Fundamentals of Photonics
Bahaa E. A. Saleh, Malvin Carl Teich
Copyright © 1991 John Wiley & Sons, Inc.
ISBNs: 0-471-83965-5 (Hardback); 0-471-2-1374-8 (Electronic)

CHAPTER

3

BEAM OPTICS

The Gaussian beam is named after the great mathematician **Karl Friedrich Gauss (1777–1855)**.

Lord Rayleigh (John W. Strutt) (1842–1919) contributed to many areas of optics, including scattering, diffraction, radiation, and image formation. The depth of focus of the Gaussian beam is named after him.

Can light be spatially confined and transported in free space without angular spread? Although the wave nature of light precludes the existence of such an idealization, light can take the form of beams that come as close as possible to spatially localized and nondiverging waves.

A plane wave and a spherical wave represent the two opposite extremes of angular and spatial confinement. The wavefront normals (rays) of a plane wave are parallel to the direction of the wave so that there is no angular spread, but the energy extends spatially over the entire space. The spherical wave, on the other hand, originates from a single point, but its wavefront normals (rays) diverge in all directions.

Waves with wavefront normals making small angles with the z axis are called paraxial waves. They must satisfy the paraxial Helmholtz equation derived in Sec. 2.2C. An important solution of this equation that exhibits the characteristics of an optical beam is a wave called the **Gaussian beam**. The beam power is principally concentrated within a small cylinder surrounding the beam axis. The intensity distribution in any transverse plane is a circularly symmetric Gaussian function centered about the beam axis. The width of this function is minimum at the beam waist and grows gradually in both directions. The wavefronts are approximately planar near the beam waist, but they gradually curve and become approximately spherical far from the waist. The angular divergence of the wavefront normals is the minimum permitted by the wave equation for a given beam width. The wavefront normals are therefore much like a thin pencil of rays. Under ideal conditions, the light from a laser takes the form of a Gaussian beam.

An expression for the complex amplitude of the Gaussian beam is derived in Sec. 3.1 and a detailed discussion of its physical properties (intensity, power, beam radius, angular divergence, depth of focus, and phase) is provided. The shaping of Gaussian beams (focusing, relaying, collimating, and expanding) by the use of various optical components is the subject of Sec. 3.2. A family of optical beams called Hermite–Gaussian beams, of which the Gaussian beam is a member, is introduced in Sec. 3.3. Laguerre–Gaussian and Bessel beams are discussed in Sec. 3.4.

3.1 THE GAUSSIAN BEAM

A. Complex Amplitude

The concept of paraxial waves was introduced in Sec. 2.2C. A paraxial wave is a plane wave e^{-jkz} (with wavenumber $k = 2\pi/\lambda$ and wavelength λ) modulated by a complex envelope $A(\mathbf{r})$ that is a slowly varying function of position (see Fig. 2.2-5). The complex amplitude is

$$U(\mathbf{r}) = A(\mathbf{r}) \exp(-jkz). \tag{3.1-1}$$

The envelope is assumed to be approximately constant within a neighborhood of size λ, so that the wave is locally like a plane wave with wavefront normals that are paraxial rays.

For the complex amplitude $U(\mathbf{r})$ to satisfy the Helmholtz equation, $\nabla^2 U + k^2 U = 0$, the complex envelope $A(\mathbf{r})$ must satisfy the paraxial Helmholtz equation (2.2-22)

$$\nabla_T^2 A - j2k\frac{\partial A}{\partial z} = 0, \qquad (3.1\text{-}2)$$

where $\nabla_T^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the transverse part of the Laplacian operator. One simple solution to the paraxial Helmholtz equation provides the paraboloidal wave for which

$$A(\mathbf{r}) = \frac{A_1}{z}\exp\left(-jk\frac{\rho^2}{2z}\right), \qquad \rho^2 = x^2 + y^2 \qquad (3.1\text{-}3)$$

(see Exercise 2.2-2) where A_1 is a constant. The paraboloidal wave is the paraxial approximation of the spherical wave $U(r) = (A_1/r)\exp(-jkr)$ when x and y are much smaller than z (see Sec. 2.2B).

Another solution of the paraxial Helmholtz equation provides the Gaussian beam. It is obtained from the paraboloidal wave by use of a simple transformation. Since the complex envelope of the paraboloidal wave (3.1-3) is a solution of the paraxial Helmholtz equation (3.1-2), a shifted version of it, with $z - \xi$ replacing z where ξ is a constant,

$$A(\mathbf{r}) = \frac{A_1}{q(z)}\exp\left[-jk\frac{\rho^2}{2q(z)}\right], \qquad q(z) = z - \xi, \qquad (3.1\text{-}4)$$

is also a solution. This provides a paraboloidal wave centered about the point $z = \xi$ instead of $z = 0$. When ξ is complex, (3.1-4) remains a solution of (3.1-2), but it acquires dramatically different properties. In particular, when ξ is purely imaginary, say $\xi = -jz_0$ where z_0 is real, (3.1-4) gives rise to the complex envelope of the Gaussian beam

$$\boxed{A(\mathbf{r}) = \frac{A_1}{q(z)}\exp\left[-jk\frac{\rho^2}{2q(z)}\right], \qquad q(z) = z + jz_0.} \qquad (3.1\text{-}5)$$

Complex Envelope

The parameter z_0 is known as the **Rayleigh range**.

To separate the amplitude and phase of this complex envelope, we write the complex function $1/q(z) = 1/(z + jz_0)$ in terms of its real and imaginary parts by defining two new real functions $R(z)$ and $W(z)$, such that

$$\boxed{\frac{1}{q(z)} = \frac{1}{R(z)} - j\frac{\lambda}{\pi W^2(z)}.} \qquad (3.1\text{-}6)$$

It will be shown subsequently that $W(z)$ and $R(z)$ are measures of the beam width and wavefront radius of curvature, respectively. Expressions for $W(z)$ and $R(z)$ as functions of z and z_0 are provided in (3.1-8) and (3.1-9). Substituting (3.1-6) into (3.1-5) and using (3.1-1), an expression for the complex amplitude $U(\mathbf{r})$ of the Gaussian beam is obtained:

$$\boxed{U(\mathbf{r}) = A_0 \frac{W_0}{W(z)} \exp\left[-\frac{\rho^2}{W^2(z)}\right] \exp\left[-jkz - jk\frac{\rho^2}{2R(z)} + j\zeta(z)\right]}$$

(3.1-7)
Gaussian-Beam
Complex Amplitude

$$W(z) = W_0\left[1 + \left(\frac{z}{z_0}\right)^2\right]^{1/2} \qquad (3.1\text{-}8)$$

$$R(z) = z\left[1 + \left(\frac{z_0}{z}\right)^2\right] \qquad (3.1\text{-}9)$$

$$\zeta(z) = \tan^{-1}\frac{z}{z_0} \qquad (3.1\text{-}10)$$

$$W_0 = \left(\frac{\lambda z_0}{\pi}\right)^{1/2} \qquad (3.1\text{-}11)$$

Beam Parameters

A new constant $A_0 = A_1/jz_0$ has been defined for convenience.

The expression for the complex amplitude of the Gaussian beam is central to this chapter. It contains two parameters, A_0 and z_0, which are determined from the boundary conditions. All other parameters are related to the Rayleigh range z_0 and the wavelength λ by (3.1-8) to (3.1-11).

B. Properties

Equations (3.1-7) to (3.1-11) will now be used to determine the properties of the Gaussian beam.

Intensity

The optical intensity $I(\mathbf{r}) = |U(\mathbf{r})|^2$ is a function of the axial and radial distances z and $\rho = (x^2 + y^2)^{1/2}$,

$$I(\rho, z) = I_0\left[\frac{W_0}{W(z)}\right]^2 \exp\left[-\frac{2\rho^2}{W^2(z)}\right], \qquad (3.1\text{-}12)$$

where $I_0 = |A_0|^2$. At each value of z the intensity is a Gaussian function of the radial distance ρ. This is why the wave is called a Gaussian beam. The Gaussian function has its peak at $\rho = 0$ (on axis) and drops monotonically with increasing ρ. The width $W(z)$ of the Gaussian distribution increases with the axial distance z as illustrated in Fig. 3.1-1.

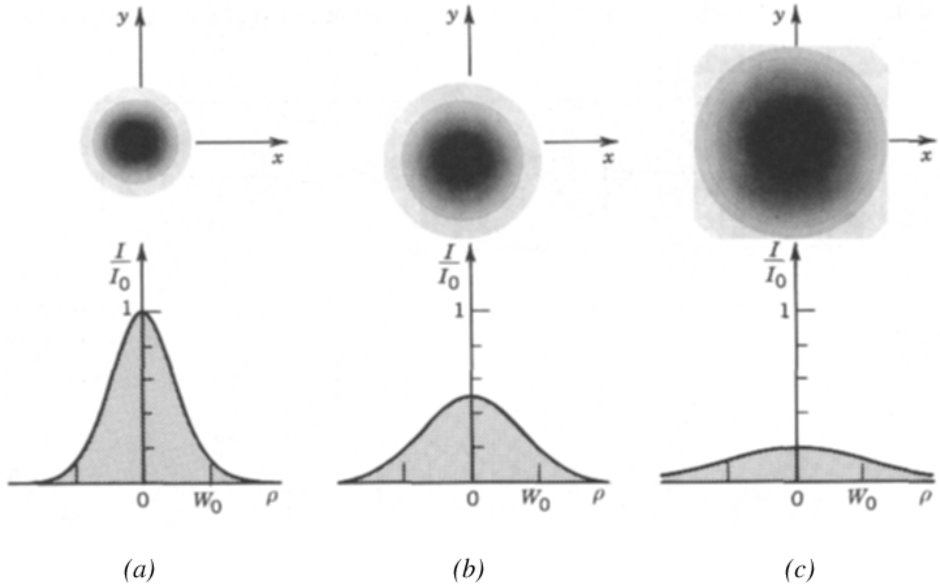

Figure 3.1-1 The normalized beam intensity I/I_0 as a function of the radial distance ρ at different axial distances: (a) $z = 0$; (b) $z = z_0$; (c) $z = 2z_0$.

On the beam axis ($\rho = 0$) the intensity

$$I(0, z) = I_0 \left[\frac{W_0}{W(z)} \right]^2 = \frac{I_0}{1 + (z/z_0)^2} \tag{3.1-13}$$

has its maximum value I_0 at $z = 0$ and drops gradually with increasing z, reaching half its peak value at $z = \pm z_0$ (Fig. 3.1-2). When $|z| \gg z_0$, $I(0, z) \approx I_0 z_0^2/z^2$, so that the intensity decreases with the distance in accordance with an inverse-square law, as for spherical and paraboloidal waves. The overall peak intensity $I(0, 0) = I_0$ occurs at the beam center ($z = 0, \rho = 0$).

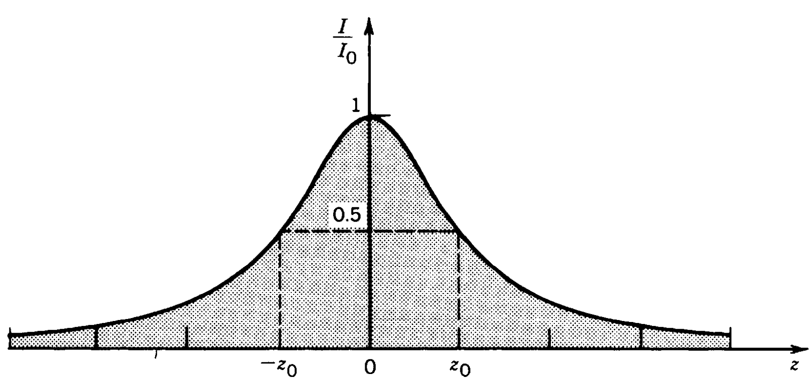

Figure 3.1-2 The normalized beam intensity I/I_0 at points on the beam axis ($\rho = 0$) as a function of z.

Power

The total optical power carried by the beam is the integral of the optical intensity over a transverse plane (say at a distance z),

$$P = \int_0^\infty I(\rho, z) 2\pi\rho \, d\rho,$$

which gives

$$P = \tfrac{1}{2} I_0 (\pi W_0^2). \tag{3.1-14}$$

The result is independent of z, as expected. Thus the beam power is one-half the peak intensity times the beam area. Since beams are often described by their power P, it is useful to express I_0 in terms of P using (3.1-14) and to rewrite (3.1-12) in the form

$$\boxed{I(\rho, z) = \frac{2P}{\pi W^2(z)} \exp\left[-\frac{2\rho^2}{W^2(z)}\right].} \tag{3.1-15}$$

Beam Intensity

The ratio of the power carried within a circle of radius ρ_0 in the transverse plane at position z to the total power is

$$\frac{1}{P}\int_0^{\rho_0} I(\rho, z) 2\pi\rho \, d\rho = 1 - \exp\left[-\frac{2\rho_0^2}{W^2(z)}\right]. \tag{3.1-16}$$

The power contained within a circle of radius $\rho_0 = W(z)$ is approximately 86% of the total power. About 99% of the power is contained within a circle of radius $1.5W(z)$.

Beam Radius

Within any transverse plane, the beam intensity assumes its peak value on the beam axis, and drops by the factor $1/e^2 \approx 0.135$ at the radial distance $\rho = W(z)$. Since 86% of the power is carried within a circle of radius $W(z)$, we regard $W(z)$ as the beam radius (also called the beam width). The rms width of the intensity distribution is $\sigma = \frac{1}{2} W(z)$ (see Appendix A, Sec. A.2, for the different definitions of width).

The dependence of the beam radius on z is governed by (3.1-8),

$$\boxed{W(z) = W_0 \left[1 + \left(\frac{z}{z_0}\right)^2\right]^{1/2}.} \tag{3.1-17}$$

Beam Radius

It assumes its minimum value W_0 in the plane $z = 0$, called the beam waist. Thus W_0 is the **waist radius**. The waist diameter $2W_0$ is called the **spot size**. The beam radius increases gradually with z, reaching $\sqrt{2}\, W_0$ at $z = z_0$, and continues increasing monotonically with z (Fig. 3.1-3). For $z \gg z_0$ the first term of (3.1-17) may be neglected, resulting in the linear relation

$$W(z) \approx \frac{W_0}{z_0} z = \theta_0 z, \tag{3.1-18}$$

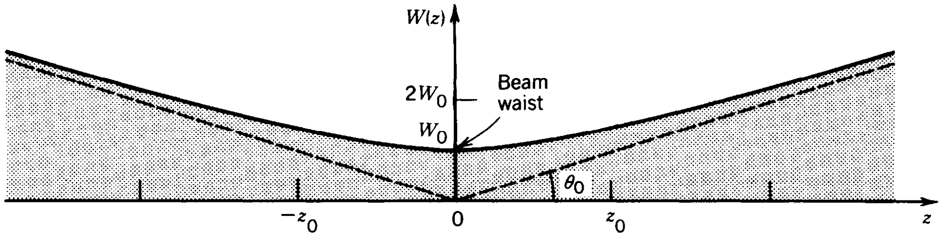

Figure 3.1-3 The beam radius $W(z)$ has its minimum value W_0 at the waist ($z = 0$), reaches $\sqrt{2}\,W_0$ at $z = \pm z_0$, and increases linearly with z for large z.

where $\theta_0 = W_0/z_0$. Using (3.1-11), we can also write

$$\theta_0 = \frac{\lambda}{\pi W_0}. \tag{3.1-19}$$

Beam Divergence

Far from the beam center, when $z \gg z_0$, the beam radius increases approximately linearly with z, defining a cone with half-angle θ_0. About 86% of the beam power is confined within this cone. The angular divergence of the beam is therefore defined by the angle

$$\boxed{\theta_0 = \frac{2}{\pi}\frac{\lambda}{2W_0}.} \tag{3.1-20}$$

Divergence Angle

The beam divergence is directly proportional to the ratio between the wavelength λ and the beam-waist diameter $2W_0$. If the waist is squeezed, the beam diverges. To obtain a highly directional beam, therefore, a short wavelength and a fat beam waist should be used.

Depth of Focus

Since the beam has its minimum width at $z = 0$, as shown in Fig. 3.1-3, it achieves its best focus at the plane $z = 0$. In either direction, the beam gradually grows "out of focus." The axial distance within which the beam radius lies within a factor $\sqrt{2}$ of its minimum value (i.e., its area lies within a factor of 2 of its minimum) is known as the **depth of focus** or **confocal parameter** (Fig. 3.1-4). It can be seen from (3.1-17) that the

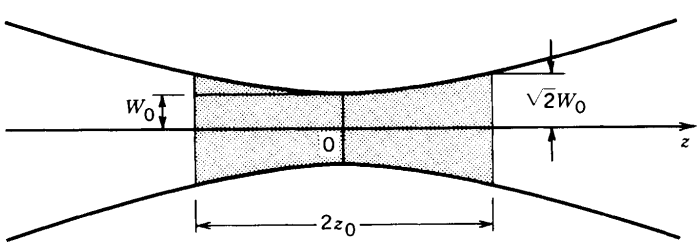

Figure 3.1-4 The depth of focus of a Gaussian beam.

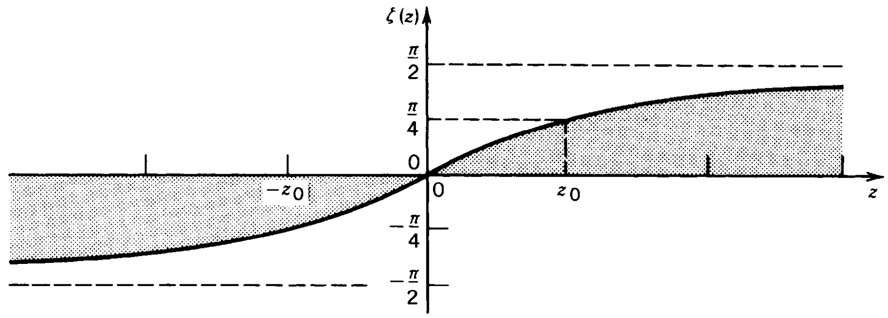

Figure 3.1-5 $\zeta(z)$ is the phase retardation of the Gaussian beam relative to a uniform plane wave at points on the beam axis.

depth of focus is twice the Rayleigh range,

$$\boxed{2z_0 = \frac{2\pi W_0^2}{\lambda}.} \tag{3.1-21}$$

Depth of Focus

The depth of focus is directly proportional to the area of the beam at its waist, and inversely proportional to the wavelength. Thus when a beam is focused to a small spot size, the depth of focus is short and the plane of focus must be located with greater accuracy. A small spot size and a long depth of focus cannot be obtained simultaneously unless the wavelength of the light is short. For $\lambda = 633$ nm (the wavelength of a He–Ne laser line), for example, a spot size $2W_0 = 2$ cm corresponds to a depth of focus $2z_0 \approx 1$ km. A much smaller spot size of 20 μm corresponds to a much shorter depth of focus of 1 mm.

Phase

The phase of the Gaussian beam is, from (3.1-7),

$$\varphi(\rho, z) = kz - \zeta(z) + \frac{k\rho^2}{2R(z)}. \tag{3.1-22}$$

On the beam axis ($\rho = 0$) the phase

$$\varphi(0, z) = kz - \zeta(z) \tag{3.1-23}$$

comprises two components. The first, kz, is the phase of a plane wave. The second represents a phase retardation $\zeta(z)$ given by (3.1-10) which ranges from $-\pi/2$ at $z = -\infty$ to $+\pi/2$ at $z = \infty$, as illustrated in Fig. 3.1-5. This phase retardation corresponds to an excess delay of the wavefront in comparison with a plane wave or a spherical wave (see also Fig. 3.1-8). The total accumulated excess retardation as the wave travels from $z = -\infty$ to $z = \infty$ is π. This phenomenon is known as the **Guoy effect.**[†]

Wavefronts

The third component in (3.1-22) is responsible for wavefront bending. It represents the deviation of the phase at off-axis points in a given transverse plane from that at the

[†]See, for example, A. E. Siegman, *Lasers*, University Science Books, Mill Valley, CA, 1986.

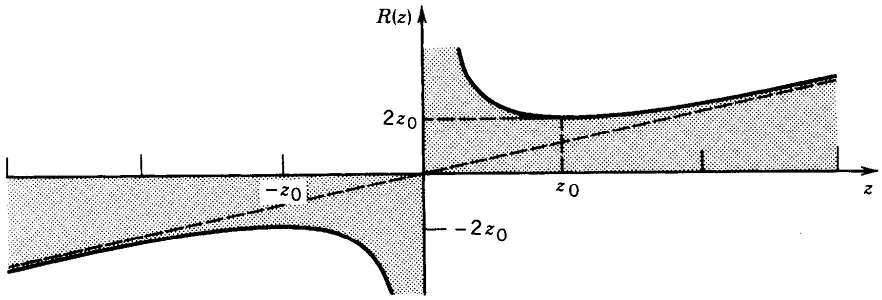

Figure 3.1-6 The radius of curvature $R(z)$ of the wavefronts of a Gaussian beam. The dashed line is the radius of curvature of a spherical wave.

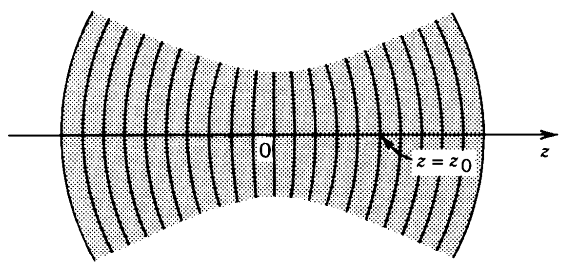

Figure 3.1-7 Wavefronts of a Gaussian beam.

axial point. The surfaces of constant phase satisfy $k[z + \rho^2/2R(z)] - \zeta(z) = 2\pi q$. Since $\zeta(z)$ and $R(z)$ are relatively slowly varying, they are approximately constant at points within the beam radius on each wavefront. We may therefore write $z + \rho^2/2R = q\lambda + \zeta\lambda/2\pi$, where $R = R(z)$ and $\zeta = \zeta(z)$. This is precisely the equation of a paraboloidal surface of radius of curvature R. Thus $R(z)$, plotted in Fig. 3.1-6, is the radius of curvature of the wavefront at position z on the beam axis.

As illustrated in Fig. 3.1-6, the radius of curvature $R(z)$ is infinite at $z = 0$, corresponding to planar wavefronts. It decreases to a minimum value of $2z_0$ at $z = z_0$. This is the point at which the wavefront has the greatest curvature (Fig. 3.1-7). The radius of curvature subsequently increases with further increase of z until $R(z) \approx z$ for $z \gg z_0$. The wavefront is then approximately the same as that of a spherical wave. For negative z the wavefronts follow an identical pattern, except for a change in sign. We have adopted the convention that a diverging wavefront has a positive radius of curvature, whereas a converging wavefront has a negative radius of curvature.

Summary: Properties of the Gaussian Beam at Special Points

- *At the plane $z = z_0$.* At an axial distance z_0 from the beam waist, the wave has the following properties:
  - (i) The beam radius is $\sqrt{2}$ times greater than the radius at the beam waist, and the area is larger by a factor of 2.
  - (ii) The intensity on the beam axis is $\frac{1}{2}$ the peak intensity.
  - (iii) The phase on the beam axis is retarded by an angle $\pi/4$ relative to the phase of a plane wave.
  - (iv) The radius of curvature of the wavefront is the smallest, so that the wavefront has the greatest curvature ($R = 2z_0$).

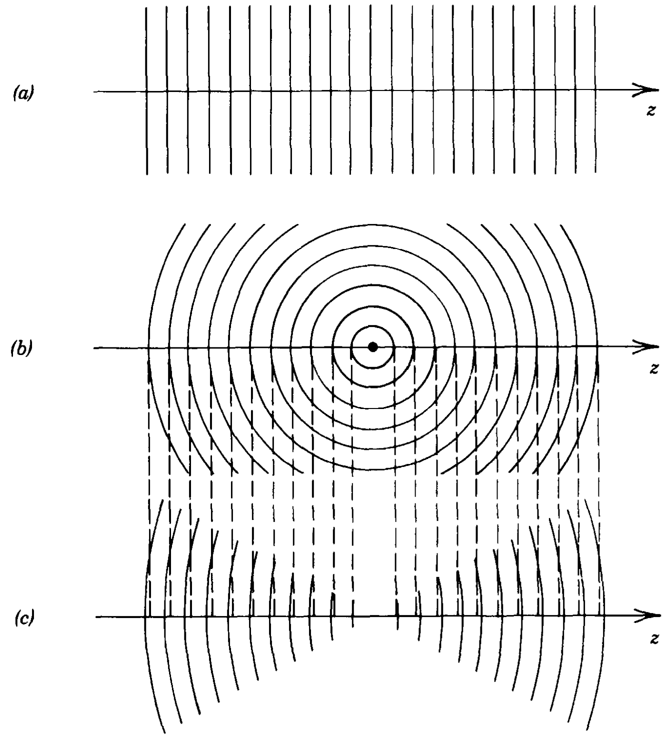

Figure 3.1-8 Wavefronts of (*a*) a uniform plane wave; (*b*) a spherical wave; (*c*) a Gaussian beam. At points near the beam center, the Gaussian beam resembles a plane wave. At large z the beam behaves like a spherical wave except that the phase is retarded by 90° (shown in this diagram by a quarter of the distance between two adjacent wavefronts).

- *Near the Beam Center.* At points for which $|z| \ll z_0$ and $\rho \ll W_0$, $\exp[-\rho^2/W^2(z)] \approx \exp(-\rho^2/W_0^2) \approx 1$, so that the beam intensity is approximately constant. Also, $R(z) \approx z_0^2/z$ and $\zeta(z) \approx 0$, so that the phase $k[z + \rho^2/2R(z)] = kz(1 + \rho^2/2z_0^2) \approx kz$. As a result, the wavefronts are approximately planar. The Gaussian beam may therefore be approximated near its center by a plane wave.
- *Far from the Beam Waist.* At points within the beam-waist radius ($\rho < W_0$) but far from the beam waist ($z \gg z_0$) the wave is approximately like a spherical wave. Since $W(z) \approx W_0 z/z_0 \gg W_0$ and $\rho < W_0$, $\exp[-\rho^2/W^2(z)] \approx 1$, so that the beam intensity is approximately uniform. Since $R(z) \approx z$ the wavefronts are approximately spherical. Thus, except for an excess phase $\zeta(z) \approx \pi/2$, the complex amplitude of the Gaussian beam approaches that of the paraboloidal wave, which in turn approaches that of the spherical wave in the paraxial approximation (Fig. 3.1-8).

EXERCISE 3.1-1

Parameters of a Gaussian Laser Beam. A 1-mW He–Ne laser produces a Gaussian beam of wavelength $\lambda = 633$ nm and a spot size $2W_0 = 0.1$ mm.

(a) Determine the angular divergence of the beam, its depth of focus, and its diameter at $z = 3.5 \times 10^5$ km (approximately the distance to the moon).
(b) What is the radius of curvature of the wavefront at $z = 0$, $z = z_0$, and $z = 2z_0$?
(c) What is the optical intensity (in $\mathrm{W/cm^2}$) at the beam center ($z = 0, \rho = 0$) and at the axial point $z = z_0$? Compare this with the intensity at $z = z_0$ of a 100-W spherical wave produced by a small isotropically emitting light source located at $z = 0$.

EXERCISE 3.1-2

Validity of the Paraxial Approximation for a Gaussian Beam. The complex envelope $A(\mathbf{r})$ of a Gaussian beam is an exact solution of the paraxial Helmholtz equation (3.1-2), but its corresponding complex amplitude $U(\mathbf{r}) = A(\mathbf{r})\exp(-jkz)$ is only an approximate solution of the Helmholtz equation (2.2-7). This is because the paraxial Helmholtz equation is itself approximate. The approximation is satisfactory if the condition (2.2-20) is satisfied. Show that if the divergence angle θ_0 of a Gaussian beam is small ($\theta_0 \ll 1$), the condition (2.2-20) for the validity of the paraxial Helmholtz equation is satisfied.

Parameters Required to Characterize a Gaussian Beam

Assuming that the wavelength λ is known, how many parameters are required to describe a plane wave, a spherical wave, and a Gaussian beam? The plane wave is completely specified by its complex amplitude and direction. The spherical wave is specified by its amplitude and the location of its origin. The Gaussian beam, in contrast, is characterized by more parameters—its peak amplitude [the parameter A_0 in (3.1-7)], its direction (the beam axis), the location of its waist, *and* one additional parameter: the waist radius W_0 *or* the Rayleigh range z_0, for example. Thus, if the beam peak amplitude and the axis are known, two additional parameters are necessary.

If the complex number $q(z) = z + jz_0$ is known, the distance z to the beam waist and the Rayleigh range z_0 are readily identified as the real and imaginary parts of $q(z)$. As an example, if the q-parameter is $3 + j4$ cm at some point on the beam axis, we conclude that the beam waist lies at a distance $z = 3$ cm to the left of that point and that the depth of focus is $2z_0 = 8$ cm. The waist radius W_0 may be determined by use of (3.1-11). The **q-parameter** $q(z)$ is therefore sufficient for characterizing a Gaussian beam of known peak amplitude and beam axis. The linear dependence of the q-parameter on z permits us to readily determine q at all points, given q at a single point. If $q(z) = q_1$ and $q(z + d) = q_2$, then $q_2 = q_1 + d$. In the present example, at $z = 13$ cm, $q = 13 + j4$.

If the beam width $W(z)$ and the radius of curvature $R(z)$ are known at an arbitrary point on the axis, the beam can be identified completely by solving (3.1-8), (3.1-9), and (3.1-11) for z, z_0, and W_0. Alternatively, the q-parameter may be determined from $W(z)$ and $R(z)$ using the relation, $1/q(z) = 1/R(z) - j\lambda/[\pi W^2(z)]$, from which the beam is identified.

EXERCISE 3.1-3

Determination of a Beam with Given Width and Curvature. Assuming that the width W and the radius of curvature R of a Gaussian beam are known at some point on the beam axis (Fig. 3.1-9), show that the beam waist is located at a distance

$$z = \frac{R}{1 + \left(\lambda R/\pi W^2\right)^2} \tag{3.1-24}$$

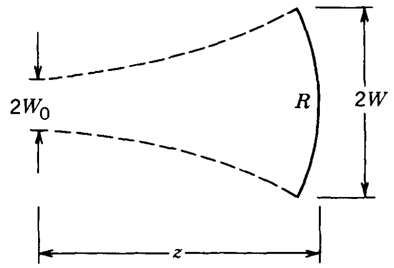

Figure 3.1-9 Given W and R, determine z and W_0.

to the left and the waist radius is

$$W_0 = \frac{W}{\left[1 + \left(\pi W^2/\lambda R\right)^2\right]^{1/2}}. \tag{3.1-25}$$

EXERCISE 3.1-4

Determination of the Width and Curvature at One Point Given the Width and Curvature at Another Point. Assume that the radius of curvature and the width of a Gaussian beam of wavelength $\lambda = 1$ μm at some point on the beam axis are $R_1 = 1$ m and $W_1 = 1$ mm, respectively (Fig. 3.1-10). Determine the beam width and the radius of curvature at a distance $d = 10$ cm to the right.

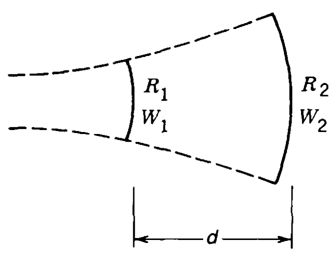

Figure 3.1-10 Given R_1, W_1, and d, determine R_2 and W_2.

EXERCISE 3.1-5

Identification of a Beam with Known Curvatures at Two Points. A Gaussian beam has radii of curvature R_1 and R_2 at two points on the beam axis separated by a distance d, as illustrated in Fig. 3.1-11. Verify that the location of the beam center and its depth of

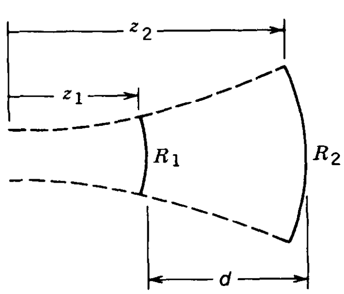

Figure 3.1-11 Given R_1, R_2, and d, determine z_1, z_2, z_0, and W_0.

focus may be determined from the relations

$$z_1 = \frac{-d(R_2 - d)}{R_2 - R_1 - 2d} \tag{3.1-26}$$

$$z_0^2 = \frac{-d(R_1 + d)(R_2 - d)(R_2 - R_1 - d)}{(R_2 - R_1 - 2d)^2} \tag{3.1-27}$$

$$W_0 = \left(\frac{\lambda z_0}{\pi}\right)^{1/2}.$$

3.2 TRANSMISSION THROUGH OPTICAL COMPONENTS

The effects of different optical components on a Gaussian beam are discussed in this section. We show that if a Gaussian beam is transmitted through a set of circularly symmetric optical components aligned with the beam axis, *the Gaussian beam remains a Gaussian beam* as long as the overall system maintains the paraxial nature of the wave. Only the beam waist and curvature are altered so that the beam is only reshaped. The results of this section are important in the design of optical instruments in which Gaussian beams are used.

A. Transmission Through a Thin Lens

The complex amplitude transmittance of a thin lens of focal length f is proportional to $\exp(jk\rho^2/2f)$ (see Sec. 2.4B). When a Gaussian beam crosses the lens its complex amplitude, given in (3.1-7), is multiplied by this phase factor. As a result, its wavefront is bent, but the beam radius is not altered.

A Gaussian beam centered at $z = 0$ with waist radius W_0 is transmitted through a thin lens located at a distance z, as illustrated in Fig. 3.2-1. The phase at the plane of the lens is $kz + k\rho^2/2R - \zeta$, where $R = R(z)$ and $\zeta = \zeta(z)$ are given by (3.1-9) and (3.1-10), respectively. The phase of the transmitted wave is altered to

$$kz + k\frac{\rho^2}{2R} - \zeta - k\frac{\rho^2}{2f} = kz + k\frac{\rho^2}{2R'} - \zeta, \tag{3.2-1}$$

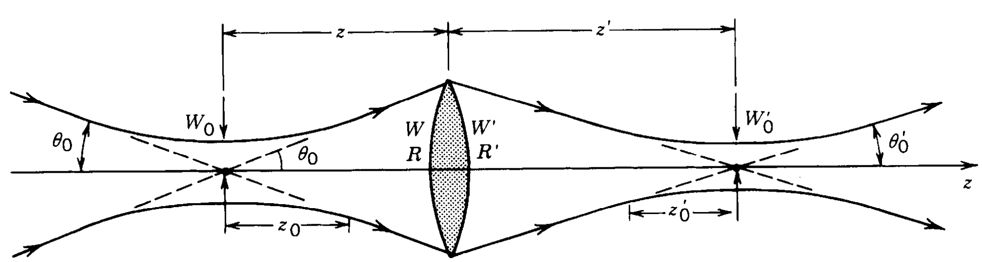

Figure 3.2-1 Transmission of a Gaussian beam through a thin lens.

where

$$\frac{1}{R'} = \frac{1}{R} - \frac{1}{f}. \tag{3.2-2}$$

We conclude that the transmitted wave is itself a Gaussian beam with width $W' = W$ and radius of curvature R', where R' satisfies the imaging equation $1/R - 1/R' = 1/f$. Note that R is positive since the wavefront of the incident beam is diverging and R' is negative since the wavefront of the transmitted beam is converging.

The parameters of the emerging beam may be determined by referring to Exercise 3.1-3, in which the parameters of a Gaussian beam were determined from its width and curvature at a given point. By use of (3.1-25) and (3.1-24) the waist radius of the new beam is

$$W_0' = \frac{W}{\left[1 + \left(\pi W^2/\lambda R'\right)^2\right]^{1/2}}, \tag{3.2-3}$$

and the center is located a distance

$$-z' = \frac{R'}{1 + \left(\lambda R'/\pi W^2\right)^2} \tag{3.2-4}$$

from the lens. A minus sign is used in (3.2-4) since the waist lies to the right of the lens. Substituting $R = z[1 + (z_0/z)^2]$ and $W = W_0[1 + (z/z_0)^2]^{1/2}$ into (3.2-2) to (3.2-4), the following expressions, which relate the parameters of the two beams, are obtained (Fig. 3.2-1):

Waist radius	$W_0' = MW_0$	(3.2-5)
Waist location	$(z' - f) = M^2(z - f)$	(3.2-6)
Depth of focus	$2z_0' = M^2(2z_0)$	(3.2-7)
Divergence	$2\theta_0' = \dfrac{2\theta_0}{M}$	(3.2-8)
Magnification	$M = \dfrac{M_r}{(1 + r^2)^{1/2}}$	(3.2-9)
$r = \dfrac{z_0}{z - f},$	$M_r = \left\lvert\dfrac{f}{z - f}\right\rvert.$	(3.2-9a)

Parameter Transformation by a Lens

The magnification factor M plays an important role. The beam waist is magnified by M, the beam depth of focus is magnified by M^2, and the angular divergence is minified by the factor M.

Limit of Ray Optics

Consider the limiting case in which $(z - f) \gg z_0$, so that the lens is well outside the depth of focus of the incident beam (Fig. 3.2-2). The beam may then be approximated by a spherical wave, and the parameter $r \ll 1$ so that $M \approx M_r$ [see (3.2-9a)]. Thus

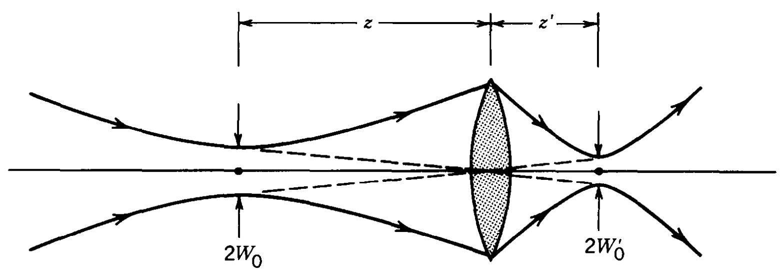

Figure 3.2-2 Beam imaging in the ray-optics limit.

(3.2-5) to (3.2-9a) reduce to

$$W_0' \approx MW_0 \tag{3.2-10}$$

$$\frac{1}{z'} + \frac{1}{z} \approx \frac{1}{f} \tag{3.2-11}$$

$$M \approx M_r = \left| \frac{f}{z - f} \right|. \tag{3.2-12}$$

Equations (3.2-10) to (3.2-12) are precisely the relations provided by ray optics for the location and size of a patch of light of diameter $2W_0$ located a distance z to the left of a thin lens (see Sec. 1.2C). The magnification factor M_r is that based on ray optics. Since (3.2-9) provides that $M < M_r$, the maximum magnification attainable is the ray-optics magnification M_r. As r^2 increases, the deviation from ray optics grows and the magnification decreases. Equations (3.2-10) to (3.2-12) also correspond to the results obtained from wave optics for the focusing of a spherical wave in the paraxial approximation (see Sec. 2.4B).

B. Beam Shaping

A lens, or sequence of lenses, may be used to reshape a Gaussian beam without compromising its Gaussian nature.

Beam Focusing

If a lens is placed at the waist of a Gaussian beam, as shown in Fig. 3.2-3, the parameters of the transmitted Gaussian beam are determined by substituting $z = 0$ in

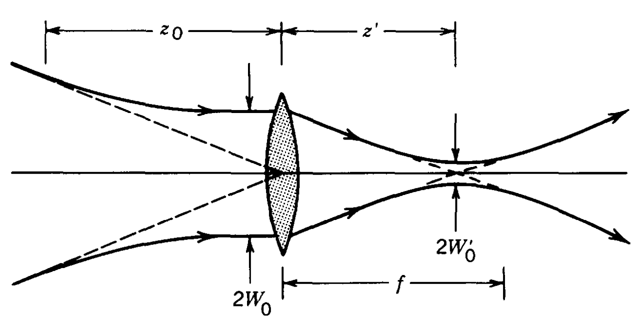

Figure 3.2-3 Focusing a beam with a lens at the beam waist.

(3.2-5) to (3.2-9a). The transmitted beam is then focused to a waist radius W_0' at a distance z' given by

$$\boxed{\begin{aligned} W_0' &= \frac{W_0}{\left[1 + (z_0/f)^2\right]^{1/2}} \\ z' &= \frac{f}{1 + (f/z_0)^2}. \end{aligned}}$$ (3.2-13) (3.2-14)

If the depth of focus of the incident beam $2z_0$ is much longer than the focal length f of the lens (Fig. 3.2-4), then $W_0' \approx (f/z_0)W_0$. Using $z_0 = \pi W_0^2/\lambda$, we obtain

$$W_0' \approx \frac{\lambda}{\pi W_0} f = \theta_0 f \tag{3.2-15}$$

$$z' \approx f. \tag{3.2-16}$$

The transmitted beam is then focused at the lens' focal plane as would be expected for parallel rays incident on a lens. This occurs because the incident Gaussian beam is well approximated by a plane wave at its waist. The spot size expected from ray optics is, of course, zero. In wave optics, however, the focused waist radius W_0' is directly proportional to the wavelength and the focal length, and inversely proportional to the radius of the incident beam. In the limit $\lambda \to 0$, the spot size does indeed approach zero in accordance with ray optics.

In many applications, such as laser scanning, laser printing, and laser fusion, it is desirable to generate the smallest possible spot size. It is clear from (3.2-15) that this may be achieved by use of the shortest possible wavelength, the thickest incident beam, and the shortest focal length. Since the lens should intercept the incident beam, its diameter D must be at least $2W_0$. Assuming that $D = 2W_0$, the diameter of the focused spot is given by

$$\boxed{2W_0' \approx \frac{4}{\pi} \lambda F_{\#},} \qquad F_{\#} = \frac{f}{D}, \tag{3.2-17}$$

Focused Spot Size

where $F_{\#}$ is the F-number of the lens. A microscope objective with small F-number is often used. Since (3.2-15) and (3.2-16) are approximate, their validity must always be confirmed before use.

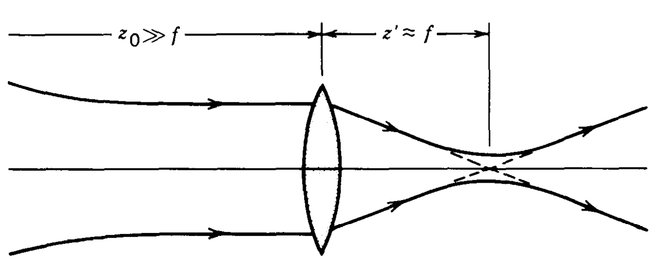

Figure 3.2-4 Focusing a collimated beam.

EXERCISE 3.2-1

Beam Relaying. A Gaussian beam of radius W_0 and wavelength λ is repeatedly focused by a sequence of identical lenses, each of focal length f and separated by distance d (Fig. 3.2-5). The focused waist radius is equal to the incident waist radius, i.e., $W_0' = W_0$. Using (3.2-6), (3.2-9), and (3.2-9a) show that this condition can arise only if the inequality $d \leq 4f$ is satisfied. Note that this is the same condition of ray confinement for a sequence of lenses derived in Sec. 1.4D using ray optics.

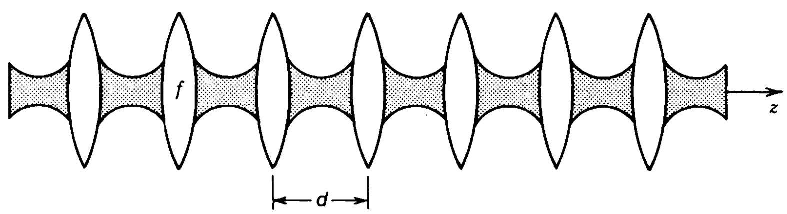

Figure 3.2-5 Beam relaying.

EXERCISE 3.2-2

Beam Collimation. A Gaussian beam is transmitted through a thin lens of focal length f.

(a) Show that the locations of the waists of the incident and transmitted beams, z and z', are related by

$$\frac{z'}{f} - 1 = \frac{z/f - 1}{(z/f - 1)^2 + (z_0/f)^2}. \tag{3.2-18}$$

This relation is plotted in Fig. 3.2-6.

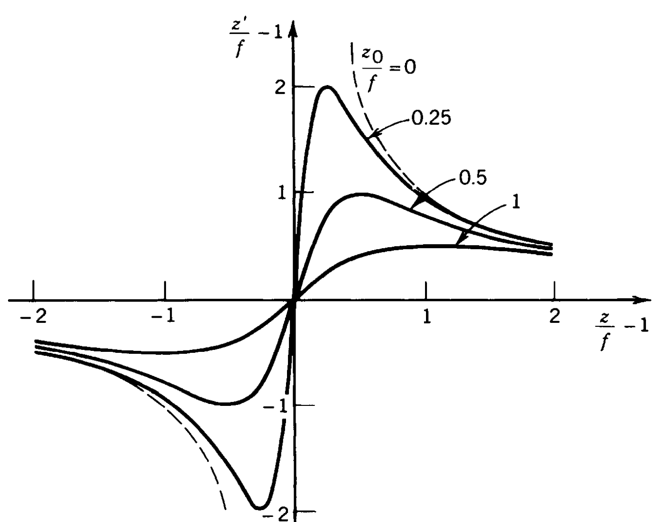

Figure 3.2-6 Relation between the waist locations of the incident and transmitted beams.

(b) The beam is collimated by making the location of the new waist z' as distant as possible from the lens. This is achieved by using the smallest ratio z_0/f (short depth of focus and long focal length). For a given ratio z_0/f, show that the optimal value of z for collimation is $z = f + z_0$.

(c) If $\lambda = 1\ \mu\text{m}$, $z_0 = 1$ cm and $f = 50$ cm, determine the optimal value of z for collimation, and the corresponding magnification M, distance z', and width W_0' of the collimated beam.

EXERCISE 3.2-3

Beam Expansion. A Gaussian beam is expanded and collimated using two lenses of focal lengths f_1 and f_2, as illustrated in Fig. 3.2-7. Parameters of the initial beam (W_0, z_0) are modified by the first lens to (W_0'', z_0'') and subsequently altered by the second lens to (W_0', z_0'). The first lens, which has a short focal length, serves to reduce the depth of focus $2z_0''$ of the beam. This prepares it for collimation by the second lens, which has a long focal length. The system functions as an inverse Keplerian telescope.

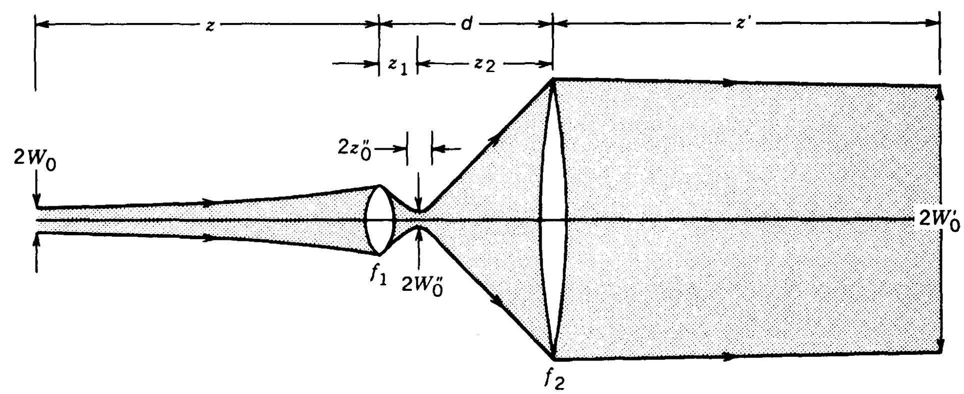

Figure 3.2-7 Beam expansion using a two-lens system.

(a) Assuming that $f_1 \ll z$ and $z - f_1 \gg z_0$, use the results of Exercise 3.2-2 to determine the optimal distance d between the lenses such that the distance z' to the waist of the final beam is as large as possible.

(b) Determine an expression for the overall magnification $M = W_0'/W_0$ of the system.

C. Reflection from a Spherical Mirror

We now examine the reflection of a Gaussian beam from a spherical mirror. Since the complex amplitude reflectance of the mirror is proportional to $\exp(-jk\rho^2/R)$, where by convention $R > 0$ for convex mirrors and $R < 0$ for concave mirrors, the action of the mirror on a Gaussian beam of width W_1 and radius of curvature R_1 is to reflect the beam and to modify its phase by the factor $-k\rho^2/R$, keeping its radius unaltered. Thus the reflected beam remains Gaussian, with parameters W_2 and R_2 given by

$$W_2 = W_1 \tag{3.2-19}$$

$$\frac{1}{R_2} = \frac{1}{R_1} + \frac{2}{R}. \tag{3.2-20}$$

Equation (3.2-20) is the same as (3.2-2) if $f = -R/2$. Thus the Gaussian beam is

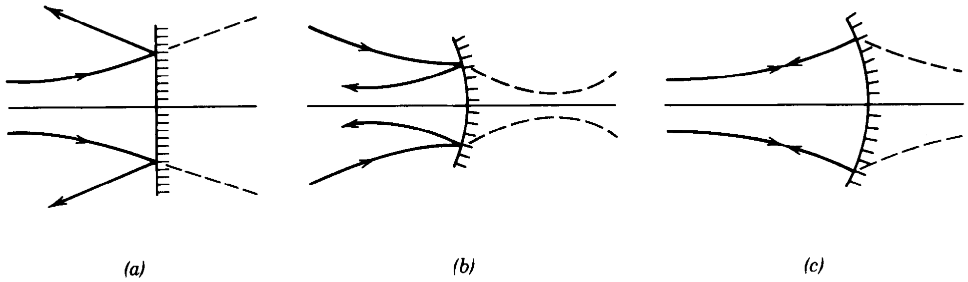

Figure 3.2-8 Reflection of a Gaussian beam of curvature R_1 from a mirror of curvature R: (*a*) $R = \infty$; (*b*) $R_1 = \infty$; (*c*) $R_1 = -R$. The dashed curves show the effects of replacing the mirror by a lens of focal length $f = -R/2$.

modified in precisely the same way as by the lens, except for a reversal of the direction of propagation.

Three special cases (illustrated in Fig. 3.2-8) are of interest:

- If the *mirror is planar*, i.e., $R = \infty$, then $R_2 = R_1$, so that the mirror reverses the direction of the beam without altering its curvature, as illustrated in Fig. 3.2-8(*a*).
- If $R_1 = \infty$, i.e., the *beam waist lies on the mirror*, then $R_2 = R/2$. If the mirror is concave ($R < 0$), $R_2 < 0$, so that the reflected beam acquires a negative curvature and the wavefronts converge. The mirror then focuses the beam to a smaller spot size, as illustrated in Fig. 3.2-8(*b*).
- If $R_1 = -R$, i.e., the incident *beam has the same curvature as the mirror*, then $R_2 = R$. The wavefronts of both the incident and reflected waves coincide with the mirror and the wave retraces its path as shown in Fig. 3.2-8(*c*). This is expected since the wavefront normals are also normal to the mirror, so that the mirror reflects the wave back onto itself. In the illustration in Fig. 3.2-8(*c*) the mirror is concave ($R < 0$); the incident wave is diverging ($R_1 > 0$) and the reflected wave is converging ($R_2 < 0$).

EXERCISE 3.2-4

Variable-Reflectance Mirrors. A spherical mirror of radius R has a variable intensity reflectance characterized by $\mathscr{R}(\rho) = |\imath(\rho)|^2 = \exp(-2\rho^2/W_m^2)$, which is a Gaussian function of the radial distance ρ. The reflectance is unity on axis and falls by a factor $1/e^2$ when $\rho = W_m$. Determine the effect of the mirror on a Gaussian beam with radius of curvature R_1 and beam radius W_1 at the mirror.

*D. Transmission Through an Arbitrary Optical System

In the paraxial approximation, an optical system is completely characterized by the 2×2 ray-transfer matrix relating the position and inclination of the transmitted ray to those of the incident ray (see Sec. 1.4). We now consider how an arbitrary paraxial optical system, characterized by a matrix $\mathbf{M}$ of elements (A, B, C, D), modifies a Gaussian beam (Fig. 3.2-9).

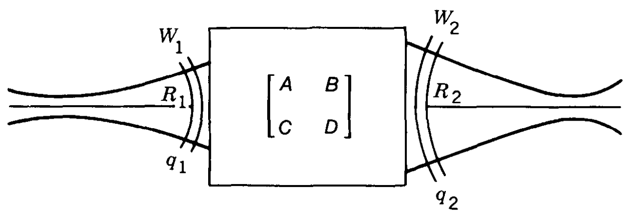

Figure 3.2-9 Modification of a Gaussian beam by an arbitrary paraxial system described by an *ABCD* matrix.

The ABCD Law

The q-parameters, q_1 and q_2, of the incident and transmitted Gaussian beams at the input and output planes of a paraxial optical system described by the (A, B, C, D) matrix are related by

$$\boxed{q_2 = \frac{Aq_1 + B}{Cq_1 + D}.} \tag{3.2-21}$$

The *ABCD* Law

Because the q parameter identifies the width W and curvature R of the Gaussian beam (see Exercise 3.1-3), this simple law, called the ***ABCD* law**, governs the effect of an arbitrary paraxial system on the Gaussian beam. The *ABCD* law will be proved by verification in special cases, and its generality will ultimately be established by induction.

Transmission Through Free Space

When the optical system is a distance d of free space (or of any homogeneous medium), the elements of the ray-transfer matrix **M** are $A = 1, B = d, C = 0, D = 1$. Since $q = z + jz_0$ in free space, the q-parameter is modified by the optical system in accordance with $q_2 = q_1 + d = (1 \cdot q_1 + d)/(0 \cdot q_1 + 1)$, so that the *ABCD* law applies.

Transmission Through a Thin Optical Component

An arbitrary thin optical component does not affect the ray position, so that

$$y_2 = y_1, \tag{3.2-22}$$

but does alter the angle in accordance with

$$\theta_2 = Cy_1 + D\theta_1, \tag{3.2-23}$$

as illustrated in Fig. 3.2-10. Thus $A = 1$ and $B = 0$, but C and D are arbitrary. In all of the thin optical components described in Sec. 1.4B, however, $D = n_1/n_2$. Since the

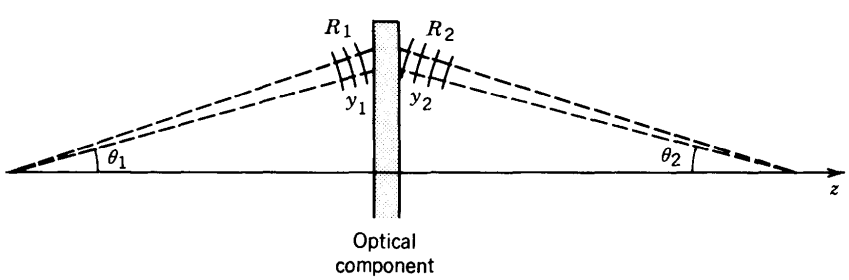

Figure 3.2-10 Modification of a Gaussian beam by a thin optical component.

optical component is thin, the beam width does not change, i.e.,

$$W_2 = W_1. \tag{3.2-24}$$

If the input and output beams are approximated by spherical waves of radii R_1 and R_2 at the input and output planes of the component, respectively, then in the paraxial approximation (small θ_1 and θ_2), $\theta_1 \approx y_1/R_1$ and $\theta_2 \approx y_2/R_2$. Substituting into (3.2-23), and using (3.2-22), we obtain

$$\frac{1}{R_2} = C + \frac{D}{R_1}. \tag{3.2-25}$$

Using (3.1-6), which is the expression for q as a function of R and W, and noting that $D = n_1/n_2 = \lambda_2/\lambda_1$, (3.2-24) and (3.2-25) can be combined into a single equation,

$$\frac{1}{q_2} = C + \frac{D}{q_1}, \tag{3.2-26}$$

from which $q_2 = (1 \cdot q_1 + 0)/(Cq_1 + D)$, so that the *ABCD* law also applies.

Invariance of the ABCD Law to Cascading

If the *ABCD* law is applicable to each of two optical systems with matrices $\mathbf{M}_i = (A_i, B_i, C_i, D_i)$, $i = 1, 2$, it must also apply to a system comprising their cascade (a system with matrix $\mathbf{M} = \mathbf{M}_2\mathbf{M}_1$). This may be shown by straightforward substitution.

Generality of the ABCD Law

Since the *ABCD* law applies to thin optical components and to propagation in a homogeneous medium, it also applies to any combination thereof. All of the paraxial optical systems of interest are combinations of propagation in homogeneous media and thin optical components such as thin lenses and mirrors. We therefore conclude that the *ABCD* law is applicable to all these systems. Since an inhomogeneous continuously varying medium may be regarded as a cascade of incremental thin elements followed by incremental distances, we conclude that the *ABCD* law applies to these systems as well, provided that all rays (wavefront normals) remain paraxial.

EXERCISE 3.2-5

Transmission of a Gaussian Beam Through a Transparent Plate. Use the *ABCD* law to examine the transmission of a Gaussian beam from air, through a transparent plate of refractive index n and thickness d, and again into air. Assume that the beam axis is normal to the plate.

3.3 HERMITE-GAUSSIAN BEAMS

The Gaussian beam is not the only beam-like solution of the paraxial Helmholtz equation (3.1-2). There are may other solutions including beams with non-Gaussian intensity distributions. Of particular interest are solutions that share the paraboloidal wavefronts of the Gaussian beam, but exhibit different intensity distributions. Beams of paraboloidal wavefronts are of importance since they match the curvatures of spherical mirrors of large radius. They can therefore reflect between two spherical mirrors that form a resonator, without being altered. Such self-reproducing waves are called the **modes** of the resonator. The optics of resonators is discussed in Chap. 9.

Consider a Gaussian beam of complex envelope

$$A_G(x, y, z) = \frac{A_1}{q(z)} \exp\left[-jk\frac{x^2 + y^2}{2q(z)}\right], \tag{3.3-1}$$

where $q(z) = z + jz_0$. The beam radius $W(z)$ is given by (3.1-8) and the wavefront radius of curvature $R(z)$ is given by (3.1-9). Consider a second wave whose complex envelope is a modulated version of the Gaussian beam,

$$A(x, y, z) = \mathscr{X}\left[\sqrt{2}\frac{x}{W(z)}\right]\mathscr{Y}\left[\sqrt{2}\frac{y}{W(z)}\right]\exp[j\mathscr{Z}(z)]A_G(x, y, z), \tag{3.3-2}$$

where $\mathscr{X}(\cdot)$, $\mathscr{Y}(\cdot)$, and $\mathscr{Z}(\cdot)$ are real functions. This wave, if it exists, has the following two properties:

- The phase is the same as that of the underlying Gaussian wave, except for an excess phase $\mathscr{Z}(z)$ that is independent of x and y. If $\mathscr{Z}(z)$ is a slowly varying function of z, the two waves have paraboloidal wavefronts with the same radius of curvature $R(z)$. These two waves are therefore focused by thin lenses and mirrors in precisely the same manner.
- The magnitude

$$A_0\mathscr{X}\left[\sqrt{2}\frac{x}{W(z)}\right]\mathscr{Y}\left[\sqrt{2}\frac{y}{W(z)}\right]\left[\frac{W_0}{W(z)}\right]\exp\left[-\frac{x^2 + y^2}{W^2(z)}\right],$$

where $A_0 = A_1/jz_0$, is a function of $x/W(z)$ and $y/W(z)$ whose widths in the x and y directions vary with z in accordance with the same scaling factor $W(z)$. As z increases, the intensity distribution in the transverse plane remains fixed, except for a magnification factor $W(z)$. This distribution is a Gaussian function modulated in the x and y directions by the functions $\mathscr{X}^2(\cdot)$ and $\mathscr{Y}^2(\cdot)$.

The modulated wave therefore represents a beam of non-Gaussian intensity distribution, but with the same wavefronts and angular divergence as the Gaussian beam.

The existence of this wave is assured if three real functions $\mathscr{X}(\cdot)$, $\mathscr{Y}(\cdot)$, and $\mathscr{Z}(z)$ can be found such that (3.3-2) satisfies the paraxial Helmholtz equation (3.1-2). Substituting (3.3-2) into (3.1-2), using the fact that A_G itself satisfies (3.1-2), and defining two new variables $u = \sqrt{2}\,x/W(z)$ and $v = \sqrt{2}\,y/W(z)$, we obtain

$$\frac{1}{\mathscr{X}}\left(\frac{\partial^2\mathscr{X}}{\partial u^2} - 2u\frac{\partial\mathscr{X}}{\partial u}\right) + \frac{1}{\mathscr{Y}}\left(\frac{\partial^2\mathscr{Y}}{\partial v^2} - 2v\frac{\partial\mathscr{Y}}{\partial v}\right) + kW^2(z)\frac{\partial\mathscr{Z}}{\partial z} = 0. \tag{3.3-3}$$

Since the left-hand side of this equation is the sum of three terms, each of which is a function of a single independent variable, u, v, or z, respectively, each of these terms must be constant. Equating the first term to the constant $-2\mu_1$ and the second to $-2\mu_2$, the third must be equal to $2(\mu_1 + \mu_2)$. This technique of "separation of variables" permits us to reduce the partial differential equation (3.3-3) into three ordinary differential equations for $\mathscr{X}(u)$, $\mathscr{Y}(v)$, and $\mathscr{Z}(z)$, respectively:

$$-\frac{1}{2}\frac{d^2\mathscr{X}}{du^2} + u\frac{d\mathscr{X}}{du} = \mu_1\mathscr{X} \tag{3.3-4a}$$

$$-\frac{1}{2}\frac{d^2\mathscr{Y}}{dv^2} + v\frac{d\mathscr{Y}}{dv} = \mu_2\mathscr{Y} \tag{3.3-4b}$$

$$z_0\left[1 + \left(\frac{z}{z_0}\right)^2\right]\frac{d\mathscr{Z}}{dz} = \mu_1 + \mu_2, \tag{3.3-4c}$$

where we have used the expression for $W(z)$ given in (3.1-8) and (3.1-11).

Equation (3.3-4a) represents an eigenvalue problem whose eigenvalues are $\mu_1 = l$, where $l = 0, 1, 2, \ldots$ and whose eigenfunctions are the **Hermite polynomials** $\mathscr{X}(u) = H_l(u)$, $l = 0, 1, 2, \ldots$. These polynomials are defined by the recurrence relation

$$H_{l+1}(u) = 2uH_l(u) - 2lH_{l-1}(u) \tag{3.3-5}$$

and

$$H_0(u) = 1, \qquad H_1(u) = 2u. \tag{3.3-6}$$

Thus

$$H_2(u) = 4u^2 - 2, \qquad H_3(u) = 8u^3 - 12u, \qquad \ldots . \tag{3.3-7}$$

Similarly, the solutions of (3.3-4b) are $\mu_2 = m$ and $\mathscr{Y}(v) = H_m(v)$, where $m = 0, 1, 2, \ldots$. There is therefore a family of solutions labeled by the indices (l, m).

Substituting $\mu_1 = l$ and $\mu_2 = m$ in (3.3-4c), and integrating, we obtain

$$\mathscr{Z}(z) = (l + m)\zeta(z), \tag{3.3-8}$$

where $\zeta(z) = \tan^{-1}(z/z_0)$. The excess phase $\mathscr{Z}(z)$ varies slowly between $-(l + m)\pi/2$ and $(l + m)\pi/2$, as z varies between $-\infty$ and ∞ (see Fig. 3.1-5).

We finally substitute into (3.3-2) to obtain an expression for the complex envelope of the beam labeled by the indices (l, m). Rearranging terms and multiplying by $\exp(-jkz)$ provides the complex amplitude

$$\boxed{\begin{aligned} U_{l,m}(x, y, z) &= A_{l,m}\left[\frac{W_0}{W(z)}\right]G_l\left[\frac{\sqrt{2}x}{W(z)}\right]G_m\left[\frac{\sqrt{2}y}{W(z)}\right] \\ &\times\exp\left[-jkz - jk\frac{x^2 + y^2}{2R(z)} + j(l + m + 1)\zeta(z)\right], \end{aligned}} \tag{3.3-9}$$

Hermite-Gaussian Beam Complex Amplitude

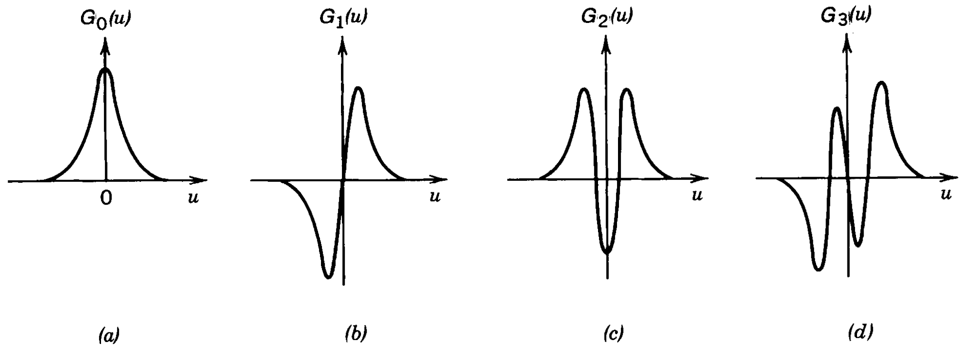

Figure 3.3-1 Several low-order Hermite-Gaussian functions: (*a*) $G_0(u)$; (*b*) $G_1(u)$; (*c*) $G_2(u)$; (*d*) $G_3(u)$.

where

$$G_l(u) = H_l(u)\exp\left(\frac{-u^2}{2}\right), \qquad l = 0, 1, 2, \ldots, \tag{3.3-10}$$

is known as the **Hermite–Gaussian function** of order l, and $A_{l,m}$ is a constant.

Since $H_0(u) = 1$, the Hermite–Gaussian function of order 0 is simply the Gaussian function. $G_1(u) = 2u\exp(-u^2/2)$ is an odd function, $G_2(u) = (4u^2 - 2)\exp(-u^2/2)$ is even, $G_3(u) = (8u^3 - 12u)\exp(-u^2/2)$ is odd, and so on. These functions are shown in Fig. 3.3-1.

An optical wave with complex amplitude given by (3.3-9) is known as the Hermite–Gaussian beam of order (l, m). The Hermite–Gaussian beam of order $(0, 0)$ is the Gaussian beam.

Intensity Distribution

The optical intensity of the (l, m) Hermite–Gaussian beam is

$$I_{l,m}(x, y, z) = |A_{l,m}|^2\left[\frac{W_0}{W(z)}\right]^2 G_l^2\left[\frac{\sqrt{2}\,x}{W(z)}\right] G_m^2\left[\frac{\sqrt{2}\,y}{W(z)}\right]. \tag{3.3-11}$$

Figure 3.3-2 illustrates the dependence of the intensity on the normalized transverse distances $u = \sqrt{2}\,x/W(z)$ and $v = \sqrt{2}\,y/W(z)$ for several values of l and m. Beams of higher order have larger widths than those of lower order as is evident from Fig. 3.3-1.

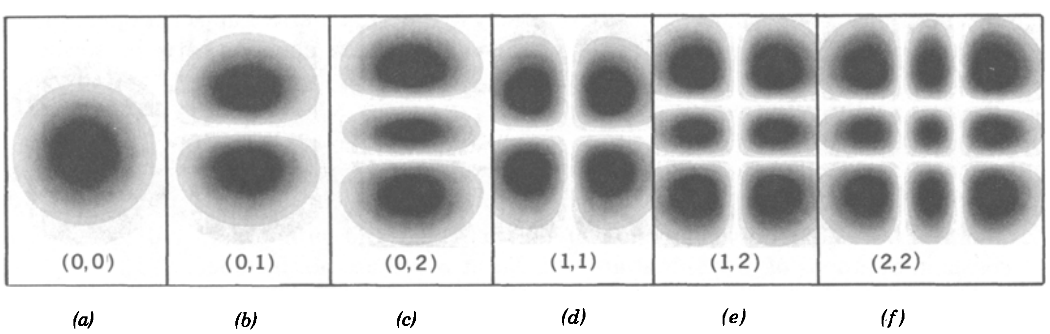

Figure 3.3-2 Intensity distributions of several low-order Hermite–Gaussian beams in the transverse plane. The order (l, m) is indicated in each case.

Regardless of the order, however, the width of the beam is proportional to $W(z)$, so that as z increases the intensity pattern is magnified by the factor $W(z)/W_0$ but otherwise maintains its profile. Among the family of Hermite–Gaussian beams, the only circularly symmetric member is the Gaussian beam.

EXERCISE 3.3-1

The Donut Beam. A wave is a superposition of two Hermite–Gaussian beams of orders $(1,0)$ and $(0,1)$ of equal intensities. The two beams have independent and random phases so that their intensities add with no interference. Show that the total intensity is a donut-shaped circularly symmetric function. Assuming that $W_0 = 1$ mm, determine the radius of the circle of peak intensity and the radii of the two circles of $1/e^2$ times the peak intensity at the beam waist.

*3.4 LAGUERRE–GAUSSIAN AND BESSEL BEAMS

Laguerre–Gaussian Beams

The Hermite–Gaussian beams form a complete set of solutions to the paraxial Helmholtz equation. Any other solution can be written as a superposition of these beams. But this family is not the only one. Another complete set of solutions, known as **Laguerre–Gaussian beams**, may be obtained by writing the paraxial Helmholtz equation in cylindrical coordinates (ρ, ϕ, z) and using separation of variables in ρ and ϕ, instead of x and y. The lowest-order Laguerre–Gaussian beam is the Gaussian beam.

Bessel Beams

In the search for beamlike waves, it is natural to examine the possibility of the existence of waves with planar wavefronts but with nonuniform intensity distributions in the transverse plane. Consider a wave with the complex amplitude

$$U(\mathbf{r}) = A(x, y)e^{-j\beta z}. \tag{3.4-1}$$

For this wave to satisfy the Helmholtz equation, $\nabla^2 U + k^2 U = 0$, $A(x, y)$ must satisfy

$$\nabla_T^2 A + k_T^2 A = 0, \tag{3.4-2}$$

where $k_T^2 + \beta^2 = k^2$ and $\nabla_T^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the transverse Laplacian operator. Equation (3.4-2), known as the two-dimensional Helmholtz equation, may be solved using the method of separation of variables. Using polar coordinates ($x = \rho\cos\phi$, $y = \rho\sin\phi$), the result is

$$A(x, y) = A_m J_m(k_T\rho)e^{jm\phi}, \qquad m = 0, \pm 1, \pm 2, \ldots, \tag{3.4-3}$$

where $J_m(\cdot)$ is the Bessel function of the first kind and mth order, and A_m is a constant. Solutions of (3.4-3) that are singular at $\rho = 0$ are not included.

For $m = 0$, the wave has a complex amplitude

$$U(\mathbf{r}) = A_0 J_0(k_T\rho)e^{-j\beta z} \tag{3.4-4}$$

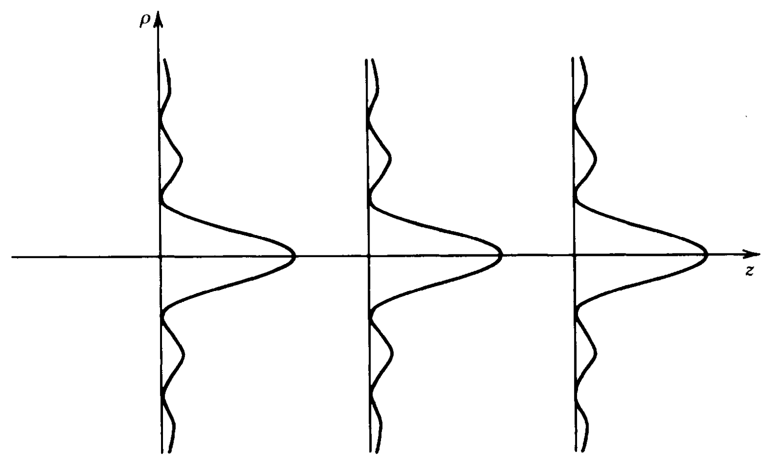

Figure 3.4-1 The intensity distribution of the Bessel beam in the transverse plane is independent of z; the beam does not diverge.

and therefore has planar wavefronts. The wavefront normals (rays) are all parallel to the z axis. The intensity distribution $I(\rho, \phi, z) = |A_0|^2 J_0^2(k_T \rho)$ is circularly symmetric, varies with ρ as illustrated in Fig. 3.4-1, and is independent of z, so that there is no spread of the optical power. This wave is called the **Bessel beam**.

It is interesting to compare the Bessel beam to the Gaussian beam. Whereas the complex amplitude of the Bessel beam is an *exact* solution of the Helmholtz equation, the complex amplitude of the Gaussian beam is only an approximate solution (its complex envelope is an exact solution of the paraxial Helmholtz equation, however). The intensity distribution of these two beams are compared in Fig. 3.4-2. The asymptotic behavior of these distributions in the limit of large radial distances is significantly different. Whereas the intensity of the Gaussian beam decreases exponentially in proportionality to $\exp[-2\rho^2/W^2(z)]$, the intensity of the Bessel beam is proportional to $J_0^2(k_T\rho) \approx (2/\pi k_T \rho)\cos^2(k_T\rho - \pi/4)$, which is an oscillatory function with slowly decaying magnitude. Whereas the rms width of the Gaussian beam, $\sigma = \frac{1}{2}W(z)$, is finite, the rms width of the Bessel beam is *infinite* at all z (see

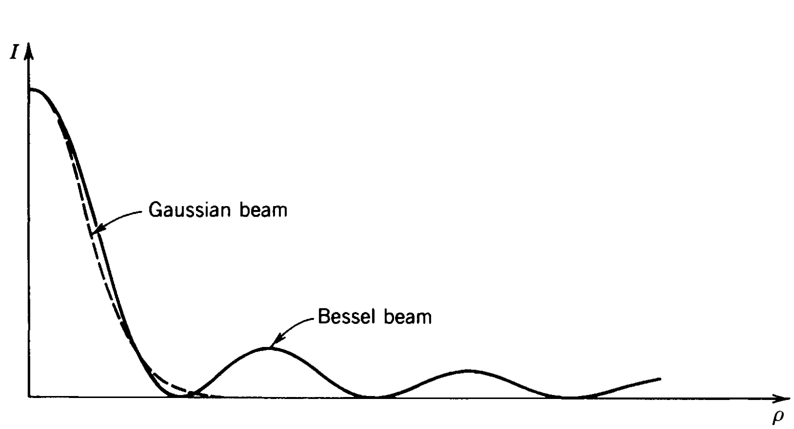

Figure 3.4-2 Comparison of the radial intensity distributions of a Gaussian beam and a Bessel beam. Parameters are selected such that the peak intensities and $1/e^2$ widths are identical in both cases.

Appendix A, Sec. A.2 for the definition of rms width). There is a tradeoff between the minimum beam size and the divergence. Thus although the divergence of the Bessel beam is zero, its rms width is infinite. The generation of Bessel beams requires special schemes.[†] Since Gaussian beams are the modes of spherical resonators, they are created naturally by lasers.

READING LIST

Books with Chapters on Optical Beams

A. Yariv, *Quantum Electronics*, Wiley, New York, 1967, 3rd ed. 1989.

J. T. Verdeyen, *Laser Electronics*, Prentice-Hall, Englewood Cliffs, NJ, 1981, 2nd ed. 1989.

P. W. Milonni and J. H. Eberly, *Lasers*, Wiley, New York, 1988.

W. Witteman, *The Laser*, Springer-Verlag, New York, 1987.

A. E. Siegman, *Lasers*, University Science Books, Mill Valley, CA, 1986.

K. Shimoda, *Introduction to Laser Physics*, Springer-Verlag, New York, 2nd ed. 1986.

S. Solimeno, B. Crosignani, and P. DiPorto, *Guiding, Diffraction and Confinement of Optical Radiation*, Academic Press, New York, 1986.

A. Yariv, *Optical Electronics*, Holt, Rinehart and Winston, New York, 1971, 3rd ed. 1985.

D. C. O'Shea, *Elements of Modern Optical Design*, Wiley, New York, 1985.

D. Marcuse, *Light Transmission Optics*, Van Nostrand Reinhold, New York, 1972, 2nd ed. 1982.

M. S. Sodha and A. K. Ghatak, *Inhomogeneous Optical Waveguides*, Plenum Press, New York, 1977.

J. A. Arnaud, *Beam and Fiber Optics*, Academic Press, New York, 1976.

A. E. Siegman, *An Introduction to Lasers and Masers*, McGraw-Hill, New York, 1971.

Special Journal Issue

Special issue on propagation and scattering of beam fields, *Journal of the Optical Society of America A*, vol. 3, no. 4, 1986.

Articles

H. Kogelnik and T. Li, Laser Beams and Resonators, *Proceedings of the IEEE*, vol. 54, pp. 1312–1329, 1966.

A. G. Fox and T. Li, Resonant Modes in a Maser Interferometer, *Bell System Technical Journal*, vol. 40, pp. 453–488, 1961.

G. D. Boyd and J. P. Gordon, Confocal Multimode Resonator for Millimeter Through Optical Wavelength Masers, *Bell System Technical Journal*, vol. 40, pp. 489–508, 1961.

PROBLEMS

3.1-1 **Beam Parameters.** The light from a Nd:YAG laser at wavelength 1.06 μm is a Gaussian beam of 1-W optical power and beam divergence $2\theta_0 = 1$ mrad. Determine the beam waist radius, the depth of focus, the maximum intensity, and the intensity on the beam axis at a distance $z = 100$ cm from the beam waist.

3.1-2 **Beam Identification by Two Widths.** A Gaussian beam of wavelength $\lambda_o = 10.6$ μm (emitted by a CO_2 laser) has widths $W_1 = 1.699$ mm and $W_2 = 3.38$ mm at two points separated by a distance $d = 10$ cm. Determine the location of the waist and the waist radius.

[†]See P. W. Milonni and J. H. Eberly, *Lasers*, Wiley, New York, 1988, Sec. 14.14.

3.1-3 **The Elliptic Gaussian Beam.** The paraxial Helmholtz equation admits a Gaussian beam with intensity $I(x, y, 0) = |A_0|^2 \exp[-2(x^2/W_{0x}^2 + y^2/W_{0y}^2)]$ in the $z = 0$ plane, with beam waist radii W_{0x} and W_{0y} in the x and y-directions respectively. The contours of constant intensity are therefore ellipses instead of circles. Write expressions for the beam depth of focus, angular divergence, and radii of curvature in the x and y directions, as functions of W_{0x}, W_{0y}, and the wavelength λ. If $W_{0x} = 2W_{0y}$, sketch the shape of the beam spot in the $z = 0$ plane and in the far field (z much greater than the depths of focus in both transverse directions).

3.2-1 **Beam Focusing.** An argon-ion laser produces a Gaussian beam of wavelength $\lambda = 488$ nm and waist radius $W_0 = 0.5$ mm. Design a single-lens optical system for focusing the light to a spot of diameter 100 μm. What is the shortest focal-length lens that may be used?

3.2-2 **Spot Size.** A Gaussian beam of Rayleigh range $z_0 = 50$ cm and wavelength $\lambda = 488$ nm is converted into a Gaussian beam of waist radius W_0' using a lens of focal length $f = 5$ cm at a distance z from its waist, as illustrated in Fig. 3.2-2. Write a computer program to plot W_0' as a function of z. Verify that in the limit $z - f \gg z_0$, (3.2-10) and (3.2-12) hold; and in the limit $z \ll z_0$ (3.2-13) holds.

3.2-3 **Beam Refraction.** A Gaussian beam is incident from air ($n = 1$) into a medium with a planar boundary and refractive index $n = 1.5$. The beam axis is normal to the boundary and the beam waist lies at the boundary. Sketch the transmitted beam. If the angular divergence of the beam in air is 1 mrad, what is the angular divergence in the medium?

*3.2-4 **Transmission of a Gaussian Beam Through a Graded-Index Slab.** The *ABCD* matrix of a SELFOC graded-index slab with quadratic refractive index (see Sec. 1.3B) $n(y) \approx n_0(1 - \frac{1}{2}\alpha^2 y^2)$ and length d is: $A = \cos \alpha d$, $B = (1/\alpha)\sin \alpha d$, $C = -\alpha \sin \alpha d$, $D = \cos \alpha d$ for paraxial rays along the z direction. A Gaussian beam of wavelength λ_o, waist radius W_0 in free space, and axis in the z direction enters the slab at its waist. Use the *ABCD* law to determine an expression for the beam width in the y direction as a function of d. Sketch the shape of the beam as it travels through the medium.

3.3-1 **Power Confinement in Hermite–Gaussian Beams.** Determine the ratio of the power contained within a circle of radius $W(z)$ in the transverse plane to the total power in the Hermite–Gaussian beams of orders (0, 0), (1, 0), (0, 1), and (1, 1). What is the ratio of the power contained within a circle of radius $W(z)/10$ to the total power for the (0, 0) and (1, 1) Hermite–Gaussian beams?

3.3-2 **Superposition of Two Beams.** Sketch the intensity of a superposition of the (1, 0) and (1, 0) Hermite–Gaussian beams assuming that the complex coefficients $A_{1,0}$ and $A_{0,1}$ in (3.3-9) are equal.

3.3-3 **Axial Phase.** Consider the Hermite–Gaussian beams of all orders (l, m) and Rayleigh range $z_0 = 30$ cm in a medium of refractive index $n = 1$. Determine the frequencies within the band $\nu = 10^{14} \pm 2 \times 10^9$ Hz for which the phase retardation between the planes $z = -z_0$ and $z = z_0$ is an integer multiple of π on the beam axis. These frequencies are the modes of a resonator made of two spherical mirrors placed at the $z = \pm z_0$ planes, as described in Sec. 9.2D.