Beam Like Solutions of the Wave Equation

The scalar wave function $\mathcal{E}(\vec{r}, t)$ (for example, the electric field amplitude of a linearly polarized beam) in a transparent homogeneous medium satisfies the source free wave equation

$$\left[\nabla^2 - \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathcal{E}(\vec{r}, t) = 0, \tag{2.1}$$

where ∇^2 is the three dimensional Laplacian operator, c is the speed of light in free space and n is the refractive index of the medium. For quasimonochromatic waves of angular frequency ω, the dominant time dependence is of the form $e^{-i\omega t}$. Then the spatial part of the wave function, introduced by writing $\mathcal{E}(\vec{r}, t) = \mathcal{E}(\vec{r})\, e^{-i\omega t}$, satisfies source-free Helmholtz equation

$$\left[\nabla^2 + \frac{n^2\omega^2}{c^2}\right]\mathcal{E}(\vec{r}) = 0\,. \tag{2.2}$$

Plane and spherical wave solutions of this equation were discussed in Chapter 1. We now explore beam-like solutions of this equation.

2.1 Gaussian Beams

Beam-like solutions should have the following characteristics:

(i) a predominant direction of propagation, and

(ii) a finite transverse cross-section (finite extent in directions perpendicular to the direction of propagation).

Since a beam has finite transverse size, wave diffraction will cause its cross-section to evolve as the beam propagates. However, we can still speak of a predominant direction of propagation if the diffractive change is not too rapid as the beam propgates. Let us now see how we can construct a solution with these charactersitics.

We recall that a quasimonochromatic plane wave propagating in the z-direction has the form

$$\mathcal{E}(\vec{r}) = \mathcal{E}_0 e^{i(kz-\omega t)}\,, \tag{2.3}$$

where the propagation constant k is related to the wavelength λ and angular frequency ω of the wave by

$$k = \frac{n\omega}{c} = \frac{2\pi n}{\lambda}\,. \tag{2.4}$$

The $e^{i(kz-\omega t)}$ factor ensures that the wave is propagating in the z-direction. For a plane wave the amplitude $\mathcal{E}_0$ is independent of the spatial coordinates. Using this solution as a guide, we look for beam-like solutions that propagate in the $z-$direction by writing the wave as

$$\mathcal{E}(\vec{r}) = \mathcal{E}_0(\vec{r}) e^{i(kz-\omega t)}, \tag{2.5}$$

The envelop function $\mathcal{E}_0(\vec{r})$ describes the variation of the wave amplitude in the x-y plane (beam profile) and its evolution with z. Substituting Eq. (2.5) in the wave equation we obtain

$$\left[\nabla^2 - \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathcal{E}(\vec{r},t) = \left[\left(\frac{\partial^2 \mathcal{E}_0}{\partial x^2} + \frac{\partial^2 \mathcal{E}_0}{\partial y^2}\right) + \frac{\partial^2 \mathcal{E}_0}{\partial z^2} + 2ik\frac{\partial \mathcal{E}_0}{\partial z} + \left(-k^2 + \frac{n^2\omega^2}{c^2}\right)\mathcal{E}_0\right] e^{i(kz-\omega t)} = 0\,. \tag{2.6}$$

Using the relation $k = n\omega/c$ [Eq.(2.4)] we find that $\mathcal{E}_0(\vec{r})$ satisfies

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\mathcal{E}_0(\vec{r}) + \left(\frac{\partial^2 \mathcal{E}_0(\vec{r})}{\partial z^2} + 2ik\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right) = 0\,. \tag{2.7}$$

This equation describes the evolution of beam profile with z due to diffraction arising from its finite transverse size. If the beam is many wavelengths broad and varies smoothly in transverse dimensions, its transverse profile will not change significantly over propagation distances of the order of a few wavelengths. Such beams are called paraxial beams. Their profile evolves slowly with z such that the following inequalities hold,

$$\begin{aligned} \frac{1}{k}\left|\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right| &= \frac{\lambda}{2\pi}\left|\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right| << |\mathcal{E}_0(\vec{r})|\,, \\ \frac{1}{k}\left|\frac{\partial}{\partial z}\left(\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right)\right| &= \frac{\lambda}{2\pi}\left|\frac{\partial}{\partial z}\left(\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right)\right| \ll \left|\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right|\,. \end{aligned} \tag{2.8}$$

The first inequality states the requirement that the change in the beam profile over distances of the order of a few wavelenths $\lambda\left|\frac{\partial \mathcal{E}_0(\vec{r})}{\partial z}\right|$ be a small fraction of $|\mathcal{E}_0(\vec{r})|$. The second inequality states that the rate at which the beam profile evolves does not significantly change over distances of the order of a few wavelenths, $\lambda\left|\frac{\partial}{\partial z}\frac{\partial\mathcal{E}_0(\vec{r})}{\partial z}\right| \ll \left|\frac{\partial\mathcal{E}_0(\vec{r})}{\partial z}\right|$. In view of the inequalities (2.8), the dominant variation of $\mathcal{E}_0$ with z in Eq.(2.7) is described by the first derivative term. Neglecting the $\partial^2\mathcal{E}_0/\partial z^2$ term, we find that the envelop function $\mathcal{E}_0(\vec{r})$ satisfies

$$\frac{\partial^2\mathcal{E}_0(\vec{r})}{\partial x^2} + \frac{\partial^2\mathcal{E}_0(\vec{r})}{\partial y^2} + 2ik\frac{\partial\mathcal{E}_0}{\partial z} = 0\,. \tag{2.9}$$

This equation is known as the paraxial wave equation. We can rewrite this equation as

$$\frac{\partial\mathcal{E}_0}{\partial z} = \frac{i}{2k}\left[\frac{\partial^2\mathcal{E}_0(\vec{r})}{\partial x^2} + \frac{\partial^2\mathcal{E}_0(\vec{r})}{\partial y^2}\right]\,. \tag{2.10}$$

In this form, this equation shows that the beam profile changes with propagation because of diffraction (finite exten in transverse dimensions). If the transverse beam profile does not change in transverse directions, we recover the plane wave result $\mathcal{E}_0 = \text{const}$.

2.1.1 Fundamental Gaussian Solution

The paraxial wave equation has many known solutions. These will be inroduced in due course. The simplest of these has circular cylindrical symmetry about the direction of propagation and is given by ($\rho = \sqrt{x^2+y^2}$)

$$\mathcal{E}_0(\rho, z) = \mathcal{A}\left[\frac{w_0}{w(z)}\right] e^{i\frac{k\rho^2}{2q(z)} - i\psi(z)} \tag{2.11a}$$

$$w(z) = w_0\sqrt{1+(z/z_R)^2} \tag{2.11b}$$

$$z_R = \frac{1}{2}kw_0^2 = \frac{\pi w_0^2 n}{\lambda} \tag{2.11c}$$

$$\frac{1}{q(z)} = \frac{1}{R(z)} + i\frac{2}{kw^2(z)} \tag{2.11d}$$

$$R(z) = z + \frac{z_R^2}{z} = z\left(1+\frac{z_R^2}{z^2}\right) \tag{2.11e}$$

$$\psi(z) = \tan^{-1}\left(\frac{z}{z_R}\right) \tag{2.11f}$$

To see the physical meaning of various terms in this equation, let us first consider the time averaged intensity $I(\rho, z) = \frac{1}{2}\epsilon_0 nc\mathcal{E}\mathcal{E}^*$, which will be given by

$$I(\rho, z) = I_0\left[\frac{w_0}{w(z)}\right]^2 \exp\left[-\frac{2\rho^2}{w^2(z)}\right]\,. \tag{2.12}$$

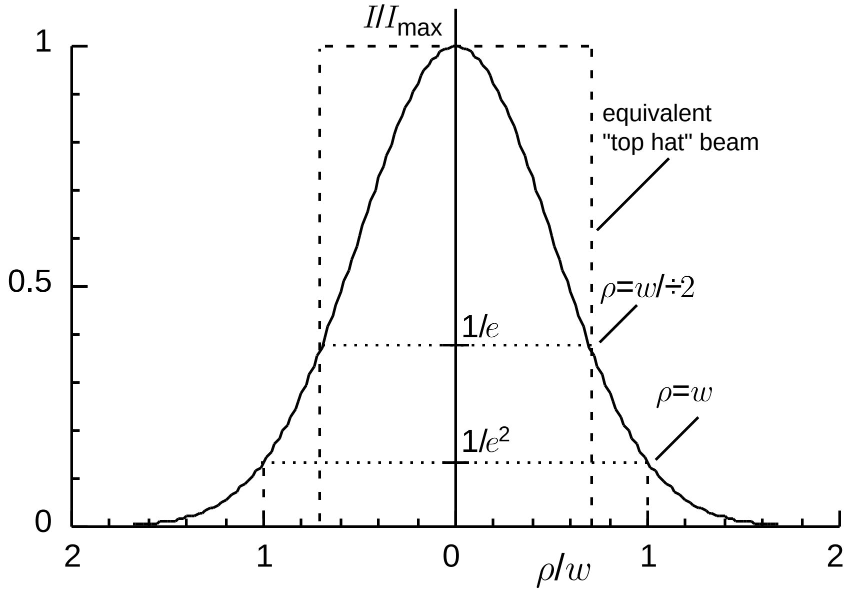

Figure 2.1: Intensity profile of a circular symmetric gaussian beam along the x (or y) axis ($\rho = x$).

This has the form of a gaussian distributions in ρ. For this reason this beam is called a gaussian beam. In terms of the total power of the beam

$$P = \int_0^{2\pi} d\varphi \int_0^{\infty} \rho d\rho I(\rho, z) = I_0 w_0^2 2\pi \times \frac{1}{4} \tag{2.13}$$

we can express the intensity as

$$I(\rho, z) = \frac{2P}{\pi w^2(z)} e^{-2\rho^2/w^2(z)} \,. \tag{2.14}$$

Figure (2.1) shows the intensity profile of the beam as a function of distance ρ (along the $x-$ axis) measured from the $z-$axis. The intensity attains its peak value $I_{\max} = 2P/\pi w^2(z)$ at the center of the beam ($\rho = 0$) and falls to $1/e^2 \approx 14\%$ of this value when $\rho = w(z)$. $w(z)$ is referred to as the beam spot size (radius). The intensity of a gaussian beam falls off rapidly as ρ increases beyond the spot size $w(z)$.

The form of the peak intensity $2P/\pi w^2(z)$ suggests another measure of beam size. If we were to imagine a circular cylindrical beam of uniform intensity and the same total power as a gaussian beam, the radius of such a beam will be

$$w_{\text{TH}} = w/\sqrt{2} \,. \tag{2.15}$$

Such a uniform beam is referred to as a "top hat beam" because the intensity distribution of such a beam will have the shape of a top hat [See Fig. (2.1)].

There are Other measures of beam size. For example, we can use a criterion based on power the trasmitted by an aperure. A circular aperture of radius a placed at the center of a gaussian beam will trasnmit a fraction of power

$$\frac{P_{\rm T}}{P} = \frac{2}{\pi w^2}\int_0^{2\pi} d\varphi \int_0^a \rho d\rho e^{-(2\rho^2/w^2)} = 1 - e^{-(2a^2/w^2)}\,. \tag{2.16}$$

This fraction as a function of aperture radius is plotted in the Fig.(2.2).

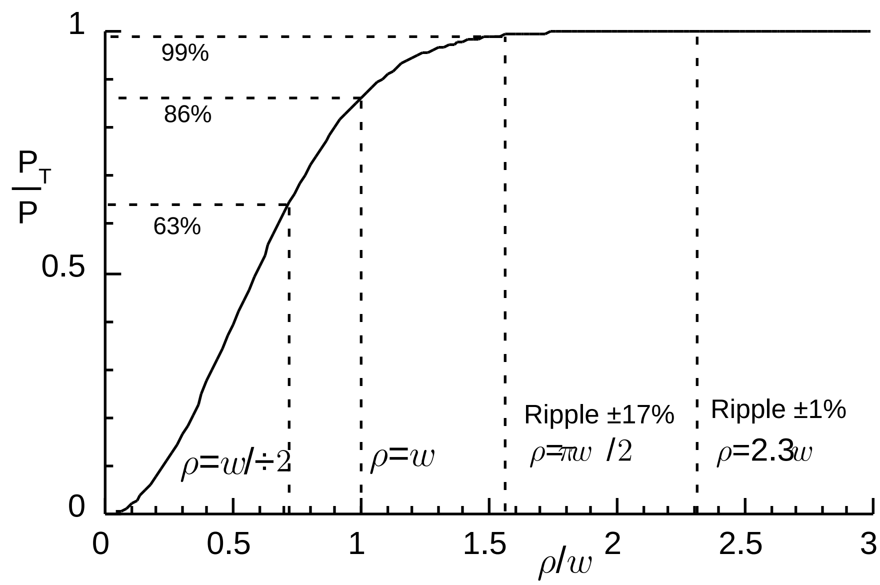

Figure 2.2: Fractional power of a gaussian beam transmitted by a circular aperture of radius ρ centered on the beam.

An aperture of radius $w_{\rm TH}$ will transmit only 63% power while an aperture of radius w will transmit 86% power. An aperture of radius $\pi w/2$ will pass 99% power. We find that considerably larger apertures than those with radius w are needed to pass the power gaussian beam of spot size w.

Beyond optimizing power transmission of a gaussian beam, we may also want to minimize diffraction ripples which will significantly distort the intensity distribution of the transmitted beam. Such ripples will be present whenever sharp edged circular apertures are used even if they pass a large fraction of the total power. A sharp circular aperture of radius $\pi w/2$, which passes 99% of the total power, will cause ripples with intensity variations of 17% in the near field and a peak intensity reduction of the same amount in the far field. To keep diffraction riplles down to 1% in the beam transmitted by a sharp edged circular aperture we must employ an aperture of radius $\approx 2.3w$. [Siegman, Chapter 18].

From the preceding discussion, it is clear that $w(z)$ is a measure of the transverse size of the beam. It measures the size of the beam spot, which of course, varies as the beam propagates. Figure (2.3) shows the variation of spot size as a function of the propagation distance z measured from the "beam waist" which is defined to be the plane where the spot size has its minimum value w_0. In writing the expression for beam spot size (2.11b) have chosen this plane to be the $z = 0$-plane. An examination of Eqs (2.11a)-(2.11f) shows that a gaussian beam is uniquely determined by the location of its waist the (minimum) spot size w_0.

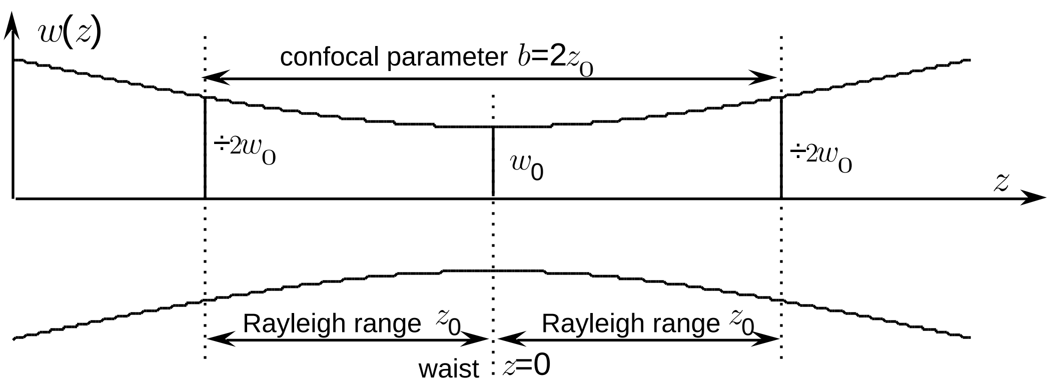

Figure 2.3: Variation of spot size with propgation near beam waist.

A gaussian beam spreads in a nonlinear fashion during its propagation. Near the waist the spread is slow so that the beam remains collimated. Far from the waist the beam spreads linearly with distance from the waist. The characteristic distance the beams travels from the waist before the spot size increase to $\sqrt{2}w_0$ (or the area doubles) is

$$z_R = kw_0^2/2 = \frac{\pi n w_0^2}{\lambda}\,. \tag{2.17}$$

This distance z_R is called the Rayleigh range. Notice that there are two points located on the opposite sides of the beam waist where the spot size has the value $\sqrt{2}w_0$. The distance between these $\sqrt{2}w_0$ spot size points is the confocal parameter

$$b = 2z_R = \frac{2n\pi w_0^2}{\lambda}\,. \tag{2.18}$$

Confocal parameter b is a measure of the distance over which a beam may be considered to have uniform cross-section near its waist. It plays an important role in the theory of laser resonators which will be discussed shortly.

In the far field $z \gg z_R$ the spot size grows linearly with z as [see Fig. (2.20)]

$$w(z) \approx w_0 \frac{z}{z_R}. \tag{2.19}$$

The far field divergence angle of the beam may be defined by the ratio of the far field spot size to the distance

$$\theta = \frac{w(z)}{z} = \frac{w_0}{z_R} = \frac{\lambda}{n\pi w_0}. \tag{2.20}$$

For paraxial approximation to be good we require $\theta < 1/\pi$ which translates to minimum spot size $w_0 > \lambda/n$. From the dependence of the confocal parameter [Eq. (2.18)] and the beam divergence angle [Eq. (2.20)] on w_0 we see that a beams with smaller waist spot size will remain collimated over shorter distances and will spread more rapidly in the far zone.

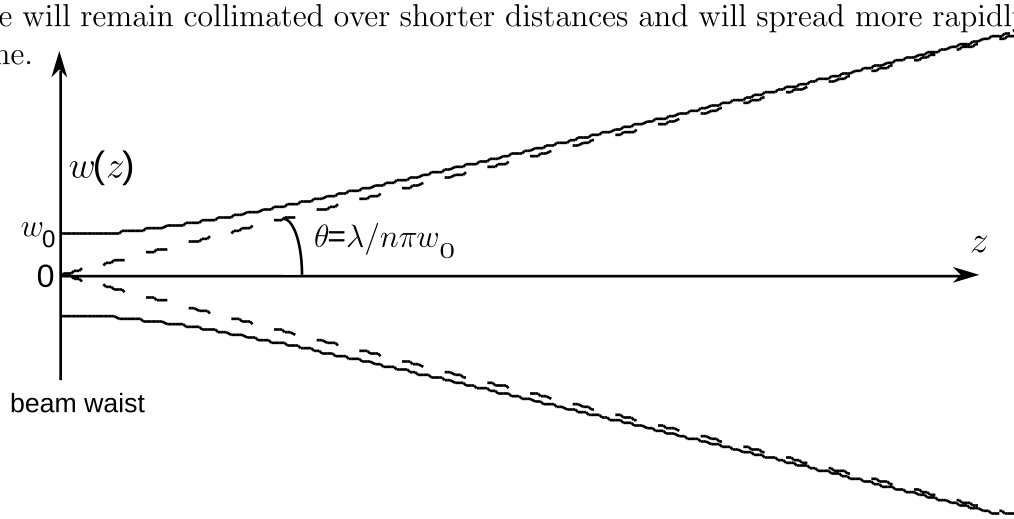

Figure 2.4: Far field divergence of a gaussian beam.

The origin of gaussian beam divergence is diffraction, which arises whenever a wave is confined to a finite transverse size. In fact, the far field beam divergence angle is of the same order as the angle associated with the Fraunhoffer (far field) diffraction of a plane wave by an aperture of radius $a \sim w_0$

$$\theta_D = 0.61 \frac{\lambda}{na}. \tag{2.21}$$

Since diffractive phenomena cannot be described by ray optics, wave optics must always be used when dealing with gaussian beams.

The parameter $R(z)$ [Eq.(2.11e)] is the radius of curvature of the very nearly spherical phasefronts at z. This can be seen by writing the gaussian beam $\mathcal{E}_0(\rho, z)e^{i(kz-\omega t)}$ [Eq.(2.11a)] in the limit $z >> z_{_R}$

$$\begin{aligned}\mathcal{E}(\vec{r}, t) &= \frac{\mathcal{A}e^{-i\psi(z)}}{\sqrt{1+z^2/z_{_R}^2}}\, e^{-\rho^2/w^2(z)}\, e^{ik(z+\frac{\rho^2}{2R(z)})}\, e^{-i\omega t} \\ &\approx -i\mathcal{A}z_{_R}\, e^{-\rho^2/w^2(z)} \left[\frac{1}{z}e^{ik(z+\frac{\rho^2}{2z})}\right] e^{-i\omega t}\,, \end{aligned} \tag{2.22}$$

where we have used the approximation

$$R(z) = z\left(1 + z_{_R}^2/z^2\right) \approx z\,, \quad z \gg z_{_R}\,. \tag{2.23}$$

Let us compare it with a spherical wave emitted by a point source on the z-axis at $z = 0$. Near the $z-$axis this wave has the form

$$\begin{aligned}\mathcal{E}_s(\vec{r}, t) &= \mathcal{E}_0\frac{e^{i(kr-\omega t)}}{r} = \mathcal{E}_0\frac{e^{ik\sqrt{\rho^2+z^2}}}{\sqrt{\rho^2+z^2}}\, e^{-i\omega t} \\ &\approx \mathcal{E}_0\left[\frac{e^{ik(z+\frac{\rho^2}{2z})}}{z}\right] e^{-i\omega t}\,, \quad \rho^2 \ll z^2\,. \end{aligned} \tag{2.24}$$

A comparison of Eqs. (2.22) and (2.24) shows that a gaussian beam has the form of a spherical wave with center of curvature located at the beam waist $z = 0$. In fact for points close to the $z-$ axis such that $\rho^2 \ll w^2(z)$, Eq. (2.22), has exactly the form of a spherical wave with spherical wavefronts

$$kz + \frac{k\rho^2}{2z} \approx k\sqrt{\rho^2+z^2} \equiv kr = \text{const}\,. \tag{2.25}$$

It is clear that $R(z)$ is the radius of curvature of constant phase surfaces which coincide with wavefronts near the $z-$axis. The sign of R is chosen according to the convention used in discussing ABCD matrices; that is, $R(z)$ is negative for a converging beam (center of curvature lies in front of the wavefront) and positive for a diverging beam (center of curvature lies behind the wavefront).

The radius of curvature $R(z)$ of the wavefronts of a gaussian beam has the following limiting forms

$$R(z) = \begin{cases} \pm\infty & |z| << z_{_R} \\ 2z_{_R} \equiv b & z = z_{_R} \\ z & |z| >> z_{_R} \end{cases} \tag{2.26}$$

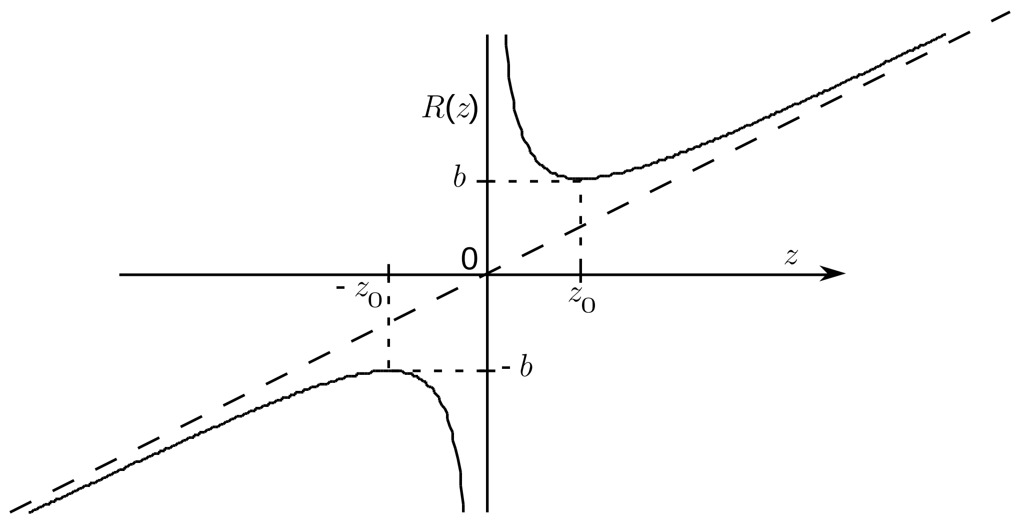

Figure 2.5: Variation of the radius of curvature of a gaussian beam with distance.

Its variation with propagation distance z measured from the waist is shown in Figs. (2.5) and (2.6). The wavefront is flat or planar at the waist corresponding to an infinite radius of curvature. As the beam propagates away from the waist the wavefronts become curved, and the radius of curvature drops until $z = z_R$. Beyond $z = z_R$ the radius of curvature increases again as $R(z) \approx z$. The minimum radius of curvature occurs for the wavefront at $z = z_R$ with the radius of curvature equal to the confocal parameter $R = 2z_R = b$. The center of curvature for the wavefront at $z = z_R$ is located at $z = -z_R$ and the center of curvature of the wavefront at $z = -z_R$ is located at $z = z_R$ as seen in Fig. (2.6). The curved wavefronts at $z = \pm z_R$ have special significance in stable resonator theory. If these wavefronts are replaced by two mirrors with matching radii of curvature, we will form a stable resonator. This resonator will have mirrors of radius of curvature R and spacing L with $R = b = 2z_R = L$. Since the focal length of a mirror of radius R is $f = R/2$, the focal points of these two mirrors will coincide at the center of the resonator. The two mirrors then form a symmetric confocal resonator, thus giving rise to the *confocal parameter* $b = 2z_R$.

The direction of energy flow in a gaussian beam is indicated in Figure (2.6) for a gaussian beam traveling from left to right. Energy is transported along rays which are the directed curves

$$\rho \equiv \sqrt{x^2 + y^2} = \rho_0 \sqrt{1 + z^2/z_R^2}\,, \qquad (2.27)$$

where ρ_0 is the ray coordinate at the waist. The Poynting vector - $\vec{S}$ which describes

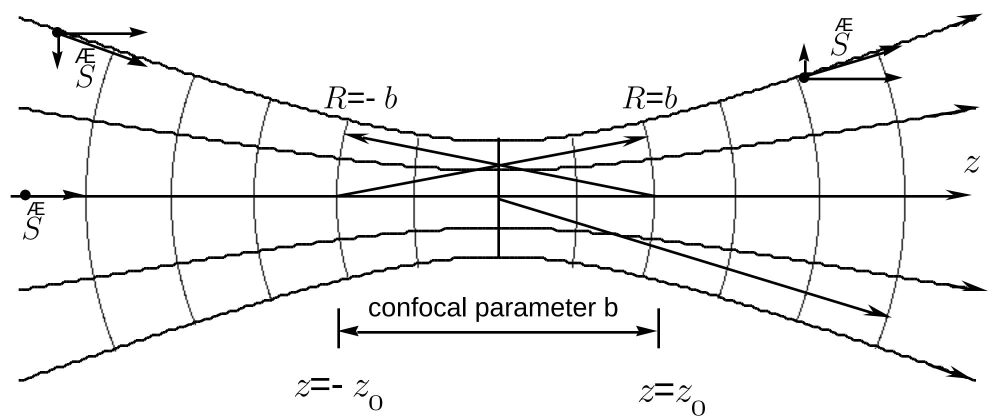

Figure 2.6: Wavefronts and ray trajectories in the waist region for a wave moving from left to right. Far from the beam waist the wavefronts are part of spheres with center of curvature at the center of the waist. The hyperbolas are rays indicating the direction of energy flow in a gaussian beam.

the flow of energy (energy flux density, W/m^2) is tangential to the rays. To the left of the waist, Poynting vector has a small radial component pointing toward the axis corresponding to a converging (focusing) beam. To the right of the waist it has a small radial component pointing away from the axis coresponding to a diverging (defocusing) beam. Energy flow across the waist is from left to right as in a plane wave moving in the z-direction.

Finally, we note that in passing the waist the phase angle $\psi(z)$ [Eq. (2.11f)] changes from $-\pi/2$ (far to the left of the waist) to $\pi/2$ (far to the right). Thus the fundamental gaussian beam acquires an extra phase of π in passing through its focal region. This phase shift, called the Guoy phase shift, is in fact a special case of the Guoy effect which says that a wave with a reasonably simple cross-section will acquire an extra phase shift of π in passing through a focal region. Higher order gaussian beams acquire larger Guoy phase shifts on account of their more complicated cross-section.

Refrences

Deviations from paraxial theory become significant for beams that are focused so tightly that beam divergence angles exceed $\theta > 1/\pi$. Corrections to these solutions are discussed in M. Lax, W. H. Louisell, and W. B. McKnight, *From Maxwell to paraxial optics*, Phys. Rev. A

11, 1365 (1975); L. W. Davis, *Theory of electromagnetic beams*, Phys. Rev. A **19**, 1177 (1975); G. P. Agarwal and D. N. Pattanayak, *Gaussian beam propagation beyond the paraxial approximation*, J. Opt. Soc. Am. **69**, 575 (1979).

A paper dealing with Gaussian beam solution of Maxwell's equations including corrections to paraxial theory and describing experimental observations of these corrections is W. L. Erikson and S. Singh, *Ploarization properties of Maxwell-Gaussian beams*, Phys. Rev. E **49**, 5778 (1994).

2.2 The ABCD Law for Gaussian Beams

During the propagation of a gaussian beam in a homogeneous medium beam spot size $w(z)$ and the radius of curvature of phasefront $R(z)$ change according to Eqs. (2.11b) and (2.11e) but the beam retains its gaussian shape. We have seen that a geometrical ray picture fails to account for this evolution.

The evolution of $R(z)$ and $w(z)$ during propagation is equivalent to an evolution of the complex beam parameter $q(z)$. We have seen that geometric ray picture cannot be used to describe the propagation of a gaussian beam. It turns out, however, that gaussian beam propagation is described in terms of a very simple propagation law for the q-parameter of a gaussian wave.

When a gaussian beam passes through an optical element whose ray matrix is given by (ABCD), its q-parameter changes. Complex beam parameter q_o just after the beam emerges from the optical element is related to the complex parameter q_i just before the beam enters the optical element by

$$q_o = \frac{q_i\, A + B}{q_i\, C + D} \tag{2.28}$$

This important relation is known as the ABCD law of a gaussian beam. This relation follows from paraxial diffraction theory (paraxial wave equation)

Refrences

1. H. Kogelnik and T. Li, *Laser Beams and Resonators*, Proc. of IEEE **54**, 1312-1329 (1966).

2. A. Siegman, *Lasers* (University Science Books, Mill Valley, CA 1986), Chapter 20.

Example 1: *Gaussian beam propagation in a homogeneous medium starting at its waist.*

Let us take the beam waist location to be the $z = 0$ plane. At the waist, the wavefronts are planar, so that the complex beam parameter is pure imaginary given by

$$q_i = -iz_{_R} = -i\frac{\pi w_0^2}{\lambda}\,. \tag{2.29}$$

The ABCD matrix for free propagation over a distance z in a homogeneous medium is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \tag{2.30}$$

Using this matrix, we find that the beam parameter at a distance z from the waist will be given by

$$q(z) = \frac{q_i\, A + B}{q_i\, C + D} = q_i + z \tag{2.31}$$

or
$$\frac{1}{q(z)} = \frac{1}{q_i + z} \tag{2.32}$$

or
$$\frac{1}{R(z)} + i\frac{2}{kw^2(z)} = \frac{1}{-iz_{_R} + z} = \frac{iz_{_R} + z}{z_{_R}^2 + z^2} \tag{2.33}$$

Equating the real and imaginary parts from the two sides, we obtain the familiar expressions $[\frac{1}{2}kw^2 = \frac{1}{2}kw_0^2(w/w_0)^2 = z_{_R}(w/w_0)^2]$ for the wavefront radius of curvature and spot size

$$R(z) = \frac{z^2 + z_{_R}^2}{z} = z + \frac{z_{_R}^2}{z} \tag{2.34}$$

$$w(z) = w_0\sqrt{\frac{z_{_R}^2 + z^2}{z_{_R}^2}} = w_0\sqrt{1 + \frac{z^2}{z_{_R}^2}} \tag{2.35}$$

Example 2: *Passage of a gaussian beam through a lens.*

The ray trasnfer matrix for the lens is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \tag{2.36}$$

Let q_i be the incident beam parameter just before the lens. Then the output beam parameter q_o (just after the lens) is given by

$$q_0 = \frac{q_i\, A + B}{q_i\, C + D} = \frac{q_i}{-(q_i/f) + 1} \tag{2.37}$$

or
$$\frac{1}{q_o} = -\frac{1}{f} + \frac{1}{q_i} \tag{2.38}$$

or
$$\frac{1}{R_o(z)} + \frac{2i}{kw_i^2} = -\frac{1}{f} + \frac{1}{R_i(z)} + \frac{2i}{kw_i^2} \tag{2.39}$$

Equating the real and imaginary parts on the two sides gives

$$\frac{1}{R_o} = \frac{1}{R_i} - \frac{1}{f} \tag{2.40}$$

$$w_o = w_i \tag{2.41}$$

Thus, a lens changes the curvature of the phasefront but leaves the spot size unaffected. A related problem is the focusing of a gaussian beam by a mirror of focal length $f = R/2$.

2.2.1 Gaussian Beam Focusing

Consider a gaussian beam incident from left on a lens of focal length f. Let the incident beam waist be located a distance d_1 from the lens and let the spot size there be w_{01}. After passing through the lens the beam has a new waist at d_2 and spot size w_{02} at the new waist. We are interested in finding d_2 and w_{02}.

The ray transfer martix for beam passage from the first waist to the second waist is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{pmatrix} \tag{2.42}$$

The complex beam parameter q_2 at the second waist is then given by

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D} = \frac{(1 - d_2/f)q_1 + (d_1 + d_2 - d_1 d_2/f)}{-q_1/f + (1 - d_1/f)}\,, \tag{2.43}$$

where q_1 is the beam parameter at the first waist. At the two waists, the complex beam parameters are pure imaginary,

$$q_1 = -iz_{R1} \equiv -i\pi n w_{01}^2/\lambda, \quad q_2 = -iz_{R2} \equiv -i\pi n w_{02}^2/\lambda. \tag{2.44}$$

Using these in the transformation equation (2.43) we obtain

$$\begin{aligned} -iz_{R2} &= \frac{-iz_{R1}A+B}{-iz_{R1}C+D} = \frac{(-iz_{R1}A+B)(iz_{R1}A+B)}{(z_{R1}C)^2+D^2} \\ &= \frac{-iz_{R1}(AD-BC)+(z_{R1}^2AC+BD)}{(z_{R1}C)^2+D^2} \end{aligned} \tag{2.45}$$

Equating the real an imaginary parts of the expression on the right hand side to the corresponding terms on the left we find

$$\Re[q_2] \equiv 0 = \frac{z_{R1}^2AC+BD}{(z_{R1}C)^2+D^2} \tag{2.46}$$

$$\Im[q_2] \equiv -z_{R2} = \frac{-z_{R1}(AD-BC)}{(z_{R1}C)^2+D^2} \tag{2.47}$$

Using the fact $AD - BC = 1$ and $z_{0i} = \pi n_i w_{0i}^2/\lambda$ we find from Eq. (2.47) that the new waist spot size is given by

$$w_{02}^2 = \frac{w_{01}^2}{(z_{R1}/f)^2+(1-d_1/f)^2} = \left(\frac{\lambda f}{\pi n w_{01}}\right)^2 \frac{1}{1+(f/z_{R1})^2(1-d_1/f)^2}\,. \tag{2.48}$$

From Eq. (2.46) we find, since the denominator is not zero, $z_{R1}^2AC+BC = 0$ which leads us to

$$z_{R1}^2\left(1-\frac{d_2}{f}\right)\left(-\frac{1}{f}\right)+\left(d_1+d_2-\frac{d_1d_2}{f}\right)\left(1-\frac{d_1}{f}\right)=0 \tag{2.49}$$

On simplifying and solving this equation for d_2 we obtain

$$d_2 = \frac{z_{R1}^2/f - d_1(1-d_1/f)}{(z_{R1}/f)^2+(1-d_1/f)^2} = f\left[1-\frac{(1-d_1/f)}{(z_{R1}/f)^2+(1-d_1/f)^2}\right] \tag{2.50}$$

A plot of exit waist position d_2 as a function of the incident waist position d_1 is shown in Figure (2.7). To see the variation of the exit waist spot size with d_1, we find it is convenient to plot the Rayleigh range $z_{R2} = \pi w_{02}^2/\lambda$ which is a measure of the spot size as a function of d_1. This is shown in Fig. (2.8).

Let us compare these results with the predictions of geometrical optics. If we consider the outgoing beam waist as the image of the incident beam waist, then geometrical optics gives the location of new beam waist d_2 and spot size w_{o2} to be

$$d_2 = \frac{fd_1}{d_1-f} = f\left[1-\frac{1}{1-d_1/f}\right] \tag{2.51}$$

$$w_{o2}^2 = w_{01}^2 \times \left(-\frac{d_2}{d_1}\right)^2 = \frac{w_{01}^2}{(1-d_1/f)^2} \tag{2.52}$$

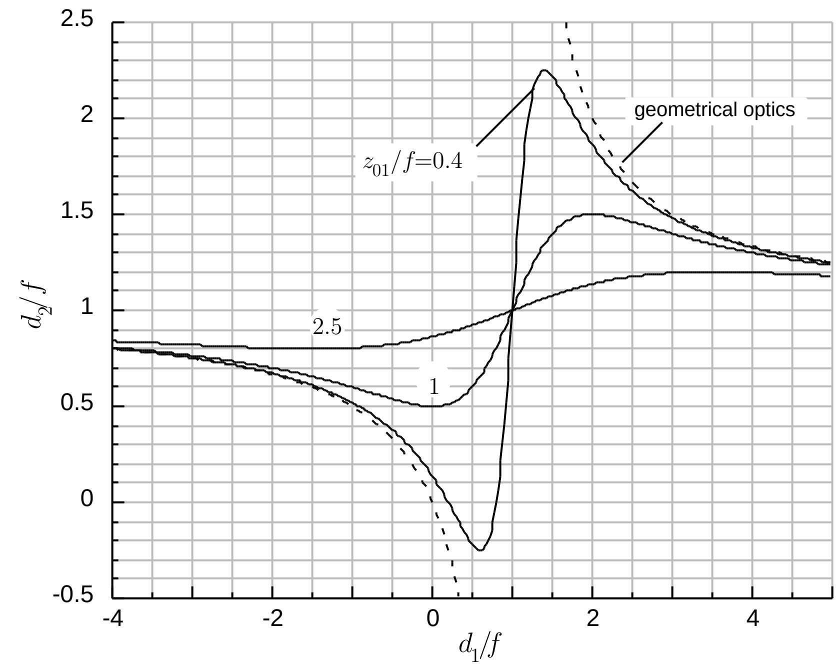

Figure 2.7: Beam waist location d_2 after a gaussian beam passes through a positive lens of focal length f as a function of the incident beam waist location d_1 for different values of the ratio z_{R1}/f.

The prediction of geometrical optics for the second beam waist location d_2 is shown by the dashed curve in Fig.(2.7). We see that gaussian beam and geometrical optics predictions agree when $|d_1/f| \gg 1$ and $z_{_{R1}}/f \ll 1$, that is, when the lens is located in the far zone of the incident beam waist. The disagreement between the two predictio ns is complete as $d_1 \to f$. In this case, geometrical optics predicts $d_2 \to \infty$ and $w_{02}^2 \to \infty$, whereas gaussian beam results are

$$d_2 \to f \quad \text{and} \quad w_{02}^2 = \left(\frac{f\lambda}{\pi n w_{01}}\right)^2 . \tag{2.53}$$

A noteworthy features of Fig. (2.7) is that the distance d_2 for the second waist from the lens has a maximum. The maximum occurs for $d_1 \approx z_{01}$. Similarly, the spot size for a gaussian beam after passing through has a finite maximum which is attained for $d_1 \approx f$.

A question of practical importance when discussing applications such as laser traps, cutting, drilling, and laser fusion is what size focal spots are possible. For a given focal length f, we can reduce the size of w_{02} by making $z_{_{R1}}/f$ large. Since $z_{_{R1}} = \pi n w_{01}^2$,

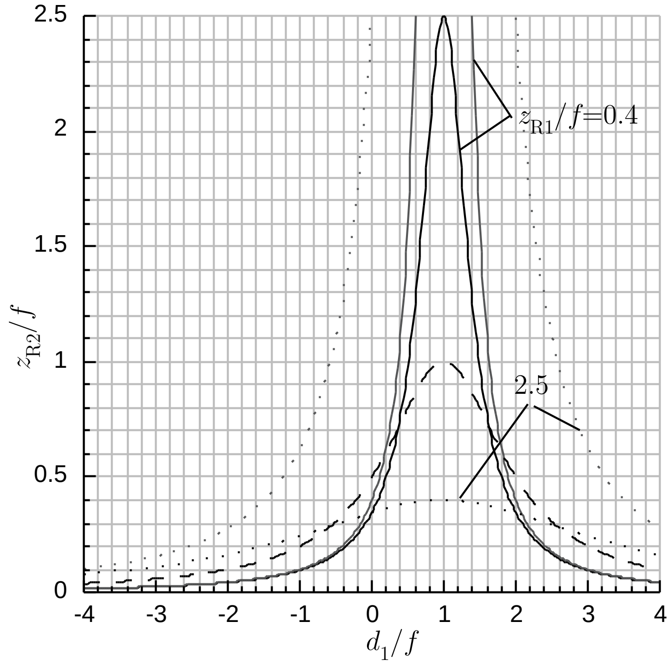

Figure 2.8: Gaussian beam spot size in terms of Rayleigh range after focusing by the lens of focal length f for several different values of z_{R1}/f. Geometrical optics predictions are shown by grey curves for $z_{R1}/f = 0.4$ and 2.5. The dashed curve is the gaussian beam result for $z_{R1}/f = 1$.

this means we need to make w_{01} large. But w_{01} is limited to be smaller than the lens aperture and even smaller if we allow for beam spreading from the first waist to the lens. One way to address this is to collimate the incident beam with a confocal parameter many focal length f long and place the lens in the near zone so that the input beam spot size does not change significantly from the first beam waist to the lens. Under these conditions w_{01} is limited by the lens aperture. If we want the lens to transmit 99% of the incident power, then w_{01} is given by

$$\frac{1}{2}\pi w_{01} = \frac{1}{2}D \tag{2.54}$$

where D is the lens aperture. Then the second spot size

$$w_{02}^2 = \left(\frac{f\lambda}{n\pi w_{01}}\right)^2 = \left(\frac{f\lambda}{nD}\right)^2 \tag{2.55}$$

$$\text{or} \qquad w_{02} = \frac{f\lambda}{nD} = \frac{\lambda}{n}\frac{f}{D} = \frac{\lambda}{n} \times (f^{\#}) \tag{2.56}$$

Here $f^{\#}$ is the f-number of the lens. A small $f^{\#} \Rightarrow$ a fast lens and a large $f^{\#} \Rightarrow$ a slow lens. Best lenses have $f^{\#} \approx 1$. Most lenses have $f^{\#} > 1$. Thus the focal spot radius is about the size of a wavelength.

The location of the focal spot d_2 is then

$$\frac{d_2}{f} = 1 - \frac{f^2}{z_{R1}^2} = 1 - \left(\frac{f\lambda}{n\pi w_{01}^2}\right)^2 = 1 - \left(\frac{n\pi w_{02}^2}{\lambda}\right)^2 \times \frac{1}{f^2} = 1 - \frac{z_{R2}^2}{f^2} \tag{2.57}$$

$$\text{or} \qquad d_2 = f - \frac{z_{R2}^2}{f}\,. \tag{2.58}$$

Since z_{R2} is the Rayleigh range for the focused beam, it is usually much less than f. Hence, the factor $z_{R2}^2/f \ll f$ so that the focal spot is very nearly in the focal plane of the lens.

The peak intensity of the focused beam is

$$I_{02} = \frac{2P}{\pi w_{02}^2} = \frac{2P}{\pi\left(\frac{\lambda f}{\pi n w_{01}}\right)^2} = \frac{2P}{f^2\pi\theta_1^2} = \frac{2P}{f^2\Omega_1} \tag{2.59}$$

where Ω_1 is the solid angle into which the incident beam waist radiates. FIGURE with equations

The peak intensity at the first beam waist

$$I_{01} = \frac{2P}{\pi w_{01}^2} = \frac{2P}{\lambda^2}\frac{n^2\lambda^2\pi}{n^2\pi^2 w_{01}^2} = \frac{2P}{(\lambda/n)^2}\pi\theta^2 \tag{2.60}$$

$$= \frac{2P}{(\lambda/n)^2}\Omega_1\,. \tag{2.61}$$

The quantity $B = I_{01}/\Omega_1$ is called the brightness of the beam (W/m^2·Sr). The brightness of a source is an invariant in the sense that linear optics elements (mirrors, lenses etc.) do not change it. Note that B is the power emitted per unit area per unit solid angle.

The result for the peak intensity in the focal spot,

$$I_{02} = \frac{2P}{f^2\Omega_1} \qquad \text{check:} \quad \frac{I_{02}}{\Omega_2} = \frac{2P}{(\lambda/n)^2}\,, \tag{2.60*}$$

then indicates that the intensity that can be obtained for a beam of given power is directly proportional its brightness. It is also inversely proportional to the solid angle divergence of the beam being focused. The directionality (smallness of Ω) of laser beams is of crucial importance for obtaining high intensities.

By contrast, a thermal source (ordinary lamp) emits in all directions (2π steradian). If it delivers a power P over an ideal lens aperture, the focal intensity is FIGURE

$$I \approx \frac{P}{f^2}\frac{1}{2\pi}\,. \tag{2.62}$$

A laser on the other hand, emits into a solid angle Ω so the focal intensity is

$$I_{02} = \frac{2P}{f^2\Omega}\,. \tag{2.63}$$

Example: Consider a He:Ne laser with $P = 1$ mW, λ=633 nm, and spot size w_{01}=1 mm. Then its divergence angle ($n = 1$), solid angle and intensity are

$$\begin{aligned}\theta &= \frac{\lambda}{\pi w_0} = \frac{0.633\times10^{-6}}{\pi\times10^{-3}} \approx 0.2\times10^{-3}\,\mathrm{rad}\\ \Omega &= \pi\theta^2 = \pi(0.2\times10^{-3})^2 \approx 1.3\times10^{-7}\,\mathrm{sr}\\ I_0 &= \frac{2P}{\pi w_0^2} = \frac{2\times10^{-3}}{\pi\times(10^{-1})^2}\,\mathrm{W/cm^2}\end{aligned}$$

If this laser is focused by a lens of $f = 2.5$ cm (the human eye), the peak intensity will be

$$I_{02} = \frac{2P}{f^2\Omega_1} = \frac{2\times10^{-3}}{6.25\times1.3\times10^{-7}} \approx 2.5\times10^3\,\mathrm{W/cm^2}\,. \tag{2.64}$$

Thus, direct viewing of even a lower-power laser beam can result in severe retinal damage. Thermal lamps would have to emit hundreds of thousands of watts to match the intensities achievable by focusing even modest power lasers.

The large intensities achievable by lasers are a direct consequence of their low divergence. While care must be exercised in dealing with laser beams, their use in repairing detached retinas and other surgical procedures has become practically routine.

2.2.2 Hermite Gaussian beam solutions

So far we have discussed only the fundamental gaussian beam solution. There are other solutions which have more complicated spatial structure. In general, these solutions of paraxial wave equation (2.9) are labeled by two indices. The well-known solutions with rectangular symmetry (separable in Cartesian coordinates) are the Hermite-Gaussian solutions given by

$$[\mathcal{E}_0]_{mn}(\vec{r}) = \mathcal{A}\frac{w_0}{w}\, H_m(\sqrt{2}x/w)H_n(\sqrt{2}y/w)e^{-i(m+n+1)\psi+ik\rho^2/2q}. \tag{2.65}$$

Here we have suppressed the $z-$dependence of $w(z)$, complex beam parameter $q(z)$ and phase $\psi(z)$ for simplicity of writing. They are independent of the beam indices and are given by Eqs. (2.11b)-(2.11f). A Hermite-gaussian beam of indices m, n is sometimes denoted by HG_{mn}.

$H_m(x)$ in Eq.(2.65) is a Hermite polynomial of degree m and argument x. Some low order Hermite polynomials and recursion relations for computing the higher order ones are listed below

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x \\ H_{m+1}(x) &= 2xH_m(x) - 2mH_{m-1}(x) \\ \frac{dH_m(x)}{dx} &= 2mH_{m-1}(x) \end{aligned} \tag{2.66}$$

Hermite-Gaussian beams maintain their form during propagation. Spot size $w(z)$ sets the length scale over which the beam profile changes significantly in transverse directions, and z_R sets the length scale over which beam profile changes significantly as the wave propagates. The intensity distribution for the beam with indices m, n is given by

$$I_{mn}(x, y) = I_0\frac{w_0^2}{w^2}H_m^2(\sqrt{2}x/w)H_n^2(\sqrt{2}y/w)e^{-2(x^2+y^2)/w^2}\,, \tag{2.67}$$

where I_0 is given by

$$I_0 = \frac{1}{2}\epsilon_0 n|\mathcal{A}|^2 c\,. \tag{2.68}$$

The total power of mode with indices m, n is

$$\begin{aligned} P &= \iint_{-\infty}^{\infty} dxdy I_{mn}(x, y, z) = I_0\frac{w_0^2}{w^2}\iint_{-\infty}^{\infty} dxdy H_m^2(\sqrt{2}x/w)H_n^2(\sqrt{2}y/w)e^{-2(x^2+y^2)/w^2} \\ &= \frac{1}{2}I_0 w_0^2\left[\int_{-\infty}^{\infty} dX H_m^2(X)e^{-X^2}\right]\cdot\left[\int_{-\infty}^{\infty} dY H_n^2(Y)e^{-Y^2}\right] \\ &= \frac{1}{2}I_0 w_0^2\left[2^m\sqrt{\pi}m!\right]\cdot\left[2^n\sqrt{\pi}n!\right] \end{aligned}$$

The intensity distribution then can be written as

$$I_{mn} = \frac{2P}{\pi w^2}\frac{1}{2^{m+n}m!n!}H_m^2(\sqrt{2}x/w)H_n^2(\sqrt{2}y/w)e^{-2(x^2+y^2)/w^2} \tag{2.69}$$

It is easy to check that for $m = 0 = n$ we recover the intensity distribution of the fundamental gaussian mode.

Intensity distributions of some low order Hermite gaussian beams are

$$\begin{aligned}
I_{00}(x,y,z) &= \frac{2P}{\pi w^2}e^{-2(x^2+y^2)/w^2} \\
I_{10}(x,y,z) &= \frac{2P}{\pi w^2}\frac{1}{2}\frac{8x^2}{w^2}e^{-2(x^2+y^2)/w^2} \\
I_{01}(x,y,z) &= \frac{2P}{\pi w^2}\frac{1}{2}\frac{8y^2}{w^2}e^{-2(x^2+y^2)/w^2} \\
I_{11}(x,y,z) &= \frac{2P}{\pi w^2}\frac{1}{4}\frac{8x^2}{w^2}\frac{8y^2}{w^2}e^{-2(x^2+y^2)/w^2} \\
I_{20}(x,y,z) &= \frac{2P}{\pi w^2}\frac{1}{8}H_2^2\left(\frac{\sqrt{2}x}{w}\right)e^{-2(x^2+y^2)/w^2}
\end{aligned} \tag{2.70}$$

In terms of scaled variables $X = \sqrt{2}x/w$, $Y = \sqrt{2}y/w$, we obtain slightly more compact expression

$$I_{mn}(X,Y) = \frac{P}{\pi 2^{m+n}m!n!}H_m^2(X)H_n^2(Y)e^{-(X^2+Y^2)}\,. \tag{2.71}$$

for the intensity. Expressed in terms of X and Y, the intensity distribution is form invariant. Note that the power normalization condition is now

$$\int I_{mn}(X,Y)dXdY = P\,.$$

2.2.3 Astigmatic Hermite Gaussian Beams

It is possible for gaussian beams to have different spot sizes in the two transverse dimensions. The fundamental beam will then have an elliptical cross-section. For this reason such a beam is called an elliptical beam. They are good models for the light emitted by semiconductor lasers. Elliptical beams will converge or diverge at different rates along the principal axes of the ellipse. Such beams will therefore be astigmatic. If a elliptic beam is passed through a lens, beam waists after the lens do not, in general, lie in the same plane. Fundamental elliptical beam has the form

$$
\begin{aligned}
\mathcal{E}_0(\vec{r}) &= \mathcal{A}\sqrt{\frac{w_{0x}w_{0y}}{w_x(z)w_y(z)}}\, e^{ik\left(\frac{x^2}{2q_x(z)}+\frac{y^2}{2q_y(z)}\right)-i\psi(z)} \\
\frac{1}{q_x(z)} &= \frac{1}{R_x(z)} + i\frac{2}{kw_x^2(z)} \\
\frac{1}{q_y(z)} &= \frac{1}{R_y(z)} + i\frac{2}{kw_y^2(z)} \\
w_x^2 &= w_{ox}^2\left[1+\left(\frac{2(z-z_x)}{kw_{0x}^2}\right)^2\right] \\
w_y^2 &= w_{oy}^2\left[1+\left(\frac{2(z-z_y)}{kw_{0y}^2}\right)^2\right] \\
R_x(z) &= (z-Z_x)\left[1+\left(\frac{kw_{0x}^2}{2(z-Z_x)}\right)^2\right] \\
R_y(z) &= (z-Z_y)\left[1+\left(\frac{kw_{0y}^2}{2(z-Z_y)}\right)^2\right] \\
\psi(z) &= \frac{1}{2}\tan^{-1}\left[\frac{2(z-Z_x)}{kw_{0x}^2}\right] + \frac{1}{2}\tan^{-1}\left[\frac{2(z-Z_y)}{kw_{0y}^2}\right]
\end{aligned}
\qquad (2.72)
$$

Beam waist occurs at $z = Z_x$ in the $x-z$ plane and at $z = Z_y$ for the $y-z$ plane. These two planes in general do not coincide. The beam in genral has elliptical profile. Semiconductor lasers emit this type of beams. Such beam can be converted into symmetric Hermite-gaussian beams by using prisms or cylindrical lenses. For $w_{0x} = w_{0y} = w_0$ (which also requires $Z_x = Z_y = Z$, we recover the fundamental circularly symmetric gaussian beam with waist at Z.

2.2.4 Laguerre Gaussian Beams

Paraxial wave equation (2.9) admits beam solutions that reflect other symmetries. For example, in the presence of cylindrical symmetry about the z-axis, Eq. (2.9) admits Laguerre-Gaussian beam solutions

$$
[\mathcal{E}_0]_{p\ell}(\vec{r}) = \mathcal{A}\sqrt{\frac{2(p!)}{\pi(p+\ell)!}}\,\frac{w_0}{w}\left(\frac{\sqrt{2}\rho}{w}\right)^{\ell} e^{i\ell\varphi}\, L_p^{|\ell|}\left(\frac{2\rho^2}{w^2}\right) e^{-i(2p+\ell+1)\psi+i\frac{k\rho^2}{2q}}. \qquad (2.73)
$$

where $L_p^{|\ell}(x)$ is the associated Laguerre polynomial and $w(z)$ and $R(z)$ are independnet of the mode indices. Some low order associated Laguerre polynomials and recursion relations for computing the higher order polynomials are $(m > 0)$

$$\begin{aligned} L_0^\ell(x) &= 1 \\ L_1^\ell(x) &= -x + \ell + 1 \\ L_2^\ell(x) &= \frac{1}{2}\left[x^2 - 2(\ell+2)x + (\ell+1)(\ell+2)\right] \\ (p+1)L_{p+1}^{|\ell|}(x) &= (2p + \ell + 1 - x)L_p^{|\ell|}(x) - (p+\ell)L_{p-1}^{|\ell|}(x) \\ x\frac{dL_p^{|\ell|}(x)}{dx} &= pL_p^{|\ell|}(x) - (p+\ell)L_{p-1}^{|\ell|}(x) \end{aligned} \tag{2.74}$$

We see that the lowest order solution ($\ell = 0 = m$) coincides with the fundamental Gaussian beam solution for the HG family. For $\ell = 0$ and $m = 1$ we obtain the intensity

$$I_{01} = I_0 \frac{w_0^2}{w^2}\frac{2\rho^2}{w^2} e^{-2\rho^2/w^2} \tag{2.75}$$

This has a dark center and is sometimes called the donut mode. Note that the "donut" shaped intensity distribution sometimes seen in lasers is most often a mixture of HG_{01} and HG_{10} modes.

Let us calculate the total power of the beam.

$$P = \int_0^\infty \rho d\rho \int_0^{2\pi} d\varphi\, I(\rho, z) \tag{2.76}$$

$$= 2\pi \left(\frac{1}{2}\epsilon_0 cn|A|^2\right) \frac{2(p!)}{\pi(\ell+p)!}\left(\frac{w_0}{w}\right)^2 \int_0^\infty \rho d\rho \left(\frac{\sqrt{2}\rho}{w}\right)^{2\ell} \left[L_p^{|\ell|}(2\rho^2/w)\right]^2 e^{-2\rho^2/w^2} \tag{2.77}$$

$$= \left(\frac{1}{2}\epsilon_0 cn|A|^2\right)\left(\frac{4\pi(p!)}{\pi(\ell+p)!}\right)\left(\frac{w_0}{w}\right)^2 \int_0^\infty dX X^\ell \left[L_p^{|\ell|}(X)\right]^2 e^{-X}\frac{w^2}{4} \tag{2.78}$$

$$= \left(\frac{1}{2}\epsilon_0 cn|A|^2\right)\left(\frac{(p!)}{(\ell+p)!}\right) w_0^2 \frac{(\ell+p)!}{(p!)} = \left(\frac{1}{2}\epsilon_0 cn|A|^2\right) w_0^2 \tag{2.79}$$

Hence we can write the intensity of $\mathrm{LG}_{p\ell}$ beam as

$$I_{p\ell} = \frac{2P}{\pi w^2}\frac{p!}{(\ell+p)!}\left(\frac{\sqrt{2}\rho}{w}\right)^{2\ell} \left[L_p^{|\ell|}(2\rho^2/w)\right]^2 e^{-2\rho^2/w^2} \tag{2.80}$$

There are also the so-called Bessel beam or nondiffracting beam solutions [3]. In practice, symmetries other than the rectangular symmetry are difficult to realize. The presence of Brewster surfaces and other nonsymmetric optical elements in laser resonators naturally leads to rectangular symmetry. For this reason we consider only the Hermite-Gaussian solutions.