Reprinted from:

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 65, NUMBER 2 FEBRUARY 1975

Application of the mechanical theory of light to fiber optics*

J. A. Arnaud

Bell Telephone Laboratories, Crawford Hill Laboratory, Holmdel, New Jersey 07733

(Received 3 July 1974)

The motion of a charged particle in a space- and time-dependent potential and the motion of an optical pulse in an inhomogeneous anisotropic medium coincide when the dispersion surfaces are the same. For the trajectories in space to coincide, it is sufficient that the wave vectors be proportional. It follows from the general expression of the average stress-energy density $(\partial\bar{L}/\partial\vec{K})\vec{K}$, where $\bar{L}$ denotes the average lagrangian density and $\vec{K}$ denotes the 4 wave vector, that radiation forces are proportional to wave vectors for charged particles as well as for optical pulses. Because of these relations, many results in mechanics are applicable to optics. In particular, the constancy of the horizontal component of the velocity of a bullet on earth has, as a counterpart in optics, the constancy of the axial component of the group velocity of optical pulses propagating in thick tapered dielectric slabs. It follows from this observation that thick tapered dielectric slabs are not suited for long-distance communication, because of the large pulse spreading that they introduce. Slabs with moderate thickness, however, may exhibit low pulse spreading.

Index Headings: Fibers; Communication optics.

We are concerned, in this paper, with the transmission of energy and information in the form of optical pulses propagating through glass fibers. Radiation forces are given detailed consideration, both for their own sake and to clarify the relationship between matter waves and optical waves. Results derived in one field of knowledge can then be applied to the other in a rather straightforward manner.

The similarity that exists between the Schrödinger equation that describes spinless particles in the nonrelativistic approximation and the scalar parabolic wave equation that describes optical or acoustical beams propagating at a small angle to some axis was discussed in detail by Fock.[1] We need only replace, in the Schrödinger equation, time t by the axial coordinate z, the ratio $m/\hbar$, where m denotes the mass of the particle and $\hbar$ the Planck constant divided by 2π, by the wave number on axis $k(0) \equiv k_0$, and $-eU(x)/\hbar$, where e denotes the electric charge of the particle, and $U(x)$, the electric potential, by the optical wave number $k(x)$. This relation was presented in Ref. 1 as purely formal; the propagation of optical pulses was not discussed. The mechanical and optical problems are, in fact, identical. Two difficulties prevented, until recently, a complete understanding of the problem. A general expression for the force exerted by a wave on an absorber located in a polarizable fluid was lacking. This result is now provided by Whitham's theory.[2] The expression of the canonical stress-energy density $(\partial\bar{L}/\partial\vec{K})\vec{K}$, where $\bar{L}$ denotes the average lagrangian density and $\vec{K}$ the 4 wave vector, shows that the force exerted by matter waves or by optical waves on absorbers is proportional to the wave vector $\vec{k}$. The second difficulty was associated with the inertia of light. Classical electrodynamics, as it is usually presented, is inconsistent with special relativity.[3] This difficulty has been solved by Penfield and Haus,[4] who have shown that the relativistic changes of mass of charge carriers in magnetizable bodies are essential to a consistent formulation (see also Ref. 5). Once a clear distinction has been made between canonical momenta (proportional to the wave vector) and mass-carrying momenta (proportional to the group velocity), no essential difficulty remains.[6–9] It is fair to say, however, that different points of view are expressed in the literature.[10–13] We have attempted to clarify the theory by discussing a number of simple but crucial experiments. New results of practical value in fiber optics are derived, in part from a comparison with similar results in mechanics.

Analogies of different natures have been drawn in the past between matter and light. We consider first the Descartes analogy, which compares trajectories in isotropic spaces. This analogy seems simple at first, but its interpretation is, in fact, difficult because time and space are entangled. The discussion in the next section is not complete. We attempt only to clarify the assumptions that are essential to the analogy. Rather deep concepts are needed for a full understanding.

I. RAYS IN ISOTROPIC MEDIA

Let us associate to any ray a vector $J\vec{k}$, called the ray canonical momentum, which has the direction of the ray. J denotes a quantity that remains constant along any given ray. The motivation for the notation $J\vec{k}$ will appear later. Similarly, the canonical energy is denoted $J\omega$. The magnitude of $J\vec{k}$ is assumed to depend only on the medium in which the ray propagates. It is further assumed that the intrinsic properties of the ray (e.g., the frequency of the disturbance) remain the same as the incidence angle i is varied. The Descartes–Snell law of refraction, which states the constancy of $\sin(i)/\sin(i')$ at a plane interface as i varies, clearly follows from the invariance of the tangential component of the vector $J\vec{k}$ (see Fig. 1b). To make the invariance of the tangential component of $J\vec{k}$ plausible, Descartes noted a similar situation in mechanics. Consider a ball traversing a breakable sheet, or rubber band (Fig. 1a). Because the rubber band does not exert any force on the ball in the horizontal direction, the horizontal component of the momentum of the ball, $J\vec{k} = m\vec{u}$, is invariant. On the other hand, the energy absorbed by the rubber band before it is broken loose is independent of the angle of incidence of the ball, as we can see from a quasistatic analysis. If we assume is-

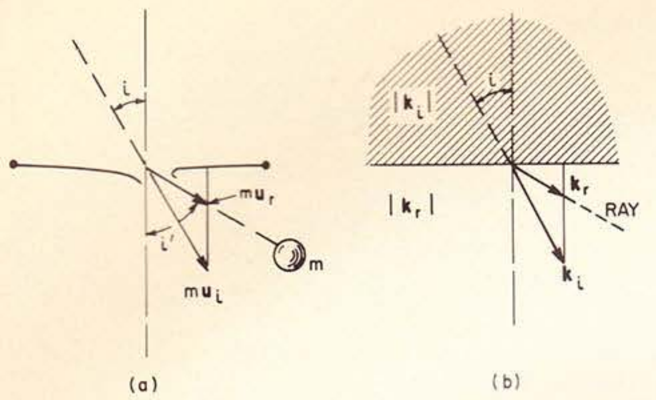

FIG. 1. Illustration of the Descartes mechanical analogy for the refraction of light rays. (a) A ball traverses a sheet that reduces its momentum $m\vec{u}$ by a factor independent of the angle of incidence. The tangential component of $m\vec{u}$ is invariant because the sheet does not exert a force on the ball in that direction. (b) A light ray is refracted away from the normal by going to a less-dense medium. The tangential component of a vector $\vec{k}$, whose magnitude is independent of the incidence angle, is invariant. $\sin(i)/\sin(i')$ is a constant in both cases.

ropy of space, the energy of the ball depends only on the magnitude of the momentum $J\vec{k}$. It follows that the ratio $|J\vec{k}'|/|J\vec{k}|$ of the magnitudes of the momenta of the ball after and before it traverses the sheet is independent of the incidence angle. From the geometric construction in Fig. 1, the law $\sin(i)/\sin(i') = \text{const}$ follows.

Let us now generalize these results to continuous media. Because only proportionality between the canonical momenta is required for the analogy to hold, and because the canonical momentum of a particle is $m\vec{u}$, where m denotes the mass and $\vec{u}$ the velocity, the mechanical analogy just described calls for the correspondence

$$u \rightleftarrows n \, , \tag{1}$$

where n denotes the refractive index of the medium, defined as $|J\vec{k}|_{\text{med}}/|J\vec{k}|_{\text{vac}}$. From Newton's nonrelativistic dynamics, the trajectory in space-time $\vec{x}(t)$ of a particle in a gravitational potential $V(\vec{x})$ obeys the differential equation

$$\frac{d^2(\vec{x})}{dt^2} = -\vec{\nabla}V(\vec{x}) \, . \tag{2}$$

If the total energy of the particle is taken equal to zero, the magnitude u of its velocity is given by $\frac{1}{2}mu^2 = -mV(\vec{x})$. Thus, if we replace u by n, according to Eq. (1), we can rewrite Eq. (2)

$$\frac{d^2(\vec{x})}{dt^2} = \vec{\nabla}[\tfrac{1}{2}n^2(\vec{x})] \, . \tag{3}$$

This is the equation for light rays (of a given color) in continuous media. This analogy is exemplified in Fig. 2, where we compare the trajectory of a ball in a shallow gutter and the trajectory of a light ray in a graded-index fiber or a nonuniform waveguide. If V (resp. k^2 or n^2) is quadratic in the transverse coordinate x but independent of the axial coordinate z, the rays are exactly sinusoidal in both systems. In Eq. (3), t does not, in general, represent the time of flight of optical pulses. What we are comparing here are trajectories in space. Time-dependent concepts have been used in the discussion because the dispersive properties of particles happen to be known from Newton's dynamics. The dispersive property of the refractive medium, however, remains undefined. That is, the dispersion of the refractive medium may or may not be the same as that of massive particles. Most likely, it will be different. The pulse velocities would then be different, too.

Let us now consider a modern version of the Descartes analogy, applicable to relativistic particles. Consider a particle with charge e traversing two closely spaced grids that have a difference of potential V between them, as shown in Fig. 3a. Let $\vec{X}(\tau)$ represent the particle trajectory, where $\vec{X} = \{\vec{x}, ict\}$ and τ is the proper time with $d\tau = i\,|d\vec{X}|$. Setting $c = 1$ for simplicity, we have, by definition

$$\left(\frac{dx_1}{d\tau}\right)^2 + \left(\frac{dx_2}{d\tau}\right)^2 + \left(\frac{dx_3}{d\tau}\right)^2 - \left(\frac{dt}{d\tau}\right)^2 + 1 = 0 \, . \tag{4}$$

The canonical momentum of the particle, whose tangential component is invariant at a plane interface, can, in the present case, be taken equal to the mass-carrying momentum

$$J\vec{k} = m(d\vec{x}/d\tau) \, . \tag{5}$$

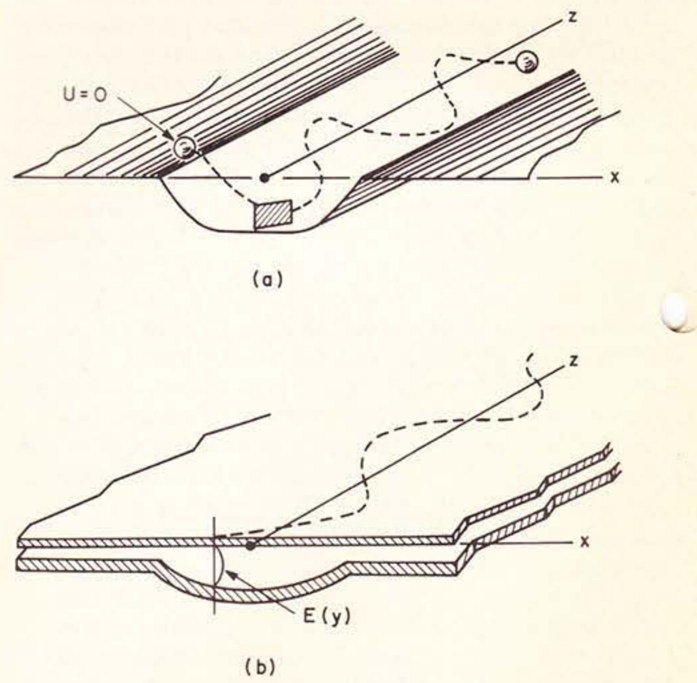

FIG. 2. (a) Motion of a steel ball in a gutter. The ball is assumed to have zero velocity (zero energy, by definition) when located on the top. The ball falls to the bottom of the gutter where it is deflected by a plate. Its subsequent motion is defined by the gravitational potential $V(x)$, which corresponds to the profile of the gutter if the gutter is sufficiently shallow. The axial velocity is a constant of motion. (b) Waveguide whose wall-to-wall spacing $2d$ varies slowly with x. The ray trajectories (in space only) are the same as in (a) when the local wave vector squared $k^2(x)$ is proportional to $V(x)$.

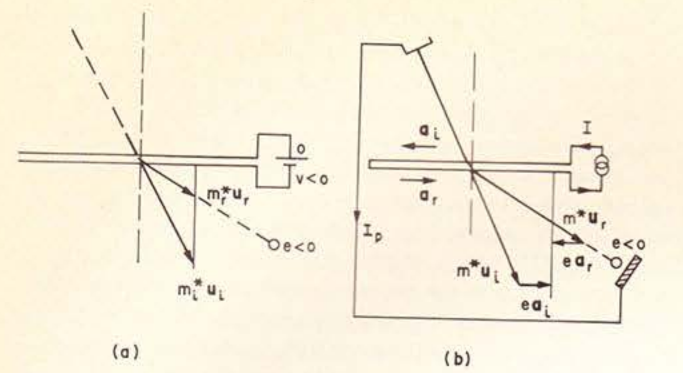

FIG. 3. (a) A modern version of the Descartes mechanical analogy, the sheet being replaced by a pair of electrodes. $m^* = m - eV$ denotes the moving mass. The ratio of the sines of the angles to the normal is a constant, as a result of isotropy. (b) The charged particle traverses a double sheet of current, creating a discontinuity in potential vector $\vec{a}$. The tangential component of the canonical momentum $J\vec{k} \equiv m^*\vec{u} + e\vec{a}$ is invariant. The momentum transferred to the sheet is opposite to the change in $m^*\vec{u}$, as can be seen by considering a steady flow of charge (current) from the emitter to the collector.

When the grids are traversed, a constant energy eV is subtracted from the canonical energy $J\omega$, taken equal to the mass energy

$$J\omega = m(dt/d\tau) \, . \tag{6}$$

Thus the ratio of the magnitudes of the canonical momenta is, by use of Eq. (4),

$$\frac{|J\vec{k}'|}{|J\vec{k}|} = \frac{|d\vec{x}'/d\tau|}{|d\vec{x}/d\tau|} = \left[\frac{(dt'/d\tau)^2 - 1}{(dt/d\tau)^2 - 1}\right]^{1/2} = \left[\frac{(J\omega - eV)^2/m^2 - 1}{(J\omega)^2/m^2 - 1}\right]^{1/2} . \tag{7}$$

Because this ratio is independent of the incidence angle, $\sin(i)/\sin(i')$ is a constant, as is the case for nonrelativistic particles. Two essential postulates were made in the above discussion: medium isotropy and invariance of the canonical momentum in directions of translational invariance of the medium. In a wave theory, the latter follows from the proportionality of the canonical momentum to the wave vector. Thus, the law of refraction follows quite generally from isotropy and translational invariance of the medium. A generalized form of the law of refraction is needed if we omit the assumption of isotropy. This is discussed in the next section.

II. RAYS IN ANISOTROPIC MEDIA

The essential concepts of ray propagation appear most clearly when the medium lacks isotropy. Let us consider a charged particle traversing two closely spaced grids carrying equal and opposite current densities, as indicated in Fig. 3b. These sheets of current create a discontinuity in the potential vector $\vec{a}$, which is otherwise uniform. The refraction of the charged particle follows from the invariance of the tangential component of the canonical momentum,

$$J\vec{k} = m\frac{d\vec{x}}{d\tau} + e\vec{a} \, . \tag{8}$$

In the present case, the canonical energy $J\omega = m(dt/d\tau)$ is a constant, and the ratio $\sin(i)/\sin(i')$ varies with i. This is because the potential vector creates an anisotropy in space. Note that, even in the absence of a magnetic field, it is permissible to add to $m(d\vec{x}/d\tau)$ in Eq. (8) a term of the form $\vec{\nabla}f(\vec{x})$, where f is an arbitrary scalar function of $\vec{x}$. This has no consequence, however, on the ray trajectories, nor on the interference patterns. In what follows, we assume that $\vec{a}$ is defined as the volume integral of the current density divided by the distance to the observation point.

III. EXPERIMENTS IN DYNAMICS

We need, first, to clarify the difference between the canonical momentum $J\vec{k}$ and the mass-carrying momentum $m(d\vec{x}/d\tau)$, introduced in the previous section. This difference is essential to understanding of the hamiltonian formalism. However, these two momenta are still sometimes confused for light waves. For example, Marcuse[12] wrote that if a light pulse goes through a dielectric slab free of loss and of Fresnel reflection, the slab is displaced toward the source of light. As noted in Ref. 3, this conclusion violates the law of mass-energy equivalence. The error is to use the canonical momentum ($J\vec{k}$) of the light pulse in a problem in which the mass-carrying momentum $m(d\vec{x}/d\tau) = E(d\vec{x}/dt)$, where E denotes the pulse energy, should be used. The magnitude of the canonical momentum increases as the light pulse enters from vacuum into a medium with $n > 1$. The magnitude of the mass-carrying momentum, on the contrary, decreases, because the group velocity $|d\vec{x}/dt|$ in a lossless medium is always less than c. The correct expression of the displacement of the dielectric slab follows from the invariance of the sum of the mass-carrying momenta of the optical pulse and of the slab, with respect to an inertial reference system. The dielectric-slab displacement is easily found to be always in the forward direction and equal to $Em_0^{-1}(u^{-1} - 1)$, where m_0 denotes the mass per unit length of the slab, and u the group velocity in the slab. The slab displacement can alternatively be evaluated, using classical electrodynamics, from the force exerted by the magnetic field of the incident wave on the polarization currents induced by the electric field. A difficulty was noted for magnetizable bodies that resulted from a basic inadequacy of classical electrodynamics. This difficulty has been resolved recently by the introduction of a force of relativistic origin, called the magnetodynamic force.[3-5] The displacement of the slab would be exceedingly difficult to measure. An experimental technique using low-loss glass fibers, however, is suggested in Ref. 7. The mechanical system analogous to the dielectric slab is the potential box, shown in Fig. 4a. If the potential in the box is negative (for an electron, with $e < 0$), the box is displaced in the forward direction as in the optical case.

If the light pulse enters under oblique incidence in the refractive medium, a momentum is transfered to the medium in the tangential direction because the tangential component of the mass-carrying momentum, unlike the tangential component of the canonical momentum, is not a constant. This tangential force is not present in

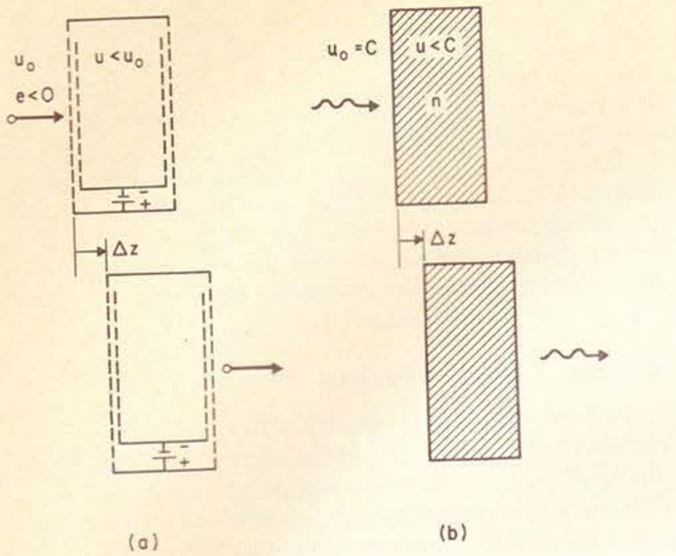

FIG. 4. Displacement of a slab traversed by a particle. Losses and reflection are neglected. (a) An electron traverses a potential box. Because u decreases as the electron enters the box for the voltage shown, the box displacement is forward. (b) In the optical case the slab displacement is always forward ($u<c$). In such arrangements, only the mass-carrying momentum is relevant.

the case of the potential box, because the mass-carrying momentum happens to equal the canonical momentum in that special case.

Unlike mass-carrying momenta, canonical momenta are defined with respect to (noninertial) reference systems attached to the medium, or to the source of the potential vector for the case of charged particles. Because the reference system is not inertial, conservation of the total momentum is not expected to hold, except in directions of translational invariance. To understand this point, consider first a simple mechanical system: with respect to earth (which cannot be considered inertial no matter how small its acceleration may be), the momentum of a bouncing ball is not invariant, because the ball can rebound off the ground. However, the horizontal component of the ball momentum is invariant if the ground is flat and smooth. For massive particles in an electric potential (or for a waveguide) the mass-carrying momentum happens to be equal to the canonical momentum. The canonical momentum can differ from the mass-carrying momentum only if a third body (a medium) is present that can absorb the difference in momenta. Such a medium is clearly absent for a free particle, a cold plasma, and a smooth uniform waveguide. This is the physical connection between these three seemingly unrelated systems.

In the presence of a potential vector, the two momenta are not the same. To clarify this point, let us go back to the situation described in Sec. II. As the charged particle moves through the current sheets, the tangential component of its canonical momentum is invariant. The momentum transferred to the current sheets, however, is opposite to the change of the mass-carrying momentum of the particle. This mass-carrying momentum has a component in the plane of the sheet generally different from zero. With respect to the inertial system, we are dealing with two objects: the charged particle on the one hand, and the current carrier on the other. No electromagnetic momentum should be considered in writing the conservation law. When the reference system is attached to the source of current, an interaction electromagnetic momentum needs to be taken into account. It is obtained by integrating $\vec{D}\times\vec{B}$, where $\vec{D}$ denotes the field of the charge and $\vec{B}$ the field of the current sheets, over the volume enclosed by the sheets ($\vec{B}$ being equal to zero outside the sheets). It is not difficult to show that this electromagnetic momentum is precisely the term $e\vec{a}$ in Eq. (8). Thus, in this noninertial coordinate system, the sum of the tangential particle and electromagnetic momenta is invariant.

Let us now discuss in some detail the problem of radiation force. Consider a light beam carrying a power P, absorbed in a liquid in which the phase velocity is $\omega/|\vec{k}|$ (see Fig. 5a). Experiments by Jones and Richard[14] have shown that the force exerted by the light beam on the absorber is correctly given in magnitude by

$$\vec{f} = (P/\omega)\vec{k} \; . \tag{9}$$

A simple derivation of Eq. (9) is through the Doppler

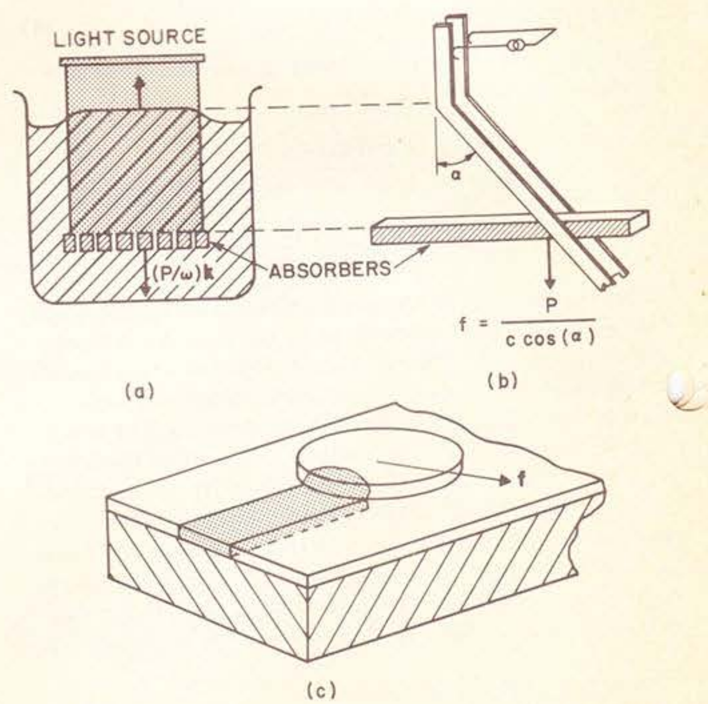

FIG. 5. Experiments in dynamics. (a) A light beam emitting a power P exerts on an absorber immersed in a fluid a force $(P/\omega)\vec{k}$ (Jones and Richards experiment). Perforations allow the absorber to move vertically in the fluid. The liquid is forced upward through the perforations and makes the surface of the liquid bulge out, in the steady state. (b) Equivalent model, in which the absorber can slide between two wires. These wires, representing the liquid with index $n=1/\cos\alpha$, make an angle α with the vertical. (c) Proposed experiment to verify that the force on an absorber need not have the direction of the beam if the medium lacks isotropy. The disk is a lossy dielectric that picks up the optical power from the film through tunneling.

effect[7] or the adiabatic invariance of the photon number in a resonator. These arguments, although correct, are not completely satisfactory because appeal must be made to the second quantization ($E=\hbar\omega$). The classical derivation, based on the Minkovski expression for the stress density is correct for incompressible fluids in the steady state.[5] For scalar waves, the result follows from the expression of the canonical stress density given in the introduction. The balance of momenta clearly shows that, in the steady state, the surface of the liquid bulges out (toward the less-dense medium), the surface of the liquid being submitted to a net force

$$f = -(P/c)(n-1) \; , \tag{10}$$

balanced by surface tension. This force also follows from the artificial dielectric model considered in Ref. 7 and illustrated in Fig. 5b. This bulging out is expected to take place after a time of the order of the time it takes for a sound wave to go from the absorber to the liquid surface. It has been suggested[6] that, for the case in which the light-beam cross section is small compared to the absorber-to-liquid surface spacing, the steady state is reached after a shorter time, of the order of the time it takes for a sound wave to move across the beam. Experiments by Ashkin *et al.*[15] suggest that indeed this shorter time is significant. Whether the final steady state is then reached is unclear to us. These transient effects have been mentioned because they have been subjects of some controversy, but they are not essential for our discussion.

Most liquids being isotropic, only the magnitudes of $\vec{f}$ and $\vec{k}$ can be compared. It is therefore interesting to consider the arrangement in Fig. 5c, where the light propagates in a thin anisotropic film, perhaps a few μm thick. An absorbing disk that can slide on top of the film is submitted, according to Eq. (9), to a force $\vec{f}=(P/\omega)\vec{k}$. This force, in general, does not have the direction of the beam. With presently available lasers, it may not be more difficult to perform this experiment than the original experiment of Jones and Richard. It may, in fact, be easier, because the radiometric forces that cause great experimental difficulties in the case of an absorber located in a gas or a liquid are absent in this new suggested experiment.

IV. WAVE PROPAGATION

We compare, in this section, the Schrödinger equation which is applicable to nonrelativistic particles, and the Fock equation, which is applicable to paraxial beams. The dispersion equation of an isotropic medium at some angular frequency ω is, ignoring for simplicity the y coordinate,

$$k_x^2 + k_z^2 = k^2(x) \; . \tag{11}$$

The (circular) dispersion curve is shown in Fig. 6a. By substitution of $i\vec{k} \rightarrow \vec{\nabla}$, the scalar Helmholtz equation is obtained,

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + k^2(x)\right]\psi = 0 \; . \tag{12}$$

The dispersion relation for particles with mass $m \equiv \hbar\overline{m}$ and charge $e \equiv \hbar\overline{e}$ in a potential vector $\vec{A} \equiv (\vec{a}, iV)$

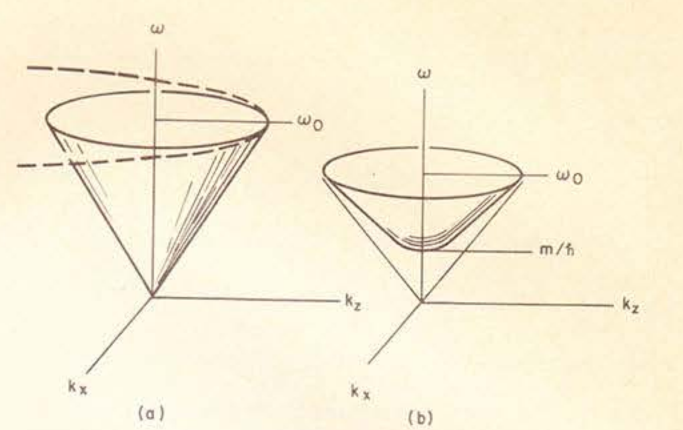

FIG. 6. Dispersion surfaces for optical waves in isotropic nondispersive media (a) and massive particle (mass m) in free space. The parabolic approximation used in beam optics consists of replacing the circle ($k_x^2 + k_z^2 - k^2 = 0$) at some $\omega=\omega_0$ by a parabola. This parabola is shown as a dotted line in (a). The nonrelativistic approximation in mechanics consists of replacing the hyperboloid in (b) by a paraboloid in the neighborhood of $\omega = m$. The laws of diffraction in free space and the law of spreading of pulses in dispersive media have the same general form. They depend on the curvatures of the dispersion surface in the k_x, k_z planes and k, ω planes, respectively.

is simply that the length of $\vec{K} - \overline{e}\vec{A}$ is invariant, $i\overline{m}$. Separating the space and time components and setting $\vec{a} = \vec{0}$, we have, at a fixed canonical energy $\omega = \overline{m}$,

$$k_x^2 + k_z^2 - \overline{e}^2 V^2(x) + 2\overline{m}\,\overline{e}\, V(x) = 0 \; . \tag{13}$$

Upon substitution of $i\vec{k} \rightarrow \vec{\nabla}$ in Eq. (13), the Klein–Gordon equation is obtained,

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \overline{e}^2 V^2(x) - 2\overline{m}\,\overline{e}\, V(x)\right]\psi = 0, \tag{14}$$

which is equivalent to Eq. (12) if

$$k^2(x) = \overline{e}^2 V^2(x) - 2\overline{m}\,\overline{e}\, V(x) \; . \tag{15}$$

For the nonrelativistic case, the term $\overline{e}^2 V^2$ can be neglected in Eq. (14). Because the total energy is equal to zero, we have $\overline{e}V + \frac{1}{2}\overline{m}u^2 = 0$. Thus, Eq. (15) requires the equality of k and $\overline{m}u$ for nonrelativistic particles, as we have indicated before.

Let us now introduce paraxial approximations in both Eq. (11) and Eq. (13). If we assume that k_x in Eq. (11) is small compared with k_z, we can write

$$k_z \approx k - (1/2k)k_x^2 \; . \tag{16}$$

With the same substitution as before, we obtain from Eq. (16) a scalar parabolic wave equation

$$-i\frac{\partial\psi}{\partial z} = \left\{k(x,z) + [1/2k_0(z)]\frac{\partial^2}{\partial x^2}\right\}\psi \; , \tag{17}$$

where, in the second term in the brackets, we have made the approximation $k(x,z) \approx k(0,z) \equiv k_0(z)$, on the ground that k does not vary very much with x. This equation is similar, though not identical, to the equation considered by Fock.[1] To make contact with the optical problem, let us indicate that in order for the reciprocity theorem to be applicable in a natural way to Eq. (17), ψ must be defined as $[k_0(z)]^{1/2}E$, where E denotes a transverse component of the electric field in the beam. For the case where k_0 is independent of z, the wave equation, Eq. (17), can be compared to the Schrödinger equation

$$-i\frac{\partial\psi}{\partial t}=\left[-\bar{e}U(x,z)+(1/2\bar{m})\frac{\partial^2}{\partial x^2}\right]\psi . \tag{18}$$

The correspondence between Eq. (17) and Eq. (18) is $t\rightarrow z$, $\bar{m}\rightarrow k_0$, and $-\bar{e}U(x,z)\rightarrow k(x,z)$. This is essentially the Fock analogy that was discussed earlier. If k_0 were a function of z, a similar correspondence could be established, merely by redefining the axial coordinate.

There is an even closer similarity between paraxial optical waves and paraxial electron waves. Let us consider a charged particle propagating in a direction close to the z axis. Because k_x is small compared with k_z, Eq. (13) can be written

$$k_z\approx(\bar{e}^2V^2-2\bar{m}\,\bar{e}V)^{1/2}-\tfrac{1}{2}(\bar{e}V^2-2\bar{m}\,\bar{e}V)^{-1/2}k_x^2 . \tag{19}$$

By use of the substitution $i\vec{k}\rightarrow\vec{\nabla}$, we obtain a parabolic wave equation. Assuming for simplicity that the axial motion is nonrelativistic, we obtain the wave equation

$$-i\frac{\partial\psi}{\partial z}=\left\{[-2\bar{m}\,\bar{e}V(x,z)]^{1/2}+\tfrac{1}{2}[-2\bar{m}\,\bar{e}V_0(z)]^{-1/2}\frac{\partial^2}{\partial x^2}\right\}\psi , \tag{20}$$

where, in the second term in the brackets, we have made the approximation $V(x,z)\approx V(0,z)\equiv V_0(z)$. Equation (20) describes the evolution of the transverse wave function ψ of paraxial particles. Its similarity with Eq. (17) is obvious. Equation (20) also has the form of the conventional Schrödinger equation (18) because $dz/dt\equiv u_0=(-2\bar{m}\,\bar{e}V_0)^{1/2}\bar{m}^{-1}$ is the axial velocity, and $-2\bar{e}(VV_0)^{1/2}\approx-2eV_0-eU$, where $U\equiv V-V_0\ll V_0$. The term $-2\bar{e}V_0$ corresponds to an unimportant phase factor on ψ that can be omitted. Our result simply means that, for nonrelativistic axial motions, the Schrödinger equation (18) is applicable in a frame of reference that moves at the mean velocity u_0 of the electron. If the axial motion were relativistic, only slight changes would be needed.

Let us illustrate the similarity of diffraction effects between optical waves and electron waves by considering the effect of grating-type devices. The discussion will be qualitative, but could easily be made quantitative with the previously derived expressions. The system shown in Fig. 7 is a sequence of metal tubes at increasing voltages. If a plane electron wave enters the tubes from the left, the electron is deflected upward, because the phase shift introduced by the voltage is (stepwise) linear in x. If a dc voltage is applied to the tubes, this phase shift results from a change of the wavelength of the electron, the frequency being constant because of the time invariance of the system. If the voltage is applied while the electron-wave packet is traveling through the tubes, the wavelength remains constant, because of the axial invariance of the system, but the frequency is changed. The phase shift, and therefore the deflection, is the same in both cases, if the voltage is applied during a time L/u, where L denotes the length of the tubes in the first experiment and u the velocity of the electron. It is interesting that in the pulsed arrangement, no classical force is exerted on the electron.

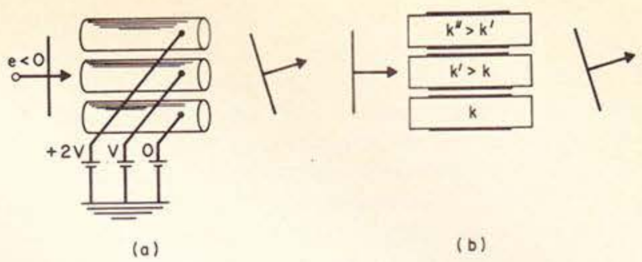

FIG. 7. Deflection of an electron or optical wave by wave-optics gratings that consist of metal tubes (a) and electro-optic crystals (b) respectively. In both the mechanical and optical systems, the wave packet is deflected upward, even though no classical force is exerted on the particle when the voltages are turned on and off while the wave packet is still traveling in the tubes or crystals (second Aharonov and Bohm effect).

Yet the electron is deflected away from its original direction. This effect is similar to the second Aharonov and Bohm effect discussed in Ref. 16. However, by considering a series of tubes rather than just two, we make it clear that a deflection of the average electron path as well as a shift of the interference pattern takes place.

Discussion of this effect for the optical deflector shown in Fig. 7b, which is essentially an optical phase array, would be identical with the foregoing. The only possible difference between the two systems would be that the dispersion of the electro-optic crystals used in the optical arrangement may not be the same as the dispersion of free space for matter waves. Thus, in the pulsed experiment, the pulse lag may be somewhat different.

V. OPTICAL FIBERS WITH MODE SELECTION

To illustrate the application of some of the results of the previous section to fiber optics, we investigate a class of optical fibers that has attractive properties for the transmission of information.

Consider the tapered dielectric slab shown in Fig. 8a. The thickness of the slab is maximum along some straight line (z axis) and decreases away from this line. The variation of thickness is assumed to be slow, so that the slab can be considered uniform in the neighborhood of some value of x. The local wave number $k(x)$ is given by the theory of uniform dielectric slabs. It is obtained by matching the tangential components of the electric and magnetic fields at the slab boundaries. As is well known, various modes can propagate in uniform slabs. We concentrate on one of them, e.g., the H_1 mode. Because of isotropy, the magnitude k of $\vec{k}$ is the same in all directions in the xz plane, the slab material itself being assumed isotropic. Once the wave number $k(x)$ has been obtained, we can make use of the results in ray optics and wave optics given in the previous sections. If $k^2(x)$ is, for example, quadratic in x, the rays are sinusoids, and they have almost all the same optical length.

A wave equation that takes diffraction into account in the xz plane can be set up, as we have seen, through

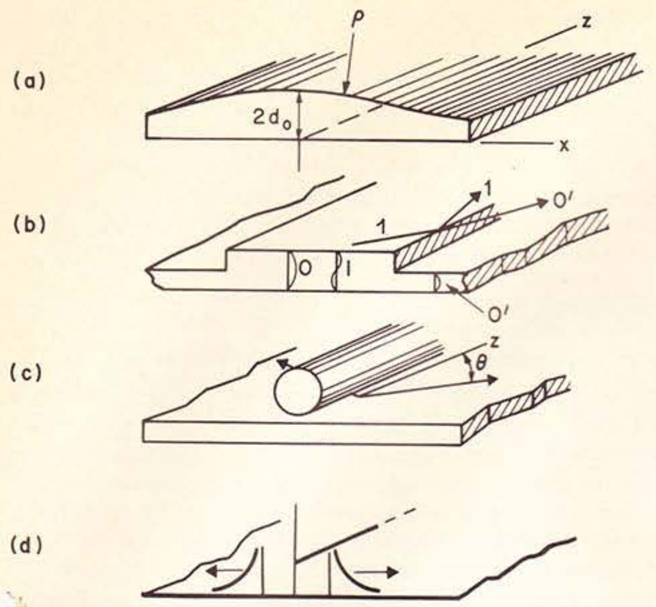

FIG. 8. Optical fibers with small pulse spreading. (a) Tapered dielectric slab. Pulse spreading is very small for certain profiles (see Fig. 11). (b) A step in the thickness introduces coupling between modes with different mode numbers in the y direction. For properly selected dimensions, only one mode is free of radiation loss (Ref. 22). (c) Coupling between a rod carrying trapped modes and a slab carrying radiation modes. (d) The configuration in (c) can be analyzed by replacing the rod by a distribution of electric and magnetic currents (only one line of current is shown) and evaluating the coupling to the surface waves.

the formal substitution $i\vec{k}\rightarrow\vec{\nabla}$ (xz plane) in the dispersion equation.[17] The variation of the field in the y direction (perpendicular to the plane of the slab), remains as given by the exact theory of the slab. The solution

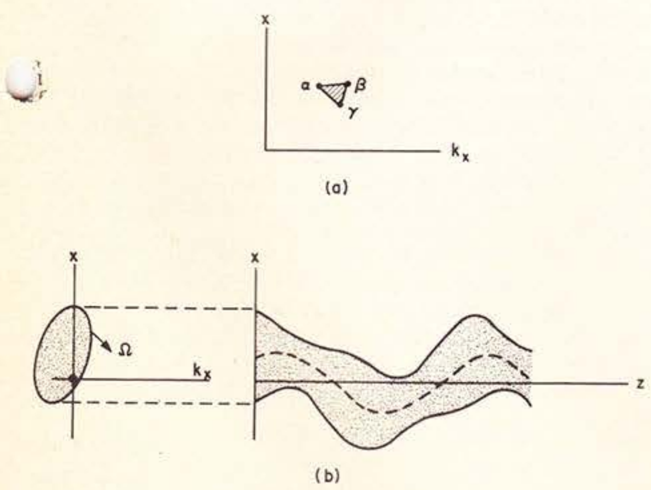

FIG. 9. (a) A ray, at some z, is represented by a point in the phase space k_x, x. The area enclosed by three neighboring rays, (α,β,γ) is invariant as z varies (Lagrange ray invariant). It follows that the density f of rays, for any continuous distribution, is invariant, $df/dz=0$ (Liouville theorem). (b) The propagation of gaussian beams in uniform square-law media is represented by an ellipse that rotates in phase space with a constant rate Ω. If the beam is injected off axis, the center of the ellipse is off-set.

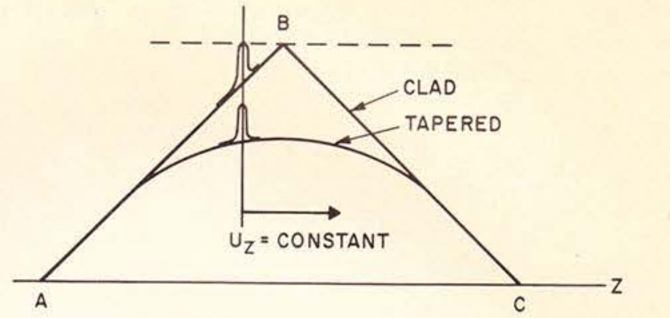

FIG. 10. For thick dielectric slabs, or metallic boundaries, and nonrelativistic particles, the theorem of Breit and Tuive, originally derived for cold plasmas, applies. The time of flight of a pulse can be obtained by extending the ray direction along a straight line from the origin, and assuming the medium homogeneous up to the reflection point. This is because the local group velocity is proportional the local k vector. Thus, $u_z\propto k_z$ is a constant of motion, that is, the pulse of light moves at a constant speed along the z axis.

of the wave equation with k^2 quadratic in x is well known in the quantum theory of harmonic oscillators. This is the product of a function of Gauss and a Hermite polynomial. The properties of square-law media need be recalled only briefly. In square-law media, the center of a beam follows a classical ray trajectory. In most text books of quantum mechanics this result is proved on the basis of the Erhenfest theorem. The complete expression for the transformed field, including the phase, obtained by Husini,[18] is much more useful for the optical problem. Conventional modes (or stationary states) have constant irradiances along the z axis. However, the size of beam modes oscillates along the axis (see Fig. 9b). A simple and useful representation of gaussian-beam modes is by a complex ray. The complex-ray representation of gaussian beams is easily generalized to space time to describe the spreading, in time as well as in space, of gaussian wave packets. To apply this representation, it is convenient to rotate the coordinate system from x,y,z,t, to x',y',z',t', with the t' axis directed along the normal to the dispersion surface $H(\vec{K})=0$ (see Fig. 6) and x',y',z' along the directions of principal curvature. In space time, the complex ray $q(z)$ becomes a 3×3 complex matrix $\mathfrak{Q}(t')$ which obeys a matrix ray equation.[19] This description ties the spreading of a gaussian pulse along a dispersive line to the diffraction of a gaussian beam in space. The formalism is the same as that given in Ref. 19 except for the increased number of transverse dimensions. A gaussian beam can alternatively be represented by an ellipse that rotates in phase space. This ellipse is the locus of the $1/e$ points of the Wigner distribution function (for a definition of the Wigner function, see Ref. 5, p. 326). Figure 9b shows that this is a convenient way to generate the beam profile.

In order to obtain information concerning the propagation of optical pulses, we need to know, not only $k(x)$, but also the variation of k with ω. If the ratio $(\omega/k)/(\partial k/\partial\omega)$ of the local phase velocity ($v=\omega/k$) to the local group velocity ($u=\partial\omega/\partial k$) happens to be independent of the x coordinate, the time of flight of a pulse along a ray

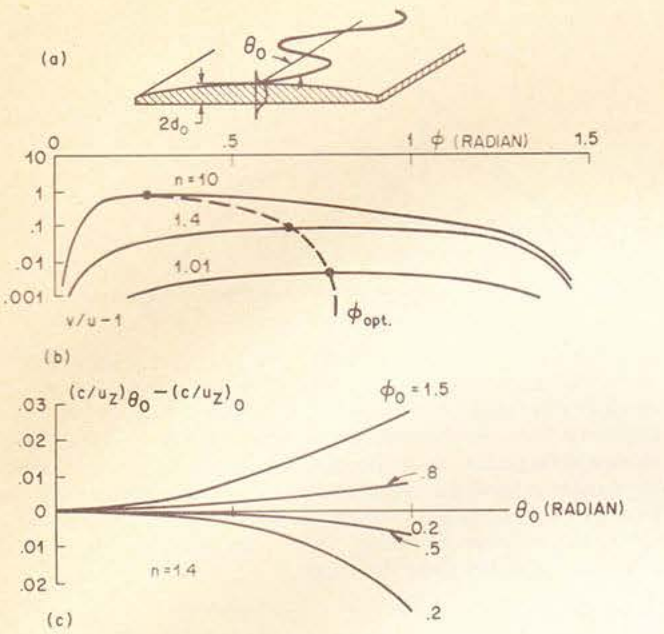

FIG. 11. (a) Quadratically tapered dielectric slab. θ_0 denotes the angle of a ray (H_0 mode in the direction perpendicular to the slab). (b) Ratio of the local magnitudes of the phase (v) to group (u) velocities as a function of the slab thickness [$\omega d/c = (n^2-1)^{-1/2}\phi/\cos\phi$] for $n=10$, 1.4, and 1.01. The ratio v/u is stationary at ϕ_{opt}. (c) Pulse delay as a function of the angle θ_0 that the ray makes at the origin with the z axis. (The number of modes is approximately proportional to θ_0^2.) For 15 modes, $\lambda_0=1$ μm, $n=1.01$, pulse spreading is less than 0.05 ns/km.

trajectory is proportional to the optical length of that ray. In that case, equal optical lengths imply equal times of flight, and therefore small pulse spreading, an essential feature for high-data-rate communication. This is not the case in general, however. For most slabs, the ratio v/u varies rapidly with the slab thickness, and therefore with x.

It is interesting that when the slab thickness, $2d$, is large compared with the wavelength, the dispersion is the same as that of a cold plasma with plasma frequency $\omega_p(x)$, or that of a relativistic particle whose mass m is a function of x, or a nonrelativistic particle in a potential $V(x)$. The kinematics and dynamics, which follow from the dispersion equation, are the same for these different media and particles. In particular, we can apply to tapered slabs the Breit and Tuive theorem,[20] originally derived for cold plasmas and also applicable to nonrelativistic particles such as bullets in the earth gravitational field, the effect of the atmosphere being neglected. This theorem says that the time of flight of a pulse can be evaluated by assuming that the path is made of two straight lines, defined from the slope on axis as shown in Fig. 10. This is another way of saying that the horizontal component of the group velocity is a constant of motion. For such dispersive systems, the canonical momentum $J\vec{k}$ is proportional to the group velocity $\vec{u}$. Thus, for a thick dielectric slab, pulse spreading is, for any profile, just as large as for a clad fiber having the same apparent width. Low pulse spreading can be obtained for different profiles (e.g., quadratic or linear dependence of k on $|x|$) only if the slab thickness is so chosen that the ratio v/u has the proper variation with x.[21] This is illustrated in Fig. 11, in which the time delay is shown as a function of the angle θ_0 that the ray makes with the z axis. Very low pulse spreading is obtained when the angle ϕ (a monotonic function of the slab thickness, defined in Fig. 11 caption) is equal to 0.65, for $n=1.45$.

CONCLUSION

The presentation of the mechanical theory of light and its application to fiber optics was sketchy. However, the works listed in the references should make up for the missing steps. One key point made in this paper is that, in many important respects, no distinction need be made between matter waves and optical or acoustical waves. The second point is that numerous results in classical or quantum mechanics can be used in fiber optics, provided that the dynamical significance of the analogy is clearly understood. We have shown, in particular, that a well-known result in mechanics, the uniform horizontal motion of a bullet on earth, has, as a counterpart in optics, the constancy of the horizontal component of the group velocity in thick dielectric slabs.

*Part of this material was presented by the author at the U.R.S.I. Symposium on Electromagnetic Theory, London, July 1974.

[1]V. A. Fock, Bull. Acad. Sci. l'URSS, Ser. Phys. 14, 70 (1950), translated in V. A. Fock, *Electromagnetic Diffraction and Propagation Problems* (Pergamon, Oxford, 1965), Ch. 14.
[2]G. B. Whitham, J. Fluid Mech. 44, 373 (1970).
[3]O. Costa de Beauregard, C. R. Acad. Sci. (Paris) B 274, 164 (1972); Nuovo Cimento, Ser. 10 63B, 611 (1969).
[4]P. Penfield and H. A. Haus, *Electrodynamics of Moving Media* (M.I.T. Press, Cambridge, 1967), p. 211.
[5]S. R. De Groot and L. G. Suttorp, *Foundations of Electrodynamics* (North–Holland, Amsterdam, 1972).
[6]H. A. Haus, Physica 43, 77 (1969).
[7]J. A. Arnaud, (a) Electr. Lett. 8, 541 (1972); (b) Opt. Commun. 7, 313 (1973); (c) Am. J. Phys. 42, 71 (1974).
[8]J. P. Gordon, Phys. Rev. A 8, 14 (1973).
[9]W. B. Joyce, Phys. Rev. D 9, 3234 (1974).
[10]R. Lucas, in *Louis de Broglie, sa Conception du Monde Physique* (Gauthier-Villars, Paris, 1973).
[11]A. Kastler, C. R. Acad. Sci. (Paris) B 278, 1013 (1974).
[12]D. Marcuse, *Engineering Quantum Electrodynamics* (Harcourt, Brace and World, New York, 1970), p. 80.
[13]D. Marcuse, *Light Transmission Optics* (Van Nostrand, New York, 1972), p. 100.
[14]R. V. Jones and J. C. S. Richard, Proc. R. Soc. (Lond.) A 221, 480 (1954).
[15]A. Ashkin, Appl. Phys. Lett. 19, 283 (1971); A. Ashkin and J. M. Diedzic, Phys. Rev. Lett. 30, 139 (1973).
[16]T. H. Boyer, Phys. Rev. D 8, 1669 (1973).
[17]J. A. Arnaud, J. Opt. Soc. Am. 62, 290 (1972) and J. A. Arnaud, in *Progress in Optics, Vol.* 11, edited by E. Wolf (North–Holland, Amsterdam, 1973), p. 247.
[18]K. Husini, Prog. Theor. Phys. 9, 381 (1953).
[19]J. A. Arnaud, Bell Syst. Tech. J. 49, 2311 (1970).
[20]J. M. Kelso, *Radio Ray Propagation in the Ionosphere* (McGraw-Hill, New York, 1964), p. 220.
[21]J. A. Arnaud, Bell Syst. Tech. J. 53, 1599 (1974).
[22]E. A. J. Marcatili, Bell Syst. Tech. J. 53, 645 (1974).