Geometrical optics in the optical length geometry

G.F. Torres del Castillo
Departamento de Física Matemática, Instituto de Ciencias Universidad Autónoma de Puebla
72570 Puebla, Pue., Mexico

C.J. Pérez Ballinas
Instituto Nacional de Astrofísica, Óptica y Electrónica
Apartado postal 216, 72000 Puebla, Pue., Mexico

Recibido el 20 de agosto de 1999; aceptado el 26 de enero de 2000

Making use of a metric conformal to the space metric, corresponding to the optical length, the Fermat principle is derived from the eikonal equation. Other characteristics of the propagation of light are also expressed in terms of this metric.

Keywords: Geometrical optics; eikonal equation; Riemannian geometry

Empleando una métrica conforme a la métrica del espacio, correspondiente al camino óptico, se deduce el principio de Fermat a partir de la ecuación iconal. Otras características de la propagación de la luz se expresan también en términos de esta métrica.

Descriptores: Óptica geométrica; ecuación iconal; geometría Riemanniana

PACS: 42.15.-i; 02.40.Ky

1. Introduction

In the short wavelength limit, the propagation of light can be associated with certain curves (the light rays), which give the direction of energy flow and, in this approximation, several aspects of the behavior of light are describable in geometrical terms (see, *e.g.*, Refs. 1 and 2). The light rays in an isotropic medium, characterized by a refractive index n, can be found making use of the Fermat principle, according to which the optical length

$$\int_A^B n\, ds$$

of an actual ray between any two points A and B is stationary as compared with the optical length of arbitrary neighboring curves joining A and B [1,2]. This means that the light rays are geodesics of the metric $n^2 ds^2$ (see also Ref. 1 and the references cited therein) and suggests that, for some purposes, the metric $n^2 ds^2$ may be more convenient than the space metric ds^2.

In this paper we make use of some results of differential geometry in the study of geometrical optics, deriving Fermat's principle from the eikonal equation and considering the effect of the curvature of the metric $n^2 ds^2$ on the propagation of light. The approach followed here allows us to obtain in a simple way the Fermat principle starting from the Maxwell equations. In Sect. 2 we summarize the derivation of the eikonal equation starting from Maxwell's equations, following Ref. 1. In Sect. 3 we review the relationship between geodesics and the eikonal equation for a constant refractive index. In Sect. 4 we derive Fermat's principle starting from the eikonal equation, proving that the light rays are geodesics of the metric $n^2 ds^2$. We also consider the effect of the curvature of this metric on the spreading of a pencil of light rays. The standard tensor formalism is employed in Sects. 3 and 4 (see, *e.g.*, Refs. 3–5).

2. The eikonal equation

The source-free Maxwell equations in an isotropic medium admit time-harmonic solutions of the form

$$\mathbf{E}(\mathbf{r},t) = \mathbf{e}(\mathbf{r}) e^{ik_0 S(\mathbf{r}) - i\omega t},$$
$$\mathbf{H}(\mathbf{r},t) = \mathbf{h}(\mathbf{r}) e^{ik_0 S(\mathbf{r}) - i\omega t}, \tag{1}$$

where $k_0 \equiv \omega/c$, $\mathbf{e}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$ are (possibly complex) vector fields and $S(\mathbf{r})$ is a real-valued function. Indeed, substituting Eqs. (1) into the Maxwell equations, with $\mathbf{D} = \varepsilon\,\mathbf{E}$ and $\mathbf{B} = \mu\,\mathbf{H}$, it follows that

$$\operatorname{grad} S \times \mathbf{h} + \varepsilon\,\mathbf{e} = \frac{i}{k_0} \operatorname{curl} \mathbf{h},$$
$$\mathbf{e} \cdot \operatorname{grad} S = \frac{i}{k_0}\frac{1}{\varepsilon} \operatorname{div}(\varepsilon\,\mathbf{e}),$$
$$\operatorname{grad} S \times \mathbf{e} - \mu\,\mathbf{h} = \frac{i}{k_0} \operatorname{curl} \mathbf{e},$$
$$\mathbf{h} \cdot \operatorname{grad} S = \frac{i}{k_0}\frac{1}{\mu} \operatorname{div}(\mu\,\mathbf{h}). \tag{2}$$

Therefore, when $\lambda_0 \equiv 2\pi/k_0$ is very small (as compared with $\varepsilon/|\operatorname{grad}\varepsilon|$ and $\mu/|\operatorname{grad}\mu|$), Eqs. (2) reduce to

$$\operatorname{grad} S \times \mathbf{h} + \varepsilon\,\mathbf{e} = 0,$$
$$\mathbf{e} \cdot \operatorname{grad} S = 0,$$
$$\operatorname{grad} S \times \mathbf{e} - \mu\,\mathbf{h} = 0,$$
$$\mathbf{h} \cdot \operatorname{grad} S = 0, \tag{3}$$

which imply that

$$(\text{grad}\, S)^2 = n^2, \tag{4}$$

where $n = \sqrt{\varepsilon\mu}$ is the refractive index. Equation (4) is known as the eikonal equation.

According to Eqs. (3), the time average of the Poynting vector, which is tangent to the light rays, is normal to the wavefronts $S = \text{const}$. By computing curl curl $\mathbf{e}$, making use of Eqs. (2) and (4), neglecting the higher-order terms in λ_0, one obtains

$$(\text{grad}\, S \cdot \text{grad})\, \mathbf{e} = \frac{1}{2}[\text{grad}\, S \cdot \text{grad}\, \ln \mu - \nabla^2 S]\, \mathbf{e} - (\mathbf{e} \cdot \text{grad}\, \ln n)\, \text{grad}\, S, \tag{5}$$

which gives the variation of $\mathbf{e}$ along the light rays. The component of the left-hand side of Eq. (5) in the direction of grad S coincides with the corresponding component of the right-hand side by virtue of the eikonal equation (4). On the other hand, the component of Eq. (5) perpendicular to grad S can be obtained by taking the cross product of each side of this equation with grad S twice; then, making use of Eqs. (3)–(5), one finds that

$$\mathbf{e} \cdot [(\text{grad}\, S \cdot \text{grad})\, \text{grad}\, S - n\, \text{grad}\, n] = 0.$$

Similarly, making use of Eqs. (2) to compute curl curl $\mathbf{h}$, it follows that

$$(\text{grad}\, S \cdot \text{grad})\, \mathbf{h} = \frac{1}{2}[\text{grad}\, S \cdot \text{grad}\, \ln \varepsilon - \nabla^2 S]\, \mathbf{h} - (\mathbf{h} \cdot \text{grad}\, \ln n)\, \text{grad}\, S. \tag{6}$$

Hence, $(\text{grad}\, S \cdot \text{grad})\, \text{grad}\, S - n\, \text{grad}\, n$ is orthogonal to $\mathbf{h}$ and, from Eq. (4), one finds that it is also orthogonal to grad S, therefore

$$(\text{grad}\, S \cdot \text{grad})\, \text{grad}\, S = n\, \text{grad}\, n. \tag{7}$$

This equation gives the changes of direction of the light rays and is equivalent to Snell's law.

From Eqs. (3), (5) and (6) it follows that the directional derivatives along the light rays of the unit vectors that determine the direction of the electric and magnetic fields, $\hat{\mathbf{e}} \equiv \mathbf{e}/\sqrt{\mathbf{e} \cdot \mathbf{e}^*}$ and $\hat{\mathbf{h}} \equiv \mathbf{h}/\sqrt{\mathbf{h} \cdot \mathbf{h}^*}$, respectively, where the * denotes complex conjugation, are given by

$$(\text{grad}\, S \cdot \text{grad})\, \hat{\mathbf{e}} = -(\hat{\mathbf{e}} \cdot \text{grad}\, \ln n)\, \text{grad}\, S,$$

$$(\text{grad}\, S \cdot \text{grad})\, \hat{\mathbf{h}} = -(\hat{\mathbf{h}} \cdot \text{grad}\, \ln n)\, \text{grad}\, S. \tag{8}$$

It will be shown in Sect. 4 that Eqs. (7)–(8) mean that the direction of each of the three mutually orthogonal vectors, grad S, $\hat{\mathbf{e}}$ and $\hat{\mathbf{h}}$ does not change along a light ray if one makes use of the metric $n^2 ds^2$.

3. Geodesics

In this section we shall consider the relationship between an equation that can be regarded as the eikonal equation for a medium of constant refractive index and the geodesics of an arbitrary (Riemannian) space. The results of this section are actually more general than what we will require in the next section and they establish several connections between differential geometry, classical mechanics and geometrical optics.

Throughout this section we shall assume that the space has an arbitrary dimension N and that its metric, $ds^2 = g_{ij} dx^i dx^j$, may not be flat. Lower case Latin indices $i, j, k, \ldots$ run from 1 to N and there is summation over repeated indices. The basic result is contained in the following proposition, which is due to Gauss [6].

Proposition 1. If S is a function such that

$$(\text{grad}\, S)^2 = \text{const.} \tag{9}$$

then the field lines of grad S are geodesics [*cf.* Eq. (4)]. (In other words, the orthogonal trajectories to the surfaces $S =$ const. are geodesics.)

Proof. Equation (9), written in terms of an arbitrary coordinate system, is

$$g^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = \text{const}, \tag{10}$$

where the matrix (g^{ij}) is the inverse of (g_{ij}). Letting

$$p_i \equiv \frac{\partial S}{\partial x^i}, \tag{11}$$

from Eq. (10) we have

$$g^{ij} p_i p_j = \text{const}. \tag{12}$$

and taking the partial derivative of this relation with respect to x^k it follows that

$$\frac{\partial g^{ij}}{\partial x^k} p_i p_j + 2 g^{ij} p_i \frac{\partial p_j}{\partial x^k} = 0, \tag{13}$$

by virtue of the symmetry of g^{ij}. Equation (11) also implies that $\partial p_j / \partial x^k = \partial p_k / \partial x^j$, therefore, from Eq. (13) we obtain

$$g^{ij} p_i \frac{\partial p_k}{\partial x^j} = -\frac{1}{2} \frac{\partial g^{ij}}{\partial x^k} p_i p_j. \tag{14}$$

Now, let $x^i = x^i(\lambda)$ be parametric equations of the "field lines" (integral curves) of grad S, *i.e.*,

$$\frac{dx^i}{d\lambda} = g^{ij} \frac{\partial S}{\partial x^j}, \tag{15}$$

which, owing to Eq. (11), amounts to

$$\frac{dx^i}{d\lambda} = g^{ij} p_j, \qquad \text{or} \qquad p_i = g_{ij} \frac{dx^j}{d\lambda}. \tag{16}$$

Hence, from Eqs. (16), using the chain rule and Eq. (14) we find that

$$\begin{aligned}\frac{d^2x^i}{d\lambda^2} &= \frac{d}{d\lambda}(g^{ij}p_j) = \frac{\partial}{\partial x^k}(g^{ij}p_j)\frac{dx^k}{d\lambda} \\ &= \frac{\partial g^{ij}}{\partial x^k}p_j\frac{dx^k}{d\lambda} + g^{ij}\frac{\partial p_j}{\partial x^k}\frac{dx^k}{d\lambda} \\ &= \frac{\partial g^{ij}}{\partial x^k}p_j\frac{dx^k}{d\lambda} + g^{ij}\frac{\partial p_j}{\partial x^k}g^{km}p_m \\ &= \frac{\partial g^{ij}}{\partial x^k}p_j\frac{dx^k}{d\lambda} - \frac{1}{2}g^{ij}\frac{\partial g^{km}}{\partial x^j}p_kp_m. \end{aligned} \tag{17}$$

The partial derivatives of the components of g^{ij} can be expressed in terms of the Christoffel symbols according to [3–5]

$$\frac{\partial g^{ij}}{\partial x^k} = -\Gamma^i_{mk}g^{mj} - \Gamma^j_{mk}g^{mi}. \tag{18}$$

Then, from Eqs. (16)–(18) one obtains

$$\frac{d^2x^i}{d\lambda^2} + \Gamma^i_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda} = 0, \tag{19}$$

which are the equations for the geodesics of the metric $ds^2 = g_{ij}dx^idx^j$ *i.e.*, the solution of Eq. (19) passing through two given points A and B makes the length $\int_A^B ds$ an extremum, or, at least, gives it a stationary value [3–5].

Remark 1. Since any electrostatic field in Euclidean space can be expressed as $\mathbf{E} = -\text{grad}\,\varphi$, Proposition 1 implies that if the magnitude of $\mathbf{E}$ is constant, then its field lines are *straight lines*. Similarly, if the magnitude of the velocity of an irrotational flow is constant, the flow lines are straight lines.

Remark 2. The converse of Proposition 1 is also true; all the geodesics of a given metric (flat or curved) are locally determined by a *complete* solution of Eq. (10) (see the examples below).

Remark 3. Equation (10) can be regarded as the Hamilton-Jacobi equation for the characteristic function, corresponding to the Hamiltonian $H = \mathbf{p}^2/2$ of a free particle in a configuration space with metric $ds^2 = g_{ij}dx^idx^j$. The orbits of the particle in configuration space are geodesics.

Remark 4. The field lines of an *arbitrary* electrostatic field, $\mathbf{E}$, are geodesics of the metric $\mathbf{E}^2g_{ij}dx^idx^j$. Indeed, from the equation $\mathbf{E}^2 = g^{ij}(\partial\varphi/\partial x^i)(\partial\varphi/\partial x^j)$, it follows that

$$\frac{g^{ij}}{\mathbf{E}^2}\frac{\partial\varphi}{\partial x^i}\frac{\partial\varphi}{\partial x^j} = 1,$$

which is of the form (10) with g^{ij} replaced by $g^{ij}/\mathbf{E}^2$. (Note, however, that not every geodesic of the metric $\mathbf{E}^2g_{ij}dx^idx^j$ is a field line.) An analogous result holds for irrotational flows.

Remark 5. From Eqs. (12) and (16) it follows that $g_{ij}(dx^i/d\lambda)(dx^j/d\lambda) = \text{const.}$, which means that the parameter λ defined by Eqs. (16) is a constant multiple of the arc length.

Remark 6. Using Eqs. (16) one finds that Eq. (14) amounts to

$$\frac{dp_k}{d\lambda} = -\frac{1}{2}\frac{\partial g^{ij}}{\partial x^k}p_ip_j. \tag{20}$$

This last equation together with the first equation (16) are the Hamilton equations for the Hamiltonian $H = \frac{1}{2}g^{ij}p_ip_j$, with λ in place of the time (*cf.* Remark 3).

As pointed out above (Remark 2), a complete solution of Eq. (10) gives locally all the geodesics of the metric g_{ij}. Indeed, if $S(x^i, \alpha_a)$ is a solution of Eq. (10) depending on $N-1$ parameters $\alpha_1, \ldots, \alpha_{N-1}$ ($i = 1, \ldots, N; a = 1, \ldots, N-1$), such that the rank of the matrix $(\partial^2 S/\partial\alpha_a\partial x^i)$ is equal to $N-1$, then taking the partial derivative of Eq. (10) with respect to α_a, it follows that

$$2g^{ij}\frac{\partial S}{\partial x^i}\frac{\partial}{\partial x^j}\left(\frac{\partial S}{\partial\alpha_a}\right) = 0$$

and making use of Eq. (15) we obtain

$$\frac{dx^j}{d\lambda}\frac{\partial}{\partial x^j}\left(\frac{\partial S}{\partial\alpha_a}\right) = 0, \tag{21}$$

which means that

$$\frac{d\beta^a}{d\lambda} = 0, \tag{22}$$

where

$$\beta^a \equiv \frac{\partial S}{\partial\alpha_a}, \tag{23}$$

i.e., the functions β^a defined by Eq. (23) are constant along the geodesics which are the orthogonal trajectories to the surfaces $S = \text{const}$. The condition that the rank of the matrix $(\partial^2 S/\partial\alpha_a\partial x^i)$ be equal to $N-1$ guarantees that, locally, for each set of values of the constants $\alpha_1, \ldots, \alpha_{N-1}$, $\beta^1, \ldots, \beta^{N-1}$, the $N-1$ equations (23) determine a curve. (Note that this procedure is analogous to that followed in classical mechanics to solve the equations of motion making use of the Hamilton-Jacobi equation [6,7].)

A very simple, but illustrative, example is given by finding a complete solution of Eq. (10) in the three-dimensional Euclidean space, using cartesian coordinates, where $g_{ij} = \delta_{ij}$. Choosing the constant on the right-hand side of Eq. (10) equal to 1, a complete solution of

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = 1 \tag{24}$$

can be obtained by separation of variables. Substituting $S = f(x) + g(y) + h(z)$ into Eq. (24) one finds that $S = \alpha_1 x + \alpha_2 y + \sqrt{1-\alpha_1^2-\alpha_2^2}\,z$, where α_1 and α_2 are separation constants. Then, according to Eq. (23) the constants β^1 and β^2 are given by

$$\begin{aligned}\beta^1 &= \frac{\partial S}{\partial\alpha_1} = x - \frac{\alpha_1}{\sqrt{1-\alpha_1^2-\alpha_2^2}}z, \\ \beta^2 &= \frac{\partial S}{\partial\alpha_2} = y - \frac{\alpha_2}{\sqrt{1-\alpha_1^2-\alpha_2^2}}z. \end{aligned} \tag{25}$$

The intersection of the two planes given by Eqs. (25) is a straight line and any straight line can be represented in the form (25). In the present case, the surfaces $S = \text{const.}$ are planes.

4. The optical metric

The eikonal equation (4) written in terms of an arbitrary system of coordinates is

$$g^{ij}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j} = n^2, \tag{26}$$

where, now, $i, j, \ldots$ run from 1 to 3. Equation (26) is equivalent to

$$\tilde{g}^{ij}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j} = 1, \tag{27}$$

where we have introduced

$$\tilde{g}^{ij} \equiv \frac{g^{ij}}{n^2}. \tag{28}$$

The inverse of the matrix $(\tilde{g}^{ij})$, given by

$$\tilde{g}_{ij} = n^2 g_{ij}, \tag{29}$$

corresponds to the metric

$$\tilde{g}_{ij}dx^i dx^j = n^2 g_{ij} dx^i dx^j, \tag{30}$$

which will be called the *optical metric*. Owing to the factor n^2 in Eq. (30), the "distance" between two points defined by the optical metric, along some curve joining these two points, is not the actual length of the curve but its *optical length*, which amounts to the vacuum velocity of light multiplied by the time needed for light to travel from one point to the other along the curve.

4.1. Derivation of the Fermat principle from the eikonal equation

According to Proposition 1, from Eq. (27) it follows that the orthogonal trajectories to the geometrical wavefronts $S = \text{const.}$ (the light rays) are geodesics with respect to the optical metric (30), *i.e.*, a light ray connecting two points A, B corresponds to a trajectory joining A and B such that its optical length has a stationary value, which demonstrates Fermat's principle. (Note that, since the space metric and the optical metric differ by a factor, the light rays are orthogonal to the wavefronts according to either metric.)

Thus, the light rays satisfy the geodesics equations for the optical metric [*cf.* Eq. (19)]

$$\frac{d^2x^i}{d\lambda^2} + \tilde{\Gamma}^i_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda} = 0, \tag{31}$$

where λ is the optical length (see Remark 5) and the $\tilde{\Gamma}^i_{jk}$ are the Christoffel symbols corresponding to the optical metric, *i.e.*,

$$\begin{aligned}\tilde{\Gamma}^i_{jk} &= \frac{1}{2}\tilde{g}^{im}\left(\frac{\partial \tilde{g}_{mk}}{\partial x^j} + \frac{\partial \tilde{g}_{mj}}{\partial x^k} - \frac{\partial \tilde{g}_{jk}}{\partial x^m}\right)\\ &= \Gamma^i_{jk} + \delta^i_k\frac{\partial \ln n}{\partial x^j} + \delta^i_j\frac{\partial \ln n}{\partial x^k} - g_{jk}g^{im}\frac{\partial \ln n}{\partial x^m},\end{aligned} \tag{32}$$

where the Γ^i_{jk} are the Christoffel symbols corresponding to the original metric g_{ij}.

The first equation (8), written in terms of components, is

$$\frac{\partial S}{\partial x^j}g^{jk}\nabla_k \hat{e}^i = -\hat{e}^k\left(\frac{\partial \ln n}{\partial x^k}\right)g^{ij}\frac{\partial S}{\partial x^j}, \tag{33}$$

where ∇_k denotes the covariant derivative compatible with the metric g_{ij} [3–5] ($\nabla_k \hat{e}^i = \partial \hat{e}^i/\partial x^k + \Gamma^i_{jk}\hat{e}^j$), hence, making use of Eqs. (32) and (3), one can express Eq. (33) in terms of the covariant derivative compatible with the metric $\tilde{g}_{ij}$, denoted by $\tilde{\nabla}_k$,

$$\begin{aligned}\frac{\partial S}{\partial x^j}g^{jk}\tilde{\nabla}_k \hat{e}^i &= \frac{\partial S}{\partial x^j}g^{jk}\left[\nabla_k \hat{e}^i + (\tilde{\Gamma}^i_{mk} - \Gamma^i_{mk})\hat{e}^m\right]\\ &= -\hat{e}^k\left(\frac{\partial \ln n}{\partial x^k}\right)g^{ij}\frac{\partial S}{\partial x^j} + \frac{\partial S}{\partial x^j}g^{jk}\left(\delta^i_k\frac{\partial \ln n}{\partial x^m} + \delta^i_m\frac{\partial \ln n}{\partial x^k} - g_{mk}g^{ir}\frac{\partial \ln n}{\partial x^r}\right)\hat{e}^m\\ &= \frac{\partial S}{\partial x^j}g^{jk}\frac{\partial \ln n}{\partial x^k}\hat{e}^i - \frac{\partial S}{\partial x^j}g^{ir}\frac{\partial \ln n}{\partial x^r}\hat{e}^j = \frac{\partial S}{\partial x^j}g^{jk}\frac{\partial \ln n}{\partial x^k}\hat{e}^i,\end{aligned}$$

thus

$$\frac{\partial S}{\partial x^j}\tilde{g}^{jk}\tilde{\nabla}_k\left(\frac{\hat{e}^i}{n}\right) = 0. \tag{34}$$

This means that, using the optical metric, $\hat{e}^i/n$ is covariantly constant along the light rays (*i.e.*, according to the geometry defined by the optical metric, $\hat{e}^i/n$ does not change along each light ray) and, in an entirely similar manner, from the second equation (8), it follows that $\hat{h}^i/n$ is also covariantly constant along the light rays (*cf.* Ref. 1 and the references cited therein).

Similarly, writing Eq. (7) in the form

$$\frac{\partial S}{\partial x^j}g^{jk}\nabla_k\left(g^{ir}\frac{\partial S}{\partial x^r}\right) = ng^{ij}\frac{\partial n}{\partial x^j},$$

making use of Eqs. (32) and (26), one finds that

$$\begin{aligned}\frac{\partial S}{\partial x^j} g^{jk} \tilde{\nabla}_k \left(g^{ir} \frac{\partial S}{\partial x^r} \right) &= \frac{\partial S}{\partial x^j} g^{jk} \left[\nabla_k \left(g^{ir} \frac{\partial S}{\partial x^r} \right) + (\tilde{\Gamma}^i_{mk} - \Gamma^i_{mk}) g^{mr} \frac{\partial S}{\partial x^r} \right] \\ &= n g^{ij} \frac{\partial n}{\partial x^j} + \frac{\partial S}{\partial x^j} g^{ji} \frac{\partial \ln n}{\partial x^m} g^{mr} \frac{\partial S}{\partial x^r} + \frac{\partial S}{\partial x^j} g^{jk} \frac{\partial \ln n}{\partial x^k} g^{ir} \frac{\partial S}{\partial x^r} - \frac{\partial S}{\partial x^j} g^{jr} \frac{\partial S}{\partial x^r} g^{is} \frac{\partial \ln n}{\partial x^s} \\ &= 2 \frac{\partial S}{\partial x^j} g^{jk} \frac{\partial \ln n}{\partial x^k} g^{ir} \frac{\partial S}{\partial x^r},\end{aligned}$$

which amounts to

$$\frac{\partial S}{\partial x^j} g^{jk} \tilde{\nabla}_k \left(\frac{g^{ir}}{n^2} \frac{\partial S}{\partial x^r} \right) = 0,$$

or, equivalently, owing to Eq. (28),

$$\frac{\partial S}{\partial x^j} \tilde{g}^{jk} \tilde{\nabla}_k \left(\tilde{g}^{ir} \frac{\partial S}{\partial x^r} \right) = 0, \tag{35}$$

this means that the (contravariant) vector $\tilde{g}^{ir} \partial S / \partial x^r$ (the gradient of S with respect to the optical metric), which is tangent to the light rays, is also covariantly constant along the light rays. Taking into account the fact that $dx^i/d\lambda = \tilde{g}^{ij} \partial S / \partial x^j$ [*cf.* Eq. (15)], one concludes that Eq. (35) is equivalent to Eq. (31).

It may be noticed that the magnitude of the three mutually orthogonal vectors $\hat{e}^i/n$, $\hat{h}^i/n$ and $\tilde{g}^{ij} \partial S/\partial x^j$, which are covariantly constant along the light rays, with respect to the optical metric is equal to 1. Indeed, from Eqs. (29) and (27) we have, $\tilde{g}_{ij} (\hat{e}^i/n)(\hat{e}^j/n) = g_{ij} \hat{e}^i \hat{e}^j = 1$, $\tilde{g}_{ij} (\hat{h}^i/n)(\hat{h}^j/n) = g_{ij} \hat{h}^i \hat{h}^j = 1$, and $\tilde{g}_{ij} (\tilde{g}^{ir} \partial S/\partial x^r) (\tilde{g}^{jm} \partial S/\partial x^m) = \tilde{g}^{jm} (\partial S/\partial x^j)(\partial S/\partial x^m) = 1$.

Denoting by v^i the components of the tangent vector to the light rays, $v^i = dx^i/d\lambda$ (which as pointed out above, is a unit vector with respect to the optical metric), Eqs. (7), (31) or (35) can be written as

$$v^k \tilde{\nabla}_k v^i = 0. \tag{36}$$

Remark 7. Applying Eqs. (16) and (20) to the optical metric and recalling that, in this case, $d\lambda = c\, dt$, one finds that the light rays are given by Hamilton's equations with the Hamiltonian $H = (c/2) \tilde{g}^{ij} p_i p_j = (c/2n^2) g^{ij} p_i p_j$, which, apart from an additive constant, is the Hamiltonian given in Ref. 8 (see also Ref. 9).

4.2. Curvature

Even if the metric g_{ij} is flat (usually the metric of the three-dimensional Euclidean space), the optical metric may have a nonvanishing curvature. In Riemannian geometry, the curvature is measured by the curvature tensor [3–5]

$$R^i{}_{jkl} = \frac{\partial \Gamma^i_{jl}}{\partial x^k} - \frac{\partial \Gamma^i_{jk}}{\partial x^l} + \Gamma^i_{mk} \Gamma^m_{jl} - \Gamma^i_{ml} \Gamma^m_{jk} \tag{37}$$

and by substituting Eq. (32) into Eq. (37) one finds that the curvature of the optical metric is given by

$$\begin{aligned}\tilde{R}^i{}_{jkl} &= R^i{}_{jkl} + \delta^i_l \nabla_k \nabla_j \ln n - \delta^i_k \nabla_l \nabla_j \ln n \\ &\quad + g^{im} (g_{jk} \nabla_m \nabla_l \ln n - g_{jl} \nabla_m \nabla_k \ln n) \\ &\quad + (\nabla_j \ln n)(\delta^i_k \nabla_l \ln n - \delta^i_l \nabla_k \ln n) \\ &\quad + g^{im} (\nabla_m \ln n)(g_{jl} \nabla_k \ln n - g_{jk} \nabla_l \ln n) \\ &\quad + (\delta^i_l g_{jk} - \delta^i_k g_{jl}) g^{mp} (\nabla_m \ln n)(\nabla_p \ln n),\end{aligned} \tag{38}$$

where $R^i{}_{jkl}$ is the curvature tensor of the original metric g_{ij} and ∇_k denotes the covariant derivative compatible with g_{ij}. (Note that $\nabla_k \ln n$ is just $\partial \ln n / \partial x^k$, but $\nabla_k \nabla_j \ln n$ involves the Christoffel symbols, $\nabla_k \nabla_j \ln n = \partial^2 \ln n / \partial x^k \partial x^l - \Gamma^m_{jk} \partial \ln n / \partial x^m$.)

Among other things, the curvature determines the behavior of bundles of geodesics in the following way. Let $x^i = x^i(\lambda, s)$ be a family of geodesics parametrized by s, *i.e.*, for a fixed value of s, $x^i = x^i(\lambda, s)$ satisfies the geodesic equations (19), and let

$$\xi^i \equiv \frac{\partial x^i}{\partial s} \tag{39}$$

be a connecting vector between neighboring geodesics and $v^i \equiv \partial x^i / \partial \lambda$ be a tangent vector to the geodesics, then

$$\begin{aligned}v^j \frac{\partial \xi^i}{\partial x^j} &= \frac{\partial x^j}{\partial \lambda} \frac{\partial \xi}{\partial x^j} = \frac{\partial \xi^i}{\partial \lambda} = \frac{\partial^2 x^i}{\partial \lambda \partial s} = \frac{\partial^2 x^i}{\partial s \partial \lambda} \\ &= \frac{\partial v^i}{\partial s} = \frac{\partial x^j}{\partial s} \frac{\partial v^i}{\partial x^j} = \xi^j \frac{\partial v^i}{\partial x^j},\end{aligned}$$

which, owing to the symmetry $\Gamma^i_{jk} = \Gamma^i_{kj}$, is equivalent to

$$v^j \nabla_j \xi^i = \xi^j \nabla_j v^i, \tag{40}$$

therefore, making use of Eqs. (40) and (36) and the identity $\nabla_k \nabla_j v^i - \nabla_j \nabla_k v^i = R^i{}_{lkj} v^l$ [3–5] one obtains

$$\begin{aligned}v^k \nabla_k (v^j \nabla_j \xi^i) &= v^k \nabla_k (\xi^j \nabla_j v^i) \\ &= (v^k \nabla_k \xi^j) \nabla_j v^i + \xi^j v^k \nabla_k \nabla_j v^i \\ &= (\xi^k \nabla_k v^j) \nabla_j v^i + \xi^j v^k \nabla_j \nabla_k v^i \\ &\quad + \xi^j v^k (\nabla_k \nabla_j - \nabla_j \nabla_k) v^i \\ &= (\xi^k \nabla_k v^j) \nabla_j v^i + \xi^j \nabla_j (v^k \nabla_k v^i) \\ &\quad - \xi^j (\nabla_j v^k) \nabla_k v^i + \xi^j v^k R^i{}_{lkj} v^l \\ &= R^i{}_{lkj} v^l v^k \xi^j,\end{aligned} \tag{41}$$

which is known as the equation of geodesic deviation and gives the relative acceleration of neighboring geodesics.

Since the light rays are geodesics of the optical metric, from Eq. (41) we obtain

$$v^k \tilde{\nabla}_k (v^j \tilde{\nabla}_j \xi^i) = \tilde{R}^i{}_{lkj} v^l v^k \xi^j, \qquad (42)$$

which governs the deviation of the light rays measured according to the optical metric; however, in most cases, it may be more relevant to know the deviation determined by the background metric (usually the Euclidean metric). As shown in the following example, a bundle of light rays may be spreading according to the optical metric, but the distance between the rays may be constant according to the Euclidean metric.

We shall consider the case where $n = 1/z$ in the half-space $z > 0$ and g_{ij} corresponds to the metric of the three-dimensional Euclidean space in cartesian coordinates (hence, $g_{ij} = \delta_{ij}$). The light rays can be found by obtaining a complete solution of

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = \frac{1}{z^2}. \qquad (43)$$

Looking for a separable solution of Eq. (43) of the form $S = f(x) + g(y) + h(z)$, one finds that

$$S = \alpha_1 x + \alpha_2 y + \int \sqrt{\frac{1}{z^2} - \alpha_1^2 - \alpha_2^2}\, dz,$$

where α_1 and α_2 are separation constants. Then

$$\beta^1 = x - \alpha_1 \int \frac{z dz}{\sqrt{1 - (\alpha_1^2 + \alpha_2^2) z^2}} = x + \frac{\alpha_1 \sqrt{1 - (\alpha_1^2 + \alpha_2^2) z^2}}{\alpha_1^2 + \alpha_2^2}$$

and

$$\beta^2 = y + \frac{\alpha_2 \sqrt{1 - (\alpha_1^2 + \alpha_2^2) z^2}}{\alpha_1^2 + \alpha_2^2},$$

therefore

$$(x - \beta^1)^2 + (y - \beta^2)^2 + z^2 = \frac{1}{\alpha_1^2 + \alpha_2^2},$$

$$\alpha_2 (x - \beta^1) = \alpha_1 (y - \beta^2),$$

which means that each ray is a semi-circle whose plane is perpendicular to the xy plane and with center on the xy plane. Furthermore, the optical length along each of these rays does not depend on the radius of the semi-circle. Indeed, making use of the parametrization $x - \beta^1 = \alpha_1 (\alpha_1^2 + \alpha_2^2)^{-1} \sin\theta$, $y - \beta^2 = \alpha_2 (\alpha_1^2 + \alpha_2^2)^{-1} \sin\theta$, $z = (\alpha_1^2 + \alpha_2^2)^{-1/2} \cos\theta$, one finds that $\int n ds = \int \sec\theta d\theta$. Hence, if two light rays are parallel to each other at some point, they remain parallel to each other and the Euclidean distance between them is constant, while, according to the optical metric, the distance between these rays grows exponentially as they approach the xy plane. The curvature of the optical metric is different from zero; in fact, from Eq. (38) one finds that $\tilde{R}^i{}_{jkl} = -(\delta^i_k \tilde{g}_{jl} - \delta^i_l \tilde{g}_{jk})$, where $\tilde{g}_{ij} = \delta_{ij}/z^2$ (in cartesian coordinates). (The simplicity of the expression for the curvature of tensor $\tilde{R}^i{}_{jkl}$ means that $\tilde{g}_{ij}$ is the metric of a space of constant curvature -1, called the three-dimensional hyperbolic space.)

Thus, by contrast with the claim made in Ref. 10, the existence of a nonvanishing curvature for the optical metric does not imply focusing or defocusing of the light rays.

Another example is provided by the function $n = a/r^2$, where a is a constant and r is the distance to the origin in the three-dimensional Euclidean space. From Eq. (38) it follows that the scalar curvature $\tilde{R}$ is equal to zero. On the other hand, the light rays in this case are circles passing through the origin, excluding the origin itself; therefore, any pencil of light rays originally parallel, converges as it approaches to the origin.

Remark 8. Given a complete solution of the eikonal equation, $S(x^i, \alpha_a)$, and two arbitrary points A, B, with coordinates (x^i_A) and (x^i_B), respectively, then

$$V(x^i_B, x^i_A) \equiv S(x^i_B, \alpha_a) - S(x^i_A, \alpha_a) \qquad (44)$$

is the *characteristic function* of the medium [1,2], provided that the constants α_a are chosen in such a way that the ray determined by $S(x^i, \alpha_a)$ joins A and B. Geometrically, $V(x^i_B, x^i_A)$ corresponds to the optical length (*i.e.*, the distance measured by the optical metric) of the ray passing through A and B.

5. Conclusions

As we have shown, in the geometrical optics approximation, some aspects of the propagation of light can be conveniently expressed making use of the optical metric and the language of differential geometry. This approach enables us to derive some properties of an optical system from general results of differential geometry, which are also applicable in classical mechanics (not to mention general relativity and some theories of unification). For instance, each Killing vector (which generates isometries), or Killing tensor, of the optical metric gives rise to a constant of the motion, which is useful to find the light rays.

Acknowledgments

The authors acknowledge one of the referees for helpful commments. One of the authors (C.J.P.B.) acknowledges the Consejo Nacional de Ciencia y Tecnología for financial support through a doctoral fellowship.

1. M. Born and E. Wolf, *Principles of Optics*, 6th edition, (Cambridge University Press, Cambridge, 1997) Chap. III.
2. H.A. Buchdahl, *An Introduction to Hamiltonian Optics*, (Dover, New York, 1993) Chap. 1.
3. J.A. Schouten, *Ricci Calculus*, 2nd edition, (Springer, Berlin, 1954).
4. L. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th edition, (Pergamon, Oxford, 1975).
5. F. de Felice and C.J.S. Clarke, *Relativity on curved manifolds*, (Cambridge University Press, Cambridge, 1990).
6. C. Lanczos, *The Variational Principles of Mechanics*, 4th edition, (Dover, New York, 1986) Chap. VIII.
7. H. Goldstein, *Classical Mechanics*, 2nd edition, (Addison-Wesley, Reading, Mass., 1980).
8. G.F. Torres del Castillo, *Rev. Mex. Fís.* **35** (1989) 301.
9. G. Krötzch and K.B. Wolf, *Rev. Mex. Fís.* **37** (1991) 136.
10. H. Guo and X. Deng, *J. Opt. Soc. Am. A* **12** (1995) 600.