Journal of Directed Energy, **2**, Summer 2006, 22–58

Appropriate Measures and Consistent Standard for High-Energy Laser Beam Quality

T. Sean Ross*

High Power Solid State Laser Branch, DELO, Laser Division, Directed Energy Directorate, Air Force Research Laboratory, 3550 Aberdeen Avenue SE, Kirtland Air Force Base, New Mexico 87117

and

William P. Latham

Tactical Systems Branch, DELS, Laser Division, Directed Energy Directorate, Air Force Research Laboratory, 3550 Aberdeen Avenue SE, Kirtland Air Force Base, New Mexico 87117

Along with power output of the laser system, laser optical quality or beam quality provides a suitable measure of performance. Power and beam quality are standards for the comparison of laser systems with each other and against mission requirements. An understanding of the meaning of beam quality is necessary to completely define laser performance capability. The current state of our community includes a multitude of different and not well understood beam quality measures: M^2, Strehl ratio, brightness, power in the bucket, "times diffraction limited," and mode content determined by a variety of beam radius measures: half-widths, second-moment radius, widths at $1/e$ or $1/e^2$ points, width of primary lobe, etc. Another complication is that different elements of the community use different measures to evaluate optical quality characteristics. We examine the assumptions behind common measures of beam quality and compare the various measures as they relate to beams from lasers employing stable resonant optical cavities. We show how the mode composition of a beam depends on prior determination of beam radius and how the term "times diffraction limited" can mean different things depending on the method used to measure beam radius. We show the ambiguities that arise between certain classes of beams and measures of beam quality and advocate for a laser beam quality standard that relates directly to mission requirements.

KEYWORDS: High-power lasers, Laser beam quality, Laser standards, M^2

Nomenclature

A	area
a	spatial limit of knife edge scan
B	brightness

Received October 10, 2005; revision received March 10, 2006.
*Corresponding author; e-mail: sean.ross@kirtland.af.mil.

C, c	mode coefficient
D	diameter
d	measured spot size
E	electric field
e	natural base, 2.7128...
f	focal length
H	Hermite polynomial
I	irradiance (radiometric nomenclature) and/or intensity (physics nomenclature)
$\hat{I}$	normalized irradiance
i	summation index
k	knife edge function
l	azimuthal mode index for Laguerre–Gaussian functions
N	number
n	x index for Hermite–Gaussian functions
M	laser mode quality
M^2	a laser beam quality measure
m	measured; y index for Hermite–Gaussian functions
P	power
P_{bi}	ideal power in the bucket
P_{bm}	measured power in the bucket
p	power fraction; radial mode index for Laguerre–Gaussian functions
p_{bi}	ideal power fraction in the bucket
p_{bm}	measured power fraction in the bucket
r	radius
S	Strehl ratio
s	source
TEM	transverse electromagnetic
TEM_{nm}	the nth, mth TEM mode
t	target
u	orthonormal basis function
W	measured mode radius
w	beam radius
w_0	beam waist
XDL	times the diffraction limit
x, y	transverse distances
z	axial distance
z_R	Rayleigh range
β	beam quality
$\Delta\nu$	bandwidth
δ	Kronecker delta function
$\Delta\phi$	rms wave-front distortion
Θ	divergence angle of envelope
θ	angle
θ_c	beam divergence for a circular profile
θ_g	beam divergence of a circular Gaussian profile
θ_{gi}	beam divergence in the x or y direction of a Gaussian profile
θ_{gl}	beam divergence measured at the lens of a Gaussian profile

θ_{gli}	beam divergence measured at the lens in the x or y direction of a Gaussian profile
θ_h	beam divergence of a higher-order Gaussian mode
θ_l	beam divergence in x or y direction
θ_m	beam divergence of a TEM_{00} Gaussian mode, same as θ_g
λ	wavelength
ν	frequency
π	3.14159...
σ	variance
ϕ	azimuthal cylindrical angle
Ω	solid angle

1. Background

The output of a laser is nearly monochromatic and extremely coherent. Narrow linewidth or long coherence length is the primary operational characteristic of a laser, and it allows the laser output beam to be focused to a tiny spot. For a thorough discussion of formal coherence theory and relationships, see Refs. 1 and 4. The optical quality or beam quality of a laser is a measure of the laser's focusability. More important, if the optical quality of the laser is excellent, the laser beam is controllable and understandable. Many solid-state lasers employ a resonant cavity that produces a beam profile made up of a series of Hermite– or Laguerre–Gaussian modes. For the case of a stable optical resonator, the laser focusability is compared to the focusability of a monochromatic field with the lowest-loss Gaussian beam spatial profile as an ideal standard to determine its optical quality. If the electromagnetic field representing the modes of the laser is exactly known, the shape and power in the beam could be evaluated anywhere within the focal volume of that beam. However, all of the parameters of the beam are not generally measured. Moreover, a solid-state laser generally includes thermally induced self-focusing effects within the laser gain material as the laser heats up. Although there are analytic ways to estimate these effects, it is important for higher laser powers to make an empirical measure of the solid-state laser's optical quality under operating conditions to evaluate the actual as-built laser performance. Laser output power and laser optical quality are the two critical performance parameters of a solid-state laser that must be determined as the directed energy community develops solid-state lasers with higher powers. The purpose of this paper is to explain the meaning of optical quality and mode quality in basic conceptual terms, so that it is understandable to a broad spectrum of the laser development community and to point out the assumptions, strengths, and weaknesses of various measures of laser beam quality to enable an informed choice of the mission-appropriate measure.

2. Measures of Optical Quality

Optical quality affects a laser beam's focusability. Measures of optical quality are divided into two categories. The first category is empirical or measurable quantities, such as total laser output power or energy, laser linewidth, focal spot size, far-field peak irradiance, and encircled power or energy in the focal spot, which is sometimes called the power in the bucket (PIB). Also included in this category are parameters, which are directly calculated using these measured values, such as beam divergence, coherence length, coherence width or

area, and brightness. These parameters quantify the laser performance empirically without comparison to a standard. A second category involves relative parameters, which compare the focusability of an actual laser beam to the focusability of an ideal standard laser beam, such as M^2, beam quality, mode quality, and Strehl ratio.[1]

2.1. Beam divergence

Diffraction is the name given to the angular spreading or divergence of light. A measure of the diffractive spread of a laser beam is the ratio of the average wavelength to the half-size of the laser beam. Various expressions for the beam divergence angles[8] are listed in Table 1. Experimentally, the focused spot diameter or spot width can be measured for a given laser. An empirical beam divergence is defined to be the ratio of the spot diameter or spot width to the distance between the focal plane and lens.

Table 1. Beam divergence angles

Expression	Description
Beam divergence	Divergence angle in the Ith direction, $I = x$ or y, for the fundamental mode.
$\theta_I = 2\lambda/D_I$	(TEM_{00} mode) due to a plane wavefront incident upon a rectangular aperture of dimension $D_x \times D_y$. About 81% of the total energy is contained in a rectangular spot defined by these divergence angles.
$\theta_c = 2.44\lambda/D$	Divergence angle for TEM_{00} mode due to a plane wavefront incident upon a circular aperture of diameter D. About 84.5% of the total energy is contained in a circular spot defined by this divergence angle.
$\theta_g = D_I/f = 2\lambda/(\pi w_0)$	Divergence angle for TEM_{00} mode in the far field of a circular Gaussian beam of waist radius w_0. About 86.5% of the total energy is contained in a spot defined by this divergence angle. θ_g is measured at the beam waist.
$\theta_{gI} = 2w_0/f = 4\lambda/(\pi D_I)$	Divergence angle for TEM_{00} mode of a circular Gaussian beam. θ_{gI} is measured at the focusing lens.
$\theta_{gi} = D_{Ii}/f = 2\lambda/(\pi w_{0i})$	Divergence angle in the Ith direction, $i = x$ or y, for TEM_{00} mode in the far field of a rectangular Gaussian beam. θ_{gi} is measured at the beam waist.
$\theta_{gIi} = 2w_0/f = 2\lambda/(\pi D_{Ii})$	Divergence angle in the Ith direction, $i = x$ or y, for TEM_{00} mode of a rectangular Gaussian beam. θ_{gIi} is measured at the focusing lens.
$\theta_m = d_s/f$	Divergence angle for any laser beam. d_s is the measured spot size, and f is the focal length of the focusing lens.
For θ_g (see above)	Divergence of the TEM_{00} Gaussian mode that is the fundamental mode for the higher-order mode of interest.
$\theta_h = M\theta_g$	Divergence of the higher-order Gaussian mode.

Table 2. Short list of measures of optical quality

Optical quality parameter	Expression
M and M^2 factors	
$M = \theta_h/\theta_g$	$M = (2p + l + 1)^{1/2}$ for circular TEM_{pl} mode
$M^2 = \theta_{hx}/\theta_{gx}$, θ_{hy}/θ_{gy}	$M_x = (2m + 1)^{1/2}$ in the x direction for rectangular TEM_{mn} mode
	$M_y = (2n + 1)^{1/2}$ in the y direction for rectangular TEM_{mn} mode
Gaussian beam quality	$\beta_g \approx M$
Beam parameter product	$W_0\Omega$
Brightness	$B = P/(A_s\Omega) = P/\lambda^2$
Total brightness	$B_t = P/\Omega = PA_s/\lambda^2$
Spectral brightness	$B_\nu = P/(\Omega\Delta\nu)$ and $B_\lambda = P/(A_s\Omega\Delta\lambda)$
Strehl ratio	$S = I_m/I_0 \sim 1 - (2\pi\Delta\phi/\lambda)^2 \sim \exp[-(2\pi\Delta\phi/\lambda)^2]$
Strehl ratio beam quality	$\beta_s = 1/S^{1/2} \sim 1 + (\pi\Delta\phi/\lambda)^2 \sim \exp[+(\pi\Delta\phi/\lambda)^2]$
Encircled power ratio in a circle of radius r_c	$\beta = (p_{\text{ideal}}/p_{\text{actual}})^{1/2}$
Total beam quality	$\beta_t = \Pi_i \beta_I$
Reduced brightness	$B_\beta = (1/\beta^2)(PA_s/\lambda^2) = f^2 I_r$

2.2. Strehl ratio

Strehl ratio (Ref. 1, Chap. 9.1.3) is most commonly used in the astronomy community. Astronomers often image point objects located in the extreme far field. Strehl is a ratio between the peak irradiance of a measured signal and the calculated peak irradiance of an aberration-free signal. In the laser community, Strehl customarily means the ratio between the peak irradiance of a laser beam and the peak irradiance of a zero-order Gaussian or other ideal beam with the same power and beam radius as in Eq. (1):

$$S = \frac{I_{\text{max-measured}}}{I_{\text{max-ideal}}}. \tag{1}$$

The Strehl ratio is commonly used in the analysis and modeling of optical system performance. In the presence of optical aberrations, mirror distortions, gain medium imperfections, and/or beam jitter, the far-field peak irradiance is degraded. If these optical distortions are small, the ratio of the actual or aberrated far-field peak irradiance I_m to the ideal or unaberrated peak irradiance I_0 is the Strehl ratio S given in Table 2. The Strehl ratio has the advantage that it is simple and can be useful in troubleshooting laser optical performance problems or during laser design when a particular distortion will potentially be introduced by the system. An empirical form of the Strehl ratio is given by measuring the actual far-field peak irradiance and calculating the ideal peak irradiance from the laser output power.

2.3. Brightness

There are many definitions [Ref. 3, p. 239, and Ref. 7, Eq. (4)] of brightness within the laser and optics community. In general, the brightness B for a coherent laser source

of output area $A_s = A_{\text{source}}$ and total power P is given by $B = P/(A_s\Omega)$, where Ω is the solid angle subtended by the area in the far field or observation plane a distance $z = f$ ($f = 1$ focal length) from the laser source as in Eq. (2). The solid angle is given by $\Omega = \Omega_{\text{target}} = A_{\text{target}}/z^2 = \lambda^2/A_s$:

$$B = \frac{P}{A_s\Omega} = \frac{P}{A_{\text{source}}\Omega_{\text{target}}} = \frac{z^2 P}{A_{\text{source}}A_{\text{target}}} = \frac{f^2 P}{A_s A_t}. \tag{2}$$

There is no ambiguity in A_{source}. It is defined in terms of the hard aperture. This is not the case with A_{target}, which can be defined in terms of second moment, first lobe, or other beam radius measure. With suitable care, brightness can be a consistent measure. If one's mission is to efficiently put power in a solid angle, such as for laser imaging, detection, and ranging (LIDAR), tracking, or communications missions, then brightness is the mission-appropriate measure. It is also customary to make brightness a relative measure by removing power, in which case it is called etendue. One advantage to brightness and etendue is that they are unaffected by the optics of a system. In the absence of aberrations, absorption, and turbulence, brightness is a conserved property of an optical system. Various expressions for brightness are given in Table 2, and a discussion of this topic is presented in Ref. 8.

2.4. Beam quality M^2 or mode quality

The primary measurable quantity that determines laser focusability performance is the encircled energy or power in a small region around the focal spot. Beams propagating from hard apertures have a well-defined central lobe that determines their focal spot. Soft apertured beams, such as Gaussian modes, do not have a well-defined focal spot, and so measures such as second moment or $1/e^2$ must be used. If the power or energy measured in the small area around the focal spot is P_s and the total laser output power or energy is P, the fraction of power or energy delivered by the laser to the focal spot is $p = P_s/P$. The fractional power p or energy in the far-field focal spot is a readily measurable and meaningful quantity. If the laser is operating in a single lowest-loss mode, the fractional power can be near the maximum obtainable for some ideal standard waveform. For a uniformly illuminated circular aperture, $p \approx 0.84$. For a rectangular beam geometry, $p \approx 0.81$. For the lowest-order Gaussian beam, $p \approx 0.86$ by the $1/e^2$ criterion. The beam quality β is defined to be the square root of the ratio of the fractional power in the far-field spot for an ideal standard beam to the fractional power in the actual laser beam, that is, $\beta = (p_{\text{ideal}}/p_{\text{act}})^{1/2}$. $\beta = 1$ if the actual beam has the same far-field power as the ideal standard. The physical mechanism that determines the mode structure of the laser beam and the focused beam spread is diffraction. For $\beta = 1$, the laser is said to be diffraction-limited. The Strehl ratio beam quality β_s is given by $\beta_s = (1/S)^{1/2}$. The bare cavity modes of a stable resonant cavity are Gaussian beams. For a single higher-order Gaussian mode, the beam waist is M times the beam waist of the lowest-loss Gaussian beam, the TEM_{00} mode. The far-field peak irradiance and the power within a small area around the focal spot is reduced by a factor of M^2, that is, the Gaussian beam quality β_g is related to M^2 by $\beta_g = (M^2)^{1/2} = M$. M^2 or β_g provides a meaningful measure of the optical quality when the beam structure consists of some combination of the Gaussian modes.[2,10] The product of beam waist times far-field divergence is known as the beam parameter product and is another measure of optical quality that can be related[16] to M^2 and is most commonly used to measure the beam quality of semiconductor lasers. If there are several sources of aberration or beam degradation, the total beam quality is equal to a product of all individual contributions, so that $\beta = \Pi\beta_i$.

The various measures of determining beam radius will be discussed later, but it is important to mention here that most measures of beam quality are highly dependent on the method chosen to determine beam radius and that there can be considerable variation in measured beam quality depending on the beam radius definition.

2.5. PIB

PIB is simply the sum of the total power within a particular area. Usually the area is circular. There are two basic means for determining the size of the circle. The first is simply to use the actual size appropriate to a mission target. If this is not known or may be variable, then one may use a size based on the diffraction properties of the laser output aperture. The radius of the bucket is called r_b. A discussion of how r_b might be chosen for a particular definition of beam quality is given in a later section. The normalized PIB (NPIB) can be calculated as the fraction of output power that ends up inside the target circle:

$$\mathrm{NPIB} = \frac{\int_0^r \int_0^{2\pi} I_{\mathrm{actual}}(r,\phi) r\,\mathrm{d}r\,\mathrm{d}\phi}{\int_0^\infty \int_0^{2\pi} I_{\mathrm{actual}}(r,\phi) r\,\mathrm{d}r\,\mathrm{d}\phi} = p_{\mathrm{actual}}. \tag{3}$$

Note that the NPIB for some "ideal" beam is given by Eq. (3) with actual replaced by ideal. If the ideal beam is chosen to have the same total power as the actual beam, the ratio comparing the power inside the target circle to an ideal beam is given by

$$\mathrm{PIBR} = \frac{\int_0^{r_b} \int_0^{2\pi} I_{\mathrm{actual}}(r,\phi) r\,\mathrm{d}r\,\mathrm{d}\phi}{\int_0^{r_b} \int_0^{2\pi} I_{\mathrm{actual}}(r,\phi) r\,\mathrm{d}r\,\mathrm{d}\phi} = \frac{p_{\mathrm{actual}}}{p_{\mathrm{ideal}}} = \frac{1}{\beta^2}. \tag{4}$$

where PIBR is relative PIB. The ratios are useful for comparing the efficiency of different systems but may lead to disagreements over the basis of comparison. Promoters of stable resonators will likely want a PIB based on an ideal low-order Gaussian beam that raises all the ambiguities in determining the characteristic radius of an appropriate comparison beam. Those working with unstable resonators will likely want a comparison with a flattop. The actual power is, of course, what ends up accomplishing the mission and represents the bottom line. In the comparisons that follow, we use Eq. (4) with a target circle equal to three times the second-moment radius.

2.6. Comparison of beam quality measures

Figures 1–3 show comparisons of M^2, Strehl ratio, brightness, and PIBR for various laser beams, first, to illustrate that these measures of beam quality are distinct and ought not to be conflated in our thinking, and second, to provide some insight to the researcher to decide which measure is appropriate for a given application. For the stable resonator modes in Figs. 1 and 2, the measures of beam quality were calculated using the $z = 0$ plane and the $z = 3{,}000$ m plane for the near and far fields. No atmospheric distortions were included. The beam itself was chosen to have a $z = 0$ (near field) second-moment radius of 3 cm. M^2 was calculated from the mode coefficients using Eq. (26). M^2 is always greater than or equal to 1 and will be shown with numbers on the right-hand scale. The Strehl ratio, brightness, and PIBR were scaled to make their maximum value 1 and are shown with numbers on the left-hand scale. The Strehl ratio was calculated from the peak irradiances

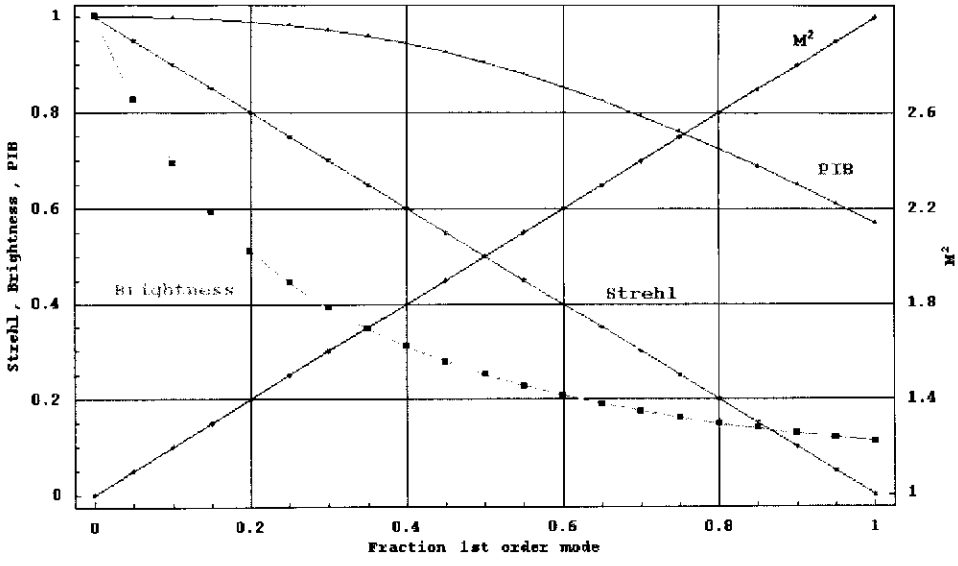

Fig. 1. Comparison of beam quality measures for a beam composed of TEM_{00} and TEM_{11} modes.

of near- and far-field beams. The brightness used is a relative brightness, sometimes called etendue, based on the physical aperture size in the near field and the second-moment beam radius in the far field. This was so that the brightness would have a maximum value of 1 and would fit on the same graph as the other measures. PIBR was calculated using the total power inside a circle with radius equal to three times the second-moment waist of the zero-order Gaussian beams using Eq. (4). This was done to provide a fair comparison since the zero-order Gaussian beams in the rest of the study were assumed to come from a hard aperture with radius three times their second-moment waist.

Figures 1 and 2 show M^2, Strehl ratio, relative brightness, and PIB as a function of percent of higher-order modes. The left-hand sides of these figures are for pure zero-order Gaussian modes, TEM_{00}, and the right-hand sides are for pure first-order, TEM_{11}, and pure $TEM_{10} + TEM_{01}$ "donut" modes as shown later in Fig. 11. Figure 1 shows that as the percentage of the beam that is first-order mode increases from 0 to 100%, M^2 increases from 1 to 3, the Strehl ratio drops from 1 to 0, and relative brightness decreases to ~13%, while relative PIB decreases to about 55%. Figure 2 shows that as the percentage of the beam that is donut mode increases from 0 to 100%, M^2 increases from 1 to 2, the Strehl ratio drops from 1 to 0, and relative brightness decreases to ~25% while PIB decreases to ~88%. Comparing two systems for suitability to a defense mission, is a donut mode 50, 0, 25, or 88% as good a beam as a zero-order Gaussian? Is a TEM_{11} beam 33, 0, 13, or 55% as good a beam as a TEM_{00} Gaussian? The differing measures of beam quality clearly do not measure the same thing.

Figure 3 shows a comparison of various beam quality measures vs. the outcoupling size. It is important to emphasize that M^2 is based on the second-moment measure of the beam radius. Theoretically, second moment requires an infinite plane: all information is significant, and information away from the centroid is quadratically more and more significant. M^2 here is calculated from Fourier propagation on a 1,024 × 1,024 numerical

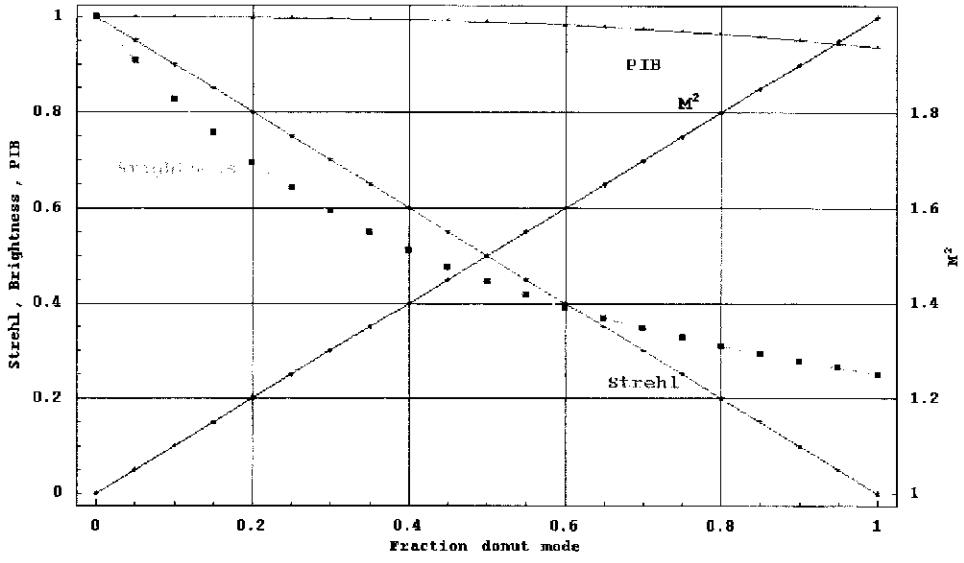

Fig. 2. Comparison of beam quality measures for a beam composed of TEM_{00} and $TEM_{01} + TEM_{10}$ (donut) modes.

array with spacing of 0.2 mm, a total field approximately 20 cm across. The far-field second-moment radius calculated is a function of this numerical field, similar to the way a measured M^2 will be a function of the detector size and noise equivalent aperture, as will be discussed in the section on experimental foundations of M^2. Different systems will measure various

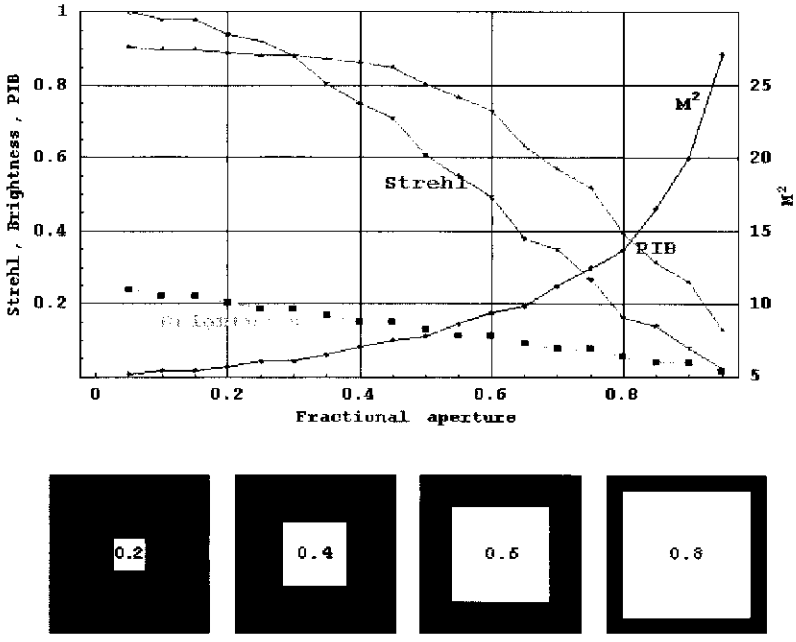

Fig. 3. Comparison of beam quality measures for a square annulus flattop beam for various apertures.

M^2 values for the same beam. We thus do not intend to convey that the M^2 of a square annulus flattop is 28. Theoretically, M^2 is infinite for all beams from hard apertures. We do intend to convey that M^2 can be measured to be 28 in conditions that match those of our Fourier propagation code. In Fig. 3, M^2 increases from ~5 to ~28 as the fractional aperture increases from 0.05 to 0.95. The Strehl ratio (compared with that of a square flattop beam) decreased from 1 to 0.0, relative brightness decreased from 0.25 to 0, and PIB decreased from 88 to 15%. In comparing a square flattop with a square annulus flattop with a fractional aperture of 0.4, is the apertured beam 70, 75, 60, or 97% as good a beam as the square flattop?

3. Foundations of M^2

The term M^2 has developed out of the theory of Gaussian modes. Gaussian modes are the bare cavity solutions for stable resonant cavities. In this section, the Laguerre–Gaussian modes that are appropriate for radial and azimuthal coordinates will be discussed. The Cartesian coordinate modes, the Hermite–Gaussian modes, are given below. The higher-order radial Gaussian modes grow in size as a function of the radial mode index. The radial size r_n of the nth mode [Ref. 12, Chap. 7.5, Eq. (44)] is approximated by

$$r_n \approx \sqrt{n} \times w_0, r_n \equiv M w = w_n, \tag{5}$$

where w_0 is the radius of the lowest-order, zeroth-order, Gaussian, the TEM_{00} mode. The Laguerre–Gaussian modes and Hermite–Gaussian modes are given in Ref. 12 [Chap. 17.1, Eq. (1); Chap. 17.5, Eq. (40); and Chap. 16.4, Eq. (64)]. We have chosen to label the scaling factor for the mode M for reasons that will be clear in the next section. Sometimes the radial mode size is labeled w_n. The lowest-order Gaussian mode beam radius has a quadratic profile as a function of axial distance along the center of the beam:

$$w = w(z) = w_0\sqrt{1 + \left[\frac{\theta_0(z - z_0)^2}{w_0}\right]} = w_0\sqrt{1 + \left[\frac{(z - z_0)^2}{z_R}\right]},$$

$$w_0 = w(z_0), \qquad \theta_0 = \frac{\lambda}{\pi w_0}, \qquad z_R = \frac{\pi w_0^2}{\lambda}, \tag{6}$$

where w_0 is the zeroth-order Gaussian beam waist radius, θ_0 is the zeroth-order Gaussian beam divergence angle, and z_R is the Rayleigh range. The higher-order modes follow a similar quadratic profile with the beam radius scaled everywhere in z by the scaling factor M. The M^2 value or the mode quality is determined by measuring the quadratic envelope of a particular actual laser beam and finding the fit parameters for a quadratic curve. The quadratic envelope can readily be calculated for the envelope $W(z)$. The expression for the quadratic that follows the Gaussian beam envelope is

$$W(z) = W_0\sqrt{1 + \left[\frac{\Theta_0(z - z_0)}{W_0}\right]^2}, \qquad W(z) = M w(z), \tag{7}$$

where the second-moment Gaussian beam radius for the zeroth-order Gaussian is $w(z)$, Θ is the divergence of the envelope, and the beam radius for the actual beam, $W(z)$, is taken to be the scaling parameter M times the zeroth-order Gaussian as given in Eq. (5). The result

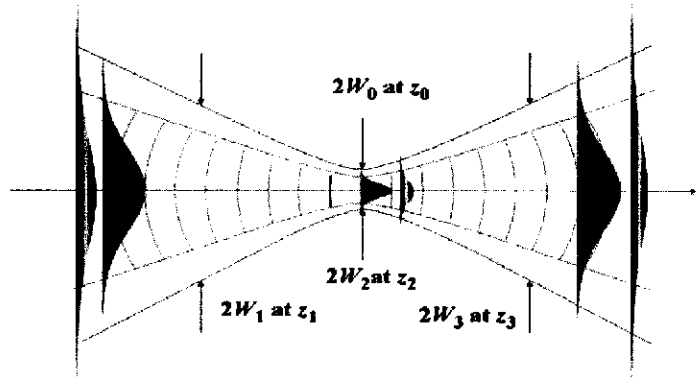

Fig. 4. Determining the parameters for the Gaussian envelope.

for the envelope quadratic function is

$$M^2 = \frac{\pi}{\lambda} W_0 \Theta_0, \qquad \Theta_0 = \frac{M\lambda}{\pi W_0} = M\theta_0. \tag{8}$$

Note that we have chosen to label the radius for the zeroth-order beam with lowercase symbols: 1) beam radius is $w(z)$ and 2) beam divergence is θ. The quadratic envelope parameters for the actual measured beam are labeled with uppercase symbols: 1) beam radius is $W(z)$ and 2) beam divergence is Θ. Determining the M^2 value depends on what particular measurement scheme is used and what choices are made for values within the fitting process. Usually, the actual beam is a mixture of a few or many Gaussian modes. There are analytic expressions for the scaling factor for single higher-order modes. However, there is not a unique Gaussian mode series solution for a given electric field profile or envelope. In the absense of any laser fluctuations or noise, three measurements of the beam radius W_i and beam positions z_i at three locations are sufficient to determine the quadratic envelope of the beam if there is no noise or beam fluctuation of any kind. In practice,[5] many measurements must be taken. If the beam consists of a single higher-order Gaussian mode, the value for M can be calculated by measuring the second-moment beam radius and multiplying by the mode coefficients as in Eq. (26).

For a multimode Gaussian beam, the beam also has a quadratic shape. M^2 is based on the comparison of the multimode beam to an ideal TEM_{00} beam. However, as the number of modes increases, the proper beam radius of that ideal TEM_{00} beam becomes increasingly hard to determine and creates inherent uncertainty in the measure of M^2, as will be discussed below. The irradiance for the one-dimensional Hermite–Gaussian modes, TEM_n modes, is shown in Fig. 5.

The formula for these modes is given in Eq. (21). Note that the modes increase in transverse size as the mode index increases, as in Eq. (5). After normalizing to the laser power P, the irradiance of the lowest-order Gaussian mode is given by

$$I(r) = \frac{2P}{\pi w^2} e^{-2r^2/w^2} = \frac{2P}{\pi w^2} e^{-[2(x^2+y^2)]/w^2}, \tag{9}$$

where the beam radius as a function of the longitudinal coordinate z and the beam divergence angle are given by Eq. (6). Since the higher-order Gaussian modes increase in transverse size

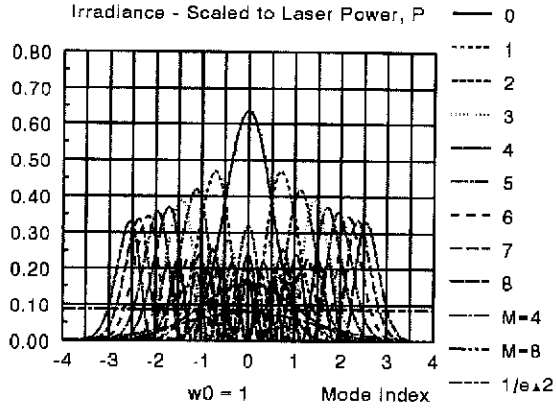

Fig. 5. One-dimensional Hermite–Gaussian mode irradiances.

with increasing index, the irradiance envelope for a higher-order mode or multimode beam can be approximated by the lowest-loss Gaussian that is larger in radius. By normalizing to the same laser power, the larger mode envelope is given by

$$I_M(r) = \frac{2P}{\pi M^2 w^2} e^{-2r^2/M^2w^2} = \frac{2P}{\pi W^2} e^{-2r^2/W^2}, \qquad W = Mw, \tag{10}$$

where the beam radius and beam divergence angle are given in Eqs. (7) and (8). Once again, the expression can readily be converted to x and y coordinates by using $r^2 = x^2 + y^2$. For nonsymmetric cases, the envelope irradiance can be written as a product of an irradiance in x and an irradiance in y that are scaled by M_x and M_y. In that case, the mode quality value M^2 is a product as follows:

$$M^2 = M_x M_y. \tag{11}$$

The mode quality or M^2 value is directly related to the quadratic envelope of the beam as discussed in the preceding section. Another measure of optical quality is the PIB as given in Eq. (3). Here, we take the bucket radius to be the $1/e^2$ point in the lowest-loss Gaussian far-field irradiance pattern. At this bucket radius, the PIB in the lowest-loss Gaussian mode is given by Ref. 12 [Chap. 17.1, Eq. (13); Chap. 17.2, Eq. (24); and Chap. 16.4, Eqs. (48)–(60)†]:

$$P_{bi} = P \times (1 - e^{-2}); \qquad p_{bi} = 1 - e^{-2} \approx 86.4\%. \tag{12}$$

That is, 86% of the total laser power is delivered into the bucket for the lowest loss mode. For the larger-radius-mode envelope, the PIB is given by

$$P_{bm} = P \times \left(1 - e^{-2/M^2}\right); \qquad p_{bm} = 1 - e^{-2/M^2}. \tag{13}$$

Note that Eq. (13) reduces to Eq. (12) for $M^2 = 1$. When $M^2 > 1$, the normalized PIB is less than 86%. Finally, the beam quality is usually taken to be the square root of the ratio

†Note that we use a slightly different normalization from those in the latter equations and that the form quoted does not include the complex factors. The full form is used in computation but was not cited in this paper for simplicity. See also Ref. 17, http://mathworld.wolfram.com/HermitePolynomial.html, Eq. (39).

of the ideal standard beam PIB to the actual beam PIB for the same bucket radius. Thus, the beam quality of the envelope beam can be given by

$$\beta = \sqrt{\frac{p_{bi}}{p_{bm}}} = \sqrt{\frac{1 - e^{-2}}{1 - e^{-2/M^2}}}. \tag{14}$$

The transverse envelope function has been used to determine a "beam quality" based on a comparison to the PIB of the lowest-loss Gaussian beam as an ideal standard. This gives a definition of beam quality for a multimode Gaussian beam. Asymmetric beams can readily be analyzed in a similar way.

An important conceptual subtlety is involved in the use of Eq. (13). As one departs from single-mode content, it becomes less and less clear what the proper radius of the lowest-order embedded Gaussian is. As one progresses to flattop shapes, there are many different, reasonable criteria on which to make this comparison, and none of them is totally convincing because a flattop has a hard aperture while Gaussian modes are all of infinite extent. By suitable comparison, a flattop can look terrible or very good in comparison to a TEM_{00} Gaussian. If we judge each beam against itself, rather than by a tenuous comparison to an ideal, then we find that a TEM_{00} Gaussian is not vastly superior to other beam shapes. For example, a TEM_{00} Gaussian has 86% of its energy inside a $1/e^2$ bucket while flattop beams have approximately 77% of their energy inside this radius, yet flattops have an infinite M^2. The general rule for the use of M^2 and Eq. (13) is that they apply only to beams that can be fully and uniquely characterized by a few Gaussian modes. Reference 6 (Sec. 6.4.1, p. 33), for example, reports that M^2 is a unique and meaningful measure for stable resonator beams with nonannular round beams up to $M^2 < 3.2$. In the following sections, we discuss some of the experimental and conceptual problems associated with the use of M^2.

4. M^2: What Your Beam Analyzer's Manual Didn't Tell You

M^2 is probably the most popular measure of laser beam quality. It is also one of the most experimentally misused and inconsistently applied measures. Many scientists have their own pet methods for measuring M^2: some use a knife edge; some take one measurement at focus and one in the far field; some measure the transmission through a single-mode optical fiber; some relate it to Strehl ratio or other beam quality measures. Many researchers rely on commercial "black-box" devices and accept the manufacturer's assurances that the device actually measures M^2. One cannot rigorously examine the methods these various commercial devices use because they are hidden behind the word "proprietary." One must trust the manufacturer's sales literature. A proper M^2 measurement is difficult to take. Fortunately, there is an international standard[5] that specifies how the measurement is to be taken. A proper M^2 measurement takes the second-moment beam radius measured in at least 10 places in the far field and through the focus and then fitted to an ideal Gaussian equation in terms of Rayleigh range, focus location, and M^2. Further, the experimenter must be able to understand the sources of uncertainty in the measurement to be able to correctly assign error values, something sadly lacking from most commercial "black-box" beam analyzers.

4.1. Tradeoffs

The heart of a beam quality measurement system is the camera. Silicon-based charge-coupled device (CCD) or charge-induced device (CID) cameras are the most common,

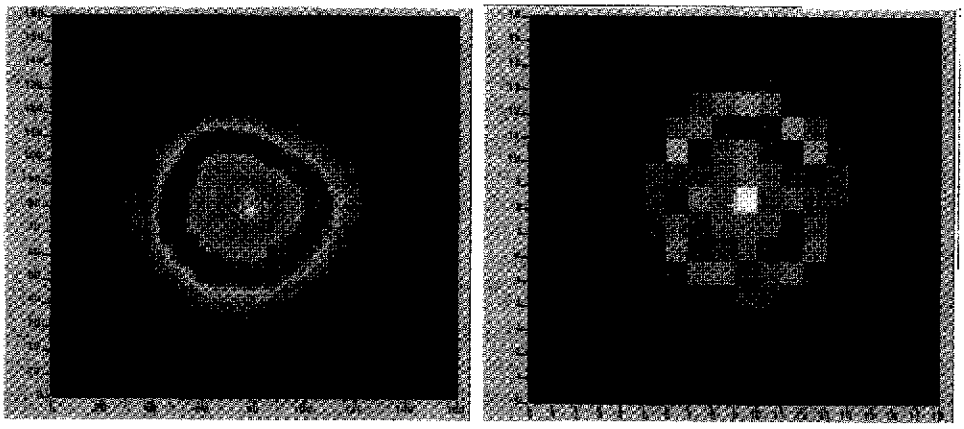

Fig. 6. Near- and far-field beam measurements.

typically have pixel sizes <10 μm, detect adequately in the visible and near infrared (NIR), and are relatively inexpensive. For the midinfrared (MIR), there are screened cameras (with pixel bleed) or pyroelectric CCD cameras with large pixels (100×100 μm). These cameras can necessitate the use of very long travel stages due to low resolution. A long travel stage allows a loose focus, a smaller variation in irradiance as the detector is moved through the beam, less variation in signal-to-noise ratio, and less stringent requirements on the pixel size of the camera. These benefits come at a price. Short travel precision stages (<10 cm) are relatively inexpensive. As the travel distance increases beyond 10 cm, the price increases significantly. These short travel stages give focused spot sizes appropriate for silicon-based CCD cameras for the visible and NIR but are often inadequate in the MIR, where screened cameras are subject to pixel bleed and pyroelectric CCD cameras that have very large pixels. A further implication of stage length is the size of the image in the CCD camera and the number of pixels across the beam. If, for example, we fill the CCD aperture in the far field, we may have 300^2 pixels across the beam. With a short travel stage, we may have to focus down to only 10^2 pixels across the beam, with a corresponding decrease in signal-to-noise ratio. The reliability of an M^2 measurement is in the near focus and far-field measurements. Poor signal to noise near focus is a source of error in the final measured value of the beam radius. Figure 6 shows a sample beam measured in the far field and near focus. The accuracy of the beam radius measurement is far lower near focus than in the far field due simply to several orders of magnitude fewer pixels across the beam. Note the change of transverse scale in Fig. 6 showing the significant portion of each beam.

4.2. Pseudo-average M^2

The measurement of M^2 is plagued by a number of theoretical–experimental disconnects. Theoretically, M^2 is an instantaneous concept. A given wavefront at a given instant can be assigned an M^2 value. Experimentally, M^2 is a time-averaged measurement. Ten measurements with intervening stage movements filter changes, and software aperture changes takes time. The International Standards Organization (ISO) standard specifies that a laser must be warmed up for at least 1 h prior to measurement. The hope is that the laser will become stable enough so that measurements taken on different parts of the wave train will yield an

average value approximately equal to the instantaneous M^2 over the entire wave train. If the laser fluctuates in power or mode content during the course of the measurement, the results are less valid. If the laser remains constant during measurement but fluctuates afterward, the results are valid, but not useful. M^2 is thus inappropriate for single-shot lasers and those with short run times or even with slow fluctuations in power or mode content. It is also inappropriate if the measurement is taken under operating conditions different from those in which the system will be used.

4.3. Second-moment beam radius

There are many ways to define or measure the width of a laser beam. M^2 is defined in terms of the second moment of the irradiance. The second moment is calculated by weighting the measured irradiance by the square of the distance from the centroid of the beam:

$$w_x = \sqrt{2\frac{\iint (x-\bar{x})^2 I(x,y)\,\mathrm{d}x\mathrm{d}y}{\iint I(x,y)\,\mathrm{d}x\mathrm{d}y}}. \tag{15}$$

The use of the second moment to measure beam radius creates another theoretical–experimental disconnect. Theoretically, the second moment is very sound and can be applied to a wide variety of beam shapes. Experimentally, it is problematic because small amounts of noise away from the beam are weighted more heavily than the actual signal one is measuring. A number of steps must be taken as a consequence of the choice of second-moment definition of beam radius. The first is due to the fact that CCD cameras return only positive signals. An artificial zero must be chosen such that the second moment of the noise is zero. The second is that no extraneous information should be taken into one's measurement. Extra information only contributes to error. Determination of what constitutes "extra information" will be discussed in the section on noise equivalent aperture.

Another implication of the second moment heavily weighting data from the wings of a beam is that annular beams such as those from unstable resonators are measured to have inordinately large beam radii and suffer in comparison with other lasers. Many high-power laser systems use unstable resonators. Defining beam quality specifications in terms of M^2 for a high-power unstable resonator may give the laser designer an impossible task that will not significantly enhance the ability of the system to accomplish its intended mission. For example, take the case of a beam with M^2 of 1.0, entirely zeroth order. Compare this with the beam of $M^2 = 2$ composed of 95% TEM_{00} mode and 5% $\mathrm{TEM}_{20} + \mathrm{TEM}_{20}$ mode and a beam of $M^2 = 2$ composed of 75% TEM_{00} mode and 25% TEM_{11}, as shown in Fig. 7. The beam in the center has some higher-order noise that will result in 5% of the energy diffracting away from the intended target. The beam on the right also has an M^2 of 2, the

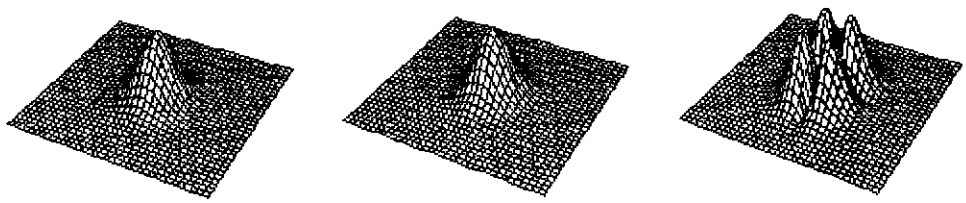

Fig. 7. $M^2 = 1$ beam (left) and $M^2 = 2$ beams (center and right) in the far field.

same as the beam in the center, but it has severe impact on the shape of the focal spot and may result in a loss of up to 25% of the energy of the laser. This discrepancy is a result of the fact that M^2 is fundamentally linked with the second-moment measure of beam radius.

4.4. Noise equivalent aperture

With a beam analysis package of sufficient flexibility, one can experimentally verify that the measured second-moment radius is a strong function of the software aperture put around the data returned by the CCD camera. This is because no matter how carefully one sets the artificial zero, the second moment of the noise will fluctuate slightly around zero. The nonzero second moment of the noise will significantly alter the results of the measurement. In the results shown in Fig. 8, the measured second-moment beam radius for the same beam from a standard, commercial Nd:YAG laser varies from 0.45 to 0.15 mm, depending on what software aperture is chosen, plotted as a fraction of the full-width at $1/e^2$ maximum beam radius.

The discrete nature of the signals returned from digital cameras has some long-reaching effects on the measurement of beam quality. If we take a hypothetical case, shown in Fig. 9, the column averages for an 8-bit camera, the signal returned will fall between 0 and 255. Further assuming that the appropriate artificial "zero" is at a pixel value of 50, this gives a maximum contrast of 205:1. A pixel value of 51 is barely significant. A pixel value of 50 is effectively "zero." Any information taken beyond the point at which the signal is less than 1/205th of the peak is extraneous and can only contain noise. This defines a noise equivalent aperture (NEA), derived in Eq. (16) assuming a TEM_{00} Gaussian, that is equal to ~2.3 beam radii in our example, which must be electronically placed around the CCD image to exclude parts that can only contain noise. This raises its own problem in that one must know the beam radius and the mode content to calculate the NEA in order to measure the beam radius prior to determining M^2, which gives only clues as to the mode content. In practice, this means that the beam radius must be first measured by a non-second-moment method to correctly set the NEA prior to measuring the second-moment beam radius. Unfortunately, the flat portion of the curve in Fig. 8 and the correct NEA do not coincide. If they did, then

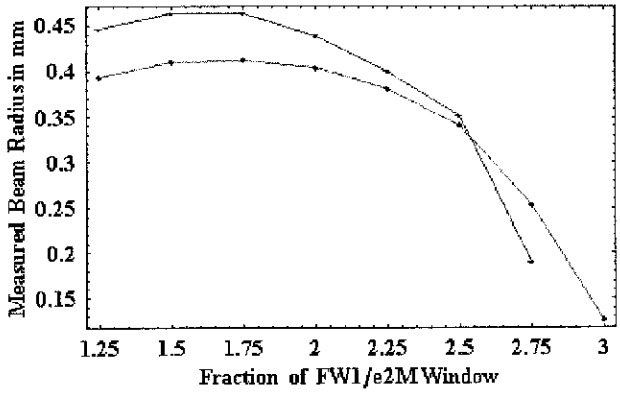

Fig. 8. Measurement of second-moment beam radius as a function of software aperture.

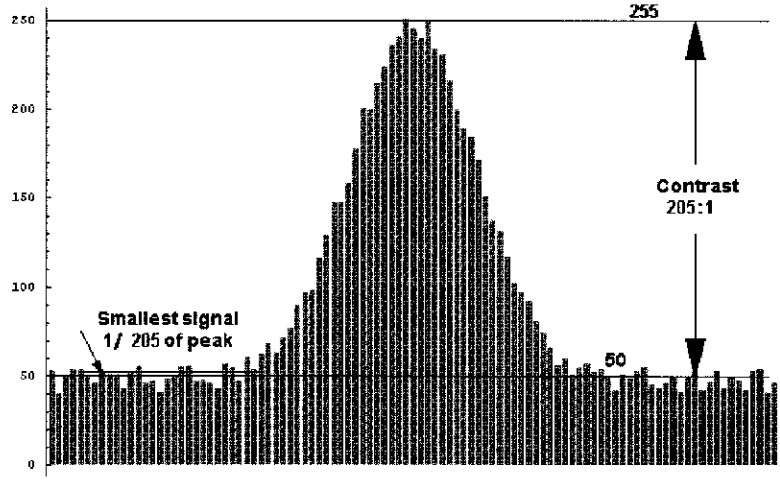

Fig. 9. Image contrast for a hypothetical CCD camera.

small errors in predetermining the beam radius would not matter. The correct NEA is on a sloped portion of the curve, and thus errors in the predetermination of the beam radius result in errors in the measured second-moment beam radius. Note that the NEA thus defined is the largest NEA appropriate. A careful characterization of the dark-current noise of the CCD camera will show a root-mean-square (rms) fluctuation that will likely be greater than the 1-pixel value used here. This will reduce the contrast accordingly. Contrast = (peak signal, artificial zero)/ max[rms noise fluctuation of camera, 1] . Note that frame averaging can be used but does not have quite the effect that might be desired; it is discussed in Appendix A: Derivation of Error Terms in M^2 Measurement.

$$e^{-2\mathrm{NEA}^2/w^2} = \frac{1}{\text{contrast}}, \qquad \mathrm{NEA} = w\sqrt{\ln(205)}$$
$$\mathrm{NEA} = w\sqrt{\ln(\text{contrast})} \qquad = 2.3 \text{ beam radii}, \tag{16}$$

Equation (16) is strictly correct only for a TEM_{00} beam. As higher-order modes enter, the NEA must be reevaluated based on the new mode content. This places the experimenter once again in the quandary of needing to know what he is attempting to measure in order to measure it properly. The measurement of M^2 is thus limited to small perturbations on the basic TEM_{00} profile. The higher M^2 is measured to be, the less meaningful it becomes.

4.5. Irradiance-dependent beam radius

One further implication of the available contrast in a CCD image is that the NEA is a function of the peak signal. If the peak of the image returns a pixel value of 200, the NEA will be different from what it would be if the peak of the image was at 250, due to the change in contrast of the image. Further, a low-contrast image will have a lower signal-to-noise ratio and measure the beam radius with greater error than a high-contrast image. This has relevance to one's filter set. Neutral-density filters are commonly used to attenuate the beam to protect the CCD camera. The ideal case would be a continuous range of attenuation to keep the peak signal near the saturation point of the CCD camera for all measurements to maximize the image contrast and keep the NEA constant. Typically, one has a discrete set

of filter values. The discontinuities in neutral density attenuation provide another source of uncertainty in the final measured value.

4.6. Curve fitting

Once the series of beam radius measurements has been taken, the data must be fitted to a quadratic equation in terms of M^2, the beam waist w_0, and the focal position z_0. Note that there are two forms of this equation, shown as Eqs. (17) and (18). The difference is in how w_0, the beam waist, is interpreted. If w_0 is interpreted as the beam radius of a pure Gaussian TEM_{00} mode embedded in the beam, then Eq. (17) is used. If w_0 is interpreted as the smallest measured beam radius, then Eq. (18) is used. Since the measurement of M^2 deals with measured, not theoretical, beam radii, Eq. (18) is the most common in curve fitting (Ref. 13, p. 8), while Eq. (17) is the most common in the scientific literature[‡]:

$$w^2(z) = M^4 w_0^2 + \left(\frac{\lambda}{\pi w_0}\right)^2 (z - z_0)^2 , \tag{17}$$

$$w^2(z) = w_0^2 + \left(\frac{M^2 \lambda}{\pi w_0}\right)^2 (z - z_0)^2 . \tag{18}$$

The ISO standard does not specify a numerical method but does recommend weighting the data points inversely by the variance of each measurement. This presumes that multiple measurements of the beam radius are taken at each position, which further slows the measurement process. Multidimensional curve fitting is not an easy numerical task, and available numerical methods are not universally robust and cannot work in an entirely automated fashion without review by a knowledgeable human being. It is fairly common for laser technicians to adjust their laser or optic train until their beam profiler gives a decent result. In other words, they adjust the laser to overcome the problems in curve fitting and detection in their beam analyzer—opposite from the ideal case of adjusting the beam analyzer to measure the laser as it is. Decisions must be made on how to weight data points, in what order to fit for the parameters, what initial guess to use, and which numerical method to employ. A Levenberg–Marquardt[9,§] method works well with the following procedure: 1) initial guesses of the smallest measured beam radius and its location for w_0 and z_0 and M^2 calculated using the divergence angle between the smallest and largest measured spot sizes; 2) an unweighted fit on all three parameters, keeping the result for z_0 and using the results for w_0 and M^2 as initial guesses for the next step; 3) a fit on w_0 and M^2 heavily weighted toward points near focus, keeping the result for w_0 and using the result for M^2 as an initial guess for the next step; and 4) an unweighted fit on M^2. While this method may not be the one "best" method, it is vastly superior to any algorithm hidden from the experimenter by the word "proprietary." A known method can be analyzed and improved. A black-box method will forever produce results whose validity is known only to the anonymous developer.

Finally, the method used should provide a means to assign error bars based on the signal-to-noise ratio of the CCD camera, variance in NEA, variance in beam radius measurements,

[‡]Equation (17) is equivalent to Eq. (7) and Eq. (18) is equivalent to Eq. (6) using the definitions of Eq. (8).

[§]Levenberg-Marquardt routines are also implemented in some fitting functions in both *Mathematica* and LabVIEW.

and discontinuities in the filter sets. M^2 is commonly quoted to two or three significant figures since that is what appears on the displays of commercial black-box devices. In fact M^2 is typically accurate to slightly better than one significant figure.

4.7. Summary of noise contributions to measurement of M^2

Noise on each pixel of a CCD camera contributes to uncertainty in measured beam radius caused, which in turn causes uncertainty in measured M^2. The primary sources of noise are as follows:

- Discretization error: A CCD camera takes discrete measurements of a continuous quantity. Measured as 1/resolution of the camera. An 8-bit camera, for example, has a discretization error of 1/256.
- CCD noise: Gaussian dark noise present on all pixels. Measured as the variance of readings around an arbitrary zero. The more pixels one has across the beam, the less of an effect this causes. The variance decreases with the number of measurements.
- Filter error: Discrete neutral-density filters cause deviation from saturation of the camera and alter the NEA. Measured as the minimum percent change in filters. If, for example, the filter set is spaced in tenths of neutral density (ND = 0.1,0.2,0.3,0.4, etc.), then this source of error is ~10%. For continuous means of attenuation, this contribution is zero.
- NEA estimation error: NEA must be estimated by using a non-second-moment method to measure beam radius. To the extent that the alternate method differs from second moment, the NEA was originally set incorrectly. Owing to dark current noise, which creates irradiance-dependent beam radius measurements, iteration to eliminate this source of error is not possible. Measured as a theoretical variance between alternate method and second moment on a noisy beam.
- Laser fluctuation: Measured as a variance. Since one can perceive the beam only via a CCD camera, this is measured the same way as CCD noise, except with the laser on.

Total variance in each beam radius measurement from the above sources is as follows:

$$\sigma^2_{\text{beam}} = \ln(\text{contrast})\left(\frac{A_{\text{pixel}}}{4\,\text{NEA}^2} + 2\pi\,\Delta\text{NEA}\right)\left(\frac{1}{\text{contrast}} + \sigma^2_{\text{dark}}\right) + \sigma^2_{\text{laser}},$$

$$\Delta\text{NEA} = \sigma^2_{\text{method}} + \sqrt{\ln(1 + \Delta\text{filter}^2)}; \qquad \sigma^2_{\text{method}} = \ln\left[1 + \frac{e^2}{2}\left(\sigma^2_{\text{dark}} + \sigma^2_{\text{laser}}\right)\right];$$

$$\sigma^2_{\text{dark}} = \sigma^2_{\text{dark-pixel}}\frac{A_{\text{pixel}}}{\pi\,\text{NEA}^2} \tag{19}$$

NEA = noise equivalent aperture, contrast = maximum contrast available on a particular camera (see Fig. 9), Δfilter = % difference between filters, dark = dark current noise averaged over the entire beam, dark-pixel = dark current noise per pixel, and w = beam radius. The derivation of Eq. (19) is presented in Appendix A.

The uncertainty in measuring beam radii causes an uncertainty in the M^2 result. The variance in M^2 is equal to the average variance in the data:

$$\sigma^2_{M^2} = \frac{1}{N}\sum_{i=1}^{N}\sigma^2_{\text{beam}\,i}. \tag{20}$$

These sources of error are far more significant than commonly reported, largely due to the widespread prevalence of commercial, black-box devices that include no estimation of error. As an example, one of the authors reported the M^2 of a tunable MIR laser source[11] as "less than 2." This was based on the above error analysis and a statistical sample of M^2 measurements. Based on the author's observation and analysis, M^2 measurements more reliable than one decimal place are exceedingly rare. More common is ± 0.3 for measurements between 1.5 and 3. Those who quote $M^2 = 1.651$, for example, were either exceedingly careful on an extraordinarily stable laser or did no error analysis at all. A more honest number might be "M^2 less than 1.8."

4.8. Summary and recommendations for the use of M^2

M^2 is a difficult measurement to take properly and is subject to several theoretical–experimental disconnects. The accuracy of a measurement is vulnerable to signal-to-noise ratio in the CCD camera, fluctuations in contrast ratio, the continuity of the filter set used, and implications of design decisions regarding camera resolution and focusing geometry. M^2 measurements from automated black-box instruments are unreliable under any nonideal conditions. M^2 is inappropriate for single-shot, multimode or annular beams and is most appropriate for laboratory, single-mode lasers with low power fluctuation and long-term stability. As a rough rule of thumb gained by experience, M^2 values above 2 are strongly suspect and M^2 values above 5 have lost most of their meaning. M^2 is best used below $\sim$1.5, at which the beam has only one or two modes in more or less predictable ratios. We recommend that a measure of beam performance directly related to the mission objectives be selected rather than automatically choosing M^2. If M^2 is chosen, then we recommend that identical measurement apparatus and procedures be used to ensure true comparability. We also recommend that black-box devices not be used without appropriate error estimation.

5. M^2: Underlying Assumptions

The preceding section outlined the experimental difficulties in measuring M^2. This section will examine the underlying assumptions M^2 is based on to enable a researcher to make a wiser decision regarding the use of M^2 for a particular application. Some measures of beam quality, such as brightness or PIB (depending on the way the "bucket" size is chosen), are absolute measurements; others, such as Strehl ratio and M^2, are comparisons. In the case of M^2, there is an implied comparison of a beam with an embedded beam composed entirely of zero-order Hermite–Gaussian profile. Hermite–Gaussian functions are important because they are self-consistent solutions of the paraxial Helmholtz equation and represent shapes that will propagate indefinitely. The mathematical form of the magnitude of normalized Hermite–Gaussian functions is [Ref. 12, Chap. 16.4, Eqs. (48)–(60)][¶]

$$u_n[x, w(z)] = \left(\frac{1}{\pi}\right)^{1/4} \sqrt{\frac{1}{2^n n! w(z)}} H_n\left[\frac{x}{w(z)}\right] e^{-[x^2/2w^2(z)]}, \tag{21}$$

[¶]Note that we use a slightly different normalization from those in the latter equations and that the form quoted does not include the complex factors. The full form is used in computation but was not cited in this paper for simplicity. See also Ref. 17, http://mathworld.wolfram.com/HermitePolynomial.html, Eq. (39).

where $H_n(x)$ are standard Hermite polynomials (Ref. 15, Chap. 24)** and $w^2(z) = w_0^2(1 + z^2/z_r^2)$ with $z_r =$ Rayleigh range $= \pi w_0^2/\lambda$. The normalized $u_n(x)$ obeys the orthogonality relation

$$\int_{\infty}^{\infty} u_n(x, w)u_m(x, w)dx = \delta_{nm}. \tag{22}$$

The orthonormal relation allows the representation of any arbitrary field profile $E(x)$ in terms of a series expansion in terms of these functions following the standard practices of linear algebra:

$$E(x) = \sum c_n u_n(x, w), \tag{23}$$

where the field coefficients c_n can be determined by the vector projection of u_n on $E(x)$:

$$c_n = \int_{-\infty}^{\infty} E(x)u_n(x, w)dx \tag{24}$$

in one dimension. In two dimensions, the c_{nm} will be the product of an integral in x and another in y. Because the modes are orthonormal, they can be used as a basis set for the expansions of Eq. (23). The set of modes u_n is only for complete, orthogonal basis functions with respect to modes based on the same mode radius w. There is no unique set of mode coefficients that will describe an arbitrary beam shape without prior determination of beam radius (Ref. 12, Chap. 16.4, p. 646).

5.1. The "best" modal decomposition

Equation (24) shows that to determine the mode content of a given arbitrary field, one must first determine the characteristic beam radius w. Restricting our discussion to the near field at the plane $z = 0$, that means we must first determine w_0, commonly called the beam waist. The ability to represent an arbitrary shape mathematically with Hermite–Gaussian modes is unaffected by our choice. There are many reasonable bases on which to determine the characteristic beam radius w_0. Each choice of w_0 will have a unique set of modes associated with it. All choices of w_0 will lead to a series representation of the field. All will propagate mathematically to the far field. No physically measurable properties of the beam will be affected by our choice of w_0. We must therefore regard the modal composition of a given beam as not entirely a property of the beam but also of our mathematics. It is the case that w_0 is the second-moment radius of the TEM_{00} order Hermite–Gaussian mode of a given expansion. This will be true no matter what definition of beam radius for the entire beam we use. The second moment of the lowest mode in an expansion will not be the second moment of the entire beam except in the case of a pure, TEM_{00} mode, and so a Hermite–Gaussian expansion only weakly suggests, but does not demand, that we use a second-moment definition of beam radius for the entire beam. It remains, therefore, to make a good determination of w_0 based on the physics of the situation. As an illustrative case, we will examine the Hermite–Gaussian representation of a beam typical to high-energy lasers. Figure 10 shows a square annulus flattop field shape, such as one might obtain from an ideal unstable resonator. In this case the field magnitude was arbitrarily set to 1, the beam has a width of 5 cm, and the inner "hole" has a width of 2 cm.

**Also see Ref. 17, http://mathworld.wolfram.com/HermitePolynomial.html.

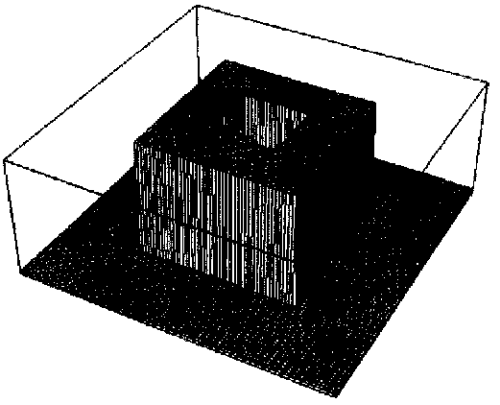

Fig. 10. Square annulus flattop.

Figure 11 shows several Hermite–Gaussian modes that might be used as a basis to represent the beam of Fig. 10. The particular modes TEM_{00}, TEM_{11}, $TEM_{02} + TEM_{20}$, and $TEM_{10} + TEM_{01}$ were chosen for this figure as they happen to be the modes with the highest field coefficients by several orders of magnitude for most choices of characteristic beam radius w_0.

Figure 12 shows the field coefficients for the modes of Fig. 11 as part of a Hermite–Gaussian modal decomposition of the square annulus flattop beam of Fig. 10 as a function of characteristic beam radius. For reference, the vertical lines in each of the four graphs of Fig. 12 show the second-moment beam radius, which is the choice that M^2 is based

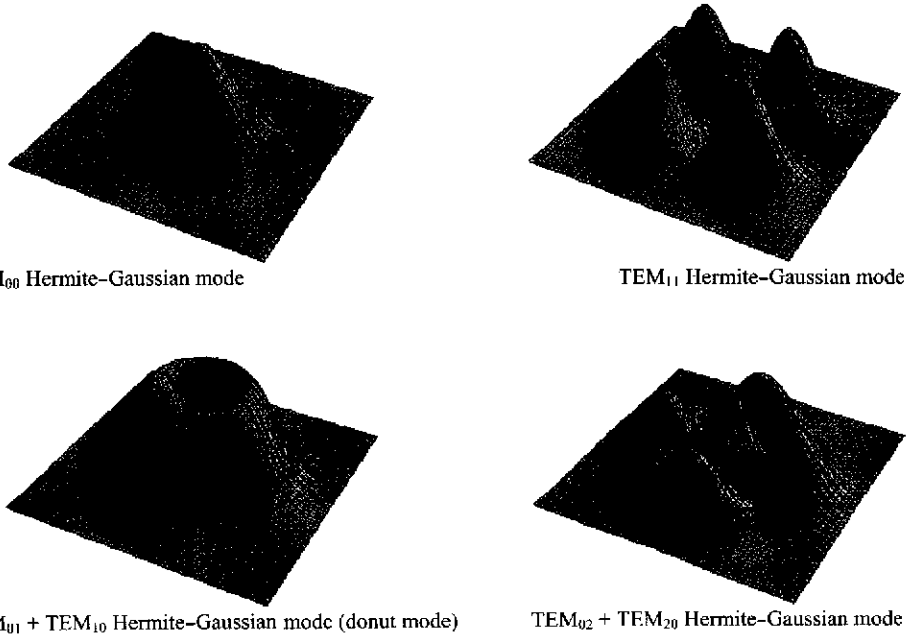

Fig. 11. Hermite–Gaussian modes.

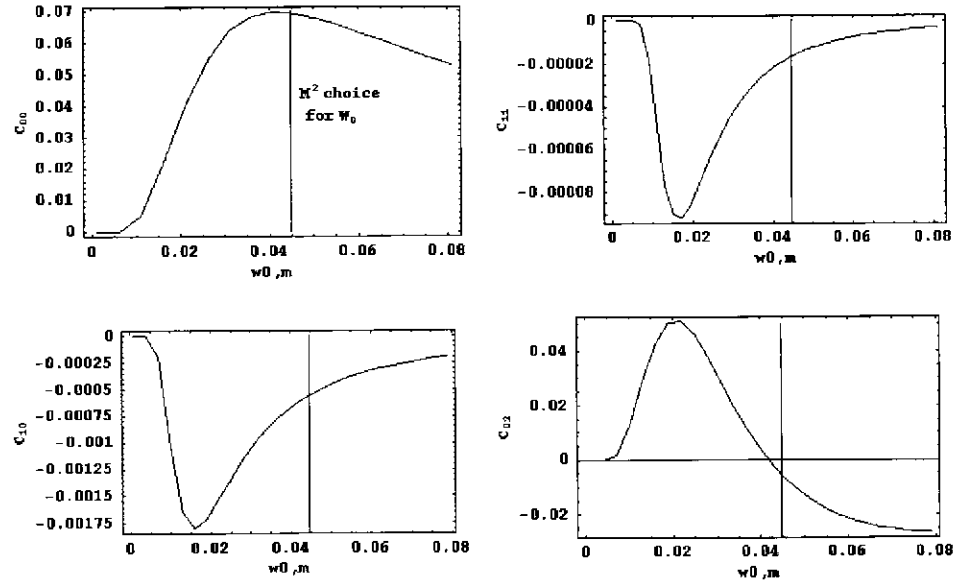

Fig. 12. Field coefficients of a square annulus flattop for several Hermite–Gaussian modes as a function of choice of characteristic beam radius w_0.

on. These coefficients were calculated for the $z = 0$ plane but, as a test, the square annulus flattop was Fourier propagated to a distance $z = 3{,}000$ m and the same modal decomposition was performed. Within the limits of numerical precision, the shapes of the curves in Fig. 12 were the same at $z = 0$ and 3,000 m. The point of this is to realize that modal composition is not entirely a property of the beam, but a product of the researcher's mathematics and choice of characteristic beam radius. Nothing physically measurable about the beam, moments, energy, fluence, propagation vectors, focal planes, etc., changes when one chooses a characteristic radius w_0 and its accompanying modal decomposition. There are four possibilities for determining the proper characteristic beam radius w_0 from the results shown: first, the peak of the curve of C_{00} vs. w_0; second, the peak of the curve of the coefficient containing the most energy, in this case C_{02} vs. w_0 ($C_{02}^2 + C_{20}^2$ has a higher peak value than C_{00}^2); third, the physical size of the aperture in the near field; and last, we could choose the second-moment or some other measure of beam radius in the near field. For comparison, Fig. 13 shows the mode composition of a TEM_{00} (at a particular radius) Gaussian beam as a function of choice of characteristic radius. As a numerical exercise, this is approaching the

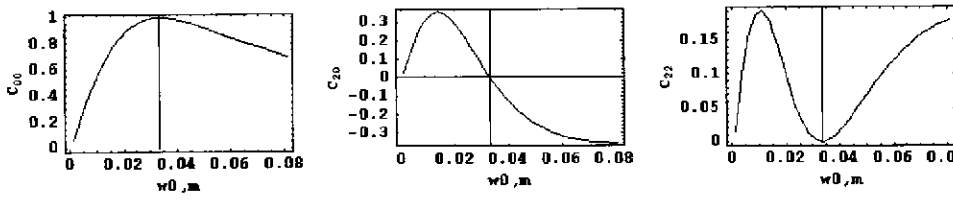

Fig. 13. Field coefficients of a TEM_{00} Gaussian beam for several Hermite–Gaussian modes as a function of choice of characteristic beam radius w_0.

absurd since we already "know" the mode content of this beam. Experimentally, however, one never "knows" the mode content in advance and must determine it from the data. In this case, we digitized a zero-order Gaussian beam with a second-moment radius of 3.33 cm from a hard rectangular aperture 20 cm wide. Depending on choice of characteristic beam radius, one can have a significant contribution from the TEM_{20} and the TEM_{22} mode as well. The feature that truly forces us to choose the "correct" radius is the fact that the C_{00} curve reaches ~1.0 at a certain radius at which all other modes drop to zero, which corresponds to the second-moment radius, indicated by the vertical lines. This is a luxury we do not have with the example of Fig. 10, and there is nothing that will tell us the "correct" characteristic radius w_0 of a square annulus flattop.

5.2. M^2 and "times diffraction limit"

Begin with Eq. (23), apply the formula for second-moment waist, and use the identity[17,††]

$$\int_{-\infty}^{\infty} x^2 u_n\left(\frac{x}{w_0}\right) u_m\left(\frac{x}{w_0}\right) = w_0^2(2n+1)\delta_{nm},$$

$$\begin{aligned} W_x^2 &= 2\int x^2 |E(x)|^2 \mathrm{d}x \\ &= \int x^2 \sum c_n c_m u_n\left(\frac{x}{w_0}\right) u_m\left(\frac{x}{w_0}\right) \mathrm{d}x \\ &= \sum c_n c_m \int x^2 \sum u_n\left(\frac{x}{w_0}\right) u_m\left(\frac{x}{w_0}\right) \mathrm{d}x \\ &= \sum c_n^2 (2n+1) w_0^2, \\ W_x^2 &= M^2 w_0^2. \end{aligned} \tag{25}$$

Thus, M^2 has an easy formula in terms of Hermite–Gaussian mode composition[‡‡]:

$$M^2 = \sum c_n^2 (2n+1). \tag{26}$$

The last line of Eq. (25) also forms the general relationship for the various "times diffraction limited" measures. A measured beam area in the far field W_x^2 is compared with an idealized or hypothesized beam area w_0^2, which is predicted from a measurement of beam area in the near field. The ratio between the two is the "times diffraction limit" of that beam. This is illustrated in the first line of Eq. (27):

$$\begin{aligned} W^2 &= \mathrm{XDL} \times w_0^2 \\ w_0 w_{\mathrm{near}} &= \frac{\lambda z}{\pi} \Rightarrow \\ W^2 &= \mathrm{XDL} \times \left(\frac{\lambda z}{\pi w_{\mathrm{near}}}\right)^2 \end{aligned} \tag{27}$$

where W is the measured beam radius in the far field, w_{near} is the measured beam radius in the near field, and w_0 is the ideal waist in the far field; z represents the focal length of the optics

[††] See http://mathworld.wolfram.com/HermitePolynomial.html, Eq. (43).

[‡‡] Reference 14, Eq. (5).

Table 3. "Times diffraction limited" numbers for square annulus flattop beam of Fig. 10 vs. various methods of measuring beam radius in near and far fields (3 km)

	$W_{\text{far field}}\downarrow$	C_{00} max[a]	Largest mode maximum[b]	Physical aperture	Second-moment radius
$W_{\text{near field}}\rightarrow$		3.58 cm	2.01 cm	2.5 cm	4.47 cm
C_{00} max[a]	2.84 cm	1	0.32	0.49	1.56
Largest mode maximum[b]	5.05 cm	3.16	1	1.54	5.0
Second-moment radius	5.37 cm	3.57	1.12	1.75	5.55
Radius of central lobe[c]	3.03 cm	1.14	0.36	0.56	1.79

[a]See Fig. 12.
[b]In the case of the square annulus flattop of Fig. 10, this is the (2,0) + (0,2) mode. See Fig. 12.
[c]Measured by examination of the far-field pattern after Fourier propagation to the far field.

and λ the wavelength. The second line of Eq. (27) comes from Ref. 12 [Chap. 17.1, Eq. (13), and Chap. 17.2, Eq. (24)] and can be used for illustrative purposes as an approximation to more formal means of mathematically propagating an aperture or radius to a focal plane. Each "times diffraction limited" measure is thus tied to one beam radius measurement in the near field and another in the far field. If the measure of beam radius is chosen to be the second moment in both the near and far fields, then the "times diffraction limit" is equal to M^2. Otherwise, it is not. It is easy to imagine that the choice of measuring technique can give rise to a number of "times diffraction limited" numbers that could describe a given beam.

Table 3 shows the wide variety of "times diffraction limited" numbers we can come up with depending on the method chosen to measure beam radius. We can argue that the square annulus flattop beam of Fig. 10 is anywhere from 0.32 to 5.55 "times diffraction limited." How can anything be better than the diffraction limit? Choose a suitable basis of comparison and one certainly can have a "better than diffraction limited" beam! Table 4 shows a similar analysis for a zero-order Gaussian beam with a near-field, second-moment radius of 2 cm.

Through suitable choice of method, we can show that a zero-order Gaussian beam is from 1 to 47.6 "times diffraction limited." The point of these examples is not to argue that a zero-order beam is better or worse than a square annulus flattop or that any of the particular methods of determining beam radius ought to be adopted, but to emphasize that "times diffraction limited" is an absolutely meaningless number unless the methods of determining beam radius are defined along with it and that "times diffraction limit" numbers generated by different methods of beam radius are not comparable. Different standards for measuring beam radius are not uncommon, and some are unconscious. In some cases, various measures of beam radius are identical for certain kinds of beams. There is no difference between half-width-half-maximum and half-width-$1/e^2$-maximum for a

Table 4. "Times diffraction limited" numbers for a zero-order Gaussian beam vs. various methods of measuring beam radius in near and far fields (3 km)

	W (far field)↓	C_{00} max or second-moment radius	Physical aperture[a]
W_{near} →		2 cm	6 cm
C_{00} max or second-moment radius	5.08 cm	1	9
NEA[b]	11.68 cm	5.3	47.6

[a]We follow the 99% criteria (Ref. 12, Chap. 17.1), which corresponds to a physical aperture three times the second-moment radius.
[b]See Eq. (16) and Fig. 9. This is chosen in lieu of the first lobe diameter used in the preceding example.

flattop beam in the aperture plane, but these measures give greatly different results in the target plane.

All hard-edged beams, such as our example square annulus flattop, have diffraction ripples extending to infinity in the focal plane. Since there is no such thing as an infinite detector, all second-moment measurements on any hard-edged beam are actually limited by the NEA of the system, thus converting the measure of M^2 to a closely related "times diffraction limited" number. It is also common to measure far-field beam radius by width of the principle lobe. This is a perfectly reasonable measure of beam radius but converts the measurement from M^2 to one of the other "times diffraction limited" numbers. Various beam radius standards are listed in Tables 3 and 4 and are associated with M^2, Strehl ratio, brightness, and "times diffraction limited." Often, these measures conflict and are not equivalent.

6. Summary, Conclusion, and Recommendations

We have shown the experimental issues related to measuring M^2 along with the theoretical underpinnings of M^2. We showed that the term "times diffraction limited" is meaningless without reference to a particular method of determining beam radius in both near and far fields. We showed that the theoretical issues that are clear for nearly zero-order Gaussian modes become very confused when dealing with beams typical of high-power unstable resonators. We therefore conclude that M^2 is an appropriate measure for nearly zero-order continuous-wave (CW) Gaussian stable resonator beams but is not an appropriate measure for pulsed, single-shot or unstable resonator beams, nor can stable resonator beams be compared on any kind of equal footing with unstable resonator beams in terms of a "times diffraction limited" number. We urge that the term "times the diffraction limit" be discarded in any context in which laser systems are compared against one another.

We also showed that the various common measures of beam quality do not measure the same aspects of the beam and that they should not become conjoined in our thought or communication. M^2, Strehl ratio, "times diffraction limit," brightness, and PIB do not measure the same thing except under very idealized conditions.

6.1. Mission consistent measure of optical beam quality

It is clear that the high-power laser community needs a measure of laser quality that is consistent with our mission. An ideal metric would satisfy at least the following criteria:

- Single-shot measurement
- Taken at either near or far field but not both
- Not heavily dependent on noise and detector fluctuation
- Not easily subject to obfuscation or argument
- Compare stable and unstable resonators on equal footing
- Easily made uniform for comparison between different systems
- Relate directly to the mission requirements

M^2 meets none of these criteria. The Strehl ratio meets several of the criteria, as does brightness. The last criterion, of relating to the mission requirements, depends on the mission. If one's mission is to have a very high peak irradiance without concern for anything else, then the Strehl ratio is appropriate. If one's mission is to illuminate a solid angle (as in LIDAR and communications), then brightness is appropriate. In the high-power laser directed energy community, our mission is most often to deliver power to a target. That is, there is a circle of some radius and we wish to get as much power in that circle as possible for a given laser source. Aside from the matter of the definition of the size of the target circle, that is the definition of the beam quality measure known as power in the bucket (PIB). Even recognizing the strong correlation between power delivery applications and PIB, there is still the problem of the proper basis of comparison. PIB is typically not quoted as a single number but shown as a curve. Figure 14 shows some sample PIB curves for the beams shown in Fig. 15. The encircled power (arbitrary units) is shown as a function of the far-field radius. In this case, all three curves were generated from beams proceeding from a 3-cm-diameter aperture without focusing or directing optics to a distance of 3 km from the aperture. The middle, thick curve represents the far-field PIB of an $M^2 = 1.2$ beam composed of zero- and first-order Hermite–Gaussian modes truncated at the $\pi/2$ w point in the near field. The lower curve is for a donut mode from the same aperture and having the same energy content as the low-order Gaussian. The upper curve is for a flattop beam from

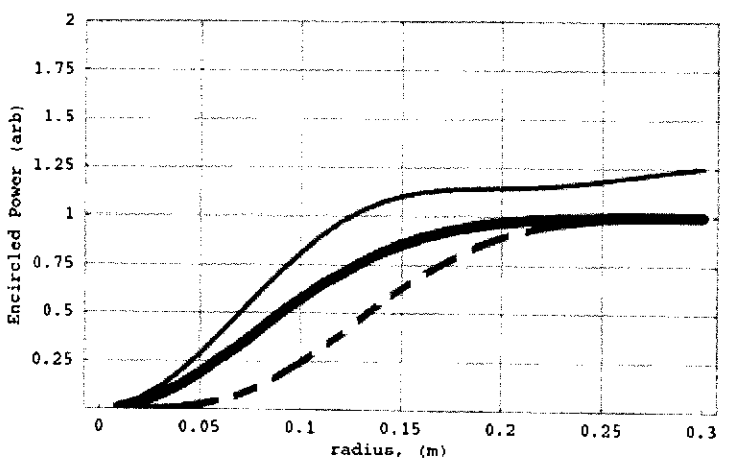

Fig. 14. PIB curves.

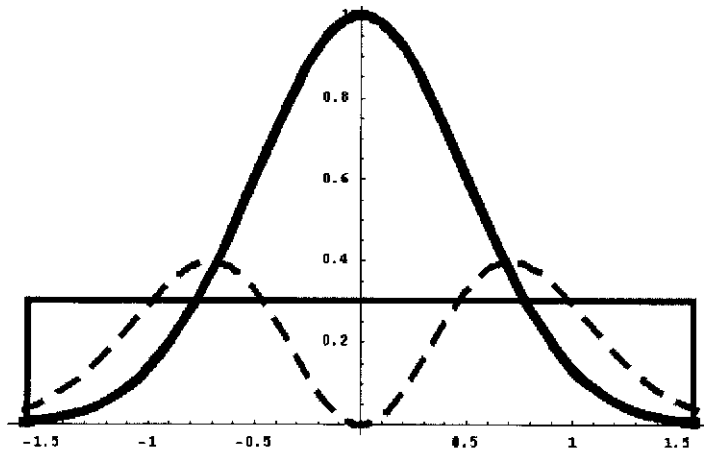

Fig. 15. Flattop and Gaussian beam comparison.

the same size aperture as the low-order beam but with 30% of its peak intensity so that it has 50% more energy content. Both stable resonator beams had $w_{00} = 1$ cm in the near field. A two-dimensional representation of this is found in Fig. 11. This basis of comparison was chosen arbitrarily. A flattop beam with the same peak intensity as a zero-order Gaussian truncated at the $\pi/2$ w points has five times its energy content and would be inconvenient to show on the same chart. This does emphasize that a TEM_{00} mode is very inefficient at energy extraction from most gain media. If a flattop beam results in a loss of 10% diffraction efficiency compared to a TEM_{00} to the far field but delivers five times more power, it is hard to argue that a TEM_{00} is the "ideal" beam shape.

The standard PIB graph is an ideal place to construct a mission requirements curve (MRC). First, we notice that lines of constant average power run horizontally and lines of constant radius run vertically. If the physical effect our mission was dependent on were a function of average power only, we might stop there. Typically, physical effects are functions of intensity or field on target. Figure 16 shows the PIB curves of Fig. 14 with lines of constant intensity or field drawn in. On a PIB curve, these lines are quadratic in shape. For a given mission, the minimum radius is decided by an analysis of jitter and atmospheric aberration. The maximum radius is decided by the target size less jitter. The physics of the laser–target interaction determines the minimum intensity or field and a MRC is generated, as in Fig. 17. Note at this point how all arguments and speculations about proper basis of comparison have been done away with. We compare each beam not with a reference ideal but with the mission requirements based on the physics of the application. Any disputes will be over the mission requirements, not the beam quality standard.

One unsatisfactory element is left in the comparison of a PIB curve to a MRC curve. That is, we would like the performance of the system to be reducible to a single number so that we can definitely state that laser system A is better than system B and by how much. One way to do this is to take the area of overlap between a laser's PIB curve and the MRC to generate a "mission compliant area" (MCA). This is notionally shown in Fig. 18. The two shaded regions show the MCAs for the flattop beam and the Gaussian beam. In the case chosen, assuming a 2-kW Gaussian beam, the flattop beam (infinite M^2) has an MCA of ~50 W/m while the $M^2 = 1.2$ beam has an MCA of 12.75 W/m. In the case chosen, we can conclude that the flattop beam is approximately four times better a beam than a low-order

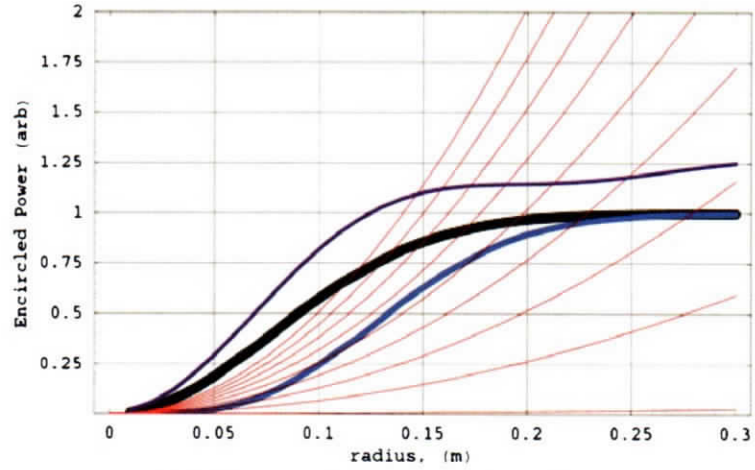

Fig. 16. PIB curve with intensity isolines.

Gaussian for the mission envisioned based on its overall greater window of operation and the power by which it meets or exceeds mission requirements. The donut mode beam, in spite of a relatively good beam quality of 2.0, has no MCA for this example.

It will be observed that the net effect of basing a laser beam standard on a previously determined mission requirements curve will displace the center of argument from the standard to the mission. Whereas we now argue about both the meaning and propriety of a laser beam quality measure and the mission requirements, a mission-related standard will ensure that our deliberations concern the mission requirements. The purpose of a laser beam quality standard is to ensure that the laser system will accomplish the mission if the standard is met. Our confusion, deliberations, and debates need to center on what the mission requirements are. Once that is over, the standard will tell us whether a given system will meet those requirements.

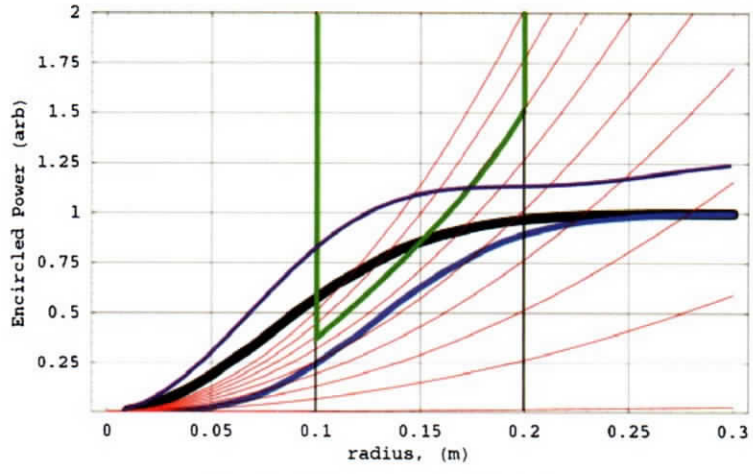

Fig. 17. PIB curve sample MRC.

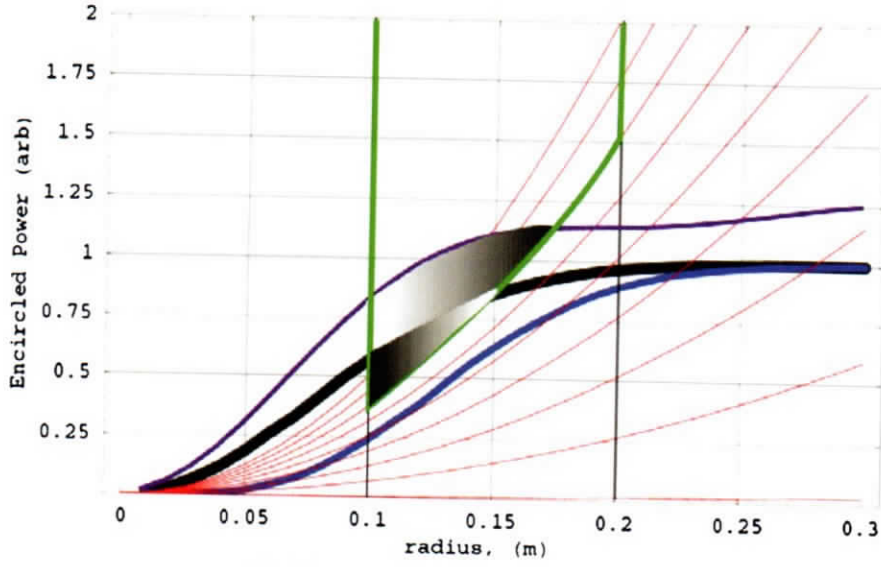

Fig. 18. MCA for two sample beams.

We realize that each service and mission may have different standards and criteria for constructing a MRC and believe that these differences are appropriate. It may become valuable for each service to determine a small number of standard test missions for general use prior to determination of exact requirements or for generic evaluation of new technologies. We have recommended only a general approach to constructing a MRC. The actual approach, determined for each mission, will include information from the laser itself, beam director, atmospheric propagation, and target interaction.

7. Appendix A: Derivation of Error Terms in M^2 Measurement

Equation (19) represents a fractional or percent variance on each beam radius measurement. This section shows the derivation of its terms except for the laser fluctuation and filter error, which must be measured directly.

7.1. Discretization and CCD noise

Take the second moment of irradiance plus some discretization error equal to 1/resolution of the camera (1/256 for an 8-bit camera). In analyzing the error, we remember that the error occurs on each pixel independently and that therefore the variances add in quadrature, which means they can be integrated as if the error were a constant under the integral:

$$\int x^2\left(\hat{I} + \frac{1}{\text{contrast } N_{\text{pixels}}\ \text{NEA}^2}\right) r\,\mathrm{d}r\,\mathrm{d}\theta = \int x^2 \hat{I} r\,\mathrm{d}r\,\mathrm{d}\theta + \int x^2 \frac{1}{w^2\ \text{contrast } N_{\text{pixels}}\ \text{NEA}^2} r\,\mathrm{d}r\,\mathrm{d}\theta. \tag{28}$$

The limits of integration are from the centroid of the beam out to the noise equivalent aperture (NEA):

$$\sigma^2_{\text{discretization}} = \int_0^{\text{NEA},2\pi} x^2 \frac{1}{w^2\ \text{contrast}\ N_{\text{pixels}}\ \text{NEA}^2} r\,dr\,d\theta$$

$$= \frac{\pi\,\text{NEA}^4 A_{\text{pixel}}}{4w^2\ \text{contrast}\ \pi\,\text{NEA}^2\ \text{NEA}^2} = \frac{\ln(\text{contrast}) A_{\text{pixel}}}{4\text{contrast}\ \text{NEA}^2}. \tag{29}$$

Likewise, the contribution to beam measurement variance due to dark current noise is

$$\sigma^2_{\text{dark-beam}} = \int_0^{\text{NEA},2\pi} x^2 \frac{\sigma^2_{\text{dark}}}{N_{\text{pixels}}\ \text{NEA}^2} r\,dr\,d\theta = \frac{\ln(\text{contrast}) A_{\text{pixel}}\ \sigma^2_{\text{dark}}}{4\ \text{NEA}^2}.$$

The number of pixels involved in a given measurement $N_{\text{pixel}} = \pi\text{NEA}^2/A_{\text{pixel}}$. The ratio of w to NEA is ln(contrast). The error term in Eq. (28) is normalized with respect to the beam radius and the radial variable.

7.2. NEA estimation error

Uncertainty in the estimation of the NEA contributes to the overall error in the measurement in several ways in Eq. (19). The error in NEA estimation due to filter error, ΔNEA, is calculated as follows, beginning with Eq. (16):

$$\frac{\text{NEA}}{w} = \sqrt{\ln(\text{contrast})},$$

$$\text{NEA}' + \Delta\text{NEA} = \sigma^2_{\text{method}} + \sqrt{\ln[\text{contrast}(1 + \Delta\text{filter}^2)]},$$

$$\Delta\text{NEA} = \sigma^2_{\text{method}} + \sqrt{\ln(1 + \Delta\text{filter}^2)}.$$

The variance in NEA estimation due to the use of a non-second-moment method depends, of course, on the alternate method chosen. The author uses the full-width at the $1/e^2$ points (FW1/e^2 M). The FW1/e^2 M is that point on a Gaussian where

$$e^{-2(x/w)^2} + \sigma^2_{\text{dark}} + \sigma^2_{\text{laser}} = e^{-2};$$

$$\sigma^2_{\text{method}} = -\frac{1}{2}\ln\left(1 - \frac{\sigma^2_{\text{dark}} + \sigma^2_{\text{laser}}}{e^{-2}}\right),$$

$$\sigma^2_{\text{method}} \approx \ln\left[1 - \frac{e^2}{2}\left(\sigma^2_{\text{dark}} + \sigma^2_{\text{laser}}\right)\right].$$

7.3. Effect of frame averaging on dark current noise

It may be hoped that averaging over several shots would reduce the effect of dark current noise, but such is not the case. We begin by noting that when two Gaussian noise distributions are added together, the resultant distribution is the convolution of the originals. The convolution of Gaussians of width σ increases as the square root of the number of shots $\sqrt{n}$.

To distinguish between averaged distributions and single-shot distributions, the subscript 1 will be used to indicate a single shot quantity:

$$\int_{-\infty}^{\infty} e^{-\alpha^2/\sigma^2} e^{-(\alpha-x)^2/\sigma^2} d\alpha = \sigma\sqrt{\frac{\pi}{2}} e^{-x^2/2\sigma^2}.$$

Thus, the $\sigma_{\text{dark-pixel}}$ increases as the $\sqrt{n}$, $\sigma_{\text{dark-pixel}} = \sqrt{n}\,\sigma_{\text{dark-pixel},1}$. This, in turn, affects the contrast, which varies as $1/\sqrt{n}$, $\text{contrast} = \text{contrast}1/\sqrt{n}$. The contrast dictates the NEA, which will expand as the fourth root of the number of shots, $\text{NEA} = w\sqrt{[\ln(\text{contrast}1/\sqrt{n})]}$. Finally, the number of pixels involved in a given measurement N_{pixels} increases as the NEA^2:

$$N_{\text{pixels}} = \frac{\pi \text{NEA}^2}{A_{\text{pixel}}} = \frac{\pi w^2 \ln(\text{contrast}1/\sqrt{n})}{A_{\text{pixel}}}.$$

We return to Eq. (29) and make the above substitutions:

$$\sigma^2_{\text{discretization}} = \frac{\pi \text{NEA}^2}{4\text{contrast}N_{\text{pixels}}} = \frac{A_{\text{pixel}}\sqrt{n}}{4\text{contrast}1}.$$

The error in the second-moment waist caused by dark current noise is thus

$$\frac{\Delta w}{w} = \sqrt{\frac{A_{\text{pixel}}\sqrt{n}}{2\text{contrast}1}}.$$

Whether it is advisable to average over many shots will depend on the noise characteristics of the laser being measured. Fortunately, the fourth root of number of shots is a slowly increasing function and the benefit of increasing the number of shots is likely to outweigh the increased error to dark current noise at least up to a point.

7.4. Sample calculation of error in M^2

Take the following sample values[§§]:

- 8-bit CCD camera with the pixel area of 12 μm $\times$ 10 μm = 120 μm^2
- Contrast = 1:200
- Dark current noise of 10% rms ($\sigma^2 = 0.01$)
- Laser noise of 10% rms ($\sigma^2 = 0.01$)
- Filter increment of 10%
- A beam radius of 20 μm at focus, for an NEA of 46 μm
- In the far field, a 1-mm beam radius for an NEA of 2.3 mm

Putting these values into Eq. (19) yields a variance of $(28\%)^2$ at focus and $(27\%)^2$ in the far field. The average variance will be approximately (27%).[2] This means that if the beam analyzer returns a result of 1.5, we have a standard deviation of ± 0.4, and so the result should be quoted as 1.5 ± 0.4. Alternately, $M^2 < 1.9$ would be appropriate. In publishing the results on the system these figures came from, we reported an M^2 of "less than 2."

[§§]These are actual values from the Fiber Pumped Optical Parametric Oscillator of Ref. 11.

8. Appendix B: Alternate M^2 Method: Knife Edge

Owing to the prevalence of alternate methods of M^2 measurement, some are included in a brief appendix of the ISO standard. This section will rigorously examine the use of one of these alternate methods. The ISO standard states that "at least for several cases there exists a correlation" between the use of a knife-edge measurement and the use of a CCD camera. The knife edge refers to the practice of scanning a surface such as a razor blade across the beam, the result being the one-dimensional integral of the signal one would see from a two-dimensional CCD camera. The ISO standard recommends scanning the knife until 16% of the energy reaches the detector and then to the 84% point and measuring the distance between them. A factor is then multiplied by this distance to calculate the "second-moment" beam radius. Perhaps this method makes some sense when taking measurements by eye and hand, but when constructing an automated system it raises some serious issues. First, every measurement taken by an automated system takes time, as does every stage motion. It may take a dozen measurements and a root-finding algorithm to locate the 16% point and the 84% point. Why throw away all the extra data? The second issue is that "the at least for several cases" actually means "very nearly single mode beams."

8.1. True second-moment knife-edge method

There is no need to attempt to approximate the second-moment beam radius using only two data points. Judicious use of the calculus method of integration by parts allows the derivation of a convenient formula that will use all the knife-edge data taken:

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 I(x,y)\,\mathrm{d}y\,\mathrm{d}x &\approx \int_{-a}^{a}\int_{-a}^{a} x^2 I(x,y)\,\mathrm{d}y\,\mathrm{d}x \\ &= a^2\left\{\int_{-a}^{a}\left[\int_{-a}^{a} x^2 I(x,y)\,\mathrm{d}y\right]\mathrm{d}x\right. \\ &\quad \left.-\int_{-a}^{a}\left[\int_{-a}^{-a} x^2 I(x,y)\,\mathrm{d}y\right]\mathrm{d}x\right\} \\ &\quad -2\int_{-a}^{a} x\int_{-a}^{x}\left[\int_{-a}^{a} I(x,y)\right]\mathrm{d}x'\mathrm{d}x \\ &= a^2[k(a)-k(-a)]-2\int_{-a}^{a} xk(x)dx. \end{aligned} \tag{30}$$

In the case of N discrete knife-edge measurements, where

$$k_i = \int_{-a}^{a} I(x_i,y)\,\mathrm{d}y, \qquad i \in [1\ldots N],$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 I(x,y)\mathrm{d}y\mathrm{d}x \approx a^2(k_N-k_1)-2\sum_{i=1}^{N-1}\frac{a}{2N}(x_i k_i + x_{i+1}k_{i+1})$$

can be derived with the trapezoid rule.

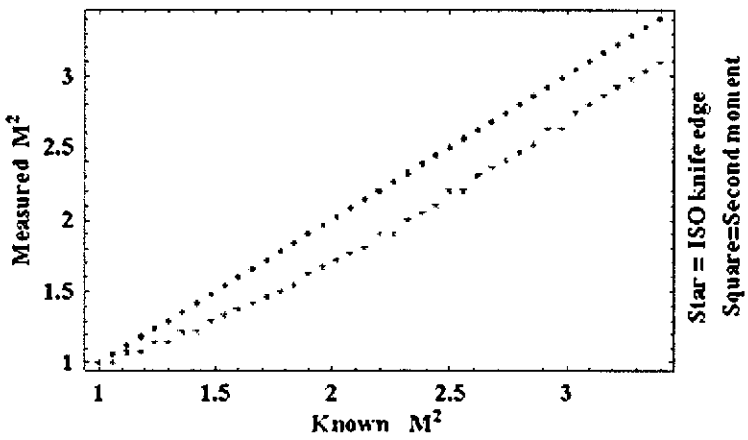

Fig. 19. Comparison of ISO 16-84 vs. second-moment knife edge.

8.2. Comparison of the ISO 16-84 method with second-moment knife-edge method

During the following discussion we will refer to the ISO recommended method of using the 16 and 84% points as the ISO 16-84 method. The method using the discrete version of Eq. (30) will be referred to as the second-moment knife-edge method. We examined a test case using a modeled beam of known mode composition and then a simulated knife-edge measurement, which will be processed with both the ISO 16-84 and the second-moment knife-edge methods to determine whether they really are equivalent.

For the test case, we used a beam composed of zeroth-, first-, and fourth-order modes. The M^2 can be easily calculated[13] for comparison by $M^2 = \sum c_n^2(2n+1)$, where the c_n^2 are the energy fractions of each mode. Figure 19 plots the M^2 as measured by the ISO 16-84 method and by the second-moment knife-edge vs. the calculated M^2 for the test case mentioned. The second-moment method used the numerical curve-fitting methods outlined in the curve-fitting section above. It can be seen that the second-moment knife edge gives the known M^2 while the ISO 16-84 method fails to do so, though it does asymptotically approach the proper slope.

Within numerical precision the ISO 16-84 method consistently undermeasures M^2 and thus fools the researcher into thinking that he has a better beam than he actually does. If the method is not specified in a beam quality standard, then the advantage in contract selection will go to the system that uses the alternate method that erroneously reports the lowest M^2 value.

9. Appendix C: M^2 and Strehl

The various measures of beam quality all converge for a pure zero-order Gaussian beam. This fact has led some to conclude that the different measures of beam quality are equivalent. One common version of this is to equate M^2 and Strehl^{-2}. The Strehl ratio is the ratio of the peak value of a beam vs. the theoretical peak value in the absence of any distortions. It is commonly used in imaging, in which the cause of deviation from zeroth-order Gaussian is due to atmospheric distortion or lens aberration. In laser propagation, it may also be due to

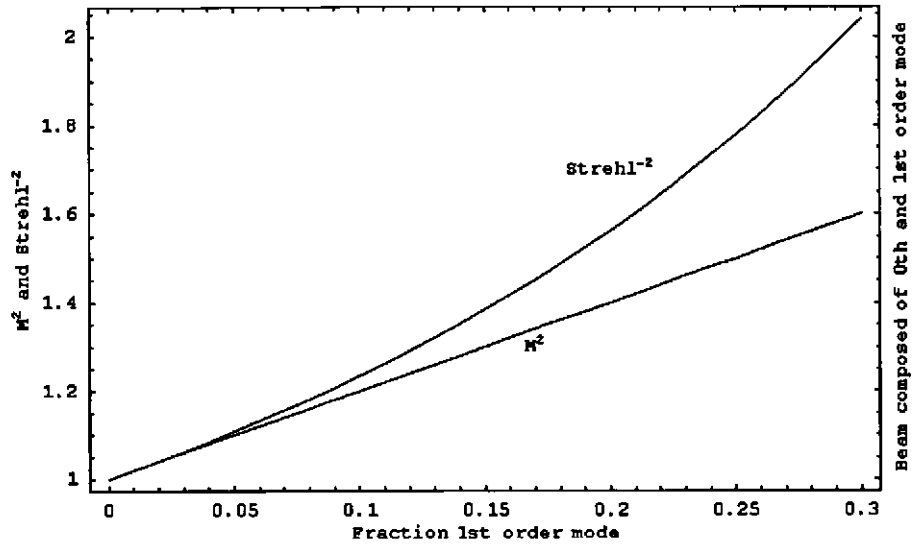

Fig. 20. Comparison of M^2 and Strehl^{-2} for a zeroth- and first-order beam.

the mode content of the laser, so that the theoretical peak value is that of a pure zeroth-order Gaussian beam. For very nearly perfect beams, Strehl^{-2} is often approximately equal to M^2. It is important to know the limitations of this approximation. Both Strehl and M^2 are easy to calculate for a hypothetical beam of known mode composition. Figure 20 shows a comparison of M^2 and Strehl^{-2} for a hypothetical beam composed of zeroth- and first-order

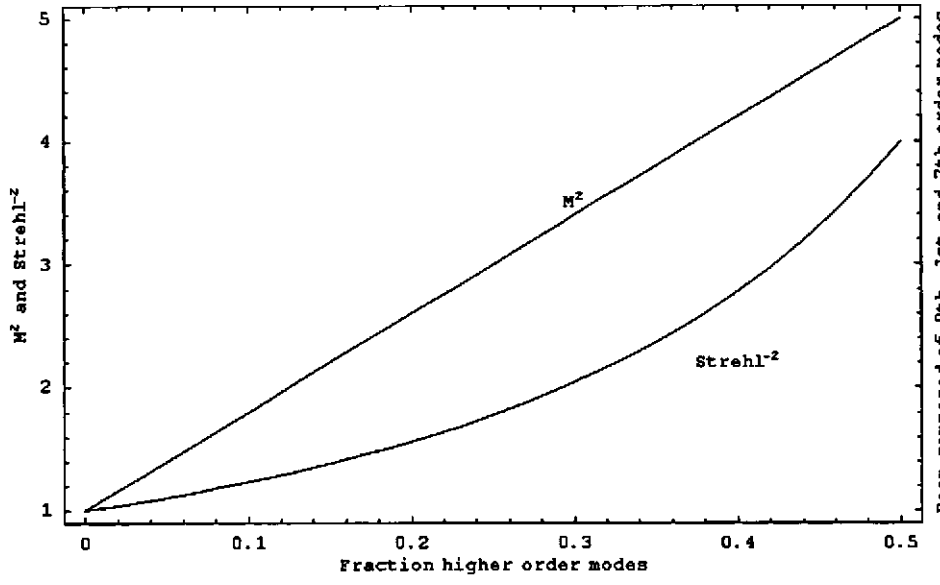

Fig. 21. Comparison of M^2 and Strehl^{-2} for a zeroth-, first-, and seventh-order beam.

modes. The horizontal axis is the energy fraction of the first-order component of the beam. In this case, M^2 and Strehl^{-2} are approximately equal for up to 5% first-order beam, out to an M^2 of ~1.1. Figure 21 shows a comparison of M^2 and Strehl^{-2} for a hypothetical beam composed of zeroth-, first-, and seventh-order modes. In this case, the first- and seventh-order modes are given equal energy content and the horizontal axis is the energy fraction of higher order modes in the beam. In this case, Strehl^{-2} sharply diverges from M^2 immediately and the two are equivalent only for $M^2 = \text{Strehl}^{-2} = 1.00$. In this particular case, if Strehl were measured and M^2 reported, then the beam quality would be reported as better than it actually is and the laser would not, in fact, perform as well as might be expected. If the phrase "times diffraction limited" were used, then the underlying method might remain obscured.

Whether M^2 is or is not approximately equal to Strehl^{-2} is entirely dependent on the mode content of the beam. In practice, we can never know the exact mode content. Each of the different measures of beam quality gives a unique view on the effect of the unknown mode content of a given beam. It is the authors' opinion that M^2 and Strehl^{-2} should never be assumed to be equivalent to one another.

References

[1]Born, M., and E. Wolf, *Principles of Optics*, Pergamon (1975).

[2]Carter, W.H., Appl. Opt. **19**, 1027 (1980).

[3]Das, P., *Lasers and Optical Engineering*, Springer-Verlag, Berlin (1991).

[4]Goodman, J.W., *Statistical Optics*, Wiley (1985).

[5]ISO Standard 11146:1999.

[6]Johnston, T.F., Jr., and M.W. Sasnett, *Handbook of Optical and Laser Scanning*, Chapter 1, "Characterization of Laser Beams: The M^2 Model."

[7]Koechner, W., *Solid-State Laser Engineering*, Springer-Verlag, Berlin, 1996.

[8]Latham, W.P., and A. Kar, Laser Optical Quality, 16th International Congress on Applications of Lasers and Electro-Optics, Nov. 12–17, 1997, San Diego, CA, LIA Proceedings, Vol. 84, Sec. A, pp. 197–206.

[9]*Numerical Recipes*, Cambridge University Press (1986).

[10]Phillips, R.L., and L.C. Andrews, Appl. Opt. **22**, 643 (1983).

[11]Ross, T.S., "3 micron Fiber Laser Pumped Optical Parametric Oscillator," Technical Digest of the Solid State and Diode Laser Technology Review, June 2004, MIR-1.

[12]Siegman, A.E., *Lasers*, University Science Books (1986).

[13]Siegman, A.E., Proc. SPIE, **1224**, 1 (1990).

[14]Siegman, A.E., and S.W. Townsend, IEEE J. Quantum Elec. **29**(4), 1212 [Eq. (5)] (1993).

[15]Spanier, J., and K.B. Oldham, *An Atlas of Functions*, Hemisphere, (1987).

[16]Weber, H., Proc. SPIE **3267**, 2 (1998).

[17]*"Mathworld"* —A Wolfram web resource http://mathworld.wolfram.com

The Authors

Dr. William P. Latham (Pete) completed his Ph.D. in Many-Body Physics in 1976. He had studied the theory of light scattering from superfluid helium in his dissertation. He then accomplished a postdoctoral research appointment in light scattering from semiconductor materials with a picosecond, pulsed laser. After being hired by the Air Force Weapons Laboratory, he began working on optical resonator theory applied to chemical lasers. He accomplished the optical cavity design for the chemical oxygen iodine laser that produced the programmatic beam quality goal. He has performed analyses for most of the large lasers within the Air Force. He has managed several large laser projects. He has also contributed his expertise in the area of technology transfer to leverage that process to

improve the programs at the Air Force Research Laboratory/Directed Energy Directorate. He has worked to establish and maintain close working relationships with universities in many areas of optical physics. He has also contributed leadership in directed energy education. He has continued to contribute scientific research in the area of light scattering from materials during his nearly 30-year tenure with AFRL.

Dr. T. Sean Ross is a graduate of the CREOL school of optics and has been with the Air Force Research Laboratory, Directed Energy Directorate, since 1998. His research projects have included pulsed and CW parametric oscillators, frequency conversion devices, and high-power solid-state lasers. He is the chair of the Solid State and Diode Laser Technology Review advisory board.