Gaussian Laser Beams via Oblate Spheroidal Waves

Kirk T. McDonald
Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544
(October 19, 2002; updated November 20, 2012)

1 Problem

Gaussian beams provide the simplest mathematical description of the essential features of a focused optical beam, by ignoring higher-order effects induced by apertures elsewhere in the system.

Wavefunctions $\psi(\mathbf{x}, t) = \psi(\mathbf{x})\, e^{-i\omega t}$ for Gaussian laser beams [1]-[13] of angular frequency ω are typically deduced in the paraxial approximation, meaning that in the far zone the functions are accurate only for angles θ with respect to the beam axis that are at most a few times the characteristic diffraction angle,

$$\theta_0 = \frac{\lambda}{\pi w_0} = \frac{2}{k w_0} = \frac{w_0}{z_0}\,, \tag{1}$$

where λ is the wavelength, $k = \omega/c = 2\pi/\lambda$ is the wave number, c is the speed of light, w_0 is the radius of the beam waist, and z_0 is the depth of focus, also called the Rayleigh range, which is related by

$$z_0 = \frac{k w_0^2}{2} = \frac{2}{k\theta_0^2}\,. \tag{2}$$

Since the angle with respect to the beam axis has unique meaning only up to a value of $\pi/2$, the paraxial approximation implies that $\theta_0 \ll 1$, and consequently that $z_0 \gg w_0 \gg \lambda$.

The question arises whether there are any "exact" solutions to the free-space wave equation,

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0, \tag{3}$$

for which the paraxial wavefunctions are a suitable approximation. For monochromatic waves, it suffices to seek "exact" solutions to the Helmholtz wave equation,

$$\nabla^2\psi + k^2\psi = 0. \tag{4}$$

This equation is known to be separable in 11 coordinate systems [14, 15], of which oblate spheroidal coordinates are well matched to the geometry of laser beams, as shown in Fig. 1.

"Exact" solutions to the Helmholtz equation in oblate spheroidal coordinates were developed in the 1930's, and are summarized in [16, 17, 18]. These solutions are, however, rather intricate and were almost forgotten at the time of the invention of the laser in 1960 [19].

This problem does not explore the "exact" solutions,[1] but rather asks you to develop a systematic set of approximate solutions to the Helmholtz equation in oblate spheroidal coordinates, which will turn out to be one representation of paraxial Gaussian laser beams.

[1] It is clear from Fig. 1 that the "exact" oblate-spheroidal wavefunctions do not apply in "free space," but for the case of a conducting plane with a circular aperture. See also [21].

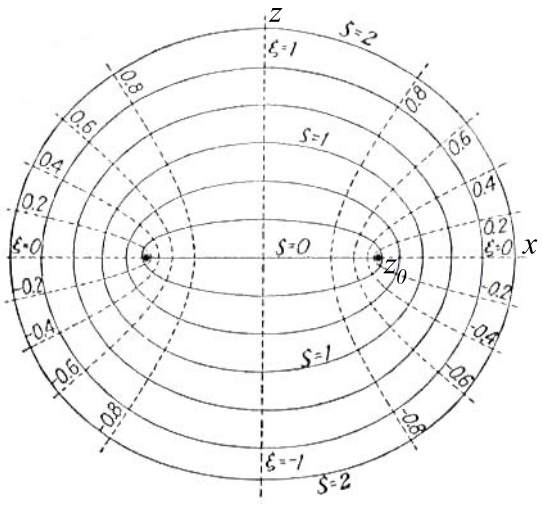

Figure 1: The x-z plane of an oblate spheroidal coordinate system (ζ, ξ, ϕ) based on hyperboloids and ellipsoids of revolution about the z axis, with foci at $(x, z) = (\pm z_0, 0)$. The coordinates have ranges $0 \leq \zeta < \infty$, $-1 \leq \xi \leq 1$, and $0 \leq \phi \leq 2\pi$).

The relation between rectangular coordinates (x, y, z) and oblate spheroidal coordinates[2] (ζ, ξ, ϕ) is,

$$x = z_0\sqrt{1+\zeta^2}\sqrt{1-\xi^2}\cos\phi, \tag{5}$$

$$y = z_0\sqrt{1+\zeta^2}\sqrt{1-\xi^2}\sin\phi, \tag{6}$$

$$z = z_0\zeta\xi, \tag{7}$$

where the length z_0 is the distance from the origin to one of the foci of the ellipses and hyperbolae whose surfaces of revolution about the z axis are surfaces of constant ζ and ξ. Coordinate ϕ is the usual azimuthal angle measured in the x-y plane. For large ζ, the oblate spheroidal coordinates are essentially identical to spherical coordinates (r, θ, ϕ) with the identification $\zeta = r/z_0$ and $\xi = \cos\theta$.

An obvious consequence of the definitions (5)-(7) is that,

$$r_\perp = \sqrt{x^2+y^2} = z_0\sqrt{1+\zeta^2}\sqrt{1-\xi^2}. \tag{8}$$

Combining eqs. (7) and (8), we find,

$$\zeta = \frac{r^2 - z_0^2 + \sqrt{(r^2-z_0^2)^2 + 4z_0^2z^2}}{2z_0^2}. \tag{9}$$

Close to the laser focus, where $r \ll z_0$, we have $\zeta \approx z/z_0 \ll 1$.

[2]Oblate spheroidal coordinates are sometimes written with $\zeta = \sinh u$ and $\xi = \cos v$ or $\sin v$.

It is clear that the oblate spheroidal wave functions will have the mathematical restriction that the entire wave crosses the plane $z = 0$ within an iris of radius z_0, the length used in the definitions (5)-(7) of the oblate spheroidal coordinates. In effect, the plane $z = 0$ is perfectly absorbing except for the iris of radius z_0.[3]

You will find that the length z_0 also has the physical significance of the Rayleigh range (depth of focus), which concept is usually associated with longitudinal rather than tranverse behavior of the waves. Since the paraxial approximation that you will explore is valid only when the beam waist w_0 is small compared to the Rayleigh range, *i.e.*, when $w_0 \ll z_0$, the paraxial wave functions are not accurate descriptions of waves of extremely short focal length, even though they will be formally defined for any value of w_0.

The wave equation (4) is separable in oblate spheroidal coordinates, with the form

$$\frac{\partial}{\partial\zeta}(1+\zeta^2)\frac{\partial\psi}{\partial\zeta} + \frac{\partial}{\partial\xi}(1-\xi^2)\frac{\partial\psi}{\partial\xi} + \frac{\zeta^2+\xi^2}{(1+\zeta^2)(1-\xi^2)}\frac{\partial^2\psi}{\partial\phi^2} + k^2 z_0^2(\zeta^2+\xi^2)\psi = 0. \tag{10}$$

It is helpful to express the wave functions in radial and transverse coordinates that are scaled by the Rayleigh range z_0 and by the diffraction angle θ_0, respectively. The oblate spheroidal coordinate ζ already has this desirable property for large values. However, the coordinate ξ is usefully replaced by,

$$\sigma = \frac{1-\xi^2}{\theta_0^2} = \frac{z_0^2}{w_0^2}(1-\xi^2) = \frac{kz_0}{2}(1-\xi^2) = \frac{r_\perp^2}{w_0^2(1+\zeta^2)}\,, \tag{11}$$

which obeys $\sigma \approx (\theta/\theta_0)^2$ for large r and small θ, and $\sigma \approx (r_\perp/w_0)^2$ near the beam waist where $\zeta \approx 0$.

To replace ξ by σ in the Helmholtz equation (10), note that $2\xi\,d\xi = -\theta_0^2\,d\sigma$. In the paraxial approximation, $\xi \approx 1$ (which implies that your solution will be restricted to waves in the hemisphere $z \geq 0$), you may suppose that,

$$d\xi \approx -\frac{\theta_0^2}{2}\,d\sigma. \tag{12}$$

Find an orthogonal set of waves,

$$\psi_n^m = Z_n^m(\zeta)S_n^m(\sigma)\,e^{\pm im\phi}, \tag{13}$$

which satisfy the Helmholtz equation in the paraxial approximation. You may anticipate that the "angular" functions $S_n^m(\sigma)$ are modulated Gaussians, containing a factor $\sigma^{m/2}\,e^{-\sigma}$. The "radial" functions Z_n^m are modulated spherical waves in the far zone, with a leading factor e^{ikr}, and it suffices to keep terms in the remaining factor that are lowest order in the small quantity θ_0.

Vector electromagnetic waves $\mathbf{E} = \mathbf{E}(\mathbf{x})\,e^{-i\omega t}$ and $\mathbf{B} = \mathbf{B}(\mathbf{x})\,e^{-i\omega t}$ that satisfy Maxwell's equations in free space can be generated from the scalar wave functions ψ_n^m by supposing the vector potential $\mathbf{A}$ has Cartesian components (for which $(\nabla^2\mathbf{A})_j = \nabla^2 A_j$ [15]) given by one

[3] Waves that have an axis must involve some physical entity that defines the axis. An iris is an obvious example of the defining structure.

or more of the scalar waves $\psi_n^m\, e^{-i\omega t}$. For these waves, the fourth Maxwell equation in free space is $c\boldsymbol{\nabla}\times\mathbf{B} = \partial\mathbf{E}/\partial t = -ik\mathbf{E}$ (Gaussian units), so both fields $\mathbf{E}$ and $\mathbf{B}$ can de derived from the vector potential $\mathbf{A}$ according to,

$$\mathbf{E} = \frac{i}{k}\boldsymbol{\nabla}\times\mathbf{B} = ik\mathbf{A} + \frac{i}{k}\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}), \qquad \mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}, \tag{14}$$

since the vector potential obeys the Helmholtz equation (4).

Calculate the ratio of the angular momentum density of the wave in the far zone to its energy density to show that quanta of these waves (photons with intrinsic spin $S = 1$) carry orbital angular momentum in addition to the intrinsic spin. Show also that lines of the Poynting flux form spirals on a cone in the far zone.

2 Solution

Aspects of this problem have been treated in [20].

2.1 The Paraxial Gaussian-Laguerre Wave Functions

2.1.1 Separation of the Approximate Helmholtz Equation

Using the approximation (12) when replacing variable ξ by σ, the Helmholtz equation (10) becomes,

$$\frac{\partial}{\partial\zeta}(1+\zeta^2)\frac{\partial\psi}{\partial\zeta} + \frac{4}{\theta_0^2}\frac{\partial}{\partial\sigma}\sigma\frac{\partial\psi}{\partial\sigma} + \frac{1+\zeta^2-\theta_0^2\sigma}{(1+\zeta^2)\theta_0^2\sigma}\frac{\partial^2\psi}{\partial\phi^2} + \frac{4}{\theta_0^4}(1+\zeta^2-\theta_0^2\sigma)\psi = 0. \tag{15}$$

This equation admits separated solutions of the form (13) for any integer m. Inserting this in eq. (15) and dividing by ψ, we find

$$\frac{1}{Z}\frac{\partial}{\partial\zeta}(1+\zeta^2)\frac{\partial Z}{\partial\zeta} + \frac{4}{\theta_0^2 S}\frac{\partial}{\partial\sigma}\sigma\frac{\partial S}{\partial\sigma} - m^2\frac{1+\zeta^2-\theta_0^2\sigma}{(1+\zeta^2)\theta_0^2\sigma} + \frac{4}{\theta_0^4}(1+\zeta^2) - \frac{4\sigma}{\theta_0^2} = 0. \tag{16}$$

The functions Z and S will be the same for integers m and $-m$, so henceforth we consider m to be non-negative, and write the azimuthal functions as $e^{\pm im\phi}$. With λ_m as the second separation constant, the ζ and σ differential equations are,

$$\frac{d}{d\zeta}(1+\zeta^2)\frac{dZ}{d\zeta} = \left(\lambda_m - \frac{4}{\theta_0^4}(1+\zeta^2) - \frac{m^2}{1+\zeta^2}\right)Z, \tag{17}$$

$$\frac{d}{d\sigma}\sigma\frac{dS}{d\sigma} = -\frac{\theta_0^2}{4}\left(\lambda_m - \frac{4\sigma}{\theta_0^2} - \frac{m^2}{\theta_0^2\sigma}\right)S. \tag{18}$$

2.1.2 The "Angular" Functions

The hint is that the wave functions have Gaussian transverse dependence, which implies that the "angular" function $S(\sigma)$ contains a factor $e^{-\sigma} = e^{-r_\perp^2/w_0^2(1+\zeta^2)}$. We therefore write $S = e^{-\sigma}T$, and eq. (18) becomes,

$$\sigma\frac{d^2T}{d\sigma^2} + (1-2\sigma)\frac{dT}{d\sigma} + \left(1 + \frac{\theta_0^2}{4}\lambda_m - \frac{m^2}{4\sigma}\right)T = 0. \tag{19}$$

The function $T(\sigma)$ cannot be represented as a polynomial, but (like the radial Shrödinger equation) this can be accomplished after a factor of $\sigma^{m/2}$ is extracted. That is, we write $T = \sigma^{m/2} L$, or $S = \sigma^{m/2}\, e^{-\sigma} L$, so that eq. (19) becomes,

$$\sigma \frac{d^2 L}{d\sigma^2} + (m + 1 - 2\sigma)\frac{dL}{d\sigma} + \nu L = 0, \tag{20}$$

where,

$$\nu = \frac{\theta_0^2}{4}\lambda_m - m - 1. \tag{21}$$

If $\nu = 2n$ for integer $n \geq 0$, this is the differential equation for generalized Laguerre polynomials $L_n^m(2\sigma)$ [22], where,

$$L_n^m(x) = m!n! \sum_{k=0}^{n} \frac{(-1)^k x^k}{(m+k)!(n-k)!k!} = 1 - \frac{nx}{m+1} + \frac{n(n-1)x^2}{2(m+1)(m+2)} - ... \tag{22}$$

By direct calculation from eq. (20) with $\nu = 2n$, we readily verify that the low-order solutions are,

$$L_0^m = 1, \qquad L_1^m(2\sigma) = 1 - \frac{2\sigma}{m+1}\,, \qquad L_2^m(2\sigma) = 1 - \frac{4\sigma}{m+1} + \frac{4\sigma^2}{(m+1)(m+2)}\,. \tag{23}$$

Note that index m can be larger than index n.

The Laguerre polynomials are normalized to 1 at $x = 0$, and obey the orthogonality relation,

$$\int_0^\infty L_n^m(x) L_{n'}^m(x) x^m\, e^{-x}\, dz = \frac{(m!)^2 n!}{(m+n)!}\delta_{nn'}. \tag{24}$$

The "angular" functions $S_n^m(\sigma)$ are thus given by,

$$S_n^m(\sigma) = \sigma^{m/2}\, e^{-\sigma} L_n^m(2\sigma), \tag{25}$$

which obey the orthogonality relation,

$$\int_0^\infty S_n^m(\sigma) S_{n'}^m(\sigma)\; d\sigma = \frac{1}{2^{m+1}} \int_0^\infty L_n^m(x) L_{n'}^m(x) x^m\, e^{-x}\, dz = \frac{(m!)^2 n!}{(m+n)! 2^{m+1}}\delta_{nn'}. \tag{26}$$

In the present application, $0 \leq \sigma \leq 1/\theta_0^2$, on which interval the functions S_n^m are only approximately orthogonal. Because of the exponential damping of the S_n^m, their orthogonality is nearly exact for $\theta_0 \lesssim 1/2$.

2.1.3 The "Radial" Functions

We now turn to the "radial" functions $Z_n^m(\zeta)$ which obey the differential equation (17) with separation constant λ_m given by,

$$\lambda_m = \frac{4}{\theta_0^2}(2n + m + 1), \tag{27}$$

using eq. (21) with $\nu = 2n$. For large r the radial functions are essentially spherical waves, and hence have leading dependence e^{ikr}. For small polar angles, where $\xi \approx 1$, the relation (7) implies that $r \approx z_0\zeta$, and $kr \approx kz_0\zeta = 2\zeta/\theta_0^2$, recalling eq. (2). Hence, we expect the radial functions to have the form,[4]

$$Z(\zeta) = e^{2i\zeta/\theta_0^2} F(\zeta). \tag{28}$$

Inserting this in eq. (17), we find that function F obeys the second-order differential equation,

$$\left(1+\zeta^2\right)\left(\frac{d^2F}{d\zeta^2} + \frac{4i}{\theta_0^2}\frac{dF}{d\zeta}\right) + 2\zeta\left(\frac{dF}{d\zeta} + \frac{2iF}{\theta_0^2}\right) = \left(\frac{4}{\theta_0^2}(2n+m+1) - \frac{m^2}{1+\zeta^2}\right)F. \tag{29}$$

In the paraxial approximation, θ_0 is small, so we keep only those terms in eq. (29) that vary as $1/\theta_0^2$, which yields the first-order differential equation,

$$\left(1+\zeta^2\right)\frac{dF}{d\zeta} = -\left(\zeta + i(2n+m+1)\right)F. \tag{30}$$

For $m = n = 0$ we write $F_0^0 = f$, in which case eq. (30) reduces to,

$$\left(1+\zeta^2\right)\frac{df}{d\zeta} = (\zeta+i)(\zeta-i)\frac{df}{d\zeta} = -(\zeta+i)f, \tag{31}$$

or,

$$\frac{df}{f} = -\frac{d\zeta}{\zeta - i}. \tag{32}$$

This integrates to $\ln f = \ln C - \ln(\zeta - i)$. We define $f(0) = 1$, so that $C = -i$ and,

$$f = \frac{1}{1+i\zeta} = \frac{1-i\zeta}{1+\zeta^2} = \frac{e^{-i\tan^{-1}\zeta}}{\sqrt{1+\zeta^2}}. \tag{33}$$

At large ζ, $f \approx 1/\zeta \propto 1/r$, as expected in the far zone for waves that have a narrow waist at $z = 0$. Indeed, we expect that $F_n^m \propto 1/\zeta$ at large ζ for all m and n. This suggests that F_n^m differs from f by only a phase change. A suitable form is,

$$F_n^m = \frac{e^{-ia_{m,n}\tan^{-1}\zeta}}{\sqrt{1+\zeta^2}} = \frac{\left(e^{-i\tan^{-1}\zeta}\right)^{a_{m,n}}}{\sqrt{1+\zeta^2}} = \frac{1}{\sqrt{1+\zeta^2}}\left(\frac{1-i\zeta}{\sqrt{1+\zeta^2}}\right)^{a_{m,n}} = \frac{(1-i\zeta)^{a_{m,n}}}{\left(1+\zeta^2\right)^{(1+a_{m,n})/2}}. \tag{34}$$

Inserting this hypothesis in the differential equation (30), we find, that it is satisfied provided

$$a_{m,n} = 2n + m + 1. \tag{35}$$

Thus, the radial functions are,

$$Z_n^m(\zeta) = e^{ikz_0\zeta}F_n^m = \frac{e^{i[kz_0\zeta - (2n+m+1)\tan^{-1}\zeta]}}{\sqrt{1+\zeta^2}}. \tag{36}$$

[4]It turns out not to be useful to extract a factor e^{ikr}/r from the radial functions, although these functions will have this form asymptotically.

2.1.4 The Gaussian-Laguerre Wave Functions in Oblate Spheroidal Coordinates

Using forms (25) and (36) in eq. (13), the paraxial Gaussian-Laguerre wave functions are,

$$\psi_n^m(\sigma,\phi,\zeta,t) = Z_n^m S_n^m\, e^{\pm im\phi}\, e^{-i\omega t} = \frac{\sigma^{m/2} L_n^m(2\sigma)\, e^{-\sigma}\, e^{i[kz_0\zeta - \omega t - (2n+m+1)\tan^{-1}\zeta \pm m\phi]}}{\sqrt{1+\zeta^2}}\,. \quad (37)$$

The factor $e^{-i(2n+m+1)\tan^{-1}\zeta}$ in the wave functions implies a phase shift of $(2n+m+1)\pi/2$ between the focal plane and the far field, as first noticed by Gouy [23], for whom this effect is named. Even the lowest mode, with $m = n = 0$, has a Gouy phase shift of $\pi/2$. This phase shift is an essential difference between a plane wave and a wave that is asymptotically plane but which has emerged from a focal region. The existence of this phase shift can be deduced from an elementary argument that applies Faraday's law to wave propagation through an aperture [24], as well as by arguments based on the Kirchhoff diffraction integral [11] as were used by Gouy.

2.1.5 The Gaussian-Laguerre Wave Functions in Cylindrical Coordinates

It is useful to relate the coordinates σ and ζ to those of a cylindrical coordinate system $(r_\perp, \phi, z)$, in the paraxial approximation that $\xi \approx 1$. For this, we recall from eqs. (7), (8) and (11) that,

$$\xi = 1 - \theta_0^2\sigma \approx 1, \quad (38)$$

so,

$$\zeta^2 = \frac{z^2}{z_0^2\xi^2} \approx \frac{z^2}{z_0^2}(1+\theta_0^2\sigma), \quad (39)$$

and hence,

$$r_\perp^2 = w_0^2\sigma(1+\zeta^2) \approx w_0^2\sigma\left(1+\frac{z^2}{z_0^2}\right) + \theta_0^4\sigma z^2 \approx w_0^2\sigma\left(1+\frac{z^2}{z_0^2}\right), \quad (40)$$

where we neglect terms in θ_0^4 in the lowest-order paraxial approximation. Then,

$$\sigma \approx \frac{r_\perp^2}{w_0^2(1+z^2/z_0^2)}\,, \quad (41)$$

and,

$$\zeta \approx \frac{z}{z_0}\left(1+\frac{\theta_0^2\sigma}{2}\right) \approx \frac{z}{z_0}\left(1+\frac{\theta_0^2 r_\perp^2}{2w_0^2(1+z^2/z_0^2)}\right) = \frac{z}{z_0}\left(1+\frac{r_\perp^2}{2(z^2+z_0^2)}\right). \quad (42)$$

For large z eq. (42) becomes,

$$\zeta \approx \frac{z}{z_0}\left(1+\frac{r_\perp^2}{2z^2}\right) \approx \frac{z}{z_0}\sqrt{1+\frac{r_\perp^2}{z^2}} = \frac{r}{z_0}\,, \quad (43)$$

as expected. That is, the factor $e^{i(kz_0\zeta-\omega t)}$ in the wave functions (37) implies that they are nearly spherical waves in the far zone.

The characteristic transverse extent of the waves at position z is sometimes called $w(z)$. For nonzero n or m the "angular" behavior of eq. (37) leads to intensity maxima, *i.e.*, rings in the radiation pattern, for which $\sigma > 1$ at the outermost ring. The leading behavior of the Laguerre polynomial $L_n^m(2\sigma)$ on this ring is therefore σ^n according to eq. (22), and the "angular" factor is $\sigma^{m/2} L_n^m(2\sigma)\, e^{-\sigma} \propto \sigma^{n+m/2}\, e^{-\sigma}$. This implies that the outermost maximum occurs for $\sigma \approx n + m/2$. Then, eq. (41) indicates that the characteristic radius of the outermost ring is

$$w_n^m(z) = \sqrt{n + m/2}\, w_0^0(z), \qquad \text{where} \qquad w_0^0(z) = w_0 \sqrt{1 + \frac{z^2}{z_0^2}} = \sqrt{\frac{2(z^2 + z_0^2)}{k z_0}}\,. \tag{44}$$

The paraxial approximation is often taken to mean that variable ζ is simply z/z_0 everywhere in eq. (37) except in the phase factor $e^{ikz_0\zeta}$, where the form (42) is required so that the waves are nearly spherical in the far zone. In this convention, we can write,

$$\psi_n^m(r_\perp, \phi, z, t) = \frac{\sigma^{m/2} L_n^m(2\sigma)\, e^{-\sigma}\, e^{i\{kz[1 + r_\perp^2/2(z^2+z_0^2)] - \omega t - (2n+m+1)\tan^{-1}(z/z_0) \pm m\phi\}}}{\sqrt{1 + z^2/z_0^2}}\,. \tag{45}$$

The wave functions may be written in a slightly more compact form if we use the scaled coordinates,

$$\rho = \frac{r_\perp}{w_0}\,, \qquad \zeta = \frac{z}{z_0}\,, \qquad \sigma = \frac{\rho^2}{1 + \zeta^2}\,. \tag{46}$$

Then, the simplest wave function is,

$$\begin{aligned} \psi_0^0(r_\perp, \phi, z, t) &= e^{-\rho^2/(1+\zeta^2)}\, e^{ikzr_\perp^2/2z_0^2(1+\zeta^2)}\, e^{i(kz - \omega t)}\, \frac{e^{-i\tan^{-1}\zeta}}{\sqrt{1+\zeta^2}} \\ &= f\, e^{-\rho^2(1-i\zeta)/(1+\zeta^2)}\, e^{i(kz-\omega t)} = f\, e^{-f\rho^2}\, e^{i(kz-\omega t)}, \end{aligned} \tag{47}$$

recalling eq. (2) and the definition of $f(\zeta)$ in eq. (33). In this manner the general, paraxial wave function can be written as,

$$\psi_n^m(r_\perp, \phi, z, t) = f^{m+1} \rho^m L_n^m(2\sigma)\, e^{-f\rho^2}\, e^{i(kz - \omega t \pm m\phi - 2n\tan^{-1}\zeta)}. \tag{48}$$

It is noteworthy that although our solution began with the hypothesis of separation of variables in oblate spheroidal coordinates, we have found wave functions that contain the factors $e^{-f\rho^2}$ and $L_n^m(\sigma)$ that are nonseparable functions of $r_\perp$ and z in cylindrical coordinates.

2.1.6 Gaussian-Laguerre Wave Packets

The wave functions found above are for a pure frequency ω. In practice one is often interested in pulses of characteristic width τ in time, whose frequency spectrum is centered on frequency ω. In this case we can replace the factor $e^{i(kz-\omega t)}$ in the wave function by $g(\varphi)\, e^{i\varphi}$, where the phase is $\varphi = kz - \omega t$, and still satisfy the wave equation (3) provided that the modulation factor g obeys [12],

$$\left| \frac{g'}{g} \right| \ll 1. \tag{49}$$

An important example of a pulse shape that satisfies eq. (49) is,

$$g(\varphi) = \mathrm{sech}\frac{\varphi}{\omega\tau}\,, \tag{50}$$

so long as $\omega\tau \gg 1$, *i.e.*, so long as the pulse is longer that a few periods of the carrier wave. Perhaps surprisingly, a Gaussian temporal profile is not consistent with condition (49). Hence, a "Gaussian beam" can have a Gaussian transverse profile, but not a Gaussian longitudinal profile as well.

2.1.7 Gaussian-Laguerre Wave Functions for Large r in Spherical Coordinates

For $z \gg z_0$ the coordinate σ is given by $\sigma \approx (r_\perp/z\theta_0)^2 \approx (\theta/\theta_0)^2$, and the wavefunctions (45) become,

$$\psi_n^m(r,\theta,\phi,t) \approx z_0 \left(\frac{\theta}{\theta_0}\right)^m L_n^m(2\theta^2/\theta_0^2)\, e^{-\theta^2/\theta_0^2}\, \frac{e^{i[kr-\omega t-(2n+m+1)\pi/2\pm m\phi]}}{r} \qquad (z \approx r \gg z_0,\ \theta_0 \ll 1), \tag{51}$$

noting that $r = \sqrt{r_\perp^2 + z^2} \approx z(1 + r_\perp^2/2z^2)$.

The wavefront surface (*i.e.*, surface of constant phase $\varphi = kr \pm m\phi$) is not spherical for nonzero m, but has the form of a "hemispherical screw" of pitch $m\lambda$ per turn. The wave vector $\mathbf{k}$, which is normal to the wavefronts, is given by,

$$\mathbf{k} = \boldsymbol{\nabla}\varphi = k\,\hat{\mathbf{r}} \pm \frac{m}{r_\perp}\,\hat{\boldsymbol{\phi}} \approx k\left(\hat{\mathbf{z}} + \theta\,\hat{\mathbf{r}}_\perp \pm \frac{m}{kr_\perp}\,\hat{\boldsymbol{\phi}}\right), \tag{52}$$

where for the paraxial waves at large r,

$$\mathbf{r} = z\,\hat{\mathbf{z}} + r_\perp\,\hat{\mathbf{r}}_\perp \approx r(\hat{\mathbf{z}} + \theta\,\hat{\mathbf{r}}_\perp). \tag{53}$$

In this case the waves are not quite TEM, and the Poynting vector (energy flow vector) is not radial but moves along a conical spiral path, as considered further in the following sections.

Since the radial spacing between wavefronts is λ, for $m > 1$ there are m interleaved "hemispherical screw" wavefront surfaces.

The dispersion relation $\omega = kc$ can be written as $\omega(\mathbf{k}) = k_r c$ in view of eq. (52), so that the group velocity vector[5] is $\mathbf{v}_g = \boldsymbol{\nabla}_\mathbf{k}\omega(\mathbf{k}) = c\,\hat{\mathbf{r}}$, whose straight lines do not exhibit the conical form of lines of the wave vector $\mathbf{k}$.

Recasting the discussion around eq. (44) in terms of the far-zone angle θ, we see that when either n or m are nonzero, the characteristic angle of the far fields is,

$$\theta \approx \sqrt{n + m/2}\,\theta_0. \tag{54}$$

2.1.8 Gaussian-Laguerre Wave Functions Close to the Laser Focus

Close the the laser focus, where $r \ll z_0$, we have,

$$\zeta \approx \frac{z}{z_0} \ll 1, \qquad \sigma \approx \rho = \frac{r_\perp}{w_0} \lesssim 1, \tag{55}$$

[5] See, for example, sec. 2.1 of [25].

and the wavefunctions (45) become,

$$\psi_n^m \approx \rho^{m/2}\, e^{-\rho^2} L_n^m(\rho^2)\, e^{i\{[k-(2n+m+1)/z_0]z-\omega t\pm m\phi\}} \qquad (r \ll z_0). \tag{56}$$

The wave vector is,

$$\mathbf{k} \approx \boldsymbol{\nabla}\{[k-(2n+m+1)/z_0]z-\omega t\pm m\phi\} = [k-(2n+m+1)/z_0]\,\hat{\mathbf{z}} \pm \frac{m}{r_\perp}\,\hat{\boldsymbol{\phi}}. \tag{57}$$

For large indices m and n the z-component, $k_z = k-(2n+m+1)/z_0$ of the wave vector close to the focus is negative.

The dispersion relation can now be written as $\omega = kc = [k_z + (2n+m+1)/z_0]c$, so the group velocity vector is $\mathbf{v}_g = \boldsymbol{\nabla}_{\mathbf{k}}\omega(\mathbf{k}) = c\,\hat{\mathbf{z}}$. Thus, for high enough m and n, the laser focus is a region where the group and phase velocities are in opposite directions.[6]

2.2 Electric and Magnetic Fields of Gaussian Beams

The scalar wave functions (48) can be used to generate vector electromagnetic fields that satisfy Maxwell's equations. For this, we use eqs. (14) with a vector potential $\mathbf{A}$ whose Cartesian components are one or more of the functions (48).

If we wish to express the electromagnetic fields in cylindrical coordinates, then we immediately obtain one family of fields from the vector potential,

$$A_x = A_y = A_\perp = A_\phi = 0, \qquad A_z = \frac{E_0}{k\theta_0}\psi_n^m(r_\perp,\phi,z,t). \tag{58}$$

The resulting magnetic field has no z component, so we may call these transverse magnetic (TM) waves. If index $m = 0$ then $\mathbf{A}$ has no ϕ dependence, and the magnetic field has no radial component; the magnetic field lines are circles about the z axis.

The lowest-order TM mode, corresponding to indices $m = n = 0$, has field components,

$$\begin{aligned}
E_\perp &= E_0\rho f^2\, e^{-f\rho^2} g\, e^{i\varphi} + \mathcal{O}(\theta_0^2),\\
E_\phi &= 0,\\
E_z &= i\theta_0 E_0 f^2(1-f\rho^2)\, e^{-f\rho^2} g\, e^{i\varphi} + \mathcal{O}(\theta_0^3).\\
B_\perp &= 0,\\
B_\phi &= E_\perp,\\
B_z &= 0,
\end{aligned} \tag{59}$$

as apparently first deduced in [27]. This is a radially polarized mode, for which $E_\perp$ necessarily vanishes along the beam axis. In the far zone the beam intensity is largest on a cone of half angle θ_0 and is very small on the axis itself; the beam appears to have a hole in the center.

The radial and longitudinal electric field of the TM_0^0 mode are illustrated in Figs. 2 and 3. Photographs of Gaussian-Laguerre laser modes from [28] are shown in Fig. 4.

As is well known, corresponding to each TM wave solution to Maxwell's equations in free space, there is a TE (transverse electric) mode obtained by the duality transformation,

$$\mathbf{E}_{\mathrm{TE}} = \mathbf{B}_{\mathrm{TM}}, \qquad \mathbf{B}_{\mathrm{TE}} = -\mathbf{E}_{\mathrm{TM}}. \tag{60}$$

[6] In the ray-optics limit of very short wavelengths, *i.e.*, very large k, this behavior can be neglected [26].

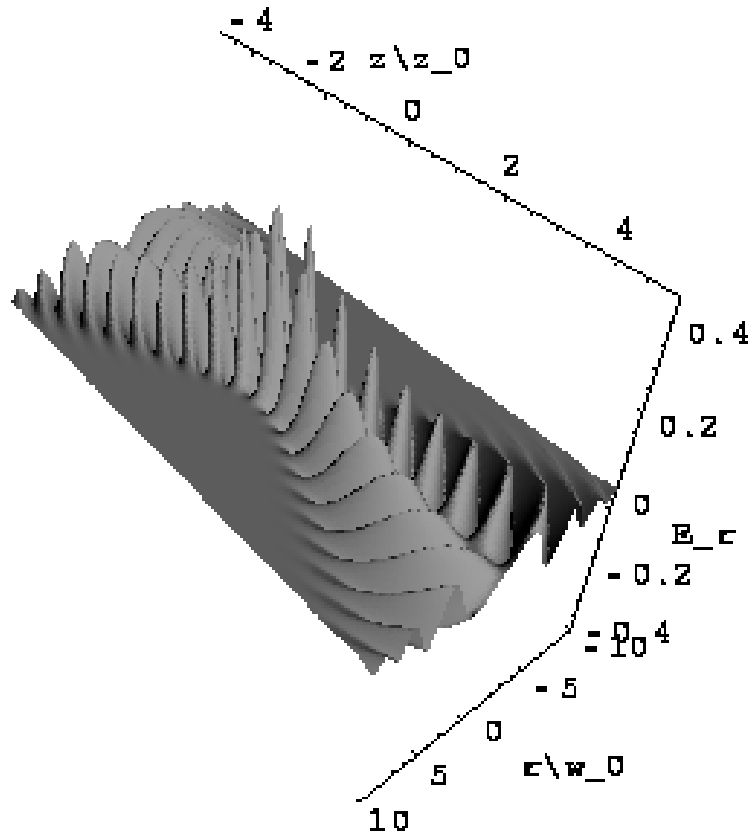

Figure 2: The electric field $E_r(r_\perp, 0, z)$ of the TM_0^0 radially polarized Gaussian beam with diffraction angle $\theta_0 = 0.45$, according to eq. (59).

Since we are considering waves in free space where $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$, the electric field could also be deduced from a vector potential, and the magnetic field from the electric field, according to the dual of eq. (14),

$$\mathbf{E} = \boldsymbol{\nabla} \times \mathbf{A}, \qquad \mathbf{B} = -\frac{i}{k} \boldsymbol{\nabla} \times \mathbf{E}. \tag{61}$$

Then, the TE modes can be obtained by use of the vector potential (58) in eq. (61).

The TM Gaussian-Laguerre modes emphasize radial polarization of the electric field, and the TE modes emphasize circular polarization. In many physical applications, linear polarization is more natural, for which the modes are well-described by Gaussian-Hermite wave functions [2, 3, 5, 9]. Formal transformations between the Gaussian-Hermite wave functions and the Gaussian-Laguerre functions have been described in [29]. Linearly polarized modes can also be obtained by supposing the nonzero component of the vector potential in eq. (61) is A_x or A_y rather than A_z.

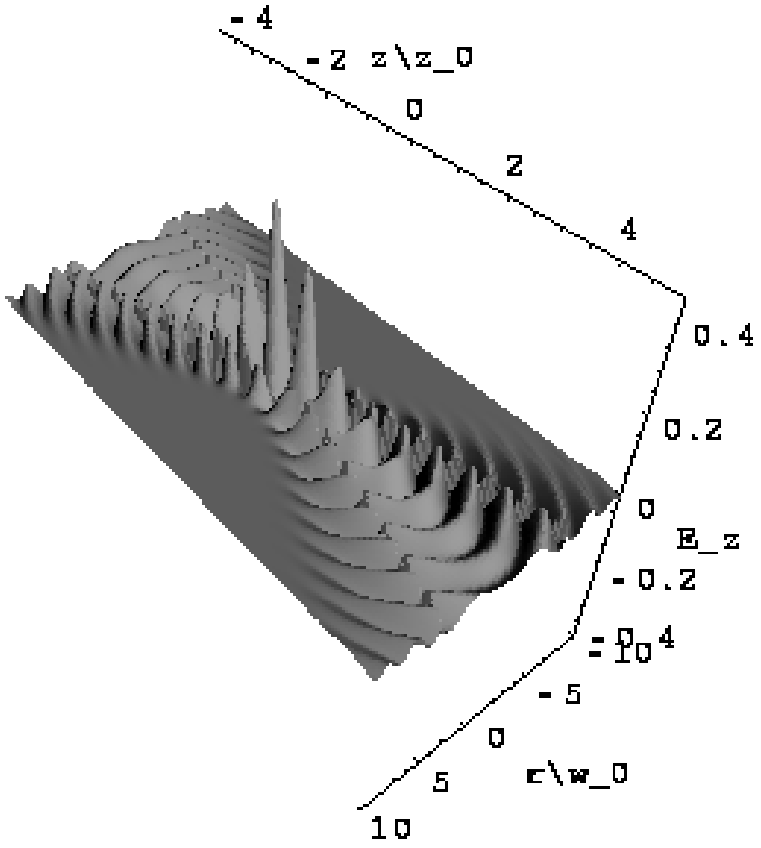

Figure 3: The electric field $E_z(r_\perp, 0, z)$ of the TM_0^0 radially polarized Gaussian beam with diffraction angle $\theta_0 = 0.45$, according to eq. (59).

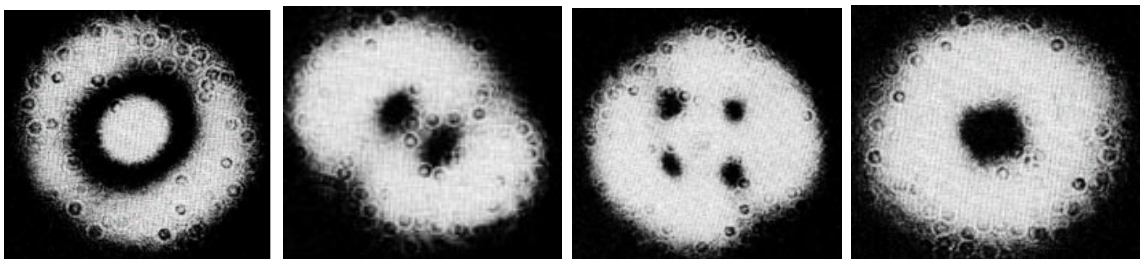

Figure 4: Photographs of Gaussian-Laguerre laser beams with $2n + m = 2$. (a) $n = 2$, $m = 0$, (b) $n = 0$, $m = 2$, $e^{2i\phi}$, (c) $n = 0$, $m = 2$, equal mixture of $e^{\pm 2i\phi}$ modes, (d) mixture of $\{n = 2, m = 0\}$ and both $\{n = 0, m = 2\}$ modes. From [28].

2.3 Energy, Momentum and Angular Momentum in the Far Zone

The electromagnetic field energy density,

$$u = \frac{E^2 + B^2}{8\pi}\,, \tag{62}$$

the field momentum density,

$$\mathbf{p} = \frac{\mathbf{E}\times\mathbf{B}}{4\pi c}\,, \tag{63}$$

and the field angular momentum density,

$$\mathbf{l} = \mathbf{r}\times\mathbf{p}, \tag{64}$$

are the same for a TM Gaussian-Laguerre mode and the TE mode related to it by the duality transformation (60).

We consider the energy, momentum and angular momentum for TM waves in the far zone, where $\zeta \approx r/z_0 \gg 1$, and $r_\perp \approx r\theta \ll r$ in terms of spherical coordinates (r,θ,ϕ). Then, the waves are nearly spherical, and so have a phase factor e^{ikr} that implies the electric field is related to the magnetic field by,[7]

$$\mathbf{E} = \frac{i}{k}\boldsymbol{\nabla}\times\mathbf{B} \approx \mathbf{B}\times\hat{\mathbf{k}} \approx \mathbf{B}\times\hat{\mathbf{r}}, \tag{65}$$

such that $E^2 = B^2$. The time-averaged densities can therefore be written,

$$\langle u\rangle = \frac{|B|^2}{8\pi}\,, \tag{66}$$

$$\langle\mathbf{p}\rangle = \frac{Re[(\mathbf{B}\times\hat{\mathbf{r}})\times\mathbf{B}^\star]}{8\pi c} = \frac{|B|^2\ \hat{\mathbf{r}} - Re[(\mathbf{B}\cdot\hat{\mathbf{r}})\mathbf{B}^\star]}{8\pi c} = \frac{\langle u\rangle}{c}\left(\hat{\mathbf{r}} - \frac{Re[(\mathbf{B}\cdot\hat{\mathbf{r}})\mathbf{B}^\star]}{|B|^2}\right). \tag{67}$$

The TM waves are derived from the vector potential (58) whose only nonzero component is A_z. Then, the magnetic field components in cylindrical coordinates are,[8]

$$B_\perp = \frac{1}{r_\perp}\frac{\partial A_z}{\partial\phi} = \frac{\pm imA_z}{r_\perp}, \qquad B_\phi = -\frac{\partial A_z}{\partial r_\perp} = -\frac{1}{w_0}\frac{\partial A_z}{\partial\rho}, \qquad B_z = 0, \tag{68}$$

where $\rho = r_\perp/w_0$. Recalling eq. (38), the factors of A_z that depend on ρ and $\sigma \approx \rho^2 z_0^2/z^2$ are,

$$A_z \propto \rho^m\, e^{-f\rho^2}\, L_n^m(2\sigma), \tag{69}$$

where in the far zone, $f(\zeta) \approx -iz_0/r$. Thus, writing $d\,L_n^m(x)/dx$ as $L'^m_n(x)$,

$$\frac{\partial A_z}{\partial\rho} = \left[\frac{m}{\rho^2} - 2f + \frac{2z_0^2}{z^2}\frac{L'^m_n(2\sigma)}{L_n^m(2\sigma)}\right]\rho A_z \approx 2i\frac{z_0}{r}\rho A_z \approx 2i\frac{\theta}{\theta_0}A_z, \tag{70}$$

since in the far zone where $z_0/r \ll 1$ the factor $\rho^m\, e^{-f\rho^2} \approx \rho^m\, e^{-\rho^2/\zeta^2}$ implies that the wave functions are large only for $\rho \approx \sqrt{m}\ \zeta \approx \sqrt{m}\ r/z_0 \gg 1$. Thus, in the far zone,

$$B_\phi \approx -\frac{2i}{w_0}\frac{\theta}{\theta_0}A_z \gg B_\perp. \tag{71}$$

[7]The approximation in eq. (65) implies that $\mathbf{E}$ is transverse to $\hat{\mathbf{r}}$ in the far field, which is not quite correct. From eq. (77) we see that $\hat{\mathbf{k}} = \langle\hat{\mathbf{p}}\rangle \approx \hat{\mathbf{r}} \pm m\,\hat{\boldsymbol{\phi}}/kr_\perp$, so there exists a small longitudinal component $E_r \approx E_z \approx \pm mB_\perp/kr_\perp \approx im^2A_z/kr_\perp^2 \propto 1/rr_\perp^2$ that affects eqs. (77)-(78) only in a higher approximation.

[8]In spherical coordinates we have $B_\theta \approx B_\perp$ and $B_r \approx \theta B_\perp \approx \pm imA_z/r \propto 1/r^2$.

Also,

$$|B|^2 = |B_\perp|^2 + |B_\phi|^2 \approx |B_\phi|^2 \approx \frac{4}{w_0^2}\frac{\theta^2}{\theta_0^2}\,|A_z|^2\,, \tag{72}$$

since $\rho = \pi r\theta_0/\lambda\theta \gg \theta_0/\theta$ in the far zone. Using this in eq. (66), we have

$$\langle u\rangle \approx \frac{1}{2\pi w_0^2}\frac{\theta^2}{\theta_0^2}\,|A_z|^2\,. \tag{73}$$

The radius vector is,

$$\mathbf{r} = z\,\hat{\mathbf{z}} + r_\perp\,\hat{\mathbf{r}}_\perp \approx r(\hat{\mathbf{z}} + \theta\,\hat{\mathbf{r}}_\perp), \tag{74}$$

so $\mathbf{B}\cdot\hat{\mathbf{r}} \approx \theta B_\perp$. Becasue this is small compared to B_ϕ, we approximate the momentum density (67) as,

$$\langle \mathbf{p}\rangle \approx \frac{\langle u\rangle}{c}\left(\hat{\mathbf{r}} - \frac{\theta Re(B_\perp B_\phi^\star)}{|B|^2}\,\hat{\boldsymbol{\phi}}\right). \tag{75}$$

From eqs. (68), (71) and (72) we have,

$$\theta Re(B_\perp B_\phi^\star) \approx \frac{\mp 2m\theta^2}{r_\perp w_0\theta_0}\,|A_z|^2 \approx \frac{\mp m w_0\theta_0}{2r_\perp}\,|B|^2 = \frac{\mp m}{kr_\perp}\,|B|^2\,, \tag{76}$$

recalling that $\theta_0 = 2/kw_0^2$. Thus, the electromagnetic momentum density is,

$$\langle \mathbf{p}\rangle \approx \frac{\langle u\rangle}{c}\left(\hat{\mathbf{r}} \pm \frac{m}{kr_\perp}\,\hat{\boldsymbol{\phi}}\right) \approx \frac{\langle u\rangle}{c}\left(\hat{\mathbf{z}} + \theta\,\hat{\mathbf{r}}_\perp \pm \frac{m}{kr_\perp}\,\hat{\boldsymbol{\phi}}\right). \tag{77}$$

Since $kr_\perp \gg 1$ in the far zone, the energy flow is largely radial outward from the focal region. The small azimuthal component for nonzero index m causes the lines of energy flow to become spirals, which lie on cones of constant polar angle θ in the far zone. The Poynting vector is normal to the wavefront surfaces that were mentioned briefly in sec. 2.1.7.

The angular momentum density is,

$$\langle \mathbf{l}\rangle = \mathbf{r}\times\langle \mathbf{p}\rangle = \mp\frac{\langle u\rangle\, mr}{kcr_\perp}\,\hat{\boldsymbol{\theta}} = \mp\frac{\langle u\rangle\, m}{\omega\theta}\,\hat{\boldsymbol{\theta}} \approx \pm\frac{\langle u\rangle\, m}{\omega}\left(\hat{\mathbf{z}} - \frac{1}{\theta}\,\hat{\mathbf{r}}_\perp\right), \tag{78}$$

noting that $\hat{\boldsymbol{\theta}} \approx \hat{\mathbf{r}}_\perp - \theta\,\hat{\mathbf{z}}$. There is no net angular momentum in any plane of fixed z coordinate. Of greater interest is the axial component of the (orbital) angular momentum, which obeys,

$$\frac{\langle l_z\rangle}{\langle u\rangle} \approx \frac{\pm m}{\omega}\,, \tag{79}$$

where the $\pm$ sign corresponds to azimuthal dependence $e^{\pm im\phi}$.

In a quantum view, the Gaussian-Laguerre mode with $A_z = \psi_n^m$ contains $N = \langle u\rangle/\hbar\omega$ photons per unit volume of energy $\hbar\omega$ each, so the classical result (79) implies that each of these photons carries orbital angular momentum component $l_z = \pm m\hbar$. Since the photons have intrinsic spin $S = 1$, with $S_z = \pm 1$, we infer that the photons of a Gaussian-Laguerre mode carry total angular momentum component $J_z = \pm m \pm 1$. Since the index n can take

on any nonnegative integer value for each value of index m, the index n is not a measure of the total angular momentum of a photon of the mode.

The angular momentum of Gaussian-Laguerre modes has also been discussed in [30], in a slightly different approximation. The first macroscopic evidence for the angular momentum of light appears to have been given in [31], following suggestions by Poynting [32] and Epstein [33].

3 Appendix: Ray Optics and Gaussian-Laguerre Beams

This Appendix was written in Jan. 2008, following e-discussions with Michael Berry.

While the Gaussian-Laguerre wave functions were deduced as approximate solutions to Maxwell's equations, it may be of interest to relate them to the older tradition of ray optics (now often called Hamiltonian optics).[9] Rays are lines of energy flow, *i.e.*, of group velocity, in the approximation that the wavelength and the extent of the source are small compared to the distance to the observer. In a homogeneous medium, such as vacuum, the rays are straight lines.

It turns out that when indices m and n are both large the Gaussian-Laguerre beams can be described in terms of skewed bundles of straight rays [26]. Far from the "focal plane" $z = 0$ the lines of Poynting vector of any Gaussian-Laguerre mode are essentially straight (and they apprear to emanate from the origin), so a kind of ray approximation holds for $z \gg |z_0|$. In the rest of this Appendix we discuss aspects of the slight depature from straightness of lines of the Poynting vector.

Consider the behavior of the Poynting vector for the outermost ring of the Gaussian-Laguerre mode of indices n and m. From eq. (44), the characteristc radius $r_\perp(z)$ of this ring is,

$$r_\perp(z) \approx w_n^m(z) = \sqrt{n + m/2}\sqrt{\frac{2(z^2 + z_0^2)}{kz_0}}\,. \tag{80}$$

From eq. (77) we see that a unit vector in the direction of the Poynting vector has an azimuthal component whose magnitude is $m/kr_\perp(z)$. Thus, as a line of the Poynting vector advances by dz it skews azimuthally by $m\,dz/kr_\perp(z)$. The corresponding change $d\phi$ in the azimuthal coordinate of the Poynting vector equals this skew divided by $r_\perp$, so we have that,

$$\frac{d\phi}{dz} = \frac{m}{kr_\perp^2(z)} \approx \frac{m}{2n+m}\frac{z_0}{z^2 + z_0^2}\,. \tag{81}$$

Integrating eq. (81) out from the "focal plane" at $z = 0$, we find the change in azimuth $\Delta\phi$ of a line of Poynting vector to be,

$$\Delta\phi(z) \approx \frac{m}{2n+m}\tan^{-1}(z/z_0) \leq \frac{\pi}{2}. \tag{82}$$

[9] The subject of ray optics is well reviewed by Landau and Lifshitz. although for the full benefit of their insights one should consult both sec. 53 of [34] and sec. 67 of [35]. An introduction by the present author is given in sec. 2.1 of [25].

Thus, while lines of the Poynting vector are not straight, the lines associated with the outermost ring of the mode skew by 90° or less between the focal plane and infinity,[10] and most of this skew occurs within a few Rayleigh ranges of $z = 0$.

Hence, lines of the Poynting vector are very close to being the straight rays of a beam as in the view of geometric optics. If the Poynting vector at a point $(r_\perp(z), \phi, z)$ in the far field is extrapolated back to the "focal plane" $z = 0$, the intercept is at radius $r_{\rm int} mz/kr_\perp(z) \approx m/k\theta$. For the outermost ring, $\theta \approx \sqrt{n + m/2}\theta_0$, so the intercept is at radius,

$$r_{\rm int} \approx \frac{m}{\sqrt{n + m/2}} \frac{1}{k\theta_0} = \sqrt{\frac{m/2}{1 + 2n/m}} w_0, \tag{83}$$

which is less than or equal to the characteristic radius $w_n^m(0)$ of the mode at $z = 0$, recalling eqs. (1) and (44). In the spirit of Keller's geometrical theory of diffraction [38], we suppose that a family of straight rays emanating from a ring of radius $r_{\rm int}$ in the plane $z = 0$ "interferes" such that the most prominent rays are those with an azimuthal skew when the index m is nonzero.[11]

It in noteworthy that the azimuthal skew (82) is proportional to the Gouy phase shift,

$$\Delta\phi_{\rm Gouy}(z) = (2n + m + 1) \tan^{-1}(z/z_0), \tag{84}$$

found in eq. (37), defining $\Delta\phi_{\rm Gouy}(0)$ to be zero.

A ray explanation of the Gouy phase has been suggested by Boyd [39]. As shown in Fig. 5, an ideal straight ray that makes angle θ to the z axis would pass through the center of the "focal plane" $z = 0$, while an actual diffracted ray follows the hyperbolic path,

$$r_\perp^2 = w^2 + \theta^2 z^2, \tag{85}$$

where $w = \theta z_0$ is the closest approach of the actual ray to the z axis.

The argument seems to be that the phase of the diffracted ray AB should be that of the ideal straight ray AC which is asymptotically the same as $z \to -\infty$. Since the length of the path AB is shorter than that of path AC by amount, say, ΔS, the phase of the diffracted ray includes an extra term $\Delta\phi_{\rm Gouy} = k\Delta S$.

Measuring from the origin, the path length of the ideal ray is,

$$S_C(z) = \frac{z}{\cos\theta} \approx z + \frac{z\theta^2}{2}\,, \tag{86}$$

while that of the diffracted ray is,

$$\begin{aligned} S_B(z) &= \int_0^z \sqrt{1 + (dr_\perp/dz)^2}\, dz \approx \int_0^z \left[1 + \frac{1}{2}\left(\frac{dr_\perp}{dz}\right)^2\right] dz = z + \frac{\theta^2}{2}\int_0^z \frac{z^2}{z^2 + z_0^2}\, dz \\ &= z + \frac{z\theta^2}{2} - \frac{\theta^2 z_0}{2} \tan^{-1}\frac{z}{z_0}\,. \end{aligned} \tag{87}$$

[10] The skew is greater on the inner rings that exist when $n > 1$, as discussed in [36, 37].

[11] As previously remarked, when indices m and n are both large, the beam can be represented in terms of a skewed family of straight rays such that the Poynting vector is everywhere tangent to the hyperboloids formed by the skewed rays [26].

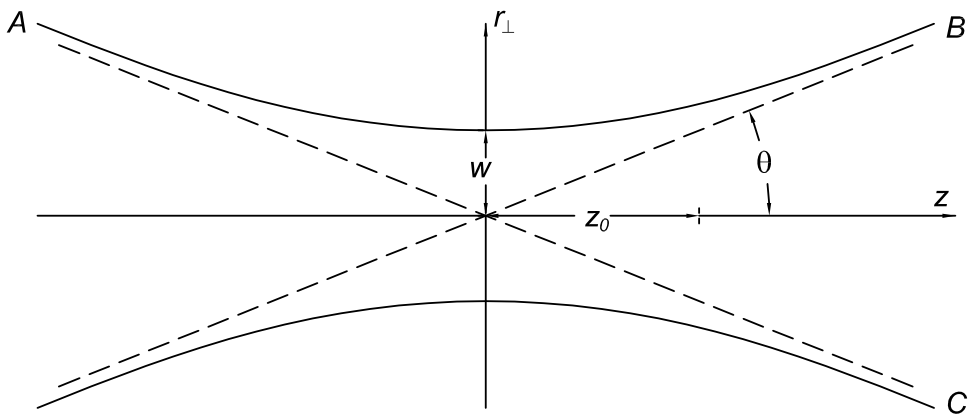

Figure 5: A diffracted ray AB and its corresponding ideal straight ray AC, both of which make angle θ to the z axis at large z. The trajectory of the diffracted ray, neglecting the small azimuthal skew, is given by eq. (85).

The estimate of the Gouy phase shift is then,

$$\Delta\phi_{\rm Gouy}(z) = k[S_B(z) - S_C(z)] = \frac{k\theta^2 z_0}{2}\tan^{-1}\frac{z}{z_0}\,. \tag{88}$$

This result has the correct dependence of the Gouy phase shift on z, although it predicts the phase shift would vary with angle. Using the angle $\theta \approx \sqrt{n+m/2}w_0/z_0$ of the outermost ring of a mode with indices n and m and recalling eq. (2), we find,

$$\Delta\phi_{\rm Gouy}(z) \approx (n+m/2)\tan^{-1}\frac{z}{z_0}\,, \tag{89}$$

which is half the nominal value (84). Thus, the ray explanation of the Gouy phase shift is suggestive rather than definitive.[12]

References

[1] Paraxial Gaussian laser beams were introduced nearly simultaneously from two different approaches in [2] and [3]. Gaussian-Laguerre functions were introduced in the Appendix of [4]. An influential review article is [5]. Corrections to the paraxial approximation were first organized in a power series in the parameter θ_0^2 in [6]. The understanding that the scalar, paraxial wave function is best thought of as a component of the vector potential was first emphasized in [7], with higher-order approximations discussed in [8]. A textbook with an extensive discussion of Gaussian beams is [9]. A recent historical review on the theory and experiment of laser beam modes is [10]. Other problems on Gaussian laser beams by the author include [11], [12] and [13].

[2] G. Goubau and F. Schwering, *On the Guided Propagation of Electromagnetic Wave Beams*, IRE Trans. Antennas and Propagation, **AP-9**, 248 (1961), http://kirkmcd.princeton.edu/examples/optics/goubau_iretap_9_248_61.pdf

[12] For another discussion of the Gouy phase shift for Gaussian-Laguerre beams, see [40].

[3] G.D. Boyd and J.P. Gordon, *Confocal Multimode Resonator for Millimeter Through Optical Wavelength Masers*, Bell Sys. Tech. J. **40**, 489 (1961), http://kirkmcd.princeton.edu/examples/optics/boyd_bstj_40_489_61.pdf

[4] G.D. Boyd and H. Kogelnik, *Generalized Confocal Resonator Theory*, Bell Sys. Tech. J. **41**, 1347 (1962), http://kirkmcd.princeton.edu/examples/optics/boyd_bstj_41_1347_62.pdf

[5] H. Kogelnik and T. Li, *Laser Beams and Resonators*, Appl. Opt. **5**, 1550 (1966), http://kirkmcd.princeton.edu/examples/optics/kogelnik_ao_5_1550_66.pdf

[6] M. Lax, W.H. Louisell and W.B. McKnight, *From Maxwell to paraxial wave optics*, Phys. Rev. A **11**, 1365 (1975), http://kirkmcd.princeton.edu/examples/optics/lax_pra_11_1365_75.pdf

[7] L.W. Davis, *Theory of electromagnetic beams*, Phys. Rev. A **19**, 1177 (1979), http://kirkmcd.princeton.edu/examples/optics/davis_pra_19_1177_79.pdf

[8] J.P. Barton and D.R. Alexander, *Fifth-order corrected electromagnetic field components for a fundamental Gaussian beam*, J. Appl. Phys. **66**, 2800 (1989), http://kirkmcd.princeton.edu/examples/optics/barton_jap_66_2800_89.pdf

[9] A.E. Siegman, *Lasers* (University Science Books, Mill Valley, CA, 1986), chaps. 16-17, http://kirkmcd.princeton.edu/examples/EM/siegman_86.pdf

[10] A.E. Siegman, *Laser Beams and Resonators: The 1960s*; and *Beyond the 1960s*, IEEE J. Sel. Topics Quant. El. **6**, 1380, 1389 (2000), http://kirkmcd.princeton.edu/examples/optics/siegman_ieeestqe_6_1380_00.pdf http://kirkmcd.princeton.edu/examples/optics/siegman_ieeestqe_6_1389_00.pdf

[11] M.S. Zolotorev and K.T. McDonald, *Time Reversed Diffraction* (Sept. 5, 1999), physics/0003058, http://kirkmcd.princeton.edu/examples/laserfocus.pdf

[12] K.T. McDonald, *Gaussian Laser Beams with Radial Polarization* (Mar. 14, 2000), physics/0003056, http://kirkmcd.princeton.edu/examples/axicon.pdf

[13] K.T. McDonald, *Bessel Beams* (June 17, 2000), physics/0006046, http://kirkmcd.princeton.edu/examples/bessel.pdf

[14] L.P. Eisenhart, *Separable Systems of Stäckel*, Ann. Math. **35**, 284 (1934), http://kirkmcd.princeton.edu/examples/EM/eisenhart_am_35_284_34.pdf

[15] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Part I (McGraw-Hill, 1953), pp. 115-116, 125-126 and 509-510, http://kirkmcd.princeton.edu/examples/EM/morse_feshbach_v1.pdf

[16] J.A. Stratton *et al.*, *Spheroidal Wave Functions* (Wiley, 1956).

[17] C. Flammer, *Spheroidal Wave Functions* (Stanford U. Press, 1957).

[18] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Wiley, 1984), chap. 21, http://kirkmcd.princeton.edu/examples/EM/abramowitz_and_stegun.pdf

[19] Secs. 8.2.1 and 8.2.1 of [17] discuss the angular and radial functions for oblate spheroidal waves in the limit of large z_0 ($= ic$ in [17]). Results closely related to our eq. (25) are obtained for the asymptotic behavior of the angular functions, but nothing like the simplicity of our eq. (36) is obtained for the asymptotic radial functions. The work of Flammer, Stratton, *et al.* seems to have been little guided by the physical significance of the parameters θ_0, w_0 and z_0 of electromagnetic waves that are strong only near an axis, and consequently had little direct impact on the later development of approximate theories of such waves.

Rather, the classic application of spheroidal wave functions was to problems in which z_0 had the physical significance of a transverse aperture. The utility of spheroidal wave functions for problems in which there is no physical aperture, but in which waves have a narrow waist, was not appreciated in [16, 17]. This oversight extends to works that emphasize the focal region of optical beams, such as M. Born and E. Wolf, *Principles of Optics*, 7$^{\rm th}$ ed. (Cambridge U. Press, 1999),
http://kirkmcd.princeton.edu/examples/EM/born_wolf_7ed.pdf
where the assumption that the beams have closely filled an aperture not in the focal plane implies non-Gaussian transverse profiles in the focal plane.

The Gaussian beams discussed here can be realized in the laboratory only if the beams do not fill any apertures in the optical transport. Prior to the invention of the laser, and the availability of very high power beams, little attention was paid to problems in which optical apertures were large compared to the beam size. Gaussian beams came into prominence in considerations of modes in a "cavity" formed by a pair of mirrors, in which the beam size should be smaller than the transverse size of the mirrors to prevent leakage beyond the mirror edges during multiple beam passes [3, 5].

[20] B.T. Landesman and H.H. Barrett, *Gaussian amplitude functions that are exact solutions to the scalar Helmholtz equation*, J. Opt. Soc. Am. A **5**, 1610 (1988),
http://kirkmcd.princeton.edu/examples/EM/landesman_josa_a5_1610_88.pdf
B.T. Landesman, *Geometrical representation of the fundamental mode of a Gaussian beam in oblate spheroidal coordinates*, J. Opt. Soc. Am. A **6**, 5 (1989),
http://kirkmcd.princeton.edu/examples/EM/landesmann_josa_a6_5_89.pdf

[21] R. Boyack and J. Lekner, *Non-existence of separable spheroidal beams*, J. Opt. **13**, 085701 (2011), http://kirkmcd.princeton.edu/examples/EM/boyak_jo_13_085701_11.pdf

[22] The generalized Laguerre polynomials $L_n^m(x)$ defined in eq. (22) are the Kummer polynomials $M(a, b, z)$ discussed in sec. 13 of [18], with $a = -n$, $b = m + 1$ and $z = x$. The orthogonality relation (24) is deducible from 22.2.12 of [18], p. 774.

[23] G. Gouy, *Sur une propreite nouvelle des ondes lumineuases*, Compt. Rendue Acad. Sci. (Paris) **110**, 1251 (1890); *Sur la propagation anomele des ondes*, *ibid.* **111**, 33 (1890),
http://kirkmcd.princeton.edu/examples/optics/gouy_cr_110_1251_90.pdf
http://kirkmcd.princeton.edu/examples/optics/gouy_cr_111_33_90.pdf

[24] M.S. Zolotorev and K.T. McDonald, *Diffraction as a Consequence of Faraday's Law*, Am. J. Phys. **68**, 674 (2000), physics/0003057, http://kirkmcd.princeton.edu/examples/diffraction.pdf

[25] K.T. McDonald, *Flow of Energy from a Localized Source in a Uniform Anisotropic Medium* (Dec. 8, 2007), http://kirkmcd.princeton.edu/examples/biaxial.pdf

[26] M. Berry and K.T. McDonald, *Exact and geometrical-optics energy trajectories in twisted beams*, J. Opt. A 10, 03505 (2008), http://kirkmcd.princeton.edu/examples/optics/berry_twistedbeams.pdf

[27] L.W. Davis and G. Patsakos, *TM and TE electromagnetic beams in free space*, Opt. Lett. **6**, 22 (1981), http://kirkmcd.princeton.edu/examples/optics/davis_ol_6_22_81.pdf

[28] M. Brambilla *et al.*, *Transverse laser patterns. I. Phase singularity crystals*, Phys. Rev. A **43**, 5090 (1991), http://kirkmcd.princeton.edu/examples/optics/brambilla_pra_43_5090_91.pdf

[29] E. Abramochkin and V. Volostnikov, *Beam transformations and nontransformed beams*, Opt. Comm. **83**, 123 (1991), http://kirkmcd.princeton.edu/examples/optics/abramochkin_oc_83_123_91.pdf

[30] L. Allen *et al.*, *Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes*, Phys. Rev. A **45**, 8185 (1992), http://kirkmcd.princeton.edu/examples/optics/allen_pra_45_8185_92.pdf

[31] R.A. Beth, *Mechanical Detection and Measurement of the Angular Momentum of Light*, Phys. Rev. **50**, 115 (1936), http://kirkmcd.princeton.edu/examples/optics/beth_pr_50_115_36.pdf

[32] J.H. Poynting, *The Wave Motion of a Revolving Shaft, and a Suggestion as to the Angular Momentum in a Beam of Circularly Polarised Light*, Proc. Roy. Soc. London A **82**, 560 (1909), http://kirkmcd.princeton.edu/examples/EM/poynting_prsla_82_560_09.pdf

[33] P.S. Epstein, *Die ponderomotorischen Drehwirkungen einer Lichtwelle und die Impulssätze der Elektronentheorie*, Ann. Phys. **44**, 593 (1914), http://kirkmcd.princeton.edu/examples/EM/epstein_ap_44_593_14.pdf

[34] L.D. Landau and E.M. Lifshitz, *Classical Theory of Fields*, 4th ed. (Butterworth-Heinemann, 1975), sec. 53, http://kirkmcd.princeton.edu/examples/EM/landau_ctf_71.pdf The first Russian edition appeared in 1941. kirkmcd.princeton.edu/examples/EM/landau_teoria_polya_41.pdf

[35] L.D. Landau and E.M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Butterworth-Heinemann, 1998), sec. 67, http://kirkmcd.princeton.edu/examples/fluids/landau_fluids_59.pdf

[36] M.J. Padgett and L. Allen, *The Poynting vector in Laguerre-Gaussian laser modes*, Opt. Comm. **121**, 36 (1995), http://kirkmcd.princeton.edu/examples/optics/padgett_oc_121_36_95.pdf

[37] L. Allen and M.J. Padgett, *The Poynting vector in Laguerre-Gaussian laser modes and the interpretation of their angular momentum density*, Opt. Comm. **184**, 67 (2000), http://kirkmcd.princeton.edu/examples/optics/allen_oc_184_67_00.pdf

[38] J.B. Keller, *Geometrical Theory of Diffraction*, J. Opt. Soc. Am. **52**, 116 (1962), http://kirkmcd.princeton.edu/examples/optics/keller_josa_52_116_62.pdf

[39] R.W. Boyd, *Intuitive explanation of the phase anomaly of focused light beams*, J. Opt. Soc. Am. **70**, 877 (1980), http://kirkmcd.princeton.edu/examples/optics/boyd_josa_70_877_80.pdf

[40] S. Feng and H.G. Winful, *Physical origin of the Gouy phase shift*, Opt. Lett. **26**, 485 (2001), http://kirkmcd.princeton.edu/examples/optics/feng_ol_26_485_01.pdf