Generally astigmatic Gaussian beam representation and optimization using skew rays

Paul D. Colbourne†

Lumentum, 61 Bill Leathem Dr., Ottawa, ON, Canada, K2J 0P7

ABSTRACT

Methods are presented of using skew rays to optimize a generally astigmatic optical system to obtain the desired Gaussian beam focus and minimize aberrations, and to calculate the propagating generally astigmatic Gaussian beam parameters at any point. The optimization method requires very little computation beyond that of a conventional ray optimization, and requires no explicit calculation of the properties of the propagating Gaussian beam. Unlike previous methods, the calculation of beam parameters does not require matrix calculations or the introduction of non-physical concepts such as imaginary rays.

Gaussian beam, general astigmatism, ray trace, lens, focus, optimize, geometric optics

1. INTRODUCTION

Although several authors have shown that propagating rays can represent propagating Gaussian beams[1-5], these techniques are not well known among optical designers. Arnaud[1] showed that Gaussian beam propagation can be described by the propagation of "complex rays" (a complex ray consisting of a "real" ray and an "imaginary" ray, essentially two rays), and he also showed that skew rays follow the envelope of a propagating circular Gaussian beam. Herloski et. al.[2] have shown that by propagating two carefully selected rays through a simply astigmatic optical system (separable into x and y components), simple calculations allow the determination of the properties of a propagating Gaussian beam. Greynolds[3], Arnaud[4], and Lü et. al.[5] have shown that by propagating four rays, complex matrix calculations can be used to determine Gaussian beam properties in the case of general astigmatism (where the optics need not be aligned to the x-y axes). In Sections 2 and 3 of this paper I show that by using four skew rays, simple calculations analogous to those of Herloski et. al. can provide the properties of a propagating Gaussian beam in the case of general astigmatism, without need for matrix calculations or non-physical concepts such as complex rays. These techniques could find application in ray-tracing software to calculate the propagation of Gaussian beams through optical systems more accurately than the widely-used formulae of Herloski et. al. which provide correct results only for systems with simple astigmatism.

Calculation of propagating Gaussian beam parameters is useful, but also of interest would be to use skew rays to optimize an optical system which uses Gaussian beams. Available techniques for this are limited and have various disadvantages. The formulae of Herloski et. al.[2] can be used to obtain the correct focus, but these formulae only work for the paraxial case, and they do not apply to systems with general astigmatism. This is a more serious restriction than may be initially thought, because even if an optical system does not contain any intentionally rotated elements, a spherical lens can induce general astigmatism if the beam is incident off-axis, or elements may become rotated during a tolerancing exercise. Using conventional rays to optimize aberrations may result in a system that is not properly optimized, even if the resulting focus errors are not important, because the rays used do not sample the actual lens apertures which the Gaussian beam sees. For example, at a particular surface the conventional rays may focus to a single point, whereas the actual beam occupies a significant region of the optic. Beam propagation methods are computationally intensive, do not work through all surface types, and tend not to converge well because calculation "noise" causes the optimization algorithms to get stuck in local minima. What is desired is a computationally efficient

† Email: paul.colbourne@lumentum.com

http://dx.doi.org/10.1117/12.2071105

ray-based optimization method applicable to generally astigmatic systems. In Section 4, I propose and demonstrate a ray-based optimization method which makes use of launched skew rays instead of the conventional method of rays diverging from a point. Incorporating these methods into ray-tracing software should enable faster, more accurate design of optical systems which use Gaussian beams.

2. SKEW RAY REPRESENTATION OF GAUSSIAN BEAMS

I will first review the skew ray representations of circular and simply astigmatic Gaussian beams, then I will extend the representation to generally astigmatic systems. As Arnaud demonstrated[1] (see his Figure 1), a set of skew rays will follow the envelope of a propagating circular Gaussian beam. This construction is of more than theoretical interest, it is useful for visualizing the propagation of a Gaussian beam through an optical system. Starting at the input beam waist, the set of skew rays is defined by

$$\begin{aligned} x &= \omega_0 \cos(\alpha) \\ y &= \omega_0 \sin(\alpha) \\ l &= -\theta_0 \sin(\alpha) \\ m &= \theta_0 \cos(\alpha) \end{aligned} \qquad (1)$$

where $l=dx/dz$ and $m=dy/dz$, and $\theta_0 = \lambda/(\pi\, \omega_0)$. If an infinite number of rays are generated, with all values of α from 0 to 2π, the resulting ray bundle will coincide with the $1/e^2$ intensity boundary of the propagating Gaussian beam at all points through the optical system, provided the beam remains circular and there are no aberrations. Propagating a subset of rays, say 16 rays with α equally spaced between 0 and 2π, allows easy visualization of Gaussian beams on layout diagrams of ray-tracing software; an example is shown in Figure 1. Note that to convert l and m into the direction cosines (a, b, c) commonly used to define ray directions, $a=l/\sqrt{(l^2+m^2+1)}$, $b=m/\sqrt{(l^2+m^2+1)}$, and $c=1/\sqrt{(l^2+m^2+1)}$. I use l and m throughout this paper rather than direction cosines because doing so produces more accurate results in the paraxial case, and also the theoretical analysis is more straightforward.

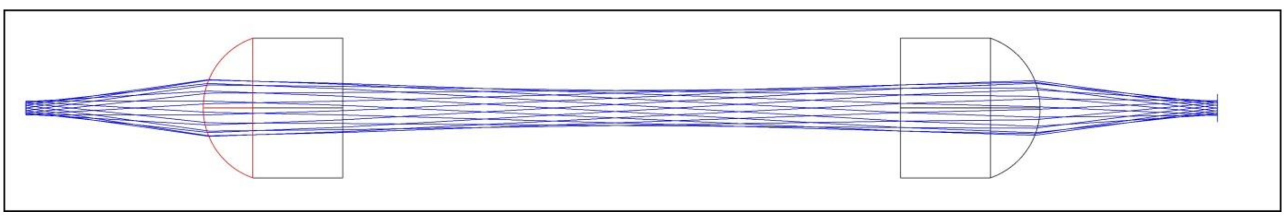

Figure 1. An example of the use of skew rays to visualize the propagation of Gaussian beams in optical systems.

Herloski et. al.'s construction[2] of a "waist ray" (originating a distance ω_0 from the z axis, parallel to the z axis) and a "divergence ray" (originating at the origin at an angle corresponding to the far-field divergence angle θ_0 of the Gaussian beam) can be interpreted as a projection of two skew rays with α=0 and α=π/2 in Eq. (1) onto the x-z plane. It is worth noting that any two skew rays with α differing by π/2 may be used as the launched rays, and all of Herloski's results will still be valid. The choice of rays to launch is not limited to the waist ray and divergence ray defined by Herloski.

If the optical system is simply astigmatic, that is, the input elliptical beam is aligned with the x-y axes and all cylindrical or astigmatic lenses are aligned with the x-y axes, one can construct a waist ray and divergence ray for the x axis and a waist and divergence ray for the y axis and compute the x and y components of the propagating Gaussian beam independently. One can realize these two waist rays and two divergence rays as projections of the same two skew rays onto the x-z plane and the y-z plane, thus by propagating two skew rays one can compute the x and y parameters of the propagating simply astigmatic Gaussian beam. A skew ray bundle defined by

$$\begin{aligned} x &= \omega_{0x} \cos(\alpha) \\ y &= \omega_{0y} \sin(\alpha) \\ l &= -\theta_{0x} \sin(\alpha) \\ m &= \theta_{0y} \cos(\alpha) \end{aligned} \qquad (2)$$

would create the same waist rays and divergence rays defined by Herloski when rays corresponding to α=0 and α=π/2 are projected onto the *x-z* and *y-z* planes. Again, any two skew rays with α differing by π/2 may be launched and all of Herloski's analysis results will still be correct.

While Herloski's analysis works for elliptical beams and simply astigmatic optical systems, the skew ray bundle defined by Eq. (2) does not follow the envelope of an elliptical beam. This is illustrated in Figure 2, which shows the evolution of a skew ray bundle representing an elliptical beam according to Eq. (2). The skew ray bundle remains elliptical in shape but the major axis rotates as the bundle propagates, however we know that a simple elliptical beam remains aligned with the *x-y* axes as it propagates. The skew ray bundle does however follow the profile of the elliptical beam if viewed only along the *x* or *y* direction, as it must according to Herloski's analysis.

The skew ray bundle may be defined with a right handed skew (that is, the rays twist as a right-handed screw as the rays propagate) or with a left handed skew, either sense would be equivalent to Herloski's construction. The rays generated by Eq. (2) are right skew rays; generating rays with a left handed skew involves using the opposite sign for either *x* and *y* or *l* and *m*. We make the interesting observation that if one propagates both left skew rays and right skew rays, the projected size of the propagating elliptical beam at any view angle φ is equal to the RMS average of the projected sizes of the left skew and right skew ray bundles (see Figure 3):

$$A(\varphi) = \sqrt{\frac{[A_R(\varphi)]^2 + [A_L(\varphi)]^2}{2}} \qquad (3)$$

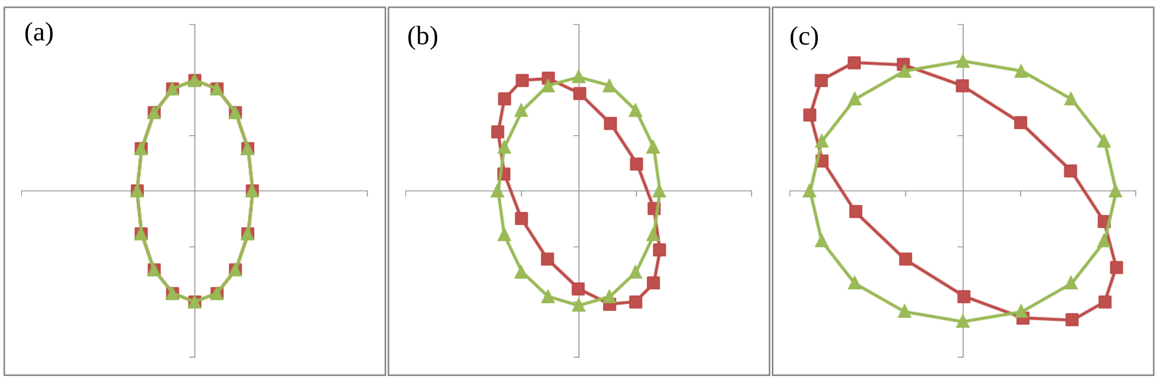

Figure 2. Evolution of a skew ray bundle corresponding to a 10 µm × 20 µm elliptical beam with λ=1.55 µm (squares), compared with the actual beam profile (triangles), (a) at beam waist, (b) 200 µm from waist, and (c) 500 µm from waist.

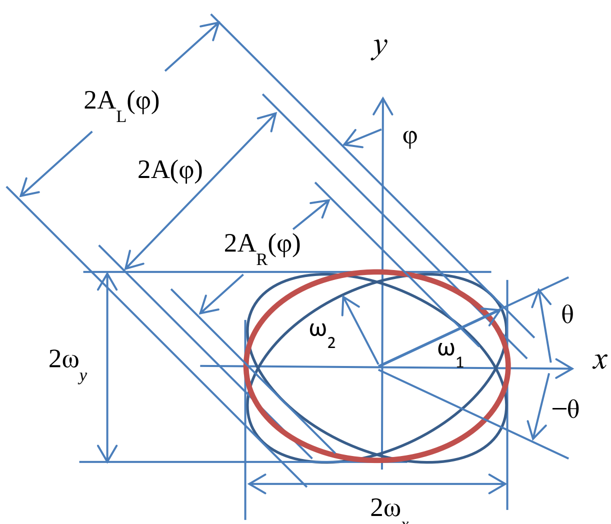

Figure 3. The projected size of the propagating elliptical beam $A(\varphi)$ is equal to the RMS average of the projected sizes of the left skew ray bundle $A_L(\varphi)$ and right skew ray bundle $A_R(\varphi)$.

It can be shown that this relationship holds through any simply astigmatic optical system by noting that, due to the symmetric way in which the left and right skew rays are generated, and the symmetric nature of simply astigmatic optical systems, the left skew ray bundle and the right skew ray bundle will have the same major and minor axes ω_1 and ω_2, and will have equal and opposite rotation angles $+\theta$ and $-\theta$ of their major axes. We also know from Herloski's analysis that the projected height and width of the left and right skew ray bundles are equal to the height and width of the propagating Gaussian beam, which can be expressed as

$$\omega_x = \sqrt{(\omega_1 \cos\theta)^2 + (\omega_2 \sin\theta)^2}$$
$$\omega_y = \sqrt{(\omega_1 \sin\theta)^2 + (\omega_2 \cos\theta)^2}. \quad (4)$$

The projected size of the Gaussian beam at a view angle φ is

$$A(\varphi) = \sqrt{(\omega_x \cos\varphi)^2 + (\omega_y \sin\varphi)^2} \quad (5)$$

and the projected sizes of the right and left skew ray bundles are

$$A_R(\varphi) = \sqrt{(\omega_1 \cos(\varphi+\theta))^2 + (\omega_2 \sin(\varphi+\theta))^2}$$
$$A_L(\varphi) = \sqrt{(\omega_1 \cos(\varphi-\theta))^2 + (\omega_2 \sin(\varphi-\theta))^2}. \quad (6)$$

Making use of the trigonometric identities $\cos(\varphi+\theta)=\cos(\varphi)\cos(\theta)-\sin(\varphi)\sin(\theta)$ and $\sin(\varphi+\theta)=\sin(\varphi)\cos(\theta)+\cos(\varphi)\sin(\theta)$ we obtain

$$\sqrt{\frac{[A_R(\varphi)]^2+[A_L(\varphi)]^2}{2}} = \sqrt{\frac{\omega_1{}^2[\cos^2(\varphi+\theta)+\cos^2(\varphi-\theta)]+\omega_2{}^2[\sin^2(\varphi+\theta)+\sin^2(\varphi-\theta)]}{2}} =$$
$$\sqrt{\omega_1{}^2[\cos^2\theta\cos^2\varphi + \sin^2\theta\sin^2\varphi] + \omega_2{}^2[\sin^2\theta\cos^2\varphi + \cos^2\theta\sin^2\varphi]} =$$
$$\sqrt{(\omega_x \cos\varphi)^2 + (\omega_y \sin\varphi)^2} = A(\varphi) \quad (7)$$

thus proving Eq. (3) for the case of simple astigmatism. In Section 3, the validity of Eq. (3) for general astigmatism is confirmed by comparing the calculated propagation of a generally astigmatic beam with previously published results.

To calculate the Gaussian beam parameters through the optical system, it is only necessary to propagate two right skew rays to define the right skew ray bundle, and two left skew rays to define the left skew ray bundle. In addition, the chief ray is propagated, launched on-axis from the center of the input beam, and all subsequent ray positions and angles are taken to be relative to the chief ray. For simplicity, the step of calculating the ray positions and angles relative to the chief ray is omitted from all following discussions. The four rays to be propagated are defined by

$$x_1 = p\omega_{0x}\cos(\alpha) \qquad x_3 = -p\omega_{0x}\cos(\beta)$$
$$y_1 = p\omega_{0y}\sin(\alpha) \qquad y_3 = -p\omega_{0y}\sin(\beta)$$
$$l_1 = -p\theta_{0x}\sin(\alpha) \qquad l_3 = -p\theta_{0x}\sin(\beta)$$
$$m_1 = p\theta_{0y}\cos(\alpha) \qquad m_3 = p\theta_{0y}\cos(\beta)$$

$$x_2 = p\omega_{0x}\cos(\alpha+\pi/2) \qquad x_4 = -p\omega_{0x}\cos(\beta+\pi/2)$$
$$y_2 = p\omega_{0y}\sin(\alpha+\pi/2) \qquad y_4 = -p\omega_{0y}\sin(\beta+\pi/2)$$
$$l_2 = -p\theta_{0x}\sin(\alpha+\pi/2) \qquad l_4 = -p\theta_{0x}\sin(\beta+\pi/2)$$
$$m_2 = p\theta_{0y}\cos(\alpha+\pi/2) \qquad m_4 = p\theta_{0y}\cos(\beta+\pi/2). \quad (8)$$

Ray 1 and ray 2 are right skew rays, and ray 3 and ray 4 are left skew rays. The pupil scaling factor *p* has been introduced to allow selection of rays closer to or further from the beam center, and the parameter β appears because an independent selection of rays can be used to represent the right and left skew ray bundles. The rays are propagated through the optical system in the normal way, and at a later point in the system we obtain new values of *x*, *y*, *l*, and *m* for each ray. For simplicity I use the same symbols for the new values. The projected size of the right skew ray bundle at a view angle φ is given by

$$A_R(\varphi) = \frac{1}{p}\sqrt{(x_1 \cos\varphi + y_1 \sin\varphi)^2 + (x_2 \cos\varphi + y_2 \sin\varphi)^2} \tag{9}$$

and the projected size of the left skew ray bundle is given by

$$A_L(\varphi) = \frac{1}{p}\sqrt{(x_3 \cos\varphi + y_3 \sin\varphi)^2 + (x_4 \cos\varphi + y_4 \sin\varphi)^2} \tag{10}$$

and therefore using Eq. (3) the projected size of the propagating Gaussian beam is given by

$$A(\varphi) = \frac{1}{p}\sqrt{\frac{(x_1 \cos\varphi + y_1 \sin\varphi)^2 + (x_2 \cos\varphi + y_2 \sin\varphi)^2 + (x_3 \cos\varphi + y_3 \sin\varphi)^2 + (x_4 \cos\varphi + y_4 \sin\varphi)^2}{2}}. \tag{11}$$

If all one requires is the projected size of the Gaussian beam at a particular view angle, for example for the purpose of drawing the Gaussian beam on a layout diagram, Eq. (11) is a very computationally efficient way to get that information.

The choice of rays to be launched is not limited to the definition of Eq. (8). The *x* and *y* components are independent, and therefore a "phase difference" γ between *x* and *y* components can be introduced without affecting the calculated value of Eq. (11). This creates left and right skew ray bundles not coincident with the intensity ellipse, as in Figure 3, even at the beam waist. But even this is not the complete set of possible launched rays. If the starting beam is circular, it can be easily seen that ray bundles rotated about the z axis should also be valid representations of the beam. Similarly for elliptical starting beams, other ray bundles may be used although they are not simple rotations. The ray definitions including such rotations are

$$\begin{aligned}
x_1 &= p\omega_{0x}[\cos(\alpha)\cos(\delta) + \sin(\alpha+\gamma)\sin(\delta)]\\
y_1 &= p\omega_{0y}[\sin(\alpha+\gamma)\cos(\delta) - \cos(\alpha)\sin(\delta)]\\
l_1 &= -p\theta_{0x}[\sin(\alpha)\cos(\delta) - \cos(\alpha+\gamma)\sin(\delta)]\\
m_1 &= p\theta_{0y}[\cos(\alpha+\gamma)\cos(\delta) + \sin(\alpha)\sin(\delta)]\\[4pt]
x_2 &= p\omega_{0x}[\cos(\alpha+\pi/2)\cos(\delta) + \sin(\alpha+\pi/2+\gamma)\sin(\delta)]\\
y_2 &= p\omega_{0y}[\sin(\alpha+\pi/2+\gamma)\cos(\delta) - \cos(\alpha+\pi/2)\sin(\delta)]\\
l_2 &= -p\theta_{0x}[\sin(\alpha+\pi/2)\cos(\delta) - \cos(\alpha+\pi/2+\gamma)\sin(\delta)]\\
m_2 &= p\theta_{0y}[\cos(\alpha+\pi/2+\gamma)\cos(\delta) + \sin(\alpha+\pi/2)\sin(\delta)]\\[4pt]
x_3 &= -p\omega_{0x}[\cos(\beta)\cos(\delta) + \sin(\beta-\gamma)\sin(\delta)]\\
y_3 &= -p\omega_{0y}[\sin(\beta-\gamma)\cos(\delta) - \cos(\beta)\sin(\delta)]\\
l_3 &= -p\theta_{0x}[\sin(\beta)\cos(\delta) - \cos(\beta-\gamma)\sin(\delta)]\\
m_3 &= p\theta_{0y}[\cos(\beta-\gamma)\cos(\delta) + \sin(\beta)\sin(\delta)]\\[4pt]
x_4 &= -p\omega_{0x}[\cos(\beta+\pi/2)\cos(\delta) + \sin(\beta+\pi/2-\gamma)\sin(\delta)]\\
y_4 &= -p\omega_{0y}[\sin(\beta+\pi/2-\gamma)\cos(\delta) - \cos(\beta+\pi/2)\sin(\delta)]\\
l_4 &= -p\theta_{0x}[\sin(\beta+\pi/2)\cos(\delta) - \cos(\beta+\pi/2-\gamma)\sin(\delta)]\\
m_4 &= p\theta_{0y}[\cos(\beta+\pi/2-\gamma)\cos(\delta) + \sin(\beta+\pi/2)\sin(\delta)].
\end{aligned} \tag{12}$$

Note that if γ=π/2, the launched right skew rays originate along a single line regardless of the value of α, and similarly the left skew rays originate along a single line. For γ=π, the "right skew" ray bundle becomes a "left skew" ray bundle, and vice versa, so the terms "right skew" and "left skew" are not strictly accurate labels. Nevertheless the term "right skew rays" is used here to mean rays 1 and 2 in Eq. (12) and "left skew rays" is used to mean rays 3 and 4, regardless of the value of γ.

Eq. (12) defines the complete set of rays which can be used with the analysis and optimization methods of this paper (Sections 3 and 4). Any values of α, β, γ, and δ may be selected and the resulting four rays will represent the propagating beam and produce the same calculated results of the beam parameters (see Section 3), assuming no aberrations.

In general, if a ray bundle (that is, rays with all possible values of α and β) corresponding to particular values of γ and δ is launched through an optical system, and the beam is subsequently re-focused at the output, the resulting ray bundle at the output will correspond to different values of γ and δ. Supposing one is using such a ray bundle to optimize the optical system, one might launch skew rays with γ=0 which sample the entire area of the corresponding Gaussian beam, but end up with rays distributed only along two diagonal lines (γ=π/2). If rays concentrated on two diagonal lines would not be considered sufficient for optimizing the system aberrations (see section 4), one can launch multiple ray bundles with different values of γ and δ to ensure sufficient sampling of the possible ray paths.

The parameters γ and δ are somewhat analogous to polarization states in terms of their effect on the shapes of the ray bundles. When γ=0, the right skew and left skew ray bundles are analogous to right circular and left circular polarization states. When γ=π/2, the ray bundles correspond to linear states of polarization, and δ=0 corresponds to ±45° linearly polarized states and δ=π/4 corresponds to 0° and 90° polarized states. A suitable choice of ray bundles to ensure complete sampling of the optical aperture under all conditions would be (1) γ=0, δ=0, (2) γ=π/2, δ=0, and (3) γ=π/2, δ=π/4. Four rays from case (3), specifically waist and divergence rays along two perpendicular axes, were used by Greynolds[3] in his Gaussian beam decomposition analysis of diffractive beam propagation; the relevance to this work is the agreement that these four rays are sufficient to represent a generally astigmatic Gaussian beam.

3. CALCULATION OF BEAM PROPERTIES

A generally astigmatic Gaussian beam has an elliptical intensity distribution, and elliptical or hyperbolic equi-phase contours which are not necessarily aligned with the intensity ellipse[6]. We have shown that the projected size of the Gaussian beam at any view angle can be calculated from four propagated skew rays. Now the task is to use this information to determine the orientation and principal axes of the intensity ellipse, and the orientation and curvatures along the principal axes of the wavefront surface.

The determination of the orientation and principal axes of the intensity ellipse from projected size calculations is equivalent to the problem of determining a polarization ellipse from radiation intensity measurements. Three projected size calculations at different view angles are required to determine the three unknown parameters of the intensity ellipse. I present a solution which takes advantage of calculations being made in the directions 0, π/2, and π/4, which are also convenient directions to perform the calculations.

When the launched rays are defined as in Eq. (12), the projected size of the generally astigmatic propagated beam at any view angle φ can be calculated using Eq. (11). The projected sizes at view angles 0, π/2, and π/4 are

$$A_0 = A(0) = \frac{1}{p}\sqrt{\frac{{x_1}^2+{x_2}^2+{x_3}^2+{x_4}^2}{2}}$$

$$A_{\pi/2} = A\left(\frac{\pi}{2}\right) = \frac{1}{p}\sqrt{\frac{{y_1}^2+{y_2}^2+{y_3}^2+{y_4}^2}{2}}$$

$$A_{\pi/4} = A\left(\frac{\pi}{4}\right) = \frac{1}{p}\sqrt{\frac{(x_1+y_1)^2+(x_2+y_2)^2+(x_3+y_3)^2+(x_4+y_4)^2}{4}}. \quad (13)$$

The orientation of the intensity ellipse is then given by

$$\theta' = \frac{1}{2}\tan^{-1}\left(\frac{2{A_{\pi/4}}^2-{A_0}^2-{A_{\pi/2}}^2}{{A_0}^2-{A_{\pi/2}}^2}\right). \quad (14)$$

If $A(0)=A(\pi/2)$ there is a division by zero in Eq. (14); in this case, we set θ′=π/4. We can define new coordinate axes (x',y') aligned with the intensity ellipse, to simplify the remaining calculations. All ray positions and angles can be mapped to the new coordinates using for example

$$
\begin{aligned}
x'_1 &= x_1\cos(\theta') + y_1\sin(\theta') \\
y'_1 &= -x_1\sin(\theta') + y_1\cos(\theta') \\
l'_1 &= l_1\cos(\theta') + m_1\sin(\theta') \\
m'_1 &= -l_1\sin(\theta') + m_1\cos(\theta').
\end{aligned} \tag{15}
$$

The principal axes $\omega_{x'}$ and $\omega_{y'}$ of the intensity ellipse are then given by

$$
\begin{aligned}
\omega_{x'} &= \frac{1}{p}\sqrt{\frac{{x'_1}^2+{x'_2}^2+{x'_3}^2+{x'_4}^2}{2}} \\
\omega_{y'} &= \frac{1}{p}\sqrt{\frac{{y'_1}^2+{y'_2}^2+{y'_3}^2+{y'_4}^2}{2}}.
\end{aligned} \tag{16}
$$

We can determine phase information by calculating *dA/dz.* The value of *dA/dz* corresponds to the slope of the wavefront surface in a direction perpendicular to the intensity ellipse, at the tangent point of the intensity ellipse, since the propagation direction is perpendicular to the wavefront. If we calculate (*dA/dz*)/*A*, this gives the curvature of the wavefront along the direction φ, provided that a line perpendicular to the intensity ellipse passes through the origin. This condition will be satisfied along the major and minor axes of the intensity ellipse. Therefore, we have that the wavefront curvatures along x' and y' are

$$
\begin{aligned}
C_{x'} &= \frac{\left(\frac{d\omega_{x'}}{dz}\right)}{\omega_{x'}} = \frac{x'_1 l'_1 + x'_2 l'_2 + x'_3 l'_3 + x'_4 l'_4}{{x'_1}^2+{x'_2}^2+{x'_3}^2+{x'_4}^2} \\
C_{y'} &= \frac{\left(\frac{d\omega_{y'}}{dz}\right)}{\omega_{y'}} = \frac{y'_1 m'_1 + y'_2 m'_2 + y'_3 m'_3 + y'_4 m'_4}{{y'_1}^2+{y'_2}^2+{y'_3}^2+{y'_4}^2}\,.
\end{aligned} \tag{17}
$$

We can also calculate *dA/dz* at an angle π/4 from the x' axis.

$$
\frac{dA'_{\pi/4}}{dz} = \frac{(x'_1+y'_1)(l'_1+m'_1)+(x'_2+y'_2)(l'_2+m'_2)+(x'_3+y'_3)(l'_3+m'_3)+(x'_4+y'_4)(l'_4+m'_4)}{2p\sqrt{(x'_1+y'_1)^2+(x'_2+y'_2)^2+(x'_3+y'_3)^2+(x'_4+y'_4)^2}} \tag{18}
$$

To complete the description of the propagating generally astigmatic Gaussian beam, we need to determine the orientation of the equi-phase contours and the curvatures along the principal axes of the wavefront surface. We can extract this information from the calculated values of $C_{x'}$, $C_{y'}$, and *dA/dz* as follows. We know that in coordinates (x'',y'') oriented along the principal axes of the wavefront surface, the sag of the wavefront surface can be defined as:

$$
S = \frac{C_{x''}x''^2 + C_{y''}y''^2}{2}. \tag{19}
$$

In the coordinate system (x',y'), rotated θ" from the coordinates (x'',y''), the wavefront surface becomes

$$
S = \frac{C_{x'}{x'}^2 + C_{y'}{y'}^2}{2} + bx'y'. \tag{20}
$$

The tangent point to the intensity ellipse at an angle π/4 is given by

$$
\begin{aligned}
u &= \tan^{-1}\left(\frac{\omega_{y'}}{\omega_{x'}}\tan\frac{\pi}{4}\right) \\
x'_t &= \omega_{x'}\cos(u) \\
y'_t &= \omega_{y'}\sin(u).
\end{aligned} \tag{21}
$$

We can calculate *dA*/*dz* (numerically equal to the slope of the wavefront surface in a direction perpendicular to the intensity ellipse) using Eq. (20):

$$\frac{dA'_{\pi/4}}{dz} = \left[\frac{dS}{dx'}\cos\frac{\pi}{4} + \frac{dS}{dy'}\sin\frac{\pi}{4}\right]_{(x'_t, y'_t)} = \frac{x'_t C_{x'} + b(x'_t + y'_t) + y'_t C_{y'}}{\sqrt{2}} \tag{22}$$

and rearranging, we obtain the value of b as

$$b = \frac{\sqrt{2}\frac{dA'_{\pi/4}}{dz} - x'_t C_{x'} - y'_t C_{y'}}{x'_t + y'_t} \tag{23}$$

where the value of *dA/dz* can be determined from Eq. (18), and $C_{x'}$, $C_{y'}$, x'_t, and y'_t. are determined from Eqs. (17) and (21). Now all constants of Eq. (20) have been determined, so we have a complete expression for the wavefront surface, and we have only to determine the coordinate axis orientation for which the wavefront expression reduces to Eq. (19). To determine the orientation, we calculate the wavefront sag around a unit circle. Three points are sufficient to determine the orientation of the wavefront surface. We choose points at angles 0, π/4, and π/2 around the unit circle.

$$S_0 = \frac{C_{x'}}{2}$$

$$S_{\pi/2} = \frac{C_{y'}}{2}$$

$$S_{\pi/4} = \frac{C_{x'} + C_{y'}}{4} + \frac{b}{2} \tag{24}$$

We calculate the quantities

$$M = \frac{S_0 + S_{\pi/2}}{2}$$

$$T = \sqrt{(S_0 - M)^2 + \left(S_{\pi/4} - M\right)^2}. \tag{25}$$

The orientation of the wavefront surface relative to the intensity ellipse is then given by

$$\theta'' = \frac{\mathrm{atan2}\left(\frac{S_0 - M}{T}\ ,\ \frac{S_{\pi/4} - M}{T}\right)}{2} \tag{26}$$

where atan2 is an arc tangent function commonly found in math libraries which can return angles from 0 to 2π. The curvatures along the major and minor axes are

$$C_{x''} = 2(M + T)$$

$$C_{y''} = 2(M - T). \tag{27}$$

We now have the complete description of the propagating generally astigmatic Gaussian beam, extracted from the propagation of four skew rays. This calculation approach is not necessarily more computationally efficient than the matrix methods described by Arnaud[4] or Lü et. al.[5], however it has been derived without reference to physically unrealistic concepts such as imaginary rays or imaginary rotation angles, and therefore has some value in furthering understanding of the propagation of generally astigmatic Gaussian beams.

As a check of the validity of the calculation results, I compare calculations of intensity ellipse orientation with calculations done by Greynolds[3] using matrix methods. Greynolds assumed a launched beam of wavelength 10 μm with waists ω_{0x} = 2 mm and ω_{0y} = 1 mm, but Greynolds' Gaussian beam definition includes a factor π in the exponent, so in the current convention of defining ω as the $1/e^2$ intensity radius, we have $\omega_{0x} = (2/\sqrt{\pi})$ mm and $\omega_{0y} = (1/\sqrt{\pi})$ mm. The beam then passes through a cylindrical lens with focal length 100 mm oriented at a 45 degree angle, which creates a generally astigmatic beam, and Greynolds calculates the orientation of the intensity ellipse at propagation distances of 0 mm (0°), 50 mm (27.5°), 100 mm (53.3°), 150 mm (74.8°), 200 mm (90°), and 250 mm (99.4°). My calculations (0°, 27.51°, 53.35°, 74.77°, 90.00° and 99.39° respectively) agree exactly with those of Greynolds. This confirms that at least Eqs. (3), (11), and (14) are valid. The wavefront calculations (culminating in Eqs. (26) and (27)) have been verified for elliptical beams immediately following a rotated cylinder lens, for which the true wavefront is easily calculated.

4. OPTIMIZATION USING SKEW RAYS

The optimization method starts with the generation of skew rays as in Eq. (12). This was accomplished in Zemax (optical design software available from Zemax, LLC) by creating a user-defined surface which offsets rays according to their angle and the beam waist radii ω_{0x} and ω_{0y} which are entered as parameters of the user-defined surface:

$$\begin{aligned} x &= x + m\omega_{0x}/|\theta_{0x}| \\ y &= y - l\omega_{0y}/|\theta_{0x}| \\ m &= m\,\omega_{0x}/\omega_{0y}. \end{aligned} \tag{28}$$

The system aperture type is set to "Object Space NA" with aperture value corresponding to θ_{0x} so that rays originate diverging from a point. It can be verified that the user-defined surface placed at the input converts each launched ray (with arbitrary pupil value) to a skew ray meeting the conditions of Eq. (8), or Eq. (12) with γ=δ=0, and appropriate values of p and α or β. The absolute value of θ_{0x} is used in Eq. (28) so that by specifying positive or negative values of ω_{0x} and ω_{0y}, right skew or left skew rays may be generated. The use of Eq. (28) requires that the input beam is elliptical and aligned with the x-y axes, with both x and y waists located at the input surface. If the desired input beam does not fit this condition, for example if it is astigmatic with x and y waists located at different planes, paraxial lenses or coordinate breaks may be inserted so that an input beam meeting the condition is transformed into the desired input beam.

If the input beam is circular ($\omega_{0x}=\omega_{0y}=\omega_0$) and remains circular throughout the optical system, there is a particularly simple method of performing optimization using skew rays. A second user-defined surface can be placed at the output surface with the opposite sign of ω_0 (but not necessarily the same magnitude as the input value if the desired output beam radius is different from the input beam radius). This reverses the skew ray offsets that were created at the input, resulting in output rays that converge to a point if the beam is focused at the output surface. Then, conventional ray optimization algorithms can be used to optimize the system and it will converge such that the Gaussian beam is focused. However, this method fails if the system contains cylindrical lenses, even if the beam returns to a circular beam at the output. The reason is that in general the output skew ray bundle will correspond to a non-zero value of γ after passing through cylindrical lenses, so implementing Eq. (28) at the output does not result in rays converging to a point.

In the more general case that the input beam is elliptical or the system contains cylindrical lenses, a different optimization method is required. Referring to Eq. (12), we note that x_1 and l_2 have the same dependence on α, γ, and δ, and similarly y_1 and m_2 have the same dependence. This suggests that if we offset the output ray 1 based on the angle of ray 2 (and not the angle of ray 1), the ray positions may converge to a point regardless of whether the beam is elliptical or has passed through cylindrical lenses. Such an optimization method cannot be implemented using a user-defined surface, because a surface operates on only one ray at a time. To implement an optimization method where the position of ray 1 can be modified based on the angle of ray 2, a Zemax macro was written, callable from within the merit function, which propagates 3 rings of 8 rays. Having access to all ray data, the macro can offset the position of ray 1 based on the angle of ray 2.

$$\begin{aligned} x_{1b} &= x_1 + l_2\omega_{0xT}/|\theta_{0xT}| \\ y_{1b} &= y_1 + m_2\omega_{0yT}/|\theta_{0yT}| \end{aligned} \tag{29}$$

Here ω_{0xT} and ω_{0yT} are the target output waist values, again set to negative values when left skew rays are being propagated The signs of ω_{0xT} and ω_{0yT} used in Eq. (29) also need to be inverted if the rays have reflected from an odd number of mirrors. Similarly, the position of ray 2 can be offset based on the angle of the ray corresponding to α+π, and continuing for each propagated ray the ray position is offset based on the angle of the ray corresponding to α+π/2 or β+π/2. The macro computes an RMS spot radius relative to the chief ray, similar to a conventional ray optimization, but using the offset ray positions (x_{1b} , y_{1b}) etc. to compute the RMS spot radius. The use of Eq. (29) assumes that the desired output is a Gaussian beam with x and y waists located at the output surface. If the desired output does not fit this condition, paraxial lenses or coordinate breaks can be inserted to transform the desired output to a beam of the required form. Two configurations are set up, one propagating right skew rays and one propagating left skew rays, and the macro is called for both configurations. If the optimization is successful and the RMS spot radius is forced to zero, then

$$x_1 = -\, l_2 \omega_{0xT} / |\theta_{0xT}|$$
$$y_1 = -\, m_2 \omega_{0yT} / \left|\theta_{0yT}\right| \qquad (30)$$

and if we assume that the ray at α+π has position and angle equal in magnitude but opposite in sign to those of ray 1 (which follows due to symmetry if there are no aberrations),

$$x_2 = l_1 \omega_{0xT} / |\theta_{0xT}|$$
$$y_2 = m_1 \omega_{0yT} / \left|\theta_{0yT}\right|. \qquad (31)$$

Translating Eqs. (42) and (43) of Herloski et. al.[2] into the notation of this paper, for the x component only, we have that for the case of simple astigmatism

$$\omega_x = \frac{1}{p} \sqrt{\, {x_1}^2 + {x_2}^2} \qquad (32)$$

$$z = \frac{x_1 l_1 + x_2 l_2}{{l_1}^2 + {l_2}^2}. \qquad (33)$$

In Eq. (32) the factor $1/p$ has been inserted since in Eq. (12) the definition of launched rays allows the pupil scaling factor p. Herloski et. al. also state in words in the paragraphs following their Eqs. (42) and (43) that

$$\theta_{0x} = \frac{1}{p} \sqrt{\, {l_1}^2 + {l_2}^2} \qquad (34)$$

where again the factor $1/p$ has been inserted. Using Eqs. (30) and (31) to replace for x_1 and x_2 in Eq. (33), we get the result that the beam waist is located at the output surface (z=0), and therefore Eq. (32) will give us the waist size ω_{0x} of the output beam. Using Eqs. (32) and (34) to calculate the Rayleigh range squared $z_{Rx}^2 = \omega_{0x}^2/\theta_{0x}^2$ and substituting for x_1 and x_2 using Eqs. (30) and (31), we get the result that $\omega_{0x}^2/\theta_{0x}^2 = \omega_{0xT}^2/\theta_{0xT}^2$ and therefore the actual beam radius ω_{0x} is equal to the desired waist size ω_{0xT}. A similar analysis can be done for the y component. We have thus proven that at least for the case of simple astigmatism where Eqs. (32) to (34) apply, the desired result of a beam with waists ω_{0xT} and ω_{0yT} located at the output surface has been achieved. Given that propagation of right skew rays and left skew rays is sufficient to represent a generally astigmatic Gaussian beam, obtaining an RMS spot size near zero at the output for both right skew rays and left skew rays ensures that even with intervening general astigmatism the correct focus is achieved. Testing with several examples shows this to be the case, and one example is described below.

To demonstrate optimization in the case of general astigmatism, a state of general astigmatism was generated by passing an elliptical beam with waists ω_{0x} = 10 μm and ω_{0y} = 20 μm and wavelength 1.55 μm through a paraxial cylinder lens oriented at a 30 degree angle. Then a series of four paraxial cylindrical lenses with variable focal lengths, rotation angles, and positions was used with a target of achieving a 10 μm radius circular output beam at a particular location 1.5 mm from the input. The only optimization operands were the Gaussian RMS spot size macro called for the right skew and left skew configurations. When the optimization was completed the optimization macro returned an RMS spot radius less than 0.02 μm for both right skew rays and left skew rays, and a Zemax Physical Optics Propagation (POP) calculation showed a 10 μm radius output spot with 99.95% coupling to a 10 μm radius output mode, confirming that the target output beam was achieved.

Just as in conventional ray optimizations, using multiple rings and arms in the optimization enables reduction of aberrations. The equivalence between rays and Gaussian beams (and indeed Gaussian beam propagation theory itself) is valid only when there are no aberrations present, but presumably the final optimized state has negligible aberrations if an RMS spot radius near zero is achieved. If the RMS spot radius is much larger than the desired spot size, the ray offsets introduced in the optimization process are negligible and the optimization is essentially equivalent to a conventional ray optimization. It is only when aberrations become small, and the optical system approaches being diffraction limited, that the ray offsets become significant, and in this state the equivalence between rays and Gaussian beams is valid.

One aberration optimization test case was an off-axis reflector, for which the optimal shape is a parabola (conic constant = −1). An input beam of waist 1 mm and wavelength 1.55 μm was incident on a biconic mirror surface 2 mm off-axis and optimized to achieve an output spot of radius 5 μm. For these conditions the optimum mirror radius is −20.26 mm (focal length 10.13 mm). The *x* and *y* radii and conic constants of the mirror were set as variables. The only merit function operands were the Gaussian RMS spot size macro called for right and left skew ray configurations and an REAY operand to force the output waist onto the axis of the reflector. The system optimized to values of radii −20.275 mm in *x* and −20.279 mm in *y* and conic constants −1.036 in *x* and −0.983 in *y* with a resulting POP coupling of 99.93%.

A second optimization test case involved two identical refractive plano-convex lenses used to couple light from one optical fiber (assumed to have a Gaussian mode shape) to another, similar to Figure 1. The focal distance and conic constant of the first lens were set as variables, the second lens picking up the same values as the first. In this case the optimized conic constant obtained when optimized with the Gaussian RMS spot size macro (−2.157) differed slightly from that obtained when optimizing with POP (−1.949), with a lower POP coupling obtained after optimizing with rays (99.5%) compared with optimization with POP (99.96%). This may be because of the skew component of the incident angle of the skew rays on the lens surface, which will affect the radial deviation of each ray by the lens. Nevertheless, because of the faster optimization using rays compared to using POP, it could still be advantageous to start with a skew ray optimization then fine-tune the optimization using POP. Including rays corresponding to $\gamma=\pi/2$ in the optimization may improve the optimized results for refractive surfaces, since these rays have no skew component.

5. CONCLUSIONS AND FURTHER WORK

I have shown that skew rays can be used to represent Gaussian beams in generally astigmatic optical systems, and can be used to calculate the properties of the propagating generally astigmatic Gaussian beam in a straightforward manner. Furthermore, skew rays can be used to optimize an optical system to minimize aberrations and obtain the correct Gaussian beam focus. The optimization method requires very little computation beyond that of a conventional ray optimization, requires no explicit calculation of the properties of the propagating Gaussian beam, and works even in optical systems with general astigmatism. While these techniques alone should improve the speed and accuracy of the design of optical systems which use Gaussian beams, it is likely that further enhancements are possible, for example methods of calculating fiber coupling efficiency using skew rays. While the optimization method demonstrated using a user-defined surface and an optimization macro is relatively simple to implement, the ease of use could be improved by incorporating into ray-tracing software so that extra configurations would be unnecessary, and including rays with $\gamma=\pi/2$ could improve the optimization performance. Optimization speed could likely be improved if the algorithms were incorporated into the software. The methods described seem to give accurate results for Gaussian beams with NA=0.15 or less, and are thus suitable for modeling beams emerging from and coupling to standard telecommunication optical fibers, but higher NA sources such as laser diodes are probably not accurately modeled, and further research could be done on appropriate ways to model high NA sources. Some gaps in the theoretical treatment could be addressed, for example rigorous proof of Eq. (3) for the case of general astigmatism. Further extensions of the analysis of Herloski et. al.[2] to the generally astigmatic case may be possible, for example while Eq. (11) is the extension of Eq. (32) and Eq. (34) can be similarly extended, an extended Eq. (33) may give the distance to the point where the projected beam size is a minimum at a particular view angle.

[1] Arnaud, J., “Representation of Gaussian beams by complex rays,” Appl. Opt. **24,** 538-543 (1985).

[2] Herloski, R., Marshall, S., and Antos, R., “Gaussian beam ray-equivalent modeling and optical design,” Appl. Opt. **22,** 1168-1174 (1983).

[3] Greynolds, A. W., “Propagation of general astigmatic Gaussian beams along skew ray paths,” Proc. SPIE **560**, 33–50 (1985).

[4] Arnaud, J. A., “Nonorthogonal Optical Waveguides and Resonators,” Bell Syst. Tech. J., **Nov. 1970**, 2311-2348 (1970).

[5] Lü, B., Feng, G., and Cai, B., "Complex ray representation of the astigmatic Gaussian beam propagation," Optical and Quantum Electronics **25**, 275-284 (1993).

[6] Arnaud, J. A., and Kogelnik, H., "Gaussian Light Beams with General Astigmatism," Appl. Opt. **8,** 1687-1693 (1969).

http://dx.doi.org/10.1117/12.2071105