Optical scalars and congruences of light rays: A link between beams and analytic aberrations

Omar de J. Cabrera-Rosas,* Paula Ortega-Vidals, Salvador Alejandro Juárez-Reyes, Citlalli Teresa Sosa-Sánchez, Ernesto Espíndola-Ramos, Israel Julián-Macías, and Gilberto Silva-Ortigoza

Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla, 72570 Puebla, Puebla, México

(Received 5 June 2017; published 24 August 2017)

Optical scalars are functions designed to analyze the behavior of geodesic congruences in general relativity. Refracted rays are three-dimensional congruences of light rays and they can be studied with this formalism. In this work we obtain the optical scalars for such congruences: the expansion Θ, the twist ω, and the shear κ. Furthermore, we apply this machinery to study the aberrations of wavefronts to establish a link between them and the aberration function W.

DOI: 10.1103/PhysRevA.96.023852

I. INTRODUCTION

Given an open region in a three-dimensional space, a congruence is a set of curves where one and only one curve passes through each point in the region [1]. Congruences of curves are very useful in general relativity because finding them enables study of the metric information on the space-time [2]. From the point of view of geometrical optics, light rays are congruences, and in fact this kind of congruence is simpler than the gravitational kinds because they are three-dimensional and the information on the wavefronts is used to study them.

Sachs developed a method to analyze the evolution of congruences (in the study of gravitational radiation) by defining the *optical scalars* [3]. These functions measure the expansion (Θ), rotation (ω), and distortion (κ) of a test circle that evolves along the congruence and they are used in gravitational lensing and cosmology and to study gravitational waves, to name a few uses. Although those functions were defined in tensor language, the problem is simpler using the spinor formalism [4].

On the other hand, the aberration theory has been a very important tool in studying the quality of optical systems [5]. Transverse aberrations provide information that is measured and used to understand the evolution of light rays and wavefronts. Rayces found the set of equations that give the relationship between the transverse aberrations and the wave aberration function [6,7]. It is important to recall that wavefronts and light rays are not observables; however, aberration theory is a powerful tool to determine the quality of the optical system under study.

Therefore, we have two approaches to study the quality of an optical system: one is given by the aberration theory, and the other one is obtained by using the optical scalars. Thus, the aim of the present work is to show that there exists an analytic relationship between the optical scalars and the aberration function $W(x,y)$. To this end, in Sec. II, we present the basic equations to compute the critical and caustic sets of the map describing the evolution of the refracted light rays. In Sec. III, using spinors we define the optical scalars. In Sec. IV, we obtain the optical scalars by using the field that describes the evolution of the refracted light rays. In Sec. V, we present the relationship between the optical scalars and the wave aberration function. Finally, we present the conclusions.

*omcbrss@gmail.com

II. GENERAL EQUATIONS

The optical system under study is shown in Fig. 1. We assume that the free space is filled out with two optical media with refraction indices n_1 and n_2, respectively. In the optical medium with refraction index n_1 we place a point light source at $\mathbf{s} = (s_1,s_2,s_3)$, which emits spherical wavefronts that are refracted at the interface $\mathbf{r} = (x,y,f(x,y))$, $\gamma = n_1/n_2$, $\hat{\mathbf{R}}$ gives the direction of the refracted light ray, $\hat{\mathbf{I}}$ gives the direction of the incident light ray, $\hat{\mathbf{N}}$ is the normal unit vector to the interface, and Ω is a function determined from the condition that $\hat{\mathbf{R}}$ be a unitary vector field [8]. That is,

$$\begin{aligned}\hat{\mathbf{N}} &= \frac{(-f_x,-f_y,1)}{\sqrt{1+f_x^2+f_y^2}},\\ \hat{\mathbf{I}} &= \frac{\mathbf{r}-\mathbf{s}}{|\mathbf{r}-\mathbf{s}|},\\ \Omega &= -\gamma(\hat{\mathbf{I}}\cdot\hat{\mathbf{N}}) + \sqrt{1-\gamma^2(\hat{\mathbf{I}}\times\hat{\mathbf{N}})^2},\\ \hat{\mathbf{R}} &= \gamma\hat{\mathbf{I}} + \Omega\hat{\mathbf{N}},\end{aligned} \tag{1}$$

where f_x and f_y represent the partial derivatives of $f(x,y)$ with respect to x and y, respectively.

The vector field that describes the ray tracing process is the $\mathbf{X}$ field given by

$$\mathbf{X} = \mathbf{r} + \ell\hat{\mathbf{R}}, \tag{2}$$

where $\ell = \tau - \gamma|\mathbf{r}-\mathbf{s}|$ and τ labels each of the wavefronts in the wavefront train. Equation (2) is a map between two subsets of $\mathcal{R}^3$, i.e., $(x,y,\ell) \mapsto (X,Y,Z)$, where (x,y,ℓ) are the local coordinates of the domain space, while (X,Y,Z) are local coordinates of the target space. Hence we see that an important quantity in the calculations is the Jacobian of this map, which is given by

$$\mathcal{N} = H_0 + \ell H_1 + \ell^2 H_2, \tag{3}$$

where

$$\begin{aligned}H_0(x,y) &= (\mathbf{r}_x\times\mathbf{r}_y)\cdot\hat{\mathbf{R}},\\ H_1(x,y) &= (\mathbf{r}_x\times\hat{\mathbf{R}}_y + \hat{\mathbf{R}}_x\times\mathbf{r}_y)\cdot\hat{\mathbf{R}},\\ H_2(x,y) &= (\hat{\mathbf{R}}_x\times\hat{\mathbf{R}}_y)\cdot\hat{\mathbf{R}}.\end{aligned} \tag{4}$$

It is useful to establish the metric coefficients and curvatures for an arbitrary surface [9]. For a surface described by the coordinate patch $\mathbf{r}$ with normal unit vector $\hat{\mathbf{N}}$, the coefficients

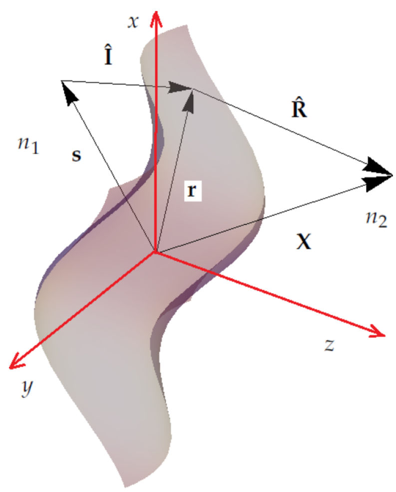


FIG. 1. Schematic of the optical system: a point source placed at $\mathbf{s}$ which emits rays that interact with an arbitrary surface $\mathbf{r}$. The incident ray $\hat{\mathbf{I}}$, the refracted ray $\hat{\mathbf{R}}$, and the vector field $\mathbf{X}$ are shown.

of the first and the second fundamental forms are given, respectively, by [10]

$$E = \mathbf{r}_x \cdot \mathbf{r}_x, \quad l = \hat{\mathbf{N}} \cdot \mathbf{r}_{xx},$$
$$F = \mathbf{r}_x \cdot \mathbf{r}_y, \quad m = \hat{\mathbf{N}} \cdot \mathbf{r}_{xy}, \tag{5}$$
$$G = \mathbf{r}_y \cdot \mathbf{r}_y, \quad n = \hat{\mathbf{N}} \cdot \mathbf{r}_{yy},$$

and the Gauss and mean curvatures are

$$K(x,y) = \frac{ln - m^2}{EG - F^2},$$
$$H(x,y) = \frac{Gl + En - 2Fm}{2(EG - F^2)}. \tag{6}$$

Observe that the (x,y,ℓ) coordinates define a moving frame that travels along the refracted wavefronts in the direction of the refracted ray $\hat{\mathbf{R}}$. In fact the metric and the inverse metric in this frame are given by

$$(g_{ab}) = \begin{pmatrix} \mathcal{E} & \mathcal{F} & 0 \\ \mathcal{F} & \mathcal{G} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and}$$
$$(g^{ab}) = \frac{1}{\mathcal{N}^2} \begin{pmatrix} \mathcal{G} & -\mathcal{F} & 0 \\ -\mathcal{F} & \mathcal{E} & 0 \\ 0 & 0 & \mathcal{N}^2 \end{pmatrix}, \tag{7}$$

where $a,b = 1,2,3$ and $\mathcal{E}$, $\mathcal{F}$, and $\mathcal{G}$ are the corresponding metric coefficients for the $\mathbf{X}$ vector field, where the unit normal vector for the surfaces described by this field (the refracted wavefronts) is the refracted ray $\hat{\mathbf{R}}$ [13].

The caustic associated with refracted light rays

The region in the optical medium with refraction index n_2, where the refracted wavefronts will be singular (the refracted light rays will focus), is the caustic associated with the refracted light rays, and by definition it is the set of points in the domain space where the map, (2), is not locally one to one [11,12]. In this case it is equivalent to the condition

$$\mathcal{N} = 0. \tag{8}$$

There are two cases for solving Eq. (8): in the case where $H_2(x,y) \neq 0$ the corresponding solution for the critical set is

$$\ell_\pm = \frac{-H_1 \pm \sqrt{H_1^2 - 4H_2H_0}}{2H_2}, \tag{9}$$

and the critical set has two branches. Therefore, the caustic set associated with the map is given by

$$\mathbf{X} = \mathbf{X}_{C\pm} = \mathbf{r} + \ell_\pm \hat{\mathbf{R}} \tag{10}$$

and also has two branches. The second case is where $H_2(x,y) = 0$; the critical set is determined by

$$\ell = -\frac{H_0}{H_1}, \tag{11}$$

the caustic set by

$$\mathbf{X}_C = \mathbf{r} - \frac{H_0}{H_1}\hat{\mathbf{R}}, \tag{12}$$

and it has only one branch.

To finish this part, we remember that the caustic is a very important object because it is a physical observable of the system, but as we see in the following sections, the discriminant of Eq. (9) not only determines the branches of the critical and caustic sets [in the case $H_2(x,y) \neq 0$], but also is related to the shear κ of the congruence of light rays.

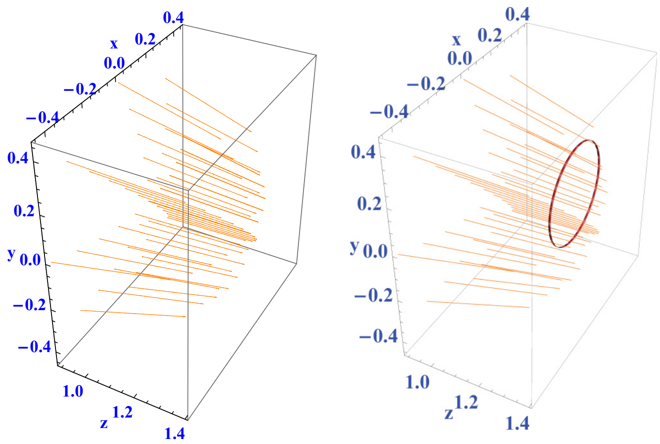
III. CONGRUENCES OF CURVES: OPTICAL SCALARS

From the geometric optics point of view, a set of light rays carries the information on the optical system. Through the process of analytic ray tracing the optical evolution of images can be described and studied [14,15]. But from the mathematical point of view, it is necessary to properly define the representation of these sets of rays. We need then to define a congruence of curves for the case under study: given an open region $\mathcal{S}$ in a three-dimensional space, a congruence in $\mathcal{S}$ is a family of curves such that through each point there passes one and only one curve from this family [1]. From the above definition, we can directly identify a set of rays as a congruence of curves that in fact is *incongruent* on the caustic [see Fig. 2(a)]. That is, the results derived here are valid out of the caustic region.

To make a proper analysis of congruences we need a formalism that directly establishes the quantities to describe the evolution of such congruences: this is the spinor formalism. Spinors are a factorization of real vector fields or, more generally, tensor fields. The advantage of this procedure is that the functions that describe the optical process can be obtained directly by analyzing congruences of geodesics, and these functions are called the *optical scalars*. In the following section we give a brief introduction to the spinor formalism in three dimensions, using only the basic definitions to have at hand the necessary tools for the physical problem. A complete guide for this procedure is given in [4].

A. Spinors

All the quantities and vector fields described in the previous section can be translated to the spinor formalism. A spinor is



(a) (b)

FIG. 2. (a) Plot of a congruence of light rays (for a system with symmetry of revolution); hereinafter plots are labeled in centimeters. (b) The same congruence, but with a test circle in a certain region of such congruence. Observe that the test circle is orthogonal to the direction of evolution of the congruence and analyzes the local information on the light rays.

a two-component complex vector ψ which transforms under rotations according to the rule

$$\psi' = Q\psi, \tag{13}$$

where Q is a unitary matrix belonging to the SU(2) group [4]. In general we can transform the components of any tensor $t^{abc\dots}$ and they are in correspondence with the components of a spinor $t^{ABCD\dots}$, where lowercase letters denote tensor indices $(a,b,c,\dots = 1,2,3)$, and capital letters spinor ones $(A,B,C,D,\dots = 1,2)$. The tensor indices are raised (or lowered) by means of the g^{ab} metric (or, correspondingly, g_{ab}), and the spinor indices with the Levi-Civita symbol ε^{AB} (or, correspondingly, with ε_{AB}). It is important to recall that any orientable three-dimensional manifold admits a spinor structure, which, however, may not be unique [4]; therefore spinors can be employed to represent points of space and, in this case, as an alternative formalism to tensor analysis in three dimensions, but they have several applications in other fields of physics. In the case of three-dimensional spaces, a single one-index spinor determines a basis. When the metric is positive definite, this relationship allows the representation of a spinor by means of an ax or a flag. Then, given a one-index spinor ψ^A different from 0, one can define the vectors $\mathbf{V}$ and $\mathbf{M}$ with components

$$V_a \equiv -\sigma_{aAB}\hat{\psi}^A\psi^B, \quad M_a \equiv \sigma_{aAB}\psi^A\psi^B, \tag{14}$$

where there is summation convention on repeated indices, σ_{aAB} are the Pauli matrices, and $\hat{\psi}^A = -\overline{\psi_A}$ (the bar denotes complex conjugation) is the mate of the spinor ψ^A [4].

In the case of a real vector field $\mathbf{T}$ we associate with it a spinor T^{AB} that, in turn, can be decomposed into a one-index spinor and its mate. This process relates the components of $\mathbf{T}$ with the components of $\mathbf{V}$ and $\mathbf{M}$ and the latter vectors have useful properties: the components of $\mathbf{V}$ are real, and its direction coincides with the direction of $\mathbf{T}$; meanwhile, the components of $\mathbf{M}$ are complex and they inhabit a plane that is always orthogonal to the direction of $\mathbf{V}$, that is, $V_aM^a = M_aM^a = 0$ and $V_aV^a = (\mathrm{Re}M_a)(\mathrm{Re}M^a) = (\mathrm{Im}M_a)(\mathrm{Im}M^a) = (\psi^A\hat{\psi}_A)^2$, and therefore if ψ^A is a normalized spinor ($\psi^A\hat{\psi}_A = 1$), then $\{\mathrm{Re}\mathbf{M},\mathrm{Im}\mathbf{M},\mathbf{V}\}$ is an orthonormal basis. On the other hand, these vectors define the operators

$$D = \frac{1}{\sqrt{2}}V^a\partial_a, \quad \delta = \frac{1}{\sqrt{2}}M^a\partial_a, \quad \overline{\delta} = \frac{1}{\sqrt{2}}\overline{M}^a\partial_a, \tag{15}$$

where $\partial_a = \partial/\partial x^a$. These operators are important because they allow us to determine the spin coefficients that describe the behavior of the congruence.

B. Spin coefficients

In a three-dimensional Riemannian manifold a unique connection (the Levi-Civita connection) can be defined. Denoting by ∇_a the covariant derivative with respect to ∂_a, the components of this connection are the real-valued functions $\Gamma^a{}_{bc}$, given by

$$\nabla_a\partial_b = \Gamma^c{}_{ba}\partial_c.$$

The spinor analog of these functions is the *spin coefficients* and they are defined by

$$\nabla_{AB}\partial_{CD} = \Gamma^R{}_{CAB}\partial_{RD} + \Gamma^R{}_{DAB}\partial_{CR}, \tag{16}$$

where ∇_{AB} denotes the covariant derivative with respect to ∂_{AB}. In fact, in a space where the metric is definite positive, the components Γ_{ABCD} are given by the complex functions

$$\kappa \equiv \Gamma_{1111}, \quad \beta \equiv \Gamma_{1211}, \quad \rho \equiv \Gamma_{2211},$$
$$\alpha \equiv \Gamma, \quad \varepsilon \equiv \Gamma_{1212}, \tag{17}$$

where ε is a purely imaginary function. The Γ functions are symmetric in the first and second pairs of indices and their complex conjugates are obtained by changing the index 1 to 2, and vice versa [4].

Now it is necessary to define the condition for a vector field to be tangent to a congruence of geodesics. Given a real vector field $t^a\partial_a$ in a space with a positive definite metric, if this field is tangent to a geodesic, we can always find locally a spinor field such that $t_{AB} = \psi_{(A}\hat{\psi}_{B)}$ and $\psi^A\hat{\psi}^B\nabla_{AB}\psi_C = 0$ (i.e., is parallelly transported along the geodesic). The previous statement is equivalent, in terms of the spin coefficients, to the condition that $\alpha = 0$, and we can always make $\varepsilon = 0$ in such a way that the triad $\{D,\delta,\overline{\delta}\}$ is parallelly transported along the geodesic. Therefore, a direct way to calculate the spin coefficients is through the commutation relations of the operators D, δ, and $\overline{\delta}$. These relations are

$$[D,\delta] = 2\alpha D + (2\varepsilon - \rho)\delta - \kappa\overline{\delta},$$
$$[\delta,\overline{\delta}] = 2(\overline{\rho} - \rho)D - 2\overline{\beta}\delta + 2\beta\overline{\delta}, \tag{18}$$

so it will be necessary now to identify the corresponding triad $\{D,\delta,\overline{\delta}\}$ for the vector field that describes the refracted congruence.

C. Optical scalars

Optical scalars were first introduced by Sachs in four dimensions. They are used to analyze the behavior of a test circle placed perpendicular to a congruence of geodesics,

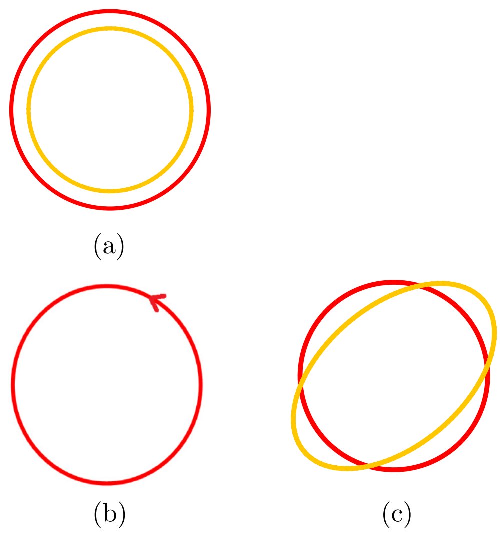


FIG. 3. (a) Test circle (red) evolving according to the congruence. The yellow (inner) circle represents the expansion Θ (in this case a contraction) of the congruence. (b) Another circle representing the twist ω of the test circle. (c) The yellow ellipse shows the deformation of the red circle. This is the geometrical representation of the shear κ of the congruence.

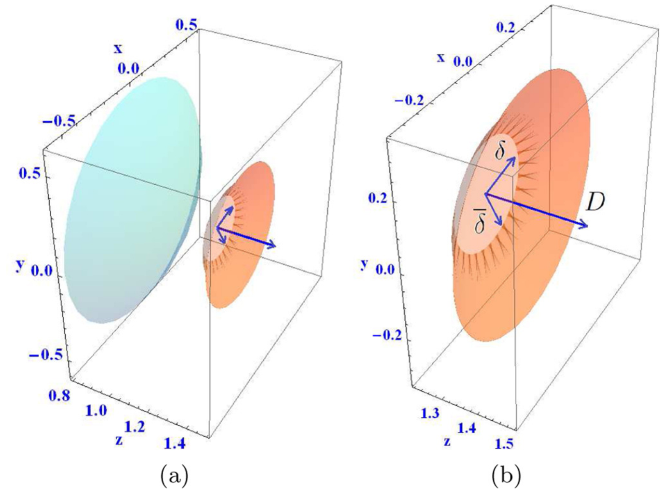


FIG. 4. (a) Schematic of a planospherical thin lens (blue) and a refracted wavefront (orange) for $\tau = 1.8$ cm. Blue arrows represent the directions of the operators δ, $\bar{\delta}$, and D. (b) Closeup of the wavefront, where the arrows denote the directions of the operators δ, $\bar{\delta}$, and D for the point $x = y = 0$ cm. Observe that δ and $\bar{\delta}$ inhabit the orthogonal plane to D. This plane is defined by the **M** vector; meanwhile the direction of the normal for that plane is the **V** vector (or, equivalently, the $\hat{\mathbf{R}}$ unit vector).

where the variations of this circle are measured [3] [Figs. 2(b) and 3]. In the three-dimensional case they are given by [4]

$$\Theta \equiv \mathrm{Re}\rho, \quad \omega \equiv \mathrm{Im}\rho, \quad \text{and} \quad \kappa, \tag{19}$$

where ρ and κ are the spin coefficients given in Eqs. (17). Note that in three dimensions, the test circle with respect to which the optical scalars are measured is placed in the plane defined by the **M** vector, and the physical meaning of each one of these functions is important: Θ measures whether this circle expands or contracts as the congruence evolves, while ω measures the rotation of the circle, and, finally, κ represents an area-preserving shear. Here $|\kappa|$ measures the magnitude of the shear, while $\arg\kappa/2$ is an angle that measures a rotation of the test circle (sheared into an ellipse) with respect to the principal axes ∂_1 and ∂_2. Therefore Θ, ω, and κ are called the *expansion*, the *twist*, and the *shear* of the congruence, respectively.

The task now is to identify the vector field that is tangent along the geodesics of the problem. A useful observation for this task is that the normal for the surfaces described by the **X** field is the refracted ray $\hat{\mathbf{R}}$. This condition will play an important role in obtaining the optical scalars that describe the congruence; we address this condition in the following section.

IV. OPTICAL SCALARS FOR THE REFRACTED CONGRUENCE OF LIGHT RAYS

Now that we have the tools to analyze the evolution of a congruence of light rays, we apply the formalism for the optical system that is our case of study. However, it is important to note that this procedure is in fact completely general, because the results obtained for this particular congruence are analogous for *any* optical congruence in three dimensions, that is, so far there is no distinction whether the congruence is incident, reflected, or refracted: by identifying the direction of evolution of the congruence and the conditions that must be fulfilled by the congruence, we can determine the corresponding optical scalars of the problem.

It has been shown that in the case of a congruence of refracted light rays (and therefore for the refracted wavefronts), the unit normal for these wavefronts is the refracted ray [13], and this fact will be very useful in our calculations. As we said before, once a vector is defined, we can find its corresponding **V** and **M** vectors (which are related to a spinor and, at the same time, the spin coefficients) and proceed to calculate the optical scalars defined in the previous section.

From the physical information on the optical system, we see that the vector field tangent to the refracted congruence of light rays is $\hat{\mathbf{R}}$, so we need to find its corresponding **V** and **M** vectors. This means that the direction of tangency is ∂_ℓ (this defines the direction of **V**) and ∂_x and ∂_y must be related to the components of **M** (Fig. 4). Observe that the triad $\{D,\delta,\bar{\delta}\}$ is parallelly transported along the refracted light ray (together with the refracted wavefronts), and from the metric information of Eqs. (7) it can be shown that the corresponding operators D, δ, and $\bar{\delta}$ for this case are

$$\begin{aligned} D &= \frac{1}{\sqrt{2}}\partial_\ell, \\ \delta &= \frac{1}{\sqrt{2}\mathcal{N}}[e^{i\mu}\sqrt{\mathcal{G}}\partial_x - e^{i\nu}\sqrt{\mathcal{E}}\partial_y], \\ \bar{\delta} &= \frac{1}{\sqrt{2}\mathcal{N}}[e^{-i\mu}\sqrt{\mathcal{G}}\partial_x - e^{-i\nu}\sqrt{\mathcal{E}}\partial_y], \end{aligned} \tag{20}$$

where μ and ν satisfy the condition $\cos(\mu-\nu) = \mathcal{F}/\sqrt{\mathcal{E}\mathcal{G}}$. If we compute these operators, the calculations to obtain the spin coefficients are cumbersome because they are expressed in terms of partial differential equations. However, we have additional information to calculate these functions: we know that $\rho = \overline{\rho}$ if and only if D is locally surface-orthogonal, i.e., there exists locally a family of two-dimensional surfaces such that, at each point, D is orthogonal to the tangent space to the surface passing through that point [4]. Furthermore, the Gauss and mean curvatures of these surfaces are given by

$$K = 2(\rho^2 - \kappa\overline{\kappa}), \quad H = -\sqrt{2}\rho. \tag{21}$$

Now we take into account for this case that D is in fact surface-orthogonal, and the surfaces orthogonal to D (the direction of the refracted light ray) are the refracted wavefronts. The curvatures of these wavefronts have been calculated [13], and from the calculations of the commutator of the operators, (20), we obtain that $\alpha = 0$ (this field is tangent to the congruence of light rays, as it should be) and, also, that κ is a complex function. This implies that κ is of the form $\kappa = |\kappa|e^{i\chi}$, where χ is an angle that, under a rotation of the form $e^{-i\chi/4}$, sends κ to its module.

Using all that information, we can now obtain the optical scalars corresponding to the optical congruence. The expansion, the twist, and the shear for the refracted congruence are

$$\begin{aligned} \Theta &= \frac{1}{2\sqrt{2}}\frac{\partial}{\partial \ell}(\ln \mathcal{N}), \\ \omega &= 0, \\ \kappa &= \frac{1}{2\sqrt{2}}\frac{\sqrt{H_1^2 - 4H_0H_2}}{\mathcal{N}}\, e^{i\chi}. \end{aligned} \tag{22}$$

From the previous equations there are several observations. First, the expansion of the congruence is measured by the change of the Jacobian of the mapping ($\mathcal{N}$): if this function remains constant with respect to the direction of propagation of the light ray, there is no expansion. Second, there is no twist, and this is because the congruence is a solution of the eikonal equation and, therefore, proceeds from a gradient field (an irrotational vector field). On the other hand, the module of the κ coefficient is a core quantity in the analysis; observe that the numerator of this module is the discriminant of the Jacobian $\mathcal{N}$ and this discriminant indicates the presence of caustic points, i.e., if it is 0, the congruence diverges; otherwise, it converges (at least locally, because the information that the shear provides is local). Finally we see that these functions have enough information about the evolution of the light rays, in other words, the optical scalars (and therefore the congruence) are another way to speak about the beam.

As a direct example to analyze the information carried by the optical scalars, we take the case of two optical media with the same refracting index. For $\gamma = 1$, $\hat{\mathbf{R}} = \hat{\mathbf{I}}$ and therefore the Jacobian $\mathcal{N}$ in this case is

$$\mathcal{N} = \frac{\sqrt{EG - F^2}\,(\hat{\mathbf{I}}\cdot\hat{\mathbf{N}})}{|\mathbf{r} - \mathbf{s}|^2}\,\tau^2, \tag{23}$$

where E, F, and G are the metric coefficients for the arbitrary refracting surface $\mathbf{r}$ and we remember that τ is the parameter that labels each wavefront in the wavefront train. With this information we can calculate the Gauss and mean curvatures of the refracted wavefronts for this case, and we obtain

$$K = \frac{1}{\tau^2}, \quad H = -\frac{1}{\tau}, \tag{24}$$

and these curvatures are the ones obtained for a sphere with its center in $\mathbf{s}$ and with radius τ. That is, when the two media are the same, the point source generates spherical wavefronts, and this *does not depend on the refracting surface*. Furthermore, using Eqs. (22) directly, the optical scalars for this case are

$$\Theta = \frac{1}{\sqrt{2}\tau}, \quad \omega = 0, \quad \kappa = 0, \tag{25}$$

so we can conclude that this congruence is expanding, and the expansion decreases as τ becomes larger. Equations (25) for large τ are equivalent to the case where the source is very far and we have plane wavefronts: this is confirmed by the curvatures K and H, because for very large τ, they are both 0. On the other hand, there is no shear, and the congruence never converges, so there are no caustic points.

V. APPLICATIONS

Optical scalars have information on the refracted light rays of the optical system, that is, information on the beam. In particular, the module of the shear $|\kappa|$ is the measure of the amount of distortion of a hypothetical test circle that is locally placed on the refracted wavefronts. Thus, we claim that this distortion could be related to aberrations of the optical system.

We focus on the case of monochromatic aberrations based on the definitions given by Rayces [6,7]. Under these definitions the wavefronts are nothing but deformations of ideal spherical wavefronts. The above means that for each point in the "base sphere" there is a corresponding deformation measured by the function $W(x,y)$: this is the aberration function of the system. In the following part we analyze this concept and we link to it the optical scalars obtained in previous sections.

A. Aberrated wavefronts and the aberration function

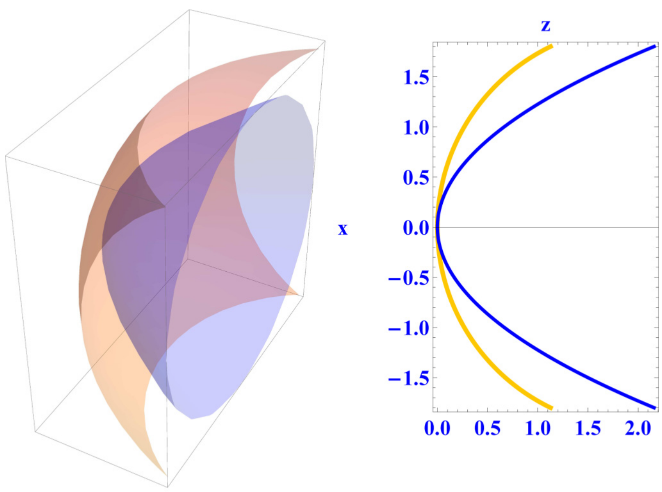
We can approach the analysis of wavefront aberrations through our formalism. A graphic description of the system is shown in Fig. 5(a). Following Rayces' analysis of the aberration function $W(x,y)$ [6], we consider a "deformed sphere" described by the coordinate patch

$$\mathbf{r}_s(x,y) = (x,y,P - \sqrt{(q - W(x,y))^2 - x^2 - y^2}), \tag{26}$$

where q is the radius of a semisphere without deformations and centered at a point P along the z axis. Therefore Eq. (26) is, in this case, the equation of an aberrated wavefront. Observe that the deformation at each point in this shell is given by the aberration function $W(x,y)$ [see Fig. 5(b)], and with this information the vector field that describes the evolution of the refracted wavefronts is

$$\mathbf{X}(x,y,\sigma) = \mathbf{r}_s + \sigma\hat{\mathbf{R}}, \tag{27}$$

where σ parameterizes the evolution of the wavefront and $\hat{\mathbf{R}}$ represents the normal to these wavefronts, and it is



(a) (b)

FIG. 5. (a) Plot representing a section of a sphere (orange surface at left) and a wavefront (blue surface at right). The difference between the two surfaces is measured by the aberration function $W(x,y)$. (b) Outlines of the sphere and the wavefront: the yellow curve at left is the sphere; the blue curve at the right is the wavefront.

given by

$$\hat{\mathbf{R}} = \frac{(\mathbf{r}_s)_x \times (\mathbf{r}_s)_y}{|(\mathbf{r}_s)_x \times (\mathbf{r}_s)_y|}. \tag{28}$$

We denote $R = |(\mathbf{r}_s)_x \times (\mathbf{r}_s)_y|$ and this quantity is important. Given the train of refracted wavefronts described by Eq. (2), we can choose a wavefront in the train by fixing ℓ. By comparing this wavefront with Eq. (27), we conclude that $\mathcal{N}$ and R are related and in fact $\mathcal{N} = R$, that is, R is the Jacobian of the optical system. An important observation is that under the parametrization, (27), R is a quantity that does not depend on the refracting surface but, instead, is obtained only with the information on the refracted congruence. Furthermore, all the quantities are functions of $W(x,y)$. Therefore, the components for the $\mathbf{X}$ vector field are

$$\begin{aligned} X(x,y,\sigma) &= x - \frac{\sigma}{R}[(q-W)W_x + x], \\ Y(x,y,\sigma) &= y - \frac{\sigma}{R}[(q-W)W_y + y], \\ Z(x,y,\sigma) &= P - \sqrt{q^2 - x^2 - y^2} + \frac{\sigma}{R}\sqrt{q^2 - x^2 - y^2}. \end{aligned} \tag{29}$$

We can now study the imaging formation process if we consider a screen placed in $Z_P = P$. This condition fixes $\sigma = R$ such that Eqs. (29) are reduced as

$$\begin{aligned} X_P(x,y,P) &= -(q-W)\frac{\partial W}{\partial x}, \\ Y_P(x,y,P) &= -(q-W)\frac{\partial W}{\partial y}, \\ Z_P(x,y,P) &= P, \end{aligned} \tag{30}$$

and we observe that the first two of these equations are the equations for the analytic transverse ray aberrations [6].

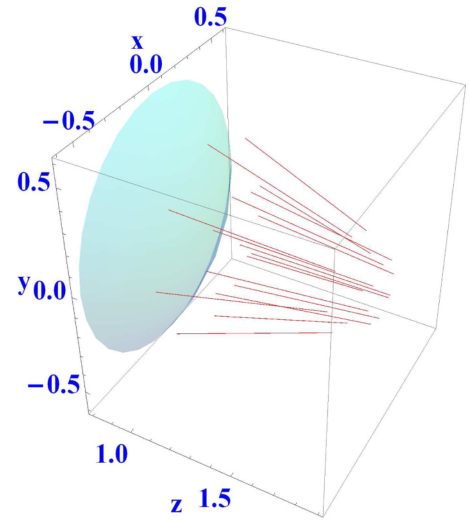


FIG. 6. Plot of a planoconic convex thin lens (a spherical one in this example). Red lines represent the congruence of refracted rays generated by this lens.

B. Relation between the shear and the aberration function

In the previous sections we have analyzed the importance of optical scalars for a congruence of light rays. It is useful to obtain a direct relation of the module of the shear κ with the aberration function $W(x,y)$. Remembering the definition of the metric coefficients, (5), for the surface, (26), we have the H's functions,

$$\begin{aligned} H_0(x,y) &= R, \\ H_1(x,y) &= -\,2RH(x,y), \\ H_2(x,y) &= RK(x,y), \end{aligned} \tag{31}$$

where $H(x,y)$ and $K(x,y)$ are the mean and Gauss curvatures associated with the refracted wavefronts, but these curvatures are now functions of the aberration function $W(x,y)$. Using Eqs. (22) a direct calculation of the optical scalars for this system gives

$$\Theta = -\frac{H}{\sqrt{2}}, \quad \omega = 0, \quad |\kappa| = \frac{1}{\sqrt{2}}\sqrt{H^2 - K}. \tag{32}$$

In this case, the magnitude of the shear is related to the principal curvatures of the wavefront, which are given by $k_{1,2} = H \pm \sqrt{H^2 - K}$. In fact, to solve the previous equations for $|\kappa| = 0$ we require that the surface that represents the wavefront consist of umbilic points (points where $k_1 = k_2 = k$) in that region, but this kind of surface is contained in either a plane or a sphere [9]: the above means that there is an equivalence between the shear and the aberration function. As before, in the case where $W(x,y) = 0$ for each point in the coordinate patch, (26), we recover the curvatures of a sphere, which implies that $|\kappa| = 0$ as expected, and this means that there is no distortion. To bear out the previous statement we study the example of a thin planoconvex lens illuminated with plane incident wavefronts (see Fig. 6) that follows the equation of conic surfaces of revolution $z(\rho) = [M + \sqrt{a^2 - (M+1)\rho^2}]/(M+1)$, where M is a parameter that is a function of the eccentricity of the conic and a is the radius of curvature of the conic [16]. The shear of this congruence is

$$|\kappa| = \frac{|\Omega(M+\gamma^2)|\rho^2\sqrt{a^2-M\rho^2}}{2\sqrt{2}\,\Xi}, \tag{33}$$

where Ω is given by

$$\Omega = \frac{\gamma\sqrt{a^2-(M+1)\rho^2}-\sqrt{a^2-(M+\gamma^2)\rho^2}}{\sqrt{a^2-M\rho^2}}, \tag{34}$$

and Ξ by

$$\Xi = |\Omega^2 a^2\ell^2 - \Omega\ell[2a^2-(M+\gamma^2)\rho^2]\sqrt{a^2-M\rho^2} + [a^2-(M+\gamma^2)\rho^2][a^2-M\rho^2]|. \tag{35}$$

In this case the condition $|\kappa| = 0$ is equivalent to $\Omega = 0$, and at the same time this means that $\gamma = 1$. So the only case with no shear is when the two media are the same, but as we showed before, this does not depend on the refracting surface.

It is important to remark that now we have at hand two ways to analyze aberrations. The first is by using the information on the refracting surface to calculate the functions, (4), and then calculate the shear. The second is using only the information on the refracted congruence (based on the refracted wavefronts) to calculate the shear and therefore the aberration function.

As an end point, we analyze the case where the aberration function has revolution symmetry, that is, where $W = W(\rho)$. This case is of special importance because it includes the common types of lenses used as optical systems.

Using Eqs. (32) the differential equation that establishes the connection between the shear and the aberration function $W(\rho)$ is

$$|\kappa| = \frac{|W'[(q-W)^3(1+W'^2)+\rho W'(3(q-W)^2-\rho^2)]+\rho(q-W)[(q-W)^2-\rho^2]W''|}{2\sqrt{2}\rho[(q-W)[(q-W)(1+W'^2)+2\rho W']]^{3/2}}, \tag{36}$$

where the apostrophe denotes the derivative with respect to ρ. Here we see explicitly that if the aberration function is equal to 0 or constant, then the shear of the congruence is 0. However, that the aberration function is constant for each point in the shell means that the deformation produced by W results in another spherical shell and this case is not aberrated. It seems that this is the only case where the aberration function is not 0 and the shear is 0.

VI. CONCLUSIONS

In this work we have studied an optical system with two refracting indices, n_1 and n_2, separated by an arbitrary refracting surface described by the **r** coordinate patch. This surface generates a refracted congruence of light rays that is described by the **X** vector field, and the normal of the refracted wavefronts is the $\hat{\mathbf{R}}$ vector. We showed that this congruence can be analyzed through the optical scalars: these functions encode the information related to the expansion, rotation, and evolution of the congruence.

We applied this formalism to study the wavefront and transverse aberrations for an optical congruence and we have shown that it is possible to make a direct calculation of such aberrations. In fact, the shear of the congruence and the aberration function are related, and the link between them is the Gauss and mean curvatures of the wavefronts. This is important because aberrations are measurable objects of the optical system that are linked with the wavefronts, which are not observable.

On the other hand, we have shown that we can relate the aberrations of the optical system only to the information on the refracted rays. That is, it is not necessary to know the refracting surface that produces the refraction, but based only on the refracted information received by an observer, we could reconstruct the physical information on the system. The previous statement is due to the fact that according to our calculations, the information encoded in the congruence of light rays allows us to establish the distortions of the images.

It is important to note that equaling Eqs. (33) and (36) the aberration function W for *all* the conic surfaces of revolution in that case could be solved. Equation (36) is a non linear differential equation that can be worked out with numerical methods, so we have another way to study aberrations in a form to complement Seidel's. Hence by studying this equation carefully, we can express the aberration $W(\rho)$ as a function of the shear of the light rays.

We believe that the information on beams characterized by congruences of light rays could be related to aberrations that give information on physical systems: that is, one has to compute the relationship between the expansion (Θ) and the aberration functions to determine the distortion of the images. For example, all this machinery could be applied to the observation of systems such as black holes, galaxies, and other kinds of cosmological objects [17] in such a way that, directly with optical data, we would be able to obtain the metric information on the space-time where these objects lie.

ACKNOWLEDGMENTS

Omar de J. Cabrera-Rosas, Paula Ortega-Vidals, Salvador Alejandro Juárez-Reyes, Citlalli Teresa Sosa-Sánchez, Ernesto Espíndola Ramos, and Israel Julián Macías were supported by CONACyT scholarships. G.S.-O. acknowledges financial support from Sistema Nacional de Investigadores (SNI). This work received partial support from VIEP-BUAP (Vicerrectoría de Investigación y Estudios de Posgrado de la Benemérita Universidad Autónoma de Puebla).

[1] E. Poisson, *A Relativist's Toolkit: The Mathematics of Black Hole Mechanics* (Cambridge University Press, Cambridge, UK, 2004).

[2] S. Chandrasekhar, *The Mathematical Theory of Blackholes* (Oxford University Press, New York, 1983).

[3] R. Sachs, Gravitational waves in general relativity. VI. The outgoing radiation condition, Proc. R. Soc. London A **264**, 309 (1961).

[4] G. F. Torres del Castillo, *3-D Spinors, Spin Weighted Functions and Their Applications. Progress in Mathematical Physics* (Birkhäuser, Boston, 2003), Vol. 20.

[5] W. T. Welford, *Aberrations of Optical Systems* (CRC Press, Boca Raton, FL, 1986).

[6] J. L. Rayces, Exact relation between wave aberration and ray aberration, Opt. Acta **11**, 85 (1964).

[7] B. R. A. Nijboer, *The Diffraction Theory of Optical Aberrations*, Ph.D. Thesis, University of Groningen (1942).

[8] M. Marciano-Melchor, E. Navarro-Morales, E. Román-Hernández, J. G. Santiago-Santiago, G. Silva-Ortigoza, R. Silva-Ortigoza, and R. Suárez-Xique, The point-characteristic function, wavefronts, and caustic of a spherical wave refracted by an arbitrary smooth surface, J. Opt. Soc. Am. A **29**, 1035 (2012).

[9] J. Oprea, *Differential Geometry and Its Applications* (Pearson, Aurora, CO, 1987).

[10] O. N. Stavroudis, *The Mathematics of Geometrical and Physical Optics, the K-Function and Its Ramifications* (Wiley-VCH, Weinheim, 2006).

[11] V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko, *Singularities of Differentiable Maps* (Birkhäuser, Boston, 1995), Vol. I.

[12] V. I. Arnold, *Catastrophe Theory* (Springer-Verlag, Berlin, 1986).

[13] O. de J. Cabrera-Rosas, P. Ortega-Vidals, S. A. Juárez-Reyes, and G. Silva-Ortigoza, Curvatures of the refracted wavefronts and ronchigrams for a plano arbitrary lens, Opt. Soc. Am. A **33**, 1760 (2016).

[14] O. N. Stavroudis and R. C. Fronczek, Generalized ray tracing and the caustic surface, Opt. Laser Technol. **10**, 185 (1978).

[15] A. M. Kassim and D. L. Shealy, Wavefront equation, caustics, and wave aberration function of simple lenses and mirrors, Appl. Opt. **27**, 516 (1988).

[16] D. Malacara, *Mathematical Representation of an Optical Surface and Its Characteristics (Optical Shop Testing)*, 3rd ed. (Wiley, New York, 2007), Chap. 9.

[17] N. Bretón *et al.*, Towards the Ronchi test for gravitational lenses: The gravitoronchigram, J. Opt. **19**, 065602 (2017).