Ray-optical Poincaré sphere for structured Gaussian beams

MIGUEL A. ALONSO1,2,* AND MARK R. DENNIS3

1The Institute of Optics, University of Rochester, Rochester, New York 14627, USA
2Center for Coherence and Quantum Optics, University of Rochester, Rochester, New York 14627, USA
3H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK
*Corresponding author: alonso@optics.rochester.edu

Received 20 December 2016; revised 26 March 2017; accepted 27 March 2017 (Doc. ID 283263); published 20 April 2017

A general family of scalar structured Gaussian beams naturally emerges from a consideration of families of rays. These ray families, with the property that their transverse profile is invariant upon propagation (except for a global rescaling), have two parameters, the first giving a position on an ellipse naturally represented by a point on a ray-family analog of the Poincaré sphere (familiar from polarization optics), and the other determining the position of a curve traced out on this Poincaré sphere. This construction naturally accounts for the well-known families of Gaussian beams, including Hermite–Gaussian, Laguerre–Gaussian, and generalized Hermite–Laguerre–Gaussian beams, but is far more general, opening the door for the design of a large variety of propagation-invariant beams. This ray-based description also provides a simple explanation for many aspects of these beams, such as “self-healing” and the Gouy and Pancharatnam–Berry phases. Further, through a conformal mapping between a projection of the Poincaré sphere and the physical space of the transverse plane of a Gaussian beam, the otherwise hidden geometric rules behind the beam’s intensity distribution are revealed. While the treatment is based on rays, a simple prescription is given for recovering exact solutions to the paraxial wave equation corresponding to these rays.

© 2017 Optical Society of America

OCIS codes: (080.7343) Wave dressing of rays; (070.2580) Paraxial wave optics; (030.4070) Modes.

https://doi.org/10.1364/OPTICA.4.000476

1. INTRODUCTION

Structured Gaussian beams are among the most familiar examples of paraxially propagating light beams. These include the Hermite–Gaussian (HG) beams [1,2], with intensity patterns resembling Cartesian grids, and Laguerre–Gaussian (LG) beams [3], whose intensities are concentric rings and whose phase can carry orbital angular momentum (OAM). A remarkable feature of Gaussian beams is that their intensity profile does not change on propagation, apart from an overall scaling; even in the far field, HG and LG modes appear the same. More recently, other self-similar beams have been studied in detail, including Airy beams [4,5], which are self-similar on propagation up to a parabolic lateral displacement, and Bessel beams [6,7] and Mathieu beams [8], whose intensity profile does not change at all on propagation. Self-similarity is more than a mathematical peculiarity, and is an important aspect of many applications of structured light. For example, because of the constant width of their main intensity lobe, approximations of Airy and Bessel beams have been the basis of several imaging techniques, whether for illumination to increase axial resolution [9,10] or for 3D shaping of the point spread function to increase depth of focus [11–15]. Given their characteristic intensity and phase profiles, structured Gaussian and other self-similar beams have been used extensively for particle manipulation [16–20], and may be eigenfunctions of natural optical operators, such as Bessel beams and LG beams of azimuthal order \( \ell \), which carry an OAM of \( \ell \hbar \) per photon [3].

Here we describe an approach to structured Gaussian beams in terms of rays. Ray optics is usually applied in situations where a light field has well-defined extended wavefronts whose wavelength is small with respect to slow amplitude variations (since these wavefronts can then be thought of as surfaces of constant eikonal or optical path length (OPL)). It turns out that Gaussian beams and their generalizations are remarkably amenable to such an analysis. Because fundamental Gaussian beams, as well as HG and LG beams, are modes of laser cavities with curved mirrors [1,2], their dynamics are well approximated by a two-dimensional isotropic harmonic oscillator representing the transverse plane, with the mirror curvature acting as the harmonic potential. Classical orbits in the two-dimensional isotropic oscillator are, of course, ellipses, and can be represented by points on the Poincaré sphere, more familiar in representing the polarization of a harmonic electric field [21]. In our analysis, a Gaussian beam is represented by a two-parameter family of rays; the rays are divided into subfamilies describing ellipses that propagate in a self-similar way, and that correspond to points on a Poincaré-like...
sphere for rays. We stress that this Poincaré sphere characterizes ray structure and not polarization; only scalar fields are considered throughout. The choice of the other parameter of the ray family then corresponds to determining a closed path of ray ellipses on this sphere, which is different for different types of beams. Consistency of the rays with wave optics forces quantization conditions on these parameters, both around the ellipse and on the Poincaré sphere path. These conditions give, for certain natural choices of path, the quantum numbers associated with HG and LG modes [22], and this approach admits arbitrarily many new kinds of structured Gaussian beams. An immediate generalization is to the generalized Hermite–Laguerre–Gaussian beams (HLG) [23–29], which interpolate between the HG and LG families on a generalized Poincaré sphere via an anamorphic fractional Fourier transformation, realized physically by transforming HG or LG beams through a beam shaping device consisting of suitably chosen pairs of cylindrical lenses [30–32].

We therefore are discussing objects very familiar in modern paraxial optics: mode families, optical operators, geometric optics, and Poincaré spheres, although they are combined in what we believe is a new way. The approach describes the general behavior of scalar beams to the level of providing interpretations of the Gouy phase (whose significance has long been disputed) [33–36] and the geometric (Pancharatnam–Berry) phase [37–39], in a way that reveals the hidden geometry behind the transverse spatial structure of these familiar light beams. Furthermore, well-established methods of approximating the wave fields from the ray family are highly efficient for this approach, and easily give the analytic forms for HG and LG beams. In a way, it forms a more complete and intuitive approach to our operator-based description of Gaussian beams in [22]. Our emphasis throughout is on recasting known properties of Gaussian beam families in terms of rays; the methods can be readily adapted as a design tool for new kinds of structured light.

Rather than derive the ray patterns from the known forms of the fields, we construct ray families for structured Gaussian beams from first principles, and later prove these to be the geometric optics analogs of the known families of beams. As in all geometric optics, we consider each ray in our construction to carry a complex amplitude whose phase increases with OPL, so that the scalar field at each point in the wave field is the sum of the amplitudes of the rays associated with that point. All Gaussian beams we consider are self-similar on propagation, and their width spreads hyperbolically; this property leads us to associate all the rays intersecting an ellipse (described by a point on the Poincaré sphere) in the transverse plane with the same weighting, and the beam is made up of a one-parameter family of such ellipses.

The structure of this paper proceeds as follows. In the next section we discuss elliptic families of rays and show how they are associated with a Poincaré sphere. This is followed in Section 3 by a discussion on their quantization, and in Section 4 by their geometric representation. Families of these ellipses and their quantization are discussed in Section 5, which are combined to give a general method of constructing approximate wave solutions (Section 6), which are then applied to the HG, LG, and GHL beams (Section 7). Properties such as Gouy and geometric phases (Section 8) and the generalization to other beam families such as Bessel and Airy beams (Section 9) are followed by a concluding discussion. Many additional proofs and derivations are presented in Supplement 1.

2. ELLIPTIC ORBITS AND THE POINCARÉ SPHERE

In ray optics, one expects the complex amplitude function representing a propagating, coherent monochromatic scalar light field to be associated with a two-parameter family of light rays [40,41]. For fields with slowly varying intensities such as plane or spherical waves, the ray directions are normal to the wavefronts and the intensity is proportional to the ray density. However, for fields with more a spatial structure, the ray–wave connection is more subtle, as several rays may pass through a given point, albeit with different directions, accounting for interference. In general, families of rays are bounded by envelopes known as caustics [42,43]. Interference-fringe-like structures can be caused by overlapping sets of rays propagating in different directions. Near caustics or other features of structured light, the rays can differ significantly from the wavefront normals, but there is still a tight link between the wave and ray descriptions, and it is possible to accurately reconstruct the wave field by associating a complex contribution to the rays [44–48]. We will see how the geometry of structured Gaussian beams can be readily understood using rays.

We assume the beam propagates in a linear, homogeneous, and isotropic medium, so the rays are straight lines. As is standard practice in paraxial ray optics, each ray is labeled by the transverse coordinate \( Q = (Q_x, Q_y) \), where it crosses the \( z = 0 \) plane, and its transverse direction vector \( P = (P_x, P_y) \), equal to the direction cosines of the ray in the \( x \) and \( y \) directions times the refractive index. (Note that the paraxial approximation implies \( |P| \ll 1 \).) The equation for the point where the ray crosses a plane of constant \( z \) is therefore \( Q + zP \). For a beam to be self-similar on propagation, the distribution of the rays in density and direction should be the same (apart from overall scaling) as \( z \) increases. As shown in Fig. 1(a), the shape and orientation of the elliptic cross section of a one-parameter family of rays remain unchanged on propagation if they conform to what is known as a ruled hyperboloid; endowing each ray in this one-dimensional subfamily with the same amplitude indeed guarantees its self-similarity on propagation. The two-parameter family of rays making up structured Gaussian beams is therefore a one-parameter superfamily of elliptic families of rays like the one in Fig. 1(a). We will call each such elliptic family an orbit of rays. It is very convenient to use the

---

![Fig. 1. Ray orbit in real space and on the Poincaré sphere. (a) The straight rays sweep out a hyperboloid whose cross sections at any constant \( z \) are ellipses with the same eccentricity and orientation. The green curve is a normal to the rays. The length of the orange ray segment must be an integer multiple of the wavelength. (b) The eccentricity and orientation of the ellipse correspond to a point on the Poincaré sphere that has Cartesian coordinates \( s_1, s_2, \) and \( s_3 \), given by the analogs of the Stokes parameters.](image-url)
parameterization of oriented ellipses afforded by the Poincaré sphere, borrowing language from polarization optics to describe these elliptic orbits of rays.

The Poincaré sphere for polarization parameterizes the two-dimensional complex Jones vectors \( \mathbf{v} \) satisfying \( \mathbf{v}^* \cdot \mathbf{v} = 1 \), and \( \mathbf{v} \) and \( \mathbf{v} \exp(-i\tau) \) are associated with the same polarization state for any real \( \tau \). \( \mathbf{v} \) is defined in terms of latitude \( \theta \) (not colatitude) and azimuth \( \phi \) on the unit Poincaré sphere:

\[
\mathbf{v}(\theta, \phi) = \cos \frac{\theta}{2} \left( \cos \frac{\phi}{2}, \sin \frac{\phi}{2} \right) + i \sin \frac{\theta}{2} \left( -\sin \frac{\phi}{2}, \cos \frac{\phi}{2} \right),
\]

where \( -\frac{1}{2} \leq \theta \leq \frac{1}{2} \pi \) and \( 0 \leq \phi < 2\pi \). As \( \tau \) varies within \( 0 \leq \tau < 2\pi \), \( \mathcal{R}[\mathbf{v} \exp(-i\tau)] \) traces out the ellipse. In polarization optics, \( \mathbf{v} \) represents the transverse harmonic electric field; with \( \tau \) evolving as time, the real part gives the ellipse; for each \( \tau \), the imaginary part is the velocity of the electric field vector [22].

An analogous characterization can be used for the elliptic orbits of rays, in which the angles \( \theta \) and \( \phi \) are fixed parameters that determine the eccentricity and orientation of the elliptic cross section. Similarly, an analog Jones vector \( \mathbf{v}(\theta, \phi) \) can be used to parameterize the position and direction of the rays in the orbit. A given ray, labeled by \( \tau \), has the \( z = 0 \) position

\[
\mathbf{Q}(\tau; \theta, \phi) = \mathbf{Q}_0 \mathcal{R}[\mathbf{v}(\theta, \phi) \exp(-i\tau)],
\]

where the constant \( \mathbf{Q}_0 \) sets the transverse scale. The ellipse’s major and minor semiaxes are \( Q_0 \cos \frac{\theta}{2} \) and \( Q_0 \sin \frac{\theta}{2} \), and its foci are \( f_{\pm} = \pm Q_0 \cos \frac{\theta}{2} \phi \sin \frac{1}{2} \phi \). The ray’s transverse direction, on the other hand, is given by the imaginary part

\[
\mathbf{P}(\tau; \theta, \phi) = P_0 \mathcal{R}[\mathbf{v}(\theta, \phi) \exp(-i\tau)],
\]

where \( P_0 \) is a constant determining the beam’s angular divergence. Thus, at any \( z \), the transverse ray coordinates given by \( \mathbf{Q} + z\mathbf{P} \) trace the same ellipse as \( \tau \) varies, up to a global hyperbolic scaling:

\[
\mathbf{Q} + z\mathbf{P} = \sqrt{Q_0^2 + z^2 P_0^2} \mathcal{R}[\mathbf{v}(\theta, \phi) \exp(-i(\tau + \zeta))],
\]

where \( \zeta = \arctan(zP_0/Q_0) \). As the beam evolves, the position of each ray around the ellipse changes with \( z \) (hence “orbit”), but the orientation and eccentricity are unchanged, as shown in Fig. 1(a) and Visualization 1.

We stress that the parameterization of elliptic orbits of rays by a Poincaré sphere is different physically from polarization. The similarity originates from the fact that mathematically, the Poincaré sphere parameterizes the classical orbits of the isotropic two-dimensional harmonic oscillator (like a transversely oscillating monochromatic electric field). Less obviously, ray families propagating back and forth in laser cavities also behave like classical harmonic oscillators, as the curvature of the spherical mirrors effectively acts as an attractive harmonic potential for the rays. Structured Gaussian beams are made up of families of orbits described by paths on the Poincaré sphere. First we discuss how the ray family is made compatible with the wave picture by a semiclassical “quantization condition.”

### 3. Quantization Condition for the Orbits

Making the ray families consistent with wave optics requires two closure conditions dictated by the field’s wavelength \( \lambda \). The allowed solutions with certain properties (such as quantized OAM) are discrete, and often can be expressed as eigenfunctions of certain operators. These conditions are mathematically analogous to those in quantum mechanics, so we refer to them as quantization conditions. The first condition applies to the orbits. Since the rays in an orbit are skewed, a curve normal to them does not close onto itself after tracing the orbit [such as the thick green curve in Fig. 1(a)]. There is a path difference along a ray between the initial and final points [represented by the orange line segment in Fig. 1(a)]. Since OPL times wavenumber corresponds to the phase of the complex amplitude along the ray, this path difference must be an integer multiple of the wavelength for the ray and wave pictures to be consistent.

This condition is expressed mathematically as follows. Let \( L_1(\tau) \) represent the OPL (also referred to as the eikonal) along each ray in the orbit, from some reference surface normal to the rays up to the \( z = 0 \) plane. The rays’ inclination is determined by \( \mathbf{P}(\tau) \), so in a small increment \( d\tau \), the infinitesimal path length \( dL_1 \) accumulated by the corresponding small step \( dQ \) along the ray family is \( dL_1 = \mathbf{P} \cdot d\mathbf{Q} \). From Eqs. (2) and (3), the OPL difference between any pair of rays labeled by \( \tau_1 \) and \( \tau_2 \) is then

\[
L_1(\tau_2) - L_1(\tau_1) = \int_{\tau_1}^{\tau_2} \mathbf{P} \cdot d\mathbf{Q} = \frac{Q_0 P_0}{2} \left( \tau_2 - \tau_1 - \frac{\sin(2\tau_2) - \sin(2\tau_1)}{2} \cos \theta \right).
\]

After tracing the entire orbit, the total OPL mismatch is \( L_1(2\pi) - L_1(0) = \pi Q_0 P_0 \), so the quantization condition yields

\[
Q_0 P_0 = (N + 1)\lambda/\pi,
\]

where \( N \) is a nonnegative integer. Significantly, this condition does not involve \( \theta \) and \( \phi \). Since \( Q_0 \) and \( P_0 \) describe the waist size and directional spread of the beam, respectively, \( Q_0 P_0 / \lambda \) is the beam quality factor \( M^2 \) [49–51], usually defined as the ratio of the product of the spatial and directional widths of a beam to the same product for a fundamental Gaussian beam. Therefore, the beam quality factor of fields made up of orbits satisfying Eq. (6) is quantized according to \( M^2 = N + 1 \). As we will discuss later, this index is also proportional to the beam’s Gouy phase shift. In the quantum mechanical analogy, the integral [Eq. (5)] plays the role of the semiclassical Bohr–Sommerfeld integral, whose quantization [Eq. (6)] corresponds to energy quantization. Constructing light beams from orbits satisfying Eq. (6) with the same \( N \) guarantees that the profile will have a well-defined beam quality factor and Gouy phase, as the beam is constructed to be an eigenfunction of the corresponding Hamiltonian operator.

### 4. Poincaré Equatorial and Physical Disks

The shape of an elliptical orbit depends on \( \theta \), with sign\( \theta \) determining the sense of twist of the rays around the ellipse under propagation in \( z \) [that is, the sign of the \( z \) component of the OAM of the orbit \( \mathbf{Q} \times \mathbf{P} \), which is positive (counterclockwise) in Fig. 1(a)]. The beam’s intensity profile is independent of this sign, so it is convenient to project both hemispheres of the Poincaré sphere, \( \theta \geq 0 \), onto the unit Poincaré equatorial disk (PED) (the term Poincaré disk is already used for a geometrical object [52]), with coordinates \( s = (s_1, s_2) = \cos \theta (\cos \phi, \sin \phi) \), and \( |s|^2 \leq 1 \). In the real space describing the transverse plane of the beam, we define the normalized ray position as
\[ q = Q/Q_0. \] The orbits are constrained to the interior of the unit disk \(|q| \leq 1\), which we call the physical disk, since it is a scaled version of a cross section of the beam (for any \(z\)). As Fig. 2 shows, a point \(s\) in the PED maps to an ellipse in the physical disk with foci \(f_{\pm} = \pm \sqrt{\cos \theta (\cos \frac{1}{2} \phi, \sin \frac{1}{2} \phi)}\), whose size is such that any rectangle in which it is inscribed is itself inscribed in the unit circle.

The mapping between the PED and physical disk can be appreciated mathematically by considering each as the unit disk in the complex plane, so any real vector \(z = (x_0, x_1)\) corresponds to the complex number \(Z(z) = x_0 + ix_1 = (1, i) \cdot z\). and, conversely, \(z(Z) = \Theta(Z), 3(Z)\). The complex numbers corresponding to the ellipse foci \(f_{\pm} = \pm \sqrt{\cos \theta (\cos \frac{1}{2} \phi, \sin \frac{1}{2} \phi)}\) are then the two square roots of the PED coordinate \(s = \cos \theta (\cos \phi, \sin \phi)\), as shown in Fig. 2:

\[ Z(f_{\pm}) = \pm \sqrt{Z(s)}. \]

This map is conformal (angle preserving) except at the origin, as shown in Fig. 2: a Cartesian grid over the PED maps onto a curvilinear orthogonal grid over the physical disk.

### 5. Families of Orbits, Caustics, and the Solid Angle Quantization Condition

The complete two-parameter ray family is constructed as a continuous one-parameter set of orbits. For the global ray structure to be preserved on propagation, all orbits must be coaxial, having a common waist plane, and have a common \(Q_0\) and \(P_0\) (and hence \(N\)), so that they all scale as \((Q_0^2 + x^2P_0^2)^{1/2}\). Such a set of orbits corresponds to a path on the Poincaré sphere, which we call a Poincaré path. The cases of interest here are those whose paths are closed loops. For simplicity, consider first a Poincaré path confined to a hemisphere, so that its projection onto the PED is a closed loop that does not touch the disk’s edge. There are two Poincaré paths (one in the upper hemisphere, one in the lower) projecting to each such PED loop, as shown in Fig. 3(a). Each point on the projected path corresponds to an ellipse in the physical disk, so the complete closed path gives rise to a family of ellipses, as shown in Fig. 3(b). Figures 3(c) and 3(d) show how the shape of their transverse ray structures is preserved under propagation (up to a hyperbolic scaling), and that this structure is the same when the loop is in the upper [Fig. 3(c)] or lower [Fig. 3(d)] hemisphere; the hemisphere only determines the handedness (the sign of the OAM). Visualization 2 shows how each point along the path corresponds to the rays in an orbit, while Visualization 3 illustrates the propagation of the rays for increasing \(x\).

Figure 3(b) represents the beam cross section as a superposition of elliptical ray orbits. This structure is determined by the path’s projection onto the PED, which we will refer to as the PED path. The envelopes of the family are caustics, here an outer one enclosing all the rays and an inner one inside of which there are no rays. The brightest intensity features of a beam are associated with these caustics, as the density of rays is highest near them. There are surprisingly simple geometrical relations between the projected Poincaré path and the caustics in the physical disk, which we now describe (the derivation can be found in Section S1 of Supplement 1).

The geometric relation is easiest to appreciate for a PED path with endpoints at the edge of the disk, such as the one shown in Fig. 4(a). The corresponding path on the full Poincaré sphere is symmetric in the upper and lower hemispheres (projecting to the same curve in the PED). The geometric prescription for finding the caustics is then as follows:

1. Find the set of circles that are tangent to both the PED path and the unit circle. There are two such sets, one on each side of the PED path [shown in pale blue and red in Fig. 4(a)]. The centers of each set of circles define a curve equidistant from the PED path and the unit circle. Each such curve, being equally close to two other curves, is a medial axis [53] (or topological skeleton), in terminology borrowed from image analysis.

2. Find the caustics by applying the square root map \(z(\pm \sqrt{1/t}(t))\) to each medial axis. Given that the square root maps each point in the PED onto two points on the physical disk, each medial axis is mapped onto two caustic segments that are identical except for a 180° rotation about the origin. Therefore, each point...
The PED path is parameterized as \( s(\eta) = (\cos \phi(\eta) \cos \theta(\eta), \sin \phi(\eta) \cos \theta(\eta)) \), so overall the ray family is parameterized by \( 0 \leq \tau, \eta < 2\pi \), topologically corresponding to a torus. Families of ellipses in the physical disk such as in Fig. 3(b) are projections of this torus, with its outline given by the caustics, consisting of either a quadrangle with corners at the boundary of the physical disk [as in Fig. 4(a)] or an outer and an inner loop [as in Fig. 4(b)].

As the Poincaré path is a closed loop, wave-optical self-consistency requires that any physical quantity (determined by the OPL) must return to its starting point on a circuit of \( \eta \). (The OPL at \( z = 0 \) for all the rays, in terms of \( \tau \) and \( \eta \), may be found in Section S2 of Supplement 1.) This gives a quantization condition around the path, just as our previous condition quantized the orbits. This condition, whose derivation is given in Section S3 of Supplement 1, is remarkably simple geometrically: the solid angle \( \Omega \) on the Poincaré sphere enclosed by the Poincaré path must be an odd multiple of \( 2\pi/(N + 1) \), namely,

\[
\Omega = \left( 2n + 1 \right) \frac{2\pi}{N + 1}, \quad n = 0, 1, \ldots, \lfloor N/2 \rfloor, \tag{8}
\]

where \( \lfloor N/2 \rfloor \) denotes the integer part of \( N/2 \).

We may appreciate the significance of this by referring to the quantum-mechanical picture. Structured light beams are usually considered as eigenfunctions of some optical operator, such as the OAM operator \( \hat{L} = -i\partial_\phi \) giving the LG modes [3] or the astigmatism operator \( \hat{M} = \frac{1}{2}(-\partial_\tau^2 + \partial_\eta^2 + x^2 - y^2) \) giving the HG modes [22]. In the completely classical, Hamiltonian picture, these quantities are functions of position \( \mathbf{Q} \) and momentum \( \mathbf{P} \), which define families of contours on the Poincaré sphere (the sphere of orbits of the isotropic two-dimensional oscillator). Thus the angular momentum \( L \) is simply the height coordinate of the Poincaré sphere \( \cos \theta \), and \( M \) is the horizontal coordinate \( \sin \theta \cos \phi \) [22]; the contours are then circles concentric to the vertical or horizontal axes of the sphere. The condition in Eq. (8) picks out a discrete set of these contours as the Poincaré paths, which correspond to the sets of ray families that are consistent with wave optics. We will discuss the LG and HG modes in detail, after having discussed how to construct approximations to the wave fields from the appropriately quantized ray families.

### 6. RAY-BASED WAVE FIELD RECONSTRUCTION

Many methods exist for estimating wave fields based purely on a ray description, which are valid even in the presence of caustics. Here we use an approach [45–48] in which a complex Gaussian field contribution is assigned to each ray, and the estimate takes the form of a double integral over \( \tau \) and \( \eta \). It is shown in Section S5 of Supplement 1 that the integral in \( \tau \) can be evaluated analytically, leading to a field estimate at the waist plane of the form

\[
U(\mathbf{x}) \approx \frac{k P_0}{\sqrt{2}} \exp \left( i \frac{\pi}{4} \right) \int A(\eta) \sqrt{\cos \theta \frac{\partial \phi}{\partial \eta} + i \frac{\partial \theta}{\partial \eta}} U_N \left( \frac{x}{Q_0}, y \right) \times \exp \left( i \left\{ kL_2 - (N + 1) \left[ T - \frac{\sin(2T) \cos \theta}{2} \right] \right\} \right) \mathrm{d}\eta, \tag{9}
\]

where \( \mathbf{x} = (x, y) \) is the transverse position at the waist plane, \( A(\eta) \) is a nonnegative amplitude function weighting the different orbits,
v is the Jones vector in Eq. (1) parameterized in terms of \( \eta, T(\eta) \) and \( L_2(\eta) \) are given in Section S2 of Supplement 1, and the \( U_N \) terms are HG elementary fields evaluated at complex values, defined as

\[
U_N(\mathbf{x}, \mathbf{v}) = \frac{1}{N!} \exp \left( -\frac{N+1}{2} \right) \left( \frac{N+1}{2} \mathbf{v} \cdot \mathbf{v} \right) \left( \frac{N+1}{2} \mathbf{v} \cdot \mathbf{v} \right)^{-1/2} \times \exp[-(N+1)|\mathbf{x}|^2]H_N\left( \sqrt{\frac{2(N+1)}{\mathbf{v} \cdot \mathbf{v}}} \cdot \mathbf{x} \cdot \mathbf{v} \right),
\]

(10)

where \( H_N \) is the \( N \)th-order Hermite polynomial and \( \mathbf{v} \cdot \mathbf{v} = \cos \theta \). Up to a complex factor, \( U_N \) is the wave contribution corresponding to an elliptical ray orbit specified by the Jones vector \( \mathbf{v} \). Figure 5 shows, for three choices of \( \mathbf{v} \), the real part and intensity of these orbit contributions, together with the corresponding ellipsoidal ray-optical orbit. \( N \) is the number of phase oscillations around the ellipse. In fact, these elementary field contributions are themselves a subset of the HLG beams, which are associated with points on a Poincaré sphere [22–25,27]. However, these contributions are expressed not as a superposition of HG or LG beams but as a single term involving a Hermite polynomial evaluated at a complex argument proportional to the Jones vector.

The expression in Eq. (9) provides a general prescription for constructing self-similar beams that are rigorous solutions to the paraxial wave equation, and that have caustics at prescribed locations. Note that the integral over the Poincaré path in Eq. (9) is assumed to be closed; the initial and final integration values are not specified, because the integrand is assumed to be periodic in \( \eta \). As shown in Sections S2 and S3 of Supplement 1, this periodicity is guaranteed by the condition in Eq. (8).

7. EXAMPLES: LG, HG, AND HLG BEAMS

We now illustrate these ideas for the two most common families of beams of this type, LG and HG beams, as well as for the more general HLG beams. Given their rotational symmetry, LG beams are separable in polar coordinates. Their PED path is then a circle of radius \( r \) centered at the origin (so \( \theta = \arccos r \)). The solid angle enclosed by this circle over the Poincaré sphere is

\[
\Omega = 2\pi(1-\sin \theta) = 2\pi[(1-(1-r^2)^{1/2})^2],
\]

which is quantized according to Eq. (8), such that \( r \) can only take the values

\[
r = \sqrt{\frac{2N+1+4n(N-n)}{N+1}}
\]

(11)

for \( n = 0, 1, \ldots, \lfloor N/2 \rfloor \). The medial axes, equidistant from the unit circle and the PED path, must also be circles centered at the origin, but with radii \( (1 \pm r)/2 \). Following the square root map onto the physical disk, the two caustics are circular as well, with radii \( Q_0 \sqrt{(1 \pm r)/2} \). More details about the ray description of these beams are given in Section S6 of Supplement 1, where it is also shown that, remarkably, the wave field estimate in Eq. (9) actually yields the exact form for LG beams with vorticity \( \ell = N - 2n \). Figure 6 shows the PED and physical disks for these beams, including the PED path, medial axes, caustics, and elliptical orbits, as well as the ray structure of the beam and the intensity cross section. The first part of Visualization 5 illustrates the ray description for varying radii of the PED path. In terms of the operator picture, all physical quantities (PED path, medial axes, caustics) must be rotation invariant, and the path quantization gives the usual angular momentum quanta \(-N \leq \ell \leq N\), quantized in integers (in steps of 2).

We now consider HG beams, which are separable in Cartesian coordinates. The PED path is a straight line terminating at the edge of the PED, as shown in Fig. 7(a). Since the PED is a projection of the sphere onto its equatorial plane, the Poincaré path on the sphere is a circle centered at the \( z \) axis with radius \( r \), equal to half the length of the straight line, and is also quantized according to Eq. (11) [22]. By simple geometry, both medial axes are
confocal parabolas with foci at the origin, which intersect each other and the PED path at the edge of the disk. These parabolic medial axes are shown in Fig. 7(a) as blue and orange curves. The caustics (square roots of the parabolas) are straight lines, as shown in Fig. 7(b): the first medial axis maps onto two vertical caustic lines (orange) at

\[ x = \pm Q_0 \sqrt{\frac{1 + \sqrt{1 - r^2}}{2}} = \pm Q_0 \sqrt{\frac{2m + 1}{2(N + 1)}} \quad (12) \]

where \( m = N - n \), while the second maps onto two horizontal caustic lines (blue) at

\[ y = \pm Q_0 \sqrt{\frac{1 - \sqrt{1 - r^2}}{2}} = \pm Q_0 \sqrt{\frac{2n + 1}{2(N + 1)}} \quad (13) \]

Thus the caustics form a rectangle enclosing the rays. Further details on the ray parameterization, and a proof that the wave field construction in Eq. (9) also gives the exact form for the HG beams, are in Section S7 of Supplement 1. The ray structure and intensity distributions are shown in Figs. 7(c) and 7(d). The last part of Visualization 5 illustrates the ray description for different positions of the PED path. HG beams are eigenfunctions of the astigmatism operator \( \hat{\mathcal{M}} \) [22], whose eigenvalues \( m - n \) are algebraically identical to those of the angular momentum operator.

Finally, we consider HLG beams, which are realized by conversion of HG or LG beams through simple combinations of cylindrical lenses [30,32] or equivalent holographic implementations, which amount to rotations of the Poincaré sphere about an axis in the equatorial plane, but which cannot be expressed simply in any separable coordinate system. As for LG and HG beams, for HLG beams the Poincaré path on the sphere is a (planar) circle whose radius \( r \) is quantized according to Eq. (11). However, the center of this circle can be at any angle \( \beta \) with respect to the vertical \( z_3 \) axis. We assume for simplicity that the center lies in the \( s_1s_3 \) plane, so that \( \beta = 0 \) gives LG beams, while \( \beta = \pi/2 \) reduces to HG beams separable in \( x \) and \( y \). Projected onto the equatorial disk, the PED path is an ellipse centered at \((1 - r^2)^{1/2} \sin \beta, 0\) with minor and major semiaxes given by \( r \cos \beta \) and \( r \), respectively, as shown in the left column of Fig. 8. The medial axes and hence the caustics (shown in the figure’s second column) can be found in parametric form, and are not conic sections. Similarly, the wave fields are no longer separable in a coordinate system, but they can still be computed from Eq. (9). Figure 8 illustrates these beams for three values of \( \beta \) intermediate between the LG and HG limits. Visualization 5 shows the complete transition from LG to HG.

In the operator picture, these beams are eigenfunctions of \( \hat{\mathcal{L}} \cos \beta + \hat{\mathcal{M}} \sin \beta \), which have the same integer eigenvalues (corresponding, in the ray picture, to rotating the spherical cap in the \( s_1s_3 \) plane). Since the operator is linear in the coordinates of the Poincaré sphere’s space, the Poincaré path is a circle with uniform weight. This simplicity of the HLG family explains why the ray-based field estimate in Eq. (9) actually yields the known exact eigenstates of the operators.

8. GOUY AND PANCHARATNAM–BERRY PHASES

In addition to revealing the hidden geometry behind the caustic structure of structured Gaussian beams, the description presented here provides a simple ray-based explanation for their Gouy and
Pancharatnam–Berry phase shifts. These two phase shifts correspond to shifts in each of the two ray parameters, $\tau$ and $\eta$, as follows.

Consider first the case of the Gouy phase shift. As shown in Eq. (4), propagation in $z$ preserves the ray structure up to a shift $\tau \rightarrow \tau + \zeta$, where $\zeta = \arctan(z P_0 / Q_0)$. Thus, any ray initially at a given location when $z = 0$ is replaced, after propagation, by another one from the same orbit, whose value of $\tau$ is larger by an amount $\zeta$. Since a variation in $\tau$ of $2\pi$ corresponds to a path length of $(N + 1) \lambda$, this shift in $\tau$ by $\zeta$ amounts to a change in the path length of $(\zeta / 2 \pi)(N + 1) \lambda$, and hence to a phase of $(N + 1) \arctan(z P_0 / Q_0)$, namely, the standard Gouy phase for a beam of this type. This effect can be appreciated from Fig. 1: all rays have roughly the same length. However, the ray that touches a given point in the orbit (say, a vertex of the ellipse) at the initial plane is not the same as the one that touches the same point at the final plane. The total phase difference is then due not only to the length of the rays but also to the OPL difference between the two rays in question.

The geometric phase for beams of the HLG family under astigmatic transformations has been studied in algebraic terms by exploiting the analogy with two-dimensional quantum harmonic oscillators [25–28] and verified experimentally for low-order beams [37–39]. Consider subjecting a HLG, LG, or more general HLG beam to a series of optical transformations that rotate the Poincaré sphere around an axis within the $s_1 s_2$ plane (through a suitable combination of cylindrical lenses) or around the $s_3$ axis (through a beam rotator such as a pair of Dove prisms or periscopes). By choosing the sequence of transformations appropriately, the circular Poincaré path for the beam can be brought back to its initial position after its center has traced a trajectory over the Poincaré sphere. However, it is easy to see that each point within the Poincaré path does not necessarily fall back onto its initial position; rather, the final state of the circle is generally rotated around its axis with respect to the initial one by some angle $\Theta$, depending on the trajectory followed. If this trajectory is composed only of segments of great circles, like that shown in Fig. 9 and Visualization 6, then the angle $\Theta$ equals the solid angle subtended by the path. In other words, this transformation reduces to a shift $\eta \rightarrow \eta - \Theta$. Such a rotation results in a phase shift for the beam that can be considered as a geometric phase, because it is not related to a change in the OPL of each ray, but to a cycling of the roles that different rays (and indeed orbits) play within the pattern. As stated in Section S3 of Supplement 1, the phase resulting from a complete rotation of the Poincaré loop is $k \Delta L_2 = (N - 2n) \pi$, so the corresponding geometric phase is $(N - 2n) \Theta / 2 = \ell \Theta / 2$, where $\ell$ is the OAM label of the LG beam within the set.

In summary, the phase accumulated under propagation can be separated into a "dynamic" phase, due to the path length traced by each ray, and a Gouy phase, due to the cycling of rays within each orbit. If the beam is also subjected to a series of transformations that rotate the Poincaré sphere but that bring the beam back to its original shape, there is a third, geometric component of the phase, due to the shifting of orbits within the beam structure. However, note that while the dynamic and Gouy phases apply to any beam, the general geometric phase can only be achieved for HLG beams, given the rotational symmetry of their Poincaré path. For beams whose Poincaré paths have $M$-fold symmetry around an axis, a more restricted version of the same phenomenon is possible.

As a final note, the subluminal propagation velocity observed for this type of beam [54,55] is easily explained by the fact that the rays making up the beam are tilted with respect to the $z$ axis, so their average length between two planes of constant $z$ is slightly longer than the distance between the planes.

9. OTHER SEPARABLE SELF-SIMILAR BEAMS AS LIMITING CASES, AND SELF-HEALING

Despite their not being explicitly Gaussian, other types of propagation-invariant beams, such as Bessel [6,7], Mathieu [8], Airy [4,5], and parabolic [56] beams, correspond to the limits of the structured Gaussian beams described here. These other beams are idealized solutions that involve infinite power, corresponding to the limit $N \rightarrow \infty$ in particular regions of the physical disk. That is, the ray families are open rather than closed loops.

Bessel and Mathieu beams correspond to a small neighborhood of the origin of the physical disk and the outer radius of the PED. For Bessel beams, the PED path is a circle centered at the origin and whose radius is nearly equal (or equal) to unity, so that one medial axis is a small circle (or point) centered at the origin, and so is the inner caustic. Mathieu beams use the same construction, except that the large circular PED path is shifted slightly from the origin but still fits within the PED. This shift decenters the small inner medial axis, and causes the resulting inner caustic to be elliptic (A shift larger than the difference between unity and the path's radius would make the inner caustic hyperbolic.) If instead the centered PED path is slightly deformed into an ellipse, the inner caustic becomes an asteroid, as in beams produced by misaligned axicons [57].

In the case of Airy and parabolic beams, on the other hand, one must focus on a small region at the edge of the PED and physical disk. Airy beams are the limit of the intersection of two medial axes (and caustics) when a locally straight PED path touches the edge of the PED. This geometry is shown in
shifts in \( \eta \) not only through shifts in \( LG \), and more general \( HLG \) beams can undergo local for structured Gaussian beams, whose ray family is compact. Where the main intensity features are located. This is not the case ultimately lost in an infinite reservoir of rays away from the region Bessel beams, the idealized ray family is open, so the shadow is a different part of its transverse profile. For beams such as Airy or quantization. The description given here is based on the partition of the two-parameter ray family, one parameter giving rays around \( \tau \) increases.

Fig. 10(a). The angle of this intersection determines the ratio of the spacing of the intensity lobes along the two caustic sheets. Parabolic beams result when the PED path is a very small circular segment starting and ending at the edge of the PED, leading to two sets of parabolic caustics, as shown in Fig. 10(b).

Propagating self-similar beams are often referred to as “self healing”; if an obstacle blocks a limited part of the beam in one plane, the blocked intensity features reappear as \( z \) increases. The effect of the block can be described to the first order in terms of the ray-optical shadow projected by the obstacle, that is, the suppression of a subset of the rays composing the field. Self-healing (which can occur more generally [58,59]) is then easily explained in terms of the cycling of rays within each ray orbit under propagation: the blocked rays are replaced by other rays leading to the same local ray structure. However, it is clear that rather than “healing,” the beam’s “wound” is simply transferred to a different part of its transverse profile. For beams such as Airy or Bessel beams, the idealized ray family is open, so the shadow is ultimately lost in an infinite reservoir of rays away from the region where the main intensity features are located. This is not the case for structured Gaussian beams, whose ray family is compact.

Due to the rotational symmetry of their Poincaré path, HG, LG, and more general HLG beams can undergo local “healing” not only through shifts in \( \tau \) under propagation but also through shifts in \( \eta \) due to rotations of the Poincaré sphere caused by the optical transformations discussed in Section 8. Such shifts would have a similar effect of displacing the blocked regions within the beam’s profile. To further abuse the already imperfect “healing” metaphor, this effect could be called “assisted healing.”

10. CONCLUDING REMARKS

We have proposed a ray-based description of structured Gaussian beams that reveals hidden geometrical restrictions in their spatial structure, particularly their caustics. For HG and LG beams, these caustics correspond to the characteristic rectangular and annular shapes of the respective beam intensities. Further, the Gouy and geometric phases that can be accumulated under propagation were also given simple explanations in terms of rays and their quantization. The description given here is based on the partition of the two-parameter ray family, one parameter giving rays around orbits with an elliptical cross section, and the other defining a curve on the Poincaré sphere representing the elliptic ray family. This develops previous work also employing Poincaré spheres to characterize the modal structure of HLG beams [23–28,37,38]. However, unlike these previous studies, where each beam is associated with a point on the Poincaré sphere, in our more general construction the beam is associated with a curve on the Poincaré sphere. The shape of this extended curve not only determines the shape of the beam but also explains (and restricts) the geometric phase resulting from beam transformations.

The approach given here also differs from other ray-based studies of structured Gaussian beams. For example, Gaussian beams have been described as bundles of complex rays [60–62], as opposed to the real rays used here. Similarly, ray-like descriptions of LG and Bessel beams have been given in terms of Wigner functions [63], but such a description uses all rays in the phase space rather than a two-parameter family, so the concept of the caustic is absent and the representation in the Poincaré sphere is not compatible with that treatment. Finally, descriptions also exist in terms of curved flux lines in addition to the rays [64]. Note that the use of a two-parameter family of rays in terms of their transverse position and direction vectors makes it easy to model the propagation of these beams not only through homogeneous media but also through optical systems described by first-order (or ABCD) matrices [65].

In the complementary operator picture of our approach [22], there is a spin-vector-like operator on the Poincaré sphere for which these HLG beams are described by circles whose centers are given by the vector direction of the operator. The operator approach, built around the \( su(2) \) Poisson algebra of the two-dimensional harmonic oscillator, reveals the algebraic connection between structured Gaussian beams, the classical and quantum harmonic oscillator, and the Poincaré sphere, contrasting with the semiclassical approach used here. We note that while we have only considered scalar fields, polarization can be incorporated. If self-similarity is to also require preservation of the polarization distribution, all rays in each orbit must carry the same polarization, but different orbits can have different polarizations. In this case, each beam is represented by two paths over two Poincaré spheres, one determining its modal structure and one its polarization distribution. The solid angles enclosed by both these paths would enter into a generalization of Eq. (8) determining the closure condition for wave-optical self-consistency. Such a generalization, however, is beyond the scope of this work.

Although our focus here has been the particular examples of HG, LG, and HLG beams, the ray-based approach can be applied to any beam with a Gaussian envelope with a well-defined Gouy phase—in fact, the approach allows such structured Gaussian beams to be designed from almost arbitrary paths on the Poincaré sphere satisfying the quantization condition [Eq. (8)]. One obvious structured Gaussian family we have not explored here is the Ince–Gaussian beams [66,67], which also interpolate between HG and LG beams, but which are separable in elliptic coordinates. From the other separable beams considered here, one might

![Fig. 11.](image-url)
expect the caustics of Ince–Gaussian ray families to be confocal ellipses and hyperbolas. Indeed, this is the case, and we defer a full discussion to a later paper. Although no other separable Gaussian beam families exist [68], the freedom of choice of curves on the Poincaré sphere allows a huge variety of Gaussian beams with new and unfamiliar properties to be designed, such as the octagonal Gaussian beam shown in Fig. 11.


Acknowledgment. We thank Michael Berry, John Hannay, Michael Morgan, and Greg Forbes for useful comments.

See Supplement 1 for supporting content.

REFERENCES