all information about the grating. But if the filter blocks all but the upper and lower pairs of nth order ($n = \pm|n|$), a cosine wave output of period $f_1 d/nf_2$ is produced. A y-directed line off the lens axis in the focal plane produces a plane wave propagating at an angle to the axis, and the two on opposite sides produce one wave upward and one downward so that the interference pattern makes the cosine wave referred to (Fig. 14.18*d*). As more and more lines are selected to pass, the output pattern approaches the original grating function, modified in period by f_2/f_1.

14.19 The Principle of Holography

In recording optical information, film and most other light-sensitive media react to intensity of the wave and not to phase. Dennis Gabor, in 1948, recognized that phase information could be captured by interfering the scattered wave from an object with a reference wave of the same frequency.[38] The concept became practical and important with the development of the laser. Leith and Upatnieks especially developed many variations of this concept.[39] Gabor first called the technique "wavefront construction," then later holography for "total recording."

Many variations are possible,[40] but to illustrate the concept, consider the arrangement of Fig. 14.19*a*. A coherent uniform plane wave, obtained for example as in Fig. 14.18*a*, is divided into two parts. One part impinges upon a desired object and the scattered wave contains amplitude and phase information about the object. This information arrives at the two-dimensional recording plane as the complex quantity $A(x, y)$. The second reference portion falls upon the recording plane as $R(x, y)$. Object and reference waves at the recording plane may be written in terms of amplitudes and phases:

$$A(x, y) = A_0(x, y)e^{-j\phi(x,y)} \tag{1}$$

$$R(x, y) = R_0(x, y)e^{-j\psi(x, y)} \tag{2}$$

[38] *D. Gabor,* Nature **161**, *777 (1948).*

[39] *E. N. Leith and J. Upatnieks,* J. Opt. Soc. Am. **52**, *1123 (1963);* **53**, *1377 (1963);* **54**, *579 (1964);* **54**, *1295 (1964).*

[40] *J. E. Kasper,* Complete Book of Holograms: How They Work and How to Make Them, *Wiley, New York, 1987.*

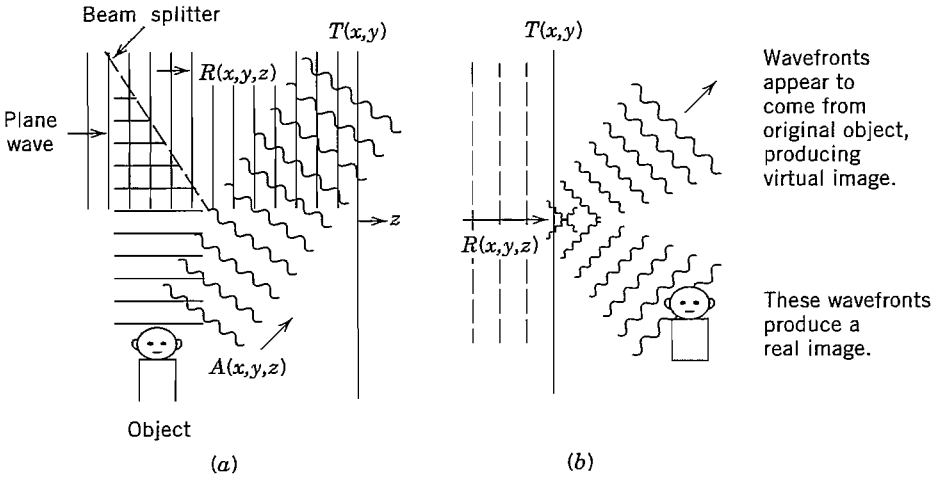


FIG. 14.19 (*a*) Plane wave splits into two parts, one part falling on the object and the other part serving as a reference wave $R(x, y, z)$. Scattered wave from object $A(x, y, z)$, interferes with reference on recording plane to make hologram $T(x, y)$. (*b*) Irradiation of hologram $T(x, y)$ by reference wave produces wavefronts appearing to come from original object, wavefronts producing a real image, and portions of reference beam (not shown).

The recording medium, when developed, is assumed to give a transmissivity proportional to total intensity. This property is well approximated by useful recording media over certain ranges. Transmissivity is then

$$\begin{aligned} T(x, y) &= K(A + R)(A^* + R^*) \\ &= K\{A_0^2 + R_0^2 + A_0R_0[e^{-j(\phi-\psi)} + e^{j(\phi-\psi)}]\} \end{aligned} \tag{3}$$

where K is a constant of proportionality. It is clear from the above that phase information $(\phi - \psi)$ is recorded. We call this recording the *hologram.*

To illustrate the reconstruction of the wavefront $A(x, y)$ we imagine the object removed and the hologram illuminated by a replica of the original reference wave $R(x, y)$. The wave transmitted through the film is then

$$\begin{aligned} B(x, y) &= T(x, y)R_0(x, y)e^{-j\psi(x, y)} \\ &= K\{(A_0^2 + R_0^2)R_0e^{-j\psi} + A_0R_0^2[e^{-j\phi} + e^{j(\phi-2\psi)}]\} \end{aligned} \tag{4}$$

The term $(A_0^2 + R_0^2)R_0e^{-j\psi}$ has the phase of the reference wave across the plane and produces a continuation of that wave to the right. The next term is proportional to the object wave and produces a wave continuing to the right as though it were coming from the original object, even though that object is no longer there. Thus it yields a virtual image of the object (Fig. 14.19*b*). The last term is proportional to the conjugate of $A(x, y)$ and continues to the right in such a way as to form a real image of the object. In order that the three transmitted waves not overlap at the viewing position and lead to confusion, relative angles between object and reference waves must be properly chosen in making the original hologram.

Example 14.19a
INTERFERENCE OF PLANE WAVES

The holographic principle will first be discussed through the simple example of plane wave interferences. Let the object wave be a uniform plane wave traveling at angle θ from the normal to the recording plane, and the reference wave a plane wave propagating in the direction of the normal (Fig. 14.19*c*). We neglect variations with y. Object and reference waves as functions of x and z are then

$$A(x, z) = A_0 e^{-jk(x \sin \theta + z \cos \theta)} \tag{5}$$

$$R(z) = R_0 e^{-j(\psi_0 + kz)} \tag{6}$$

At the recording plane $z = 0$ with assumptions as above, transmissivity from (3) is

$$T(x) = K\{(A_0^2 + R_0^2) + A_0 R_0[e^{-j(kx \sin \theta - \psi_0)} + e^{j(kx \sin \theta - \psi_0)}]\} \tag{7}$$

The original object wave is now removed and the film irradiated with a replica of the reference wave (6), resulting in a transmitted wave

$$B(x) = K\{(A_0^2 + R_0^2)R_0 e^{-j\psi_0} + A_0 R_0^2[e^{-jkx \sin \theta} + e^{j(kx \sin \theta - 2\psi_0)}]\} \tag{8}$$

The first term is uniform in phase over the plane and continues to the right as a part of the reference wave R_0. The second term has a phase variation in x proper to excite a wave to the right proportional to the continuation of (5). The last term has the conjugate phase and excites a plane wave traveling with angle θ from the normal. All these are illustrated in Fig. 14.19*d*.

In this example we may consider that the interference of the reference and object plane waves has produced a grating in the film. When irradiated, the waves at angles $\pm\theta$ correspond to the first Bragg orders of this diffraction grating (Sec. 6.10).

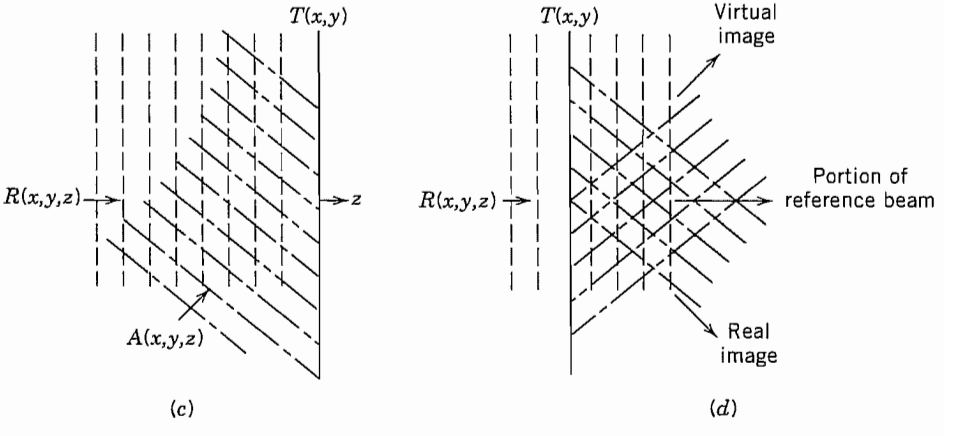
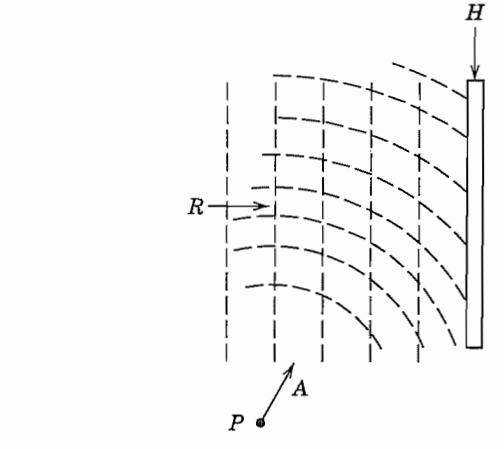


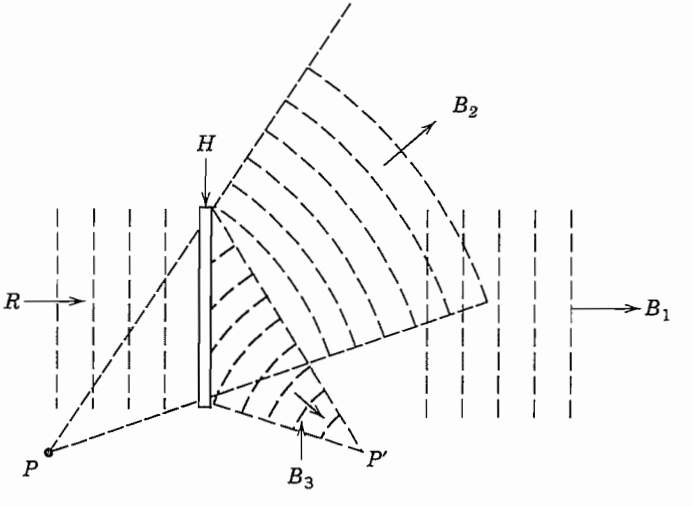
FIG. 14.19 (*c*) Interference pattern between two plane waves R and A recorded as hologram $T(x,y)$. (*d*) Irradiation of T by reference wave R results in three plane waves, one of which appears to be a continuation of A, though that is now removed.

Example 14.19b
HOLOGRAM OF A POINT SOURCE

As a second example consider the hologram of a point source. This is especially interesting since the scattered wave from a general object may be considered as made up of a superposition of waves from such point sources. Consider Fig. 14.19*e* with point



(*e*)



(*f*)

FIG. 14.19 (*e*) Interference between spherical waves from P and reference plane wave R recorded as hologram H. (*f*) Irradiation of H by reference R produces plane wave B_1; diverging spherical wave B_2, which appears to come from P (now removed); and converging spherical wave B_3, converging to real image P' of P.

source P located at (x_0, y_0, z_0). The idealized spherical wave (Prob. 14.19c) emanating from the source may be written as

$$A(x, y, z) = \frac{A_0}{r(x, y, z)} e^{-jkr(x,y,z)} \tag{9}$$

where

$$r(x, y, z) = [(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{1/2} \tag{10}$$

After interfering with a plane reference wave of the form (6), and recording with linearity assumed as in the previous examples, transmissivity at $z = 0$ is

$$T(x, y) = K\left\{\left(\frac{A_0^2}{r^2} + R_0^2\right) + \frac{A_0 R_0}{r}\left[e^{j(\psi_0 - kr(x,y,z_0)} + e^{-j(\psi_0 - kr(x,y,z_0)}\right]\right\} \tag{11}$$

The original point source is removed and the hologram irradiated with reference wave (6), yielding a transmitted wave

$$B(x, y) = B_1 + B_2 + B_3 \tag{12}$$

where

$$B_1(x, y) = KR_0\left(\frac{A_0^2}{r^2} + R_0^2\right)\exp[-j(kz_0 + \psi)] \tag{13}$$

$$B_2(x, y) = \frac{KR_0^2 A_0}{r}\exp\{-jk[(x - x_0)^2 + (y - y_0)^2 + z_0^2]^{1/2}\} \tag{14}$$

$$B_3(x, y) = \frac{KR_0^2 A_0}{r}\exp\{-2j\psi_0 + jk[(x - x_0)^2 + (y - y_0)^2 + z_0^2]^{1/2}\} \tag{15}$$

Following previous arguments, B_2 continues to the right as a diverging spherical wave as though emanating from the original source P (now missing), and B_3 as a converging spherical wave resulting in a real image of P at P'. B_1 is uniform in phase over the plane but varying in amplitude through the dependence on r. This can be made small and the continuation of B_1 to the right is approximately a plane wave, with B_2 the desired reconstruction.

PROBLEMS

14.2a Obtain from Eq. 14.2(4) the approximate spread of focal length Δf from $r = 0$ to r_{max} when $r_{max}/R \ll 1$. For a spherical mirror of radius of curvature 1 m, used with a laser of wavelength $\lambda_0 = 1$ μm, give the maximum radius of rays from the axis if spread in focal length is not to be more than a wavelength.

14.2b The f-number of a lens is defined as the ratio of focal length to diameter. (Here we will denote by F to avoid confusion with focal length.) Give the spread of focal

length in terms of F when $r_{max} \ll R$ for the spherical mirror. Explain why "stopping down" a lens to higher values of F should increase sharpness of an image.

14.2c Carry out the derivation corresponding to that of Ex. 14.2c for a diverging thin lens with plane surface at $z = 0$ and a parabolic surface $z = d + r^2/4g$, $0 < r < a$, (i) by ray analysis and (ii) by phase considerations.

14.2d For the doubly convex lens of Fig. 14.2*e*, derive the equation for focal length, Eq. 14.2 (21), by considering refraction at the surface.

14.2e Consider the imaging for a spherical mirror by ray reflection. That is, for an object point on the axis distance d_1 from a mirror with radius of curvature R, find the image distance d_2 from the mirror at which a ray reflected at radius r crosses the axis. Under what approximations does Eq. 14.2(24) apply?

14.3a For a uniform plane wave, polarized with e_z only, propagating at angle α from the x axis with $\cos\gamma = 0$, give all quantities $\mathbf{e}$, $\mathbf{h}$, S, u_e, and u_h used in the general formulation. Material has refractive index n.

14.3b We will find later that many useful beams have gaussian forms. Assume one with form in the transverse plane, $\mathbf{e} = \hat{\mathbf{x}} \exp[-(x^2 + y^2)/w_0^2]$ propagating substantially in the z direction, $S \approx nz$. Compare the neglected term on the right side of Eq. 14.3(5) with the term $\nabla S \times \mathbf{e}$ in magnitude and direction and state under what condition it is negligible.

14.3c Utilizing Eq. 14.3(12) for the eikonal of a plane wave, show that $\mathbf{E}$ and $\mathbf{H}$ are related as in an arbitrarily propagating plane wave in a homogeneous medium.

14.4a Find the required form of refractive index variation to maintain a ray in a circular path of constant radius R.

14.4b Very low absorptions can be measured by using the thermal lens effect resulting from passing a beam through a cell containing the sample, as in Ex. 14.4a. It has been shown [J. R. Whinnery, *Acc. Chem. Res.* 7, 225 (1974)] that the heating effect results, in steady state, in a quadratic index variation near the axis as in Eq. 14.4(11), with

$$\Delta = -\frac{\alpha P(dn/dT)}{4\pi k n_0}$$

where α is absorption coefficient, P power in the laser beam, dn/dT the variation of refractive index with temperature, k the thermal conductivity, and n_0 the refractive index on the axis. Find the expected focal length of a cell 1 cm long filled with carbon disulfide ($n_0 = 1.63$, $dn/dT = 7.9 \times 10^{-4}\,\mathrm{K}^{-1}$, $k = 6.82 \times 10^{-4}$ J/cm-s-K, $\alpha = 6 \times 10^{-4}\,\mathrm{cm}^{-1}$) for a laser beam with 0.5 W power and a beam radius $a = 0.5$ mm.

14.4c Derive Eq. 14.4(4) from Eq. 14.4(10).

14.4d A typical graded index fiber used in an optical communications system with moderate data rates has $n_0 = 1.5$, $\Delta = 0.005$, and $a = 25\ \mu$m. For a ray crossing the axis at an angle r_0' find the distance at which it returns to the axis. What is the maximum angle of crossing to maintain r_{max} less than a?

14.5a A ray passes from a medium of index n_1 into a slab of thickness d and index n_2, then exits into n_1. Compare radius and slope at the output from the ray matrix and an exact Snell's law calculation if $n_1 = 1$, $n_2 = 2$, $d = 1$ cm, and angle of input ray is (i) 20 degrees and (ii) 40 degrees from the perpendicular to the slab.

14.5b Slabs of dielectric with index n_1 and thickness d_1 alternate with slabs having index n_2 and thickness d_2. Using the paraxial (small-angle) approximation, find the ray matrix for one period of this combination, then two periods, and induce the result for N periods. Interpret the result.

14.5c Derive Eq. 14.5(5) with the approximations stated.

14.5d For the doubly convex lens of Fig. 14.2*e*, derive the expression for focal length, Eq. 14.2(21), by considering the tandem combination of two spherical dielectric interfaces, with proper attention to sign.

14.6a An alternate form of the solution in Eq. 14.6(7) to the difference equation for the periodic lens system is

$$r_m = C_1 \cos m\theta + C_2 \sin m\theta$$

Relate C_1 and C_2 to the values of initial radius and slope of the rays. The ray matrix for one period is assumed known.

14.6b A periodic lens system has $d_1 = d_2$ and $f_1 = f_2 = d_1/5$. Check stability and sketch the ray paths through a few lenses starting with zero slope and $r = 0.1f_1$ at the first lens.

14.6c Repeat Prob. 14.6b for $d_1 = d_2$, $f_1 = f_2 = d_1$.

14.6d Repeat Prob. 14.6b for the case of alternate converging and diverging lenses, $d_1 = d_2$, $f_1 = -f_2 = d_1$.

14.7a Explain why sinusoidal rather than hyperbolic solutions are required in the film of Fig. 14.7*a* for guided modes with exponential decay away from the film in cover and substrate regions.

14.7b In the example of the glass film guide given in Sec. 14.7, thickness d is increased to 2 μm. Use Fig. 14.7*b* to find n_{eff} for all of the guided TE and TM modes.

14.7c In a certain semiconductor laser, the active region consists of a layer of GaAs 0.3 μm thick, with $n_2 = 3.35$. This is surrounded above and below with GaAlAs with $n_1 = n_3 = 3.23$. Use Fig. 14.7*b* to find the modes that can be guided by this active layer, and their values of n_{eff}. Wavelength λ_0 is 0.85 μm.

14.7d* To illustrate the graphical method of solution for Eq. 14.7(7), consider the symmetric case with $n_1 = n_3$ so that $q = p$. Show that (5) is then satisfied either by $pd = hd \tan(hd/2)$ or $pd = -hd \cot(hd/2)$. Sketch some curves of pd versus hd from these equations. Then show from (4) that the loci for constant v [defined by (6)] in the pd versus hd plane are circles and sketch these for $v = 2, 5, 8$. Solutions (modes) are given by points of intersection between the circles and the first sets of curves plotted. (Note that p must be positive for guided modes.) Check the modes predicted for the three values of v from Fig. 14.7*b*.

14.7e* Find the approximation to Eq. 14.7(7) when $n_3 - n_1 \gg n_2 - n_3$, as is the case for many guides with dielectric substrates but air above. Show that a graphical solution similar to that of Prob. 14.7d may be utilized for this case also.

14.8a A guide is fabricated as in Fig. 14.8*a* with the guiding material GaAs with $n_1 = 3.59$, width 1.8 μm, and depth 0.3 μm. The surrounding material is AlGaAs with $n_3 = 3.385$ except for the top surface, which is protected by silica with $n_2 = 1.5$. Free-space wavelength is 0.85 μm. Find approximate values of n_{eff} for this guide by the effective index method, for the lowest-order mode.

14.8b We are to design a dielectric guide by in-diffusing titanium into lithium niobate, much as in Fig. 14.8*b*. Assume index change Δn and diffusion depth d as in H. Naitoh *et al.*, *Appl. Opt.* **16**, 2546 (1977): $\Delta n = 0.078t^{-0.85}$ and $d = 0.43t$, where t is diffusion time in hours and d is in μm.

(i) As a first approximation, use the model of a step index guide as in Fig. 14.7*a* with air above and estimate the maximum diffusion time for single-mode operation. Take $\lambda = 0.633$ μm and the refractive index of lithium niobate as 2.05.

(ii) For operation with a v value 0.75 of that above, give diffusion time and effective index as based on this model.

(iii) Now, taking into account lateral confinement with a width 1.5 μm, use effective index method to give the next correction to n_{eff}.

14.9a For a step index fiber of silica ($n = 1.500$) cladding on a glass ($n = 1.505$) core of 6.0-μm diameter, find the cutoff wavelengths of the two lowest axially symmetric TE modes.

14.9b A fiber with $a = 10$ μm has $n_1 = 1.51$ and $n_2 = 1.50$. Use Fig. 14.9*d* to estimate the number of modes that may propagate at $\lambda_0 = 1.3$ μm. Give the values of n_{eff} for the highest and lowest order of these modes. What radius would be required for only one propagating mode?

14.10a For a fiber with quadratic index variation having $n(0) = 1.5$, $a = 10$ μm, and $\lambda_0 = 1$ μm, what Δ is required for a beam radius $w = 3$ μm? Find the percentage difference of phase velocity from that of a plane wave in material with index $n(0)$ for a fundamental mode and Hermite–gaussian modes of order m, p.

14.10b For the numerical values given in Prob. 14.10a, estimate the term on the right of Eq. 14.10(6) in comparison with the second term on the left.

14.10c Using Eq. 14.10(17) show that (16) does satisfy (15) with the conditions on w and β_{mp} as given.

14.10d* Using Eq. 14.10(23) show that (22) satisfies (21) with the conditions on w and β_{mp} as given.

14.11a A multimode fiber has $n_1 = 1.51$ and $n_2 = 1.50$. About what maximum data rate could be used with this fiber over a distance of 7 km?

14.11b If one uses the expression 14.10(19) rather than the approximation (20), group velocity does depend upon mode order. Find group velocity from (19) and the intermode dispersion based upon this result. (In a practical case, the boundary effect at $r = a$ and the departure from the ideal profile may be more important, but this is at least one component of dispersion.)

14.11c For a GaAlAs laser source at $\lambda_0 = 0.85$ μm with a single-mode fiber, material dispersion is likely to be dominant. If $d^2n/d\lambda^2 \approx 3.2 \times 10^{10}\ \text{m}^{-2}$ at this wavelength, what approximate data rate is usable over a length of 20 km if the laser is (i) an ideal coherent source? (ii) Has a spectral width $\Delta\lambda_0 = 0.2$ nm?

14.11d The wavelength of minimum attenuation for silica is about 1.55 μm. Use Fig. 14.11 to estimate the usable distance for a data rate of 1 Gb/s, using an unshifted fiber at this wavelength assuming perfectly coherent sources.

14.11e Under what conditions does Eq. 14.11(8) reduce to the simpler Eq. 14.11(4)? For $\tau = 10$ ps and $\lambda = 1.5$ μm, how small must the $\Delta\lambda$ of the laser be for the signal spectrum to be dominant over the source spectrum?

14.12a Show that Eq. 14.12(13) is a solution of (12) for the conditions stated.

14.12b For a silica fiber of core diameter 10 μm and refractive index 1.5, estimate the power required to maintain a fundamental soliton if wavelength is 1.4 μm with dispersion estimated as 5 ps/nm-km.

14.13a Verify the form in Eq. 14.13(12).

14.13b A typical helium–neon laser with $\lambda_0 = 633$ nm has a beam radius of about 0.5 mm. Taking this as w_0, find beam radius at the other side of a bay 10 km away. Similarly find beam radius on the moon of a Nd–YAG laser ($\lambda_0 = 1.06$ μm) 3.84×10^3 km from the earth where it starts with $w_0 = 5$ mm.

14.13c For the examples of Prob. 14.13b, there is an optimum w_0 to produce minimum beam radius at the receiver a given distance away. Find the optimum w_0 and the corresponding $w(z)$ at the targets, for the two examples of that problem.

14.13d** Show that Eq. 14.13(18) is a solution of (2) in rectangular coordinates.

14.13e** Show that Eq. 14.13(19) is a solution of (2) in circular cylindrical coordinates.

14.13f* Note that in contrast to the gaussian beam, a beam with Bessel function variation in radius,

$$E = E_0 J_0(\alpha r) e^{j(\omega t - \beta z)}$$

does not vary as it propagates in z and has been proposed for some applications [J. Durnin, *J. Opt. Soc. Am.* **A4**, 651, (1987)]. Show under what conditions it is a solution of the wave equation. Find the power propagating in an annular ring between the zeros m and $m + 1$ of the Bessel function and show that this is independent of m for large m. Discuss the advantages and disadvantages of the profile as compared with the gaussian.

14.14a A gaussian beam of beam radius w_1 and radius of curvature R_1 passes from dielectric with index n_1 to one with index n_2, the plane interface being normal to the beam axis. Find gaussian beam properties in medium 2.

14.14b One practical problem is that of focusing the output of a laser, assumed to be of fundamental gaussian beam form, onto a fiber distance L away. If beam radius of the laser (assumed a waist) is w_1 and that at the fiber (also a waist) is w_2, find position d from the laser and focal length f of a thin lens for the desired focusing. (*Hint:* Work forward from the laser and backward from the fiber until gaussian beams intersect.)

14.14c This is a variation of Prob. 14.14b in which there is a specific lens with given focal length f, but its placing and the laser-fiber spacing L are variable. Find d and L for the given w_1, w_2, and f in this case.

14.14d In the rod with quadratic index variation described in Sec. 14.10 [$n(0) = 1.5$, $\Delta = 0.01$, $\lambda_0 = 1$ μm, $a = 50$ μm], a gaussian beam is introduced with zero slope but $w(0) = 12$ μm. Describe the beam propagation for $z > 0$.

14.14e This is similar to Prob. 14.14d except that the gaussian beam is introduced into the graded index fiber at the equilibrium radius but with slope $dw/dz = 0.2$.

14.15a Derive Eq. 14.15(8) from Eqs. (5), (6), and (7).

14.15b For a helium–neon laser, the discharge tube has a diameter of 5 mm and length of 40 cm. A plane mirror is placed at one end and it is desired to keep gaussian beam diameter not more than 3 mm at the other end to minimize wall losses. Find the

radius of curvature of a mirror to be placed close to the second end. Wavelength is 633 nm. Comment on the two solutions.

14.15c For a typical solid-state laser such as ruby, Nd–YAG, or alexandrite, $a_1 = a_2 \approx 5$ mm, $d \approx 10$ cm, and $\lambda_0 \approx 1$ μm with refractive index ≈ 1.7. Check the condition on Fresnel number N for stability of such resonators. (Note that dielectric discontinuity at the crystal side boundary further confines the beam.)

14.15d The resonator for a CO_2 laser with $\lambda_0 = 10.6$ μm has radii of curvature $R_1 = 10$ m, $R_2 = 20$ m (sign convention as in Fig. 14.14*a*) and spacing $d = 1.5$ m. Show location on the stability diagram in Fig. 14.6*c* and find radius and position of the beam waist and the beam radii at the two mirrors.

14.15e A half-confocal resonator is made by inserting a plane mirror at the midplane of a symmetric confocal resonator. The resulting resonator has $R_1 = 2d$, $R_2 = \infty$, and by an image argument, would seem to be equivalent to the original symmetric resonator. Check the *ABCD* matrix for a round trip of this resonator and comment on the differences from the symmetric confocal resonator.

14.16a Show location on a stability diagram of the resonators of Probs. 14.15b and d and of the half-confocal resonator of Prob. 14.15e.

14.16b A ruby rod 10 cm long has refractive index $n = 1.77$ at free-space wavelength $\lambda_0 = 0.6943$ μm. If the plane ends form the resonant reflectors, find the longitudinal mode number l nearest to the given wavelength and the frequency separation between longitudinal modes.

14.16c The ruby rod of Prob. 14.16b now has its ends ground to form a confocal resonator. Give the radius of curvature needed and calculate minimum spot size and spot size at the ends. For a given mode number l, how much is frequency shifted from the value for the plane mirrors?

14.16d Find longitudinal mode separation and transverse mode separation for the resonators of Probs. 14.15b and d.

14.16e Show that there are stable configurations in which the mirror curvatures are in the same direction. That is, the mode exists between a concave and convex surface.

14.17a Supply the details of the derivation of Eq. 14.17(9) with the definitions noted in the text.

14.17b Show specifically by use of Eq. 14.17(9) that a gaussian beam at the input focal plane does transform to a gaussian at the output focal plane, as stated in the text.

14.17c Convert the integral of Eq. 14.17(6) to polar coordinates (r, ϕ) and find the form of ψ at the output focal plane if that at the input focal plane corresponds to a uniformly illuminated circle of radius a, centered on the axis.

14.18a By using two successive Fourier transforms of form Eq. 14.17(6), show that when $T_2(x_2) = 1$, the $\psi_3(x_3)$ of Fig. 14.18*b* is like the input function but inverted and changed in scale.

14.18b In Ex. 14.18 suppose that the spatial filter passes only lines $\pm n$ but with a π-phase shift for $-n$ and zero-phase shift for $+n$. Describe the output function $\psi_3(x_3)$.

14.18c Extend Ex. 14.18 to a two-dimensional screen with both horizontal and vertical gratings. Describe the pattern for the Fourier plane. Design a spatial filter so that only a horizontal grating appears at plane 3.

14.19a In holography the reference wave need not be a plane wave. Consider Ex. 14.19a with the object wave a plane wave incident at an angle on the recording plane and the reference a spherical wave:

$$R(x, y, z) = R_0 \exp\{-jk[x^2 + y^2 + (z - z_1)^2]^{1/2}\}$$

Find $T(x, y)$ in the plane $z = 0$ and the transmitted wave when the resulting hologram is irradiated with this R.

14.19b As in Prob. 14.19a but with the object wave a spherical wave as in Ex. 14.19b.

14.19c It was shown in Chapter 12 that a spherically symmetric wave of the form 14.19(9) is not a solution of Maxwell's equations. Discuss the concept of a point source in optics and the conditions under which (9) may be a useful approximation.

14.19d A hologram may be made of the Fourier transform of a two-dimensional function by combining concepts of this section and the preceding two. Sketch an arrangement for making such a Fourier transform hologram.