1

Characterization of Laser Beams: The M^2 Model

Thomas F. Johnston, Jr.
Optical Physics Solutions
Grass Valley, California, USA

Michael W. Sasnett
Optical System Engineering
Los Altos, California, USA

CONTENTS

1.1 INTRODUCTION

The M^2 model is currently the preferred way of quantitatively describing a laser beam, including its propagation through free space and lenses; specifically, as ratios of its parameters with respect to the simplest theoretical gaussian laser beam. The present chapter describes the model and measuring techniques for reliably determining—in each of the two orthogonal propagation planes—the key spatial parameters of a laser beam; namely, the beam waist diameter $2W_0$, the Rayleigh range z_R, the beam divergence Θ, and waist location z_0.

1.2 HISTORICAL DEVELOPMENT OF LASER-BEAM CHARACTERIZATION

In 1966, six years after the first laser was demonstrated, a classic review paper[1] by Kogelnik and Li of Bell Telephone Laboratories was published, which served as the standard reference on the description of laser beams for many years. Here the $1/e^2$ diameter definition[1,2] for the width of the fundamental-mode gaussian beam was used. The more complex transverse irradiance patterns, or transverse modes, of laser beams were identified with sets of eigenfunction solutions to the wave equation, including diffraction, describing the electric fields of the beam modes. These solutions came in two forms: those with rectangular symmetry were described mathematically by Hermite–Gaussian functions, those with cylindrical symmetry by Laguerre–Gaussian functions. So with the appropriate basis set, any beam could be decomposed into a weighted sum of the electric fields of these modes, at least in principle. Mathematically, for this expansion to be unique the phases of the electric fields must be known. This is difficult at optical frequencies. Irradiance measurements alone, where the phase information is lost in squaring the E-fields, does not allow determination of the expansion coefficients. This "in principle but not in practice" description of light beams was all that was available and seemed to be all that was needed for several succeeding years.

Workers often measured beam diameters by scanning an aperture across the beam to detect the transmitted power profile. Apertures used were pinholes, slits, or knife-edges, and the beam diameters were (and still are) defined based on the measurement effect that would be produced on a fundamental-mode beam. Commercial laser beams were specified as being pure fundamental mode, the lowest order or zero–zero transverse electromagnetic wave eigenfunction, "TEM_{00}."

In 1971, Marshall[3] published a short note introducing the M^2 factor, indicating M ($=\sqrt{M^2}$) as the multiplying factor by which the diameter of a beam is larger than that of the fundamental mode of the same laser resonator. Marshall's interest lay with the effects produced by industrial lasers and since they depend on focused spot size, he pointed out that they depend on M^2. No discussion was given of how to measure M^2 and the concept languished thereafter for several years.

From the late 1970s and into the 1980s, Bastiaans,[4] Siegman,[5,6] and others developed theories of bundles of light rays at narrow angles to an axis based on the Fourier transform relationship between the irradiance and the spatial frequency (or ray-angle) distributions to account for the propagation of the bundle. Such a bundle of rays is a beam. The beam diameter was defined as the standard deviation of the irradiance distribution (now called the second-moment diameter, when multiplied by four), and the square of this diameter was shown to increase as the square of the propagation distance—an expansion law for the diameter of hyperbolic form. These theories could be tested by measuring just the beam's irradiance profile along the propagation path.

In about 1987, one of us designed a telescope to locate a beam waist for an industrial CO_2 laser at a particular place in the external optical system. The design was based on measurements showing where the input beam waist was located and on blind faith that the laser datasheet claim for a "TEM_{00}" beam was correct. This telescope provided nothing like the expected result. Out of despair and disorientation came the energy to make more beam measurements and from these measurements came the realization that the factor that limited the maximum distance between the telescope and the beam waist it produced was exactly the same factor by which actual focus-spot diameter at the work surface exceeded

the calculated TEM_{00} spot diameter. That factor was M^2 and when used in modified Kogelnik and Li equations, design of optical systems for multimode beams became possible.[7] This ignited some interest in knowing more about laser beams than had previously been considered sufficient. Laser datasheets that claimed "TEM_{00}" were no longer adequate.

In the 1980s, commercial profilers[8] reporting a beam's $1/e^2$ diameter became ubiquitous. By the end of the 1980s, experience with commercial profilers and these theories converged with the development[6] of the theoretical M^2 model and a commercial instrument[9] to measure the beam quality based on it, which first became available in 1990. The time to determine a beam's M^2 value dropped from half a day to half a minute.

With high accuracy M^2 measurements more readily available in the early 1990s, the reporting of a beam's M^2 value became commonplace, and commercial lasers with good beams were now specified[10] as having $M^2 < 1.1$. The International Organization for Standards began committee meetings to define standards for the spatial characterization of laser beams, ultimately deciding on the beam quality M^2 value based on the second-moment diameter as the standard.[11] This diameter definition has the best theoretical support, in the form of the Fourier transform theories of the 1980s, but suffers from being sensitive to noise on the profile signal, which often makes the measured diameters unreliable.[12,13] That led to the development in 1993 of rules[14] to convert diameters measured with the more forgiving methods into second-moment diameters for a large class of beams.

The M^2 model as commercially implemented does not cover beams that twist as they propagate in space, those with general astigmatism.[15,16] The earlier Fourier transform theories and their more recent extensions do, however, and allow for ten constants[17] needed to fully characterize a beam (adding to the six used in the M^2 model). Recently, in 2001, the first natural beam[18] (as opposed to a test beam artificially constructed) was measured by Nemes et al. that required all ten constants for its complete description.

Several recommendations can be made for characterizing a beam. Model the beam only to the level of complexity appropriate to your needs: three constants suffice if the beam spot is round at all propagation distances; six constants cover beams with simple astigmatism, divergence asymmetry, or waist asymmetry; ten constants are needed for beams with elliptical spots whose orientation twist in space (general astigmatism). Measure your beams with a reliable method, and when required, convert those values at the end into ISO standard units. Lastly, stay appraised of developments in instrumentation that may meet your need with more convenience, speed, and accuracy.

1.3 ORGANIZATION OF THIS CHAPTER

Section 1.2 provides an historical introduction to the field, outlining how the field developed to its present state.

The technical discussion begins in Section 1.4 by explaining the M^2 model. This mathematical model built around the quantity M^2 (variously called the beam quality, times-diffraction-limit number, or the beam propagation factor) describes the real, multimode beams that lasers produce and how their properties change when propagating in free space.

This discussion is continued in Section 1.5 covering the transformation of a beam through a lens. Section 1.6 explains the different methods used to define and measure beam diameters, and how measurements made with one method can be converted into the values measured with one of the other methods. This includes the standard diameter definition

adopted by the International Organization for Standardization (ISO), the second-moment diameter, and the experimental difficulties encountered with this method.

The technical development continues in Section 1.7 where the logic and precautions needed in measuring the beam quality M^2 are presented. Thoroughly discussed is the "four-cuts" method (a cut is a measurement of a beam diameter), the simplest way to obtain an accurate M^2 value.

Section 1.8 discusses the common and possible types of beam asymmetry that may be encountered in three dimensions when the propagation constants for the two orthogonal (and usually independent) propagation planes are combined. The concept of the "equivalent cylindrical beam" is introduced to complete the technical development of the M^2 model. Propagation plots for beams with combinations of asymmetries are illustrated. A short discussion follows of "twisted beams," those with general astigmatism, which are not covered in the M^2 model, and require a beam matrix of ten moments of second order for their complete description. This second-order beam matrix theory is a part of the underpinnings of the ISO's choice of the noise-sensitive second-moment diameter as the "standard."

Section 1.9 applies the M^2 model to an analysis of a stereolithography laser-scanning system. Using results of earlier sections, by working backward from assumed perturbations or defects in the scanned beam at the work surface, the deviations in beam constants at the laser head that would produce them are found. An overview of the M^2 model, in Section 1.10, concludes the text.

A glossary follows explaining the technical terms used in the field, with the references ending the chapter.

1.4 THE M^2 MODEL FOR MIXED-MODE BEAMS

In laser beam-scanning applications, the main concern is having knowledge of the beam spot-size—the transverse dimensions of the beam—at any point along the beam path. The mixed-mode ($M^2 > 1$) propagation equations are derived as extensions of those for the fundamental mode, so pure modes and particularly the fundamental mode are the starting point.

1.4.1 Pure Transverse Modes: The Hermite–Gaussian and Laguerre–Gaussian Functions

Lasers emit beams in a variety of characteristic patterns or transverse modes that can occur as a pure single mode or more often, as a mixture of several superposed pure modes. The transverse irradiance distribution of a pure mode is the square of the electric field amplitude versus the transverse distance from the beam axis, which when measured is termed a transverse profile. This amplitude is described mathematically by Hermite–Gaussian functions if it has rectangular symmetry, or by a Laguerre–Gaussian function if it has circular symmetry.[1,2,5,19] These functions when plotted reproduce the familiar spot patterns—the appearance of a beam on an inserted card—first photographed in Reference 20 and shown in References 1 and 19. Computed spot patterns are displayed here in Figure 1.1. The computations were done in Mathematica for the first six cylindrically symmetric modes, in order of increasing diffraction loss for a circular limiting aperture. These modes are the solutions to the wave equation for a bundle of rays propagating at small angles (paraxial rays) to the z-axis, under the influence of diffraction and are of the general forms[1,2,7]

$$U_{mn}(x,y,z) = H_m(x/w)H_n(y/w)u(x,y,z) \tag{1.1a}$$

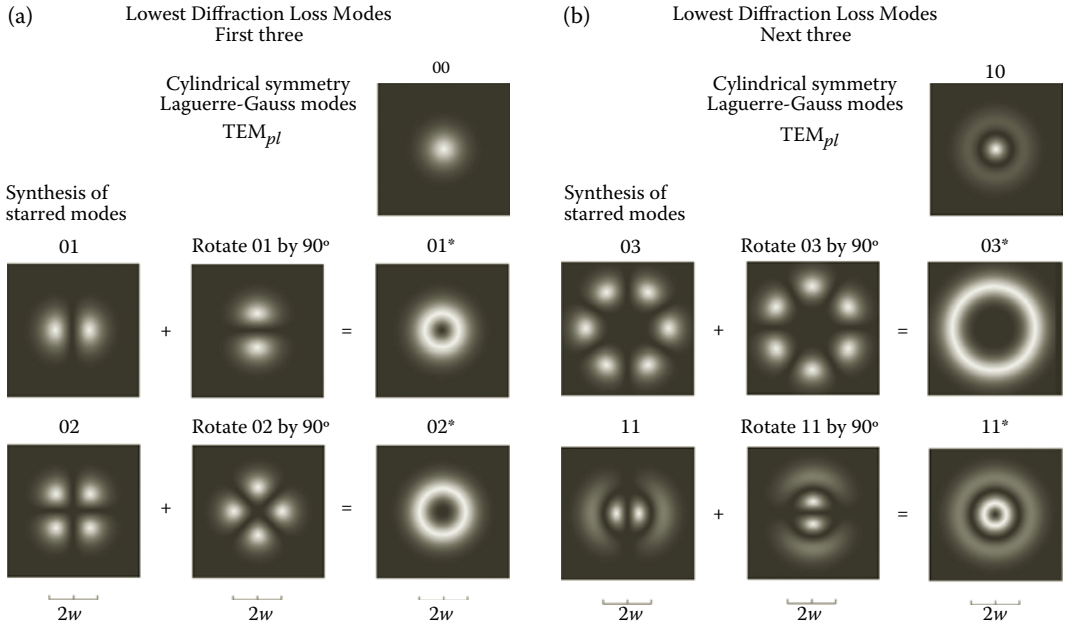


FIGURE 1.1
Computed spot patterns for cylindrically symmetric modes in order of increasing diffraction loss for a circular limiting aperture. The subscript numbers *pl* above each image indicate the mode order. Starred modes are constructed as shown, as the sum of a pattern with a copy of itself rotated by 90°: (a) First three modes. (b) Next three modes.

or

$$U_{pl}(r, j, z) = L_{pl}\left(r / w, j\right)u(r, z). \tag{1.1b}$$

In Equation 1.1a, $H_m(x/w)H_n(y/w)$ represents a pair of Hermite polynomials, one a function of x/w, the other of y/w, where x, y are orthogonal transverse coordinates and w is the radial scale parameter. In Equation 1.1b, $L_{pl}(r/w, \varphi)$ represents a generalized Laguerre polynomial, a function of the r, φ transverse radial and angular coordinates. These polynomials have no dependence on the propagation distance z other than through the dependence $w(z)$ in x/w, y/w, or r/w. The $w(z)$ dependence describes the beam convergence or divergence. The other function u is the gaussian

$$u = \left(\frac{2}{p}\right)^{1/2} \exp\left[\frac{-(x^2 + y^2)}{w^2}\right] = \left(\frac{2}{p}\right)^{1/2} \exp\left[\frac{-r^2}{w^2}\right]. \tag{1.2}$$

Because the radial gaussian function splits into a product of two gaussians, one a function of x, the other of y, the Hermite–Gaussian function splits into the product of two functions, one in x/w only and the other in y/w only, each of which is independently a solution to the wave equation. This has the consequence that beams can have independent propagation parameters in the two orthogonal planes (x, z) and (y, z).

These functions of the transverse space coordinates consist of a damping gaussian factor, limiting the beam diameter, times a modulating polynomial that pushes light energy out radially as polynomial orders increase. The order numbers m, n of the Hermite polynomials, or p, l of the Laguerre polynomial of the pure mode also determine the number of nodes in the spot pattern, for which the modes are named. They are designated as

transverse electromagnetic modes, or $TEM_{m,n}$ for a mode with m nodes in the horizontal direction and n nodes in the vertical direction, or $TEM_{p,l}$ for a mode with p nodes in a radial direction—not counting the null at the center if there is one—and l nodes in going angularly around half of a circumference. Figure 1.2a through f, show the theoretical beam irradiance profiles for the six pure modes from Figure 1.1. Because these are the six lowest loss modes,[21,22] they are commonly found in real laser beams. The modes as shown all originate in the same resonator—they all have the same radial scale parameter $w(z)$. The addition of an asterisk to the mode designation—a "starred mode"—signifies a composite of two degenerate (same frequency) Hermite–Gaussian modes or as here, Laguerre–Gaussian modes in space and phase quadrature to form a mode of radial symmetry. This is explained in Reference 20, discussed in Reference 5, p. 689, and shown in Figure 1.1 for a mode pattern with an azimuthal variation ($l \neq 0$) as the addition of the mode with a copy of itself after a 90° rotation, to produce a smooth ring-shaped pattern.

The simplest mode is the TEM_{00} mode, also called the lowest order mode or fundamental mode of Figures 1.1 and 1.2a, and consists of a single spot with a gaussian profile (here L_{pl}

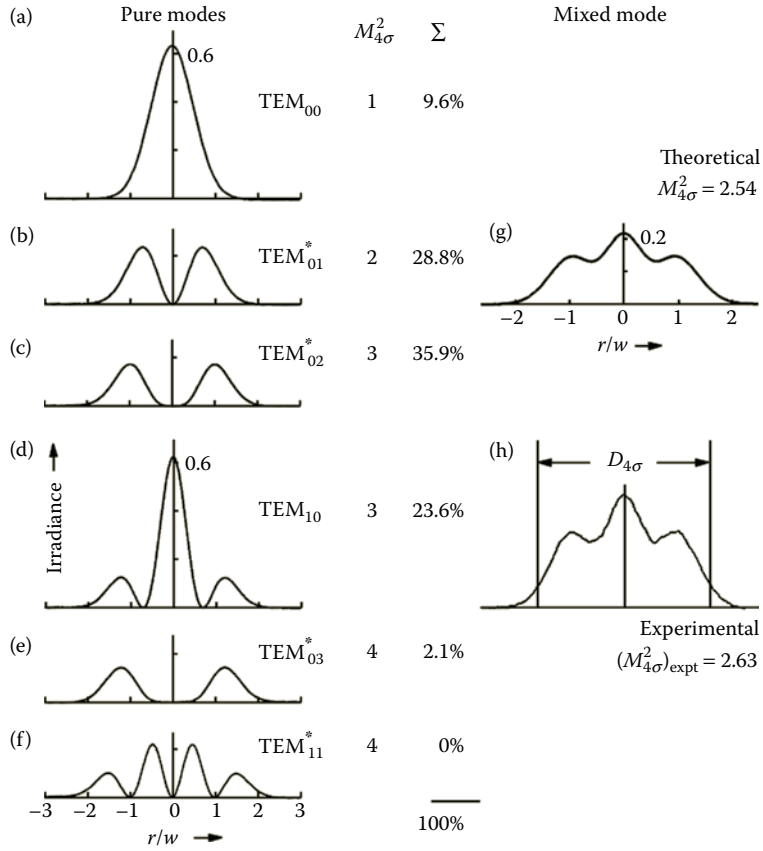


FIGURE 1.2
Synthesis of a mixed-mode as the weighted sum of pure modes. The theoretical pinhole profiles (a) to (f) for the six pure radial modes from Figure 1.1, shown in the first column, are summed with weighting fraction Σ of the third column to produce the mixed-mode profile (g). The beam qualities $M^2_{4\sigma}$ for each mode, in the second column, are similarly summed with weight Σ to produce the mixed-mode beam quality also shown in (g). The matching experimental pinhole profile is shown in (h).

is unity). The next higher-order mode has a single node (Figures 1.1 and 1.2b) and is appropriately called the "donut" mode, symbol TEM^*_{01}. The next two "starred" mode spots look like a donut with larger holes, the spot pattern of the TEM_{10} mode looks like a target with a bright center, and the TEM^*_{11} mode spot looks like a target with a dark center (Figures 1.1 and 1.2). All higher-order modes have a larger beam diameter than the fundamental mode. The six pure modes of Figure 1.2 are shown with the vertical scale normalized such that when integrated over the transverse plane, each contains unit power.

The physical reason that Hermite–Gaussian and Laguerre–Gaussian functions describe the transverse modes of laser beams is straightforward. Laser beams are generated in resonators by the constructive interference of waves multiply reflected back and forth along the beam axis. For this interference to be a maximum, permitting a large stored energy to saturate the available gain, the returned wave after a round trip of the resonator should match the transverse profile of the initial wave. The functions that do this are the eigenfunctions of the Fresnel–Kirchhoff integral equation used to calculate the propagation of a paraxial rays with diffraction included.[5,19] In other words, these are precisely the beam irradiance profiles that in propagating and diffracting maintain a self-similar profile, allowing after a round trip, maximum constructive interference and gain dominance.

1.4.2 Mixed Modes: The Incoherent Superposition of Pure Modes

While a laser may operate in a close approximation to a pure higher-order mode, for example, by a scratch or dust mote on a mirror forcing a node and suppressing a lower-order mode with an irradiance maximum at that location, actual lasers tend to operate with a mixture of several high-order modes oscillating simultaneously. The one major exception is lasing in the pure fundamental mode in a resonator with a circular limiting aperture, where the aperture diameter is critically adjusted to exclude the next higher-order (donut) mode. Each pure transverse mode has a unique frequency different from that for adjacent modes by tens or hundreds of MHz. This is usually beyond the response bandwidth of profile measuring instruments so any mode interference effects are invisible in such measurements.

Figure 1.2g shows a higher-order mode synthesized by mixing the five lowest order modes of Figure 1.2a through e in a sum with the weightings shown in the column labeled Σ. These weights—also called mode fractions—were chosen by a fitting program to match the result to the experimental pinhole profile (see Section 1.6.4.2) of Figure 1.2h. In the experiment[14] the number of transverse modes oscillating and their orders were known (by detecting the radio-frequency transverse mode beat notes in a fast photodiode). This information was used in the fitting procedure. The laser was a typical 1-m-long argon ion laser operating at a wavelength of 514 nm, except that a larger than normal intracavity limiting aperture diameter was used to produce this mode mixture.

Because the polynomials of Equation 1.1 have no explicit dependence on z, the profiles and widths of the modes in a mixture remain in the same ratio to each other and specifically to the fundamental mode as the beam propagates. This means that however the diameter $2W$ of a mixed-mode beam is defined (several alternatives are discussed in Section 1.6), if this diameter is M times larger than the fundamental-mode diameter at one propagation distance, it will remain so at any distance:

$$W(z) = Mw(z). \tag{1.3}$$

This equation introduces the convention that upper case letters are used for the attributes of high-order and mixed modes and lower case letters used for the underlying fundamental mode.

1.4.3 Properties of the Fundamental Mode Related to the Beam Diameter

The attributes of the simplest beam, a fundamental mode with a round spot (a cylindrically symmetric or stigmatic beam) are reviewed in Figures 1.3 and 1.4. The beam profile varies as the transverse irradiance distribution and is given by the function of gaussian form[1,2] (Figure 1.3a):

$$I\left(\frac{r}{w}\right) = I_0 \exp\left[-2\left(\frac{r}{w}\right)^2\right]. \tag{1.4}$$

The symbol I denotes a detector signal proportional to irradiance (and by using I instead of E, the recommended symbol for irradiance, avoids confusion with the electric field of the beam). The peak irradiance is I_0, and the radial scale parameter w introduced in Equation 1.1 can now be identified as the distance transverse to the beam axis at which the irradiance value falls to $1/e^2$ (13.5%) of the peak irradiance. This $1/e^2$ diameter definition, introduced[1,2] in the early 1960s, has been universally used since with one exception. (The one exception is in the field of biology where the fundamental-mode diameter is defined as the radial distance to drop to $1/e$ (36.8%) of the central peak value, making beams in biological references a diameter $2w' = \sqrt{2}w$ instead of $2w$.) Many different beam diameter definitions have been used subsequently for higher-order modes (these are discussed in Section 1.6) but they all share one common property: when applied to the fundamental-mode, they reduce to the traditional $1/e^2$ diameter.

Tables of the gaussian function are usually listed under the heading of the normal distribution, normal curve of error, or Gauss distribution and are of the form (see p. 763 of Reference 23)

$$I(x) = \left[\frac{1}{s(2p)^{1/2}}\right]\exp\left(\frac{-x^2}{2s^2}\right) \tag{1.5}$$

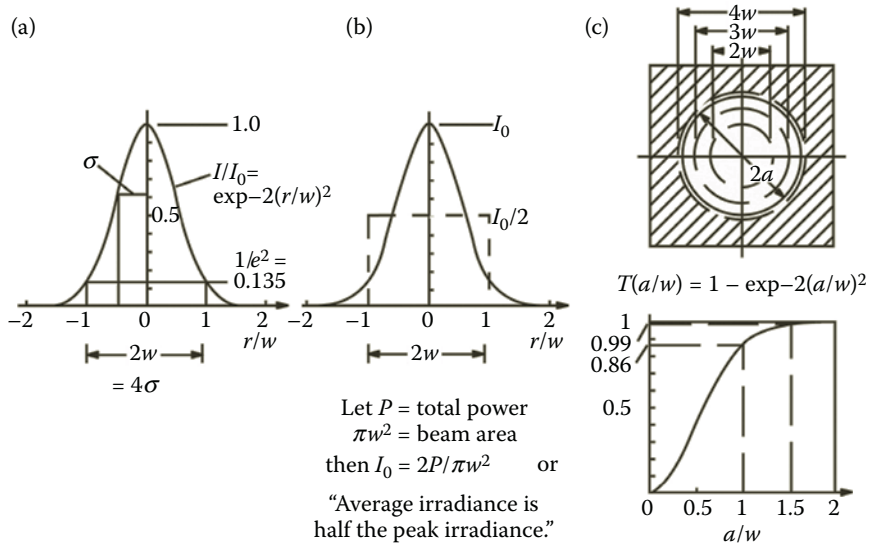


FIGURE 1.3
Properties of the fundamental mode related to the beam diameter, explained in the text; (a) definition of the $1/e^2$ diameter as the distance between the 13.5% levels on the pinhole profile; (b) relation between the peak irradiance and average irradiance; (c) transmission fraction through a circular aperture.

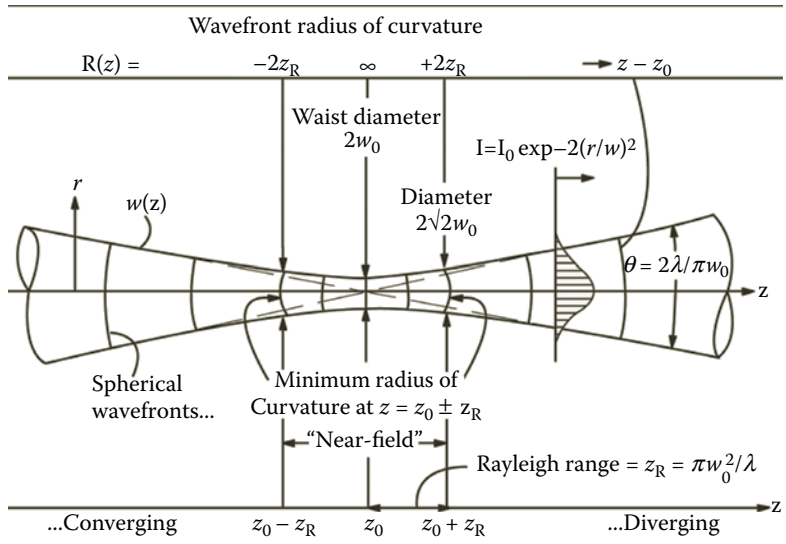


FIGURE 1.4
Propagation properties of the pure gaussian, fundamental-mode beam. The wavefront curvatures are exaggerated to show their variation with propagation distance.

where σ is the standard deviation of the gaussian distribution. Comparing Equation 1.4 and Equation 1.5 shows that the $1/e^2$ diameter is related to the standard deviation σ of the irradiance profile, as defined in Equation 1.5, as

$$2w = 4\sigma. \tag{1.6}$$

For a beam of total power P, the value of the peak irradiance I_0 is found[5] by integrating Equation 1.4 over the transverse plane (yielding I_0 times an area of $\pi w^2/2$) and equating this to P. The result

$$I_0 = \frac{2P}{\pi w^2} \tag{1.7}$$

is easily remembered by noting that "the average irradiance is half the peak irradiance." This is a handy, often-used simplification allowing the actual beam profile to be replaced by a round flat-topped profile of diameter $2w$ for back-of-the-envelope conceptualizations (see Figure 1.3b).

If the gaussian beam is centered on a circular aperture of diameter $2a$ the transmitted fraction $T(a/w)$ of the total beam power is given by a similar integration[5] over the cross-sectional area as (see Figure 1.3c):

$$T\left(\frac{a}{w}\right) = 1 - \exp\left[-2\left(\frac{a}{w}\right)^2\right]. \tag{1.8}$$

This gives a transmission fraction of 86.5% for an aperture of diameter $2w$, and 98.9% for one of diameter $3w$. In practice, a minimum diameter for an optic or other aperture to pass the beam and leave it unaffected is $4.6w$ to $5w$ to reduce the sharp edge diffraction ripples overlaid on the beam profile to an amplitude of <1%.[5] It is interesting to note that for a low power, visible, fundamental-mode beam, the spot appears to be a diameter of about $4w$ to the human eye viewing the spot on a card.

The transmission of a fundamental-mode beam past a vertical knife-edge is also readily computed. The knife-edge transmission function is $T(x/w) = 0$ for $x' \le x$, $T = 1$ for $x' > x$, where x is the horizontal distance of the knife-edge from the beam axis and x' is the horizontal integration variable. In Equation 1.4, the substitution $r^2 = x^2 + y^2$ is made, the integration over y yields multiplication by a constant, and the final integration over x' is expressed in terms of the error function as:

$$T\left(\frac{x}{w}\right) = \left(\frac{1}{2}\right)\left[1 \pm \operatorname{erf}\left(\frac{\sqrt{2}\,x}{w}\right)\right], \quad + \text{if } x < 0, \quad - \text{if } x > 0. \tag{1.9}$$

The error function of probability theory in Equation 1.9 is defined (see p. 745 of Reference 23) as

$$\operatorname{erf}(t) = \left(\frac{2}{\sqrt{p}}\int_0^t \exp(-u^2)\mathrm{du}\right) \tag{1.10}$$

and is tabulated in many mathematical tables. The $1/e^2$ diameter of a fundamental-mode beam is measured with a translating knife-edge by noting the difference in translation distances of the edge $(x_1 - x_2)$ that yield transmissions of 84.1% and 15.9%. By Equation 1.9 this separation equals w, and the beam diameter is twice this difference.*

1.4.4 Propagation Properties of the Fundamental-Mode Beam

The general properties expected for the propagation of a gaussian beam can be outlined from simple physical principles. As predicted by solving the wave equation with diffraction, a bundle of focused paraxial rays converges to a *finite* minimum diameter $2w_0$, called the waist diameter. The full angular spread θ of the converging and, on the other side, diverging beam is proportional to the beam's wavelength λ divided by the minimum diameter,[10] $\theta \propto \lambda/2w_0$. A scale length z_R for spread of the beam, is the propagation distance for the beam diameter to grow an amount comparable to the waist diameter, or $z_R\theta \sim w_0$, giving $z_R \propto w_0^2/\lambda$. Because the rays of the bundle propagate perpendicularly to the wavefronts (surfaces of constant phase), at the minimum's location the rays are parallel by symmetry and the wavefront there is planar. At large distances $z - z_0$ from the waist diameter location at z_0—the propagation axis is z—the wavefronts become Huygen's wavelets diverging from z_0 with wavefront radii of curvature $R(z)$, and eventually become plane waves. Since the wavefronts are plane at the minimum diameter at the waist and at large distances on either side, but converge and diverge through the waist, there must be points of maximum wavefront curvature (minimum radius of curvature) to either side of z_0.

The actual beam propagation equations describing the change in beam radius $w(z)$ and radius of curvature $R(z)$ with z, are derived[1,2,5] as solutions to the wave equation in the complex plane and show all of these features. They are (see Figure 1.4):

$$w(z) = w_0\sqrt{1 + \frac{(z - z_0)^2}{z_R^2}} \tag{1.11}$$

$$R(z) = (z - z_0)\left[1 + \frac{z_R^2}{(z - z_0)^2}\right] \tag{1.12}$$

$$z_R = \frac{pw_0^2}{l} \tag{1.13}$$

$$q = \frac{2l}{pw_0} = \frac{2w_0}{z_R} \tag{1.14}$$

* The knife-edge transmission function is illustrated later in Figure 1.8c and 1.8f of Section 1.6.

and

$$\psi(z) = -\tan^{-1}\left(\frac{z}{z_R}\right). \tag{1.15}$$

In these equations, the minimum beam diameter $2w_0$ (the waist diameter) is located at z_0 along the propagation axis z. A plot of $w(z)$ versus z, beam radius versus propagation distance [Equation 1.11] is termed the axial profile or propagation plot and is a hyperbola. The scale length for beam expansion, z_R, is termed the Rayleigh range [Equation 1.13] and has the expected dependence on λ and w_0. The radius of curvature $R(z)$ of the beam wavefront, as given by Equation 1.12, has the expected behavior. At large distances from the waist—the region termed the "far-field"—and where $|z - z_0| >> z_R$ the radius of curvature first becomes $R \to (z - z_0)$ and then becomes plane when $|R| \to \infty$ as $|z - z_0| \to \infty$, and also is plane at $(z - z_0) = 0$. By differentiating Equation 1.12 and equating the result to zero the points of minimum absolute value of the radius of curvature are found to occur at $z - z_0 = \pm z_R$ and have the values $R_{min} = \pm 2z_R$. The full divergence angle θ develops in the far-field, the beam envelope is asymptotic to two straight lines crossing the axis at the waist location (Figure 1.4). Finally, $\psi(z)$ is the phase shift[5,24] of the laser beam relative to that of an ideal plane wave. It is a consequence of the beam going through a focus (the waist), the gaussian beam version of the Gouy phase shift.[24]

By Equation 1.11, the diameter $2w(z)$ of the beam increases by the factor $\sqrt{2}$ (and for a round beam the cross-sectional area doubles) for a propagation distance $\pm z_R$ away from the waist (Figure 1.4). This condition is often used to define the Rayleigh range z_R,[5,25] but another significant condition is that at these two propagation distances the wavefront radius of curvature goes through its extreme values ($|R| = R_{min}$). The Rayleigh range can be defined as half the distance between these curvature extremes. The region within a Rayleigh range of the waist is defined as the "near-field" region. Within this region wavefronts flatten as the waist is approached and outside they flatten as they recede from the waist. A positive lens placed in a diverging beam and moved back towards the source waist will encounter ever-steeper wavefront curvatures so long as the lens remains out of the near-field. On the lens output side, the transformed waist moves away from the lens, moving qualitatively as a geometrical optics image would. When the lens enters the near-field region still approaching the source waist, ever-flatter wavefronts are encountered and then the transformed waist *also* approaches the lens. The laser system designer who misunderstands this unusual property of beams will have unpleasant surprises. Many laser systems have undergone emergency redesign when prototype testing revealed this counter-intuitive focusing behavior! In many ways, Rayleigh range is the single most important quantity in characterizing a beam (notice that this is a factor in all of Equations 1.11 through 1.15). It will be shown in the next section that measurement of a beam's Rayleigh range is the basis for measuring the beam quality M^2 of a mixed-mode beam.

As the lowest order solution to the wave equation, the fundamental-mode with a gaussian irradiance profile of a given waist diameter $2w_0$ is the beam of lowest divergence, at the limit set by diffraction,[10] of any paraxial bundle with that minimum diameter. Confining a bundle to a smaller diameter proportionally increases—by diffraction—the divergence angle of the bundle, and the product $2w_0\theta$ is an invariant for any mode. The smallest possible value, $4\lambda/\pi$, is achieved only by the fundamental mode. This is just the Uncertainty Principle for photons—laterally confining a photon in the bundle increases the spread of its transverse momentum and correspondingly the divergence angle of the bundle. This limit cannot be achieved by real-world lasers but sometimes it is closely approached. Helium–neon lasers, especially the low-cost versions with internal mirrors (no Brewster windows),

are wonderful sources of beams within 1% or 2% of this limit. Aside from the wavelength, which must be known to specify any beam, the ideal, round, (stigmatic) fundamental-mode beam is specified by only two constants: the waist diameter $2w_0$ and its location z_0 (or equivalents such as z_R and z_0). This will no longer be true when mixed modes are considered.

As noted at the beginning of this section the propagation constants for the (x, z) and (y, z) planes are independent and can be different. In each plane, the rays obey equations exactly of the same form[6] as Equations 1.11 through 1.15 with subscripts added indicating the x or y plane. For beams with pure (but different) gaussian profiles in each plane, two more constants are introduced for a total of four required to specify the beam. If $z_{0x} \neq z_{0y}$ (different waist locations in the two principal propagation planes) the beam exhibits simple astigmatism; if $2w_{0x} \neq 2w_{0y}$ (different waist diameters) the beam has asymmetric waists.*

1.4.5 Propagation Properties of the Mixed-Mode Beam: The Embedded Gaussian and the M^2 Model

In Section 1.4.2 a mixed mode was defined as the power-weighted superposition of several higher-order modes originating in the same resonator, each with the same underlying gaussian waist radius w_0 determining the radial scale length $w(z)$ in their mode functions [Equations 1.1 and 1.2]. This underlying fundamental mode, with w_0 fixed[5] by the radii of curvature and spacing of the resonator mirrors, is called the embedded gaussian for that resonator regardless of whether or not the mixed mode actually contains some fundamental-mode power. To treat the mixed-mode case, use is made[7] of the fact that its diameter is everywhere (for all z) proportional to the embedded gaussian diameter. From Equation 1.3 the substitution $w(z) = W(z)/M$ in Equations 1.11 through 1.15 yields the mixed-mode propagation equations:

$$W(z) = W_0\sqrt{1 + \frac{(z - z_0)^2}{z_R^2}} \tag{1.16a}$$

$$R(z) = (z - z_0)\left[1 + \frac{z_R^2}{(z - z_0)^2}\right] \tag{1.17a}$$

$$Z_R = \frac{\pi W_0^2}{M^2 \lambda} = z_R \tag{1.18}$$

and

$$\Theta = \frac{2M^2\lambda}{\pi W_0} = \frac{2W_0}{z_R} = Mq. \tag{1.19}$$

The mixed mode, a sum of transverse modes with different optical frequencies, no longer has a simple expression for the Gouy phase shift analogous to Equation 1.15. The convention followed here is that upper case quantities refer to the mixed mode and lower case quantities refer to the embedded gaussian.

Also useful are the inverse forms of Equation 1.16a and Equation 1.17a expressing the waist radius W_0 and waist location z_0 in terms of the beam radius $W(z)$ and wavefront curvature $R(z)$ at propagation distance z:

$$W_0 = \frac{W(z)}{\sqrt{1 + [\pi W(z)^2/M^2\lambda R(z)]^2}} \tag{1.16b}$$

* These beam asymmetries are illustrated later in Figure 1.15 of Section 1.8.

and

$$z_0 = \frac{R(z)}{1+[M^2 \lambda R(z)/\pi W(z)^2]^2} \tag{1.17b}$$

These forms are obtained from Reference 1, with the substitution $w = W/M$ in their Equations 24 and 25.

Many of the properties of the fundamental-mode beam carry over to the mixed-mode one (Figure 1.5). Since $W_0 = Mw_0$, substitution of this in the middle part of Equation 1.19 yields the last part, the mixed-mode divergence is M times that of the embedded gaussian. Similarly, the beam propagation profile $W(z)$ also has the form of a hyperbola (one M times larger) with asymptotes crossing at the waist location. The Rayleigh ranges are the same for both mixed and embedded gaussian modes as substituting $W_0 = Mw_0$ in the middle of Equation 1.18 shows, so the radii of curvature and the limits of the near-field region are the same for both. The mixed-mode beam diameter still expands by a factor of $\sqrt{2}$ in a propagation distance of z_R away from the waist location z_0, the starting diameter W_0 is just M times larger.

In considering propagation in the independent (x, z) and (y, z) planes, there are now two new constants needed to specify the beam, M_x^2 and M_y^2, for a total of six required constants. In making up the mixed mode, the Hermite–Gaussian functions summed in the two planes need not be the same or have the same distribution of weights, making $M_x^2 \neq M_y^2$ a possibility. In this case the beam is said to have divergence asymmetry since $\Theta \propto M^2$ by the first part of Equation 1.19.

It might be asked, why are these Equations 1.16 through 1.19 termed the "M^2 model" (and not the "M model")? There are two reasons. The first is that the embedded gaussian is buried in the mixed-mode profile, and cannot be measured independently, making it difficult to directly determine M. The mixed-mode diameter still grows by $\sqrt{2}$ in a

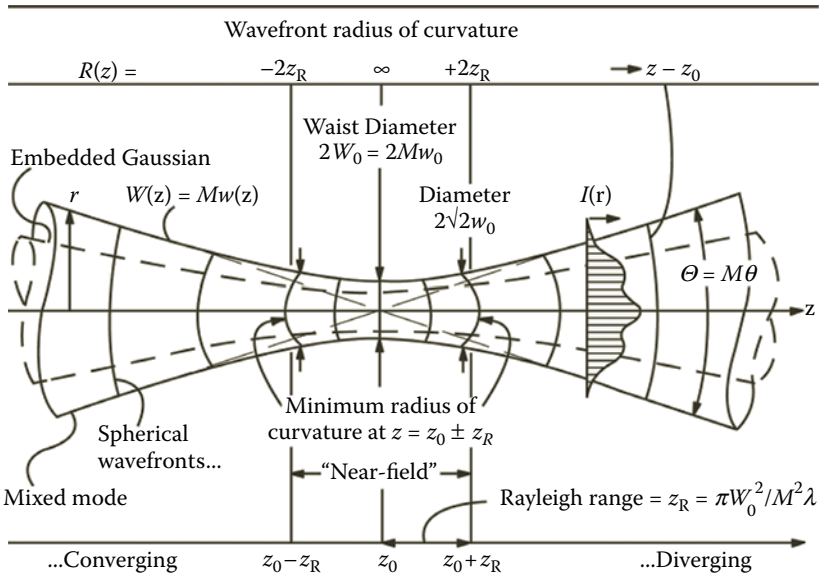


FIGURE 1.5
Propagation properties of the mixed-mode beam drawn for $M^2 = 2.63$. The embedded gaussian is the fundamental-mode beam originating in the same resonator. The wavefront curvatures are exaggerated to show their variation with propagation distance.

propagation distance z_R from the waist location, so z_R can be found from several diameter measurements fitted to a hyperbolic form. The waist diameter $2W_0$ can also be measured, thus giving directly, by Equation 1.18,

$$M^2 = \frac{\pi W_0^2}{\lambda z_R}. \tag{1.20}$$

This is how M^2 is in fact measured, the practical aspects of which will be discussed in Section 1.7. (As an aside, notice that Equation 1.20 shows that M^2 scales as the square of the beam diameter; this is used later in the discussion of conversions between different diameter definitions in Section 1.6.4.)

The second reason is the more important one: M^2 is an invariant of the beam, and is conserved[26] as the beam propagates through ordinary nonaberrating optical elements. Like the fundamental-mode beam whose waist diameter-divergence product was conserved, the same product for the mixed-mode beam is

$$(2W_0)\Theta = \frac{(2W_0)2M^2\lambda}{\pi W_0} = M^2\frac{4\lambda}{\pi}. \tag{1.21}$$

This is larger by the factor M^2 than the invariant product for a fundamental mode.

Equation 1.21 can be rearranged to read

$$M^2 = \frac{\Theta}{(2\lambda/\pi W_0)} = \frac{\Theta}{\theta_n}. \tag{1.22}$$

Here $\theta_n = 2\lambda/\pi W_0$ is recognized as the divergence of a fundamental-mode beam with a waist diameter $2W_0$, the same as the mixed-mode beam. This is called the normalizing gaussian; it has an M times larger scale constant $W_0 = Mw_0$ in its exponential term than the embedded gaussian and it would *not* be generated in the resonator of the mixed-mode beam. It does represent the diffraction-limited minimum divergence for a ray bundle constricted to the diameter $2W_0$. Thus by Equation 1.22 the invariant factor M^2 can be seen to be the "times-diffraction-limit" number referred to in the literature.[5] This also identifies M^2 as the inverse beam quality number, the highest quality beam being an idealized diffraction-limited one with $M^2 = 1$, while all real beams are at least slightly imperfect and have $M^2 > 1$.

The value of the M^2 model is twofold. Once the six constants of the beam are accurately determined (by fitting propagation plot data for each of the two independent propagation planes) they can be applied by the system designer to accurately predict the behavior of the beam throughout the optical system before it is built. The spot diameters, aperture transmissions, focus locations, depths of field, and so forth can all be found for the vast majority of existing commercial lasers. The second value is that there are commercial instruments available that efficiently measure and document a beam's constants in the M^2 model. This permits quality control inspection of the lasers at final test, or whenever there is a system problem and the laser is the suspected cause. Defective optics can introduce aberrations in the beam wavefronts. If inside the laser, they increase M^2 by forcing larger amounts of high-divergence, high-order modes in the mixed-mode sum. If outside the resonator, they also adversely affect M^2. Measurement of the beam quality during system assembly, after each optic is added to detect a downstream increase in M^2, can aid in quality control of the overall optical system.

Beams excluded from the model as described are those whose orthogonal axes rotate or twist about the propagation axis (called beams with general astigmatism[15,16,27]) such as might come from lasers with nonplanar ring or out-of-plane folded resonators. The

symmetry of the beam is determined by the symmetry of the resonator. Fortunately, few commercial lasers produce beams having these characteristics. An overview of the full range of symmetry possibilities for laser beams is discussed in Section 1.8.3.

The fact that M^2 is not unique, that is, that a given value of M^2 can be arrived at by a variety of different higher-order modes or mode weights in the mixed mode is sometimes stated to be a deficiency of the M^2 model. This is also its strength. It is a simple predictive model that does not require measurement and analysis to determine the mode content in a beam. In the evolution of beam models, the original discussion[1,2] pointed out that as eigenfunctions of the wave equation, the full (infinite) set of Hermite–Gaussian or Laguerre–Gaussian functions (Equation 1.1) describing the electric field of the beam modes form an orthonormal set. As such they could model an arbitrary paraxial light bundle with a weighted sum. This is true only if the phases of the E-fields are kept in the sum, and measuring the phase of an optical wave generally is a difficult matter. Summing the irradiances (the square of the E-fields) breaks the orthonormality condition and for years it was not obvious that a simple model relying only on irradiance measurements was possible. Then in the 1980s, methods based on Fourier transforms of irradiance and ray angular distributions of light bundles were introduced,[4,6] which showed that as far as predictions of beam diameters in an optical system were concerned, irradiance profile measurements would (usually) suffice. The M^2 model was born, and commercial instruments[10] for its application soon followed. Later we realized that modes "turn on" in a characteristic sequence as diffraction losses are reduced in the generating resonator. This makes a given M^2 correspond to a unique mode mix in many common cases after all (see Section 1.6.4).

1.5 TRANSFORMATION BY A LENS OF FUNDAMENTAL AND MIXED-MODE BEAMS

Knowledge of how a beam is transformed by a lens is not only useful in general, but in particular, a lens is used to gain an accessible region around the waist for the measurements of diameters that are analyzed to produce M^2 (see Section 1.7). This transformation is discussed next.

In geometrical optics a point source at a distance s_1 from a thin lens produces a spherical wave whose radius of curvature is R_1 at the lens (and whose curvature is $1/R_1$), where $R_1 = s_1$. In traversing the lens, this curvature is reduced by the power $1/f$ of the lens (f is the effective focal length of the lens) to produce an exiting spherical wave of curvature $1/R_2$ according to the thin lens formula:

$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}. \tag{1.23}$$

An image of the source point forms at the distance R_2 from the lens from convergence of this spherical wave. Note that the conventions used in Equation 1.23 are the same as in Equation 1.17, namely, the beam always travels from left to right, converging wavefronts with center of curvature to the right have negative radii, and diverging wavefronts with centers to the left have positive radii. [The usual convention in geometrical optics[28] is that converging wavefronts leaving the lens are assigned positive radii, which would put a minus sign on the term $1/R_2$ of Equation 1.23.]

The quantities used in the beam-lens transform are defined in Figure 1.6. Following Kogelnik[1] the beam parameters on the input side of the lens are designated with a subscript 1 (for "1-space") and on the output side with a subscript 2 (for "2-space"). The principal plane description[28] of a real (thick) lens is used, in which the thick lens is replaced by a thin one acting at the lens principal planes H1, H2. Rays between H1 and H2 are drawn parallel to the axis by convention, and waist locations z_{01} and z_{02} are measured from H1 and H2 respectively (with distances to the right as positive for z_{02} and distances to the left as positive for z_{01}).

A lens inserted in a beam makes the same change in wavefront curvature as it did in geometrical optics [Equation 1.23], but the wavefront R_2 converges to a waist of finite diameter $2W_{02}$ at a distance z_{02} given by Equation 1.17b. For each of the two independent propagation planes, there are three constants required to specify the transformed beam, and three constraints needed to determine them. The lens should be aberration-free (typically, used at $f/20$ or smaller aperture) and, if so, the beam quality is not changed in passing through it, giving the first condition $M_2^2 = M_1^2$. The second constraint is that the wavefront curvatures match, between the input curvature modified by the lens [Equation 1.23], and the transformed beam at the same location as specified by the transformed beam constants through Equation 1.17a. A beam actually has two points with the same magnitude and sign of the curvature, one inside the near-field region of that sign and one outside, which differ in beam diameters. The ambiguity as to which point is matched is removed by the third constraint, that the beam diameter is unchanged in traversing the (thin) lens.

These three constraints define three equations that next are solved for the transformed waist diameter and location. This is facilitated by Equations 1.16b and 1.17b for W_0 and z_0 as functions of $W(z)$ and $R(z)$. The solution[1,29–31] is written in terms of the transformation constant Γ (using the modern symbols from a commercial M^2 measuring instrument[9]) as follows:

$$\Gamma = \frac{f^2}{[(z_{01} - f)^2 + z_{R1}^2]} \tag{1.24}$$

$$M_1^2 = M_2^2 = M^2 \tag{1.25}$$

$$W_{02} = \sqrt{\Gamma} W_{01} \tag{1.26}$$

$$z_{R2} = \Gamma z_{R1} \tag{1.27}$$

$$z_{02} = f + \Gamma(z_{01} - f) \tag{1.28}$$

A set of these equations apply to each of the two principal propagation planes (x, z) and (y, z).

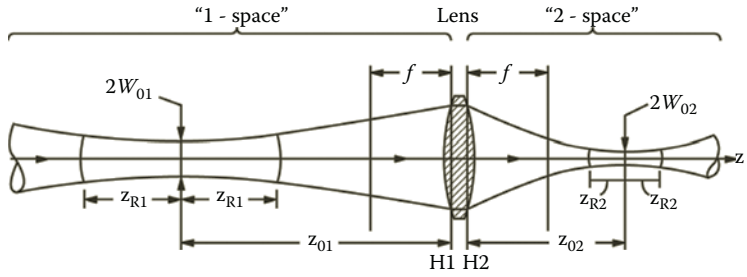


FIGURE 1.6
Definitions of quantities used in the beam-lens transform.

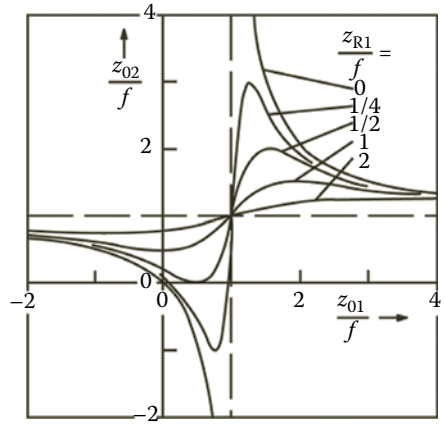


FIGURE 1.7
Parametric plots of the transformed waist location as a function of the input waist location for the beam-lens transform, with f as the lens focal length and the Rayleigh range z_{R1} of the input beam as parameters.

The transform equations [Equations 1.24 through 1.28] are not as simple as in geometrical optics because of the complexity of the way the beam wavefront curvatures change with propagation distance, Equation 1.17a. Like the image and object distances in geometrical optics, the transformed beam waist location depends on the input waist location—but also depends, as does the wavefront curvature, on the Rayleigh range of the input beam. The most peculiar behavior as the waist-to-lens distance varies is when the input focal plane of the lens moves within the near-field of the incident beam, $|z_{01} - f| < z_{R1}$. Then the slope of the z_{02} versus z_{01} curve turns from negative to positive (in geometrical optics the slope of the object to image distance curve is always negative). This sign change can be demonstrated by substituting Equation 1.24 into Equation 1.28 and differentiating the result with respect to z_{01}. As the lens continues to move closer to the input waist, the transformed waist location also moves closer to the lens, exactly opposite to what happens in geometrical optics. In the beam-lens transform, the input and transformed waists are *not* images of each other (in the geometrical optics sense). Despite the intransigence of beam waists, the object–image relationship of beam diameters at conjugate planes on each side of the lens does apply just as in geometrical optics. A good modern discussion of the beam-lens transform is presented in O'Shea's textbook[32] (where his parameter $\alpha^2 = \Gamma$ here).

A pictorial description of the beam-lens transform is given by a figure in Reference 30, redrawn here as Figure 1.7. Variables normalized to the lens focal length f are used to show how the transformed waist location z_{02}/f varies with the input waist location z_{01}/f. The input Rayleigh range z_{R1}/f (also normalized) is used as a parameter and several curves are plotted for different values. The anomalous slope regions of the plot are evident. The geometrical optics thin lens result, Equation 1.23, is recovered when the input Rayleigh range becomes negligible, $z_{R1}/f = 0$ (the condition for a point source), and the slopes of both wings of the curve are then always negative.

1.5.1 Application of the Beam-Lens Transform to the Measurement of Divergence

An initial application of the beam-lens transform equations is to show that the divergence of the input beam Θ_1 in 1-space of Figure 1.6 can be determined by measuring the

beam diameter $2W_f$ at precisely one focal length behind the lens exit plane H2 in 2-space from the equation:

$$\Theta_1 = \frac{2W_f}{f}. \tag{1.29}$$

This result is independent of where the lens is placed in the input beam. This follows by finding in 2-space the diameter $2W_f$ at $z_2 = f$ [from Equation 1.16a] and substituting Equations 1.19, 1.24, and 1.28:

$$2W_f = 2W_{02}\left[1 + \frac{(f - z_{02})^2}{z_{R2}^2}\right]^{1/2} = 2W_{02}\left(\frac{f}{z_{R2}}\right)\left(\frac{1}{\Gamma^{1/2}}\right)$$
$$= 2W_{01}\left(\frac{f}{z_{R2}}\right)\left(\frac{1}{\Gamma}\right) = 2W_{01}\left(\frac{f}{z_{R1}}\right) = \Theta_1 f$$

which is Equation 1.29. In Figure 3b of Reference 25 there is an illustration showing how the transform equations operate to keep the output beam diameter one focal length from the lens fixed at the value $\Theta_1 f$ despite variations in the input waist location, z_{01}. The measurement method implied by Equation 1.29 is the simplest way to get a good value for the beam divergence Θ_1. Care should be taken to pick a long enough focal length lens that the beam diameter is large enough for the precision of the diameter-measurement method in use.

1.5.2 Applications of the Beam-Lens Transform: The Limit of Tight Focusing

When the aperture of a short focal length lens is filled on the input side, the smallest possible diameter output waist is reached and this is called the limit of tight focusing. This limit is characterized by (1) the beam diameter at the lens being given by $2W_{lens} = \Theta_2 f$; (2) the output waist being near the focal plane $z_{02} = f$; and (3) there being a short depth of field at the focus, $z_{R2}/f \ll 1$. Applying Equation 1.29 in the reverse direction gives the 2-space divergence as the ratio of the beam diameter $2W_{1f}$ at f to the left of the lens, to the focal length, $\Theta_2 f = 2W_{1f}$. By condition (1) this means $2W_{1f} = 2W_{lens}$ or that there is little change in the input beam diameter over a propagation distance f. That makes the first condition characterizing the tight focusing case equivalent to $z_{R1}/f \gg 1$. Then from Equation 1.19,

$$2W_{lens} = \frac{2\lambda M^2 f}{\pi W_{02}}$$

or

$$2W_{02} = 2\lambda M^2\left(\frac{f}{\pi W_{lens}}\right) = 2\lambda M^2 (f/\#) \tag{1.30}$$

for the tight focusing limit. Here Siegman's definition[5] is used that a lens of diameter D_{lens} is filled for a fundamental-mode beam of diameter πW_{lens} (this degree of aperture filling produces <1% clipping of the beam). Thus $f/\pi W_{lens} = f/D_{lens} = (f/\#)$. The depth of field of the focus is $z_{R2} = \pi W_{02}^2/M^2\lambda = \pi M^2\lambda(f/\#)^2$. This generalizes a familiar result[5] for a fundamental-mode beam to the $M^2 \neq 1$ case.

Marshall's point[3] (from 1971) is made by Equation 1.30, that a higher-order mode beam focuses to a larger spot by a factor of M^2, with less depth of field, and therefore cuts and welds less well than a fundamental-mode beam.

1.5.3 The Inverse Transform Constant

The transform equations work equally well going from 2-space to 1-space, with one transformation constant the inverse of the other,

$$\Gamma_{21} = \frac{1}{\Gamma_{12}}. \tag{1.31}$$

This obviously is true by symmetry but the algebraic proof is left to the reader.

1.6 BEAM DIAMETER DEFINITIONS FOR FUNDAMENTAL AND MIXED-MODE BEAMS

It has been said that the problem of measuring the cross-sectional diameter of a laser beam is like trying to measure the diameter of a cotton ball with a pair of calipers. The difficulty is not in the precision of the measuring instrument, but in deciding what is an acceptable definition of the edges.

Unlike the fundamental-mode beam where the $1/e^2$ diameter definition is universally understood and applied, for mixed modes a number of different diameter definitions[7] have been employed. The different definitions have in common that they all reduce to the $1/e^2$ diameter when applied to an $M^2 = 1$ fundamental-mode beam, but when applied to a mixed mode with higher-order-mode content they in general give different numerical values. As M^2 always depends on a product of two measured diameters, its numerical value changes also as the square of that for diameters. It is all the same beam, but different methods provide results in different currencies; one has to specify what currency is in use and know the exchange rate.

Since the adoption[11] by the ISO committee on beam widths of the second-moment diameter as the standard definition for beam diameters, there has been increasing effort among laser users to put this into practice. This definition, discussed in Section 1.6.3.5, has the best analytical and theoretical support but is difficult experimentally to measure reproducibly because of sensitivity to small amounts of noise in the data. The older methods therefore persist and the best strategy[25] at present is to use the more forgiving methods for the multiple diameter measurements needed to determine M^2. Then at one propagation distance, do a careful diameter measurement by the second-moment definition to provide a conversion factor. This conversion factor can then be applied to obtain standardized diameters at any distance z in the beam. This strategy will likely evolve in the future if and when instrument makers respond to the ISO Committee's choice and devise algorithms and direct methods for ready and accurate computations of second-moment diameters.

1.6.1 Determining Beam Diameters from Irradiance Profiles

Beam diameters are determined from irradiance profiles, the record of the power transmitted through a mask as a function of the mask's translation coordinate transverse to

the beam. A sufficiently large linear power detector is inserted in the beam, with a uniformly sensitive area to capture the total power of the beam. Detection sensitivity should be adequate to measure ~1% of the total power, and response speed should allow faithful reproduction of the time-varying transmitted power. The mask is mounted on a translation stage, placed in front of the detector, and moved or scanned perpendicularly to the beam axis to record a profile. An instrument that performs these functions is called a beam profiler. In a useful version based on a charge-coupled-device (CCD) camera, the masking is done on electronic pixel data under software control.

The beam propagation direction defines the z-axis. The scan direction is usually along one of the principal diameters of the beam spot and commercial profilers are mounted to provide rotation about the beam axis to facilitate alignment of the scan in these directions. The principal diameters for an elliptical spot are the major and minor axes of the ellipse (or the rectangular axes for a Hermite–Gaussian mode). The principal propagation planes (x, z) and (y, z) are defined as those containing the principal spot diameters. The beam orientation is arbitrary and in general may require rotation of coordinates to tie it to the laboratory reference frame. It is assumed this rotation is known, and without loss of generality to give simple descriptive terminology in this discussion, here the z-axis is taken to be horizontal, the principal propagation planes as the horizontal and vertical planes in the laboratory, with the scan along the x-axis. If the mask requires centering in the beam (e.g., a pinhole) to find the principal diameter, it is mounted on a y-axis stage as well and x-scans at different y-heights taken to determine the widest one at the beam center. Alternatively, a mirror directs the beam onto the profiler and the spot is put at different heights to find the beam center by tipping the mirror about a horizontal rotation axis. If the beam is repetitively pulsed and detected with an energy meter, the stage is moved in increments between pulses. If a CCD camera is the detector, a scan line is the readout of sequential pixels and no external mask is required in front of the camera. A CCD camera generally requires a variable attenuator[33] inserted before the camera to set the peak irradiance level just below the saturation level of the camera for optimum resolution of the irradiance value on the ordinate axis of the profile.

The results of this process are irradiance profiles such as shown in Figure 1.8 for two pure modes, the fundamental mode in the first row and the donut mode in the second, where three scans are calculated for each, one for a pinhole (first column), a slit (second column), and a knife-edge (third column) as masks. The traditional definitions used to extract diameters from these profiles are the same for the pinhole and slit. This is to normalize the scan to the highest peak as 100%, then to come down on the scan to an ordinate level at $1/e^2$ (13.5%) and measure the diameter—or clip width—as the scan width between these crossing points (called clip levels or clip points and shown as dots in Figure 1.8). The symbols D_{pin} and D_{slit} are used for these two diameters. For the knife-edge diameter (symbol D_{ke}) the definition is to take the scan width between the 15.9% and 84.1% clip points and double it, as this rule produces the $1/e^2$ diameter when applied to the fundamental mode.

As shown in Figure 1.8 the diameter results for the donut mode (TEM^*_{01}) are all larger than the $2w$ diameter of the fundamental mode, as expected. However, the answers for the three different methods for the donut mode—and in general, for all higher-order modes—are all different! The ratio of the donut mode to fundamental-mode diameter is 1.51, 1.42, and 1.53 by the pinhole, slit, and knife-edge methods, respectively. The reason, obviously, is that traces of different shapes are produced by the different methods. The pinhole cuts the donut right across the hole and records a null at the center; the slit extends vertically across the whole spot and records a transmission dip in crossing the hole but never reaches zero due to the contribution of the light above and below the hole. Even higher transmission

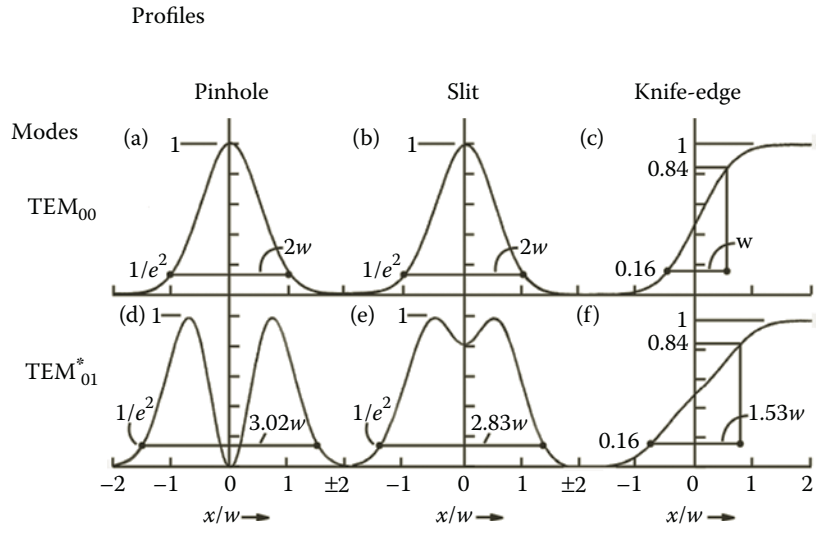


FIGURE 1.8
Theoretical beam profiles (irradiance vs. translation distance) from a scanning pinhole (a) and (d), slit (b) and (e), and knife-edge (c) and (f) cutting the fundamental and donut modes, illustrating that different methods give different diameters for higher-order mode beams. The knife-edge diameter is defined as *twice* the translation distance between the 15.9% and 84.1% cut points.

results with the knife-edge and here the donut profile differs from the fundamental one only in being less steeply sloped (the spot is wider) and having slight inflections of the slope around the hole at the 50% clip point, the beam center.

There are two other two common definitions. The first is the diameter of a circular aperture giving 86.5% transmission when centered on the beam. It is variously called the variable-aperture diameter, the encircled power diameter, or the "power-in-the-bucket" method, and designated by the symbol D_{86}. The last is the second-moment diameter, defined as four times the standard deviation of the radial irradiance distribution recorded by a pinhole scan, and designated by the symbol $D_{4\sigma}$. For the ratio of donut mode to fundamental-mode diameters, these definitions give 1.32 and 1.41 respectively, also different from the three other values above.

After the discussion of some common considerations (Section 1.6.2), these five diameter definitions are evaluated in Section 1.6.3 leading to the summary given in Table 1.1.

1.6.2 General Considerations in Obtaining Useable Beam Profiles

Five questions are important in evaluating what beam diameter method is best for a given application:

1. *How important is it to resolve the full range of irradiance variations?* Only a pinhole scan (or its near equivalent, a CCD camera snapshot read out pixel by pixel) shows the full range, but this is not of significance in some applications, for example, where the total dose of light delivered is integrated in an absorber.
2. *How important is it to use a method that is insensitive to the alignment of the beam into the profiler?* If the test technician cannot be relied on to carefully center the beam on the profiler, the slit or knife-edge methods still give reliable results,

TABLE 1.1

Properties of Mixed-Mode Diameter Definitions

Diameter Symbol	Scan Aperture and Name	Diameter Definition	Conversion Constant $c_{i\sigma}$ to $D_{4\sigma}$	Alignment Sensitive?	Resolution of I(r) Peaks	Convolution Error?	Signal-to-Noise Ratio	Comments
D_{pin}	Pinhole (dia. H)	Separation of clip points $1/e^2$ down from the highest peak	0.805	Yes	High	Yes, if $H/2w > 1/6$	Low	Shows best details of irradiance peaks
D_{slit}	Slit (width S)	Separation of clip points $1/e^2$ down from the highest peak	0.950	No	Medium	Yes if $S/2w > 1/8$	Medium	Directly measured diameter is close to $D_{4\sigma}$
D_{ke}	Knife-edge	Twice the separation of 15.9%, 84.1% clip points	0.813	No	Low	None	High	Most robust diameter experimentally (vs. noise and spot structure)
D_{86}	Variable aperture ("power in the bucket")	Diameter of centered circular aperture passing 86.5% of total power	1.136	Yes	Low	None	High	Works well only on round beams. Computed readily on CCD cameras. Used on kW lasers
$D_{4\sigma}$ or $D_{2\sqrt{2}s}$	Second-moment (linear or radial)	Four times the standard deviation of irradiance distribution from a pinhole scan	1	Yes	High	Yes, as for pinhole scan	Low	*ISO standard diameter.* Susceptible to error from noise on wings of the profile. Supported best by theory. Computed readily on CCD cameras
N/A	CCD camera	Various custom algorithms	N/A	No	Medium	Yes	Low	Computes all of above diameter definitions with appropriate software

but not the other methods. With a CCD camera there is a trade-off between alignment sensitivity and accuracy. For best accuracy, a magnifying lens—of known magnification—can be placed in front of the camera to fill the maximum number of pixels, but then the camera becomes somewhat alignment sensitive.

3. *With what accuracy and repeatability is the diameter determined?* The amount of light transmitted by the mask determines the signal-to-noise ratio of the profile and ultimately answers the question. The methods based on a pinhole scan (D_{pin}, $D_{4\sigma}$, and CCD cameras) suffer from low light levels in this regard. On the other hand, a laser beam is generated in a resonator subject to microphonic perturbations, making the beam jitter in position and the profile distort typically by about 1% of the beam diameter, so that a greater instrument measurement accuracy is usually not significant.
4. *Is the convolution error associated with the method significant?* The convolution error is the contribution to the measured diameter due to the finite dimensions of the scan aperture, either the diameter H of a pinhole or width S of a slit. A 10-micron focused spot cannot be accurately measured with a pinhole of 50-micron diameter. The distortion of a pinhole profile of a fundamental mode is shown in Figure 1.9a as a function of the ratio of hole diameter to the mode width $H/2w$. The peak amplitude drops and a slight broadening occurs as $H/2w$ increases. The central 100% peak amplitude point is "washed out" or averaged to a lower value in the profile by the sampling of lower amplitude regions nearby as the finite diameter pinhole scans across the center as Figure 1.9b indicates. The reduction in peak amplitude of the convoluted profile is like lowering the clip level below 13.5% on the original profile: the measured diameter becomes larger. Very similar profile distortions occur with a slit scan as a function of $S/2w$; here S is the silt width. The ratio of the measured width including this convolution error to the correct width is plotted in Figure 1.9c for the pinhole (H) and slit (S). This gives the rule of thumb for pinhole scans: to keep the error in the measured diameter to 1% or less, keep the pinhole diameter H to one-sixth or less of $2w$, that is, $H < w/3$. The corresponding rule[34] for slits is the measured diameter is in error by <1% if the width S is 1/8 or less of $2w$. For modes like TEM_{10} of Figure 1.2d with a feature (the central peak) narrower than that of the fundamental mode, the aperture widths H or S should be no bigger than these same fractions of the narrow feature's width. (Note, McCally[34] uses the biologist's definition of $1/e$ clip points for the fundamental-mode diameter, a factor $1/\sqrt{2}$ smaller than our $1/e^2$ diameter; his results require conversion.)

Distortion of the profile can be a more subtle effect and can give misleading results. When measuring a predominantly TEM^*_{01} focused beam through the waist region, for example, a pinhole profiler will at first show the expected trace, with a dip in the middle like Figure 1.8d or e. This will change to one with a central peak as in Figure 1.8a at the propagation distance along the beam where the pinhole is no longer small compared to the beam diameter. The donut hole can fall through the pinhole!

Convolution errors are a concern normally only when working with focused beams, as when measuring divergence by the method of Section 1.5.1. Generally, however, it is desirable to go to the far-field, reached by working in 2-space at the focal plane behind an inserted lens, to obtain a true (undistorted) profile. The beam coming out of the laser often has "diffractive

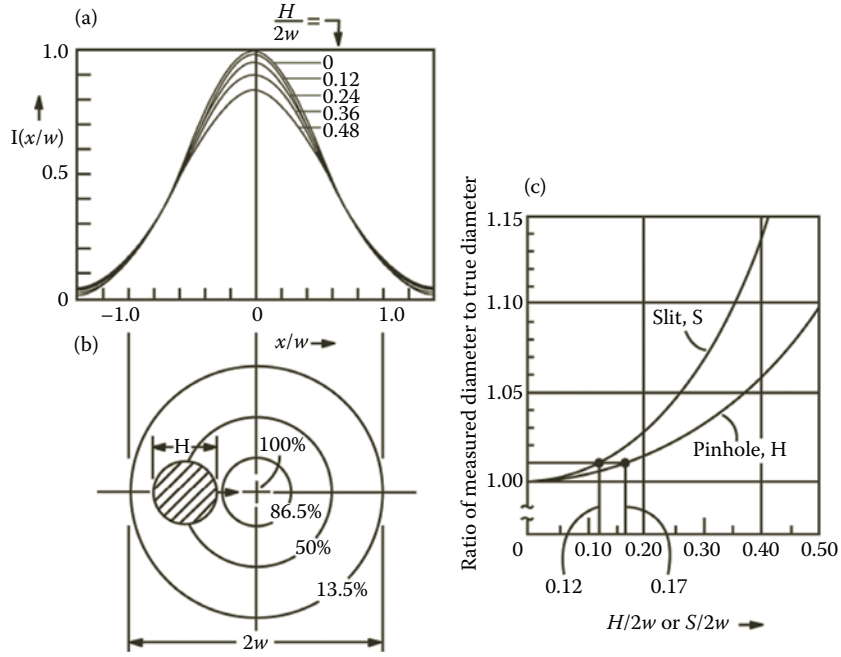


FIGURE 1.9
Convolution of the theoretical fundamental-mode profile in a scan with a pinhole or slit of finite dimensions (H, diameter of the pinhole; S, width of the slit; $2w$, the $1/e^2$ diameter of the mode). (a) Distortion of the shape and width of the pinhole profile as $H/2w$ increases. (b) Plan view of the pinhole scan showing "washout" of the 100% amplitude point. For the pinhole shown, $H/2w = 0.24$, corresponding to the third curve down from the top in (a). (c) Convolution error, or ratio of the measured diameter $2w_{meas}$ to the true diameter $2w$, as a function of $H/2w$ for the pinhole and $S/2w$ for the slit.

overlay," low-amplitude high-divergence light diffracted from the mode-limiting internal aperture, overlaid on the main beam. The resulting interference can significantly distort the profile, even at <1% amplitude of the diffracted light. It is the E-fields that interfere; for an irradiance $I = E^2$ overlaid by a 0.01 E^2 distorting component, the E-fields add and subtract as $E \pm 0.1\ E$ at the interference peaks and valleys. The resulting fringe contrast ratio, $I_{peak}/I_{valley} = [(1.1)/(0.9)]^2 = 1.49$ is a significant distortion to the profile even though the power in the diffractive overlay is insignificant. Moving the profiler some distance away from the output end of the laser spreads the diffractive overlay rapidly compared to the beam expansion, but often several meters additional distance is required. This leaves the use of a lens to reach the far-field as the answer, and convolution distortion then must be dealt with.

Aligning a small-diameter (e.g., 10 micron) pinhole to a small (e.g., 100 micron)-focused spot is another problem. The search time to achieve overlap and some transmitted signal for peaking alignment can be very long if done manually, so having a fast update rate—10 scans a second is good—provided by commercial instruments can be a major aide. Some instruments[9] have electronic alignment systems to facilitate finding the overlap of small pinhole and small beam.

Knife-edges have no convolution error to the extent that they are straight (razor blades are straight[8] to <2 microns deviation over 1000 microns length). The circular aperture of the encircled power method is usually a precision drilled hole and has no convolution

error so long as it is accurately round and made in a material much thinner than the hole diameter (to avoid occultation error).

5. *Are the diameter measurements along the propagation path free of discontinuities and abrupt changes?* Consider making many diameter measurements along the propagation axis, and fitting the data to a hyperbola to find the beam's Rayleigh range and beam quality. Discontinuities in the data will make a poor fit and final result. Such discontinuities can arise[35] with the $1/e^2$ clip-level diameter definitions with mixed modes with low peaks on the edges, as in Figure 1.2g, only lower. As the mode mixture changes to bring the outer peaks near the clip level, the measured diameter can jump from the separation of the outer peaks of the profile to the width of the central peak as amplitude noise perturbs the profile. Similarly, for a mixed mode with rectangular symmetry, as azimuth is continuously changed from the major principal plane direction towards the minor one, the relative amplitude of the outermost peaks of the profile can drop.[35] The clip point then can jump discontinuously with perturbing noise when the height is near the clip level. Only D_{pin} and D_{slit} are subject to this difficulty.

This last question can be rephrased to ask, is the diameter definition readable by a machine? A human observer will notice an outer peak of height near the clip level causing the profiler readout to fluctuate, and correct the situation by adjusting the mode mixture, the azimuth, or the clip level. A machine will take the bad data in, and produce unreliable results. When a lot of diameter data needs to be gathered, as in measuring a propagation plot to determine M^2, automated machine data acquisition is desirable. In this regard, the knife-edge diameter is best, as it always produces an unambiguous monotonic trace for all higher-order and mixed modes.

1.6.2.1 How Commercial Scanning Aperture Profilers Work

Commercial profilers[8] typically use the $1/e^2$ diameter definition with pinhole and slit masks, and occasionally will report an incorrect diameter due to the "not entirely machine readable" defect of these definitions. These profilers use a rotating drum to carry a slit or pinhole mask smoothly and rapidly (typically at a 10 Hz repetition rate) in front of a large area detector inserted into the drum. On the first pass through the laser spot, the electronics remembers the 100% signal level, and on the second pass when the 13.5% clip level is crossed as the signal rises, a counter is started. This counts the angular increments of drum motion from an angular encoder, which when multiplied by the known drum radius, provides the mask translation in spatial increments of 0.2 microns. (In newer, high precision designs discussed in the next paragraph this increment has been reduced to 0.01 microns.) When the clip level is passed as the signal falls, the counter is stopped and the value of the beam diameter—total counts times spatial increment—is reported. Actually, what is reported on the digital readout is an average selected by the user of the last two to 20 measurements, to slow the report rate down to what can be read visually. If a pure donut mode is scanned with the pinhole version of this instrument [the profile of Figure 1.8d], the counter starts at the clip-level dot on the left ($x/w = -1.51$) but stops as the falling clip level is met at the left edge of the donut hole ($x/w = -0.16$). The scan continues and the counter turns on again as the clip level is passed with the rising signal at the right edge of the donut hole ($x/w = +0.16$), because the drum has not completed a revolution to reset the counter for a new measurement. Finally, the counter turns off again at the rightmost clip-level dot

(x/w = +1.51), and the diameter reported is the actual diameter minus the width of the hole at the clip-level height, an error of about –11%. This possible error usually goes unnoticed because the dips in mixed-mode profiles do not often go as low as 13.5%.

In recent years scanning aperture profilers have been mechanically upgraded to provide more precision (0.01 micron spatial resolution) and interfaced with PC controllers to provide more features in addition to beam diameter: full 12-bit digitized profiles and the $D_{4\sigma}$ diameters calculated from them (not just clip widths and analog traces), profile peak position, centroid position, spot ellipticity (with slit or knife-edge profilers carrying two orthogonal apertures), and even absolute power (when so calibrated). With micron-sized apertures and submicron sampling, beam diameters of 5 microns can be measured to 2% accuracy. As before, different detector types (silicon, germanium, or pyroelectric) cover wavelengths from UV to Far IR. Beams pulsed at repetition frequencies down to 1 kHz can be measured with profilers having user-controlled variable scan speed (drum speeds are slowed to intercept enough pulses to build up the profile). In addition they can measure beams without attenuation, as compared to camera-based systems that typically require six to nine orders of magnitude attenuation. Infrared beams at power levels of 3 kW focused to diameters of 175 μm have been directly measured with cooled profilers fitted with copper apertures.

Commercial profilers, because of their speed and accuracy, are a major improvement for frequent beam diameter measurements over the traditional practice of a manually driven translation stage carrying a razor blade (or slit) across the beam. Focused beams in particular need high instrument accuracy to resolve the small spot and provide the real time update rate to acquire a signal by overlapping the aperture with the beam. With a signal linearity range of 10^4 and a spatial resolution (if convolution error is neglected) of 0.01 microns over a 9-mm scan range (10^6 spatial resolution elements) one of these small, new profilers brings an impressive potential of 10^{10} information bits to the problem of measuring a beam diameter. Compare this to a modern CCD camera of 9-mm sensor width, 5-micron pixel spacing (2×10^3 spatial resolution elements), and 12-bit (4×10^3) linearity range, for a total of 10^7 information bits. It is understandable why in measuring beam quality M^2, profiler-based instruments surpass camera-based ones in speed and accuracy. The camera, of course, has its own advantages of giving a two-dimensional map of all the irradiance peaks in the laser spot and its ability to measure beams from low repetition rate pulsed lasers.

1.6.3 Comparing the Five Common Methods for Defining and Measuring Beam Diameters

The discussion that follows and Table 1.1 summarize the properties of the five diameter definitions.

1.6.3.1 D_{pin}, Separation of 1/e² Clip Points of a Pinhole Profile

The pinhole scan reveals the structure of the irradiance variations across the beam spot with the greatest accuracy and detail, but does so working with a low light signal level and it is subject to convolution error with focused spots. To minimize convolution error, several pinholes of diameters H (10-micron and 50-micron pinholes are common) are used to keep $H < w/3$ where w here is the fundamental-mode radius or smallest feature size for a higher-order mode beam. The pinhole method requires accurate centering of the beam on the scan line of the pinhole and this makes it less adaptable to a machine measurement. This diameter definition also can give ambiguous results if the profile contains secondary

peaks of a height close to the clip level. The pinhole profile provides the basic data from which the second-moment diameter is calculated. Be sure the rule for the profile to be free of convolution error is met first!

1.6.3.2 D_{slit}, Separation of 1/e² Clip Points of a Slit Profile

The slit scan does not require centering of the beam spot and works at a medium light signal level, but does not reveal as much detail of the irradiance variations [compare Figure 1.8d and e]. This method is subject to convolution error with focused spots; the slit width S should satisfy $S/2w < 1/8$ with $2w$ as the smallest feature size of the profile. It too can give ambiguous results on profiles with secondary peaks near the clip level. This diameter definition produces a direct result (that is, without applying the conversion rules explained in Section 1.6.4.3) closest to the ISO standard second-moment diameter of the three other methods.

1.6.3.3 D_{ke}, Twice the Separation of the 15.9% and 84.1% Clip Points of a Knife-Edge Scan

The knife-edge does not require centering of the beam spot and works at a high light signal level, but reveals almost no detail of the irradiance variations [compare Figure 1.8d and f], only the slight inflection points in the slope of the knife-edge profile show that there are any irradiance peaks at all. All modes give a simple slanted S-shaped profile. There generally is no convolution error with this method, and there are no diameter ambiguities when secondary peaks are present. Experimentally, it is the most robust diameter measurement and is least affected by beam-pointing jitter and power fluctuations, making this method fully machine readable. This diameter is the basic one measured in the most common commercial instrument[9] designed to automatically measure propagation plots and all six beam parameters.

1.6.3.4 D_{86}, Diameter of a Centered Circular Aperture Passing 86.5% of the Total Beam Power

Unlike the other diameter measurements, the variable-aperture diameter passes light in both the x- and y-transverse planes simultaneously and cannot be used to separately measure the two principal diameters; it works best with round beams. It must also be centered in the beam for accurate results. While an iris or variable aperture can be used, more frequently sets of precision fixed apertures are used instead. A metal plate drill gauge, with some of the plate milled away on the back side of the gauge to reduce its thickness to less than the smallest aperture size to eliminate occultation error, is a convenient tool. The two diameters bracketing the 86.5% transmission point are first found, and the final result computed by interpolation. Alternatively, if there is a long propagation length available, an aperture with a transmission near 86.5% may be moved along the beam to locate the distance where that diameter produces precisely this transmission. This diameter definition is used mainly for two reasons. For high power lasers—for instance CO_2 lasers in the kilowatt range—little diagnostic analytical instrumentation is available that can absorb this power. A water-cooled copper aperture, however, can still be safely inserted in front of a power meter to give some quantification of the beam diameter. The second reason is that this diameter is readily computed from the output of a CCD camera and is available on camera instrumentation, with the computation locating the beam centroid, making physical centering of the camera unnecessary.

1.6.3.5 $D_{4\sigma}$, Four Times the Standard Deviation of the Pinhole Irradiance Profile

This diameter is computed from a pinhole irradiance profile, which for accuracy should be free of convolution error and diffractive overlay. For a beam with a rectangular cross-sectional symmetry described by a weighted sum of Hermite–Gaussian modes the calculation proceeds by finding the rectangular moments of the profile treated as a distribution function. The zeroth moment gives the total power P of the beam, the first moment the centroid, and the second moment leads to the variance σ^2 of the distribution:

$$\text{Zeroth moment or total power } P = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x,y)\mathrm{d}x\mathrm{d}y \tag{1.32}$$

$$\text{First moment or centroid } \langle x\rangle = \left(\frac{1}{P}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xI(x,y)\mathrm{d}x\mathrm{d}y \tag{1.33}$$

$$\text{Second moment } \left\langle x^2\right\rangle = \left(\frac{1}{P}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2I(x,y)\mathrm{d}x\mathrm{d}y \tag{1.34}$$

$$\text{Variance of the distribution } \sigma_x^2 = \left\langle x^2\right\rangle - \langle x\rangle^2 \tag{1.35}$$

$$\text{Linear second-moment diameter } \mathrm{D}_{4\sigma x} = 4\sigma_x \tag{1.36}$$

This last equation comes from the requirement that the second-moment diameter reduce to the $1/e^2$ diameter when applied to a fundamental-mode beam, as explained in arriving at Equation 1.6. A precisely similar set of equations holds for the moments in the vertical plane (y, z) to define a vertical principal plane centroid and diameter [Equations 1.33 through 1.36 with x and y interchanged]:

$$\text{Linear second moment diameter } D_{4\sigma y} = 4\sigma_y. \tag{1.37}$$

A similar set of moment equations defines a radial second-moment diameter, applicable to beams with cylindrical symmetry described by a weighted sum of Laguerre–Gaussian functions. Here the pinhole x-scan profile is split in half at the centroid point $\langle x\rangle$, and the half profile is taken as the radial variation of the cylindrically symmetric beam. In the transverse radial coordinate plane (r, θ), the origin is the center of the beam spot defined by the centroid ($\langle x\rangle$, $\langle y\rangle$) given by the rectangular first moments, Equation 1.33.

$$\text{Zeroth – moment or total power } P = \int_0^{2p}\int_0^{\infty} I(r,\theta)r\mathrm{d}r\mathrm{d}\theta \tag{1.38}$$

$$\text{Radial second moment } \left\langle r^2\right\rangle = \left(\frac{1}{P}\right)\int_0^{2p}\int_0^{\infty} r^3I(r,\theta)\mathrm{d}r\mathrm{d}\theta \tag{1.39}$$

$$\text{Variance of the distribution } \sigma_r^2 \equiv \left\langle r^2\right\rangle \tag{1.40}$$

$$\text{Radial second-moment diameter } D_{2\sqrt{2}\sigma r} = 2\sqrt{2}\sigma_r \tag{1.41}$$

This last equation derives from the requirement that the linear and radial variances are related[6] by:

$$\sigma_x^2 + \sigma_y^2 = \sigma_r^2. \tag{1.42}$$

Then for a cylindrically symmetric mode $\sigma_x = \sigma_y$, yielding $2\sigma_x^2 = \sigma_r^2$ or $\sigma_x = (1/\sqrt{2})\sigma_r$. Since for a fundamental-mode beam $2w = 4\sigma_x$, from the radial mode description of that beam, there results[6] $2w = 4(1/\sqrt{2})\sigma_r = 2\sqrt{2}\,\sigma_r$, which is Equation 1.41. By mixing modes, combinations of Hermite–Gaussian modes can be made to have the same irradiance profiles as Laguerre–Gaussian modes, and vice versa. Therefore, for compactness the symbols $D_{4\sigma}$ or M_{4s}^2 will be used for either linear or radial second-moment quantities unless there is a need to specifically distinguish a quantity as a radial moment.

1.6.3.6 Sensitivity of $D_{4\sigma}$ to the Signal-to-Noise Ratio of the Profile

The experimental difficulties in evaluating these integrals with noise on the profile signal come from the weighting by a high power of the transverse coordinate in the second-moment calculation, by the square in the linear case [Equation 1.34], and by the cube in the radial case [Equation 1.39]. Take as an example a measurement of a fundamental-mode spot with a CCD camera, using 256 counts to digitize the irradiance values, and 128 counts used to digitize half the integration range of the transverse coordinate. In the linear case, one noise count (0.4% noise) at the edge of the range—at the 128th transverse count—is weighted by the factor $1 \times (128)^2 = 16{,}384$ in the integration, versus 256×1 counts for the central peak. The contribution of this single noise count is 64 times that of the pixel at the central peak in the integration. In the radial case, the one noise count at the limiting transverse pixel makes a contribution $(128)^3/256 = 8192$ times that of the pixel at the central peak. A discussion of the high sensitivity of the second-moment diameter to noise on the wings of the profile is given in Reference 12. There the second-moment and knife-edge methods are compared for five simulated modes, and the knife-edge found to be considerably more forgiving and in agreement with common expectations.

To manage this sensitivity to noise, it is essential that both some measure of the detector's background illumination and noise be subtracted from the signal, and that the integration from the beam centroid outward be truncated at the edges of the illuminated region. Both means reduce the effect of noise on the wings of the profile.

A distinction is made between subtraction of background, the detector's readout with the beam blocked, and subtraction of the baseline, the noise floor of the dark detector. Because of the high directionality of laser beams, typically the background can (and should) be reduced to insignificance by inserting an aperture near the laser (blocking concomitant light) and adding a light-shielding tube to the detector (blocking ambient light).

There are differences of opinion as to the best method for subtracting the noise floor with CCD cameras, but recommended here is what is termed "thresholding." From either a dark camera frame or preferably, from the nonilluminated corners of the signal frame, a standard deviation is computed for this measured noise, and three times this value subtracted uniformly from the signal frame before data analysis. This avoids taking the difference between one random noise frame (the background frame) from another (the noise on the signal frame), which often just adds noise.

To set the integration truncation limit, the beam radius is estimated (typically by a diameter-measurement method less sensitive to noise) and the integration is carried out over the range of from three to four (estimated) beam radii. The constancy of the computed second-moment diameter is observed over this range. Then integration limits are set just wide enough to yield a stable second-moment value. When the width setting is judged to be correct, the measurement should be repeated to check reproducibility.

Other problems with CCD cameras that can look like noise are that they are subject to drift, response nonlinearity and nonuniformity, "bleeding" of signal to adjacent pixels, and low damage threshold requiring attenuation not only to prevent signal saturation but to protect them as well. For these reasons, coupled with the need for analysis software to read them out, cameras are best purchased from dealers who have assessed these problems and will stand behind their instrument's measurement accuracy.

In one commercial instrument[9] two additional checks are provided to assess the effect of noise on the radial second-moment calculation done on a pinhole single line scan. The first check compares the second-moment diameter calculated from the right half profile, to that from the left half profile. If the beam is indeed cylindrically symmetric and the contribution from noise on the profile is negligible, the ratio of these two results should be near unity. The second check is an option in the calculation called "noise-clip ON/OFF." In the wings of the 256 count wide profile where the signal is near zero, noise counts vary the trace above and below the average dark level, and the lowest noise pixels acquire a negative sign when the linear baseline (between the means of the 20 points on either end of the scan line) is subtracted. This is desirable, these negative noise pixels help cancel positive ones, but it is straightforward for the processor in the instrument to clip these pixels to a zero value with the "noise-clip" option turned ON. The size of the resulting change in the calculated second-moment diameter provides a test of how large the contribution is from noise in the wings.

It is also recommended when measuring a second-moment diameter to vary the sources of noise on the laser beam. Check that the resonator alignment is peaked, the sources of microphonics impinging on the laser are minimized, the laser is warmed up and bolted down to the stable table, and so forth, and watch for variations in the second-moment diameter. A more complete analysis[9] of the effect of noise on diameter measurements showed that the standard deviation over the mean of ten repeated second-moment diameter measurements was five to ten times larger than that for knife-edge measurements of the same beam at (low) signal-to-noise levels from 50 down to ten. With these precautions required in interpreting $D_{4\sigma}$ results, it is fair to say that the second moment as currently implemented is not a "machine readable" diameter definition.

1.6.3.7 Reasons for $D_{4\sigma}$ Being the ISO Choice of Standard Diameter

Since there is considerable experimental difficulty in measuring second-moment diameters, why is this definition the one adopted[11] as the standard by the International Organization for Standards? The primary answer is that this definition is the one best supported by theory. The general theories of the propagation of ray bundles[4,6,19] are based on the Fourier transform relationship[6] between the irradiance distribution and angular spatial-frequency distribution. These show two essential requirements are met if the beam width is defined by the second-moment diameter [Equation 1.36]. The beam width is rigorously defined[6] for all realizable beams [excluding only those with discontinuous edges,[6] for which the integration Equation 1.34 may not converge] and the square of this width (the variance) increases as a quadratic function of the free space propagation distance away from the waist. That is, $D_{4\sigma}(z)$ increases with z according to the hyperbolic form (Equation 1.16a). All other diameter definitions gain legitimacy in propagation theory by being shown to be proportional to the second-moment diameter.

Another important feature of the second-moment diameter is that the beam quality (M^2 values) calculated using it turn out to be integers for either the pure, rectangular-symmetry Hermite–Gaussian modes, or the pure, cylindrical-symmetry Laguerre–Gaussian modes. Thus not only for the fundamental mode is $M_{4s}^2 = 1$, which happens by definition,

but for the next higher-order mode, the donut mode, $M_{4s}^2 = 2$, and so on counting up by unity each time the mode order increases. In general[6] the formulas are:

$$\text{Hermite – Gaussian modes} \quad M_{4s}^2 = (m + n + 1) \tag{1.43}$$

$$\text{Laguerre – gaussianmodes} \quad M_{2\sqrt{2}s}^2 = (2p + l + 1) \tag{1.44}$$

where m, n are the order numbers of the Hermite polynomials, and p, l the order numbers for the generalized Laguerre polynomials associated with the modes as before (Equation 1.1). For the six modes shown in Figure 1.2, of increasing order from (a) to (f), the values are $M_{4\sigma}^2 = 1, 2, 3, 3, 4, 4$ respectively. The integers $(m + n + 1)$ or $(2p + l + 1)$ are termed the mode order numbers, and they determine as well the mode's optical oscillating frequency. Modes with the same frequency are termed degenerate. As the mode order number increases, the degree of degeneracy increases, there being three degenerate pure modes each for $(2p + l + 1) = M^2 = 5$ or 6, four for $M^2 = 7$ or 8, five for $M^2 = 9$ or 10, and so on. The diameters of the pure modes in second-moment units are just the square root of the mode order numbers times the fundamental-mode diameter (by Equation 1.3):

$$\text{Pure Hermite – Gaussian modes} \quad \frac{D_{4s}}{2w} = \sqrt{m + n + 1} \tag{1.45}$$

$$\text{Pure Laguerre – Gaussian modes} \quad \frac{D_{2\sqrt{2}s}}{2w} = \sqrt{2p + l + 1} \tag{1.46}$$

Another consequence of the pure modes having integer values of beam quality is that for mixed modes, the M_{4s}^2 value is a simple power-weighted sum of the integer M_{4s}^2 values of the component modes. Finding integers like this in a physical theory is strong indication that the quantities have been defined and measured "the way nature intended."

Another reason for the ISO Committee's choice of $D_{4\sigma}$ as the diameter standard is that the committee members were aware that conversion formulae were available to permit diameters measured according to the other definitions to be put in standard form. These formulae are discussed in the next section.

The last line of Table 1.1 refers to CCD camera properties. A CCD camera together with frame-grabber electronics and appropriate software can be a universal instrument capable of providing diameter measurements according to any or all of the definitions. Affordable cameras do not provide as large a dynamic range for irradiance levels (useful range ~1000:1) compared to that for a silicon detector (~10^4) but good variable attenuators are readily available[33] to allow camera operation just below saturation to make the most of the range that exists. Spatial resolution of 5 micron per pixel may be inadequate for direct measurement of focused beams but flexibility, ease of use, and quick access to colorful 2-D irradiance maps make it an attractive choice for beam diameters large enough to fill an adequate number of pixels. Imaging optics can be used if necessary to measure smaller beams. If improvements in CCD cameras continue at their recent pace, they are likely to become superior to all the older methods of measuring beam diameters.

1.6.3.8 Diameter Definitions: Final Note

It is important to emphasize that the M^2 model can be applied using any reasonable definition of beam diameter as long as the definition is used consistently both in making measurements and interpreting calculated values. Results will then be meaningful and reliable.

In fact, there can be cases where it is important to use a "nonstandard" diameter definition. For example, there is a trend toward steeper sides and flattened tops as M^2 increases. The effect becomes pronounced for M^2 values above ten and at 50 or more, profiles can be aptly described[5] as a "top hat" shape. The diameter of such a beam becomes unambiguous and it makes sense to abandon the standard definitions ($D_{4\sigma}$, D_{86}, etc.) and just measure the diameter of the "top hat" cylinder. The good news is that for such beams, pinhole scans would show the diameter at half-maximum irradiance to be insignificantly different from that at the $1/e^2$ level. The aperture size that passes 86.5% of the total power will not provide as meaningful a result in this situation as the aperture that transmits 95% of the power. The latter would likely be little different in size from the one that passes 98%. Curve fitting to a series of D_{95} measurements will yield a set of valid parameters describing the beam but this defines a new "currency" and one must stay consistent and not mix these diameters with those arrived at by a different method or definition.

1.6.4 Conversions between Diameter Definitions

For a diameter conversion algorithm to be widely applied, it must be normalized, with the natural normalization being the diameter of the fundamental mode generated in the same resonator as the measured beam, the embedded gaussian. Using Equation 1.3, this essentially changes the problem of converting diameters into one of converting M^2 values.

The conversion rules that are now part of the ISO beam widths document[11] were first derived empirically and later found to have theoretical support. They apply to cylindrically symmetric modes generated in a resonator with a circular limiting aperture and an approximately uniform gain medium. In this case, if $M^2_{2\sqrt{2}s}$ is known, then the mixture and relative amplitudes of the modes oscillating can also be reasonably estimated.

1.6.4.1 Is M^2 Unique?

Determining the fractions of the pure modes in a mixture for a cylindrically symmetric beam from the beam quality alone seems unlikely at first, because the beam quality M^2 is not unique in the mathematical sense. Consider the case of a beam with $M^2 = 1.1$ in second-moment units. An experienced laser engineer might guess the likely composition is 90% fundamental mode ($M^2 = 1$) and 10% donut mode ($M^2 = 2$) to give $M^2 = (0.9) + 2(0.1) = 1.1$ for the mixed mode, and she/he would be right. For a beam of $M^2 = 5$ however, the problem is much harder. The number of possible modes above threshold makes for a large range of possible mix fractions within the $M^2 = 5$ constraint.

Our empirical results showed, however, that for the class of lasers with round beams described just, M^2 was unique at least up to values of $M^2_{4s} = 3.2$.[14] In these resonators, diffraction losses and spatial mode competition in saturating the gain determine the mixed-mode composition. As the circular limiting aperture is opened—as the Fresnel number of the resonator is increased—some modes grow and others decrease in a predictable and reproducible way, such that for each M^2 there is a unique known mode mixture. Furthermore, this knowledge has allowed us to establish mathematical rules for interconversion of beam diameters between the various measurement definitions.

1.6.4.2 Emprical Basis for the Conversion Rules

We acquired the empirical data[14] by using an argon ion laser set up to give beams with a large range of M^2 values as a function of the diameter of the circular mode-limiting

aperture. By varying this aperture diameter and the gain—the latter by adjusting the laser tube's current—values of $M^2_{2\sqrt{2}s}$ from 1 to 2.5 were covered with the green line at 514 nm; the upper limit was increased to 3.2 by changing to the higher gain of the 488 nm blue line. As the blue line was generated in the same resonator, the blue beam diameters here could be scaled by multiplying the square root of the ratio of the wavelengths, a factor of 1.027, for comparison to the green line diameters. The beam from this laser was split to feed an array of monitoring equipment. A radio frequency photodiode and rf spectrum analyzer gave how many modes and what mode orders were oscillating. Profiles were recorded with a commercial slit and pinhole profiler[8] and a commercial beam propagation analyzer[9] to obtain knife-edge diameters, M^2_{ke}, and radial second-moment diameters. A CCD camera and software computed the variable aperture diameter. In front of the camera, a lens provided a known (1.47 times) magnification to fill an adequate number of pixels, and a variable attenuator set the light level.

As the laser's internal aperture was opened and the beam diameter enlarged, the mode spot alternated from one with a peak in the center to one with a dip at the center in over one-and-a-half cycles as shown in the profiles of Figure 1.10b. Seven aperture settings were chosen spanning the range of M^2 values, two giving the highest central peaks (A and E), two at the deepest dips (C and F), and three transitional ones (AP, named the "perturbed A-mode", B, and D). The full set of diagnostic data at these settings was recorded. Knowing the number of modes oscillating and the mode orders at each setting from the rf spectrum, trial mode mixtures were assumed. The resulting theoretical profiles were adjusted[14] to match the experimental pinhole profiles. An example is Figure 1.2, where the theoretical mixed-mode profile, (g), is matched to experimental profile, (h), which is the same as Mode E in Figure 1.10b.

Once the TEM^*_{0n} modes were included[14] in the mode mix, good matches of profiles were found. These modes are like the donut mode, for which $n = 1$, but with increasingly larger holes in the center as their order $(n + 1)$ increases. Because they have $p = 0$ they are "all null" (nearly zero in amplitude) in the middle. They make the most of the r^3 weighting factor in the second-moment integral to reach a given second-moment diameter $Mw = \sqrt{(2p + l + 1)}\, w$ at the smallest radius, resulting in the lowest tails[14] to their profiles of all modes of the same-order number. They thus have the lowest diffraction loss for a limiting circular aperture and always oscillate first among pure modes of the same order as the Fresnel number of the resonator is increased. It was noted in Reference 20 that in this aperture-opening process there was a gradual extinction of a mode of lower order soon after a mode of next higher order reached threshold. This is clearly a gain competition effect won by the higher-order mode. A possible physical reason of general applicability discussed in Reference 20 was that the larger spatial extent of the higher-order mode provided access to a region of gain not addressed by the competing lower-order mode.

The final mode fractions for the seven mixed modes were determined using a Mathematica function called SimpleFit made available by Wolfram Research. These fractions are plotted in Figure 1.10a as a function of the resultant beam quality M^2_{4s} for the mixed modes. The modes turn on in the order of decreasing diffraction loss as shown by McCumber[21] and then gradually extinguish, as predicted in the preceding paragraph. At each value of M^2_{4s} for this argon ion laser there is a characteristic set of oscillating modes, mode fractions, and mode profiles (Figure 1.10). Here for every M^2 value there is a unique mixture of modes. From all the data gathered, simple conversion rules given in the next section between diameter definitions were derived. Over the range measured of $M^2_{4s} = 1$ to 3.2, the error to convert knife-edge, slit, and variable aperture diameters to second-moment diameters was ±2% (one standard deviation). This is a ±4% error in converting M^2. The error was ±4% for conversion of pinhole diameters to second-moment diameters.

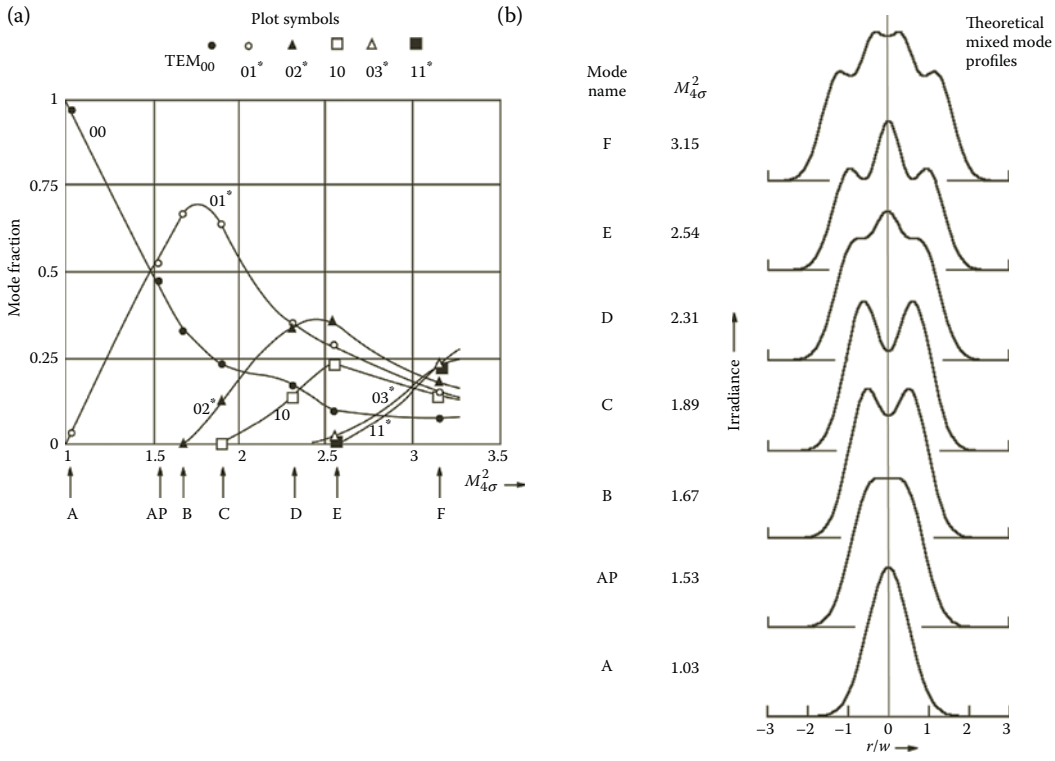


FIGURE 1.10
Observed mode fractions for a beam from a resonator with a limiting circular aperture. As the aperture diameter increases M^2_{4s} follows, with the mode fractions changing in a characteristic fashion as higher-order modes come above threshold. (a) The mode fractions as a function of M^2_{4s}. (b) The computed pinhole profiles and their $M^2_{4\sigma}$ values for the characteristic set of mixed modes A to F measured to determine the mode fractions.

We then tested the rules on other lasers[14] within this M^2 range and found that knife-edge diameter measurements converted to second-moment diameters agreed with directly measured second-moment diameters within ±2%. The conversion error is defined as the fraction in excess of unity of the $D_{4\sigma}$ diameter obtained by the conversion rule, over that obtained directly from the variance of the irradiance profile, expressed in percent. The knife-edge diameter conversion subsequently was tested on three other gas lasers at M^2_{4s} = 4.2, 7.5, and 7.7 and found to remain valid to ±2%. However, a test[25] on a pulsed Ho:YAG laser at M^2_{4s} = 13.8 gave a conversion error of −9%; this is thought to be due to the strong transient thermal lensing in this medium affecting the spatial gain saturation. This consistency in the face of an extrapolation by a factor of two indicates that these conversion rules are fairly robust, valid to the stated accuracy, and that the mixed modes on which they are based exist in this large class of lasers. Apparently, for many lasers, M^2 is unique.

1.6.4.3 Rules for Converting Diameters between Different Definitions

The empirical results showed there was a linear relationship between $M_i = \sqrt{M_i^2}$ and the square root of the second-moment beam quality $M_{4s} = \sqrt{M^2_{4s}}$, where M_i is the square root

of the beam quality obtained by method "*i*" and *i* can signify any of the other definitions. Since all the diameter definitions give the same result for the fundamental-mode beam (for which the beam quality is unity) the linear relationship can be expressed with a single proportionality constant $c_{i\sigma}$ in the form:

$$M_{4s} - 1 = c_{is}(M_i - 1) \tag{1.47}$$

for the conversion from the method "*i*" to second-moment quantities. This form ensures that the linear plot of $M_{4\sigma}$ versus M_i passes through the origin with no offset term and that only the slope constant *c* is required to define the relationship.

In the same resonator, the fundamental-mode diameter is given by the ratio of the mixed-mode diameter to *M*. This is true independent of what diameter definition is used, and thus a second relationship is:

$$\frac{D_i}{M_i} = 2w = \frac{D_{4s}}{M_{4s}}. \tag{1.48}$$

Here D_i is the diameter obtained by method "*i*." Substituting Equation 1.48 into Equation 1.47 yields:

$$D_{4s} = \left(\frac{D_i}{M_i}\right)[c_{is}(M_i - 1) + 1]. \tag{1.49}$$

The values of the conversion constants $c_{i\sigma}$ are listed in Table 1.1 to convert from the diameter definitions summarized there to the second-moment diameter, $D_{4\sigma}$.

Since each of the other diameter methods is linearly related to the second-moment diameter, they all are linearly related. The conversion constants between the other methods can be obtained from those for the second-moment conversions. Let one of the other methods be denoted by subscript "*j*." From Equation 1.47 there results:

$$(M_{4s} - 1) = c_{is}(M_i - 1) = c_{js}(M_j - 1)$$

therefore

$$(M_i - 1) = \left(\frac{c_{js}}{c_{is}}\right)(M_j - 1).$$

By definition of a conversion constant for method $i \rightarrow j$,

$$(M_i - 1) = c_{ji}(M_j - 1).$$

Hence:

$$c_{ji} = \left(\frac{c_{js}}{c_{is}}\right). \tag{1.50}$$

This gives the conversion constants between any two methods in Table 1.1, by taking the ratios of their constants for conversion to the second-moment values. Note that Equation 1.50 also implies that $c_{ji} = 1/c_{ij}$, which is also useful.

The values for the $c_{i\sigma}$ constants in Table 1.1 are an improvement over our earlier results[14] that were incorporated in the ISO beam-test document.[11] More experimental data later became available, but also it was realized once the mode fractions were determined experimentally that the conversion constants could then be calculated from theory alone. From the mixed-mode set A to F defined by the mode fractions of Figure 1.10a, each of the

theoretical diameters D_i for the different methods was calculated. By Equation 1.3, these were converted to M_i's. Then plots of $M_{4\sigma} - 1$ versus $M_i - 1$ were least-squares curve fit to determine by Equation 1.47 the values of $c_{i\sigma}$ listed in Table 1.1. The fit for the slope $c_{i\sigma}$, was for one parameter only with the intercept forced to be zero. This gives an internally consistent set of $c_{i\sigma}$'s so that Equation 1.50 is valid.

1.7 PRACTICAL ASPECTS OF BEAM QUALITY M^2 MEASUREMENT: THE FOUR-CUTS METHOD

The four-cuts method means measuring the beam diameter at four judicious axial positions, the minimum number—as explained in this section—to permit an accurate determination of M^2. To execute this method well, several subtleties should first be understood.

The simplest way to measure M would be to take the ratio of the mixed-mode beam diameter to that of the embedded gaussian as by Equation 1.3, $M = W/w$, except that the embedded gaussian is inaccessible by being enclosed inside the mixed mode. However, both beams have the same Rayleigh range. By measuring z_R and the waist diameter $2W_0$ for the accessible mixed mode, the beam quality is determined through Equation 1.20:

$$M^2 = \frac{\pi W_0^2}{\lambda z_R}. \tag{1.20}$$

The general approach is to measure beam diameters $2W_i$ at multiple locations z_i along the propagation path and least-squares curve fit this data to a hyperbolic form to determine z_R and $2W_0$. But even by taking this computer-intensive approach, unreliable values will sometimes result unless a number of subtle pitfalls[25] (often ignored) are avoided on the way to good (±5%) M^2 values. The pitfalls are highlighted in *italics* as they are encountered in this discussion.

Well-designed commercial instruments[9] avoid these pitfalls, and a button push yields a good answer. For the engineer performing the measurement on his or her own, and who can start by roughly estimating the beam's waist diameter and location (using burn paper, a card inserted in the beam, or a profiler slid along the propagation axis) a minimum effort, logical, quick method exists, which circumvents the subtle difficulties. This is the method[25] of "four cuts," the subject of this section.

The first pitfall is avoided by realizing that in the M^2 model the beam divergence is no longer determined by the inverse of the waist diameter alone (as it is for a fundamental mode) but has the additional proportionality factor M^2:

$$\Theta = \frac{2M^2\lambda}{\pi W_0}. \tag{1.19}$$

The first implication of this additional degree of freedom is that the beam waist must be measured directly, not inferred from a divergence measurement. Consider the propagation plots shown in Figure 1.11a. Several beams are plotted, all with the same values of the ratio M^2/W_0 and therefore the same divergence, but with different M^2 [accomplished by having the Rayleigh range proportional to W_0, see the second form of Equation 1.19 in Section 1.4]. From measurements all far from the waist it would be impossible to distinguish between

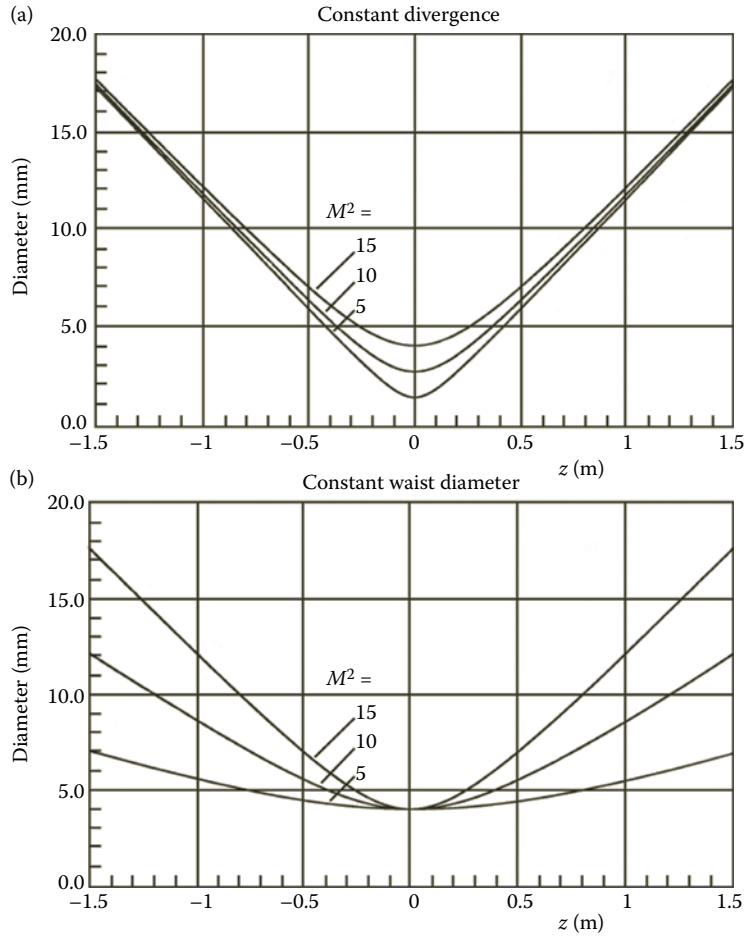


FIGURE 1.11
Beams of constant divergence (a) and constant waist diameter (b) to illustrate the consequences of $M^2 \neq 1$. The beam must be sampled in both near- and far-fields to distinguish these possibilities. The curves are drawn with values appropriate for a beam of λ = 2.1 microns. (Redrawn from Johnston, T.F., Jr. *Appl. Opt.* 1998, *37,* 4840–4850.)

these curves to determine M^2. On the other hand, in Figure 1.11b are propagation plots for several beams with the same waist diameters but different M^2 and therefore divergences. Here $\Theta \propto M^2$ and by Equation 1.18, $z_R \propto 1/M^2$. Measurements all near the waist could not distinguish these curves to determine the divergences. Both near- and far-field diameter measurements are needed to measure M^2.

Any of the diameter-measurement methods can be used to define an M^2 value, and *the next pitfall is avoided by staying in one currency,* and do not mix, for instance, the knife-edge divergence measurement with the laser manufacturer's quoted $D_{4\sigma}$ (second-moment) waist value. Consistently use the most reliable diameter-measurement method you have available, and in the end convert your results to values in the standard $D_{4\sigma}$ currency.

1.7.1 The Logic of the Four-Cuts Method

The four-cuts method starts with the error estimate for your best method for measuring diameters, and uses that to set the tolerances on all other measurements. Let diameters be determined to a fractional error g,

$$g = \left(\frac{2W_{\text{meas}}}{2W}\right) - 1 \tag{1.51}$$

where $2W_{\text{meas}}$ is the measured diameter, and $2W$ the correct diameter. It is assumed g is small, usually 1%–2%. This will yield a fractional precision h for the beam quality of h = 3%–5% since M^2 varies as the product of two diameters, with a small error added for a required lens transform (discussed in Section 1.7.1.1). The term "cut" is used for a diameter measurement, after the common use of a knife-edge scan cutting across the beam to determine a diameter. Let us define the normalized or fractional propagation distance from the waist as:

$$h(z) = \frac{(z - z_0)}{z_{\text{R}}}. \tag{1.52}$$

Let the fractional error in locating the waist be η_0. For this miss in cut placement in measuring the waist diameter $2W_0$ to cause an error of less than g, Equation 1.16a gives:

$$\sqrt{1 + h_0^2} < g + 1 \quad \text{or} \quad h_0 < \sqrt{2g} \tag{1.53}$$

for $g \ll 1$. If $g = 0.01$, then $\eta_0 < \sqrt{(0.02)} \cong 1/7$. *The tolerable error in locating z_0 is one-seventh of a Rayleigh range for a 1% precision in diameter measurements.*

To locate the waist to this precision, beam cuts must be taken far enough away from the waist to detect the growth in beam diameter with distance. At the waist location the diameter change with propagation is nil; to precisely locate a waist requires observations far from it where the diameter variation can be reliably detected. On both sides away from the waist, cuts must be made at distances of a sizeable fraction of the Rayleigh range.

To find the optimum cut distances, look at the fractional change Q in beam diameter versus normalized propagation distance:

$$Q \equiv \left(\frac{1}{W}\right)\frac{\mathrm{d}W}{\mathrm{d}h} = \frac{h}{(1 + h^2)}. \tag{1.54}$$

Figure 1.12 is a plot of this function, Equation 1.54, in which it is easy to see that the maximum fractional change of Q occurs at $\eta = \pm 1$. By making cuts within −2 to −0.5 and +0.5 to +2.0 Rayleigh ranges from the waist corresponding to η within these numerical values, 80% of the maximum fractional change is available. This will significantly enhance the reliability of the position determination over that made using diameters from less than 0.5 z_{R} away from the waist. *An accessible span of at least a Rayleigh range centered on the waist is needed for diameter measurements.*

Note that Figure 1.12 highlights the physical significance of the propagation locations one Rayleigh range to either side of the waist. The wavefront curvature is largest in absolute magnitude there, resulting in the fractional change in diameter Q with propagation coordinate z reaching extremes of ±0.5 there as well.

1.7.1.1 Requirement of an Auxiliary Lens to Make an Accessible Waist

Most lasers have their beam waists located internally where they are inaccessible. Therefore, an accessible auxiliary waist related to the inaccessible one is achieved by inserting a lens

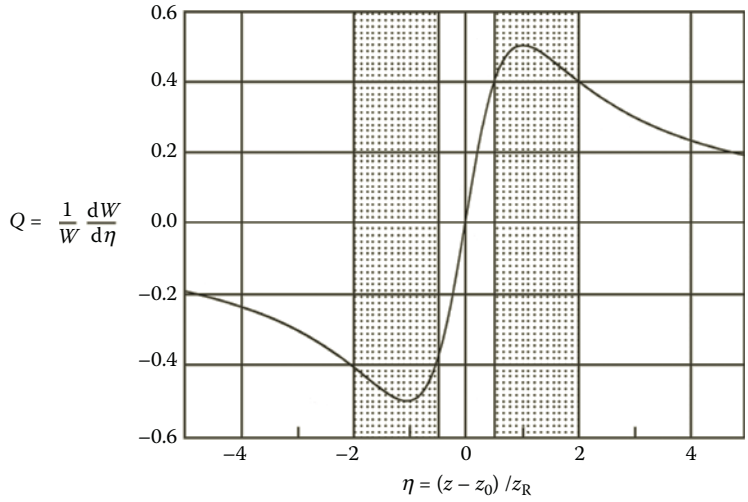


FIGURE 1.12
The fractional change Q in beam diameter as a function of the normalized propagation distance from the waist. Cuts made to locate the waist in the shaded regions benefit from a fractional change of 80% or more of the maximum change. This requires a minimum of one Rayleigh range of access to the beam around the waist location. (Redrawn from Johnston, T.F., Jr. *Appl. Opt.* 1998, *37*, 4840–4850.)

or concave mirror into the beam, and making the M^2 measurement on the new beam. Then the constants found are transformed back through the lens to determine the constants for the original beam. *This requirement to insert a lens, and then transform through the lens back to the original beam constants, is an often-ignored pitfall in making accurate beam measurements.*

The temptation is to use what is available, and just measure the beam on the output side of the output coupler. Usually this means the data is all on the diverging side of the waist. The problem is that nothing in this data constrains the waist location very well. In the curve fit, small errors in the measured diameters will send the waist location skittering back and forth to the detriment of the extrapolation to find the waist diameter. Inserting a lens and making a beam that is accessible on both sides of its waist is a significantly more reliable procedure.

There are three constants (z_{02}, $2W_{02}$, M^2) needed to fix the 2-space beam shown in Figure 1.6 for one of the principal propagation planes, so, in principle, only three cuts should suffice, but then one of them would have to be within the range $|\eta_0| < 1/7$. The location of this narrow range $z_{02} \pm z_{R2}/7$ is at this point unknown. Therefore four cuts are used, the first an estimated Rayleigh range z_{R2} to one side of the estimated waist location z_{02}, the second and third at about 0.9 and 1.1 times this estimated Rayleigh range to the other side (see Figure 1.13). These cut locations and the diameters determined there are labeled by their cut numbers i = 1, 2, 3. Between z_2 and z_3 there is a diameter that matches $2W_1$ and the location z_{match} of this is determined by interpolation:

$$z_{match} = z_2 + \frac{(z_3 - z_2)(W_1 - W_2)}{(W_3 - W_2)} \tag{1.55}$$

$$z_4 = z_{02} = \frac{(z_1 + z_{match})}{2} \tag{1.56}$$

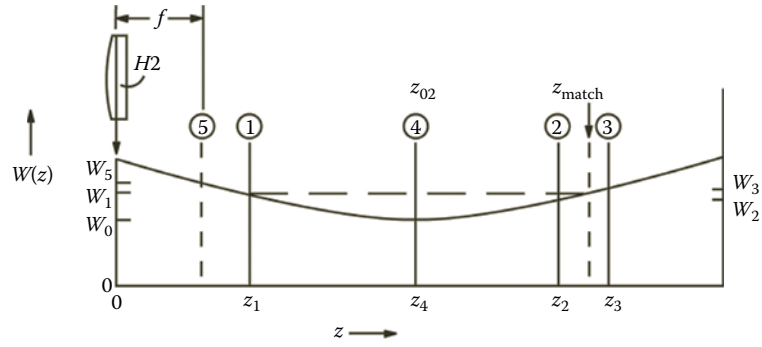


FIGURE 1.13
The four-cuts method. Shown is the beam propagation plot in 2-space, behind the inserted auxiliary lens; the circled numbers indicate the order of the cuts made to locate the waist. The propagation distance z_{match} of the diameter matching that at the first cut at z_1 determines the waist location z_{02} as halfway between these equal diameters. (Redrawn from Johnston, T.F., Jr. *Appl. Opt.* 1998, *37*, 4840–4850.)

The waist is located exactly halfway between z_1 and z_{match}, and the fourth cut is made there at z_4 to directly measure the waist diameter $2W_{02} = 2W_4$ of the 2-space beam and complete the minimum data to determine M.²

1.7.1.2 Accuracy of the Location Found for the Waist

If the locating cuts (1, 2, and 3 of Figure 1.13) are within the ranges specified from $|Q| > 0.4$ and the diameters are measured to the fractional error g, then the error in the normalized waist location η_0 is no worse than $g/Q = 2.5\%$. This is much less (since g is small) than the tolerance $\sqrt{2g} = 14.1\% = 1/7$ determined from inequality Equation 1.53. The measured waist diameter is then correct to the fractional error g.

The fractional error in measurement of diameters g when divided by the fractional change in diameter with normalized propagation Q, yields the fractional error in normalized waist location $\eta_0 = g/Q$. The plot of Figure 1.12 is thus actually a quantitative version of the statement "to precisely locate a null requires observations far from the null" when locating the waist. Diameter measurements inside the range $z_{02} \pm z_R/2$ quickly lose any ability to contribute precision in locating the waist as here Q drops to zero.

There is much value in locating the waist as accurately as the diameter-measurement tolerance will allow in that it reduces the number of unknown constants to be determined by curve fitting from three to two. The number of terms in the curve fit drops by a factor of four, and the remaining terms are made more accurate. Some of these terms depend on the distance from the waist to the ith-cut location, $z_i - z_{02}$, either squared or raised to the fourth power. It is often useful to take a fifth cut at $z_5 = f$ as shown by the vertical dashed line in Figure 1.13. This cross checks the input beam divergence by Equation 1.29 and balances the number of points on either side of the auxiliary waist at z_{02} to improve the curve fit.

1.7.2 Graphical Analysis of the Data

The data, which consists of a table of four- or five-cut locations and their beam diameters for each of the two independent principal propagation planes, is next plotted. A sample plot for the $\lambda = 2.1$ micron Ho:YAG laser beam analyzed in Reference 25 is shown in Figure 1.14. There it was found that with as few data points as required in the four-cuts method, and

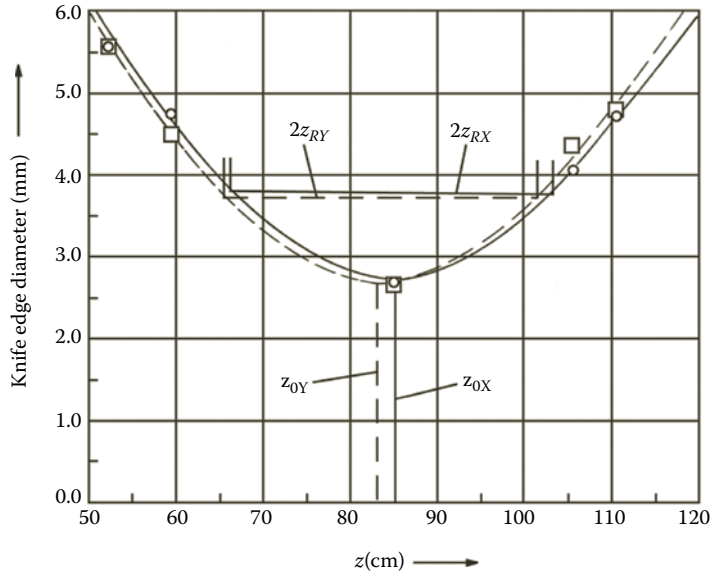


FIGURE 1.14
An example of graphical analysis of propagation data for the auxiliary beam in 2-space. The chords give the Rayleigh ranges for the *x*- and *y*-planes. They are drawn at ordinates on the plot √2 larger than the waist diameters located at z_{0x} and z_{0y}. (Redrawn from Johnston, T.F., Jr. *Appl. Opt.* 1998, *37*, 4840–4850.)

with the initial waist location and Rayleigh range estimates close to the final values (within ~10%), a simple and quick graphical analysis is as accurate as a curve fit.

Generally, with more points as in commercial instrumentation, a weighted least-squares curve fit of the data to a hyperbolic form is required,[25] discussed in Section 1.7.3. The curve fit also generates a sum of residuals for a statistical measure of the goodness of fit.

In the graphical analysis after the points are plotted, smooth curves of approximately hyperbolic form are laid in symmetrically about the known waist locations for each principal propagation plane, here in Figure 1.14 with a French curve. Next, horizontal chords are marked off at heights √2 times the waist diameters $2W_4$ to intersect the smooth curves. The distance between these intersection points on each curve are twice the Rayleigh ranges $2z_{Rx}$, $2z_{Ry}$ respectively, and these lengths are measured off the plot for use in Equation 1.20 with $2W_{0x} = 2W_{4x}$ (and $2W_{0y} = 2W_{4y}$) to determine M_x^2 (and M_y^2) for the auxiliary 2-space beam. For the data of Figure 1.14 the results were $z_{Rx} = 17.6$ cm and $z_{Ry} = 17.8$ cm, resulting in knife-edge beam qualities $M_x^2 = 15.4$ and $M_y^2 = 14.9$.

These results are termed the initial graphical solution and can be improved to give the corrected graphical solution by using the fact that a better estimate of the waist diameter is available than just the closest measured point. By the propagation law, Equation 1.16, if the miss distance of the closest point (Cut 4) is η_0 then the best estimate of the corrected waist diameter is:

$$2W_{02} = \frac{2W_4}{\sqrt{(1 + h_0^2)}}. \tag{1.57}$$

The corrected solution uses the Rayleigh range and waist values from the initial graphical solution in Equation 1.57 to obtain a corrected waist diameter, and plots a chord at a height of √2 times this diameter to determine a corrected length $2z_R$ and M^2 from Equation 1.20. In the example of Figure 1.14, the chords shown are the corrected chords; only the *y*-axis data changed slightly to $z_{Ry} = 17.3$ cm and $M_y^2 = 15.2$. After curve fitting the same data, the fractional rms error (goodness of fit) for the five diameter points were the same at <1.9%.

This good accuracy is a consequence of the four-cuts strategy. The waist diameter is directly measured and if the initial estimate for the Rayleigh range is close, the other cuts give data points near the intersection points of the chords fixing the $2z_R$ values on the plot. The graphical analysis then amounts to an analog interpolation to find the best positions for the intersection points.

There are two last steps. The first is to transform the 2-space data back to 1-space to get the constants for the original beam, using Equations 1.24 through 1.28. This adds a small fractional error to the end result due to the uncertainties in z_{02} and z_{R2}, which contribute a slight uncertainty to the transformation constant Γ of Equation 1.24 (in the example of Reference 25, a 2% error in Γ, 1% additional error in transformed diameters).

The second step is to convert these knife-edge measurements of Figure 1.14 to standard second-moment units as done in Table 3 of Reference 25. The beam of Reference 25 is the one that did not work well with the conversion rules of Section 1.1.5. Instead the conversion of $M_{ke}^2 = 15.4$ to $M_{4s}^2 = 13.8$ was done by comparing measurements at cut 5, the focal plane of the auxiliary lens, of the knife-edge diameter to the second-moment diameter calculated from a pinhole scan. This gave the ratio $D_{ke}/D_{4\sigma} = 1.055$ or a factor of $1/(1.055)^2 = 0.897$ for the M^2 conversion.

1.7.3 Discussion of Curve-Fit Analysis of the Data

A complete numerical example of a full weighted least-squares curve fit to analyze the four-cuts data, or a larger data set, is given in Reference 25 and need not be repeated. There are some subtle pitfalls to avoid in using curve fits on beam propagation data and these are briefly discussed.

A least-squares curve fit is the only general way to account for all the data properly. A common mistake *is to use the wrong function for the curve fit*, which necessitates a discussion of what is the correct one. The fit should be to a hyperbolic form, Equation 1.16, but that is not all. It also should be a weighted curve fit, with the weight of the ith squared residual in the least-squares sum being the inverse square power of the measured diameter $2W_i$.

There are three reasons for this choice of weighting. The first is that in general in a weighted curve fit, the weights[36] should be the inverse squares of the uncertainties in the original measurements. For many lasers, the fractional error in the measured diameter is observed to increase with the diameter; this is probably due to the longer time it takes to scan a larger diameter. The spectrum of both amplitude noise and pointing jitter on a beam tends to increase towards lower frequencies and longer measurement times give this noise a greater influence.

The second reason arises from an empirical study[25] of different weightings one of us did during the development of a commercial M^2 measuring instrument.[9] Amplitude noise was impressed on the beam of a fundamental-mode ion laser with a known $M_{4s}^2 = 1.03$, by rapid manual dithering of the tube current while the instrument's data gathering run[9] was underway. (Note, the ModeMaster[9] collects 260 knife-edge cuts in each of two orthogonal planes in a 30-s "focus" run, generating the beam propagation plots.) The same data was then fitted to a hyperbola five times, with five different weighting factors.

The weights were the measured diameter raised to the nth power, $(2W_i)^n$, with $n = -1$, -0.5, 0, $+0.5$, or $+1$. The weight with $n = 0$ is unity or equal weight for all data points. Data runs were repeated many times with increasing noise amplitude, and the resulting M^2 values for all five weighting schemes were compared each time. The equal or negative power weightings gave stable M^2 values within 3% of the correct value up to 5% peak-to-peak amplitude noise. The positive power weightings $n = +0.5$ gave 4%–5% and $n = +1$ gave 12%–19% errors in M^2 respectively at this noise level. With larger noise amplitudes, the positive power weightings gave errors that grew rapidly and nonlinearly.

A common curve-fitting technique is to use a polynomial fit for the square of the beam diameter versus propagation distance. This may be convenient but it could give an unsatisfactory result. This technique takes advantage of the wide availability of polynomial curve-fit software, and the fact that the square of Equation 1.16 gives a quadratic for $W(z)^2$ as a function of z. However, look at what this does. Let $2W_i$ be the measured ith diameter, and $2W_i'$ be the exact diameter with the small deviation between them $2\delta_i = 2W_i - 2W_i'$. In the W^2 polynomial curve fit, the ith term is

$$(W_i)^2 = (W_i' + d_i)^2 = (W_i')^2 + 2W_i' d_i$$

making the residual

$$(W_i)^2 - (W_i')^2 = 2W_i' d_i.$$

The residual from the exact polynomial curve is weighted in the fit by $2W_i'$, a positive power (+1) of W_i', and so will give unstable results if there is more than a few percent amplitude noise on the beam. At the time of completion of the 1995 ISO document on beam-test procedures, this difficulty with a polynomial curve fit was unrecognized, and a polynomial fit was (incorrectly) recommended there; the 2004 version correctly recommends fitting to a hyperbolic form.[11]

The third reason for an inverse-power weighting is that mathematically the least fractional error results for a ratio quantity like $M^2 = \Theta/\theta_n$ in Equation 1.22 if the fractional errors from the denominator and numerator roughly balance. The residuals from the more numerous cuts far from the waist—the points giving the measurement of divergence Θ, or numerator—would swamp with equal weighting the fewer (or single) cut at the waist—the point(s) giving the divergence of the normalizing gaussian, or denominator. An inverse square weighting approximately halves the influence of the three or four far points, compared to the unity weighting at the waist in the four-cuts method, giving the desired rough balance.

1.7.4 Commercial Instruments and Software Packages

There are three main commercial instruments for measuring beam quality and a host of less well developed others. The first is the original[9,35] system designed as a beam propagation analyzer and believed at this time to be the most fully developed, the ModeMaster™ from Coherent, Inc. The cylindrical scan head (10 cm diameter by 31 cm length) mounts through angle and translation alignment stages to a heavy stable-table post. The basic diameter measurements are achieved with two orthogonal knife-edge cuts. Both principal propagation planes are measured nearly simultaneously on a drum spinning at 10 Hz behind an auxiliary lens. Measurements are restricted to continuous-wave laser beams or high repetition rate (>100 kHz) pulsed lasers. The lens moves to carry the auxiliary beam through the plane of the knife-edges to assemble 260 cuts in each principal propagation plane making up a

pair of propagation plots for the auxiliary beam in a 30-s "focus" pass. A curve fit, with an inverse-diameter weighting, to a hyperbola is done and the fitted parameters are transformed through the lens by the on-board processor to present a data report for the original beam.* The M^2 measurement accuracy is specified at 5% and the waist diameter accuracy at 2%, with a minimum of 100 sample points taken across the profile. The drum also carries two pinholes, each of different diameters, giving pinhole profiles that are processed to give direct second-moment diameters. The instrument also measures beam-pointing stability. Electronic alignment aids are included. In the original instrument operation was controlled through a dedicated electronics console; the current version is driven by a laptop PC.

The second is the ModeScan™ family of instruments from Photon, Inc. Originally intended as an upgrade of a user's existing 10 Hz rotating drum profiler to a beam propagation analyzer, the simplest version is a modular package consisting of a 0.5 m rail to manually translate the profiler in the beam behind a fixed input lens, with software for the user's PC computer that prompts the user for input of position data. When the data fields are filled, the software calculates the M^2 of this auxiliary beam (the same as the input beam's) and transforms the data through the lens for the other input beam constants. Later versions automate the stage drive and data acquisition, expand the software features, and include a new profiler with a selection of five rotation speeds to measure pulsed beams. The latest instrument, the ModeScan™ 1780 model, is a new design in a 26 cm × 18 cm × 8 cm housing gimbal mounted to a 1/2" diameter stable-table post. It incorporates beam splitters behind a fixed input lens to pick off ten sample beams, and direct them to a CCD array camera. The diameters at these ten different throw distances are measured simultaneously and this data fit to a hyperbola to generate the beam constants. Placing ten spots at once on the camera sensor reduces the number of illuminated pixels per spot (a minimum of 15 pixels per diameter measurement is recommended), but enough are lit that the M^2 measurement accuracy of this and the rest of these systems is listed as 4%–5%. This latest model is the only commercial instrument capable of determining beam constants from a single pulse, and therefore of showing pulse-to-pulse beam variations.

The third instrument is the CCD camera-based M^2–200 Beam Propagation Analyzer from Ophir Spiricon, Inc., operating with pulsed or cw laser beams. The original instrument with a 500-mm focal length input lens occupied a 28 cm × 82 cm footprint on the stable table; a new model, the M^2-200s, uses a 300-mm lens and is a half-size version fitting into 26 cm × 44 cm. Here a stepper motor and translation stage on a rail moves an optical delay line behind the fixed input lens to effectively scan the detector surface through the auxiliary beam. The PC computer attached to the system automatically adjusts filter wheel attenuation, subtracts background, sets spot truncation, and calculates the second-moment diameters[37] directly from the CCD profiles. A curve fit to a hyperbola is done[37] with the results transformed back through the input lens to present the constants of the original beam. The M^2 measurement accuracy is given as 5%. This product, with its long focal length input lenses, is positioned to measure the large beams of industrial lasers for process monitoring and control.

1.8 TYPES OF BEAM ASYMMETRY

In the previous sections, the means for the spatial characterization of laser beams were established. This section looks at commonly found beam shapes and others that are

* An example of a data report is shown later in Figure 1.17 of Section 1.9.

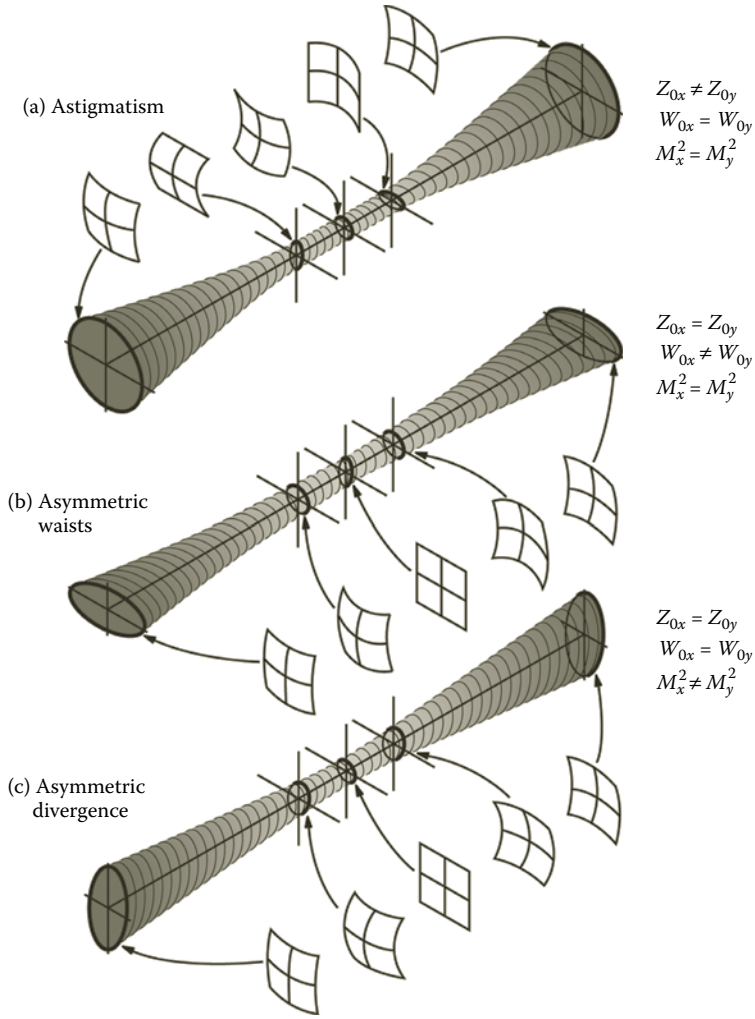


FIGURE 1.15
Depiction of the three-dimensional appearance of the three basic types of asymmetry for a mixed-mode beam: (a) astigmatism, (b) asymmetric waist diameters, and (c) asymmetric divergence. The window insets show the wavefront curvatures along the beam path. (Redrawn from Johnston, T.F., Jr. *Appl. Opt.* 1998, *37,* 4840–4850.)

possible. The three common types of beam asymmetry are depicted in Figure 1.15. These are the pure forms but mixtures of all three are common in real beams.

1.8.1 Common Types of Beam Asymmetry

The first is *simple astigmatism* (Figure 1.15a), where the waist locations for the two orthogonal principal propagation planes do not coincide, $z_{0x} \neq z_{0y}$, but $W_{0x} = W_{0y}$, and $M_x^2 = M_y^2$. Because here the waist diameters and beam qualities are the same for the principal propagation planes, so are the divergences, $\Theta_x \propto M_x^2/W_{0x} = M_y^2/W_{0y} \propto \Theta_y$ [see Equation 1.19]. This makes the beams round in the converging and diverging far-fields. At the two waist planes the beam cross sections are elliptical, one oriented in the vertical and the other the

horizontal plane, with the minor diameters equal. Midway between the waists, the beam becomes round like the "circle of least confusion" point in the treatment of astigmatism[28] in geometrical optics. The simple astigmatic beam is characterized by three round cross sections, at the distant ends and midpoint, with orthogonally oriented elliptical cross sections in between.

The window frame insets of Figure 1.15 show the wavefront curvatures, which are spherical in the far-field, cylindrical at the waist planes with one cylindrical axis horizontal, the other vertical, and saddle-shaped at the midpoint between the waists. The wavefront curvatures determine the nature of the focus when a lens is inserted.

Simple astigmatic beams can be generated in resonators with three spherical mirrors, with one used off-axis to give an internal focus,[38] unless there is astigmatic compensation built in as with a Brewster plate of the correct thickness[38] added to the focusing arm. Many diode lasers are astigmatic but with the other two types of asymmetry as well because the channeling effects in the plane parallel to the junction differ from those in the plane perpendicular to it, giving two different effective source points for the parallel and perpendicular wavefronts. Beams formed using angle-matched second harmonic generation can be astigmatic due to walk-off in the phase matching plane of the beam in the birefringent doubling crystal. The diode lasers in laser pointers frequently have a large astigmatism, as large as the Rayleigh range for the high-divergence axis.

The next is *asymmetric waists* (Figure 1.15b), where the waist diameters are unequal. Because of the different waist diameters but with equal beam qualities, in the far-fields where divergence dominates, the cross sections are elliptical with the long axes of the ellipses (shown as horizontal) perpendicular to the long axis of the waist ellipse (here vertical). In between there are round cross sections at planes symmetrically placed around the waist location—the same geometry as in Figure 1.15a, with the ellipses and circles interchanged. The wavefronts are plane at the waist, and ellipsoidal everywhere else, with curvatures at the round cross sections in the ratio of the square of the waist diameters.

Lasers having an out-of-round gain medium are likely to produce beams with asymmetric waists. A solid-state laser pumped from the end by an elliptical beam from a diode laser is an example. Mode selection is by the combined effects of gain aperturing and absorption in the unpumped regions. The resonant beam shape will mimic the geometry of the pumped region. Beam walk-off from angle-matched nonlinear processes can also produce asymmetric waists.

The third type is *asymmetric divergence* (Figure 1.15) where the beam qualities differ in the principal propagation planes to give proportionally different divergence angles, $\Theta_x \propto M_x^2 \neq M_y^2 \propto \Theta_y$, but $W_{0x} = W_{0y}$, and $z_{0x} = z_{0y}$. The simplest description of this beam is that the mode in one principal propagation plane is of higher order than in the other. In the far-field, cross sections are elliptical as in case (b), but the beam is round only at the waist plane. The wavefronts are plane at the waist and ellipsoidal everywhere else and the Rayleigh ranges are different in the two principal propagation planes.

A CW dye laser using a high-viscosity dye jet provides an example of pure asymmetric divergence.[39] The pump-beam spot was round, but the heat it deposited in the dye stream was cooled differentially by the flow. In the flow direction the temperature gradient was smoothed by the forced convection but in the other direction a more severe thermal gradient existed, causing an aberration that resulted in $M_{4sy}^2 = 1.51$ for that plane compared to $M_{4sx}^2 = 1.06$ for the plane parallel to the flow with negligible aberration. Because of the round pump beam, waist asymmetry was only $2W_{0y}/2W_{0x} = 1.06$.

1.8.2 The Equivalent Cylindrical Beam Concept

Beams with combinations of these asymmetries can be depicted with superposed (x, z) plane and (y, z) plane propagation plots as shown in Figure 1.16a. More generally there is a propagation plot $W(\alpha, z)$ for each azimuth angle α around the propagation axis z. The angle α is measured from the x-axis and $W(\alpha, z)$ lies in the plane containing α and the z-axis. The three-dimensional beam envelope shown in Figure 1.15a, b, or c is called the beam caustic

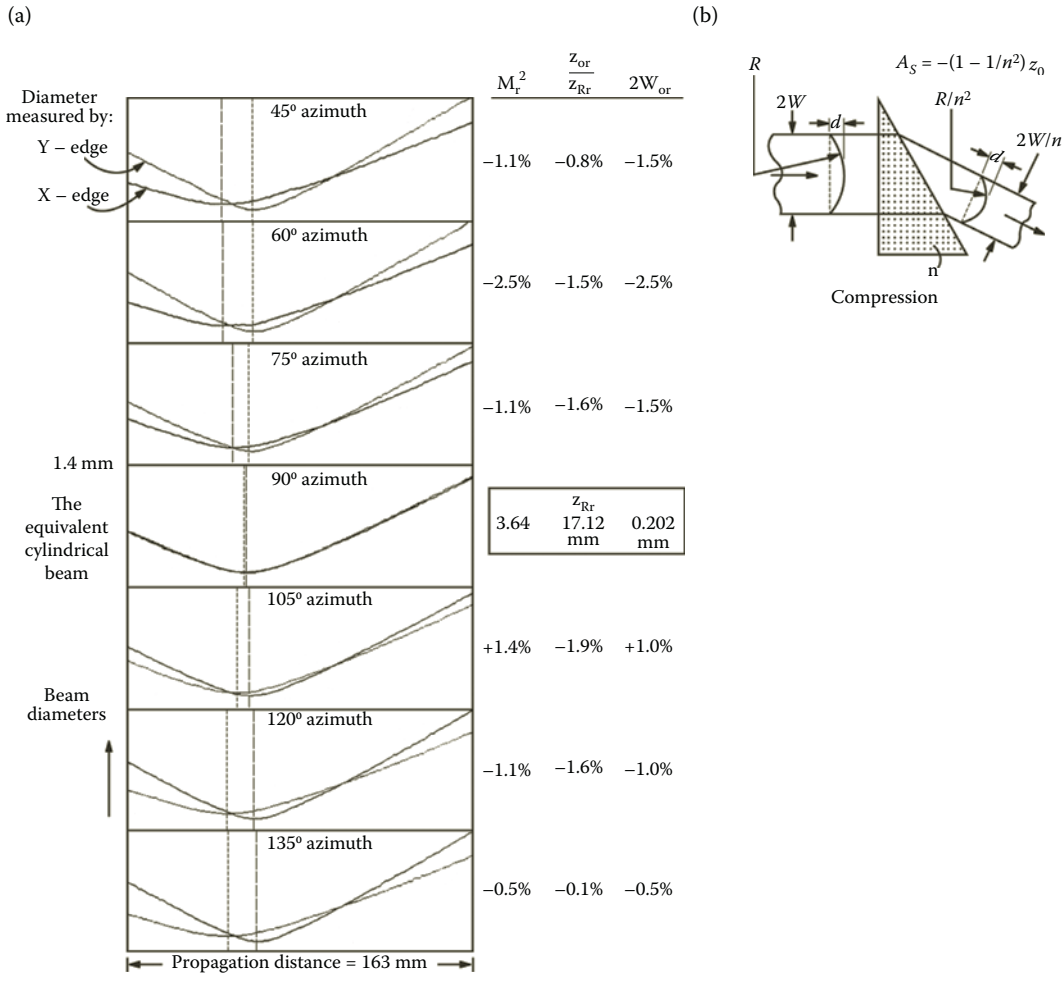


FIGURE 1.16
(a) Experimental propagation plots with beam diameters measured by orthogonal knife-edges for a beam with both astigmatism and waist asymmetry. The percentage variation of the constants of the equivalent cylindrical beam, computed from the plots for each instrument azimuth, is listed in the right-hand columns. The small variations demonstrate the constants are independent of the azimuth of the two orthogonal cutting planes intersecting the beam caustic surface. The constants of the equivalent cylindrical beam, in the box, correspond to the cuts at an instrument azimuth of 90°. (b) Diagram showing how a half-Brewster prism introduces both astigmatism A_s and waist asymmetry $W_{0y}/W_{0x} = n$ to the beam.

surface, and is swept out by $W(\alpha, z)$ as the azimuth angle α rotates through a full circle from 0 to 2π.

For beams with combinations of moderate asymmetries, it is convenient to define an *equivalent cylindrical* beam. This is a beam with cylindrical symmetry—with a round spot for all z—and the real-beam asymmetries are treated as deviations from this round beam. The constants defining this equivalent cylindrical beam are the best average of the beam constants for the two independent principal propagation planes. Many problems can be treated with just this simpler equivalent beam. In particular, it has been predicted theoretically (A.E. Siegman, personal communication, 1990) and demonstrated experimentally[40] that the centered circular aperture computed to give 86% transmission for the equivalent cylindrical beam, has this same transmission for the out-of-round real beam. The minimum aperture sizes for the real beam after propagation in free space can be computed using just the three equivalent cylindrical beam constants. Because the equivalent cylindrical beam is round for all z like the radial modes discussed in Section 1.4, the subscript r is used to denote its constants, and the beam is sometimes called the *equivalent radial mode*.

The equivalent cylindrical beam is best understood by considering the plots of Figure 1.16a. These were measured with the ModeMaster beam propagation analyzer.[9] The profiler built into this instrument uses two knife-edge masks at right angles to each other. They are mounted on a rotating drum at 45° to the scan direction of the drum. This arrangement is equivalent to a vertical and a horizontal knife-edge, each scanned at $1/\sqrt{2}$ times the actual scan speed of the drum, when the analyzer head's azimuth angle is set to 45° to align one knife-edge with the horizontal. Each run to measure beam diameters versus propagation distance produces two propagation plots for the diameters at right angles to the two edges. Normally the analyzer head's azimuth angle is adjusted to record the propagation plots in the two principal planes of the beam. For Figure 1.16a, the analyzer head's azimuth angle was incremented in 15° steps through 90° and new sets of propagation plots recorded for each increment, generating the seven plots shown.

The asymmetric beam of Figure 1.16a was formed by inserting a Brewster-angle half prism[41] in the cylindrically symmetric beam Mode E of Figure 1.10b and Figures 1.2g and h. The prism was oriented as in Figure 1.16b to produce a compression of the beam diameter in one dimension in the (x, z) plane. The prism thus introduces astigmatism and waist asymmetry to the beam. From Figure 1.16b the incoming wavefront of radius of curvature R has a sagitta of the arc, $d = W^2/2R$, which remains unchanged upon the transit of the prism while the beam diameter is compressed. For the Brewster-angle prism, it can be shown[41] that the exiting beam diameter is smaller by the factor $1/n$, where n is the index of refraction of the prism material, here silica with $n = 1.46$. The radius of curvature exiting the prism is thus R/n^2. The M^2 of the beam is unchanged in traversing the prism. From these three conditions both the reduced waist diameter in the x-direction and the astigmatic distance introduced in the exiting beam can be determined [using Equations 1.16b and 1.17b and a little algebra] to be $2W_0/n$ and $A_s = (z_{0y} - z_{0x}) = -(1 - 1/n^2)z_0$, where z_0 is the propagation distance from the input waist location to the prism.

The propagation plots of Figure 1.16a are for the directly measured internal beam, behind the lens of the beam propagation analyzer. These were used because the beam diameter and propagation distance scales of the internal plots remain the same as the instrument azimuth is varied and this facilitates comparison of the plots. Notice in the top plot [45° instrument azimuth], because the internal propagation plots are shown, the axis with the n-times larger divergence and $1/n$-times smaller waist is the y-axis, interchanged with the compressed x-axis of the external beam in the beam-lens transform of Section 1.5.

As the instrument azimuth angle moves around from the initial 45° value (which measures the principal propagation planes for this beam) to 90°, the plots from the two orthogonal edges coalesce into a single "average" curve, then separate with continuing azimuth increments. The plots at 135° are identical to the 45° plots with the *x*-edge and *y*-edge curves interchanged. The dashed and dotted vertical lines on each plot locate the waists for the *x*-edge and *y*-edge curves, respectively. The beam constants for the symmetric, 90°-azimuth plots are those for the equivalent cylindrical beam.

To visualize this process of cutting the beam caustic surface with two orthogonal planes, then rotating the azimuth of the cutting planes, look at Figure 1.15c. The initially vertical (*y*-edge) plane is cutting the caustic in its highest divergence plane, and moves towards a lower divergence $W(\alpha, z)$ plot as the azimuth is incremented. The initially horizontal (*x*-edge) plane is cutting the caustic in its lowest divergence plane, and moves toward a higher divergence $W(\alpha, z)$ plot as the azimuth is incremented. When the cutting planes reach 45° azimuth to the principal planes of the beam, the orthogonal propagation plots match as they would for a round beam with no asymmetries.

Siegman[42] gives the following expressions for the beam constants of the equivalent cylindrical beam in terms of the six constants of the real beam:

$$z_{0r} = \left(\frac{M_x^4 W_{0y}^2}{M_x^4 W_{0y}^2 + M_y^4 W_{0x}^2} \right) z_{0x} + \left(\frac{M_y^4 W_{0x}^2}{M_x^4 W_{0y}^2 + M_y^4 W_{0x}^2} \right) z_{0y} \tag{1.58}$$

$$2W_{0r}^2 = W_{0x}^2 + W_{0y}^2 + \left(\frac{1}{p^2} \right) \left(\frac{M_x^4 M_{0y}^4}{M_x^4 M_{0y}^2 + M_y^4 M_{0x}^2} \right) I^2 (z_{0x} - z_{0y})^2 \tag{1.59}$$

and

$$M_r^4 = \frac{W_{0r}^2}{4} \left(\frac{M_x^4}{W_{0x}^2} + \frac{M_y^4}{W_{0y}^2} \right). \tag{1.60}$$

The columns of numbers in Figure 1.16a demonstrate that the beam constants of the equivalent cylindrical beam are the same when computed from plots for any azimuth, a necessary condition for the equivalent cylindrical beam concept to be useful. The equivalent cylindrical beam quality, waist location, and waist diameter were computed for each azimuth increment from the plots shown, and normalized to the constants measured for the 90° azimuth shown in the box. The percentage errors for these measurements are given in the three columns; the magnitudes of all errors are no larger than 2.5% and are within the instrument measurement tolerances.

From Equations 1.58 and 1.59 for an astigmatic beam the equivalent cylindrical waist lies between the two astigmatic waists, and the square of the cylindrical waist diameter exceeds the sum of the squares of the two astigmatic waist diameters. For a beam with no astigmatism ($z_{0x} = z_{0y}$) the equivalent cylindrical constants become:

$$2W_{0r}^2 = W_{0x}^2 + W_{0y}^2 \tag{1.61}$$

$$M_r^4 = \left(\frac{W_{0x}^2 + W_{0y}^2}{4W_{0x}^2} \right) M_x^4 + \left(\frac{W_{0x}^2 + W_{0y}^2}{4W_{0y}^2} \right) M_y^4. \tag{1.62}$$

For a beam with no astigmatism and no waist asymmetry the equivalent cylindrical beam quality is:

$$M_r^4 = \frac{(M_x^4 + M_y^4)}{2}. \tag{1.63}$$

A beam of this type with different values of M_x^2 and M_y^2 will have a round spot at the waist plane, but not in the far-field as illustrated in Figure 1.15c.

1.8.3 Other Beam Asymmetries: Twisted Beams, General Astigmatism

The shape of a beam caustic surface is determined by the straight-line paths of rays where they emerge at the margin of the particular beam. Such shapes are all examples of ruled surfaces and those depicted in Figure 1.15 are hyperboloids. In principle, any paraxial ensemble of light rays (i.e., a beam) will be enclosed by a ruled surface. Another example is a taut ribbon, and these surfaces can be twisted. Imagine the shapes of Figure 1.15 as taut, flexible membranes. Start with a shape similar to Figure 1.15b except with *all* of the cross sections being horizontally elongated ellipses (a beam with *both* asymmetric waists and divergence). Mentally rotate the far-field ellipses to vertical, the distant one by +90° and the foreground one by −90° azimuth, while keeping the waist ellipse horizontal. In propagating from $z = -\infty$ to $+\infty$ the elliptical cross sections of this beam twist through 180° of azimuth. Such a twisted beam can be physically realized and is said to have *general astigmatism*.[15,16] Here all spots can be ellipses,[15] a waist location is defined by a cross section having a uniform phasefront,[16] and the Rayleigh range is defined as the distance of propagation away from the waist that increments[16] the Gouy phase by $\pi/4$.

Such beams are produced by nonorthogonal[5] optical systems, for example, two astigmatic elements in cascade with azimuth angles that differ by something other than 0° or 90°. Rays in the (x, z) and the (y, z) planes are coupled and cannot be independently analyzed. The general theory for spatial characterization of such beams uses ray matrices weighted by the Wigner density function[4,17] averaged over a four-dimensional geometrical optics "phase space." Rays are described by 4 × 1 column vectors; each vector gives the position x, y and slope u (= θ_x), v (= θ_y) of the ray at the location z along the propagation axis. There are 16 possible second-order moments of these variables; they propagate in free space with a quadratic expansion law.[6,26] The square of the second-moment diameter D_{4s}^2 is such a second-order moment and this is theoretical support this diameter definition enjoys. The beam matrix $\boldsymbol{P}$, the 4 × 4 array of these 16 second-order moments, then fully characterizes the beam with general astigmatism.

The 16 possible second moments can be listed as

$$\langle x^2\rangle;\langle xy\rangle;\langle xu\rangle;\langle xv\rangle; \quad \langle y^2\rangle;'\langle yx\rangle';\langle yu\rangle;\langle yv\rangle;$$

$$\langle u^2\rangle;'\langle ux\rangle';'\langle uy\rangle';\langle uv\rangle; \quad \langle v^2\rangle;'\langle vx\rangle';'\langle vy\rangle';'\langle vu\rangle'.$$

However, by symmetry $\langle xy\rangle = \langle yx\rangle$, and so on; so that only ten of these are independent and in the list those in single quotes are redundant. The moments containing only spatial variables $\langle x^2\rangle$, $\langle xy\rangle$, $\langle y^2\rangle$ can be evaluated as the variances of irradiance pinhole profiles in the proper direction; the $\langle xy\rangle$ profile is at 45° to the x- or y-axes. The moments containing the angular variables cannot be evaluated directly, but are found by inserting optics (usually a cylindrical lens) and measuring downstream irradiance moments at appropriate propagation distances.

From these moments, beam constants are calculated. The first six are the familiar set $2W_{0x}$, $2W_{0y}$, z_{0x}, z_{0y}, M^2_{0x}, M^2_{0y}. The other four address the rate of twist of the phasefront and spot pattern with propagation distance, the generalized radii of curvature of the wavefronts and the orbital angular momentum[43] carried per photon by the beam. Beam classes are defined[44] by values of invariants calculated from the ten second-order moments. A simple association of the resultant shapes for the beam caustic envelopes with each class is *not* immediately available from these definitions.

Twisted-phase beams have been generated by inserting an appropriate computer designed diffractive optical element into an ordinary beam.[45,46] The first report of a beam from an "ordinary" laser (not one deliberately perturbed to produce a twisted phase) that required all ten matrix elements for its characterization, appeared in 2001.[18] Nonorthogonal beams can arise in nonorthogonal resonators,[5] such as in twisted-ring resonators. Until instruments are developed[17] to measure all elements of the beam matrix P and then are applied to characterize beams from many of lasers, the fraction of laser beams with general astigmatism will not be clear. The techniques discussed in previous sections are the methods that can be used together with various auxiliary optics to measure these second-order moments.

1.9 APPLICATIONS OF THE M^2 MODEL TO LASER BEAM SCANNERS

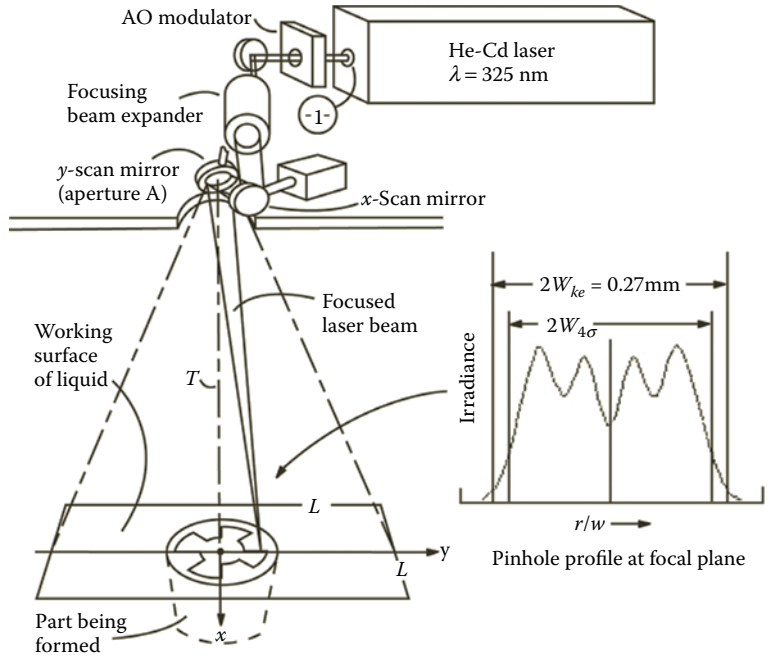
This section applies previous concepts and results to determine appropriate specifications for a laser used in an industrial scanning system, by working backward from the beam properties needed at the work surface. This example shows how parts of the M^2 model interact in the design of a system and how the model can be applied to solve simpler individual problems.

1.9.1 A Stereolithography Scanner

The example analyzes an actual stereolithography scanning system, shown in Figure 1.17. A multimode ultraviolet beam (of 325 nm wavelength) writes on a liquid photopolymer surface under computer control, selectively hardening tiny volume elements of plastic to build up a three-dimensional part. After a 1/4-mm thick slice (cross section) of the part is completed, a jack supporting the growing part inside the vat of liquid lowers the part and brings it back up to 1/4 mm below the surface for the next slice to be written. Parts of great complexity can be formed overnight directly from their CAD-file specifications. This process is called stereolithography and has created what is termed the "rapid-prototyping" industry.

Beam characteristics for the laser are shown in the data report of Figure 1.17. Much of the scanning system design elements used in this analysis are available in the literature.[47–50] The beam from the laser is expanded in an adjustable telescope that also focuses the spot on the liquid surface at the optimum beam spot size[47] for the solidification process, $2W_{02}$ = 0.25 mm ± 10%, measured with a slit profiler.

Notice first that the system geometry defines a maximum M^2 for the laser beam in this application. The rapidly moving y-scan mirror benefits from a low moment of inertia and has a small diameter A. This is the minimum diameter needed to just pass the expanded beam incident on the mirror, making the beam diameter at the mirror $2W_A$ be smaller than A only by some safety factor γ or $2W_A = A/\gamma$. From this mirror, the beam is focused on the



-1-

MODEMASTER (tm) Beam Propagation Analyzer REPORT

Scan Head	s/n D630	Detr–1	silicon s/n D429
Lens–1S singlet:	s/n D650	f=109.16mm	@ 325.0nm
Clip Levels	16.0%/84.0%	(Adjust 2.00)	
Curve Fit	Special	Software Rev.	3.00

DATA

03–03–95 18: 49: 14 Azimuth 45 degrees

	X	Y	R
Msqd	5.33	5.16	5.24
EXTERNAL			
2Wo (mm)	1. 482	1. 460	1. 471
2We (mm)	1. 488	1. 463	1. 476
Zo (m)	–0.089	–0.063	–0.076
Zr (m)	0.993	0.997	0.995
Diverg. (mr)	1.49	1.47	1.48
Astigmatism	(Zoy–Zox) / Zrr = 2.6%		
Waist Asymmetry	2Woy / 2Wox = 0.986		
Divergence Asym	Theta y/Theta x = 0.982		

FIGURE 1.17
A stereolithography scanning system based on a helium–cadmium ultraviolet multimode laser. The pinhole focal plane profile (upper inset) shows the irradiance profile at the surface of the liquid photopolymer. The printout from the commercial beam propagation analyzer (lower inset) applies to the beam at the laser output, location (–1-). Laser data courtesy of CVI Melles Griot, Inc.

liquid surface below at the throw distance T shown in Figure 1.17. The maximum convergence angle of the beam focused on the vat surface is therefore $\Theta_2|_{\max} = A/\gamma T$ (a larger angle would overfill the y-scan mirror). The focused beam waist diameter, given in the previous paragraph is $2W_{02}$; a diffraction-limited beam of that waist diameter—a normalizing gaussian—has a divergence angle of $\theta_n = 2\lambda/\pi W_{02}$. This defines a maximum M^2 for this application by Equation 1.22 of $M^2|_{\max} = \Theta_2|_{\max}/\theta_n = \pi W_{02}A/2\lambda\gamma T$.

This may be evaluated in two different ways. From scaling a photograph of the system,[49] an estimate for T can be made as between 0.6 m and 0.7 m, or a reasonable value is

$T = 0.65$ m. The y-scan mirror diameter A is likely that of a small standard substrate, such as $A \sim 7.75$ mm, and a likely safety factor is about $\gamma \sim 1.5$, yielding $2W_A \sim 5.2$ mm and $M_{\text{slit}}^2 \sim 4.8$. This rough estimate is refined in Section 1.9.5. This beam quality is given in slit units because this is the currency for the focal diameter at the vat; the assumption being made that the value of $2W_A$ used is also in slit units for this estimate.

Alternatively, the beam diameter A/γ can be determined working back from the vat to the y-scan mirror since it is known that the laser of Figure 1.17 is designed for this application and that the measured data (given in Figure 1.17) are within the nominal beam specifications. Those measurements are in knife-edge currency[9] (see Section 1.9.4). Once the knife-edge waist diameter at the vat is found, so is $\theta_n = 2\lambda/\pi W_{02}$ and $2W_A = T\Theta = TM^2\theta_n$, all in knife-edge units. A diameter conversion is thus required to bring the diameter at the waist into knife-edge units for a consistent currency.

1.9.2 Conversion to a Consistent Knife-Edge Currency

By Equation 1.48, for any diameter definition i, the ratio D_i/M_i equals the embedded gaussian diameter $2w$ and therefore the conversion from slit to knife-edge diameters at the vat is just $D_{\text{ke}} = D_{\text{slit}}(M_{\text{ke}}/M_{\text{slit}})$. The square root of the beam quality M_{ke} is known from the report (Figure 1.17), $M_{\text{ke}} = \sqrt{(5.24)} = 2.289$. Here the R or "round beam" column value was used, the equivalent cylindrical beam constants as discussed in Section 1.9.4. To determine M_{slit}, use is made of the expression just above Equation 1.50 relating any M_i to any M_j for different diameter definitions i and j. This conversion formula requires knowledge of the M^2 of the starting currency j; M^2 is known here only for knife-edge units, so j = knife-edge. The desired ending currency is in slit units, i = slit. Then Equation 1.50 gives the required conversion constant, in terms of the conversion constants to second-moment diameters from Table 1.1, as:

$$c_{\text{ke}\to\text{slit}} = c_{ji} = \frac{c_{js}}{c_{is}} = \frac{c_{\text{ke}\to s}}{c_{\text{slit}\to s}} = \frac{(0.813)}{(0.950)} = 0.856.$$

This gives $(M_{\text{slit}} - 1) = 0.856(M_{\text{ke}} - 1) = 1.103$, thus $M_{\text{slit}} = 2.103$ and $M_{\text{slit}}^2 = 4.423$.

Then Equation 1.48 yields the focal diameter at the vat in knife-edge units, $2W_{02\,\text{ke}} = 0.272$ mm, a knife edge to slit diameter ratio of 1.088 for this beam. The "normalizing gaussian" divergence angle above is then evaluated as $\theta_n = 1.521$ mr, the maximum convergence angle is larger than θ_n by $M^2 = 5.24$, making the beam diameter at the y-scan mirror be $2W_A = TM^2\theta_n = 5.180$ mm, all in knife-edge units.

For comparison, using the knife-edge to second-moment conversion constant from Table 1.1 and Equation 1.47 gives the second-moment beam quality and beam diameter at the vat of $M_{4\sigma}^2 = 4.19$ and $2W_{02}|_{4\sigma} = 0.243$ mm. The irradiance profile in Figure 1.17 shows the relative size of the second-moment diameter to the knife-edge diameter. It is evident that the former would require a larger safety factor γ than the latter if used in estimating a safe minimum mirror aperture.

For the remainder of this section, diameters are all in knife-edge units and for simplicity the subscripts indicating this are suppressed.

1.9.3 Why Use a Multimode Laser?

What is the advantage of a multimode laser in this application? First, the critical optic, the scan mirror of diameter A, required for the larger multimode beam diameter is

of reasonable size, so it is possible to use such a larger beam here. The significant advantage is seen from the product data sheet for this laser (CVI-Melles Griot Model 74 Helium–Cadmium laser): with single isotope cadmium used in the laser (the X models on the data sheet) the multimode power is 55 mW, the fundamental-mode power is 13 mW, *a ratio of 4.2 times*. With natural isotopic mix cadmium, the numbers are 40 mW and 8 mW, *a ratio of 5 times*. Thus the laser's output power is roughly proportional to its M^2, making the multimode laser considerably smaller and less expensive than a fundamental-mode laser would be at the power level required for this application.

1.9.4 How to Read the Laser Test Report

Notice that the beam quality number used in Section 1.9.2 was from the "*R*" column (for radial or round mode) of the REPORT shown in Figure 1.17. These are the beam constants for the equivalent cylindrical beam discussed in Section 1.8.2, the best theoretical average[6,40,42] of the X and Y column constants for the two principal propagation planes on the report. Since there is less than 4% difference between M_x^2 and M_y^2, it is appropriate to use the average values in the *R* column and treat the beam as round for this exercise. The fact that the report is all in knife-edge units is signified by the "clip-levels" line reading 16%/84% (adjust: times 2.00) as explained in Reference 9. The EXTERNAL label means these constants are for the original beam external to the instrument, after the lens transform has been done from the constants measured for the INTERNAL auxiliary beam inside. Next, listed for the two principal propagation planes in the X and Y columns, and the equivalent cylindrical beam in the *R* column, are: the external beam waist diameter $2W_0$; the beam diameter $2W_e$ at the instrument's reference surface (the beam entrance plane at the front bezel, designated plane B); the waist location z_0 measured from plane B (the zero point of the beam propagation axis) with negative values being back towards the laser;[9] the Rayleigh range z_R; and the beam divergence. Lastly, the significant beam asymmetry ratios are listed, with the astigmatism normalized to the equivalent cylindrical beam's Rayleigh range.

The whole report is readily converted into a different currency with a diameter a factor of τ larger, if desired, by multiplying the M^2 values by τ^2, the diameters and the divergence by τ, and leaving the z_0, z_R, values and the asymmetry ratios unchanged.

1.9.5 Replacing the Focusing Beam Expander with an Equivalent Lens

The beam expander of Figure 1.17, when left at a fixed focus setting, can be replaced with an equivalent thin lens placed at the *y*-scan mirror location, with the laser moved back a distance z_{01} from the lens, as shown in Figure 1.18b. The propagation over z_{01} expands the beam to match the spot size at this mirror. To find z_{01}, Equation 1.16a is used to find the propagation distance for the required beam expansion of $\rho = 2W_A/2W_{01} = (5.180 \text{ mm})/(1.471 \text{ mm}) = 3.521$. Here $2W_{01}$ is the laser's waist diameter in 1-space, on the input side of the equivalent lens, from the report of Figure 1.17. From Equation 1.16a, $\rho = \sqrt{[1 + (z_{01}/z_{R1})^2]}$, yielding $z_{01}/z_{R1} = \sqrt{[\rho^2 - 1]}$, and with the 1-space Rayleigh range $z_{R1} = 0.995$ m taken from the report, there results $z_{01} = 3.359$ m, also shown in Figure 1.18b.

The equivalent lens focal length $f_{\text{equiv}} \equiv f$ is next properly chosen to bring this beam to a focus at the vat. Since the waist diameters on either side of the equivalent lens are known, by Equation 1.26 the required transformation constant Γ is also known. This then gives by

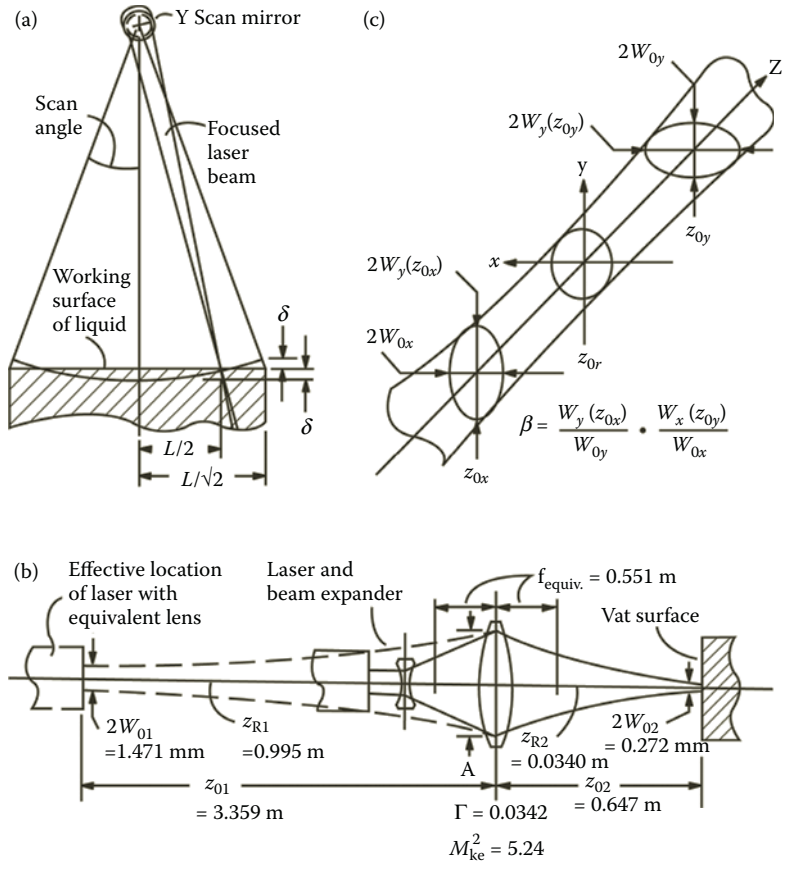


FIGURE 1.18
Analysis of the stereolithography system. (a) Optimum focus to minimize spot-size change over the working surface. For clarity the scan angle shown is larger than the actual scan angle of ±15°. (b) Replacement of the focusing beam expander with an equivalent thin lens of focal length f_{equiv}. Parameters for the equivalent lens transform of the unperturbed beam are shown. (c) Definition of an out-of-roundness parameter $\beta \equiv$ (quadratic ratio of astigmatic diameters) for the focal region of an astigmatic beam. The quantities shown are all for the vat-side focal region or 2-space.

Equation 1.24 a quadratic equation solvable for f:

$$\Gamma = \left(\frac{2W_{02}}{2W_{01}}\right)^2 = \frac{f^2}{[(z_{01} - f)^2 + z_{R1}^2]^2} \tag{1.64}$$

yielding

$$f = z_{01}\left[\frac{\Gamma}{\Gamma - 1}\right]\left\{1 \pm \sqrt{1 - \left(\frac{\Gamma - 1}{\Gamma}\right)\left[1 + \frac{z_{R1}^2}{z_{01}^2}\right]}\right\}. \tag{1.65}$$

Inserting $2W_{02} = 0.272$ mm and $2W_{01} = 1.471$ mm in Equation 1.64 gives $\Gamma = 0.03422$, and this in Equation 1.65 produces $f = f_{equiv} = 0.5511$ m. In what follows, a precise value of z_{02} that corresponds to T in Figure 1.17 is needed and the only value at hand is the previous estimate of $T = 0.65$ m for the y-scan mirror to vat distance. A precise value is needed

consistent with the quantities used in the lens transform z_{01} and f_{equiv} above. This is given by Equation 1.28, now that Γ, z_{01}, and f_{equiv} are known, as $z_{02} = 0.6472$ m. This also shows that the original estimate of T was reasonable. The nominal values for the quantities involved in the equivalent lens transform are shown in Figure 1.18b. The effects of perturbing the nominal values are studied in Section 1.9.7.

1.9.6 Depth of Field and Spot-Size Variation at the Scanned Surface

With the equivalent lens transform defined, questions relating the input beam to the scanned beam can be answered. First, what is the amount of defocus over the scanned field? From Equation 1.27, the Rayleigh range in 2-space at the vat is $z_{R2} = \Gamma\, z_{R1} = 3.404$ cm. The longest radial scan distance is to the corner of the square vat of side length $L = 250$ mm, Figure 1.17, a distance of $\sqrt{2}L/2 = L/\sqrt{2}$. The variation in distance from the y-scan mirror to the corner of the square vat's working surface is $\sqrt{[T^2 + (L/\sqrt{2})^2]} - T = 2.371$ cm, or 0.696 times the vat side Rayleigh range. By Equation 1.16a, from the center of the vat to the far corner, the spot size of the beam will grow by a factor of $\sqrt{[1 + (0.696)^2]} = 1.219$. The simplest way to reduce this range is to focus the beam at the middle of a side edge of the vat, at a radial distance of $L/2$ [see Figures 1.17 and 18a]. This splits the defocus amount 2δ equally among the corners and the middle, to 11% maximum change in spot diameter on the liquid surface over the scanned field. Equivalently, the liquid level could be raised by δ. However, for simplicity the focal distance will be left at T here for the remainder of this analysis.

1.9.7 Laser Specifications to Limit Spot Out-of-Roundness on the Scanned Surface

Next, the inverse lens transform, from the vat side back to the laser side, is applied to transfer scanning beam specifications into laser-beam specifications. From Equation 1.31, for the transform equations going from 2-space to 1-space, use the inverse transformation constant $\Gamma_{21} = 1/\Gamma_{12}$.

Since the lens transformation constant depends on both the input waist location and the Rayleigh range, in general beams without astigmatism but with some other asymmetry when transformed become astigmatic as the results in the remainder of this section show. The plan, starting from the nominal, round, equivalent cylindrical beam of Figure 1.17 transformed to a round beam at the vat, is to perturb the beam at the vat to have a ~10% out-of-round spot. This beam is then transformed back to the laser to see which 1-space variables change and by how much, to account for the perturbation on the scanned side of the lens. The 10% out-of-roundness of the scanned spot is deemed acceptable because that amount of growth in spot diameter over the field was found acceptable above.

The perturbations are made as equal changes of opposite sign in the two independent propagation planes. For example, 10% out-of-roundness due to waist asymmetry is accomplished with a +5% change in W_{02y} and a −5% change in W_{02x}. The resulting changes in the 1-space beam constants are not completely symmetrical, due to the nonlinearity of the beam-lens transform. The effect of perturbing a constant in only one principal propagation plane is given directly in the tables by the percentage changes, the column in parentheses, for 1-space shown for that plane. Because the propagation planes are independent, so are the percentage changes in each plane.

1.9.7.1 Case A: 10% Waist Asymmetry

Assume $2W_{02x}$ is reduced 5% (to 0.259 mm), and $2W_{02y}$ is raised 5% (to 0.286 mm) to give a *waist asymmetry* different from unity by 10%. To calculate the effect on the input beam, first

the new Rayleigh ranges for the beam at the vat are found as z_{R2x} = 3.088 cm (reduced 10%) and z_{R2y} = 3.753 cm (increased 10%). For each of these a new $1/\Gamma$ for the inverse transform is computed from Equation 1.24, followed by the remaining constants through Equations 1.26 through 1.28. The results for the 1-space beam constants, and their percentage change shown in parentheses are summarized in Table 1.2. The initial value of $1/\Gamma$ is 29.2259. In the table A_s/z_{Rr} stands for the normalized astigmatism $A_s/z_{Rr} = (z_{01y} - z_{02x})/z_{R1r}$, where z_{R1r} is the Rayleigh range in 1-space of the equivalent cylindrical beam.

The +10% pure waist asymmetry at the vat (i.e., accompanied by no astigmatism or divergence asymmetry) for the most part transforms through the lens to a corresponding +8% waist asymmetry at the laser. The same is true for the divergence asymmetry. The different waist diameters at the vat, generate different Rayleigh ranges there and in the lens transform produce a –12% normalized astigmatism at the laser. Specify the laser to have less than these asymmetries to keep the scanned beam out-of-roundness below 10%.

1.9.7.2 Case B: 10% Divergence Asymmetry

Here it is assumed M_x^2 is reduced 5% and M_y^2 is increased by 5% to give a +10% change in the 2-space *divergence asymmetry* without changing the waist asymmetry $W_{02y} / W_{02x} = 1$. By Equation 1.18 or 1.19 the Rayleigh ranges on the vat side of the lens change inversely with their M^2 to make them z_{R2x} = 3.088 cm, z_{R2y} = 3.753 cm. Applying Equations 1.24 through 1.28 to each principal plane produces Table 1.3, the results for the 1-space beam constants and their percentage change.

The divergence asymmetry of the beam at the vat carries through the lens to the laser, and implies some astigmatism is necessary at the laser (but half as much as Case A) to get pure divergence asymmetry at the vat.

TABLE 1.2

Laser Constants Corresponding to *10% Waist Asymmetry* at the Scanned Surface

Quantity	*x*	*y*	Ratios, *y*/*x*	Ratio was:
$1/\Gamma$	29.816 (+2.0 %)	28.540 (–2.4%)	0.957 (–4.3%)	1
$2W_{01}$ (mm)	1.415 (–3.8%)	1.526 (+3.8%)	1.079 (+7.9%)	1
z_{01} (m)	3.416 (+1.7%)	3.294 (–2.0%)	A_s/z_{Rr} = –12.3%	0
z_{R1} (m)	0.921 (–7.5%)	1.071 (+7.7%)	1.163 (+16.4%)	1
Θ_1 (mr)	1.537 (+3.8%)	1.425 (–3.7%)	0.927 (–7.3%)	1

TABLE 1.3

Laser Constants Corresponding to *10% Divergence Asymmetry* at the Scanned Surface

Quantity	*x*	*y*	Ratios, *y*/*x*	Ratio was:
$1/\Gamma$	28.896 (–1.1%)	29.532 (+1.0%)	1.022 (+2.2%)	1
$2W_{01}$ (mm)	1.463 (–0.6%)	1.478 (+0.5%)	1.011 (+1.1%)	1
z_{01} (m)	3.328 (–0.9%)	3.389 (+0.9%)	A_s/z_{Rr} = +6.1%	0
z_{R1} (m)	1.033 (+3.8%)	0.958 (–3.8%)	0.927 (–7.3%)	1
Θ_1 (mr)	1.416 (–4.3%)	1.544 (+4.3%)	1.091 (+9.1%)	1

1.9.7.3 Case C: 12% Out-of-Roundness across the Scanned Surface Due to Astigmatism

A little discussion is required to define an out-of-roundness parameter for the focal region of an astigmatic beam in general before applying the concept to the focus at the vat. It has already been shown (Section 1.9.6) that the path length to the liquid surface changes over the scanned field by 2.37 cm. This path change causes a spot-size variation of 21.9% if the spot is focused at the center of the vat, and 11% if focused at the middle of a vat edge to reduce the variation across the scanned surface. On top of this, there is an out-of-round change in the spot, if the beam is astigmatic. The fastest change of shape with z of the elliptical spots in a beam with pure astigmatism, Figure 1.16a, takes place between the two astigmatic waists in the focal region, which is the working region for the beam in the stereolithography system. Suppose the astigmatic distance $z_{02y} - z_{02x}$ is matched to the path length change of 2.37 cm but the edge focus is used to split this distance (see Figure 1.18a). This makes the largest path between an astigmatic waist and the liquid working surface anywhere in the field be 1.19 cm. Then from Equation 1.16a and Figure 1.18c

$$\frac{W_{2x}(z_{02r})}{W_{02x}} = \sqrt{1 + \left(\frac{1.19}{3.40}\right)^2} = 1.059$$

where z_{02r} is the equivalent cylindrical beam waist location halfway between the x and y waist locations. The spot at the vat only goes to 5.9% out-of-round, but the orientation of the out-of-round ellipse is along the y-axis in the corners where the liquid is below z_{02r} and along the x-axis in the middle of the field where the liquid is above z_{02r}. This can have an unpleasant effect on the part, because the textures of the x- and y-formed surfaces differ. Therefore, an adequate out-of-round parameter for the focal region of an astigmatic beam can be defined (Figure 1.18c) as $\beta \equiv$ (quadratic ratio of astigmatic diameters), where the product of the x-direction and y-direction out-of-round diameter ratios is

$$b = \left[\frac{W_{2y}(z_{02x})}{W_{02y}}\right]\left[\frac{W_{2x}(z_{02y})}{W_{02x}}\right]. \tag{1.66}$$

The ratios are evaluated at the two astigmatic waist locations as indicated in Equation 1.66. From the above, $\beta = (1.059)^2 = 1.12$ for an astigmatic distance equal to the scanned depth of field; this is taken here as "12% out-of-roundness due to astigmatism" for the final example.

The calculations proceed in this example with $z_{02x} = (0.6472 - 0.0119)$ m = 0.6353 m and $z_{02y} = (0.6472 + 0.0119)$ m = 0.6591 m, with the other 2-space beam parameters left at their unperturbed values of Figure 1.18b. Table 1.4 gives the results for the 1-space beam.

TABLE 1.4

Laser Constants Corresponding to a *12% Out-of-Roundness* due to Astigmatism ($\beta = 1.121$) across the Scanned Surface

Quantity	x	y	Ratios, y/x	Ratio was:
$1/\Gamma$	36.794(+25.9%)	23.708 (−18.9%)	0.644 (−35.6%)	1
$2W_{01}$ (mm)	1.651 (+12.2%)	1.345 (−9.9%)	0.803 (−19.7%)	1
z_{01} (m)	3.651 (+8.7%)	3.110 (−7.4%)	$A_s/z_{Rr} = -52.4\%$	0
z_{R1} (m)	1.253 (+25.9%)	0.807 (−18.9%)	0.644 (−35.6%)	1
Θ_1 (mr)	1.318 (−11.0%)	1.641 (+10.9%)	1.216 (+24.6%)	1

This type of asymmetry, transformed back to the laser side of the equivalent lens, is devastating to the 1-space beam constants. More correctly, it would take devastating input beam characteristics to produce this large a "quadratic-ratio-of-astigmatic-diameters" parameter. There are large percentage changes in 1-space waist asymmetry, astigmatism, and divergence asymmetry. Actual lasers with asymmetries this large would be rejected by the laser manufacturer, and the scanner manufacturer would not have to deal with them. Lasers with sufficient beam asymmetry to give $\beta = 1.12$ at the scanned surface would not make it into the field.

In conclusion, the strictest specifications found for the laser to meet upon incoming testing were from Case A, yielding 10% waist asymmetry at the vat surface. To stay below this out-of-roundness at the vat, the analysis gave bounds at the laser of less than 12% normalized astigmatism and less than 8% waist and divergence asymmetry. These values were easily met by the laser tested and reported in Figure 1.17. In an actual situation of setting laser specifications, several more examples should be run, including cases starting on the laser side and calculating the asymmetries that result in the scanning beam. Readers journeying this far into this applications section should now have sufficient analytical tools provided by the M^2 model to complete those calculations themselves.

1.10 CONCLUSION: OVERVIEW OF THE M^2 MODEL

The M^2 model provides description of real beams by generalizing the equations for the ideal fundamental-mode beam. Any real beam, whether made up of modes from a stable laser resonator or not, is larger in diameter by the factor M—for all propagation distances z—than the embedded gaussian beam implicit within it. Thus the change in equations takes the form of replacing the $1/e^2$ radius w of the embedded gaussian beam by W/M, where W is the radius of the real beam. This replacement generalizes both the beam propagation and beam-lens transform equations.

The real beam, with waist diameter $2W_0$ being larger than the embedded gaussian by the factor M for all z, diverges at an M times larger rate. All diffraction-limited beams have a gaussian irradiance profile, and one of waist diameter $2W_0$, being M times larger than the embedded gaussian diverges at a rate $1/M$ as fast as the embedded gaussian. Hence the real beam divergence is M^2 times larger than the diffraction limit. This identifies M^2 as a beam invariant unchanging in free space propagation or transmission through lenses, and as a measure of the beam quality. An M^2 of unity is the highest quality, a diffraction-limited beam, and real beams with larger values have increasing degrees of higher-order mode content and wavefront aberration (and hence are also called mixed-mode beams).

To apply this analytical description of a mixed-mode beam, its M^2 must first be measured, and here the simplicity of the ideas becomes more complex. The measurement requires finding the scale length for expansion of the beam diameter with propagation, z_R, the Rayleigh range. Several diameters at well-chosen z locations on both sides of the waist are determined and this data is fit to the correct hyperbolic form. The fit gives three beam constants—the beam quality, the waist diameter, and the waist location—for each independent and orthogonal principal propagation plane.

The first additional complexity is that different definitions give different numerical values for the diameters for the mixed mode and the higher-order modes it contains. Beam diameters are still measured from beam irradiance profiles, but different profiling masks

(pinholes, slits, knife-edges, or centered circular apertures) all give different shaped profiles for higher-order modes and hence different diameters. Care is required to keep track of which measurement "currency" is in use and to not mix different currencies. A standard diameter definition—the second-moment diameter, four times the standard deviation of a pinhole irradiance distribution of the beam—has been adopted by the ISO. However, this diameter is computation intensive and difficult to measure reliably because of sensitivity to noise on the wings of the profile. Therefore, conversion rules have been developed applicable to cylindrically symmetric mixed-mode beams permitting measurements done in one diameter currency to be changed into another. The basis of these rules is our observation that higher-order modes turn on and off in a characteristic sequence as the diameter of the circular limiting aperture in the resonator is opened. This associates with the increasing second-moment M^2 a unique set of mode fractions, allowing accurate conversion rules to be derived.

The second additional complexity is that the diameter measurements and curve fits done to determine M^2 may give unreliable answers unless several pitfalls in the process are avoided. Chief among these is that the mixed-mode waist must be accurately located and its diameter physically measured and not inferred or assumed. Since the waists of most lasers are buried inside the resonator, this requires the use of an auxiliary lens to form an auxiliary beam with an accessible waist. The constants determined for this auxiliary beam then are transformed back to those for the original beam by means of the beam-lens transform equations. Commercial instruments that do this automatically are available.

Beams with pure forms of the classic asymmetries have been illustrated, those with only astigmatism, or waist asymmetry, or divergence asymmetry. Beams with combinations of asymmetries may be represented by pairs of propagation plots, one for each principal propagation plane. Beam asymmetries can also be interpreted as deviations from a theoretical "best weighted average" round beam, the equivalent cylindrical beam. There are also beams not directly covered in the M^2 model whose principal propagation planes twist in space like a twisted ribbon—beams with "general astigmatism."

Lastly, the M^2 model was demonstrated by analyzing an actual laser-beam scanning system used in stereolithography. Asymmetries causing out-of-round spots on the scanned surface were analytically projected back through the delivery system to determine the size of the corresponding asymmetries at the laser source.

There are many applications of the M^2 model. It quantifies the mode specifications for commercial lasers, with the means to test to these specifications. Its use permits design of multimode lasers and their beam delivery systems. The beam transformations occurring in nonlinear optics can be analyzed. The divergence of a high M^2 laser beam can be matched into the acceptance angle a high numerical aperture fiber, to take advantage of the lower cost per unit of output power of a multimode laser. These are just a few of many other applications, all with the backup of commercial instrumentation to make the beam measurement process easy and efficient. This chapter has provided the analytical tools to make these applications realities.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Emeritus A. E. Siegman, Stanford University, for many years of enlightening interactions on this subject; and thank for helpful

discussions: Gerald F. Marshall, the original editor of this book, who always has another intriguing question; and G. Nemes of Astigmat, who taught us about beams with general astigmatism. In this revised edition, Jeff Guttman of Photon, Inc. updated us on recent developments in cameras and profilers. Lastly, David Bacher and John O'Shaughnessy of CVI Melles Griot, Inc., and especially Gerald F. Marshall contributed very helpful and constructive reviews of the manuscript.

GLOSSARY

Astigmatism, general: The property of beams having elliptical cross sections for all z, with the principal axes of the ellipses rotating with propagation along the beam axis (nonorthogonal beams; "twisted" beams).

Astigmatism, normalized: The difference in waist locations for the two independent principal propagation planes divided by the Rayleigh range of the equivalent cylindrical beam, $A_s/z_{Rr} = (z_{0y} - z_{0x})/z_{Rr}$, usually expressed in percent.

Astigmatism, simple: Having different waist locations in the two principal propagation planes, $z_{0x} \neq z_{0y}$.

Asymmetric divergence: Having different divergence angles $\Theta_x \neq \Theta_y$ in the two principal propagation planes.

Asymmetric waists: Having different waist diameters in the two principal propagation planes, $2W_{0x} \neq 2W_{0y}$.

Beam caustic surface: The envelope of the beam swept out by rotating the curve of the beam radius $W(z)$ versus propagation distance about the propagation axis z. When a plane containing the z-axis and at an angle α to the x-axis cuts the caustic surface the intersection gives the propagation plot for azimuth α. See the discussion of Figure 1.16a.

Beam, equivalent cylindrical: A cylindrically symmetric beam constructed mathematically in the M^2 model from the beam constants measured in the two principal propagation planes of an asymmetric beam, see the explanation of Figure 1.16a. The propagation plot for the equivalent cylindrical beam is obtained from the beams of Figure 1.15 by slicing their caustic surfaces along the z-axis at a 45° inclination to the x- or y-axes. This is the best cylindrically symmetric average beam for a beam with asymmetry. The subscript r is used to denote the constants for this beam, for round or radial symmetry.

Beam, gaussian: A uniphase beam with spherical wavefronts whose transverse irradiance profiles everywhere have the form of a gaussian function. Such an idealized beam is diffraction limited with $M^2 = 1$, a condition that can only be approached by real beams.

Beam, idealized: The abstract mathematical description of a beam (which can have $M^2 = 1$).

Beam propagation analyzer: An instrument that measures beam diameters as a function of propagation distance, displays the $2W(z)$ versus z propagation plot, and curve fits this data to a hyperbola to determine beam quality M^2, waist location z_0, and waist diameter $2W_0$.

Beam propagation constant M^2: So called because replacing the fundamental-mode radius $w(z)$ in its propagation equation by $w(z) = W(z)/M$ predicts the propagation of the mixed mode, of radius $W(z)$.

Beam quality: The quantity M^2, so called because a real beam has M^2 times the divergence of a diffraction-limited beam of the same waist diameter; see also "normalizing gaussian."

Beam, real: An actual beam; all have at least slight imperfections and thus an M^2 greater than one.

Clip width: Distance (on the mask translation axis) between the points on an irradiance profile at a specified fraction (such as 13.5%) of the height of the highest peak.

Conversions, beam diameter: Empirical rules derived for beams of cylindrical symmetry, to convert diameters measured by one method to those measured by another, such as slit diameters to knife-edge diameters.

Convolution error: Contribution to the measured diameter from the finite size of the scanning aperture; minimizing this is an important consideration for pinhole and slit measurements.

Cut: A beam diameter measurement, from the cutting action of a profiler's scanning aperture.

Diameter, $1/e^2$: Beam diameter defined by the aperture translation distance between clip points on an irradiance profile at a height of 13.5% = $1/e^2$ relative to the highest peak at 100%.

Diffractive overlay: Interference from high angle rays overlapping the beam, diffracted from the limiting aperture in the resonator. This can distort profiles taken close to the laser output coupler (within a Rayleigh range).

Eigenfunctions: A set of functions f_n associated with a linear operator Q satisfying $Qf_n = c_n f_n$, where the c_n are scalar constants (the eigenvalues). Because of this self-replicating property these functions occur in many physical problems, for example, the laser mode functions that also describe the harmonic oscillator and hydrogen wave functions in quantum mechanics.

Embedded gaussian: The fundamental mode of the resonator that generates a mixed-mode beam. The mixed-mode beam diameter is M times larger than the embedded gaussian beam diameter at all propagation distances z.

Far-field: The propagation region(s) of a beam many Rayleigh ranges away from the waist locations. In the far-field, the transverse extent of the beam grows linearly with increasing distance from the waist.

Four-cuts method: The simplest method for determining M^2, requiring only four well-chosen diameter measurements both straddling the waist location and at the waist location.

Fresnel number: The square of the radius of the limiting aperture in a resonator, divided by the mirror separation and the wavelength. As the aperture is opened and this number increases, modes of higher order oscillate and join the mix of modes.

Gaussian: A mathematical function of the form $\exp(-x^2)$; see also "beam, gaussian."

Hermite–Gaussian function: An eigenfunction of the wave equation including diffraction, which describes beams of rectangular symmetry, of the form of a gaussian function times a pair of Hermite polynomials of orders (m, n).

Invariant, beam: A quantity that is unchanged by propagation in free space or transmission through ordinary, nonaberrating, optical elements (lenses, Brewster windows, etc.).

Irradiance: The power per unit cross-sectional area of the beam.

Laguerre–Gaussian function: An eigenfunction of the wave equation including diffraction, that describes beams of cylindrical symmetry of the form of a gaussian function times a generalized Laguerre polynomial of order (p, l).

M^2: The ratio of the waist diameter-divergence angle product of the mixed-mode beam, to that for the embedded gaussian of that beam. A beam invariant, this is also called the "times-diffraction-limit" number, the beam quality, and the beam propagation factor.

Mode: The characteristic frequencies and transverse irradiance patterns of beams formed in laser oscillators, described by Hermite–Gaussian and Laguerre–Gaussian functions, denoted by the symbols $TEM_{m,n}$, $TEM_{p,l}$ with m, n or p, l the order numbers of the function's polynomials.

Mode, degenerate: Two modes with the same optical frequency, and therefore, order numbers.

Mode, donut: A starred mode, TEM^*_{01}, with the second-lowest diffraction loss through a circular limiting aperture, and an irradiance profile with a hole (null) in the center (see Figure 1.1).

Mode, fundamental: The TEM_{00} mode, with a gaussian irradiance distribution, a single-spot peaked profile, the lowest mode order and smallest beam diameter from a given resonator, and with $M^2 = 1$ in the limit of perfection. Thus this mode has the lowest diffraction loss through a centered circular limiting aperture.

Mode, higher order: Any mode of order number greater than that of the fundamental mode.

Mode, longitudinal: A mode of frequency $q(c/2L)$, where c is the speed of light and q is a large integer equal to the number of beam wavelengths that fit in the round trip path $2L$ of the resonator. The $(q + 1)^{th}$ longitudinal mode has a frequency $(c/2L)$ higher than the q^{th}; each longitudinal mode is associated with a given transverse mode.

Mode, lowest order: The fundamental mode, of order number one.

Mode, mixed: An incoherent superposition of pure modes, all from the same resonator, with a diameter $2W$ that is M times larger for all z than $2w$, the fundamental-mode diameter from the set. Also called a real beam as only idealized beams have $M^2 = 1$ (indicating zero higher-order mode content).

Mode order number: For Hermite–Gaussian modes, $(m + n + 1)$; for Laguerre–Gaussian modes, $(2p + l + 1)$; the order numbers determine the mode frequencies and phase shifts, and give the mode's beam quality M^2_{4s} measured in second-moment units.

Mode or spot pattern: The two-dimensional pattern of the irradiance distribution as would be viewed on a flat surface inserted normally in the beam.

Mode, pure: Any transverse mode that is *not* a mixture of modes of different orders.

Mode, starred: A circularly symmetric mode that is a composite of two degenerate modes combined in space and phase quadrature, that is, superposed with a copy of itself after a 90° rotation (see Figure 1.1).

Mode, transverse: A mode, designated by the symbols $TEM_{m,n}$, $TEM_{p,l}$, whose transverse irradiance distribution is described by the Hermite–Gaussian or Laguerre–Gaussian functions of m, n or p, l order numbers.

Near-field: The beam propagation region(s) within a Rayleigh range from the waist location.

Noise-clip option: A test of the sensitivity to noise of the second-moment diameter computed from a pinhole profile, consisting of discarding any profile data with negative values after subtraction of the background to see the change this makes in the computed diameter.

Normalizing Gaussian: A diffraction-limited, idealized gaussian beam of the same waist diameter as a mixed-mode real beam, whose divergence is used as the denominator in a ratio with the real beam's divergence to compute the real beam's M^2.

Paraxial: Meaning close to the beam axis, this refers to a ray (or bundle of rays) propagating at an angle small enough with respect to the central axis that this angle and its tangent are essentially equal.

Power-in-the-bucket: Alternate term for D_{86}, the variable-aperture beam diameter definition.

Principal diameters (of an elliptical spot): The diameters along the major and minor axes of the ellipse.

Principal propagation planes, independent: The two perpendicular planes containing the major and minor axes of an elliptical beam spot (x- and y-axes) and the propagation axis (z). In the M^2 model the three propagation constants for each of these two planes are independent.

Profile: The record of transmitted power versus translation distance of a small aperture or other mask scanned across the beam.

Profile, knife-edge: A profile taken with a knife edge mask, yielding a tilted S-shaped curve.

Profile, pinhole: A profile taken with a pinhole aperture and capable of showing all the irradiance highs and lows but requiring careful centering of the beam to the scanned track of the pinhole. Signal-to-noise ratio and convolution error are inversely dependent on the pinhole diameter making the hole diameter an important consideration.

Profiler: An instrument for measuring beam diameters, that scans a mask (pinhole, slit, or knife-edge) through the beam, displays the resulting profile, and (usually) reports the beam diameter on a digital readout as the scan distance—or clip width—between preset clip points on the profile.

Profile, slit: A profile taken with a slit aperture, showing something of the irradiance highs and lows and *not* requiring centering of the beam to the scanned track. Signal-to-noise ratio and convolution error are counterdependent on the slit width, making it an important consideration.

Propagation constants: The set of parameters: waist diameter $2W_0$, waist location z_0, and beam quality M^2, in each of the two principal propagation planes that define how the transverse extent of a beam changes as it propagates.

Propagation plot: The plot of beam diameter versus propagation distance, $2W(z)$ versus z. For the beams covered in the M^2 model, the form of this plot is a hyperbola.

Rayleigh range: In the M^2 model, the propagation distance z_R from the waist location to where the wavefront reaches maximum curvature, also the distance from the waist to where the beam diameter has increased by $\sqrt{2}$, and the scale length for beam expansion with propagation, $z_R = \pi W_0^2/M^2\lambda$.

Resonator: The aligned set of mirrors providing light feedback in a closed path through the gain medium in a laser. Since the wavefront curvatures and surface curvatures must match at the mirrors, the resonator determines the mode properties of the beam.

Scan: Movement of a mask or aperture transversely across a beam while recording the transmitted power; see "cut."

Second-moment diameter: $D_{4\sigma}$, equal to four times the standard deviation, σ, of the transverse irradiance distribution obtained from a pinhole profile.

Second-moment, linear: The integral over the transverse plane of the square of the linear coordinate times the irradiance distribution, for example, $\langle x^2 \rangle$, used in calculating the variance of the distribution $\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$.

Second-moment, radial: The integral over the transverse plane of the irradiance distribution times the square of the radial coordinate measured outwardly from the

centroid of the spot, for example $\langle r^2 \rangle$, used in calculating the variance of the distribution $\sigma_r^2 = \langle r^2 \rangle$. In the integration the distribution is weighted by r^3 since the area element is $dA = rdr\, d\theta$.

Spot: The two-dimensional irradiance distribution or cross section of a beam as seen on a flat surface normal to the beam axis.

Stigmatic: Describes a beam that maintains a round cross section as it propagates, or more formally, a beam that maintains a rotationally symmetric irradiance distribution in free space. (The opposite term is *astigmatic*, where cross sections are elliptical at some propagation distances z.)

TDL, times-diffraction-limit number: The number of times the divergence of a real beam is larger than that of a diffraction-limited beam (called the normalizing gaussian) of the same waist diameter; TDL = $\Theta/\theta_n = M^2$. Also the factor by which a real-beam waist diameter is larger than that for a gaussian beam ($M^2 = 1$) converging at the same numerical aperture (NA).

***TEM*$_{mn}$:** (For Transverse ElectroMagnetic wave). A symbol used to designate a transverse mode of rectangular symmetry described by a Hermite–Gaussian function with polynomial orders m, n.

***TEM*$_{pl}$:** (For Transverse ElectroMagnetic wave). A symbol used to designate a transverse mode of cylindrical symmetry described by a Laguerre–Gaussian function with polynomial orders p, l.

Thresholding: A method for noise reduction in the readout of a CCD camera frame by measuring the noise level in nonilluminated portions of the frame (such as the corners) from which a standard deviation σ is calculated, then subtracting a uniform noise floor level (typically of 3σ amplitude) from the entire frame before processing the signal.

Variable-aperture diameter: D_{86}, (or D_{xx}) the diameter of a centered circular aperture passing 86.5% (or xx%) of the total power in the beam.

Waist, beam: The location on the beam propagation plot where the beam diameter is at a minimum; also used for the value of this minimum diameter.

Waist diameters: $2W_{0x}$, $2W_{0y}$, the minimum diameters in each principal propagation plane.

Waist locations: z_{0x}, z_{0y}, the points along the propagation axis where the minimum (waist) diameter(s) of the beam in each of the independent principal propagation planes are located.

Wave equation: Propagation of paraxial rays including the effect of diffraction are described by either the Fresnel–Kirchhoff diffraction integral equation of Boyd and Gordon[2] or the simple scalar wave equation used by Kogelnik and Li;[1] both have the Hermite–Gaussian and Laguerre–Gaussian functions as eigenfunction solutions.

REFERENCES

1. Kogelnik, H.; Li, T. Laser beams and resonators. *Applied Optics* 1966, *5*, 1550–1567.
2. Boyd, G.D.; Gordon, J.P. Confocal multimode resonator for millimeter through optical wavelength masers. *Bell System Technical Journal* 1961, *40*, 489–508.
3. Marshall, L. Applications à la mode. *Laser Focus* 1971, *7*(4), 26–29.

4. Bastiaans, M.J. Wigner distribution function and its application to first-order optics. *Journal of the Optical Society of America* 1979, *69,* 1710–1716.
5. Siegman, A.E. *Lasers;* University Science Books: Sausalito, CA, 1986; ISBN 0-935702-11-3.
6. Siegman, A.E. New developments in laser resonators. *Proceedings of SPIE* 1990, *1224*, 2–14.
7. Sasnett, M.W. Propagation of multimode laser beams—The M² factor. In *The Physics and Technology of Laser Resonators;* Hall, D.R., Jackson, P.E., Eds.; Adam Hilger: New York, 1989; Chapter 9, ISBN 0–85274-117–0.
8. Johnston, T.F., Jr.; Fleischer, J.M. Calibration standard for laser beam profilers: method for absolute accuracy measurement with a Fresnel diffraction test pattern. *Applied Optics* 1996, *35,* 1719–1734.
9. The Coherent, Inc., ModeMaster™. The manual for the PC version of this instrument containing much useful information is available upon request or on their website from Coherent Laser Measurement and Control, 27650 SW 95th Avenue, Wilsonville, OR 97070.
10. Johnston, T.F., Jr. M² concept characterizes beam quality. *Laser Focus* 1990, *26*(5), 173–183.
11. Test methods for laser beam widths, divergence angles, and beam propagation ratios, ISO/FDIS 11146:2004 in three parts: -1, Stigmatic and simple astigmatic beams; -2, General astigmatic beams; -3 Intrinsic and geometrical laser beam classification, propagation and details of test methods; available from Deutsches Institut fur Normung, Pforzheim, Germany.
12. Lawrence, G.N. Proposed international standard for laser-beam quality falls short. *Laser Focus World* 1994, *30*(7), 109–114.
13. Sasnett, M. et al. Toward an ISO beam geometry standard. *Laser Focus World* 1994, *30*(9), 53.
14. Johnston, T.F., Jr.; Sasnett, M.W.; Austin, L.W. Measurement of "standard" beam diameters. In *Laser Beam Characterization;* Mejias, P.M., Weber, H., Martinez-Herrero, R., Gonzales-Urena, A., Eds.; SEDO: Madrid, 1993; 111–121.
15. Arnaud, J. A.; Kogelnik, H. Gaussian light beams with general astigmatism. *Applied Optics* 1969, *8,* 1687–1693.
16. Mansuripur, M. Gaussian beam optics. *Optics and Photonics News* 2001, *12*(1), 44–47.
17. Nemes, G.; Siegman, A.E. Measurement of all ten second-order moments of an astigmatic beam by the use of rotating simple astigmatic (anamorphic) optics. *Journal of the Optical Society of America* 1994, *11,* 2257–2264.
18. Serna, J.; Encinas-Sanz, F.; Nemes, G. Complete spatial characterization of a pulsed doughnut-type beam by use of spherical optics and a cylinder lens. *Journal of the Optical Society of America* 2001, *18,* 1726–1733.
19. Silfvast, W.T. *Laser Fundamentals;* Cambridge University Press: New York, 1996; Chapter 10, ISBN 0-521-55617-1.
20. Rigrod, W.W. Isolation of axi-symmetric optical resonator modes. *Applied Physics Letters* 1963, *2*, 51–53.
21. McCumber, D.E. Eigenmodes of a symmetric cylindrical confocal laser resonator and their perturbation by output-coupling apertures. *Bell System Technical Journal* 1965, *44,* 333–363.
22. Koechner, W. *Solid–State Laser Engineering,* 5th Ed.; Springer-Verlag: New York, 1999; Figure 5.10.
23. Wolfram, S. *The Mathematica Book,* 3rd Ed.; Cambridge University Press: Cambridge, UK, 1996; ISBN 0-521-58889-8.
24. Feng, S.; Winful, H.G. Physical origin of the Gouy phase shift. *Optics Letters* 2001, *26,* 485–489.
25. Johnston, T.F., Jr. Beam propagation (M²) measurement made as easy as it gets: The four-cuts method. *Applied Optics* 1998, *37,* 4840–4850.
26. Belanger, P.A. Beam propagation and the ABCD ray matrices. *Optics Letters* 1991, *16,* 196–198.
27. Serna, J.; Nemes, G. Decoupling of coherent Gaussian beams with general astigmatism. *Optics Letters* 1993, *18*, 1774–1776.
28. Hecht, E. *Optics,* 2nd Ed.; Addison–Wesley Publishing Co.: Menlo Park, CA, 1987; ISBN 0-201-11609-X.
29. Kogelnik, H. Imaging of optical modes—Resonators with internal lenses. *Bell System Technical Journal* 1965, *44,* 455–494.

30. Self, S.A. Focusing of spherical Gaussian beams. *Applied Optics* 1983, *22*, 658–661.
31. Herman, R.M.; Wiggins, T.A. Focusing and magnification in Gaussian beams. *Applied Optics* 1986, *25,* 2473–2474.
32. O'Shea, D.C. *Elements of Modern Optical Design;* John Wiley & Sons: New York, 1985; ISBN 0-471-07796-8, 235–237.
33. Wright, D.L.; Fleischer, J.M. Measuring Laser Beam Parameters Using Non-Distorting Attenuation and Multiple Simultaneous Samples. US Patent No. 5,329,350, 1994.
34. McCally, R.L. Measurement of Gaussian beam parameters. *Optics Letters* 1984, *23,* 2227.
35. Sasnett, M.W.; Johnston, T.F., Jr. Apparatus for Measuring the Mode Quality of a Laser Beam. US Patent No. 5,100,231, March 31, 1992.
36. Taylor, J.R. *An Introduction to Error Analysis;* University Science: Mill Valley, CA, 1982; ISBN 0–935702-10–5.
37. Green, L. Automated measurement tool enhances beam consistency. *Laser Focus World* 2001, *37*(3), 165–166.
38. Kogelnik, H.; Ippen, E.P.; Dienes, A.; Shank, C.V. Astigmatically compensated cavities for CW dye lasers. *IEEE Journal of Quantum Electronics* 1972, *3,* 373–379.
39. Johnston, T.F., Jr.; Sasnett, M.W. The effect of pump laser mode quality on the mode quality of the CW dye laser. SPIE Proceedings 1992, *1834,* Optcon Conference, Boston, 1992, Paper #29.
40. Johnston, T.F., Jr.; Sasnett, M.W. Modeling multimode CW laser beams with the beam quality meter. OPTCON, Boston, MA, 5 November 1990, Paper OSM 2.4.
41. Firester, A.H.; Gayeski, T.E.; Heller, M.E. Efficient generation of laser beams with an elliptic cross section. *Applied Optics* 1972, *11,* 1648–1649.
42. Siegman, A.E. Laser beam propagation and beam quality formulas using spatial-frequency and intensity-moment analysis, distributed to the ISO Committee on test methods for laser beam parameters, August 1990, 32.
43. Simpson, N.B.; Dholakia, K.; Allen, L.; Padgett, M.J. Mechanical equivalence of spin and orbital angular momentum of light: and optical spanner. *Optics Letters* 1997, *22,* 52–54.
44. Nemes, G.; Serna, J. Laser beam characterization with use of second order moments: An overview. In *DPSS Lasers: Applications and Issues, OSA TOPS;* Dowley, M. W., Ed.; 1998; 17, 200–207.
45. Piestun, R. Multidimensional synthesis of light fields. *Optics and Photonics News* 2001, *12*(11), 28–32.
46. Kivsharand, Y.S.; Ostrovskaya, E.A. Optical vortices. *Optics and Photonics News* 2002, *13*(4), 24–28.
47. Partanen, J.P.; Jacobs, P.F. Lasers for stereolithography. In *OSA TOPS on Lasers and Optics for Manufacturing;* Tam, A.C., Ed.; Optical Society of America: Washington, DC, 1997; Vol. 9, 9–13.
48. Partanen, J. Lasers for solid imaging. *Optics and Photonics News* 2002, *13*(5), 44–48.
49. Ibbs, K.; Iverson, N.J. Rapid prototyping: New lasers make better parts, faster. *Photonics Spectra* 1997, *31*(6), 4 pages.
50. SLA 250/30 Product Data Sheet from 3D Systems, 26081 Avenue Hall: Valencia, CA 91355.