

2

Gaussian Beam Propagation

Gaussian beams play such an important role in optical lasers as well as in longer wavelength systems that they have been extensively analyzed, starting with some of the classic treatments mentioned in Chapter 1. Almost every text on optical systems discusses Gaussian beam propagation in some detail, and several comprehensive review articles are available. However, for millimeter and submillimeter wavelength systems there are naturally certain aspects that deserve special attention, and we emphasize aspects of quasioptical propagation that have proven to be of greatest importance at these relatively long wavelengths.

In the following sections we first give a derivation of Gaussian beam formulas based on the paraxial wave equation, in cylindrical and in rectangular coordinates. We discuss normalization, beam truncation, and interpretation of the Gaussian beam propagation formulas. We next cover higher order modes in different coordinate systems and consider the effective size of Gaussian beam modes. We then present inverse formulas for Gaussian beam propagation, which are of considerable use in system design. Finally, we consider the paraxial approximation in more detail and present an alternative derivation of Gaussian beam propagation based on diffraction integrals.

2.1 DERIVATION OF BASIC GAUSSIAN BEAM PROPAGATION

2.1.1 The Paraxial Wave Equation

Only in very special cases does the propagation of an electromagnetic wave result in a distribution of field amplitudes that is independent of position: the most familiar example is a plane wave. If we restrict the region over which there is initially a nonzero field, wave propagation becomes a problem of diffraction, which in its most general form is an extremely complex vector problem. We treat here a simplified problem encountered when a beam of

radiation that is largely collimated; that is, it has a well-defined direction of propagation but has also some transverse variation (unlike in a plane wave). We thus develop the **paraxial wave equation**, which forms the basis for Gaussian beam propagation. Thus, a Gaussian beam does have limited transverse variation compared to a plane wave. It is different from a beam originating from a source in geometrical optics in that it originates from a region of finite extent, rather than from an infinitesimal **point source**.

A single component, ψ , of an electromagnetic wave propagating in a uniform medium satisfies the Helmholtz (wave) equation

$$(\nabla^2 + k^2)\psi = 0, \quad (2.1)$$

where ψ represents any component of \mathbf{E} or \mathbf{H} . We have assumed a time variation at angular frequency ω of the form $\exp(j\omega t)$. The wave number k is equal to $2\pi/\lambda$, so that $k = \omega(\epsilon_r\mu_r)^{0.5}/c$, where ϵ_r and μ_r are the relative permittivity and permeability of the medium, respectively. For a plane wave, the amplitudes of the electric and magnetic fields are constant; and their directions are mutually perpendicular, and perpendicular to the propagation vector. For a beam of radiation that is similar to a plane wave but for which we will allow some variation perpendicular to the axis of propagation, we can still assume that the electric and magnetic fields are (mutually perpendicular and) perpendicular to the direction of propagation. Letting the direction of propagation be in the positive z direction, we can write the distribution for any component of the electric field (suppressing the time dependence) as

$$E(x, y, z) = u(x, y, z) \exp(-jkz), \quad (2.2)$$

where u is a complex scalar function that defines the non-plane wave part of the beam. In rectangular coordinates, the Helmholtz equation is

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} + k^2 E = 0. \quad (2.3)$$

If we substitute our quasi-plane wave solution, we obtain

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - 2jk \frac{\partial u}{\partial z} = 0, \quad (2.4)$$

which is sometimes called the **reduced wave equation**.

The paraxial approximation consists of assuming that the variation along the direction of propagation of the amplitude u (due to diffraction) will be small over a distance comparable to a wavelength, and that the axial variation will be small compared to the variation perpendicular to this direction. The first statement implies that (in magnitude) $[\Delta(\partial u/\partial z)/\Delta z]\lambda \ll \partial u/\partial z$, which enables us to conclude that the third term in equation 2.4 is small compared to the fourth term. The second statement allows us to conclude that the third term is small compared to the first two. Consequently, we may drop the third term, obtaining finally the **paraxial wave equation** in rectangular coordinates

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2jk \frac{\partial u}{\partial z} = 0. \quad (2.5)$$

Solutions to the paraxial wave equation are the Gaussian beam modes that form the basis of quasioptical system design. There is no rigorous ‘‘cutoff’’ for the application of the paraxial approximation, but it is generally reasonably good as long as the angular divergence of the beam is confined (or largely confined) to within 0.5 radian (or about 30 degrees) of the

z axis. Errors introduced by the paraxial approximation are shown explicitly by [MART93]; extension beyond the paraxial approximation is further discussed in Section 2.8, and other references can be found there.

2.1.2 The Fundamental Gaussian Beam Mode Solution in Cylindrical Coordinates

Solutions to the paraxial wave equation can be obtained in various coordinate systems; in addition to the rectangular coordinate system used above, the axial symmetry that characterizes many situations encountered in practice (e.g., corrugated feed horns and lenses) makes cylindrical coordinates the natural choice. In cylindrical coordinates, r represents the perpendicular distance from the axis of propagation, taken again to be the z axis, and the angular coordinate is represented by φ . In this coordinate system the paraxial wave equation is

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r} \frac{\partial^2 u}{\partial \varphi^2} - 2jk \frac{\partial u}{\partial z} = 0, \quad (2.6)$$

where $u \equiv u(r, \varphi, z)$. For the moment, we will assume axial symmetry, that is, u is independent of φ , which makes the third term in equation 2.6 equal to zero, whereupon we obtain the **axially symmetric paraxial wave equation**

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} - 2jk \frac{\partial u}{\partial z} = 0. \quad (2.7)$$

From prior work, we note that the simplest solution of the axially symmetric paraxial wave equation can be written in the form

$$u(r, z) = A(z) \exp \left[\frac{-jkr^2}{2q(z)} \right], \quad (2.8)$$

where A and q are two complex functions (of z only), which remain to be determined. Obviously, this expression for u looks something like a Gaussian distribution. To obtain the unknown terms in equation 2.8, we substitute this expression for u into the axially symmetric paraxial wave equation 2.7 and obtain

$$-2jk \left(\frac{A}{q} + \frac{\partial A}{\partial z} \right) + \frac{k^2 r^2 A}{q^2} \left(\frac{\partial q}{\partial z} - 1 \right) = 0. \quad (2.9)$$

Since this equation must be satisfied for all r as well as all z , and given that the first part depends only on z while the second part depends on r and z , the two parts must individually be equal to zero. This gives us two relationships that must be simultaneously satisfied:

$$\frac{\partial q}{\partial z} = 1 \quad (2.10a)$$

and

$$\frac{\partial A}{\partial z} = -\frac{A}{q}. \quad (2.10b)$$

Equation 2.10a has the solution

$$q(z) = q(z_0) + (z - z_0). \quad (2.11a)$$

Without loss of generality, we define the reference position along the z axis to be $z_0 = 0$, which yields

$$q(z) = q(0) + z. \quad (2.11b)$$

The function q is called the **complex beam parameter** (since it is complex), but it is often referred to simply as the **beam parameter** or **Gaussian beam parameter**. Since it appears in equation 2.8 as $1/q$, it is reasonable to write

$$\frac{1}{q} = \left(\frac{1}{q}\right)_r - j \left(\frac{1}{q}\right)_i, \quad (2.12)$$

where the subscripted terms are the real and imaginary parts of the quantity $1/q$, respectively. Substituting into equation 2.8, the exponential term becomes

$$\exp\left(\frac{-jkr^2}{2q}\right) = \exp\left[\left(\frac{-jkr^2}{2}\right) \left(\frac{1}{q}\right)_r - \left(\frac{kr^2}{2}\right) \left(\frac{1}{q}\right)_i\right]. \quad (2.13)$$

The imaginary term has the form of the phase variation produced by a spherical wave front in the paraxial limit. We can see this starting with an equiphase surface having **radius of curvature** R and defining $\phi(r)$ to be the **phase variation relative to a plane for a fixed value of z as a function of r** as shown in Figure 2.1. In the limit $r \ll R$, the phase delay incurred is approximately equal to

$$\phi(r) \cong \frac{\pi r^2}{\lambda R} = \frac{kr^2}{2R}. \quad (2.14)$$

We thus make the important identification of the real part of $1/q$ with the radius of curvature of the beam

$$\left(\frac{1}{q}\right)_r = \frac{1}{R}. \quad (2.15)$$

Since q is a function of z , it is evident that the radius of curvature of the beam will depend on the position along the axis of propagation. It is important not to confuse the phase shift ϕ (which we shall see depends on z) with the azimuthal coordinate φ .

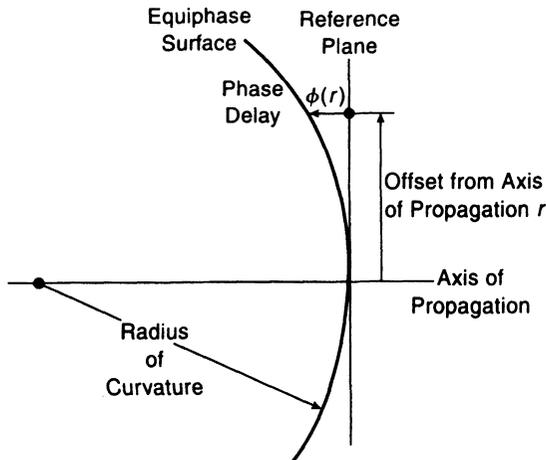


Figure 2.1 Phase shift of spherical wave relative to plane wave. The phase delay of the spherical wave, at distance r from axis defined by propagation direction of plane wave, is $\phi(r)$.

The second part of the exponential in equation 2.13 is real and has a Gaussian variation as a function of the distance from the axis of propagation. Taking the standard form for a

Gaussian distribution to be

$$f(r) = f(0) \exp \left[- \left(\frac{r}{r_0} \right)^2 \right], \quad (2.16)$$

we see that the quantity r_0 represents the distance to the $1/e$ point relative to the on-axis value. To make the second part of equation 2.13 have this form we take

$$\left(\frac{1}{q} \right)_i = \frac{2}{kw^2(z)} = \frac{\lambda}{\pi w^2}, \quad (2.17)$$

and thus define the **beam radius** w , which is the value of the radius at which the field falls to $1/e$ relative to its on-axis value. Since q is a function of z , the beam radius as well as the radius of curvature will depend on the position along the axis of propagation.

With these definitions, we see that the function q is given by

$$\frac{1}{q} = \frac{1}{R} - \frac{j\lambda}{\pi w^2}, \quad (2.18)$$

where both R and w are functions of z .

At $z = 0$ we have from equation 2.8, $u(r, 0) = A(0) \exp[-jkr^2/2q(0)]$, and if we choose w_0 such that $w_0 = [\lambda q(0)/j\pi]^{0.5}$, we find the relative field distribution at $z = 0$ to be

$$u(r, 0) = u(0, 0) \exp \left(\frac{-r^2}{w_0^2} \right), \quad (2.19)$$

where w_0 denotes the beam radius at $z = 0$, which is called the **beam waist radius**. With this definition, we obtain from equation 2.11b a second important expression for q :

$$q = \frac{j\pi w_0^2}{\lambda} + z. \quad (2.20)$$

Equations 2.18 and 2.20 together allow us to obtain the radius of curvature and the beam radius as a function of position along the axis of propagation:

$$R = z + \frac{1}{z} \left(\frac{\pi w_0^2}{\lambda} \right)^2 \quad (2.21a)$$

$$w = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{0.5} \quad (2.21b)$$

We see that the the beam waist radius is the minimum value of the beam radius and that it occurs at the beam waist, where the radius of curvature is infinite, characteristic of a plane wave front. The transverse spreading of a Gaussian beam as it propagates, together with drop in on-axis amplitude, are illustrated in Figure 2.2a, while the behavior of the radius of curvature is shown schematically in Figure 2.2b. The relationships given in equations 2.21a and 2.21b are fundamental for Gaussian beam propagation, and we will return to them in subsequent sections. In particular, the quantity $\pi w_0^2/\lambda$, called the **confocal distance**, plays a prominent role and is discussed further in Section 2.2.4.

To complete our analysis of the basic Gaussian beam equation, we must use the second of the pair of equations obtained from substituting our trial solution in the paraxial wave equation. Rewriting equation 2.10b, we find $dA/A = -dz/q$, and from equation 2.10a we

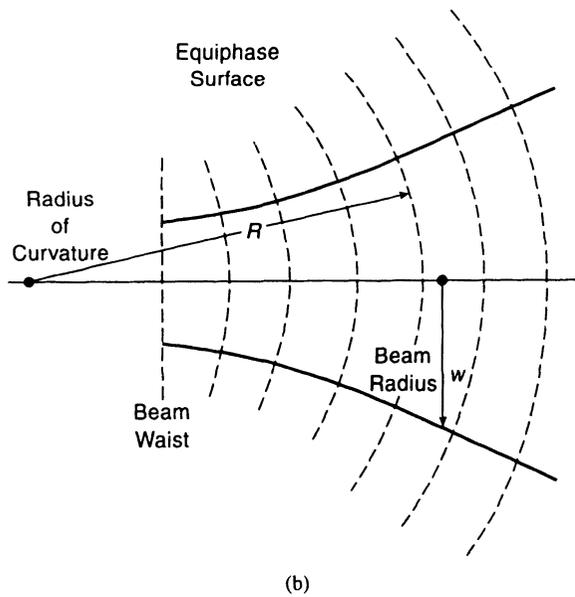
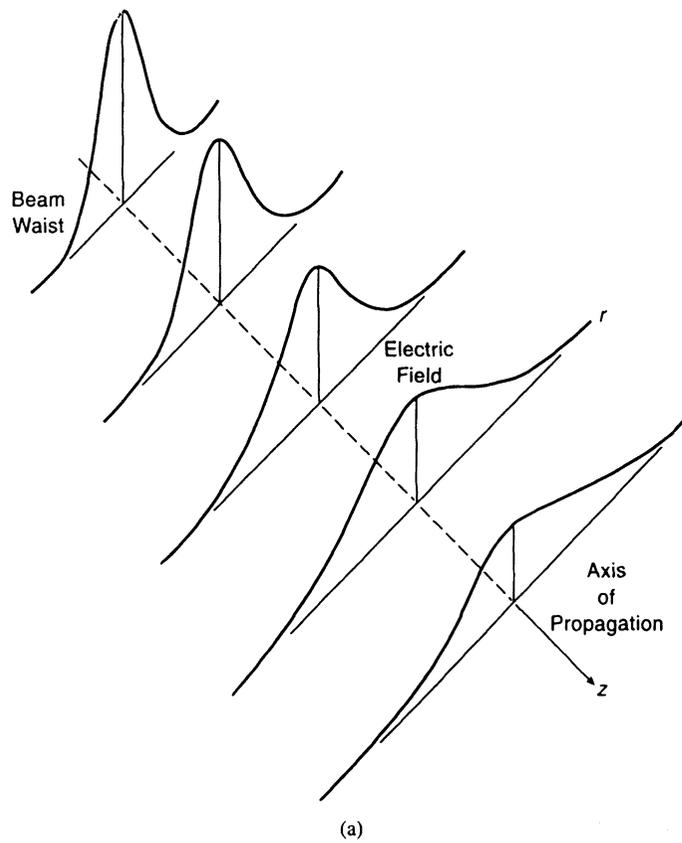


Figure 2.2 Schematic diagram of Gaussian beam propagation. (a) Propagating beam indicating increase in beam radius and diminution of peak amplitude as distance from waist increases. (b) Cut through beam showing equiphase surfaces (broken lines), beam radius w , and radius of curvature R .

have $dz = dq$ so that we can write $dA/A = -dq/q$. Hence, $A(z)/A(0) = q(0)/q(z)$, and substituting q from equation 2.20, we find

$$\frac{A(z)}{A(0)} = \frac{1 + j\lambda z/\pi w_0^2}{1 + (\lambda z/\pi w_0^2)^2}. \quad (2.22)$$

It is convenient to express this in terms of a phasor, and defining

$$\tan \phi_0 = \frac{\lambda z}{\pi w_0^2}, \quad (2.23)$$

we see that

$$\frac{A(z)}{A(0)} = \frac{w_0}{w} \exp(j\phi_0). \quad (2.24)$$

The **Gaussian beam phase shift**, ϕ_0 , also is discussed in more detail below. If we take the amplitude on-axis at the beam waist to be unity, we have the complete expression for the fundamental Gaussian beam mode

$$u(r, z) = \frac{w_0}{w} \exp\left(\frac{-r^2}{w^2} - \frac{j\pi r^2}{\lambda R} + j\phi_0\right). \quad (2.25a)$$

The expression for the electric field can be obtained immediately using equation 2.2, and differs only owing to the plane wave phase factor, so we find

$$E(r, z) = \left(\frac{w_0}{w}\right) \exp\left(\frac{-r^2}{w^2} - jkz - \frac{j\pi r^2}{\lambda R} + j\phi_0\right), \quad (2.25b)$$

with the variation in w , R , and ϕ_0 as a function of z being given by equations 2.21 and 2.23.

2.1.3 Normalization

To relate the expression for the electric field given above to the total power in a propagating Gaussian beam, we assume (again in the paraxial limit) that the electric and magnetic field components are related to each other like those in a plane wave. Thus, the total power is proportional to the square of the electric field integrated over the area of the beam. A convenient normalization is to set the integral (extending from radius 0 to ∞) to unity, namely, $\int |E|^2 \cdot 2\pi r dr = 1$. Using the electric field distribution from equation 2.25b, we find that this integral, evaluated at $z = 0$, gives $\pi w_0^2/2$. Consequently, the normalized electric field distribution at any distance along the axis of propagation is given by

$$E(r, z) = \left(\frac{2}{\pi w^2}\right)^{0.5} \exp\left(\frac{-r^2}{w^2} - jkz - \frac{j\pi r^2}{\lambda R} + j\phi_0\right). \quad (2.26a)$$

Relating this numerically to the power flow depends on the system of units employed. The normalized form for the electric field distribution will be that used here, unless otherwise indicated. Together with the equations

$$R = z + \frac{1}{z} \left(\frac{\pi w_0^2}{\lambda}\right)^2, \quad (2.26b)$$

$$w = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2}\right)^2\right]^{0.5}, \quad (2.26c)$$

$$\tan \phi_0 = \frac{\lambda z}{\pi w_0^2}, \quad (2.26d)$$

we have completely described the behavior of the fundamental Gaussian beam mode that satisfies the paraxial wave equation.

2.1.4 Fundamental Gaussian Beam Mode in Rectangular Coordinates: One Dimension

It is possible to consider a beam that has variation in one coordinate perpendicular to the axis of propagation but is uniform in the other coordinate. Then, the paraxial wave equation (equation 2.5) for variation along the x axis only reduces to

$$\frac{\partial^2 u}{\partial x^2} - 2jk \frac{\partial u}{\partial z} = 0. \quad (2.27)$$

A trial solution of the form $u(x, z) = A_x(z) \exp[-jkx^2/2q_x(z)]$ together with the requirement that the solution be valid for all values of x and z , leads to the conditions

$$\frac{\partial q_x}{\partial z} = 1 \quad (2.28a)$$

and

$$\frac{\partial A_x}{\partial z} = -\frac{1}{2} \frac{A_x}{q_x}. \quad (2.28b)$$

The first of this pair of equations is identical to equation 2.10a, suggesting a solution similar to that used before (equation 2.20)

$$q_x = \frac{j\pi w_{0x}^2}{\lambda} + z, \quad (2.29a)$$

and we find this to be an appropriate choice. This leads to analogous definitions of the real and imaginary parts of q_x

$$\frac{1}{q_x} = \frac{1}{R_x} - \frac{j\lambda}{\pi w_x^2}, \quad (2.29b)$$

and we find that the solution has the same form as in the axially symmetric case, in terms of beam radius, radius of curvature, and the variation of w_x and R_x as a function of distance along the axis of propagation. The solution to equation 2.28b has the form $A_x(z)/A(0) = [q_x(0)/q_x(z)]^{0.5}$. The real part of the solution now has a square root dependence on w , as is appropriate for variation in one dimension, and a phase shift half as large as in the preceding case. The normalized form of the electric field distribution is

$$E(x, z) = \left(\frac{2}{\pi w_x^2} \right)^{0.25} \exp \left(-\frac{x^2}{w_x^2} - jkz - \frac{j\pi x^2}{\lambda R_x} + \frac{j\phi_{0x}}{2} \right), \quad (2.30)$$

with ϕ_{0x} defined analogously to ϕ_0 in equation 2.26 and the variation of R_x , w_x , and ϕ_{0x} given by equations 2.26b through 2.26d.

2.1.5 Fundamental Gaussian Beam Mode in Rectangular Coordinates: Two Dimensions

We use a similar approach to solve the paraxial wave equation in this case, employing a trial solution of the form $u(x, y, z) = A_x(z)A_y(z) \exp(-jkx^2/2q_x) \exp(-jky^2/2q_y)$. This form is motivated by our desire to keep the solution independent in the two orthogonal

coordinates. The solution separates, and with the requirement that it be valid independently for all x and y , we obtain the conditions

$$\frac{\partial q_x}{\partial z} = 1 \quad \text{and} \quad \frac{\partial q_y}{\partial z} = 1, \quad (2.31a)$$

together with

$$\frac{\partial A_x}{\partial z} = -\frac{1}{2} \frac{A_x}{q_x} \quad \text{and} \quad \frac{\partial A_y}{\partial z} = -\frac{1}{2} \frac{A_y}{q_y}. \quad (2.31b)$$

The field distribution is just the product of x and y portions, and the normalized form is

$$E(x, y, z) = \left(\frac{2}{\pi w_x w_y} \right)^{0.5} \cdot \exp \left(-\frac{x^2}{w_x^2} - \frac{y^2}{w_y^2} - \frac{j\pi x^2}{\lambda R_x} - \frac{j\pi y^2}{\lambda R_y} + \frac{j\phi_{0x}}{2} + \frac{j\phi_{0y}}{2} \right), \quad (2.32a)$$

where

$$w_x = w_{0x} \left[1 + \left(\frac{\lambda z}{\pi w_{0x}^2} \right)^2 \right]^{0.5}, \quad (2.32b)$$

$$w_y = w_{0y} \left[1 + \left(\frac{\lambda z}{\pi w_{0y}^2} \right)^2 \right]^{0.5}, \quad (2.32c)$$

$$R_x = z + \frac{1}{z} \left(\frac{\pi w_{0x}^2}{\lambda} \right)^2, \quad (2.32d)$$

$$R_y = z + \frac{1}{z} \left(\frac{\pi w_{0y}^2}{\lambda} \right)^2, \quad (2.32e)$$

$$\phi_{0x} = \tan^{-1} \left(\frac{\lambda z}{\pi w_{0x}^2} \right), \quad (2.32f)$$

$$\phi_{0y} = \tan^{-1} \left(\frac{\lambda z}{\pi w_{0y}^2} \right). \quad (2.32g)$$

In addition to the independence of the beam waist radii along the orthogonal coordinates, we can choose the reference positions along the z axis, for the complex beam parameters q_x and q_y , to be different (which is just equivalent to adding an arbitrary relative phase shift). The critical parameters describing variation of the Gaussian beam in the two directions perpendicular to its axis of propagation are entirely independent. This means that we can deal with asymmetric Gaussian beams, if these are appropriate to the situation, and we can consider focusing (transformation) of a Gaussian beam along a single axis independent of its variation in the orthogonal direction.

In the special case that (1) the beam waist radii w_{0x} and w_{0y} are equal and (2) the beam waist radii are located at the same value of z , we regain the symmetric fundamental mode

Gaussian beam (e.g., for $w_0 = w_{0x} = w_{0y}$, $R = R_x = R_y$); and noting that $r^2 = x^2 + y^2$, we see that equation 2.32 becomes identical to equation 2.26.

2.2 DESCRIPTION OF GAUSSIAN BEAM PROPAGATION

2.2.1 Concentration of the Fundamental Mode Gaussian Beam Near the Beam Waist

The field distribution and the power density of the fundamental Gaussian beam mode are both maximum on the axis of propagation ($r = 0$) at the beam waist ($z = 0$). As indicated by equation 2.26a, the field amplitude and power density diminish as z and r vary from zero. Figure 2.3 shows contours of power density relative to maximum value. The power density always drops monotonically as a function of r for fixed z , reflecting its Gaussian form. For $r/w_0 \leq 1/\sqrt{2}$, the relative power density decreases monotonically as z increases. For any fixed value of $r > w_0/\sqrt{2}$ corresponding to $p_{rel} < e^{-1}$, there is a maximum as a function of z , which occurs at $z = (\pi w_0^2/\lambda)[2(r/w_0)^2 - 1]^{0.5}$. This maximum, which results in the “dog bone” shape of the lower contours in the figure, is a consequence of the enhancement of the power density at a fixed distance from the axis of propagation that is due to the broadening of the beam (cf. [MOOS91]).

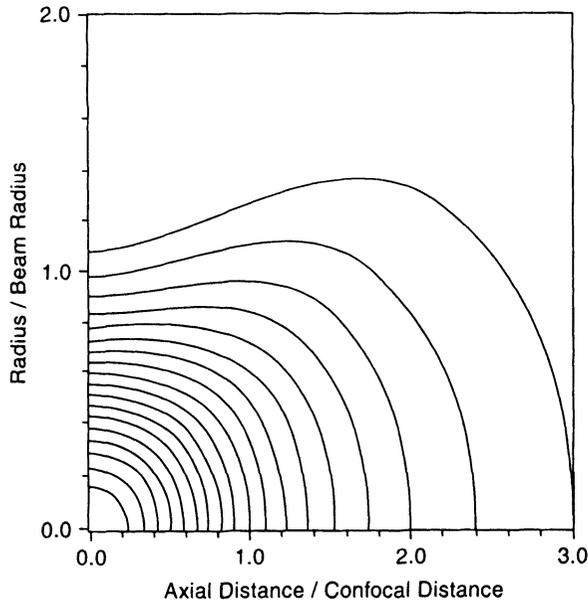


Figure 2.3 Contours of relative power density in propagating Gaussian beam normalized to peak on the axis of propagation ($r = 0$) at the beam waist ($z = 0$). The contours are at values 0.10, 0.15, 0.20, 0.25, ... relative to the maximum value, which reflect the diminution of on-axis peak power density and increasing beam radius as the beam propagates from the beam waist.

2.2.2 Fundamental Mode Gaussian Beam and Edge Taper

The fundamental Gaussian beam mode (described by equations 2.26, 2.30, or 2.32 depending on the coordinate system) has a Gaussian distribution of the electric field per-

pendicular to the axis of propagation, and at all distances along this axis:

$$\frac{|E(r, z)|}{|E(0, z)|} = \exp \left[- \left(\frac{r}{w} \right)^2 \right], \quad (2.33a)$$

where r is the distance from the propagation axis. The distribution of power density is proportional to this quantity squared:

$$\frac{P(r)}{P(0)} = \exp \left[-2 \left(\frac{r}{w} \right)^2 \right], \quad (2.33b)$$

and is likewise a Gaussian, which is an extremely convenient feature but one that can lead to some confusion. Since the basic description of the Gaussian beam mode is in terms of its electric field distribution, it is most natural to use the width of the field distribution to characterize the beam, although it is true that the power distribution is more often directly measured. The latter consideration has led some authors to define the Gaussian beam in terms of the width of the distribution of the power (cf. [ARNA76]), but we will use the quantity w throughout this book to denote the distance from the propagation axis at which the field has fallen to $1/e$ of its on-axis value.

It is straightforward to characterize the fundamental mode Gaussian beam in terms of the relative power level at a specified radius. The **edge taper** T_e is the relative power density at a radius r_e , which is given by

$$T_e = \frac{P(r_e)}{P(0)}. \quad (2.34a)$$

With the power distribution given by equation 2.33b we see that

$$T_e(r_e) = \exp \left(\frac{-2r_e^2}{w^2} \right). \quad (2.34b)$$

The edge taper is often expressed in decibels to accommodate efficiently a large dynamic range, with

$$T_e \text{ (dB)} = -10 \log_{10}(T_e). \quad (2.35a)$$

The fundamental mode Gaussian of the electric field distribution in linear coordinates and the power distribution in logarithmic form are shown in Figure 2.4.

The edge radius of a beam is obtained from the edge taper (or the radius from any specified power level relative to that on the axis of propagation) using

$$\frac{r_e}{w} = 0.3393 [T_e \text{ (dB)}]^{0.5}. \quad (2.35b)$$

Some reference values are provided in Table 2.1. Note that the full width to half-maximum (fwhm) of the beam is just twice the radius for 3 dB taper, which is equal to $1.175w$. A diameter of $4w$ truncates the beam at a level 34.7 dB below that on the axis of propagation and includes 99.97% of the power in the fundamental mode Gaussian beam. This is generally sufficient to make the effects of diffraction by the truncation quite small. The subject of truncation is discussed further in Chapters 6 and 11.

For the fundamental mode Gaussian in cylindrical coordinates, the fraction of the total power contained within a circle of radius r_e centered on the beam axis is found using

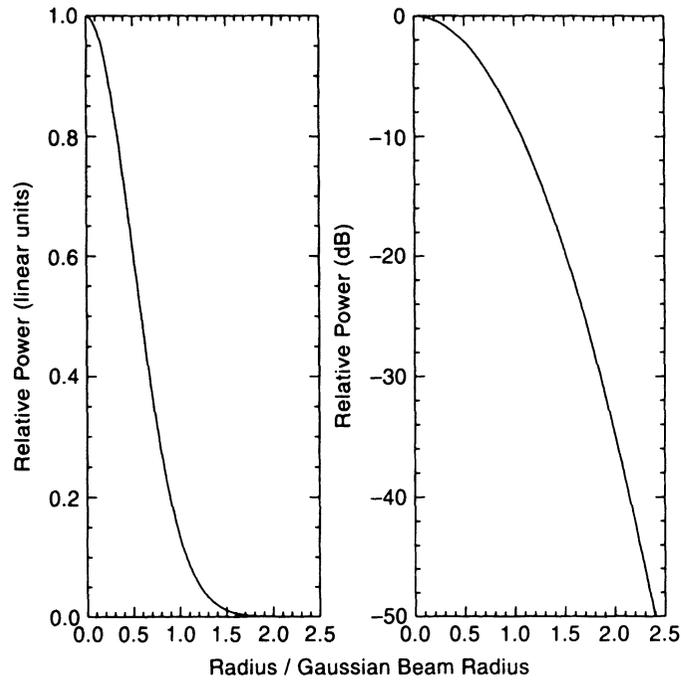


Figure 2.4 Fundamental mode Gaussian beam field distribution in linear units (left) and power distribution in logarithmic units (right). The horizontal axis is the radius expressed in terms of the beam radius, w .

TABLE 2.1 Fundamental Mode Gaussian Beam and Edge Taper

r_e/w	$T_e(r_e)$	$F(r_e)$	T_e (dB)
0.0	1.0000	0.0000	0.0
0.2	0.9231	0.0769	0.4
0.4	0.7262	0.2739	1.4
0.6	0.4868	0.5133	3.1
0.8	0.2780	0.7220	5.6
1.0	0.1353	0.8647	8.7
1.2	0.0561	0.9439	12.5
1.4	0.0198	0.9802	17.0
1.6	0.0060	0.9940	22.2
1.8	0.0015	0.9985	28.1
2.0	0.0003	0.9997	34.7
2.2	0.0001	0.9999	42.0

equation 2.33 to be

$$F_e(r_e) = \int_{r=0}^{r=r_e} |E(r)|^2 \cdot 2\pi r \, dr = 1 - T_e(r_e). \quad (2.36)$$

Thus, the fractional power of a fundamental mode Gaussian that falls outside radius r_e is just equal to the edge taper of the beam at that radius. Values for the fraction of the total

power propagating in a fundamental mode Gaussian beam as a function of radius of a circle centered on the beam axis are also given in Table 2.1 and shown in Figure 2.5.

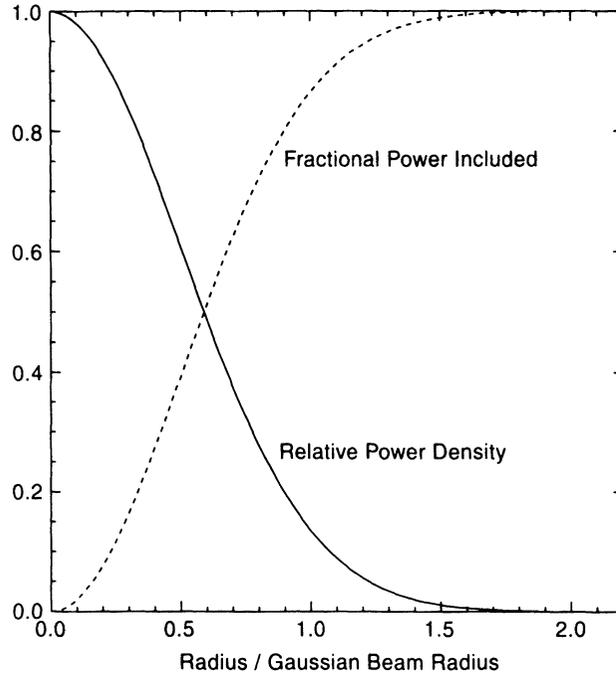


Figure 2.5 Fundamental mode Gaussian beam and fractional power contained included in circular area of specified radius.

In addition to the beam radius describing the Gaussian beam amplitude and power distributions, the Gaussian beam mode is defined by its radius of curvature. In the paraxial limit, the equiphase surfaces are spherical caps of radius R , as indicated in Figure 2.2b. As described above (Section 2.1.2), we have a quadratic variation of phase perpendicular to the axis of propagation at a fixed value of z . The radius of curvature defines the center of curvature of the beam, which varies as a function of the distance from the beam waist.

2.2.3 Average and Peak Power Density in a Gaussian Beam

The Gaussian beam formulas used here (e.g., equation 2.26a) are normalized in the sense that we assume unit total power propagating. This is elegant and efficient, but in some cases—high power radar systems are one example—it is important to know the actual power density. Since one of the main advantages of quasioptical propagation is the ability to reduce the power density by spreading the beam over a controlled region in space, we often wish to know how the peak power density depends on the actual beam size. From equation 2.26a we can write the expression for the actual power density P_{act} in a beam with total propagating power P_{tot} as

$$P_{\text{act}}(r) = P_{\text{tot}} \frac{2}{\pi w^2} \exp \left[-2 \left(\frac{r}{w} \right)^2 \right]. \quad (2.37)$$

Using equation 2.35b to relate the beam radius to the edge taper T_e at a specific radius r_e , we find

$$P_{\max} = P_{\text{act}}(0) = \left[\frac{T_e(\text{dB})}{4.343} \right] \frac{P_{\text{tot}}}{\pi r_e^2}. \quad (2.38)$$

This expression is useful if the relative power density or taper is known at any particular radius r_e . If we consider r_e to be the “edge” of the system defined by some focusing element or aperture, and as long as there has not been too much spillover, the second term on the right-hand side is the average power density,

$$P_{\text{av}} = \frac{P_{\text{tot}}}{\pi r_e^2}, \quad (2.39)$$

and we can relate the peak and average power densities through

$$P_{\max} = \left[\frac{T_e(\text{dB})}{4.343} \right] P_{\text{av}} = \frac{2r_e^2}{w^2} P_{\text{av}}. \quad (2.40)$$

For a strong edge taper of 34.7 dB produced by taking $r_e = 2w$, we find $P_{\max} = 8P_{\text{av}}$. On the other hand, for the very mild edge taper of 8.69 dB, obtained from $r_e = w$ (a taper that generally is not suitable for quasioptical system elements but is close to the value used for radiating antenna illumination, as discussed in Chapter 6), $P_{\max} = 2P_{\text{av}}$. This range of 2 to 8 includes the ratios of peak to average power density generally encountered in Gaussian beam systems.

2.2.4 Confocal Distance: Near and Far Fields

The variation of the descriptive parameters of a Gaussian beam has a particularly simple form when expressed in terms of the **confocal distance** or **confocal parameter**

$$z_c = \frac{\pi w_0^2}{\lambda}; \quad (2.41)$$

note that this parameter could be defined in a one-dimensional coordinate system in terms of w_{0x} or w_{0y} . This terminology derives from resonator theory, where z_c plays a major role. The confocal distance is sometimes called the **Rayleigh range** and is denoted z_0 by some authors and \hat{z} by others. Using the foregoing definition for confocal distance, the Gaussian beam parameters can be rewritten as

$$R = z + \frac{z_c^2}{z}, \quad (2.42a)$$

$$w = w_0 \left[1 + \left(\frac{z}{z_c} \right)^2 \right]^{0.5}, \quad (2.42b)$$

$$\phi_0 = \tan^{-1} \left(\frac{z}{z_c} \right). \quad (2.42c)$$

For example, for a wavelength of 0.3 cm and beam waist radius w_0 equal to 1 cm, the confocal distance is equal to 10.5 cm. We see that the radius of curvature R , the beam radius w , and the Gaussian beam phase shift ϕ_0 all change appreciably between the beam waist, located at $z = 0$, and the confocal distance at $z = z_c$.

One of the beauties of the Gaussian beam mode solutions to the paraxial wave equation is that a simple set of equations (e.g., equations 2.42) describes the behavior of the beam parameters at all distances from the beam waist. It is still natural to divide the propagating beam into a “near field,” defined by $z \ll z_c$ and a “far field,” defined by $z \gg z_c$, in analogy with more general diffraction calculations. The “transition region” occurs at the confocal distance z_c .

At the beam waist, the beam radius w attains its minimum value w_0 , and the electric field distribution is most concentrated, as shown in Figure 2.2a. As required by conservation of energy, the electric field and power distributions have their maximum on-axis values at the beam waist. The radius of curvature of the Gaussian beam is infinite there, since the phase front is planar at the beam waist. The phase shift ϕ_0 , which is the on-axis phase of a Gaussian beam relative to a plane wave, is, by definition zero at the beam waist.

Away from the beam waist, the beam radius increases monotonically. As described by equation 2.42b and as shown in Figure 2.6, the variation of w with z is seen to be hyperbolic. In the near field, the beam radius is essentially unchanged from its value at the beam waist; $w \leq \sqrt{2} w_0$. Thus, we can say that the confocal distance defines the distance over which the Gaussian beam propagates without significant growth—meaning that it remains essentially **collimated**. As we move away from the waist, the radius of curvature, as described by equation 2.42a and shown in Figure 2.6, decreases until we reach distance z_c .

At a distance from the waist equal to z_c , the beam radius is equal to $\sqrt{2} w_0$, the radius of curvature attains its minimum value equal to $2z_c$, and the phase shift is equal to $\pi/4$. At

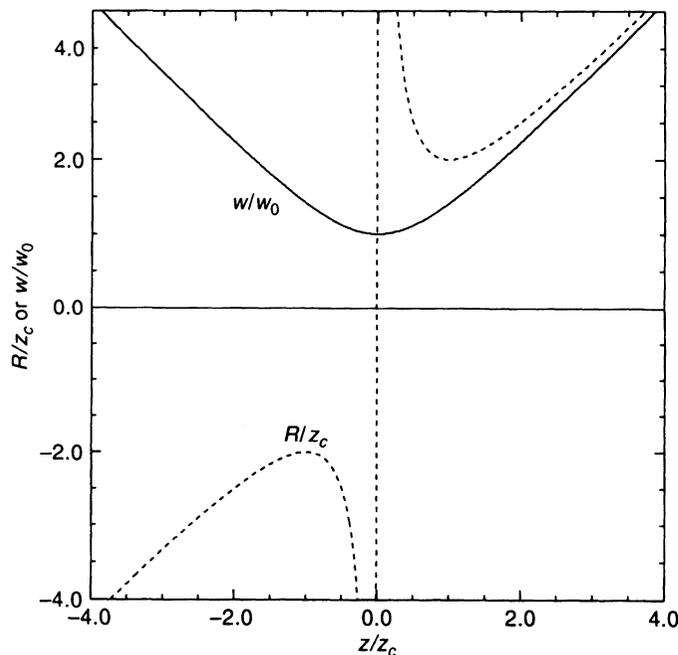


Figure 2.6 Variation of beam radius w and radius of curvature R of Gaussian beam as a function of distance z from beam waist. The beam radius is normalized to the value at the beam waist—the beam waist radius w_0 , while the radius of curvature is normalized to the confocal distance $z_c = \pi w_0^2/\lambda$.

distances from the waist greater than z_c , the beam radius grows significantly, and the radius of curvature increases.

In the far field, $z \gg z_c$, the beam radius grows linearly with distance. The growth of the $1/e$ radius of the electric field can be defined in terms of an angle $\theta = \tan^{-1}(w/z)$, and in the far-field limit we obtain the **asymptotic beam growth angle** θ_0 , given by

$$\theta_0 = \lim_{z \gg z_c} \left[\tan^{-1} \left(\frac{w}{z} \right) \right] = \tan^{-1} \left(\frac{\lambda}{\pi w_0} \right), \quad (2.43a)$$

as shown in Figure 2.7. As a numerical example, we see that for $\lambda = 0.3$ cm and $w_0 = 1$ cm, $\theta_0 \cong 0.1$ radian. The small-angle approximation can generally be used satisfactorily in the paraxial limit, giving

$$\theta_0 \cong \frac{\lambda}{\pi w_0}. \quad (2.43b)$$

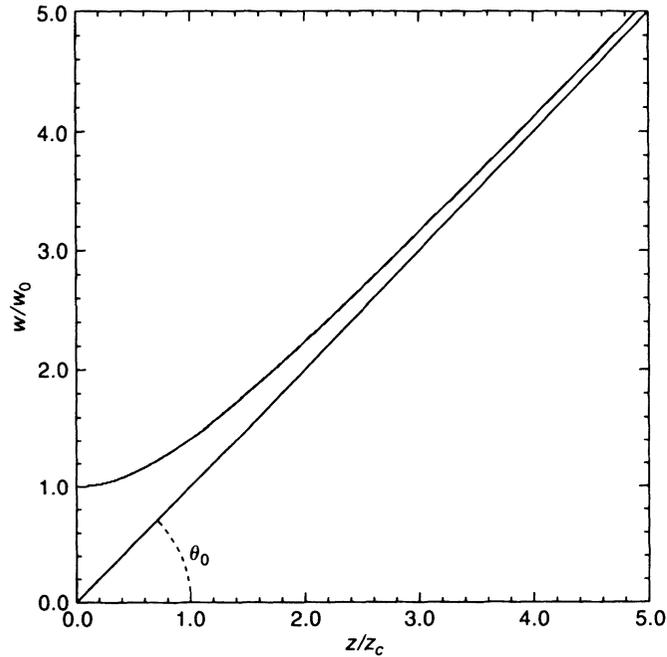


Figure 2.7 Divergence angle, θ_0 , of Gaussian beam illustrated in terms of the asymptotic growth angle of the beam radius as a function of distance from the beam waist.

In the far field, it is convenient to express the electric field distribution as a function of angle away from the propagation axis. The usual field distribution as a function of distance from the axis of propagation becomes a Gaussian function of the off-axis angle θ :

$$\frac{E(\theta)}{E(0)} = \exp \left[- \left(\frac{\theta}{\theta_0} \right)^2 \right]. \quad (2.44)$$

This is, of course, a reflection of the constancy of the form of the Gaussian beam. It is also a convenient feature in that, for example, the fraction of a power outside a specified angle,

θ_e , is given by an expression of the same form used for the distribution as a function of radius (equation 2.36), but with θ_e and θ_0 substituted for r_e and w_0 .

From equation 2.42a we see that in the far field the radius of curvature also increases linearly with distance, since for $z \gg z_c$, $R \rightarrow z$. In this limit, the radius of curvature is just equal to the distance from the beam waist. The phase shift has the asymptotic limit $\phi_0 = \pi/2$ in two dimensions. This is an example of the Gouy phase shift, which occurs for any focused beam of radiation ([SIEG86], Section 17.4, pp. 682–684; [BOYD80]), but note that the phase shift is only half this value for a Gaussian beam in one dimension.

Useful formulas that summarize the propagation of a symmetric fundamental mode Gaussian beam in a cylindrical coordinate system are collected for convenient reference in Table 2.2.

TABLE 2.2 Summary of Fundamental Mode Gaussian Beam Formulas¹

$E(r, z) = \left[\frac{2}{\pi w^2(z)} \right]^{0.5} \exp \left[\frac{-r^2}{w^2(z)} - jkz - \frac{j\pi r^2}{\lambda R(z)} + j\phi_0(z) \right]$	Transverse field distribution ²
$w(z) = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{0.5}$	Beam radius
$\frac{P(r)}{P(0)} = \exp \left[-2 \left(\frac{r}{w(z)} \right)^2 \right]$	Relative power distribution transverse to axis of propagation
$T_e \text{ (dB)} = 8.686 \left(\frac{r_e}{w} \right)^2$	Edge taper
$\theta_0 = \frac{\lambda}{\pi w_0}$	Far-field divergence angle
$\theta_{\text{rwhm}} = 1.18 \theta_0$	Far-field beam width of power distribution to half-maximum
$R(z) = z + \frac{(\pi w_0^2 / \lambda)^2}{z}$	Radius of curvature
$\phi_0(z) = \tan^{-1} \left(\frac{\lambda z}{\pi w_0^2} \right)$	Phase shift

¹ Symmetric beam having waist radius w_0 located at $z = 0$ along axis of propagation z . The transverse coordinate is r , which is limited by edge radius r_e for truncated beam.

² Normalized so that $\int_0^\infty |E|^2 2\pi r dr = 1$.

2.3 GEOMETRICAL OPTICS LIMITS OF GAUSSIAN BEAM PROPAGATION

The geometrical optics limit is that in which $\lambda \rightarrow 0$, so that effects of diffraction become unimportant. Some caution is necessary to apply this to Gaussian beam formulas, since taking the limit $\lambda \rightarrow 0$ for fixed value of w_0 is equivalent to making $z_c \rightarrow \infty$, and the region of interest is always in the near field of the beam waist. The resulting asymptotic behavior

$w \rightarrow w_0$, $R \rightarrow \infty$, and $\theta_0 \rightarrow 0$ is what we would expect from a perfectly collimated beam that suffers no diffraction effects.

If we wish to maintain a finite value of z_c , one convenient way is to let the waist radius approach zero along with the wavelength. In this situation, we have $\theta_0 \rightarrow \text{constant}$, and $w \rightarrow \theta_0 z$ while $R \rightarrow z$. This behavior is just what we expect for a geometrical beam diverging from a point source.

2.4 HIGHER ORDER GAUSSIAN BEAM MODE SOLUTIONS OF THE PARAXIAL WAVE EQUATION

The Gaussian beam solutions of the paraxial wave equation for the different coordinate systems presented in Section 2.1 were indicated to be the simplest solutions of this equation describing propagation of a quasi-collimated beam of radiation. While certainly the most important and most widely used, they are not the only solutions. In certain situations we need to deal with solutions that have a more complex variation of the electric field perpendicular to the axis of propagation: these are the **higher order Gaussian beam mode solutions**. Such solutions have polynomials of different kinds superimposed on the fundamental Gaussian field distribution. The higher order beam modes are characterized by a beam radius and a radius of curvature that have the same behavior as that of the fundamental mode presented above, while their phase shifts are different. Higher order Gaussian beam modes in cylindrical coordinates must be included to deal with radiating systems that have a high degree of axial symmetry but do not have perfectly Gaussian radiation patterns (e.g., corrugated feed horns). Higher order beam modes in rectangular coordinates can be produced by an off-axis mirror, as discussed in Chapter 5, or they can be the result of the non-Gaussian field distribution in a horn (such as a rectangular feed horn; cf. Chapter 7).

2.4.1 Higher Order Modes in Cylindrical Coordinates

In a cylindrical coordinate system, a general solution must allow variation of the electric field as a function of the polar angle φ . In addition, a trial solution need not be limited to the purely Gaussian form employed earlier (equation 2.8), but may contain terms with additional radial variation. A plausible trial solution for such a higher order solution is

$$u(r, \varphi, z) = A(z) \exp\left[-\frac{jk r^2}{2q(z)}\right] S(r) \exp(jm\varphi), \quad (2.45)$$

where the complex amplitude $A(z)$ and the complex beam parameter $q(z)$ depend only on distance along the propagation axis, $S(r)$ is an unknown radial function, and m is an integer. Assuming the same form for q as obtained for the fundamental Gaussian beam mode in Section 2.1.2, we find that the paraxial wave equation reduces to a differential equation for S . The solutions obtained are

$$S(r) = \left(\frac{\sqrt{2r}}{w}\right)^m L_{pm}\left(\frac{2r^2}{w^2}\right), \quad (2.46)$$

where w is the beam radius as defined and used previously and L_{pm} is the generalized Laguerre polynomial. In the Gaussian beam context, p is the radial index and m is the angular index. The polynomials $L_{pm}(u)$ are solutions to Laguerre's differential equation [MARG56]

$$u \frac{d^2 L_{pm}}{du^2} + (m+1-u) \frac{dL_{pm}}{du} + pL_{pm} = 0, \quad (2.47)$$

and can conveniently be obtained from the expression [GOUB69]

$$L_{pm}(u) = \frac{e^u u^{-m}}{p!} \frac{d^p}{du^p} (e^{-u} u^{p+m}). \quad (2.48)$$

They can also be obtained from direct series representations ([ABRA65], [MART89])

$$L_{pm}(u) = \sum_{l=0}^{l=p} \frac{(p+m)!(-u)^l}{(m+l)!(p-l)!l!}. \quad (2.49)$$

Some of the low order Laguerre polynomials are

$$L_{0m}(u) = 1 \quad (2.50)$$

$$L_{1m}(u) = 1 + m - u$$

$$L_{2m}(u) = \frac{1}{2}[(2+m)(1+m) - 2(2+m)u + u^2]$$

$$L_{3m}(u) = \frac{1}{6}[(3+m)(2+m)(1+m) - 3(3+m)(2+m)u + 3(3+m)u^2 - u^3].$$

A solution to the paraxial wave equation in cylindrical coordinates with the Laguerre polynomial having indices p and m is generally called the **pm Gaussian beam mode** or simply the **pm mode**, and the normalized electric field distribution is given by

$$\begin{aligned} E_{pm}(r, \varphi, z) = & \left[\frac{2p!}{\pi(p+m)!} \right]^{0.5} \frac{1}{w(z)} \left[\frac{\sqrt{2r}}{w(z)} \right]^m L_{pm} \left(\frac{2r^2}{w^2(z)} \right) \\ & \cdot \exp \left[\frac{-r^2}{w^2(z)} - jkz - \frac{j\pi r^2}{\lambda R(z)} - j(2p+m+1)\phi_0(z) \right] \\ & \cdot \exp(jm\varphi), \end{aligned} \quad (2.51)$$

where the beam radius w , the radius of curvature R , and the phase shift ϕ_0 are exactly the same as for the fundamental Gaussian beam mode. Aside from the angular dependence and the more complex radial dependence, the only significant difference in the electric field distribution is that the phase shift is greater than for the fundamental mode by an amount that depends on the mode parameters.

These higher order Gaussian beam mode solutions are normalized so that each represents unit power flow (cf. Section 2.1.3), and they obey the orthogonality relationship

$$\iint r dr d\varphi E_{pm}(r, \varphi, z) E_{qn}^*(r, \varphi, z) = \delta_{pq} \delta_{mn}. \quad (2.52)$$

It is sometimes convenient to make combinations of these higher order Gaussian beam modes that are real functions of φ . This can be done straightforwardly by combining $\exp(jm\varphi)$ and $\exp(-jm\varphi)$ terms into $\cos(m\varphi)$ and $\sin(m\varphi)$ beam mode functions. To preserve the correct normalization, the beam mode amplitudes must be multiplied by a factor equal to 1 for $m = 0$ and equal to $\sqrt{2}$ otherwise.

If we wish to consider modes that are axially symmetric (independent of φ), we choose from those defined by equation 2.51 the subset having $m = 0$. These are often used

in describing systems that are azimuthally symmetric but are not exactly described by the fundamental Gaussian beam mode, such as a corrugated feedhorn (cf. Chapter 7). These modes can be written as

$$E_{p0}(r, z) = \left[\frac{2}{\pi w^2} \right]^{0.5} L_{p0} \left(\frac{2r^2}{w^2} \right) \exp \left[-\frac{r^2}{w^2} - jkz - \frac{j\pi r^2}{\lambda R} + j(2p + 1)\phi_0 \right], \quad (2.53)$$

where we have omitted explicit dependence of the various quantities on distance along the axis of propagation. The functions L_{p0} are the ordinary Laguerre polynomials that can be obtained from equations 2.47 to 2.49 with $m = 0$ since $L_p(u) \equiv L_{p0}(u)$. They are given by

$$L_p(u) = \frac{e^u}{p!} \frac{d^p}{du^p} (e^{-u} u^p), \quad (2.54)$$

or by the series representation

$$L_p(u) = \sum_{l=0}^{l=p} \frac{p!(-u)^l}{(p-l)!!!!}. \quad (2.55)$$

The amplitude distributions transverse to the axis of propagation of some Gauss-Laguerre beams of low order are shown in Figure 2.8. Two-dimensional representations of the E_0 and E_2 modes are shown in Figures 2.9a and 2.9b, respectively. The axially symmetric beam mode of order p has p zero crossings for $0 \leq r \leq \infty$, with the sign of the electric field reversing itself in each successive annular region. The power density distribution thus has $p + 1$ “bright rings,” including the central “spot.” The non-axially

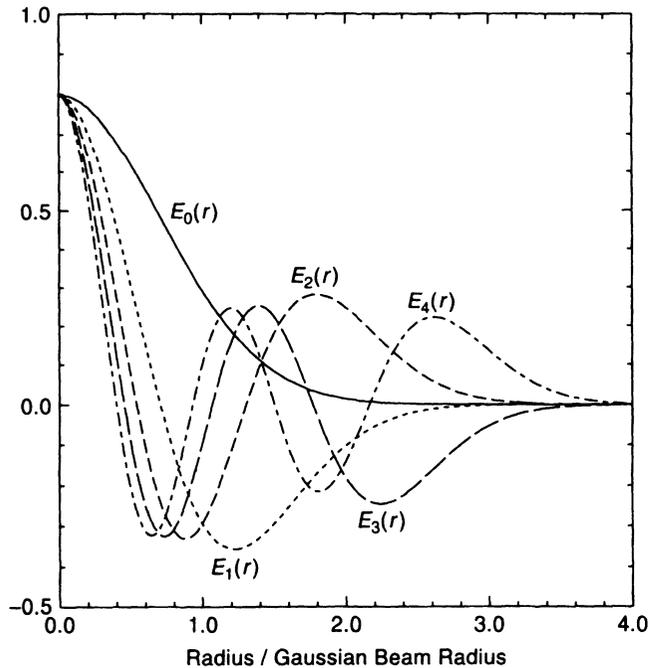
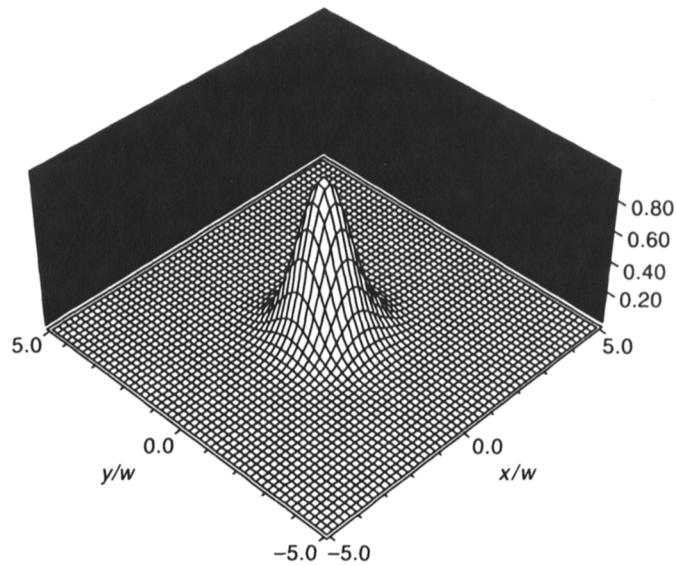
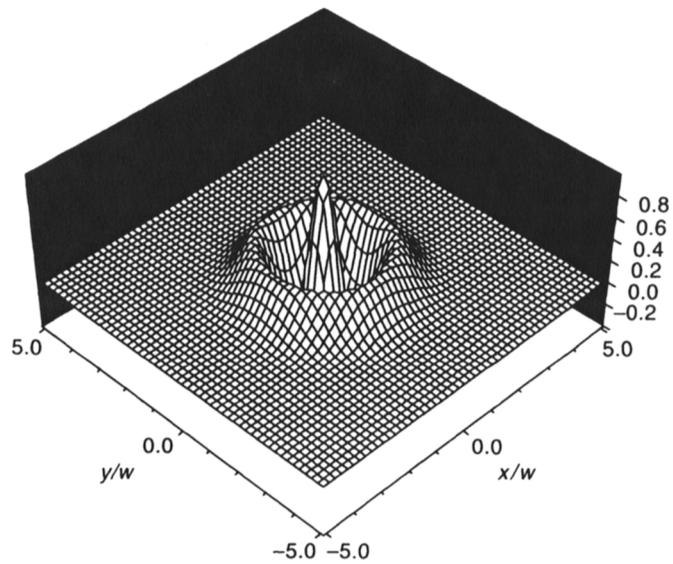


Figure 2.8 Electric field distribution transverse to axis of propagation, of axially symmetric Gauss-Laguerre beam modes E_0 (fundamental mode) through E_4 .



(a)



(b)

Figure 2.9 Two-dimensional representations of axially symmetric Gauss-Laguerre beam modes: (a) fundamental E_0 mode and (b) E_2 mode.

symmetric modes are more complex; the pm mode (with $\cos m\varphi$ or $\sin m\varphi$) has each annular region broken up into $2m + \delta_{0m}$ zones with alternating signs, for $0 \leq \varphi \leq 2\pi$. Thus, the power density has $(2m + \delta_{0m})(p + 1)$ bright regions.

2.4.2 Higher Order Modes in Rectangular Coordinates

When a rectangular coordinate system is used for the higher order modes, the general two-dimensional Gaussian beam mode is simply the product of two one-dimensional functions. Each of these is a more general solution to the paraxial wave equation (equation 2.5) above. Considering the x coordinate alone for the moment, we include an additional x -dependent function H to obtain the higher order modes. A trial solution of the form

$$u(x, z) = A(z) H \left(\frac{\sqrt{2}x}{w(z)} \right) \exp \left[-\frac{jkx^2}{2q(z)} \right] \quad (2.56)$$

is successful if we take the beam radius w and the complex beam parameter q to be the same as for the fundamental mode discussed above. The function H satisfies Hermite's differential equation [MARG56]

$$\frac{d^2 H(u)}{du^2} - 2u \frac{dH(u)}{du} + 2mH(u) = 0, \quad (2.57)$$

where m is a positive integer. This is the defining equation for the Hermite polynomial of order m , denoted $H_m(u)$. $H_0(u) = 1$ and $H_1(u) = 2u$; the remaining polynomials are easily obtained from the recursion relation

$$H_{n+1}(u) = 2[uH_n(u) - nH_{n-1}(u)], \quad (2.58)$$

and can also be found from direct series expansion or from the expression [MARG56]

$$H_n(u) = (-1)^n e^{u^2} \frac{d^n}{du^n} (e^{-u^2}). \quad (2.59)$$

The Hermite polynomials through order 4 are:

$$H_0(u) = 1 \quad (2.60)$$

$$H_1(u) = 2u$$

$$H_2(u) = 4u^2 - 2$$

$$H_3(u) = 8u^3 - 12u$$

$$H_4(u) = 16u^4 - 48u^2 + 12.$$

With the same convention for normalization used earlier, we find the expression for the one-dimensional Gaussian beam mode of order m to be

$$E_m(x, z) = \left(\frac{2}{\pi} \right)^{0.25} \left[\frac{1}{w_x 2^m m!} \right]^{0.5} H_m \left(\frac{\sqrt{2}x}{w_x} \right) \cdot \exp \left[-\frac{x^2}{w_x^2} - jkz - \frac{j\pi x^2}{\lambda R_x} + \frac{j(2m+1)\phi_{0x}}{2} \right]. \quad (2.61)$$

The variation of the beam radius, the radius of curvature, and the phase shift are the same as for the fundamental mode (equations 2.26b–d), but we note that the phase shift is greater for the higher order modes. The E_0 mode is of course identical to the fundamental mode in one dimension (equation 2.30).

In dealing with the two-dimensional case, the paraxial wave equation for $u(x, y, z)$ separates with the appropriate trial solution formed from the product of functions like those of equation 2.61. We have the ability to deal with higher order modes having unequal beam

waist radii and different beam waist locations. Normalizing to unit power flow results in the expression for the mn Gauss–Hermite beam mode

$$E_{mn}(x, y, z) = \left(\frac{1}{\pi w_x w_y 2^{m+n-1} m! n!} \right)^{0.5} H_m \left(\frac{\sqrt{2}x}{w_x} \right) H_n \left(\frac{\sqrt{2}y}{w_y} \right) \cdot \exp \left[-\frac{x^2}{w_x^2} - \frac{y^2}{w_y^2} - jkz - \frac{j\pi x^2}{\lambda R_x} - \frac{j\pi y^2}{\lambda R_y} + \frac{j(2m+1)\phi_{0x}}{2} + \frac{j(2n+1)\phi_{0y}}{2} \right]. \quad (2.62)$$

The higher order modes in rectangular coordinates obey the orthogonality relationship

$$\iint_{-\infty}^{\infty} E_{mn}(x, y, z) E_{pq}^*(x, y, z) dx dy = \delta_{mp} \delta_{nq}. \quad (2.63)$$

Some Gauss–Hermite beams of low order are shown in Figure 2.10. The Gauss–Hermite beam mode $E_m(x)$ has m zero crossings in the interval $-\infty \leq x \leq \infty$. Thus, the power distribution has $m + 1$ regions with local intensity maxima along the x axis, while the $E_{mn}(x, y)$ beam mode in two dimensions has $(m + 1)(n + 1)$ “bright spots.”

One special situation is that in which beams in x and y with equal beam waist radii are located at the same value of z . In this case we obtain (taking $w_x = w_y \equiv w$, $R_x = R_y \equiv R$, and $\phi_{0x} = \phi_{0y} \equiv \phi_0$)

$$E_{mn}(x, y, z) = \left(\frac{1}{\pi w^2 2^{m+n-1} m! n!} \right)^{0.5} H_m \left(\frac{\sqrt{2}x}{w} \right) H_n \left(\frac{\sqrt{2}y}{w} \right) \cdot \exp \left[-\frac{(x^2 + y^2)}{w^2} - jkz - \frac{j\pi(x^2 + y^2)}{\lambda R} + j(m + n + 1)\phi_0 \right]. \quad (2.64)$$

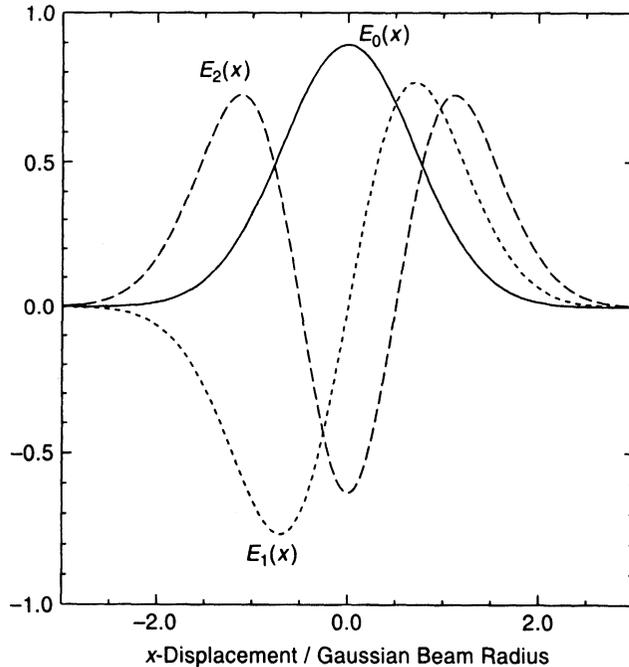


Figure 2.10 Electric field distribution of Gauss–Hermite beam modes E_0 , E_1 , and E_2 .

This expression can be useful if we have equal waist radii in the two coordinates, but the beam of interest is not simply the fundamental Gaussian mode. For $m = n = 0$, we again obtain the fundamental Gaussian beam mode with purely Gaussian distribution.

2.5 THE SIZE OF GAUSSIAN BEAM MODES

Although we carry out calculations primarily with the field distributions, we most often measure the power distribution of a Gaussian beam. This convention is of practical importance in determining the beam radius at a particular point along the beam's axis of propagation, or in verifying the beam waist radius in an actual system. For a fundamental mode Gaussian, the fraction of power included within a circle of radius r_0 increases smoothly with increasing r_0 as discussed in Section 2.2.1. For the higher order modes, the behavior is not so simple, since it is evident from Section 2.4 that power is concentrated away from the axis of propagation. Consequently, the beam radius w is not an accurate indication of the transverse extent of higher order Gaussian beam modes.

It is convenient to have a good measure of the “size” of a Gaussian beam for arbitrary mode order; this is also referred to as the “spot size.” An appealing definition for the size of the Gaussian beam pm mode in cylindrical coordinates is [PHIL83]

$$\rho_{r-pm}^2 = 2 \iint I_{pm}(r, \varphi) r^2 dS = 2 \iint r^3 dr d\varphi |E_{pm}(r, \varphi)|^2, \quad (2.65)$$

where we employ the normalized form of the field distribution (equation 2.51) or normalize by dividing by $\iint I_{pm}(r, \varphi) dS$. Evaluation of this integral yields

$$\rho_{r-pm} = w[2p + m + 1]^{0.5}, \quad (2.66)$$

where w is the beam radius at the position of interest along the axis of propagation, and ρ_{r-pm} , given by equation 2.66, is just equal to the beam radius for the fundamental mode with $p = m = 0$.

The analogous definition for the m mode in one dimension in a Cartesian coordinate system is

$$\rho_{x-m}^2 = 2 \int |E_m(x)|^2 x^2 dx = w_x^2 \left[m + \frac{1}{2} \right]^{0.5}, \quad (2.67)$$

where we have adapted the discussion in [CART80] to conform to our notation. While it might appear that these modifications give inconsistent results for the fundamental mode, this is not really the case, since we need to consider a two-dimensional case in rectangular geometry for comparison with the cylindrical case. For the n mode in the y direction, we obtain

$$\rho_{y-n} = w_y \left[n + \frac{1}{2} \right]^{0.5}. \quad (2.68)$$

The two-dimensional beam size is defined as $\rho_{xy}^2 = \rho_x^2 + \rho_y^2$, which for a symmetric beam with $w_x = w_y = w$, becomes

$$\rho_{xy-mn} = w[m + n + 1]^{0.5}, \quad (2.69)$$

and for the fundamental mode gives $\rho_{xy-00} = w$, in agreement with the result obtained from equation 2.66. The size of the Gauss–Laguerre and Gauss–Hermite beam modes thus

grows as the square root of the mode number for high order modes. This is in accord with the picture that a higher order mode has power concentrated at a larger distance from the axis of propagation, for a given w , than does the fundamental mode. It is particularly important that high order beam modes are “effectively larger” than the fundamental mode having the same beam radius when the fundamental mode is not a satisfactory description of the propagating beam, and we want to avoid truncation of the beam. The guidelines given in Section 2.2.2 apply specifically to the fundamental mode, and the focusing elements, components, and apertures must be increased in size if the higher order modes are to be accommodated without excessive truncation.

2.6 GAUSSIAN BEAM MEASUREMENTS

It is naturally of interest for the design engineer to be able to verify that a quasioptical system that has been designed and constructed actually operates in a manner that can be accurately described by the expected Gaussian beam parameters. This is important not only to ensure overall high efficiency, but to be able to predict accurately the performance of certain quasioptical components (discussed in more detail in Chapter 9), which depend critically on the parameters of the Gaussian beam employed.

A variety of techniques for measuring power distribution in a quasioptical beam have been developed. Work on optical fibers and Gaussian beams of small transverse dimensions at optical frequencies has encouraged approaches that measure power transmitted through a grating with regions of varying opacity; the fractional transmission is related to the relative size of the beam radius and the grating period. It may be more convenient to measure the maximum and minimum transmission through such a grating as it is scanned across the beam than to determine the beam profile by scanning a pinhole or knife edge (cf. discussion in [CHER92]).

However, at millimeter and submillimeter wavelengths, beam sizes are generally large enough that beams can be effectively and accurately scanned with a small detector (cf. [GOLD77]). This technique assumes the availability of a reasonably strong signal, as is often provided by the local oscillator in a heterodyne radiometric system. Best results are obtained by interposing a sheet of absorbing material to minimize reflections from the measurement system.

An alternative for probing the beam profile is to employ a high sensitivity radiometric system and to move a small piece of absorbing material transversely in the beam. If the overall beam is terminated in a cooled load (e.g., at the temperature of liquid nitrogen), the moving absorber can be at ambient temperature, which is an added convenience. To obtain high spatial resolution, only a small fraction of the beam can be filled by the load at the different temperature. Thus the signal produced is necessarily a small fraction of the maximum that can be obtained for a given temperature difference and good sensitivity is critical. If the beam is symmetric, the moving sample can be made into a strip filling the beam in one dimension, without sacrificing spatial resolution. A half-plane can also be used and the actual beam shape obtained by deconvolution; this approach can also be utilized for asymmetric beams, although a more elaborate analysis of the data is necessary to obtain the relevant beam parameters [BILG85].

Another good method, which is particularly effective for small systems, is to let the beam propagate and measure the angular distribution of radiation at a distance $z \gg z_c$.

Then, following the discussion in Section 2.2.4, the beam waist radius can be determined. Note that a precise measurement requires knowledge of the beam waist location, which may or may not be available. In practice, however, this technique works well to verify the size of the beam waist as long as its location is reasonably well known. It is basically the convenience of a measurement of angular power distribution (i.e., using an antenna positioner system) that makes this approach more attractive than transverse beam scanning, and the choice of which method to employ will largely depend on the details of the system being measured and the equipment available.

Relatively little work has been done on measuring the phase distribution of Gaussian beams; the usual assumption is that if the intensity distribution follows a smooth Gaussian, the phase will be that of the expected spherical wave. On the other hand, “ripples” in the transverse intensity distribution are generally indicative of the presence of multiple modes with different phase distributions, which are symptomatic of truncation, misalignment, or other problems. An interesting method for measurement of the phase distribution of coherent optical beams described by [RUSC66] could be applied to quasioptical systems at longer wavelengths. If the phase and amplitude of the far field pattern are measured (as is possible with many antenna pattern measurement systems), then the amplitude and phase of the radiating beam can be recovered. While the quadratic phase variation characterizing the spherical wave front is difficult to distinguish from an error in location of the reference plane, higher order phase variations can be measured with high reliability.

2.7 INVERSE FORMULAS FOR GAUSSIAN BEAM PROPAGATION

In the discussion to this point it has been assumed that we know the size of the beam waist radius and its location and that it is possible to calculate (using, e.g., equation 2.21) the beam radius and radius of curvature at some specified position along the axis of propagation. We can represent this calculation by $\{w_0, z\} \rightarrow \{w, R\}$. In practice we may know only the size of a Gaussian beam, and the distance to its waist—this might come about, for example, by measurement of the size of a beam and knowledge that it was produced by a feed horn at a specified location. Or, we might be able to measure the beam radius and the radius of curvature (if phase measurements can be carried out). In these cases, we need to have “inverse” formulas, in the sense of working back to the beam waist, to allow us to determine the unknown parameters of the beam.

The most elegant of these inverse formulas is obtained directly from the two different definitions of the complex beam parameter (equations 2.29a and 2.29b). By taking the inverse of either of these, rationalizing, and equating real and imaginary parts, we obtain the transformation for $\{w, R\} \rightarrow \{w_0, z\}$; the resulting expressions are given in Table 2.3. This is a special case, because the two pairs of parameters are related to the imaginary and real parts of q and q^{-1} . If we have other pairs of parameters, such as w and z or w_0 and R , we have to solve fourth-order equations, and obtain pairs of solutions. In the other cases it is straightforward to invert the standard equations (2.26b and 2.26c) to obtain the desired relationships.

The set of six pairs of known parameters (including the conventional one in which the beam waist radius and location are known), together with the relevant equations to obtain

TABLE 2.3 Formulas for Determining Gaussian Beam Quantities Starting with Different Pairs of Known Parameters

Known Parameter Pairs			
w_0	z	$w = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{0.5}$	$R = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right]$
R	z	$w_0^2 = \frac{\lambda}{\pi} [z(R-z)]^{0.5}$	w from w_0 and z
w	z	$w_0^2 = \frac{w^2}{2} \left\{ 1 \pm \left[1 - \left(\frac{2\lambda z}{\pi w^2} \right)^2 \right]^{0.5} \right\}$	R from w_0 and z
w_0	w	$z = \frac{\pi w_0}{\lambda} [w^2 - w_0^2]^{0.5}$	R from w_0 and z
w_0	R	$z = \frac{R}{2} \left\{ 1 \pm \left[1 - \left(\frac{2\pi w_0^2}{\lambda R} \right)^2 \right]^{0.5} \right\}$	w from w_0 and z
w	R	$w = \frac{w}{\left[1 + \left(\frac{\pi w^2}{\lambda R} \right)^2 \right]^{0.5}}$	$z = \frac{R}{1 + \left(\frac{\lambda R}{\pi w^2} \right)^2}$

unknown parameters, are given in Table 2.3. In using these, it is assumed that once we have solved for the beam waist radius and its location (i.e., once we know w_0 and z), we can use the standard equations to obtain other information desired about the Gaussian beam. We note again that these formulas apply to the higher order as well as to the fundamental Gaussian beam mode, but care must be taken in determining w from measurements of the field distribution of a higher order mode.

2.8 THE PARAXIAL LIMIT AND IMPROVED SOLUTIONS TO THE WAVE EQUATION

The preceding discussion in this chapter has been based on solutions to the paraxial wave equation (equations 2.5–2.7). Since the paraxial wave equation is a satisfactory approximation to the complete wave equation only for reasonably well-collimated beams, it is appropriate to ask how divergent a beam can be before the Gaussian beam mode solutions cease to be acceptably accurate. For a highly divergent beam, the electric field distribution at the beam waist is concentrated within a very small region, on the order of a wavelength or less. In this situation, the approximation that variations will occur on a scale that is large compared to a wavelength is unlikely to be satisfactory. In fact, a solution to the wave equation cannot have transverse variations on such a small scale and still have an electric field that is purely transverse to the axis of propagation. In addition, it is not possible to have an electric field that is purely linearly polarized, as has been assumed to be the case in the preceding discussion.

Thus, when we consider a beam waist that is on the order of a wavelength in size or smaller, we find that the actual solution for the electric field has longitudinal and cross-polarized components. In addition, the variation of the beam size and its amplitude as

a function of distance from the beam waist do not follow the basic Gaussian beam formulas developed above. This topic has received considerable attention in recent years. Approximate solutions based on a series expansion of the field in terms of a parameter proportional to w_0/λ have been developed, and recursion relations found to allow computation (cf. [VANN64], [LAX75], [AGAR79], [COUT81], [AGAR88]). These solutions include a longitudinal component as well as modifications to the transverse distribution.

Corrections for higher order beam modes have also been studied [TAKE85]. As indicated in figures presented by [NEMO90], if we force at the waist a solution that is a fundamental Gaussian distribution transverse to the axis of propagation, the beam diverges more rapidly than expected from the Gaussian beam mode equations, and the on-axis amplitude decreases more rapidly in consequence. The phase variation is also affected. [NEMO90] defines four different regimes. For $w_0/\lambda \geq 0.9$ the paraxial approximation itself is valid, while for $0.5 \leq w_0/\lambda \leq 0.9$ the paraxial and exact solutions differ, but the first-order correction is effective. For $0.25 \leq w_0/\lambda \leq 0.5$, the first-order correction is not sufficient, while for $w_0/\lambda < 0.25$ the paraxial approximation completely fails and the corrections are ineffective. Similar criteria have been derived by [MART93], based on a plane wave expansion of a propagating beam. They find that for $w_0/\lambda \geq 1.6$ corrections to the paraxial approximation are negligible, but for $w_0/\lambda \leq 0.95$ the paraxial approximation introduces significant error.

The criterion $w_0/\lambda \geq 0.9$ (which is in reasonable agreement with limits fixed in earlier treatments, e.g., [VANN64]), is a very useful one for defining the range of applicability of the paraxial approximation. It corresponds to a value of the far-field divergence angle $\theta_0 \leq 0.35$ rad or 20° . Thus (using equation 2.36 or Table 2.1) approximately 99% of the power in the fundamental mode Gaussian beam is within 30° of the axis of propagation for this limiting value of θ_0 . While, as suggested above, this is not a hard limit for the application of the paraxial approximation, it represents a limit for using it with good confidence. Employing the paraxial approximation for angles up to 45° will give essentially correct answers, but there will inevitably be errors as we approach the upper limit of this range.

Unfortunately, the first-order corrections as given explicitly by [NEMO90] are so complex that they have not seen any significant use, and they are unlikely to be very helpful in general design procedures. They could profitably be applied, however, in a specific situation involving large angles once an initial but insufficiently accurate design had been obtained by means of the paraxial approximation.

A different approach by [TUOV92] is based on finding an improved “quasi-Gaussian” solution, which is exact at the beam waist and does a better job of satisfying the full-wave equation than do the Gaussian beam modes, which are solutions of the paraxial wave equation. This improved solution has the (un-normalized) form in cylindrical coordinates

$$E(r, z) = \frac{w_0}{w} \frac{1}{F''^2} \exp \left[-\frac{(r/F'')^2}{w^2} - jkz - jkR(F'' - 1) + j\phi_0 \right], \quad (2.70)$$

where $F'' = [1 + (r/R)^2]^{0.5}$. This is obviously very similar to equation 2.25b, and in fact for $r \ll R$, we can take $F'' = 1$ in the amplitude term while keeping only terms to second order in the phase. This yields the standard fundamental Gaussian beam mode solution to the paraxial wave equation. This solution is derived and analyzed extensively in [FRIB92], and it appears to be an improvement, except possibly in the region $z \cong z_c$. It may be useful for improving the Gaussian beam analysis of systems with very small effective waist radii

(e.g., feed horns having very small apertures). The transformation properties of such a modified beam remain to be studied in detail.

2.9 ALTERNATIVE DERIVATION OF THE GAUSSIAN BEAM PROPAGATION FORMULA

It is illuminating to consider the propagation of a Gaussian beam in the context of a diffraction integral. With the assumption of small angles so that obliquity factors can be set to unity, the familiar Huygens–Fresnel diffraction integral for the field produced by a planar phase distribution and amplitude illumination function E_0 can be written (cf. [SIEG86] Section 16.2, pp. 630–637)

$$E(x', y', z') = \frac{j}{\lambda z'} \exp(-jkz') \iint E_0(x, y, 0) \exp\left[\frac{-jk(x' - x)^2 + (y' - y)^2}{2z'}\right] dx dy. \quad (2.71)$$

We have assumed that the illuminated plane is defined by coordinates $(x, y, z = 0)$, while the observation plane is defined by (x', y', z') . Consider the incident illumination to be an axially symmetric Gaussian beam with a planar phase front, $E_0 = \exp[-(x^2 + y^2)/w_0^2]$. We can then separate the x and y integrals, with each providing an expression of the form (ignoring the plane wave phase factor)

$$E_x(x', z') = \left(\frac{j}{\lambda z'}\right)^{0.5} \int \exp\left\{-\left[\frac{x^2}{w_0^2} + \frac{jk(x' - x)^2}{2z'}\right]\right\} dx, \quad (2.72)$$

where the integral extends over the range $-\infty \leq x \leq \infty$. Completing the square and taking advantage of the definite integral

$$\int_{-\infty}^{\infty} \exp(-ax^2 + bx) dx = \left[\frac{\pi}{a}\right]^{0.5} \exp\left(\frac{b^2}{4a}\right); \quad a > 0 \quad (2.73)$$

(which turns out to be a very useful expression for analysis of Gaussian beam propagation), we obtain the expression

$$E_x(x', z') = \left(\frac{j}{\lambda z'}\right)^{0.5} \left(\frac{2\pi w_0^2 z'}{2z' + jkw_0^2}\right)^{0.5} \exp\left[\frac{-k^2 x'^2 w_0^2 - 2jkz'x'^2}{4z'^2 + (kw_0^2)^2}\right] \quad (2.74)$$

The real and imaginary parts of the exponential are suggestive, and after some manipulation, we find that

$$E_x(x', z') = \left(\frac{w_0}{w}\right)^{0.5} \exp\left(\frac{-x'^2}{w^2} - \frac{j\pi x'^2}{\lambda R} - \frac{j\phi_0}{2}\right), \quad (2.75)$$

together with the variation of w , R , and ϕ_0 given by equations 2.26b to 2.26d. Combining the x and y integrals and the plane wave phase factor, we see that the propagation of the fundamental mode Gaussian beam can be directly obtained from a diffraction integral approach. The same is true of the higher order Gaussian beam modes, but this involves considerably greater mathematical complexity.

2.10 BIBLIOGRAPHIC NOTES

Since almost every text on optics and optical engineering covers Gaussian beam propagation at some level, it is impossible to give a complete list of these references. However, texts which have been particularly useful to the author are [ARNA76], [MARC75], [SIEG86] (Chapters 16 and 17, pp. 626–697), and [YARI71]. Some of the more comprehensive review articles that cover fundamental and higher order Gaussian beam modes are [KOG66] and [MART89].

Diffraction theory is covered extensively in the texts [BORN65] and [SIEG86], as well as many named in the other references on Gaussian beams.

With the idea of being helpful to the reader, I point out that the discussions of higher order Gaussian beam modes, in particular, seem to be fraught with typographical errors. In equation (64) of [SIEG86] the factor $(1 + \delta_{0m})$ should be omitted, and the last exponential should be $\exp(jm\varphi)$. Equation 3.3 of [MART89] should have the term $(-R^2)^l$ rather than $(-R)^l$, and the terms in equation 3.4 should have an additional factor $n!$. In equation 3.11 of this reference, the delta function should be δ_{0m} . The present work is hopefully free of these errors, but almost inevitably will contain others. The author would be grateful to any reader identifying such problems and bringing them to his attention.

Gaussian beam propagation is also discussed extensively in some of the references given in Chapter 1, particularly those by [GOUB68] and [GOUB69]. Other useful references include the articles [CHU66], [KOG65], [KOG66], [MART78], and [MART89]. The last reference also includes an interesting discussion of the paraxial limit.

Depictions of the higher order Gaussian beam modes can be found in a number of places, with a relatively complete presentation being given by [MOOS91]. The behavior of higher order modes with $p = 0$ is discussed by [PAXT84].

A variety of alternative approaches have been developed for analysis of Gaussian beam propagation. These include the use of a complex argument for the beam modes ([SIEG73], [SIEG86]), representation of a Gaussian beam at a specified distance from its waist as point on a complex circle diagram developed by [COLL64] and by [DESC64], and geometrical constructions to describe the propagation ([LAUR67]). Gaussian beams can also be considered as complex rays, as described by [DESC71], [PRAT77], and [ARNA85]. The availability of computers makes it practical to perform numerical analyses, such as Fourier transformations and expansion in plane waves ([SIEG86], Section 16.7, pp. 656–662) in situations where Gaussian beam propagation is not effective. These alternative methods of considering Gaussian beam propagation remain valuable for the increased understanding that they provide.

The spot size of Gaussian beams is specifically discussed in articles [BRID75], [CART80], [CART82], and [PHIL83].

Gaussian beams in anisotropic media are discussed in [ERME70], and in certain conditions solutions similar to those discussed here can be obtained. In addition to the references given in Section 2.6, [CART72] discusses properties of Gaussian beams with elliptical cross sections.

A technique for recovering the complex Gaussian beam mode coefficients in a propagating beam from intensity measurements alone is presented by [ISAA93].

Alternative derivations of Gaussian beam propagation formulas are given in the texts by Siegman and by Marcuse, already cited, and in [WILL73].