The optical–mechanical analogy in general relativity: New methods for the paths of light and of the planets

James Evans[a)]
Department of Physics, University of Puget Sound, Tacoma, Washington 98416

Kamal K. Nandi
Department of Mathematics, University of North Bengal, Darjeeling (W.B.) 734 430, India

Anwarul Islam
Department of Mathematics, Government Tolaram College, Narayanganj, Dhaka, Bangladesh

(Received 10 October 1994; accepted 5 April 1996)

The optical–mechanical analogy involves the expression of geometrical optics and particle mechanics in the same mathematical language. In this paper, an especially simple version of the optical–mechanical analogy is extended to general relativity. A variational principle for the trajectories of photons and particles is obtained which is applicable to a broad class of metrics. This permits us to cast the exact equations of motion for both massive and massless particles into the form of Newtonian mechanics. The new equations of motion are illustrated by applications to the Schwarzschild metric.

I. INTRODUCTION

Geometrical optics may be cast into the form of Newtonian mechanics, with substantial advantages for practical calculation.[1,2] The shape of the optical ray is governed by a differential equation formally identical to Newton's law of motion (acceleration= −gradient of potential energy):

$$\frac{d^2\mathbf{r}}{dA^2} = \nabla\left(\frac{c_0^2 n^2}{2}\right), \tag{1}$$

where $\mathbf{r}$ is the position of a light pulse moving along the ray, $n(\mathbf{r})$ is the index of refraction, and c_0 is the vacuum speed of light. The independent variable A is the stepping parameter, or optical action. In this formulation of geometrical optics, A is analogous to the time t in mechanics. (A is related to t by $dA=dt/n^2$.) The "potential energy function" is $-c_0^2n^2/2$. All the usual force and energy methods of elementary mechanics can be brought to bear on geometrical optics.

Now, a gravitational field can be represented as an optical medium with an effective index of refraction. This idea was exploited long ago by Eddington[3] and has been developed in more detail by other writers.[4,5] In the most general metric, the effective medium has rather complicated properties. But many metrics that are widely studied for physical or cosmological applications are reasonably simple. Take for example the Schwarzschild metric, which describes the spacetime structure around a stationary spherical object such as a star or black hole. With a suitable coordinate transformation, the space part of the Schwarzschild line element can be rendered isotropic. This makes it possible to identify an isotropic coordinate speed of light and hence an effective index of refraction. This index of refraction may be used without modification in any formulation of geometrical optics that is convenient—including the "$F=ma$" formulation based on Eq. (1).[6] That is, the calculation of light orbits in the Schwarzschild metric can be handled, exactly, by the methods of Newtonian mechanics. Can the same methods be applied to the motion of massive particles?

This paper supplies an affirmative answer to the question.[7] Starting from the geodesic property of the particle orbits, we derive a variational principle that can be considered analogous to Fermat's principle or to Maupertuis's principle. The analogy between these principles allows us to express the problem of planetary motion in general relativity in terms of a formalism identical to Newtonian mechanics. The new equations of motion are of Newtonian form, cover both light and massive particles, and are exact. The use of the new equations will be illustrated by examples and their physical meaning discussed in detail.

II. THE GRAVITATIONAL FIELD AS AN OPTICAL MEDIUM

The methods to be studied here are applicable to static, isotropic metrics. In such a case the line element can be written in the form

$$\begin{aligned} ds^2 &= \Omega^2(\mathbf{r})c_0^2\,dt^2-\Phi^{-2}(\mathbf{r})[dr^2+r^2\,d\theta^2 \\ &\quad +r^2\sin^2\theta\,d\phi^2] \\ &= \Omega^2(\mathbf{r})c_0^2\,dt^2-\Phi^{-2}(\mathbf{r})|d\mathbf{r}|^2, \end{aligned} \tag{2}$$

where Ω and Φ are functions of the spatial coordinates. The isotropic coordinate speed of light[8] $c(\mathbf{r})$ may be obtained by putting $ds=0$:

$$c(\mathbf{r})=\left|\frac{d\mathbf{r}}{dt}\right|=c_0\Phi(\mathbf{r})\Omega(\mathbf{r}). \tag{3}$$

Thus the effective index of refraction n (defined by $n=c_0/c$) is

$$n(\mathbf{r})=\Phi^{-1}\Omega^{-1}. \tag{4}$$

Many line elements of physical and cosmological significance can be put into the form of Eq. (2). Throughout this paper we shall use the Schwarzschild metric as our example because of its familiarity and its importance for solar-system dynamics. In standard coordinates (r',θ,ϕ,t), the Schwarzschild line element is

$$\begin{aligned} ds^2 &= \left(1-\frac{2m}{r'}\right)c_0^2\,dt^2-\left(1-\frac{2m}{r'}\right)^{-1}dr'^2 \\ &\quad -r'^2(d\theta^2+\sin^2\theta\,d\phi^2), \end{aligned} \tag{5}$$

where $m=GM/c_0^2$, G is the universal gravitation constant, c_0 is the speed of light in vacuum, and M is the mass of the central gravitating body. Define a new radial coordinate r by[9]

$$r'=r\Phi^{-1}(r)=r\left(1+\frac{m}{2r}\right)^2. \tag{6}$$

The line element can then be written in the form of Eq. (2) using the so-called isotropic coordinates (r,θ,ϕ,t), with

$$\Omega(r)=\left(1+\frac{m}{2r}\right)^{-1}\left(1-\frac{m}{2r}\right). \tag{7}$$

Thus the index of refraction turns out to be

$$n(r)=\left(1+\frac{m}{2r}\right)^3\left(1-\frac{m}{2r}\right)^{-1}, \tag{8}$$

an expression that is valid for $r>m/2$. As shown in Ref. 6, this index of refraction can be used with the classical optical formalism based on Eq. (1) to derive the usual equations for light orbits in the Schwarzschild metric. The goal of the present paper is to extend the treatment to massive particles.

III. DERIVATION OF THE VARIATIONAL PRINCIPLE

The particle orbits may be obtained by requiring that they be geodesics:

$$\delta\int_{\mathbf{x}_1,t_1}^{\mathbf{x}_2,t_2} ds=0, \tag{9}$$

where the symbol δ indicates a variation in the path of integration between two fixed points in spacetime, $(\mathbf{x}_1,t_1)$ and $(\mathbf{x}_2,t_2)$. We restrict ourselves to static metrics that can be put into isotropic form. Then, by use of Eqs. (2) and (4), the geodesic condition may be written

$$\delta\int_{\mathbf{x}_1,t_1}^{\mathbf{x}_2,t_2} \Omega c_0\left[1-\frac{v^2n^2}{c_0^2}\right]^{1/2} dt=0, \tag{10}$$

a form analogous to Hamilton's principle. Effective Lagrangians are standard tools in general relativity. In this case the Lagrangian is

$$L(x_i,\dot{x}_i)=-c_0^2\Omega\left[1-\frac{v^2n^2}{c_0^2}\right]^{1/2}, \tag{11}$$

where Ω and n are functions of the coordinates alone, where $\dot{x}_i\equiv dx_i/dt$, and where $v^2=\Sigma_{i=1}^3(dx_i/dt)^2$, if we choose to work in Cartesian coordinates. The expression for the Lagrangian has been multiplied by an extra factor of $-c_0$ for later convenience.

Form the Hamiltonian in the usual way:

$$H=\sum_{i=1}^{3} p_i\dot{x}_i-L, \tag{12}$$

where the canonical momenta p_i are defined by

$$p_i \equiv \frac{\partial L}{\partial \dot{x}_i}. \tag{13}$$

With the Lagrangian from Eq. (11), this results in

$$p_i = \Omega n^2[1 - v^2 n^2/c_0^2]^{-1/2} \dot{x}_i \tag{14}$$

and

$$H = c_0^2 \Omega [1 - v^2 n^2/c_0^2]^{-1/2}, \tag{15}$$

or, if H is expressed in terms of the canonical momenta,

$$H = c_0^2[\Omega^2 + p^2/n^2 c_0^2]^{1/2}. \tag{16}$$

Because $\partial L/\partial t = 0$, H is a constant of the motion.

Now, starting from Hamilton's principle,

$$\delta \int_{\mathbf{x}_1, t_1}^{\mathbf{x}_2, t_2} L\, dt = 0, \tag{17}$$

one may derive in the usual way the corresponding action principle (Jacobi's form of Maupertuis's principle):[10]

$$\delta \int_{\mathbf{x}_1}^{\mathbf{x}_2} \left(\sum_i p_i \dot{x}_i \right) dt = 0, \tag{18}$$

where now the path of integration is varied, subject to conservation of energy, between two fixed points in space, $\mathbf{x}_1$ and $\mathbf{x}_2$, but the times at the end points need not be held fixed. With the canonical momenta from Eq. (14), Eq. (18) becomes

$$\delta \int_{\mathbf{x}_1}^{\mathbf{x}_2} n^2 v^2 \Omega \left[1 - \frac{v^2 n^2}{c_0^2} \right]^{-1/2} dt = 0.$$

We constrain the varied paths to those that satisfy energy conservation by substituting H from Eq. (15), with the result

$$\delta \int_{\mathbf{x}_1}^{\mathbf{x}_2} n^2 v^2\, dt = 0.$$

Now, t is not really a suitable variable for parametrizing the path, since the values of t at the end points are not fixed. Let us instead integrate (symbolically) over the arc length dl $(=|d\mathbf{r}| = (\Sigma_{i=1}^3 dx_i^2)^{1/2})$ from one fixed space point, $\mathbf{x}_1$, to another, $\mathbf{x}_2$. Since $dt = dl/v$, we have immediately

$$\delta \int_{\mathbf{x}_1}^{\mathbf{x}_2} n^2 v\, dl = 0. \tag{19}$$

This is a variational principle on which an analogy to geometrical optics or to classical mechanics can be constructed. Eq. (19) is the foundation upon which the remainder of this paper is based.[11]

IV. REMARKS ON THE VARIATIONAL PRINCIPLE

In Eq. (19), $n^2 v$ is to be considered a function of position alone (in the isotropic version of the Schwarzschild metric, a function of r only). The path of integration is varied between the fixed end points $\mathbf{x}_1$ and $\mathbf{x}_2$, and the value of H is held constant during the variation. The path that minimizes the integral will be the path taken by a massive particle in the Schwarzschild field. Thus, Eq. (19) is of exactly the same form as Fermat's principle, which forms a basis for geometrical optics, and Maupertuis's principle, which can be taken as a basis for classical mechanics (as long as the force can be derived from a velocity-independent potential):

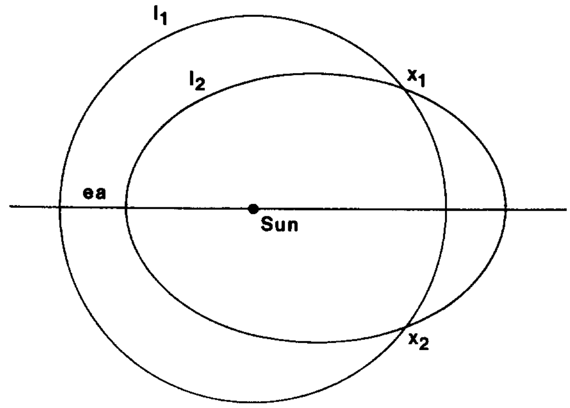

Fig. 1. A circular planetary orbit and an elliptical orbit of the same energy, considered as a variation of the path of integration.

optics and mechanics in GR	geometrical optics (Fermat)	classical mechanics (Maupertuis)
$\delta \int n^2 v\, dl = 0,$	$\delta \int n\, dl = 0,$	$\delta \int v\, dl = 0.$

In the case of light orbits in general relativity, since we may put $v = c_0/n$, Eq. (19) actually reduces to Fermat's principle.[12] Let us note also that Maupertuis's principle (and hence ordinary Newtonian gravitational motion) arises as a limiting case of Eq. (19). Begin by writing

$$\delta \int n^2 v\, dl = \int (\delta n^2) v\, dl + \int n^2 (\delta v)\, dl + \int n^2 v (\delta\, dl).$$

It is easy to convince oneself that the first term on the right is vastly smaller than the second and third terms. Thus, in ordinary solar-system dynamics, n^2 may be treated as approximately constant in the variational calculation and we obtain Maupertuis's principle as the classical limit of Eq. (19).

To see this, consider a circular orbit and a superimposed elliptical orbit of the same energy, and hence of the same semi-major axis a. (See Fig. 1) The Sun is at the center of the circular orbit and at one focus of the ellipse. The perihelion of the ellipse lies inside the circle by a distance ea, where e is the eccentricity of the ellipse. We may consider path l_2 as a variation from path l_1 in the variation of the integral of Eq. (19).

Let v_0 denote the speed of the particle on the circular orbit and δv the maximum variation from v_0, which occurs at the perihelion of the varied path. $\delta v/v_0$ is of order e. It is clear that path l_2 is shorter than path l_1 by something of order ea. Hence also $(\delta\, dl)/dl$ is of order e. Now, from Eq. (8),

$$n^2 \approx 1 + 4m/r,$$

from which we may calculate

$$\delta n^2 \approx \left(\frac{dn^2}{dr} \right) dr \approx \left(\frac{dn^2}{dr} \right) ea.$$

Thus

$$\frac{\delta n^2}{n^2} \approx \frac{e4m}{a},$$

which is smaller than e by the factor $4m/a = 4(v_0/c_0)^2$, roughly 4×10^{-8} for the orbit of the Earth. Thus, in most solar-system mechanics, Maupertuis's principle $\delta\int v\, dl = 0$ is a perfectly satisfactory approximation to $\delta\int n^2 v\, dl = 0$.

V. EQUATION OF MOTION

Let us suppose that some portion of the path can be parametrized by one of the Cartesian coordinates, say z. Then we may write Eq. (19) in the form

$$\delta \int_{z_1}^{z_2} n^2 v \left|\frac{d\mathbf{r}}{dz}\right| dz = 0, \tag{20}$$

where

$$\left|\frac{d\mathbf{r}}{dz}\right| = \left(\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right)^{1/2} \tag{21}$$

and where the path must begin at z_1 and end at z_2. The Euler conditions that the integral be stationary are of the form

$$\left|\frac{d\mathbf{r}}{dz}\right| \frac{\partial}{\partial x}(n^2 v) = \frac{d}{dz}\left[n^2 v \left|\frac{d\mathbf{r}}{dz}\right|^{-1} \frac{dx}{dz}\right]. \tag{22}$$

This is the equation of motion for the x coordinate. There is a similar equation of motion for the y coordinate.

Now let us parametrize the path by some parameter A other than z. We need not commit ourselves yet to a definition of A. Instead we shall choose A to give the simplest equation of motion. To pass over from z to A as independent variable, we multiply both sides of Eq. (22) by dz/dA. We also multiply and divide by dz/dA inside the square bracket on the right side of the equation. The result is

$$\left|\frac{d\mathbf{r}}{dA}\right| \frac{\partial}{\partial x}(n^2 v) = \frac{d}{dA}\left(n^2 v \left|\frac{d\mathbf{r}}{dA}\right|^{-1} \frac{dx}{dA}\right). \tag{23}$$

There is nothing special about the x coordinate, so similar equations of motion must apply to y and to z. Thus, in vector form, the equation of motion is

$$\left|\frac{d\mathbf{r}}{dA}\right| \nabla(n^2 v) = \frac{d}{dA}\left(n^2 v \left|\frac{d\mathbf{r}}{dA}\right|^{-1} \frac{d\mathbf{r}}{dA}\right). \tag{24}$$

To give the equation of motion the simplest possible form, and to take advantage of the analogy to Newtonian mechanics, let us now define A by

$$\left|\frac{d\mathbf{r}}{dA}\right| \equiv n^2 v. \tag{25}$$

To find the resulting connection between A and t, note that

$$\left|\frac{d\mathbf{r}}{dA}\right| = \left|\frac{d\mathbf{r}}{dt}\right| \frac{dt}{dA} = v\, \frac{dt}{dA}.$$

Making use of Eq. (25), we get

$$dA = \frac{dt}{n^2}. \tag{26}$$

Thus the stepping parameter is the same as that used in the $F=ma$ formulation of geometrical optics: A is the *optical action*.[13]

With the definition of A given by Eq. (25), the equation of motion [Eq. (24)] becomes

$$\frac{d^2\mathbf{r}}{dA^2} = \nabla\left(\frac{1}{2} n^4 v^2\right). \tag{27}$$

The left-hand side of Eq. (27) is of the form of an acceleration: it is the second derivative of the position vector with respect to the independent variable. The right-hand side of the equation is of the form of a force: it is the negative gradient of a "potential energy function:"[14]

$$\text{"potential energy"} = -\tfrac{1}{2} n^4 v^2. \tag{28}$$

The analog of the velocity is $d\mathbf{r}/dA$. Thus

$$\text{"kinetic energy"} = \frac{1}{2}\left|\frac{d\mathbf{r}}{dA}\right|^2. \tag{29}$$

The analogue of the "total energy" is the sum of the potential and the kinetic. But, by virtue of Eq. (25), these two are guaranteed to sum to zero:

$$\text{"total energy"} = \frac{1}{2}\left|\frac{d\mathbf{r}}{dA}\right|^2 - \frac{1}{2} n^4 v^2 = 0. \tag{30}$$

Thus the calculation of the paths of light and of the planets reduces to the usual zero-energy $F=ma$ optics.

VI. SUMMARY AND DISCUSSION OF RESULTS

In Secs. III and V, we have provided a reformulation of the trajectory problem for both massive and massless particles in general relativity, using the classical optical–mechanical analogy as our model. This formulation is exact. Thus the equation of motion (27), although of very simple form, is equivalent to any other formulation of the geodesic equations of motion in static metrics of suitable type. The Newtonian form of Eq. (27) should perhaps be thought of as coming from $F=ma$ optics (which is, after all, exact) and not from Newtonian mechanics (which is, of course, only approximate).

It is also easy to see that Eq. (27) reduces to the correct Newtonian limit. We need only put $n \approx 1$ (so that $dA \approx dt$) and put $v^2/2 \approx E - U$ (the difference between the total and the potential energy per unit mass). Then Eq. (27) reduces to Newton's law of motion, $d^2\mathbf{r}/dt^2 = -\nabla U$.

Although the examples we shall explore below are based on the Schwarzschild metric, these methods have a reasonably high degree of generality. None of the principal equations produced above depends on the *detailed* form of the metric. The method is applicable as long as (1) the metric is static, so that the effective Lagrangian does not depend explicitly on the time, and (2) a coordinate transformation can be found that renders the space part of the metric isotropic, so that the line element can be written in the form of Eq. (2). There are many other metrics of physical and cosmological interest that satisfy these conditions.[15]

In problem solving, one may attack the motion of particles and the path of light in exactly the same way. Both the motion of particles and the path of light in the gravitational field are described by the same equation of motion, Eq. (27). In solving problems with Eq. (27), one may use without modification all the force methods that are familiar from Newtonian mechanics. (One simply thinks of A as if it were the time.) The other important problem-solving technique of classical mechanics is based on the conservation of energy. Here, the analogous equation is Eq. (30), which applies, again, for both light and particles.

Moreover, one need not always begin from the fundamental equations (27) and (30). Indeed, one may begin from any valid classical–mechanical formula then simply make the transcriptions

$$t \to A, \quad U \to -n^4 v^2/2, \quad E \to 0, \tag{31}$$

and a correct general-relativistic formula (valid in the isotropic coordinates) will be obtained.

For both light and particles, the stepping parameter A is defined by Eq. (25). In many calculations, e.g., in finding the shape of an orbit, the stepping parameter is ultimately eliminated. Equation (25) will suffice for this purpose. In other situations, it may be necessary to supply an explicit connection between A and t. Equation (26) provides the necessary connection—again, for both light and particles.[16] After the equations governing the situation are obtained, or after they are solved, one may transform back to the original metric, if desired.

The only difference between the treatment of light and that of particles resides in the choice of $v(\mathbf{r})$:

$$v = c_0 n^{-1} \quad \text{(light)}, \tag{32}$$

$$v = c_0 n^{-1}[1 - c_0^4 \Omega^2/H^2]^{1/2} \quad \text{(particles)}. \tag{33}$$

Equation (33) is obtained by solving Eq. (15) for v. H is a constant of the motion. Ω and n are determined by the metric—by Eqs. (7) and (8) in the case of the Schwarzschild metric.

Because the particle expression for $v(\mathbf{r})$ contains the constant of the motion H, the particle problem has an extra degree of freedom: We may specify the initial speed of the particle. Thus, in general, more types of orbits exist for particles than for light in the same metric. Let us evaluate H for two common situations. First, for a classical-limit planetary orbit, $v/c_0 \ll 1$ and $m/r \ll 1$. Thus $\Omega \approx 1 - m/r$ and Eq. (15) becomes

$$H \approx c_0^2 + \tfrac{1}{2} v^2 - m/r. \tag{34}$$

Here, H is approximately equal to the rest-mass energy plus the classical kinetic and potential energy per unit mass. Consider now a particle in space devoid of gravitational influences. Then $\Omega = 1$, $n = 1$, and Eq. (15) becomes

$$H = c_0^2 (1 - v^2/c_0^2)^{-1/2} = c_0^2 \gamma, \tag{35}$$

the relativistic energy of a free particle.

VII. EXAMPLES OF FORCE METHODS

A. Deflection of starlight or an ultra-relativistic particle by the Sun

Let a ray of light or an ultra-relativistic ($v \approx c_0$) particle graze the Sun, as in Fig. 2. Let the original line of motion be parallel to the x axis, as shown. We wish to calculate the angular deflection, assumed small. The fundamental equation for force methods is Eq. (27), which we rewrite in the following form:

$$\frac{d\mathbf{p}}{dA} = \nabla\left(\frac{1}{2} n^4 v^2\right), \tag{36}$$

where the "momentum" $\mathbf{p}$ is defined[17] by

$$\mathbf{p} \equiv \frac{d\mathbf{r}}{dA}. \tag{37}$$

Since the "potential energy" is a function of the radial coordinate r alone, the "force" points in the radial direction, and the y component of the equation of motion is

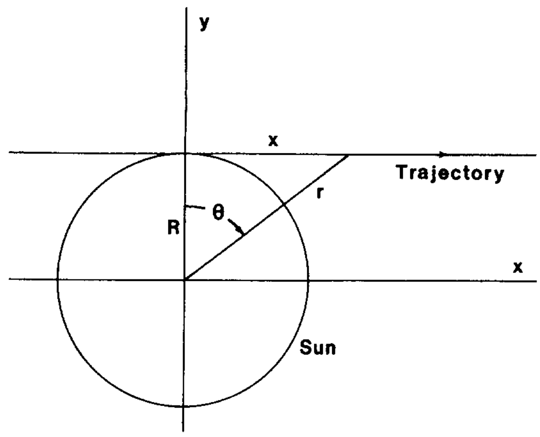

Fig. 2. A ray of light, or the trajectory of an ultra-relativistic particle, grazing the sun.

$$\frac{dp_y}{dA} = \cos\theta \frac{d}{dr}\left(\frac{1}{2} n^4 v^2\right) = \cos\theta\, n^2 v \frac{d}{dr}(n^2 v). \tag{38}$$

The change in p_y as the particle or photon travels from $x = -\infty$ to $x = \infty$ is obtained by integrating Eq. (38):

$$\Delta p_y = \int \cos\theta\, n^2 v \frac{d}{dr}(n^2 v)\, dA. \tag{39}$$

The classical analog of Eq. (39) is the impulse-momentum theorem.

Because the deflection will be very small, $\mathbf{p}$ will point almost entirely in the x direction during the whole course of the motion. That is $|\mathbf{p}| \approx p_x$. Moreover, p_x will undergo but a tiny fractional change and can, in fact, be taken as constant. Thus, with excellent accuracy, the angular deflection β is given by

$$\beta = \frac{\Delta p_y}{p_x}. \tag{40}$$

β is the angle between the incoming and the outgoing asymptotes. The strategy we shall use in evaluating β is then simply to calculate the y component of the impulse accumulated over the whole trajectory.

To carry out the integration, it is simplest to put

$$dA = \frac{dA}{dx}\frac{dx}{d\theta}\, d\theta \tag{41}$$

and to integrate from $\theta = -\pi/2$ to $\pi/2$.

Now, dx/dA is simply p_x. Moreover, since $|\mathbf{p}| \equiv n^2 v$ and since, as already argued, $|\mathbf{p}| \approx p_x$, we may put $dx/dA \approx n^2 v$. From the geometry of Fig. 1,

$$\frac{dx}{d\theta} = \frac{R}{\cos^2\theta}. \tag{42}$$

Putting all of this into Eq. (39) we have

$$\Delta p_y = \int_{-\pi/2}^{\pi/2} \frac{R}{\cos\theta} \frac{d}{dr}(n^2 v)\, d\theta. \tag{43}$$

So far, Eq. (43) applies to any spherically symmetric index of refraction. And, of course, it is equally valid for massive and massless particles. The only approximation made so far is that of small deflection, which justifies the impulse approximation.

Now, let us specialize to the Schwarzschild metric. In the weak field of the Sun, Ω and n given by Eq. (7) and (8) may be approximated by

$$n \approx 1 + 2m/r, \tag{44}$$

$$\Omega \approx 1 - m/r. \tag{45}$$

Then, for light, with the use of Eq. (32), we have

$$\frac{d}{dr}(n^2 v) = \frac{-2mc_0}{r^2} \quad \text{(light)}. \tag{46}$$

For a massive particle, note that, from Eq. (33),

$$n^2 v = c_0 n[1 - c_0^4 \Omega^2 / H^2]^{1/2} \quad \text{(particle)}. \tag{47}$$

Also, for the ultra-relativistic particle (i.e., for the case of small deflection), Eq. (35) gives

$$H \approx c_0^2 (1 - v_0^2/c_0^2)^{-1/2} = c_0^2 \gamma_0, \tag{48}$$

where v_0 is the particle's speed when it is infinitely far from the Sun. Substituting Eqs. (44), (45), and (48) into Eq. (47), differentiating and expanding to first order in m/r, we obtain

$$\frac{d}{dr}(n^2 v) = \frac{-mc_0}{r^2} \frac{c_0}{v_0} \left(1 + \frac{v_0^2}{c_0^2}\right) \quad \text{(particle)}, \tag{49}$$

which differs from Eq. (46) only in the constant coefficient multiplying mc_0/r^2.

Let us first calculate the deflection of light. Putting Eq. (46) into Eq. (43) and noting from Fig. 2 that

$$1/r^2 = \cos^2\theta / R^2, \tag{50}$$

we obtain for the change in the y "momentum"

$$\Delta p_y = \frac{-2mc_0}{R} \int_{-\pi/2}^{\pi/2} \cos\theta \, d\theta = \frac{-4mc_0}{R} \quad \text{(light)}. \tag{51}$$

The angular β deflection is given by Eq. (40). Since $p = nc_0$ never differs appreciably from its value at $x = \infty$, and since the motion is almost entirely in the x direction, $p_x \approx c_0$. Thus we obtain the usual expression for the deflection of light:

$$\beta = \frac{4m}{R} \quad \text{(light)}, \tag{52}$$

where we have dropped the minus sign.

For the massive particle the calculation goes in exactly the same way since Eq. (46) and Eq. (49) differ only by a constant multiplicative factor. Thus

$$\Delta p_y = \frac{-2mc_0}{R} \frac{c_0}{v_0} \left(1 + \frac{v_0^2}{c_0^2}\right) \quad \text{(particle)}. \tag{53}$$

Since $p = n^2 v$ never differs appreciably from its value at infinity and since the motion is almost entirely in the x direction, we may put $p_x = v_0$. Thus the angular deflection is[18]

$$\beta = \frac{2m}{R} \left(1 + \frac{c_0^2}{v_0^2}\right). \quad \text{(particle)} \tag{54}$$

For $v_0 < c_0$, the deflection (54) of the particle is greater than the deflection (52) of light. But as $v_0 \to c_0$, the deflection of the particle becomes equal to that of light.

It is sometimes said that Newtonian mechanics implies a deflection of starlight by the Sun, since the deflection of a particle infinitely less massive than the Sun does not depend on the mass of the particle, and we could think of the photon as a Newtonian particle in the limit of very small mass. We then need only put $v_0 = c_0$ and calculate the deflection using Newtonian principles. As is well known, the result is just one half of that given in Eq. (52).[19] We see here that in general relativity, too, the deflections of light and of a material particle coming in at approximately the speed of light are the same (both being twice the Newtonian result).

B. Circular motion

In the preceding example, because we were interested only in the asymptotic behavior of the trajectory, there was no need to convert the final answer from the isotropic radial coordinate r to the original radial coordinate r'. Uniform circular motion provides an example that can be handled exactly, but that does require us to convert back to the original coordinate system to interpret our answers.

Let n and Ω be functions of the radial coordinate alone. The dynamical condition for uniform circular motion may then be written down by analogy to the classical–mechanical rule $-dU/dr = -v^2/r$. That is, we require

$$\frac{d}{dr}\left(\frac{1}{2} n^4 v^2\right) = \frac{-1}{r} \left|\frac{d\mathbf{r}}{dA}\right|^2. \tag{55}$$

Using Eq. (25) on the right side, we obtain

$$\frac{d}{dr}\left(\frac{1}{2} n^4 v^2\right) = \frac{-n^4 v^2}{r}. \tag{56}$$

This differential equation is the exact general-relativistic condition for uniform circular motion in a central-force situation. It applies equally to massive particles and to light.

For the special case of circular light orbits, we invoke Eq. (32) and then Eq. (56) becomes

$$\frac{dn}{dr} = -\frac{n}{r}. \tag{57}$$

There are two interesting questions we could ask at this stage. First we might ask which function $n(r)$ will permit a circular light orbit at any value of the radius r. A simple integration shows that n must be proportional to r^{-1}.

A second question we might pose for solution using Eq. (57) is this: In the Schwarzschild metric, for what value of r will a circular orbit exist? We substitute the Schwarzschild form of $n(r)$ [Eq. (8)] into Eq. (57) and solve for r with the result

$$r = m(1 \pm \sqrt{3}/2). \tag{58}$$

We must select the upper sign, since our index of refraction is defined only for $r > m/2$. Finally, it is useful to transform back to the original coordinate system. Substituting $r = m(1 + \sqrt{3}/2)$ into Eq. (6) and solving for r', we find

$$r' = 3m = 3MG/c_0^2, \tag{59}$$

the usual result. Thus only one circular light orbit exists. It is worth remarking that this calculation has been an exact one.

The existence of circular massive-particle orbits in the Schwarzschild field can be investigated in the same way, by starting from Eqs. (56) and (33), together with the special forms of Eqs. (7) and (8) for the Schwarzschild metric.

Note that we wrote down the exact general-relativistic condition for uniform circular motion [Eq. (55)] by analogy to classical mechanics, using the transcriptions (31) as a mnemonic device. But the exact relativistic condition also contains the classical–mechanical condition as a limiting case. We can recover the classical–mechanical rule by letting $n\to 1$ in Eq. (55): $d(\frac{1}{2}v^2)/dr=-v^2/r$.

VIII. AN EXAMPLE OF ENERGY METHODS

The preceding examples involved the geodesic motion of light and particles in the Schwarzschild field, but both were devoted to special cases. As our final example, let us derive the general equation of motion for light and particles. Although we could continue with force methods based upon Eq. (27), we choose instead to illustrate energy methods by using Eq. (30), which plays the role of a statement of conservation of energy.

If we write out the "kinetic energy" in polar coordinates, Eq. (30) becomes

$$\left(\frac{dr}{dA}\right)^2+r^2\left(\frac{d\theta}{dA}\right)^2-n^4v^2=0. \tag{60}$$

The fact that the "potential energy" is a function of the radial coordinate alone leads (just as in classical central-force motion) to a conserved "angular momentum":

$$h=r^2\frac{d\theta}{dA}=\text{constant}. \tag{61}$$

[This may be seen by writing out the θ component of Eq. (27).] We use Eq. (61) to pass over from A to θ as independent variable in Eq. (60), obtaining

$$r^{-4}\left(\frac{dr}{d\theta}\right)^2+r^{-2}-h^{-2}n^4v^2=0. \tag{62}$$

In the usual way, define

$$u=r^{-1}. \tag{63}$$

Then $dr/d\theta=-u^{-2}\,du/d\theta$ and Eq. (62) may be written

$$\left(\frac{du}{d\theta}\right)^2+u^2-n^4v^2/h^2=0. \tag{64}$$

The classical analog of Eq. (64), for the motion of a massive particle in a central potential $U(r)$, is

$$\left(\frac{du}{d\theta}\right)^2+u^2-\frac{2(E-U)}{h_0^2}=0,$$

where h_0 is the classical–mechanical angular momentum per unit mass, $r^2\dot{\theta}$. Note that we could have simply written down the exact general-relativistic formula [Eq. (64)] by analogy to the classical formula and the use of the transcriptions (31).

Equation (64) still applies both to planets and to light. To specialize to the path of a planet, we invoke Eq. (33), so that Eq. (64) becomes

$$\left(\frac{du}{d\theta}\right)^2+u^2-n^2c_0^2h^{-2}\left[1-\frac{c_0^4\Omega^2}{H^2}\right]=0 \quad \text{(planet)}. \tag{65}$$

To calculate the path of light, we use Eq. (32). Thus Eq. (64) becomes

$$\left(\frac{du}{d\theta}\right)^2+u^2-n^2c_0^2h^{-2}=0 \quad \text{(light)}. \tag{66}$$

The only fact about the metric we have used so far is its spherical symmetry. (That is, n^4v^2 is a function of r alone.) Let us now focus on the Schwarzschild problem: Φ, Ω, and n are given by Eq. (6), (7), and (8). To return to the original (nonisotropic) metric, we need to invert the coordinate transformation given by Eq. (6):

$$u=u'/\Phi, \tag{67}$$

where $u'=1/r'$ and r' is the standard radial coordinate. It is not hard to show that

$$\frac{du}{du'}=\Phi^{-1}\Omega^{-1}=n. \tag{68}$$

Also, it will eventually be helpful to have explicit forms for Φ, Ω and n as functions of u' rather than u. A little algebra gives

$$\Phi=\tfrac{1}{4}[1+(1-2mu')^{1/2}]^2, \tag{69}$$

$$\Omega=(1-2mu')^{1/2}, \tag{70}$$

$$n=4(1-2mu')^{-1/2}[1+(1-2mu')^{1/2}]^{-2}. \tag{71}$$

We do not need all of these relations for the present calculation, but it is convenient to group them in one place.

With the use of Eqs. (67) and (68), Eq. (65) becomes

$$\left(\frac{du'}{d\theta}\right)^2+\Omega^2u'^2-c_0^2h^{-2}\left[1-\frac{c_0^4\Omega^2}{H^2}\right]=0 \quad \text{(planet)}. \tag{72}$$

Only now do we need the explicit form of Eq. (70) for Ω. Equation (72) becomes

$$\left(\frac{du'}{d\theta}\right)^2-2mc_0^6h^{-2}H^{-2}u'+u'^2-2mu'^3+c_0^6h^{-2}H^{-2}-c_0^2h^{-2}=0.$$

Differentiating with respect to θ, we get

$$\frac{d^2u'}{d\theta^2}+u'-3mu'^2-mc_0^6h^{-2}H^{-2}=0. \tag{73}$$

Equation (73) is exact and may be handled as is. However, in solar-system applications, we will introduce no appreciable error if we approximate the constants of the motion H and h by their classical limits. Note that the constant term in Eq. (73) is already of first order in m. Thus, by virtue of Eq. (34), we may put $H\approx c_0^2$, and ignore terms of order $(m/r)^2$ (weak field). The other constant of the motion is $h=r^2(d\theta/dA)$. With use of Eq. (26), this may be written

$$h=\frac{r^2n^2\,d\theta}{dt}=n^2h_0,$$

where $h_0=r^2\,d\theta/dt$ is the angular momentum per unit mass from classical mechanics. Because the constant term in Eq. (73) is already first order in m, we may replace n by unity. Thus, Eq. (73) becomes

$$\frac{d^2u'}{d\theta^2}+u'=\frac{mc_0^2}{h_0^2}+3mu'^2 \quad \text{(planet)}. \tag{74}$$

This is the usual equation obtained in general relativity for a precessing elliptical orbit. The perihelion advance per revolution on the orbit is[20] $6\pi m^2c_0^2/h_0^2=6\pi G^2M^2/(h_0c_0)^2$. No approximations have been made in deriving Eq. (74), except, of course, in evaluating the constants of the motion H and h.

The usual light-orbit equation for the Schwarzschild metric is obtained in a similar fashion, by using Eqs. (67), (68), and (70) in Eq. (66), with the result

$$\frac{d^2u'}{d\theta^2}+u'=3mu'^2. \quad \text{(light)}. \tag{75}$$

We could also have written this down immediately after having solved the particle case, by noting that for light we let H approach ∞ [see Eq. (15)]. Thus, Eq. (74) reduces to Eq. (75).

IX. ON THE PHYSICAL CONTENT OF THE OPTICAL–MECHANICAL ANALOGY

Equation (27)—with A defined by Eq. (25)—covers a lot of physics: It applies to the geometrical optics of isotropic media, to classical mechanics in velocity-independent potentials (for which we put $n=1$, so that $A\to t$), as well as to the orbits of light and particles in static metrics of sufficiently good symmetry in general relativity. The expression of particle mechanics and geometrical optics in the same mathematical language is an aspect of the optical–mechanical analogy. There are several different formulations of this analogy, but Eq. (1) [with its generalization, Eq. (27)] is the simplest.

When the equations governing two domains of physics can be cast into similar forms, two possibilities arise. The similarity might be due to mere contrivance. In this case, the mathematical analogy can be exploited to transfer methods and results from the more to the less familiar field, but the underlying physics of the two fields could be quite different. The second possibility is that the mathematical analogy results from a deeper connection between the two apparently disparate fields. In the case of the optical–mechanical analogy, this is certainly the situation.

The classical optical–mechanical analogy is based upon the similarity of two variational principles:

$$\delta\int v^{-1}\,dl=0 \quad \text{Fermat, least time (optics)}, \tag{76}$$

$$\delta\int v\,dl=0 \quad \text{Maupertuis, least action (mechanics)}. \tag{77}$$

The roots of Eq. (76) go back to the 17th century, when Pierre de Fermat proposed that light travels from a point in one medium to a point in another by the path requiring the least time.[21] From this, Fermat was able to obtain the constancy of the ratio of the sines of the angles of incidence and refraction. Fermat was attacked by the followers of Descartes for advocating a principle which implied that light travels more rapidly in air than in water when the master had taught that just the opposite was true.[22] Worse, Fermat appeared to be reintroducing teleological arguments and final causes into natural philosophy.[23] Fermat replied bitterly that he would leave to Mr. Descartes the glory of having explained the refraction of light and would content himself with offering an abstract mathematical proposition, without asserting that it applied to light. The originator of the principle of least time retreated from the position that it had any physical content at all.

In 1690, Christiaan Huygens published a *Treatise on Light*, in which he developed a wave theory of light, and applied it to rectilinear propagation, reflection, refraction, and double refraction. Huygens also proved that, if light were a wave, it would obey Fermat's principle of least time. Thus, at the close of the 17th century, Fermat's principle was attached to the wave theory.[24]

Newton, in his *Opticks* of 1704, advocated a corpuscular theory of light—although his corpuscles had to be endowed with a vibratory character to explain the inflections of light (diffraction and interference phenomena). The corpuscularity of light was practically the only proposition of physics on which Newtonians and Cartesians could agree. It is not surprising that, under the combined influence of Newton and Descartes, the wave theory practically disappeared. And Fermat's principle of least time fell into disrepute: hardly anyone believed in it. It is not to be found in most of the 18th-century textbooks, for example. There were in the 18th century always a few advocates of the wave theory, among whom Euler was the most prominent. But the overwhelming majority of physical thinkers believed that light is a stream of particles and that, as both Descartes and Newton had shown, it must therefore travel more quickly in water than in air.

Equation (77) is a product of 18th-century thought. In 1744, Maupertuis announced a minimization principle intended to encompass both mechanics and optics: the principle of least action.[25] Maupertuis called *action* the product of mass, speed and distance. His definition of the action was vague and was applied inconsistently, but he succeeded in deducing from his principle the rules governing elastic and inelastic collisions of particles. He also applied his principle to light and managed to derive the law of refraction. Maupertuis's principle was consistent with the view of light as a particle and the opinion of Descartes and Newton that the speed of light was greater in denser media. With this fact of nature re-established, Maupertuis exclaimed, "the whole edifice that Fermat had constructed is destroyed: light, when it traverses different media, goes neither by the shortest route, nor by the route of least time... ."[26] Maupertuis also pointed out that for the cases of reflection and straight-line propagation—cases in which the speed of light does not change—the path of least time is the same as the path of least action, which, according to Maupertuis, explained how Fermat had gone wrong.

Thus, by the middle of the 18th century, there were two competing variational principles, least time and least action, each of which could serve as a basis for geometrical optics. One of these was associated with the wave theory and one with the particle theory. The *shape* of a ray in a given situation was predicted to be exactly the same by the two theories. This is easily demonstrated. In the particle theory, the index of refraction n is directly proportional to the speed v of the particles. In the wave theory, n is inversely proportional to the (phase) speed v of the waves. Thus, Eqs. (76) and (77) can *both* be expressed by

$$\delta\int n\,dl=0, \quad \text{theory-neutral (optics)}. \tag{78}$$

$n(\mathbf{r})$ is the same function of the coordinates, regardless of the theory we prefer. The two theories thus predict the same path and differ only in the progress in time of light along the ray; but, of course, this could not be measured.

The major issue at stake was the nature of light. Proponents of the particle theory were not simply making an analogy between geometrical optics and mechanics, but rather claiming that they were one and the same. At the close of the 18th century, majority opinion held that one variational principle—that of Maupertuis—covered both. The minority school of wave theorists recognized two principles—those of Fermat and Maupertuis—which applied to different domains of physics.

The principle of least action *as it applies to mechanics* was elaborated and clarified by Euler, and especially by Lagrange. Lagrange formulated the least action principle more precisely, developed a calculus of variations for exploring its consequences, and demonstrated that it could serve as a foundation for mechanics, a welcome alternative to the philosophically disturbing force-based physics of Newton. Among other things, Lagrange showed that the principle of least action and conservation of energy were together equivalent to Newton's law of motion ($F=ma$).[27] This also served to strip Maupertuis's principle of its metaphysical significance. For Maupertuis, the fact that both light and particles follow the paths of least action had been a sure sign of the wisdom of the Creator. For Lagrange, the principle of least action was just an alternative foundation for mechanics.

The principle of "least action" (really, least time) *as it applies to optics* was explored by Laplace,[28] but most notably by William Rowan Hamilton, who was largely responsible for the creation of the optical–mechanical analogy.[29] Hamilton's early work in mathematical physics was devoted to geometrical optics. When Hamilton began this work, in the 1820s, French physicists were already abandoning the particle theory of light, as a result of the success of Fresnel's mathematical wave theory. In England and Ireland (Hamilton's native land), where Newton's influence was stronger, the conversion required another decade or so. Thus, in his first papers on systems of rays (1824–1830), Hamilton always called the integral $\int n\,dl$ the *action*, in analogy to the action integral $\int v\,dl$ of mechanics. The analogy would be most straightforward for the particle theory, in which n is directly proportional to v. However, Hamilton did not necessarily subscribe to the particle theory. In his first paper, he refrained from stating a position on the nature of light and wrote that he used the term *action*, "intending only to express a remarkable analogy, and not assuming any hypothesis about the nature of light." [30] For the integrand of the "action" integral (essentially the index of refraction), Hamilton sometimes used m, sometimes μ, and sometimes v. The use of v for the refractive index might seem to imply acceptance of the particle theory, but this was not the case.

Indeed, by the time (1832) of his "Third Supplement," Hamilton had definitely adopted the wave theory. This is apparent in the softening of his vocabulary; for he now refered to a "law of least action, or of swiftest propagation," i.e., he began to adopt the language of the wave theory and to use it side-by-side with, and as an alternative to, the language of the particle theory.[31] The integrand v he called the *medium-function* and characterized it as "a molecular velocity or an undulatory slowness." In the "Third Supplement," moreover, Hamilton applied his methods directly to the wave theory of Fresnel. The most spectacular result was Hamilton's theoretical prediction of the phenomenon of conical refraction, which was observed shortly afterward by Lloyd.

The essence of Hamilton's geometrical optics is his *characteristic function*

$$V(a,b,c,x,y,z)\equiv\int_{a,b,c}^{x,y,z} n\,dl, \tag{79}$$

which is nothing other than Hamilton's "action" integral or (as we would say) the optical path length between point (a,b,c) and point (x,y,z) on the same ray. Hamilton showed that V satisfies the partial differential equation

$$\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2=n^2(x,y,z). \tag{80}$$

If this differential equation can be solved, knowledge of V constitutes a complete knowledge of the optical system. In particular, the direction of the ray at any point can be found by differentiation of V. Hamilton's characteristic function V is today usually called the *eikonal*, a term introduced by H. Bruns in 1895. Equation (80) is called the *eikonal equation*, and is one of the fundamental equations of modern geometrical optics.[32] The surfaces V=constant are surfaces of equal travel time—the wave fronts.

In the same way, a central feature of Hamilton's mechanics is the characteristic function

$$V(a,b,c,x,y,z)\equiv\int_{a,b,c}^{x,y,z} v\,dl, \tag{81}$$

which is the action for a particle moving between points (a,b,c) and (x,y,z). Hamilton showed that, for a particle of unit mass, V satisfies the partial differential equation

$$\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2=2(E-U), \tag{82}$$

where U is the potential energy function and E is the energy, which is constant on a given trajectory. Note that since $2(E-U)=v^2(x,y,z)$, Eq. (82) for particle mechanics has exactly the same form as Eq. (80) for geometrical optics. Equation (82) is called the time-independent form of the Hamilton–Jacobi equation. If Eq. (82) can be solved, the characteristic function V constitutes a complete description of the motion of the particle. For example, the velocity components of the particle at (x,y,z) can be found by differentiation.[33] Thus, in Hamilton's formulation of mechanics, it is the characteristic function that embodies the optical–mechanical analogy. In mechanics, the surfaces V=constant are surfaces of constant action, analogous to the wave fronts in optics. The particle trajectories correspond to the optical rays. Hamilton's paper on the optical–mechanical analogy, "On a General Method of Expressing the Paths of Light, and of the Planets, by the Coefficients of a Characteristic Function," was published, obscurely, in the *Dublin University Review* for October, 1833.[34]

Hamilton had struggled long and hard with geometrical optics. The development of his general theory of rays occupied him from 1824 to 1832. By contrast, the work on mechanics came quickly and easily in the years 1833–1834, for it was largely but an extension of the optical theory. It is one of the many ironies of this story that there exists today a Hamiltonian system of dynamics because Hamilton happened to develop his geometrical optics at a time when the debate over the particle and wave theories of light had not yet been resolved.[35]

Hamilton's papers on dynamics won him lasting recognition. His method in dynamics was further developed by Jacobi and remained a part of the main stream of 19th-century theoretical physics. In contrast, Hamilton's optical papers were read and understood by very few and soon dropped almost completely out of sight. There were several reasons for this. Hamilton's Irish national sentiment had led him to publish his optical papers in an obscure and little-circulated journal, the *Transactions* of the Royal Irish Academy, even though his uncle had warned him that this would be little better than committing them to a tomb.[36] (The papers on mechanics appeared in the *Philosophical Transactions* of the Royal Society of London). Perhaps an even greater obstacle was Hamilton's difficult style. The papers are very densely written. Moreover, Hamilton strove constantly for higher abstraction and greater generality, and rarely concerned himself with practical applications. He created a general system of geometrical optics, but failed to teach his prospective readers how to use it.

It is a further irony of this story that Bruns, the modern rediscoverer of Hamilton's method in optics, had never seen Hamilton's optical papers. Bruns essentially worked backward from Hamilton's theory of the characteristic function in mechanics to reach an optical analogy—the reverse of the route that Hamilton had traveled some 70 years before.[37]

By the 1830s, the wave theory had triumped. Fermat was vindicated at last, and one could confidently express geometrical optics in the form of Eq. (76). Now there were two well-established variational principles, Eq. (76) for light and Eq. (77) for particles. The fact that they were so similar in form was only a curiosity, and Hamilton's optical–mechanical analogy was regarded as an elaborate and sterile oddity, until the development of quantum theory in the first few decades of our own century.

The action integral of Eq. (77)—expressed in Jacobi's form, Eq. (18)—played a major role in the old quantum theory of Bohr and Sommerfeld. But it was Louis de Broglie who first brought the principles of Fermat and Maupertuis together and showed that they were one. de Broglie argued that atoms displayed effects involving integer numbers (as in the Balmer formula for the frequencies of light emitted by the hydrogen atom). Practically the only place in physics where such numbers turn up is in wave phenomena. But the orbiting electrons are also particles. Thus, de Broglie began by trying to make the electron in the atom simultaneously obey both the principle of Fermat and the principle of Maupertuis. The result of this attempt was the de Broglie relation

$$mv = \hbar k, \tag{83}$$

where k is the wave number of the electron wave, $\hbar$ is Planck's constant divided by 2π, m is the mass of the electron (now considered as a particle), and v is the particle's speed (which de Broglie was also able to identify with the group velocity of the waves).[38] Working backwards from the de Broglie relation, it is easy to see that the two variational principles (as applied to electrons) are really one and the same. Let us express Fermat's principle in terms of the wave number k (rather than n or the phase velocity v):

$$\delta \int k \, dl = 0 \quad \text{(optics)}. \tag{84}$$

Equation (84) expresses the condition that, in the geometrical optics limit, small variations in the path do not affect the optical path length. This is why the ray is where it is: Only along the actual ray can neighboring bundles of virtual paths interfere constructively. Equation (84) must apply, in the geometrical optics limit, to all wave disturbances that obey the superposition principle. To apply Eq. (84), which is very general, to the particular case of matter waves, we use the de Broglie relation, Eq. (83). Equation (84) immediately becomes Maupertuis's principle, Eq. (77). Thus, classical mechanics can be understood as the geometrical-optics limit of wave mechanics: the particle trajectories are the rays of the matter waves. Maupertuis's principle therefore amounts to a special case of Fermat's principle—the special case involving the matter waves of quantum mechanics.

While de Broglie began directly from the analogy between the principles of Fermat and Maupertuis, Schrödinger chose to begin from Hamilton's form of the optical–mechanical anology. In particular, Schrödinger succeeded in deriving the time-independent form of the electron-wave equation (Schrödinger's equation) from the time-independent Hamilton–Jacobi equation, Eq. (82).[39] Thus, the origin of wave mechanics is a part of the history of the optical–mechanical analogy. We can hardly deny that the optical–mechanical analogy possesses a deep physical significance.

In retrospect, Hamilton's optical–mechanical analogy can be seen as providing a hint about the wavelike nature of material particles, a hint that lay unnoticed until de Broglie took it up a century later. A great, and so far unrealized, goal of late 20th-century physics is the unification of gravity with the other forces of nature. The program of unification has been profoundly successful, beginning from the 19th-century unification of optics, radiant heat, electricity, and magnetism. The quantum field theorist, confronted with effective Lagrangians in general relativity, naturally seeks to apply the techniques that have worked so well with other fields. But one could still maintain the possibility that this is only wishful thinking and that gravity has nothing to do with the other forces.

The skeptical position is, of course, logically possible—but then what is one to make of variational principles such as Eq. (19)? In the case of the propagation of light and the motion of particles in a gravitational field, both Fermat's principle and Maupertuis's principle may be regarded as special cases of Eq. (19), which, in turn, follows from the geodesic property of the trajectories. There is nothing of waves involved in Eq. (19), which follows from considerations of the behavior of particles in general relativity. How can it be the case that (1) Maupertuis's principle is a special case of Fermat's in the context of wave mechanics and that (2) both Maupertuis's principle and Fermat's principle are special cases of a variational principle deduced from general relativity? The fact that Maupertuis's principle can simultaneously be the geometrical-optics limit of wave mechanics and a special case of the geodesic equation in general relativity seems to argue strongly for an underlying unity of quantum mechanics and gravitation theory.

APPENDIX: ON THE RELATION OF "$F=MA$" OPTICS TO HAMILTON'S FORMULATION OF THE OPTICAL–MECHANICAL ANALOGY

We can, if we wish, formulate the motion of photons and massive particles in the Schwarzschild field in the same way as Hamilton formulated the motion of classical particles and the paths of light in refractive media. In analogy to Eq. (79) or Eq. (81), we define a characteristic function

$$V(\mathbf{y},\mathbf{x}) \equiv \int_{\mathbf{y}}^{\mathbf{x}} n^2 v \, dl, \tag{85}$$

where $\mathbf{y}$ and $\mathbf{x}$ are two points in space. V will then satisfy the partial differential equation

$$\sum_i \left(\frac{\partial V}{\partial x_i} \right)^2 = n^4 v^2. \tag{86}$$

If Eq. (86) can be solved, V will contain all the information we require about the trajectories of photons and particles in the Schwarzschild field. This would lead us into a realm of high abstraction, with no increase in calculating power over Eq. (27).

But it is worth a moment's reflection to see that Eq. (86) is actually equivalent to Eq. (27). We may write

$$dl=\sum_i \cos\theta_i\, dx_i, \tag{87}$$

where the $\cos\theta_i$ are the direction cosines of **dl**, i.e., the cosines of the angles **dl** makes with the Cartesian axes. If we now parametrize the path by a stepping parameter A (yet to be defined),

$$\cos\theta_i=\frac{(dx_i/dA)}{(dl/dA)}. \tag{88}$$

Then V can be written

$$V=\sum_i \int n^2 v\left(\frac{dl}{dA}\right)^{-1}\frac{dx_i}{dA}\, dx_i, \tag{89}$$

from which it follows that

$$\frac{\partial V}{\partial x_i}=n^2 v\left(\frac{dl}{dA}\right)^{-1}\frac{dx_i}{dA}. \tag{90}$$

Substituting Eq. (90) into Eq. (86), then differentiating both sides of Eq. (86) with respect to A gives

$$\sum_i \frac{d}{dA}\left[n^2 v\left(\frac{dl}{dA}\right)^{-1}\frac{dx_i}{dA}\right]^2=\sum_i \frac{\partial}{\partial x_i}(n^4v^2)\,\frac{dx_i}{dA}. \tag{91}$$

The left-hand side of Eq. (91) cries out for us to define A by

$$\frac{dl}{dA}=n^2 v, \tag{92}$$

just as in Eq. (25). For then Eq. (91) becomes

$$2\sum_i \frac{d^2x_i}{dA^2}\frac{dx_i}{dA}=\sum_i \frac{\partial}{\partial x_i}(n^4v^2)\,\frac{dx_i}{dA}. \tag{93}$$

The direction of the "velocity" $d\mathbf{x}/dA$ at a given point can be chosen at will, in the form of initial conditions. Thus, for Eq. (93) to hold in all circumstances, the coefficients of the "velocity" components dx_i/dA must be equal term by term:

$$\frac{d^2x_i}{dA^2}=\frac{\partial(\frac{1}{2}n^4v^2)}{\partial x_i},$$

or, in vector form, Eq. (27) precisely.

Thus, Eq. (1), and its generalization, Eq. (27), can be regarded as equivalent to Hamilton's formulation of the optical–mechanical analogy: we need only make the right choice for the stepping parameter. Again, it should be noted that both Eqs. (86) and (27) include classical mechanics (for which $n\to 1$ and $A\to t$) and classical geometrical optics (for which $v\to c_0/n$) as special cases.

By virtue of Eq. (90) and Eq. (92), $dx_i/dA=\partial V/\partial x_i$. Thus the generalized version of Hamilton's differential equation [Eq. (86)] becomes, for us, a first integral of the motion [Eq. (30)], equivalent to "conservation of energy." Hamilton, of course, would have regarded $dx_i/dA=\partial V/\partial x_i$ as the prescription for extracting answers, in the form of velocity components (dx_i/dA) by differentiation of the solution V to the differential equation!

[a]E-mail address: JCEvans@UPS.edu

[1] J. Evans and M. Rosenquist, "'$F=ma$' optics," Am. J. Phys. **54**, 876–883 (1986).

[2] J. Evans, "Simple forms for equations of rays in gradient-index lenses," Am. J. Phys. **58**, 773–778 (1990).

[3] A. S. Eddington, *Space, Time and Gravitation* (Cambridge U. P., Cambridge, 1920), p. 109.

[4] J. Plebanski, "Electromagnetic waves in gravitational fields," Phys. Rev. **118**, 1396–1408 (1960).

[5] F. de Felice, "On the gravitational field acting as an optical medium," Gen. Rel. Grav. **2**, 347–357 (1971).

[6] K. K. Nandi and A. Islam, "On the optical–mechanical analogy in general relativity," Am. J. Phys. **63**, 251–256 (1995).

[7] For more detail see James Evans, Kamal K. Nandi, and Anwarul Islam, "The optical–mechanical analogy in general relativity: Exact Newtonian forms for the equations of motion of particles and photons," Gen. Rel. Grav. **28**, 413–439 (1996).

[8] The coordinate speed of light is not, of course, susceptible of direct measurement. Each observer in a locally inertial frame will measure c_0 for the speed of light.

[9] The isotropic form of the Schwarzschild line element is mentioned in many standard texts, e.g., W. Misner, K. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Co., San Francisco, 1973), p. 840. For a systematic technique for finding isotropic coordinates see R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity and Gravitation* (McGraw-Hill, New York, 1965), pp. 174–177. Effective indices of refraction for a number of metrics are given in Ref. 7. Note in particular that any spherically symmetric static metric can be put into isotropic form.

[10] For passing between Hamilton's principle and Jacobi's form of Maupertuis's principle, see Cornelius Lanczos, *The Variational Principles of Mechanics* (Dover, New York, 1986), 4th ed., pp. 125–128 and 132–134.

[11] We have obtained Eq. (19) directly from the geodesic condition. But it may also be considered a special case of the three-dimensional variational principle of Weyl and Levi-Civita. Hermann Weyl, "Zur Gravitationstheorie," Ann. Phys. **54**, 117–145 (1917). T. Levi-Civita, "Statica einsteiniana," Atti della Reale Accademia Dei Lincei, Ser. 5, Rendiconti, Classe di scienze fisiche, matematiche e naturali **26**, 458–470 (1917). To obtain Eq. (19) from the variational principle of Weyl and Levi-Civita, one goes into isotropic coordinates, introduces the index of refraction, and then invokes conservation of energy on the varied path.

[12] That light obeys Fermat's principle in static metrics is, of course, a well-known fact. That is, the coordinate time of travel between two fixed points is invariant under infinitesimal deformations of the path. T. Levi-Civita, "La teoria di Einstein e il principio di Fermat," Nuovo Cimento **16**, 105–114 (1918). For a good recent discussion see N. Straumann, *General Relativity and Relativistic Astrophysics* (Springer-Verlag, Berlin, 1984), pp. 99–100.

[13] Reference 2, p. 774. The stepping parameter used in Refs. 1 and 2 differs from the stepping parameter adopted here by a constant factor c_0. This is only a matter of convention. We call A the optical action because dA is proportional to $c(\mathbf{r})\, dl$, and thus is analogous to the action $v(\mathbf{r})\, dl$ of classical mechanics.

[14] The classical limit of this expression is $-\frac{1}{2}v^2$, which is equal to $U(\mathbf{r})-E$. In the classical situation, it might be more appropriate to refer to $-\frac{1}{2}v^2(\mathbf{r})$ as the "speed function," since it does differ from the potential energy by an additive constant. In classical geometrical optics (Ref. 1), the "potential energy" is $-\frac{1}{2}n^2c_0^2$ and the additive constant is always zero. In the geometrical optics of the Schwarzschild field, by virtue of Eq. (32), the "potential energy" is again $-\frac{1}{2}n^2c_0^2$, with no additive constant. In the case of particle motion in general relativity, the "potential energy" is given by Eq. (28)—with no additive constant, but with a constant of integration (corresponding to the energy) buried in the expression for $v(\mathbf{r})$. Thus the term "potential energy" for $-\frac{1}{2}n^4v^2$ should be regarded as coming from the optical situations, in which it is perfectly apt. For particle dynamics, whether classical or relativistic, the term is perhaps not a perfect choice.

[15] Other metrics amenable to the same treatment include the de Sitter metric, the Reissner–Nordström metric, the Bertotti–Robinson metric and the Halilsoy metric. These are all discussed in Ref. 7.

[16]For an example of a radar-echo delay calculation which requires the use of Eq. (26), see Ref. 7.

[17]$\mathbf{p}$ is not the canonical momentum used above, but rather the *optical analog* to the classical Newtonian momentum.

[18]The same result is obtained by standard methods in Misner, Thorne and Wheeler (Ref. 9), p. 671.

[19]To begin the classical calculation, write $d\mathbf{p}/dt=\boldsymbol{\nabla}(\frac{1}{2}v^2)$, where $\frac{1}{2}v^2=\frac{1}{2}v_0^2+MG/r$. The angular deflection turns out to be $(2m/R)(c_0^2/v_0^2)$.

[20]For a detailed solution of Eq. (74), see, for example, Jerry B. Marion, *Classical Dynamics of Particles and Systems* (Academic, New York, 1970), 2nd ed., pp. 266–270.

[21]Fermat made his principle of least time known through correspondence, starting around 1662. He wrote two short pieces for this purpose, which were never published in a regular way during his lifetime. They first appeared in print in the published correspondence of Descartes. These two pieces can be found in the standard edition of Fermat's works: *Oeuvres de Fermat*, edited by Paul Tannery and Charles Henry (Gautier-Villars et Fils, Paris, 1891–1912). See "Analysis ad refractiones," in Vol. 1, pp. 170–172, with French translation in Vol. 3, pp. 149–151; and "Synthesis ad refractiones," Vol. 1, pp. 173–179, with French translation in Vol. 3, pp. 151–156. For an English translation of the first piece, see William Francis Magie, *A Source Book in Physics* (McGraw-Hill, New York, 1935), pp. 278–280.

[22]Descartes had derived the sine law from a particle model of light. The component of the velocity parallel to the interface was assumed to remain constant when light crossed from one medium to another. Since the ray lies closer to the normal in water than in air, this implies that it travels faster in the water. René Descartes, *La dioptrique* (published as a supplement to *Discours de la méthode* (Leiden, 1637). *Oeuvres de Descartes*, edited by Charles Adam and Paul Tannery (J. Vrin, Paris, 1874–1882), 2nd ed., Vol. 6., pp. 93–105. *Discourse on Method, Optics, Geometry, and Meteorology*, translated by Paul J. Olscamp (Bobbs–Merrill, Indianapolis, 1965).

[23]On Fermat versus the Cartesians, see René Dugas, *A History of Mechanics*, translated by J. R. Maddox (Dover, New York, 1988), pp. 258–259.

[24]Christiaan Huygens, *Traité de la lumière* (Leiden, 1690). *Oeuvres complète de Christiaan Huygens* (Martinus Nijhoff, La Haye, 1888–1950), Vol. 19, pp. 451–537. English translation: *Treatise on Light...by Christiaan Huygens*, translated by Silvanus P. Thompson (Dover, New York, 1962; first published in 1912). For Huygens's treatment of Fermat's principle, see *Oeuvres*, Vol. 19, pp. 489–490; Thompson, pp. 42–45.

[25]Pierre-Louis Moreau de Maupertuis, *Oeuvres* (Georg Olms, Hildesheim and New York, 1965–1974), 4 Vols. (This is a reprint based on the editions of Lyon, 1768 and Berlin, 1758.) Maupertuis announced his principle of least action in a paper read in 1744 to the Paris Academy of Sciences, and titled, "The Agreement between Different Laws of Nature, Which Had until Now Appeared Incompatible," (*Oeuvres*, Vol. 4, pp. 3–28). This first paper addressed the refraction of light. A second paper, "Research on the Laws of Motion," read to the Royal Academy of Sciences of Berlin in 1746, treated the collisions of elastic and inelastic bodies by the method of least action (*Oeuvres*, Vol. 4, pp. 31–42). It is one of the many ironies of this story that Maupertuis applied least action first to light, and only later to mechanics: the principle was applied first to the domain in which it is invalid.

[26]Fermat, *Oeuvres* (Ref. 21), Vol. 4, p. 15.

[27]For an account of Euler's and Lagrange's development of Maupertuis's principle, see Wolfgang Yourgrau and Stanley Mandelstam, *Variational Principles in Dynamics and Quantum Mechanics* (Dover, New York, 1979).

[28]Pierre-Simon de Laplace, *Oeuvres complète de Laplace*, (Gauthier-Villars, Paris, 1878–1904), 13 vols., Vol. 12, pp. 267–298.

[29]Hamilton's work is most easily consulted in *The Mathematical Papers of Sir William Rowan Hamilton* (Cambridge U.P., Cambridge, 1931–1967), 3 vols.; Vol. 1, *Geometrical Optics*, edited by A. W. Conway and J. L. Synge; Vol. 2, *Dynamics*, edited by A. W. Conway and A. J. McConnell; Vol. 3, *Algebra*, edited by H. Halberstam and R. E. Ingram.

[30]Hamilton, Ref. 29, Vol. 1, p. 14.

[31]Hamilton, Ref. 29, Vol. 1, pp. 168–169.

[32]See, for example, Max Born and Emil Wolf, *Principles of Optics* (Pergamon, Oxford, 1964), 2nd ed., p. 112; or Miles V. Klein, *Optics* (Wiley, New York, 1970), p. 28.

[33]This is clearer if we write V in the form $V=\int(v_x\,dx+v_y\,dy+v_z\,dz)$. Then $\partial V/\partial x=v_x$.

[34]Reprinted in Hamilton, Ref. 29, Vol. 1, pp. 311–332. For an excellent summary of Hamilton's further development of his dynamical theory (leading to the canonical equations of motion and the Hamilton–Jacobi equation) see Thomas L. Hankins, *Sir William Rowan Hamilton* (Johns Hopkins U.P., Baltimore and London, 1980), pp. 181–198. Hankins also provides a good summary of Hamilton's work in geometrical optics (pp. 61–81).

[35]Cf. editors' introduction to Hamilton, Ref. 29, Vol. 1, p. xxi.

[36]Hankins, Ref. 34, p. 86.

[37]Hankins, Ref. 34, p. 87.

[38]Louis de Broglie, *Recherches sur la théorie des quanta* (Masson, Paris, 1963). (A reprint of the thesis of 1924). Here, we have expressed the de Broglie relation nonrelativistically. de Broglie's relation does, of course, hold relativistically: We need only replace mv by the relativistic momentum p.

[39]E. Schrödinger, "Quantisierung als Eigenwertproblem," Ann. Phys. **79**, 361–376 (1926). This paper is available in English translation in J. F. Shearer and W. M. Deans (translators), *Collected Papers on Wave Mechanics by E. Schrödinger* (Blackie and Son, Glasgow, 1928). For an account of Schrödinger's procedure, see Ref. 27, pp. 118–119, or W. Moore, *Schrödinger: Life and Thought* (Cambridge U.P., Cambridge, 1989), pp. 200–207.