"$F = m\,a$" optics

James Evans
University of Puget Sound, Tacoma, Washington 98416
Mark Rosenquist
Shell Western Exploration & Production, Inc., Rocky Mountain Division, P. O. Box 831, Houston, Texas 77001

(Received 10 May 1985; accepted for publication 9 October 1985)

Fermat's principle may be used to derive an equation, of the form $F = ma$, governing the shape of a light ray in a medium of varying refractive index. Many interesting problems in gradient-index optics that ordinarily require considerable computation may therefore be solved very simply by analogy to familiar mechanical problems. This approach also provides the means of thoroughly exploring the optical-mechanical analogy at a much more elementary level than is usual.

I. INTRODUCTION

In this paper we present several instances in which experience and insights developed in mechanics can be used to solve classic and interesting problems from the geometrical optics of media with varying index of refraction. This approach to optics fits very well into a mechanics course at the level of the books by Marion[1] and Symon.[2] Such an excursion into gradient-index optics can strengthen a mechanics course by bolstering the treatment of Hamilton's principle with a second example of a variational principle (Fermat's), and by showing how methods of solving $F = m\,a$ problems can be used in a different context. In addition, "$F = m\,a$" optics produces a great reduction in the mathematical complexity of many interesting optics problems and presents the optical-mechanical analogy in a way that is easily accessible to undergraduates.

II. DEVELOPMENT OF THE FUNDAMENTAL EQUATION

A. Formulation of a variational principle

We begin our presentation of geometrical optics in media with varying index of refraction with the statement of a variational principle. The principle is essentially Fermat's: that rays from one point to another are paths of minimum (or stationary) phase. Although Fermat's principle is quite familiar, we briefly recount its justification here: one of the chief virtues of Fermat's principle for intermediate-level mechanics students is that—in contrast to Hamilton's principle— it may be justified by a simple physical argument. To simplify the discussion, we will always assume that the waves are continuous and of a single frequency.

We regard the wave disturbance as a sum of the contributions propagated along all possible paths from the source to the receiver. Since neighboring paths traverse nearly the same ground in the medium, the magnitudes of the disturbances arriving along each path must be virtually identical. However, since we are considering wave motion in the geometrical limit, in which the wavelength is much smaller than any other feature of the problem, even a slight difference in paths means that the phase of the wavelets contributed by neighboring paths will be many cycles different. (See Fig. 1.) A bundle of neighboring virtual paths from the source to the receiver will therefore make contributions with equal magnitudes but with a random assortment of phases, and will therefore interfere destructively. The only circumstance in which neighboring paths will not collectively cancel one another's contributions occurs when the phase along the path is stationary, e.g., minimum. In this single instance, the phases of the contributions along neighboring paths are nearly equal and the wavelets interfere constructively. We may therefore characterize light trajectories with the following principle:

δ(the phase along the path) = 0.

In a medium of varying index of refraction n, the phase along a path is calculated as the integral of the wave number times the displacement along the path:

$$\text{phase} = \int k\,|d\mathbf{r}| = \frac{\omega}{c}\int n(\mathbf{r})\,|d\mathbf{r}|. \quad (1)$$

To insist that the phase be stationary, of course, we must vary the path slightly and require that the variation in the phase when this is done be zero:

$$\delta\,\text{phase} = \frac{\omega}{c}\,\delta\int n(\mathbf{r})\,|d\mathbf{r}| = 0,$$

$$\delta\int n(\mathbf{r})\,|d\mathbf{r}| = 0. \quad (2)$$

The typical physics undergraduate encounters only one variational principle: Hamilton's principle of least action in a mechanics course. It is usually presented without justification. Because of its global nature and unspecified origins, it often takes on a mystical character. Indeed, Maupertuis himself, the eighteenth-century originator of a version of the least action principle, was more than a little mystical in his enthusiasm:

"Our principle, more consistent with the ideas we ought to have of things, leaves the world in the continual need of the power of the Creator, and is a necessary consequence of the wisest use of this power... . These laws, so beautiful and so simple, are perhaps the only ones that the Creator and Ordainer of things has established in matter to effect all the phenomena of this visible world."[3]

By contrast, the principle of stationary phase in geomet-

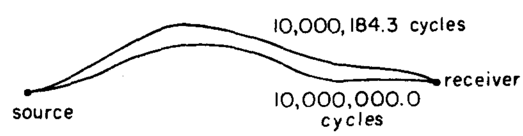

Fig. 1. Random phase relation between neighboring virtual paths.

rical optics can be obtained by simple arguments based solely on the concepts of the wave and interference. If we expose students to a second instance of a variational principle in physics, and support this principle with physical reasoning, perhaps Hamilton's principle will seem less esoteric.

B. Differential equation for the ray

In order to find ray shapes in various media, we derive a differential equation from the variational principle. This calculation parallels the derivation of the Euler–Lagrange equations of motion from Hamilton's principle, and its main steps will therefore be familiar.

We specify the position along a path in three dimensions as a function of a single variable a that we call the stepping parameter. As the variable a increases, the point specified by $\mathbf{r}(a)$ moves smoothly along the path;

$$\mathbf{r} = [x(a), y(a), z(a)].$$

The variational principle [Eq. (2)] can be expressed in terms of an ordinary integral with respect to the stepping parameter a:

$$\delta \int n(\mathbf{r}) \left| \frac{d\mathbf{r}}{da} \right| da = 0. \tag{3}$$

The integral is varied by integrating along a slightly different path. The second path differs from the first by a variation function $\mathbf{e}(a)$, which is small because the paths are neighboring. (See Fig. 2.) Thus, in performing the variation, we make the substitution

$$\mathbf{r} \to \mathbf{r} + \mathbf{e}. \tag{4}$$

Moreover, $\mathbf{e}(a)$ vanishes at the end points, because the disturbances must still orginate from the source and be observed at the receiver.

The variation indicated in Eq. (3) may be expressed in terms of a variation in n and a variation in $|d\mathbf{r}/da|$:

$$\int \left[(\delta n) \left| \frac{d\mathbf{r}}{da} \right| + n\left(\delta \left| \frac{d\mathbf{r}}{da} \right|\right) \right] da = 0. \tag{5}$$

The index of refraction is a function only of $\mathbf{r}$ and thus, to the first order in $\mathbf{e}$, the variation in n is

$$\delta n = (\mathbf{grad}\, n) \cdot \mathbf{e}. \tag{6}$$

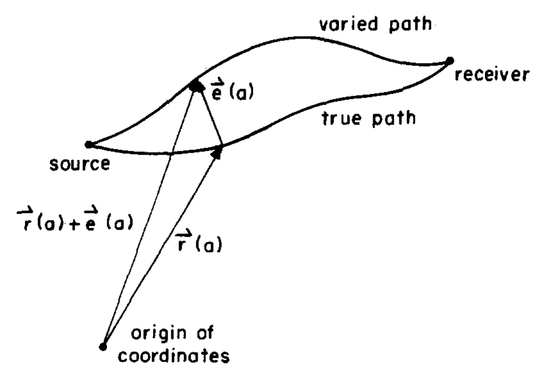

Fig. 2. Variation of the path of integration.

We now calculate the variation in $|d\mathbf{r}/da|$:

$$\delta \left| \frac{d\mathbf{r}}{da} \right| = \left| \frac{d(\mathbf{r}+\mathbf{e})}{da} \right| - \left| \frac{d\mathbf{r}}{da} \right|$$
$$\simeq \left[\left| \frac{d\mathbf{r}}{da} \right|^2 + 2 \frac{d\mathbf{r}}{da} \cdot \frac{d\mathbf{e}}{da} \right]^{1/2} - \left| \frac{d\mathbf{r}}{da} \right|$$
$$\simeq \left(\frac{d\mathbf{r}}{da} \cdot \frac{d\mathbf{e}}{da} \right) \Big/ \left| \frac{d\mathbf{r}}{da} \right|,$$

correct to the first order in $\mathbf{e}$.

We now introduce the prime to denote differentiation with respect to the stepping parameter a: $\mathbf{r}' \equiv d\mathbf{r}/da$. Thus we have

$$\delta \left| \frac{d\mathbf{r}}{da} \right| = \frac{\mathbf{r}' \cdot \mathbf{e}'}{|\mathbf{r}'|}. \tag{7}$$

Substitution of Eqs. (6) and (7) into Eq. (5) gives

$$\int \left(\mathbf{grad}\, n \cdot \mathbf{e} |\mathbf{r}'| + \frac{n\mathbf{r}' \cdot \mathbf{e}'}{|\mathbf{r}'|} \right) da = 0. \tag{8}$$

Just as in mechanics, we perform a partial integration of the second term in Eq. (8). The integrated term vanishes because the variation $\mathbf{e}$ is zero at the source and the receiver. We obtain

$$\int \left[(\mathbf{grad}\, n) |\mathbf{r}'| - \frac{d}{da} \left(\frac{n\mathbf{r}'}{|\mathbf{r}'|} \right) \right] \cdot \mathbf{e}\, da = 0.$$

Since we require the integral to be zero for all neighboring paths, i.e., for any infinitesimal $\mathbf{e}$, the term in brackets must be identically zero. This gives us a differential equation for the rays:

$$(\mathbf{grad}\, n) |\mathbf{r}'| = \frac{d}{da} \left(\frac{n\mathbf{r}'}{|\mathbf{r}'|} \right). \tag{9}$$

The usual form of this equation, derived from the eikonal equation in geometrical optics,[4] takes the stepping parameter to be the arc length s. Then $d\mathbf{r}/ds$ is a unit vector along the path and $|d\mathbf{r}/ds| = 1$. The differential equation [Eq. (9)] becomes:

$$\mathbf{grad}\, n = \frac{d}{ds} \left(n \frac{d\mathbf{r}}{ds} \right), \quad (s = \text{arc length})$$

or

$$\mathbf{grad}\, n = \left[(\mathbf{grad}\, n) \cdot \frac{d\mathbf{r}}{ds} \right] \frac{d\mathbf{r}}{ds} - n \frac{d^2\mathbf{r}}{ds^2}.$$

This equation is unfortunately nonlinear and, furthermore, the various components of $\mathbf{r}(a)$ are coupled. A far greater simplification of Eq. (9) results if we choose the stepping parameter to be something other than the arc length. If we choose the stepping parameter a such that $|\mathbf{r}'| = n$, the equation for the ray [Eq. (9)] becomes

$$n\, \mathbf{grad}\, n = \mathbf{r}'', \quad \left(\left| \frac{d\mathbf{r}}{da} \right| = n \right)$$

or

$$\mathbf{grad} \left(\frac{n^2}{2} \right) = \mathbf{r}''. \tag{10}$$

The physics student has no older friend than this equation, which has exactly the same form as $\mathbf{F} = m\, \mathbf{a}$. The familiarity of this equation allows the student to call on all of his experience at problem solving in mechanics in the new domain of optics.

III. A TABLE OF ANALOGIES

The optical-mechanical analogy has been discussed often and from several different perspectives.[5] The analogy continues to provide a useful point of departure for investigations of mechanical and optical systems.[6] The sophistication of most of these treatments places them well beyond the reach of the typical undergraduate. A principal goal of the present paper is to offer an approach to the optical-mechanical analogy that is simple enough and useful enough to be included in an undergraduate mechanics or optics course. At the deepest level, the physical significance of the optical-mechanical analogy rests on the common wavelike properties of light and material particles. Light is, however, not "just like particles," and one must refrain from pushing the analogy too far in an excess of enthusiasm. For example, the motions of particles and light pulses, even in corresponding potentials, are not entirely the same—a point that will be illustrated below.[7] In our short course on gradient-index optics we do not stress, therefore, the physical analogy. Rather, we exploit the formal similarity between equations to permit the rapid solution of optical problems cast into the same form as familiar mechanical problems.

It is convenient formally to identify several mechanical quantities with the analogous quantities in "$F = m\,a$" optics. (Refer to Table I.) The mechanical quantities—position, velocity, and so on—are presumed to be associated with a traveling material point particle. The corresponding optical quantities are associated with a moving pulse of light. Our derivation of the variational principle depended on neighboring rays interfering destructively. In order that this interference take place, disturbances emitted at different times must meet at the receiver. The "pulse" of light must therefore consist of many cycles. Also, the derivation was based on the assumption that the disturbance was of a single frequency. If there is little dispersion, all frequencies will follow nearly the same path but significant dispersion, as always, will cause a finite pulse to spread.

The position of the light pulse corresponds to the position of the material particle. But where the latter is generally regarded as a function of time, we treat the position of the light pulse as a function of the stepping parameter a. (See Table I.)

Thus the role of the time t is played in the optical formalism by the stepping parameter a. The side condition $|d\mathbf{r}/da| = n$ allows one to pass over to a description in terms of the time, if necessary. A second form of the side condition is obtained by the following simple manipulation:

$$n = \left|\frac{d\mathbf{r}}{da}\right| \tag{11}$$

$$= \left|\frac{d\mathbf{r}}{dt}\right|\frac{dt}{da}$$

$$= \frac{c}{n}\frac{dt}{da}.$$

And thus

$$da = \frac{c}{n^2}\,dt. \tag{12}$$

This second form of the side condition will frequently be useful.

Corresponding to $d\mathbf{r}/dt$, the velocity of the material particle, we have the optical quantity $d\mathbf{r}/da$. (We will denote derivatives with respect to time by a dot, and derivatives with respect to the stepping parameter a by a prime. Thus $\dot{\mathbf{r}} \equiv d\mathbf{r}/dt$ and $\mathbf{r}' \equiv d\mathbf{r}/da$.) Although $\mathbf{r}'$ plays the role of a velocity in the optical formalism, it is not really a velocity. The stepping parameter a has dimensions of length, not time, so $\mathbf{r}'$ is actually dimensionless.

From Eq. (10), the optical equivalent of Newton's second law, $m\ddot{\mathbf{r}} = -\,\text{grad}\,U$, we may identify the analog of the potential energy as $-n^2/2$, and the analog of the mass as the dimensionless number 1.

The analog of the kinetic energy may be constructed by combining the quantities analogous to $\dot{\mathbf{r}}$ and m. $T = \frac{1}{2}m|\dot{\mathbf{r}}|^2$ passes over to $\frac{1}{2}|\mathbf{r}'|^2$.

The analog of the total energy is

$$E = T + U$$

$$= \tfrac{1}{2}|\mathbf{r}'|^2 - \tfrac{1}{2}n^2.$$

But the side condition [Eq. (11)] requires that $|\mathbf{r}'| = n$. Hence

$$E = 0.$$

Thus $F = m\,a$ optics is analogous to mechanics at zero energy. This result follows from the fact that there are fewer initial conditions to be specified in the optical system than in the corresponding mechanical case. Imagine that we release a material particle in a mechanical potential. We are free to choose the particle's initial position, its initial speed and the initial direction of its motion. With these quantities specified, the potential function and Newton's second law determine the particle's future motion. The corresponding optical situation is this: imagine a region of varying index of refraction. Imagine turning on a flashlight briefly. The trajectory of the light pulse corresponds to the trajectory of the material particle. We are free to choose the initial position of the light pulse, (i.e., we may place the flashlight wherever we please). We may also choose the initial direction of the motion of the light pulse (by aiming the flashlight as we please). We may not, however, arbitrarily choose the initial speed of the light pulse, for that is determined by the index of refraction, which is itself presumed to be a function of the position alone. Thus, in the optical situation, the position of the light pulse determines not only its potential, but also its kinetic "energy." In the case that we choose $U = -n^2/2$, with no additive constant, the "kinetic energy" turns out to be the negative of the "potential energy," so that the "total energy" is always zero. The optical situation is therefore somewhat more restrictive than

Table I. Corresponding quantities in mechanics and $F = m\,a$ optics.

Quantity	Mechanics	Optics
Position	$\mathbf{r}(t)$	$\mathbf{r}(a)$
"Time"	t	a
"Velocity"	$\frac{d\mathbf{r}}{dt} \equiv \dot{\mathbf{r}}$	$\frac{d\mathbf{r}}{da} \equiv \mathbf{r}'$
"Potential energy"	$U(\mathbf{r})$	$\frac{-n^2(\mathbf{r})}{2}$
"Mass"	m	1
"Kinetic energy"	$T = \frac{m}{2}\left\|\frac{d\mathbf{r}}{dt}\right\|^2$	$\frac{1}{2}\left\|\frac{d\mathbf{r}}{da}\right\|^2$
"Total energy"	$\frac{m}{2}\left\|\frac{d\mathbf{r}}{dt}\right\|^2 + U$	$\frac{1}{2}\left\|\frac{d\mathbf{r}}{da}\right\|^2 - \frac{n^2}{2} = 0$

the mechanical one. A given mechanical problem may admit of several solutions, corresponding to different energies. The analogous optical problem will in general have but a single solution—the one corresponding to zero "energy."

The prescription $E = 0$ results from the identification of $-n^2/2$ with the potential energy, i.e., with no additive constant. In most situations this convention leads to optical results which parallel familiar mechanical results. In a few cases, however, this convention for the "optical potential energy" differs from the convention normally followed in the corresponding mechanical situation. An example of such a difference will be found in our treatment of the optical analog of the harmonic oscillator in Sec. V F below.

IV. SOLVING PROBLEMS WITH "$F = m\,a$" OPTICS

In this and the following section, we will give several examples of optical problems that can be solved using the methods and intuitions of mechanics. We will begin with a simple derivation of Snell's law and progress to some of the classic problems of gradient-index optics. The last few of these are normal fare in a graduate level optics course and usually involve considerable computation. In "$F = m\,a$" optics, however, these problems are well within the reach of undergraduates.

A. Snell's law

Consider two regions (called 1 and 2), separated by a plane boundary, as in Fig. 3. Let the index of refraction in these two regions be n_1 and n_2, respectively. Choose coordinate axes as shown, with the x axis lying along the boundary.

As the index of refraction does not vary in the x direction, our fundamental equation [Eq. (10)] requires

$$\frac{d^2x}{da^2} = 0,$$

so that

$$\frac{dx}{da} = \text{constant}.$$

That is, dx/da is the same in the two media. The problem is somewhat analogous to a free-fall problem in mechanics:

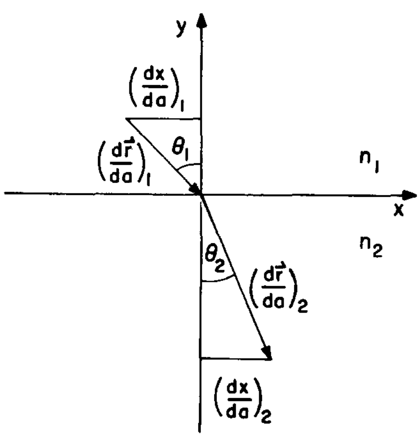

Fig. 3. Derivation of Snell's law.

the x component of the "velocity" is constant, while the y component of the "velocity" and total "speed" are not. Applying the equality $(dx/da)_1 = (dx/da)_2$ to Fig. 3 gives

$$\left|\frac{d\mathbf{r}}{da}\right|_1 \sin\theta_1 = \left|\frac{d\mathbf{r}}{da}\right|_2 \sin\theta_2.$$

The side condition [Eq. (11)] then yields

$$\left|\frac{d\mathbf{r}}{da}\right|_1 = n_1 \quad \text{and} \quad \left|\frac{d\mathbf{r}}{da}\right|_2 = n_2.$$

Thus we have Snell's Law:

$$n_2 \sin\theta_2 = n_1 \sin\theta_1,$$

as a consequence of the constancy of the x component of $d\mathbf{r}/da$ across the boundary.[8]

B. Road surface mirage

For a less trivial example we consider a linearly varying index of refraction. This might be a reasonable first approximation to the optical properties of the air above a road surface on a sunny day, when the air next to the road is less dense than the air higher up. Such variation in the density of the air leads to the familiar road surface mirage in which a dry road appears to be wet. To solve this problem, we let $n = n_0 + \alpha y$, where y is the vertical distance above the road surface, and where n_0 and α are constants. Our "equations of motion" [$\mathbf{r}'' = \mathbf{grad}(n^2/2)$] are then

$$\frac{d^2x}{da^2} = 0,$$

$$\frac{d^2y}{da^2} = \alpha(n_0 + \alpha y).$$

This corresponds to uniform motion in the x direction and exponential motion in the y direction. Integration of the x equation yields

$$x = Aa + B.$$

where A and B are constants of integration. The solution of the differential equation for $y(a)$ is a hyperbolic sine or cosine plus a particular integral:

$$y = C\cosh(\alpha a) + D\sinh(\alpha a) - (n_0/\alpha).$$

The constants A, B, C, and D are to be determined by application of the "initial conditions."

Choose coordinate axes as in Fig. 4 with the y axis passing through the lowest point of the ray, which is a distance h above the road surface. Further, let $a = 0$ when $x = 0$. (This corresponds to defining the moment $t = 0$ in a mechanics problem.) The x and y positions at $a = 0$ are then

$$x_{a=0} = 0, \quad y_{a=0} = h.$$

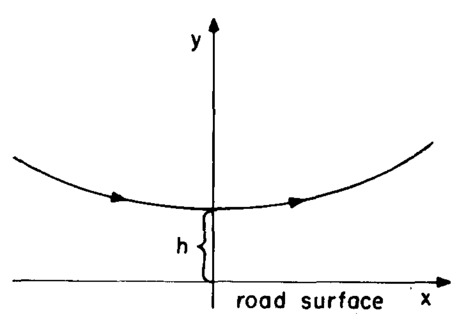

Fig. 4. Refraction of light above a warm road surface.

These conditions require

$$B = 0, \quad C = h + \frac{n_0}{\alpha}.$$

The analog of the velocity is $d\mathbf{r}/da$. The magnitude of this quantity is equal to n. At $x = 0$, the "velocity" is horizontal, i.e., wholly in the x direction. Thus,

$$\left.\frac{dx}{da}\right|_{a=0} = n_h, \quad \left.\frac{dy}{da}\right|_{a=0} = 0,$$

where n_h stands for $n_0 + \alpha h$, the index of refraction at height h above the road surface. Application of these conditions to our solution gives

$$A = n_h, \quad D = 0.$$

Thus the parametric equations for the ray are

$$x = (n_0 + \alpha h)a,$$

$$y = \left(\frac{n_0}{\alpha} + h\right)\cosh(\alpha a) - \frac{n_0}{\alpha}.$$

Eliminating a yields

$$y = h + \frac{n_h}{\alpha}\left[\cosh\left(\frac{\alpha x}{n_h}\right) - 1\right].$$

The ray is thus a catenary: the ray hangs in the warm air just as a chain hangs under its own weight.[9]

V. CENTRAL-FIELD MOTION

The formalism for dealing with light trajectories in media with spherical symmetry is a standard part of graduate-level optics texts.[10–12] We treat the optical central-field formalism here because it runs exactly parallel to the corresponding formalism in mechanics but is unlikely to be familiar to those who are not opticians. Again, our goal is to formulate the analogy in simpler terms than usually encountered and to illustrate its applications to interesting problems which are nevertheless easy enough for undergraduates.

A. "Angular momentum"

Consider a spherically symmetric ball or cloud of gas (planetary or stellar atmospheres provide a mental picture of such a situation). In such a case the index of refraction is a function only of r, the radial distance from the center of symmetry. The "force" is then a central "force;" each light orbit lies in a plane containing the "force center;" moreover, the analog of the angular momentum is a constant of the motion. Expressed in terms of plane polar coordinates (r,θ), the angular momentum of a material particle of mass m is $L = mr^2\, d\theta/dt$. The corresponding optical quantity is

$$L = r^2 \frac{d\theta}{da} = r^2\theta'. \tag{13}$$

This expression for the "optical angular momentum" may be cast into an alternative and frequently useful form. Refer to Fig. 5(a). Draw the radius vector from the "force" center 0 to an arbitrary point P on the curvilinear ray. Define φ as the angle between the radius vector and the tangent to the ray. At point P the instantaneous "velocity" is $d\mathbf{r}/da$, and the component of the "velocity" transverse to the radius vector is $|d\mathbf{r}/da|\sin\varphi$, as shown in Fig. 5(b). Now, the magnitude of this "transverse velocity" is also just $r\theta'$, in analogy with the familiar $r\dot{\theta}$ of mechanics. Thus,

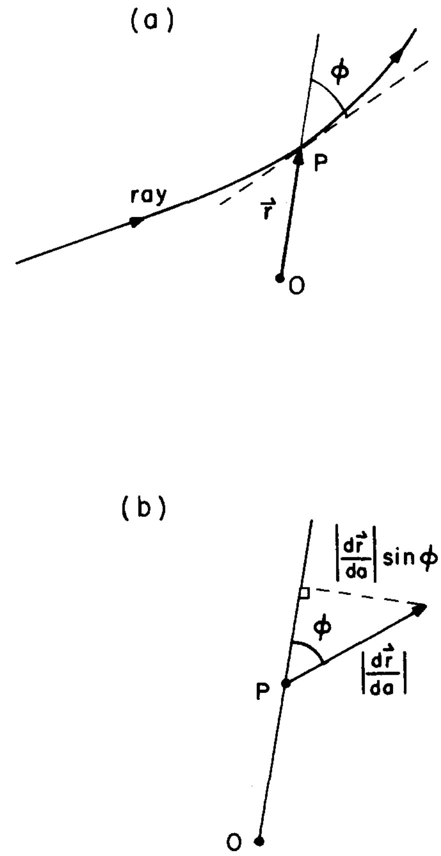

Fig. 5. Illustrating the formula of Bouguer.

Eq. (13) may be written

$$L = r\left|\frac{d\mathbf{r}}{da}\right|\sin\varphi,$$

or by use of the side condition [Eq. (11)],

$$L = r\,n\sin\varphi = \text{constant}, \tag{14}$$

a relation sometimes known as the formula of Bouguer. This form of the "angular momentum" is useful whenever one wishes to avoid any explicit use of the stepping parameter and to express L solely in terms of ordinary optical quantities.

B. Circular orbits

We seek the form of $n(r)$ that will permit circular orbits centered on the symmetry center. By analogy to the mechanical case, we may write down the condition for uniform circular motion ($-mv^2/r = -dU/dr$):

$$-\frac{1}{r}\left|\frac{d\mathbf{r}}{da}\right|^2 = -\frac{1}{2}\frac{dn^2}{dr}$$

$$= n\frac{dn}{dr}.$$

Upon substitution of the side condition $|d\mathbf{r}/da| = n$, we obtain

$$-\frac{dr}{r} = \frac{dn}{n},$$

which may be integrated to yield

$$n = k/r.$$

The "potential energy" is then $U = -k^2/2r^2$. And the "force law" is $F = -dU/dr = -k^2/r^3$. Thus, in the optical case, circular orbits are possible only for the r^{-3} "force law."[13]

This result requires a comment. In mechanics, any attractive central-force law will permit circular orbits, which may or may not be stable. The condition for the existence of the orbit, $mv^2/r = F(r)$, can be met simply by starting the particle at the correct speed for the radius of the chosen orbit. In the case of optics, the speed is already determined by the index of refraction, which is itself a function of r alone. Thus it develops that only in the case of the r^{-3} "force" can the condition for circular orbits be satisfied.

As a point of interest, we now calculate the orbital period associated with the circular orbits for the $n = k/r$ situation. The speed of light is $v = c/n = cr/k$. The circumference of the orbit is $2\pi r$. The orbital period is thus

$$\tau = \frac{2\pi k}{c},$$

independent of the radius of the orbit. Here we can see Fermat's principle directly—if smaller circles took less time, the light rays would dive inward to minimize the time of flight. A circular light orbit can only exist when the orbital period is stationary with respect to a change in the radius. If circular orbits are to be possible at all radii, then the period can have no variation with radius. (Note that in the mechanical case, in the r^{-3} force field, the period of circular orbits is proportional to the square of the radius.)

C. General orbit for $n = k/r$

As the $n = k/r$ case is of some interest, we calculate the most general orbit allowed. The calculation runs parallel to that for central-force motion for material particles. The analog of the energy is

$$E = \tfrac{1}{2}|\mathbf{r}'|^2 - \tfrac{1}{2}n^2 = 0,$$

where, as always, the prime denotes differentiation with respect to the stepping parameter a. Writing $|\mathbf{r}'|^2$ out in plane polar coordinates, we have

$$r'^2 + r^2\theta'^2 - n^2 = 0. \qquad (15)$$

Another constant of the motion is the "angular momentum," obtained above:

$$L = r^2\theta'.$$

Using L to eliminate θ' in Eq. (15) yields:

$$r' = \left(n^2 - \frac{L^2}{r^2}\right)^{1/2}.$$

Then, writing $r' = (dr/d\theta)(d\theta/da) = \theta' dr/d\theta$, we obtain

$$\int d\theta = \int \frac{L\,dr}{r^2(n^2 - L^2/r^2)^{1/2}}, \qquad (16)$$

in close analogy with a familiar mechanical result.

Now we suppose that $n = k/r$, the case at hand. Then

$$\theta = \int_{r_0}^{r} \frac{L\,dr}{r(k^2 - L^2)^{1/2}} + \theta_0.$$

Upon integration, and putting $\theta_0 = 0$ so that $\theta = 0$ when $r = r_0$, we have

$$r = r_0 \exp\left(\frac{(k^2 - L^2)^{1/2}}{L}\,\theta\right), \qquad (17)$$

an exponential spiral, a familiar result from mechanics. It remains to reexpress L in terms of k and the launching angle β defined in Fig. 6. By the formula of Bouguer [Eq. (14)],

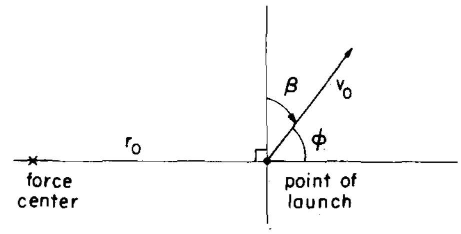

Fig. 6. Geometrical quantities involved in the central-force formalism.

$$L = r_0 n_0 \sin\varphi_0$$
$$= r_0 n_0 \cos\beta,$$

where n_0 denotes the index of refraction at the point of release, where the radius is r_0. For the case at hand $n = k/r$, so that $L = k\cos\beta$. The formula [Eq. (17)] for the exponential spiral then becomes, with a little manipulation,

$$r = r_0 \exp(\theta \tan\beta).$$

Note that if $\beta = 0$, $r = r_0$ and we recover the circular orbit derived above. If $0 < \beta < \pi/2$, $\tan\beta$ is positive and the trajectory is an expanding spiral. If $-\pi/2 < \beta < 0$, $\tan\beta$ is negative and the trajectory is a collapsing spiral. The circular orbit is therefore not stable, but lies between the expanding and collapsing spirals.

D. Light orbits in the r^{-2} *"force field"*

The case of the r^{-2} "force law" is instructive because of the analogy to planetary motion. We have then

$$n = (2k/r)^{1/2},$$

so that $U = -n^2/2$ becomes $U = -k/r$, as required.

From the general result of Eq. (16) we have

$$\theta = \int \frac{L\,dr}{r^2\left(\frac{2k}{r} - \frac{L^2}{r^2}\right)^{1/2}} + \text{const.}$$

This familiar integral is easily handled with the substitution $u = 1/r$. (See any intermediate-level treatment of planetary orbits.) The result is

$$r = \frac{L^2}{k(1 + \cos\theta)}.$$

The orbit of the light is thus a parabola with its focus at the origin. The constant of integration has been chosen to make r minimum at $\theta = 0$.

In the mechanical problem, all the conic sections are obtained—hyperbolas and ellipses as well as parabolas. The unbound hyperbolic orbits correspond to positive total energy; the bound elliptical orbits to negative energy (using the convention in which the potential energy vanishes at infinity). In the optical problem, only the zero-"energy" parabolic orbits are allowed, which confirms a statement made above.

E. A remark on similarities and differences

That the optical and the zero-energy mechanical orbits must have the same form for the same "force law" follows immediately from the fact that in seeking r as a function of θ, we eliminate the stepping parameter from the problem.

As the optical formalism differs from the mechanical only by the substitution of a for t, once the stepping parameter is eliminated it matters not at all whether this parameter was a or t.

The case is different, however, when we seek the position as a function of *time*. Here we see that the light pulse moves along its trajectory at a different rate than does the particle—although the trajectories are identical in form. For an illustration, take the central-field problem. In the case of a material particle the motion is governed by Kepler's second law or, equivalently, by the conservation of angular momentum:

$$\tfrac{1}{2}r^2\dot{\theta} = \text{areal velocity} = \text{constant}, \qquad \text{(mechanics)}$$

the constant being $L/2m$. For the optical case, the constant of the motion corresponding to the angular momentum is $L = r^2\theta'$. But we want to know how θ varies with t and not with a. Using the side condition [Eq. (12)] to pass over from $d\theta/da$ to $d\theta/dt$ in the expressions for the "angular momentum" gives

$$\tfrac{1}{2}r^2\dot{\theta} = \frac{Lc}{2n^2}, \qquad \text{(optics)}$$

so the areal velocity is not constant if n varies with r.

In the case of the r^{-3} "force" ($n = k/r$), the conservation of "angular momentum" results in

$$\dot{\theta} = \frac{Lc}{k^2} = \text{const},$$

so that the light pulse moves on its trajectory at constant angular speed, a result far different from constant areal velocity. In the r^{-2} "force field," the transverse velocity $r\dot{\theta}$ is constant.

F. Luneberg lens

The Luneberg lens is a sphere of radius r_0 having an index of refraction that is 1 at the edge and increases toward the center according to the formula $n = (2 - r^2/r_0^2)^{1/2}$. As first shown by Luneberg[14] in 1944, this lens has the property of focusing all parallel rays at the same point—there is no spherical aberration. Since the index of refraction depends only on the radius, this medium corresponds to a central force, and in fact the "force," $n\,\mathbf{grad}\,n$, is

$$n\,\mathbf{grad}\,n = -\frac{1}{r_0^2}\mathbf{r}.$$

In "$F = m\,a$" optics, the Luneberg Lens is evidently the analog of a harmonic oscillator, with "force constant" $1/r_0^2$. As usual, the "mass" is the dimensionless number 1. In cartesian coordinates, the general solution for a ray trajectory is thus

$$x(a) = A\sin\left(\frac{a}{r_0} + \alpha\right),$$

$$y(a) = B\sin\left(\frac{a}{r_0} + \beta\right),$$

where the "amplitudes of vibration" A and B and the phases α and β must be determined by matching initial conditions.

We consider a ray incident parallel to the x axis and a distance b from it, as shown in Fig. 7. We set $a = 0$ as the light enters the lens. The initial conditions for the y motion

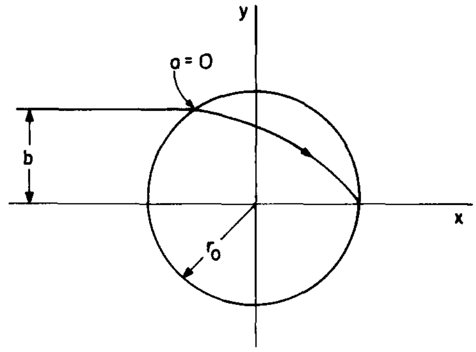

Fig. 7. Luneberg lens.

are

$$\left.\frac{dy}{da}\right|_{a=0} = 0, \quad y_{a=0} = b,$$

from which follow

$$\beta = \pi/2, \quad B = b.$$

The initial condition for the x "velocity" is

$$\left.\frac{dx}{da}\right|_{a=0} = 1,$$

which yields

$$r_0 = A\cos\alpha.$$

Our solution is thus,

$$x = A\sin\left(\frac{a}{r_0} + \alpha\right) \quad \text{with} \quad r_0 = A\cos\alpha, \tag{18}$$

$$y = b\cos\left(\frac{a}{r_0}\right). \tag{19}$$

We now ask where the ray crosses the x axis—what is x when $y = 0$? From Eq. (19) we find that $y = 0$ when the stepping parameter $a = \pi r_0/2$. Then $x = A\sin(\pi/2 + \alpha) = A\cos\alpha$, which the initial conditions have told us is r_0. Thus the ray crosses the x axis at the edge of the sphere, independent of how far the ray was from the x axis when it entered the lens. All parallel rays thus converge to the same point at the back edge of the lens.[15]

VI. TRIAL BY COMBAT

One of the authors taught this material recently in an junior-level mechanics class for which the text was the well known book by Marion. The "$F = m\,a$" optics was introduced following the textbook chapters on the calculus of variations, Lagrangian dynamics, and central-force motion. Some of the problems above were presented in class and others were assigned for homework. As we have shown, many problems involving ray optics can be rendered formally quite similar to very familiar problems in classical dynamics. A very small investment in new mathematicals tools therefore extends the student's powers from material particles to light rays—a fact that all but the sleepiest students seem to appreciate. The resolution of familiar problems in an unfamiliar context proved to be an interesting exercise.

[1] J. B. Marion, *Classical Dynamics of Particles and Systems*, 2nd ed. (Academic, New York, 1970).

[2]K. R. Symon, *Mechanics*, 3rd ed. (Addison-Wesley, Reading, MA, 1971).

[3]Pierre Louis Moreau de Maupertuis, *Essai de Cosmologie*, 1759, quoted by W. Yourgrau and S. Mandelstam in *Variational Principles in Dynamics and Quantum Mechanics*, 3rd ed. (Dover, New York, 1968), p. 20.

[4]See, for example, Max Born and Emil Wolf, *Principles of Optics*, 6th ed. (Pergamon, Oxford, 1980), pp. 101–132.

[5]Born and Wolf (Ref. 4), pp. 738–746, provide a standard treatment from the optical point of view. A standard treatment on the mechanical side is that of H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA, 1980), pp. 484–492. See also Yourgrau and Mandelstam (Ref. 3), pp. 58–64.

[6]Three recent examples are provided by the following. R. J. Black and A. Ankiewicz, "Fiber-optic analogies with mechanics," Am. J. Phys. **53**, 554–563 (1985). J. W. Blaker and M. A. Tavel, "The application of Noether's theorem to optical systems," Am. J. Phys. **42**, 857–861 (1974). W. B. Joyce, "Classical particle description of photons and phonons," Phys. Rev. D **9**, 3234–3256 (1974).

[7]This point has also been stressed by J. A. Arnaud, "Analogy between optical rays and nonrelativistic particle trajectories: A comment," Am. J. Phys. **44**, 1067–1069 (1976).

[8]We have derived Snell's law using the constancy of the component of $d\mathbf{r}/da$ parallel to the boundary, i.e., we have used $(dx/da)_1 = (dx/da)_2$. It is worth remembering that, although these quantities play the roles of velocities, they are not really velocities. The actual x component of the velocity is $v_x = (dx/da)(da/dt) = (c/n^2)dx/da$. The constancy of dx/da across the boundary therefore implies $v_{x1}n_1^2 = v_{x2}n_2^2$, so that the horizontal velocities are definitely not the same in the two media.

[9]In this problem we have supposed the *index of refraction* to vary linearly with the height above the road; as shown, the ray is a catenary. Alternatively, one may take the *speed of light* to vary linearly with the height. The results then are that the ray is a circular arc, with the center located at the height, where the speed of light goes to zero. (Our thanks to F. Danes for pointing this out.) This alternative problem is most easily solved by starting from the conservation of "energy," i.e., the equation that is the last entry in Table I.

[10]Born and Wolf (Ref. 4).

[11]M. V. Klein, *Optics* (Wiley, New York, 1970).

[12]E. W. Marchand, *Gradient Index Optics* (Academic, New York, 1978).

[13]The question we have posed is, more strictly, this: For what function $n(r)$ will there exist a circular orbit centered on the origin *for every r*. The condition we obtained, $rdn = -ndr$, is satisfied for all r only by the function $n = k/r$. Only in this particular case, then, do such circular orbits exist for all radii. There are, however, an infinity of possible functions $n(r)$ for which the condition $rdn = -ndr$ is satisfied *at one or more particular values for r*. In such a case, a circular orbit centered on the orgin will be possible, but only at particular, isolated radii. An example of such a case is provided by the Maxwell "fish-eye," i.e., by the function $n(r) = n_0[1 + (r/b)^2]^{-1}$. As is well known, the general light orbit in this system is a circle whose center is displaced from the origin. The off-centeredness of the circular orbit depends upon n_0 and b, as well as upon the initial conditions. [Born and Wolf (Ref. 4), pp. 147–149.] However, for the one particular case $r = b$, the off-centeredness vanishes. Indeed, it may be verified by direct calculation that for the Maxwell fish-eye, $dn/dr = -n/r$ only at $r = b$.

[14]R. P. Luneberg, *Mathematical Theory of Optics*, Brown University mimeographed notes, 1944 (University of California, Berkely, CA, 1964).

[15]This problem provides an instance in which the optical "potential energy," $-n^2/2$, does not correspond to the usual choice in mechanics. We have $U = r^2/(2r_0^2) - 1$ in the optical case, while the usual choice in mechanics is $kr^2/2$. This constant shift in the scale of the "potential energy" does not, of course, alter the trajectories. However, our usual convention of taking $-n^2/2$ as the analog of the potential energy results, as usual, in a "total energy" of zero, thus producing a paradox—how can there be motion in the case of the harmonic oscillator if the "total energy" is zero? The resolution of the paradox simply involves the choice of the zero of "potential energy." If the optical "potential energy" were defined exactly as is customary in the mechanical case, the "total energy" would be 1.