Reflection and refraction of optical beams at dielectric interfaces

Jeffery J. Regan, and David R. Andersen

Citation: Computers in Physics **5**, 49 (1991); doi: 10.1063/1.168406
View online: https://doi.org/10.1063/1.168406
View Table of Contents: http://aip.scitation.org/toc/cip/5/1
Published by the American Institute of Physics

Reflection and refraction of optical beams at dielectric interfaces

Jeffrey J. Regan[a)] and David R. Andersen
Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, Iowa 52240

(Received 2 March 1990; accepted 25 June 1990)

The phenomena associated with the reflection and refraction of light have been widely studied. As the level of dimensionality of the model used in the analysis increases, so does the variety of interesting phenomena that emerge from the analysis. For example, plane-wave calculations give rise to Brewster angle and total internal reflection effects. If a finite two-dimensional slab beam is considered, Goos–Hänchen, focal, and angular shifts become apparent. As the problem is generalized to higher dimensions, possibly including the temporal dimension, additional phenomena may be expected as well. This paper begins by defining the terms and general notation used with the interface problem by reviewing plane-wave (infinite field) solutions. This is followed by a literature review of some two-dimensional results, including a discussion of the Goos–Hänchen, focal, and angular shift. Then numerical solutions to the full three-dimensional problem of finite beams (under a paraxial approximation) are presented using modern visualization techniques. The numerical results provide an intuitive understanding of interface phenomena from a new perspective, graphically underscoring the difference between the geometrical ray model of reflection and refraction and the complicated field interaction that actually occurs. Finally, the numerical model is extended to include results for the nonlinear interface.

INTRODUCTION

The reflection and refraction of light at an interface between two dissimilar dielectric materials is a fascinating problem. The problem is important from fundamental as well as applications aspects. For example, dielectric waveguide devices as well as various leaky wave radiation structures can only be understood in terms of the reflection of the electromagnetic waves from the dielectric boundaries inherent in the device. The problem can be broken into two broad classes: (1) the set of phenomena associated with plane-wave propagation; and (2) the set of phenomena associated with bounded beam propagation. While a simple analytical procedure can yield substantial results and understanding in the plane-wave case, the analysis and understanding of the bounded beam case is far more difficult. One of the purposes of this paper is, through the use of visualization, to provide a coupling between the available two-dimensional analytical theory and numerically obtained solutions to the three-dimensional paraxial wave equation. In this way, the reader can obtain a clearer picture of some of the effects suggested by the theoretical results.

In addition to questions associated with the linear theory discussed in the first two sections of the paper, one may wish to consider the more general interface problem where one or more of the media comprising the interface possess dielectric functions that respond in some fashion to an applied optical field. This is a case where the superposition integral over a set of plane waves, so often used in the linear problem, cannot be used. Thus numerical solutions to the propagation equation are required. Some examples of such solutions are presented in order to introduce the reader to the rich variety of behaviors that can come from this simple problem.

I. ANALYTICAL TECHNIQUE

The interface problem is essentially the problem of solving the Helmholtz equation for the boundary conditions depicted in Fig. 1. The geometry convention used has a beam incident in the $x = 0$ plane, focused on the origin, coming in from the $+y, -z$ quarter space. Space is separated into two different, semi-infinite, isotropic, lossless materials by the $y = 0$ plane. The Helmholtz equation is

$$(\nabla^2 + \omega^2 \epsilon \mu)\mathbf{E} = 0. \tag{1}$$

The parameters ϵ and μ are the electric permittivity and magnetic permeability of the medium (possibly dependent on the electromagnetic field), and ω is the frequency of the time dependence of the solution, assumed to be $\exp(i\omega t)$. The vector $\mathbf{E}$ is a member of a linear vector space associated with a complex field, and it represents the electric field. A similar equation governs the magnetic field, and its solution is of course directly related to the solution for the electric field.

The first line of attack in the analytical explanation of reflection and refraction of light takes advantage of the linearity of the problem by breaking up the aggregate electric field into three parts; the incident, reflected, and transmitted fields. Each of these can be found [as a solution to

[a)] Currently at Department of Physics, 0319, University of California—San Digeo, La Jolla, CA 92093.

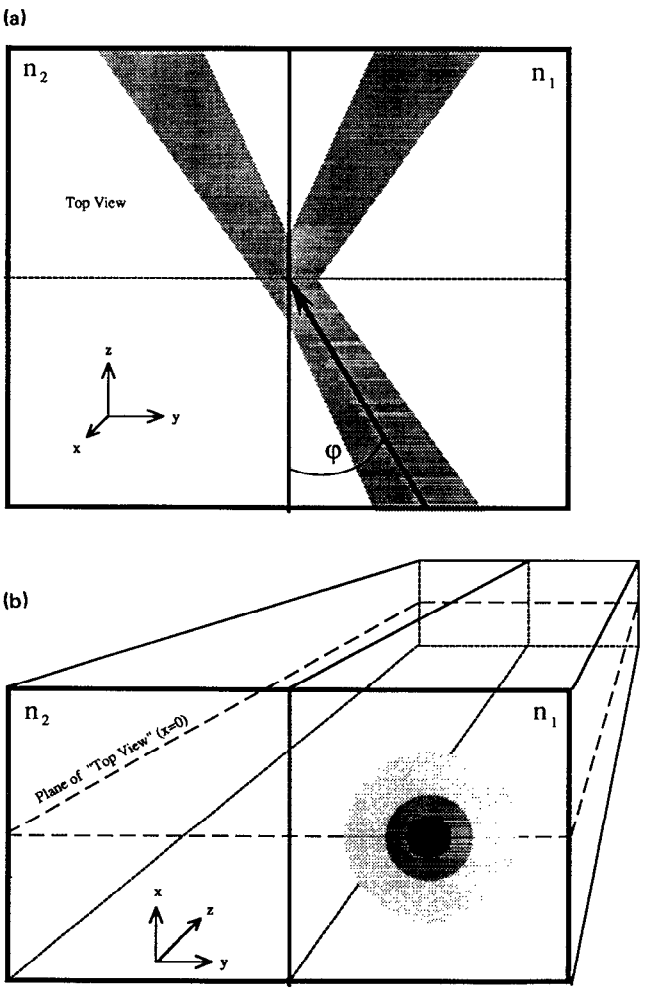


FIG. 1. Geometry of the interface; (a) top view from positive x perspective; (b) end view from negative z perspective. The interface is the dark vertical line. The beam is always focused on the origin. The calculation propagates an initial field through the positive z direction, with a periodic sampling of slices.

Eq. (1)] almost independently of the others, and the entire solution is then a superposition of the three. Upon dividing up the problem, the simplest solutions to try are plane waves—uniform fields of infinite extent possessing planar phase fronts. Snell's laws of reflection and refraction, the fundamental results of geometrical optics used to predict the direction of reflecting and refracting rays, can be deduced from the plane-wave analysis (and, indeed, are strictly valid only in the case of plane waves). The same analysis also yields the Fresnel relations, which quantify just how much of an incident plane wave is transmitted or reflected. Despite the fact that a true plane wave is physically unrealizable, these results are very useful. One reason is that in cases where one considers only a limited region, a plane wave can be a very reasonable approximation to the local field. Another reason is that finite fields (*beams*) of light can be treated as superpositions of plane waves, each wave uniquely identified by its propagation vector (which fixes both the wave's direction and wavelength), with some associated intensity and relative phase. In this way, it is often possible to use the results of plane-wave theory in a superposition integral to explore whole classes of solutions, as long as the problem is linear.

II. INFINITE FIELDS (PLANE WAVES)

The usual geometry for the plane-wave analysis is presented in Fig. 1. The interface is the $y = 0$ plane; on the positive y side is a medium characterized by the permittivity ϵ_1 and permeability μ_1, while the negative y side is characterized by ϵ_2 and μ_2. The materials are isotropic and lossless (ϵ and μ are real scalars). For the linear plane-wave results desired here, an incident wave can be completely specified by three parameters: the three components of its propagation vector, or the total magnitude of the propagation vector and any two components, or two direction angles and a wavelength. To simplify the discussion below, one of these three parameters will be eliminated by assuming all incoming rays lie in the same plane of incidence, the $x = 0$ plane. Because of this, the results one finds will only make sense in an appropriate two-dimensional subspace of this three-dimensional parameter space, so care must be taken in their application. Under these assumptions, any transversely, linearly polarized incident plane wave can be decomposed into two waves, one polarized perpendicular to the plane of incidence, and the other polarized in the plane of incidence. These cases can be considered independently, then scaled and recombined for any desired polarization. Perpendicular polarization shall be considered first.

The electric field of a perpendicularly polarized plane wave solving Eq. (1) has the form $\mathbf{E} = \hat{\mathbf{x}}E_0 \exp(-i\mathbf{k}\cdot\mathbf{r})$. The vector $\hat{\mathbf{x}}$ is a unit vector in the x direction (parallel to the interface), the vector $\mathbf{r}$ is the position in space for which the value of the electric field is sought, and the vector $\mathbf{k}$ is the propagation (or wave number) vector. The propagation vector is permitted to have complex components in order to represent decaying solutions. The vector formed from the real parts of its components is normal to the phase front, defining the direction of propagation. Here, E_0 is a complex constant representing the wave amplitude, with some arbitrary phase factor. A perpendicularly polarized plane wave incident on the interface traveling in the negative y and positive z directions, with signs chosen to yield positive wave numbers, can be written

$$\mathbf{E}_I = \hat{\mathbf{x}}E_0 \exp(ik_{Iy}y - ik_{Iz}z). \tag{2}$$

Its associated reflected and transmitted waves are assumed to have similar forms:

$$\mathbf{E}_R = \hat{\mathbf{x}}RE_0 \exp(-ik_{Ry}y - ik_{Rz}z), \tag{3}$$

$$\mathbf{E}_T = \hat{\mathbf{x}}TE_0 \exp(ik_{Ty}y - ik_{Tz}z). \tag{4}$$

The propagation vector for the reflected wave is written $\mathbf{k}_R = k_{Ry}\hat{\mathbf{y}} + k_{Rz}\hat{\mathbf{z}}$, implying a wave propagating in the positive y and positive z directions. The complex coefficients R and T are inserted to permit the introduction of some phase shift and scaling relative to the incident wave. The coefficients that satisfy these relations are only valid for perpendicular polarization; different coefficients are found for parallel polarization. In Eqs. (2)–(4), the im-

portant unknowns are the y and z components of the three propagation vectors and the two coefficients R and T. The consideration of a plane wave at a linear interface is essentially a consideration of the relationships between these values.

Substitution of the three electric field solutions into the Helmholtz equation shows that the following must hold true for the components of the propagation vectors, where the total wave numbers k_I, k_R, and k_T are defined:

$$k_I^2 \equiv k_{Iz}^2 + k_{Iy}^2 = \omega^2 \epsilon_1 \mu_1,$$
$$k_R^2 \equiv k_{Rz}^2 + k_{Ry}^2 = \omega^2 \epsilon_1 \mu_1,$$
$$k_T^2 \equiv k_{Tz}^2 + k_{Ty}^2 = \omega^2 \epsilon_2 \mu_2.$$

Since ϵ and μ have been assumed to be real, the components of k must each be purely real or purely imaginary. It is now useful to define the index of refraction n and the vacuum wave number k_0: $\omega^2 \epsilon \mu = (\omega^2 \epsilon_0 \mu_0)(\epsilon_r \mu_r) = k_0^2 n^2$. The interface can then be characterized by the index of refraction ratio it represents:

$$n_{12} = n_1/n_2 = \sqrt{\epsilon_{1r}\mu_{1r}}/\sqrt{\epsilon_{2r}\mu_{2r}}.$$

Thus $k_T = k_I/n_{12}$.

Assuming no conductivity, boundary conditions require the tangential components of **E** and **H** to be continuous across the interface; $E_{Ix} + E_{Rx} = E_{Tx}$ and $H_{Iz} + H_{Rz} = H_{Tz}$. Applying this condition to the electric fields in Eqs. (2)–(4), and their associated magnetic fields found from Faraday's law, one finds that at the interface ($y = 0$),

$$\exp(-ik_{Iz}z) + R\exp(-ik_{Rz}z) = T\exp(-ik_{Tz}z), \tag{5a}$$

$$k_{Iy}(1/\mu_1)\exp(-ik_{Iz}z) - k_{Ry}(1/\mu_1)R\exp(-ik_{Rz}z) = k_{Ty}(1/\mu_2)T\exp(-ik_{Tz}z). \tag{5b}$$

Since these equations must hold for all possible values of z along the interface, one concludes that $k_{Iz} = k_{Rz} = k_{Tz}$. This is the phase-matching condition. At all points along the interface the z components of the propagation vectors of the incident, reflected, and transmitted field must match. This also implies that $k_{Iy} = k_{Ry}$.

Snell's laws describe directions of propagation with angles, usually defined between the field's propagation vector and a normal to the interface. Here, the glancing angle (the angle between the propagation vector and the interface itself) will be used since the discussion will soon center on beams propagating nearly parallel to the interface, making the glancing angle a more intuitive choice. If *propagating* solutions are assumed (in which the components of the propagation vectors are real), then from Fig. 1 one can write $k_{Iz} = \cos(\varphi_I)k_I$, $k_{Rz} = \cos(\varphi_R)k_R$, and $k_{Tz} = \cos(\varphi_T)k_T$. From this and the relationships listed above one can deduce Snell's laws for glancing angles; $\varphi_I = \varphi_R$ and $n_1\cos(\varphi_I) = n_2\cos(\varphi_T)$.

The other half of the plane-wave analysis, parallel polarization of the incident field, proceeds in a similar fashion. The initial expressions involved are the same as in Eqs. (2)–(4), after substituting **E** for **H**, $-$**H** for **E**, μ for ϵ, and ϵ for μ. The phase-matching condition and Snell's laws come out exactly the same way, but the Fresnel reflection coefficient R and the transmission coefficient T are slightly different for the two cases. By applying the relationships revealed about the components of the propagation vectors to Eq. (5) and solving for R and T, then doing the same for the parallel polarization case one finds that

$$R = \frac{(1 - mK)}{(1 + mK)},$$
$$T = 2/(1 + mK), \tag{6}$$

where

$$K = \sqrt{(k_I/n_{12})^2 - k_{Iz}^2}/\sqrt{k_I^2 - k_{Iz}^2},$$
$$m = \begin{cases} \mu_1/\mu_2, & \text{perpendicular polarization,} \\ \epsilon_1/\epsilon_2, & \text{parallel polarization.} \end{cases}$$

When $K^* = -K$ (K pure imaginary), $RR^* = 1$, and the wave is totally reflected, independent of its polarization. R contributes only a phase shift to make the reflected wave different from the incident wave. This condition on K requires that $n_{12} > 1$, implying that the target index must be less than the source index for TIR to occur. This condition on K also requires that $n_{12}^2 k_{Iz}^2 > k_I^2$, which is equivalent to the condition that $\cos^2(\varphi_I) > 1/n_{12}^2$. Defining the critical angle $\varphi_C = \cos^{-1}(1/n_{12})$, it is clear that any plane wave with a glancing angle in the range 0 to φ_C will totally reflect. Conversely, when $mK = 1$, $RR^* = 0$, and the wave is totally transmitted. Since m appears, this condition depends upon the polarization. When relating the condition back to the incident angle, one finds that there will only be two angles (a unique angle for each of the two polarizations) where the condition is satisfied. Further, the angle is real (achievable) only when m is in a certain range. Since in practice $\mu = \mu_0$ for most nonmagnetic materials, m is unity for perpendicular polarization and for this value of m the zero reflection angle is unachievable. But unlike μ, the permittivity ϵ varies considerably from material to material, so for parallel polarization m is often in the proper range. The corresponding incident glancing angle, where no *parallel* polarized light reflects, is called the Brewster angle, denoted φ_B.

The results in Eq. (6) have a wider range of applicability than Snell's laws because they take into account two of the three parameters necessary to define a plane wave. If one can expand the incident beam as a summation of plane waves, each having the same propagation vector k_I but different k_{Iz}, one can then insert the expression for R or T above into an integral (over k_{Iz}) and obtain an expression for the reflected or transmitted beam. This is a powerful tool, but it is important to remember that any results obtained through the use of Eq. (6) are necessarily two-dimensional. The propagation vector of each plane wave must lie in the same plane of incidence [k_{Ix} is not a parameter in Eq. (6)]. Thus one is really only considering "slab" beams, beams that are uniform to infinity along the direction perpendicular to the plane of incidence.

For the remainder of this discussion it will be assumed that $\mu = \mu_0$ for all materials, thus m is 1 for perpendicular polarization and n_{12}^2 for parallel polarization.

III. FINITE FIELDS (BEAMS)

There are three related effects involved in the reflection and refraction of finite beams from an interface that are either meaningless in or not predicted by the plane-wave/geometrical ray model. They each involve beam structure changes or shifts that could not exist for an infinite plane wave. In what follows, the plane-wave results are used with two free parameters, the z component of the incident beam's propagation vector, and the total magnitude of that vector. As pointed out, this will permit the construction of semi-infinite slab beams, beams that are limited in the y and z directions but extend uniformly to infinity in the x dimension.

Some simplifying assumptions can be made to ease the task of finding a solution to the Helmholtz equation. Viewing a bounded beam as a superposition of plane waves, one defines the beam's direction of travel as the direction of the propagation vector of the beam's strongest plane-wave component (hereafter referred to as the beam's propagation vector). It is the quickly oscillating phase dependence of this dominant plane wave that is mostly responsible for the beam's field variation in its direction of travel. Considering a beam propagating nearly parallel to the z axis (at a small glancing angle to an interface parallel to the z axis), one can neglect this phase dependence by considering only the "z envelope" of the beam. This is done by seeking solutions of the form

$$E(x,y,z) = \Phi(x,y,z)\exp(-ik_{Pz}z); \quad k_{Pz} = \cos(\varphi_I)k_0 n_1. \tag{7}$$

The z envelope variable is introduced here as Φ. The value of k_{Pz} selected for extraction is the z component of the propagation vector of a plane wave that is incident at a (small) glancing angle φ_I on the interface. For a beam, an incident angle is of course defined relative to the beam's propagation vector. By removing this particular value from the phase, one achieves an envelope with a very small z variation along the interface, where the beam's magnitude is highest. This fact is exploited below to simplify the Helmholtz equation. Further, by considering only transversely polarized beams, one can use the scalar Helmholtz equation and neglect writing the vector components. The solution found can then be used to describe the x and y vectorial components.

Substituting the form of the solution in Eq. (7) into the scalar Helmholtz equation and neglecting the small (by assumption) second partial derivative in z yields the paraxial wave equation

$$\{\nabla_t^2 + [(nk_0)^2 - k_{Pz}^2]\}\Phi = i2k_{Pz}\frac{\partial}{\partial z}\Phi. \tag{8}$$

A relatively simple (rotationally invariant about z) solution to the polar coordinate version of this equation is the fundamental Gaussian beam, the simplest expression one can associate with the field of a laser beam. For $\varphi_I = 0$ this solution is

$$\Phi(x,y,z) = \left[E_0\frac{\omega_0}{\omega(z)}\exp\left(\frac{-r^2}{\omega^2(z)}\right)\right] \times\exp\left(\frac{-ikr^2}{2R(z)}\right)\exp[i\eta(z)], \tag{9}$$

where

$r^2 = x^2 + y^2$	(radial symmetry),
$\omega^2(z) = \omega_0^2[1 + (z/z_0)^2]$	[spot size (a characteristic beam radius)],
$R(z) = z[1 + (z_0/z)^2]$	(radial curvature of phase front),
$\eta(z) = \tan^{-1}(z/z_0)$	(represents a 90° phase shift as $z \to \pm\infty$),
$z_0 = k\omega_0^2/2$	(a characteristic beam length),
$k = nk_0 = n(2\pi/\lambda_0)$	(vacuum wave number scaled by the index).

The beam parameters are the axial vacuum wavelength λ_0, and ω_0 and E_0, the spot size and magnitude of the beam at its focus (spot size is at a minimum, magnitude is at a maximum). The first term in Eq. (9) shows the magnitude dependence on r and z. The beam is radially symmetric around the z axis, and its magnitude decreases with r to bound the beam radially. Since the spot size $\omega(z) \to \infty$ as $z \to \pm\infty$, the magnitude decreases with z as well, bounding the beam axially. The second term shows the phase dependence on both r and z. For constant r, as one moves along the propagation axis z, the radial curvature R quickly decreases from $\pm\infty$, reaches a minimum, and slowly tends back to $\pm\infty$. The phase fronts are very good approximations to spherical waves if observed close to the axis. For the remainder of this paper, the subscript "I" will be dropped from k and φ with the assumption that the incident beam is being discussed.

With the introduction of a more realistic solution, the three interface effects not apparent in the plane-wave analysis can be explored. But, as shall be seen, even for this relatively simple model of a true light beam it is not an easy task to arrive at results which predict detailed beam behavior in an obvious or intuitive way which are valid for a wide range of situations.

IV. THE GOOS–HÄNCHEN SHIFT

The first nongeometrical effect to be recognized, its existence was suspected by Newton but not experimentally confirmed until the work of Goos and Hänchen.[1] The Goos–Hänchen (GH) shift is the distance between the point at which the incident beam appears to strike the interface and the point from which the reflected beam appears to reemerge from the interface. In the ray optics model of this shift, the incident ray penetrates the real interface and reflects off of a "virtual" interface within the target medium, to exit the real interface at a point shifted from its entrance point. The GH shift is only appreciable for incidence in the TIR regime, reaching a maximum on the TIR side of the critical angle. Lotsch[2] made a detailed study of this shift, by finding expressions for the time-averaged Poynting vector in both media to determine the flow of beam energy. Some energy is found to enter the target media, travel along the interface, and emerge again to join the reflected beam, effectively shifting the "center of gravity" of the reflected beam. The classical expression (consistent with the notation here) for the GH shift as found in the early work by Lotsch and others is

$$D = \frac{m\cot(\varphi)(n_{12}^2 - 1)}{n_{12}^3 \pi \sqrt{\cos^2(\varphi) - \cos^2(\varphi_C)}\{\sin^2(\varphi) + m^2[\cos^2(\varphi) - \cos^2(\varphi_C)]\}} \lambda_0, \tag{10}$$

where D is the shift along the interface. The problem with this expression is that D is infinite for beams incident at the critical angle, suggesting that the incident beam is guided by the interface and propagates along it for an infinite distance before reflecting. Another problem is that the expression is independent of the beam waist, a factor that should play a role since the waist serves to define the beam's location.

Horowitz and Tamir[3] attack the problem of the critical angle by using the plane-wave results above. For a beam they use a well-collimated ($\omega_0 \gg \lambda_0$) Gaussian beam. Their approach is to break up the incident beam into its constituent plane-wave components, then employ the Fresnel reflectance [Eq. (6)] to find the reflected field. They make a slice parallel to the interface through the incident beam along the plane $y = h (h > 0)$, then take the Fourier transform of the slice to realize an expression that is a function of the spatial frequency variable k_z. The expression is the amplitude A of the plane wave having k_z as the z component of its propagation vector. This amplitude is inserted into an inverse Fourier transform expression to yield the incident beam as an integral sum (over k_z) of plane waves:

$$E_I(y,z) = \int_{-\infty}^{\infty} dk_z\, A(k_z)\exp[i(k_z z + k_y y)].$$

The reflected field is then simply this integral with the signs on y reversed and the Fresnel reflectance R (also a function of k_z) inserted to scale the amplitude;

$$E_R(y,z) = \int_{-\infty}^{\infty} dk_z\, A(k_z)R(k_z)\exp[i(k_z z - k_y y)]. \tag{11}$$

A constant phase shift of $\exp(ik_y h)$, accounting for the fact that the incident field was found from its values at $y = h$, has been neglected from these equations. To extract the GH shift from the expression, the Fresnel reflectance is split into a sum of two terms, one giving the plane-wave reflectance at the incident angle φ of interest, the other containing the remainder of the reflectances. This has the effect of splitting the integral expression for the reflected field into two terms also; a "geometrically" reflected field arising from the plane wave incident at φ and the reflected field due to all the other plane waves. This second term can then be viewed as a correction to the geometrical theory. Performing the integrals, an expression for the reflected field is achieved in which z is replaced by $z - D$, thus isolating the shift. The authors show that the resulting expression for D reduces to Eq. (10) under the assumption

$$\left|\frac{k\omega_0\delta}{\sqrt{2}}\right| \gg 1,$$

$$\delta = \left|\frac{\cos(\varphi) - \cos(\varphi_C)}{\sin(\varphi)}\right| = \left|\frac{\cos(\varphi) - (n_2/n_1)}{\sin(\varphi)}\right|.$$

Thus, for a well-collimated beam ($k\omega_0$ large), the classical expression works well down to very small differences between the incident and critical angles (small δ).

As the incident glancing angle increases toward the critical angle (in the TIR regime), the actual shift must at some point leave the curve described by Eq. (10), since that curve predicts that the shift will grow larger until the reflected beam shifts off to infinity when $\varphi = \varphi_C$. The true shift remains on that curve longer for large beam waists, but with large waists the GH is less apparent. After leaving the curve the true shift immediately takes on the largest value it will achieve, then it drops off slightly to remain roughly constant as the incident angle increases through the critical angle. Finally, in the partial transmission regime, the shift quickly falls to zero. The peak shift is achieved not at the critical incident angle, as one might expect, but at slightly below the critical glancing angle.

V. THE FOCAL SHIFT

McGuirk and Carniglia[4] simplify the above approach for TIR by considering only those plane-wave components of the beam which are incident in the range $0 < \varphi < \varphi_C$. To apply their results to a general collimated beam, one must be confident that the plane waves available in that range can create the desired beam. For a Gaussian beam, this is not difficult, since the plane-wave angular spectrum is also a localized Gaussian. By assuming all plane waves experience TIR, the Fresnel reflectance has unit magnitude, and only a phase change is introduced upon reflection, thus $R(k_z) = \exp[i\theta(k_z)]$. The phase of the reflectance is expanded in a Taylor series about $k_z = 0$;

$$\theta(k_z) = \theta_0 + k_z D + k_z^2(F/2k) + \cdots,$$

where

$$D = \left.\frac{\partial\theta}{\partial k_z}\right|_{k_z=0} \quad \text{and} \quad F = k\left.\frac{\partial^2\theta}{\partial k_z^2}\right|_{k_z=0}.$$

One can expand $k_y = \sqrt{k^2 - k_z^2}$ around $k_z = 0$ to realize that $k_z^2/2k = k - k_y + O(k_z^4/k^3)$. Using this and dropping terms of higher order than k_z^2 yields the approximation

$$\theta(k_z) = \theta_0 + k_z D + (k - k_y)F.$$

This expression for the phase of the reflectance is then inserted into Eq. (11) to yield an approximation of the reflected field,

$$E_R(y,z) = \exp[i(\theta_0 + kF)]\int_{-\infty}^{\infty} dk_z\, A(k_z) \times \exp[i(k_z(z + D) - k_y(y + F)].$$

Comparing this expression to the geometrically reflected field [Eq. (11) with $R = 1$], the GH shift along the interface is embodied in D and is proportional to the rate of change of the phase of the reflectance with respect to the

incident angle. Another shift, embodied in F, is found along y. McGuirk and Carniglia provide a more detailed analysis in the coordinates of the reflected beam, and this shift is found as a shift along the beam axis, which permits one to call it a focal shift. It is proportional to the 2nd derivative of the phase of the reflectance with respect to the incident angle. This shift along the beam path is *different* from the beam path length increase associated with the previously mentioned ray model of the GH shift, and is seen here from their arguments to be independent of the GH shift. The focal shift was given a ray optics basis by Carniglia and Brownstein.[5] In the zeroth-order geometrical optics picture, two rays (plane waves) incident at different angles onto different points on an interface will reflect off the interface and intersect at some focus. Including the first order, both of these plane waves reflect off of a virtual interface within the target material, thus the GH shift is introduced and the focus moves with it. In the second order, the plane waves reflect off *different* virtual surfaces at *different* depths in the target, moving the focus along the path predicted at first order.

The focal shift would be very difficult to observe with a well-collimated beam, even though it can exceed the GH shift for angles near the critical angle. It can only be made apparent with a focused beam.

VI. THE ANGULAR SHIFT

The angular shift of a reflected or transmitted beam is a deviation from the angles predicted by Snell's laws. The angular shift was first recognized as such by Ra, Bertoni, and Felsen,[6] in a consideration of the interface problem using complex source coordinates. The dominant plane wave of an incoming beam defines its incident angle. Due to the odd relationship between the incident angle and the reflectivity, the dominant plane wave may be supplanted upon reflection (or transmission) by one of its neighbors in the plane-wave spectrum of the incident beam. This other plane wave, not suffering as much from transmission loss, becomes the new dominant plane wave in the reflected beam—resulting in an angular shift.

Antar and Boerner[7] begin their analysis in a manner similar to those mentioned above in that they express an incoming Gaussian beam as a sum of plane waves, then use the Fresnel relations to describe the reflected and transmitted beams. They expand the entire reflectance (not just the phase) in a series and consider the reflected fields due to each of the terms as different Gaussian beam modes. The lowest-order mode (arising from the first term in the series) is called the geometrically reflected mode. From this perspective, they derive both the GH and angular shift. In a later paper White *et al.*[8] provide a simpler picture of the angular shift. They start in the same fashion as the others, using a Fourier transform to find amplitudes of incoming plane waves, but they write the amplitudes as a function of incident angle instead of as a function of k_z. The amplitudes $A(\varphi)$ are then squared to use as magnitudes on incoming geometrical rays. The reflectance is also written as a function of angle $R(\varphi)$ (for a given interface index n_{12} and wavelength) and the direction φ of the reflected beam is found by solving

$$\frac{\partial}{\partial \varphi} |R(\varphi)|^2 |A(\varphi)|^2 = 0.$$

Consideration of the solution of this equation shows that for perpendicular polarization the shift is positive (the glancing angle of reflection is smaller than the angle of incidence) but small. For parallel polarization, the behavior is more complicated, due to the behavior at the Brewster angle. By definition, the Fresnel reflectance is zero for a parallel polarized plane wave incident at φ_B. But for a beam composed of a group of plane waves, there are strong components coming in on either side of φ_B, so one can expect to see two peaks in the reflected field around a zero in the middle (at φ_B). Chan and Tamir[9] give a detailed discussion of this phenomena. As the beam's incident angle is decreased from φ_B, the peak appearing at the lower reflected angle will grow while the higher peak shrinks. This constitutes a forward shift. Conversely, when φ increases from φ_B, the other peak will dominate, yielding a backward shift. In general, the shift depends on the square of wavelength/waist ratio (thus is small for collimated beams), and is only appreciable at Brewster's angle. There is also an equally slight angular shift of the transmitted beam. Naturally, with two peaks in the reflected beam, the concept of an angle of reflection is somewhat ambiguous.

In summary, the three shifting effects described above can be attributed to the distinctly nonlinear relationship between the Fresnel coefficients and the incident angle. For tightly collimated beams, the plane-wave spectrum is well distributed over this nonlinear range, enhancing these effects. The shifts are most in evidence at the critical angle for TIR and at the Brewster angle of zero parallel polarization reflection, the two angles where the reflectance changes most rapidly. If more than one interface is present (as in multilayered media), the shifts can be even more profound and can occur in either direction. Also, as Tamir[10] shows, a change in the beam waist is manifest as yet a fourth effect.

VII. NUMERICAL TECHNIQUE

The paraxial wave equation [Eq. (8)] can be solved in a reasonable amount of time on a fast computer. The magnitude of the resulting solution Φ is equal (under the assumption of slow z variation) to the magnitude of the actual variable of interest, E. There is also a correspondence between the phase of Φ and the phase of E, when considering the phase as a function of x and y. The relationship between the phases of the two variables as a function of z must be interpreted with the paraxial approximation of Eq. (7) in mind. Equation (8) is isomorphic to the Schrödinger equation with the potential $[(nk_0)^2 - k_{Pz}^2]$. The refractive index is a parameter in the potential, and must be set appropriately depending on which side of the interface one is working. The algorithm used does not care how this index is set, so one could use a variety of index profiles (not just one that presents a single interface) or even permit the index to vary as a function of the local field intensity.

Since $\Phi = \Phi_{Re} + i\Phi_{Im}$, Eq. (8) can be written as the real system

$$\begin{bmatrix}\Phi'_{\mathrm{Re}}\\ \Phi'_{\mathrm{Im}}\end{bmatrix}=\begin{bmatrix}0 & \frac{1}{2k_z}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+[(nk_0)^2-k_{Pz}^2]\right)\\ \frac{-1}{2k_z}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+[(nk_0)^2-k_{Pz}^2]\right) & 0\end{bmatrix}\begin{bmatrix}\Phi_{\mathrm{Re}}\\ \Phi_{\mathrm{Im}}\end{bmatrix}.$$

The derivative on the left-hand side is with respect to z. To develop a discrete analog of this equation for presentation to a computer, the x and y dimensions are discretized into a grid and the second partial derivatives for both Φ_{Re} and Φ_{Im} are replaced by fifth-order finite-difference approximations. The partial differential equation with two dependent variables and three independent variables becomes a coupled system of ODEs with more dependent variables but only one independent variable (z). Each point on the two grids for Φ_{Re} and Φ_{Im} is a new dependent variable representing the local field, and the number of these dependent variables depends on the grid resolution. With this discretization adopted, an appropriate solver can then be used to propagate any initial 2-D field distribution through the z dimension. In the case described here, a fifth-order predictor–corrector algorithm was used.

During computation, the total power at any slice through the beam should not change, and any symmetries (as characterized by the moments of the field) should be preserved. The resolution was chosen to keep these conserved quantities within a few tenths of a percent of their initial values. For computation on a vector machine, the grid dimensions were chosen by first selecting a size required to achieve the desired resolution (for a given λ_0 and beam waist), and then increasing that size to be as close as possible to the next integer multiple of the machine vector length. This results in the highest resolution possible for the price already paid in vector computation overhead. In a typical run the grid dimensions were such that $\sim 150\,000$ coupled equations were solved.

A given beam is inserted into the computation via the initial condition required by the algorithm. The first grids of Φ are initialized using the Gaussian beam of Eq. (9), after the application of an appropriate transformation of the coordinates to account for the glancing angle. Figure 1(b) shows the initial condition end of the lattice, where the first grid is initialized by computing a Gaussian beam incident at a small glancing angle.

At the x and y boundaries of the problem space, the electric field is held at zero, modeling perfectly reflecting boundaries. The confusion that could arise from reflections off these boundaries can be avoided by making the problem space large enough, relative to the beam waist, that the field intensity at the boundaries is down by at least 40 dB from its maximum value.

The solution as provided by the computer is the value of the complex variable Φ sampled in planes of constant z at regular intervals throughout the volume. To view the solution, the magnitude and phase of Φ at various planar slices through the volume are mapped to colors on a display device. These images can be viewed individually, as presented here, or they can be used as frames in a movie to permit one to see the evolution of the field as it propagates. To provide a reference for the visualization technique employed, a Gaussian beam [computed from Eq. (7) with Eq. (9)] is presented in two images in Fig. 2. The beam is focused on the origin, incident on the "interface" at an angle of 45°. In the top half of each picture the field magnitude is presented in green. The magnitude is always mapped logarithmically to the color intensity, so that low-intensity structure remains visible. In the bottom half of each image is the phase, where red indicates a phase of 0, yellow $= \pi/2$, blue $= \pi$, and cyan $= 3\pi/2$. The coordinate system for these images (and all others used) may be found simply by overlaying the system presented in Fig. 1. The origin is always in the center of the magnitude and phase representations. The horizontal dimension (piercing the interface) is always y, while the vertical dimension is either z or x depending on the perspective. Figure 2(a) is a slice through the space corresponding to the $x = 0$ plane, as viewed from a positive x perspective, so the vertical dimension is z. Hereafter, this slice will be called a "top view." In this case the interface line marking the $y = 0$ plane is shown for reference only. The index of refraction on both sides of it is 1.5, thus there is no interface. In Fig. 2(b), the end (or initial condition) view of this same reference beam is presented by a slice through the z axis, and the vertical dimension is x.

The vacuum wavelength (λ_0) of the beam defines the scale of the problem. For this reason, all lengths will be given in terms of the vacuum wavelength used. The views in Fig. 2 are $6\lambda_0$ on a side, and the beam's waist is $\frac{2}{3}\lambda_0$. Normally the beam waist exceeds the wavelength, but this was not done here so that the phase structure would be easy to see. The diagonal distance is $\sim 8.5\lambda_0$, and since $\lambda_0 = n\lambda$, one expects to see $\sim 12.7\lambda$ (12.7 cycles) across the diagonal in Fig. 2(a). There are only ~ 12 because of the η term in Eq. (9). At the focus the phase front is flat and zero (red), and becomes curved as one leaves the origin in the direction of propagation. The end view of Fig. 2(b) is taken at $z = -160\lambda_0$. In this view, the spherical wave nature of the Gaussian beam is apparent. The concentric rings are distorted since this is an oblique slice (at 45°).

VIII. THE LINEAR INTERFACE

The initial condition for the modeling of a Gaussian beam incident at an 8° glancing angle, focused on the origin, is shown in Fig. 3. A small glancing angle ($\varphi < 10°$) must be used under the constraints of the paraxial approximation. In this picture, the vertical (x) dimension is $44\lambda_0$, the horizontal (y) is $82\lambda_0$, and the slice represents the plane $z = -160\lambda_0$. The beam waist is $10\lambda_0$. Since the glancing angle is small, the phase presented here comes from the spherical wave region close to the beam waist; thus there is little of the "distortion" seen in Fig. 2(b). Also in contrast to Fig. 2(b), this view is many beam lengths (z_0) from the origin; thus the radius of curvature of the phase fronts is larger and only a portion of the circle is visible. The phase seen here is exactly the magnitude of the y component of the beam's propagation vector.

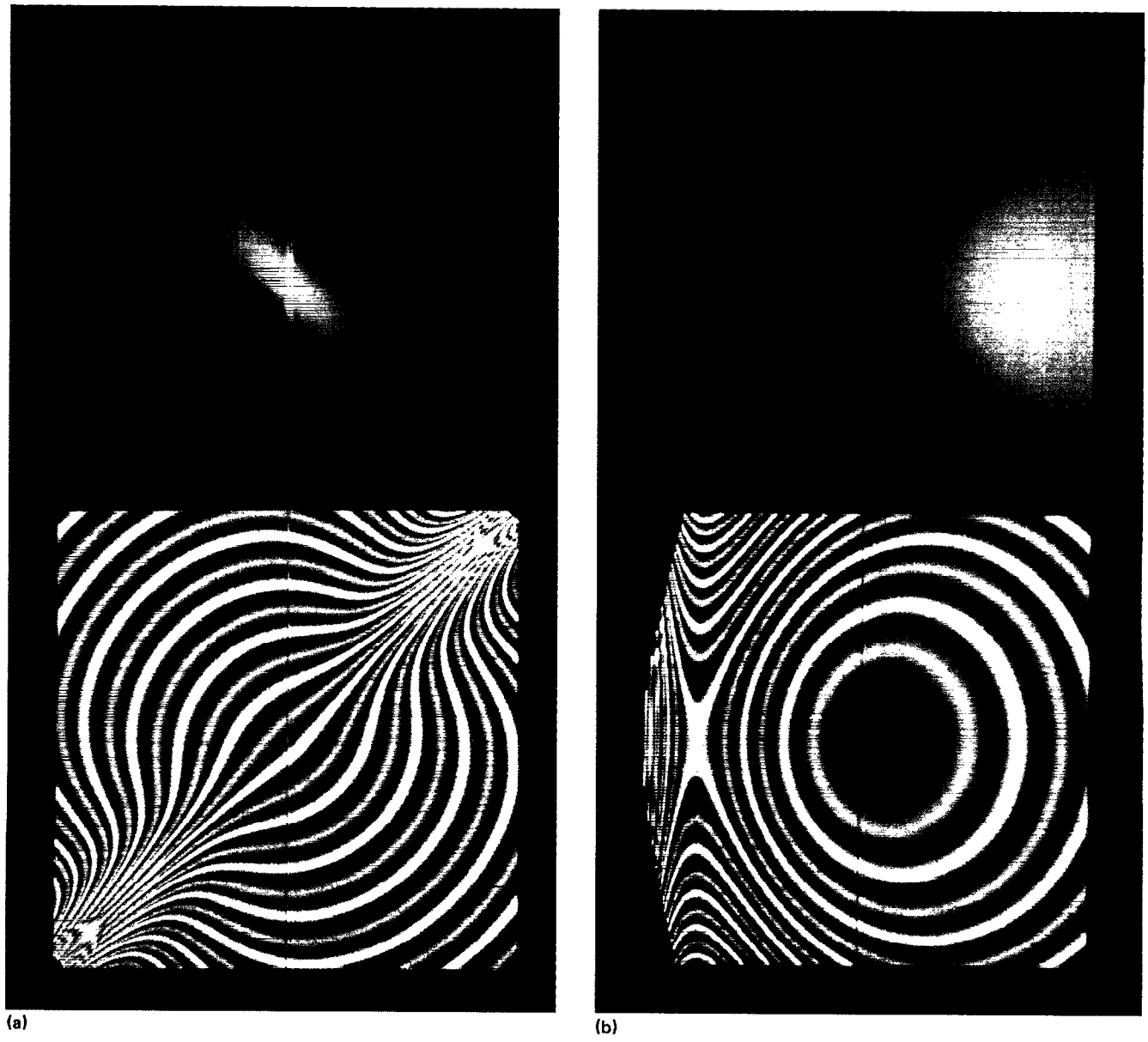


FIG. 2. The Gaussian beam: (a) top view of $x = 0$ slice; (b) end view of $z = -160\lambda_0$. The top half of each image shows the field magnitude logarithmically mapped to the color green. The bottom half shows the field phase; red indicates a phase of 0, yellow = $\pi/2$, blue = π, and cyan = $3\pi/2$. The geometry of this image and all other images corresponds to the geometry depicted in Fig. 1. The beam waist is $\frac{3}{2}\lambda_0$, and the end view is $6\lambda_0$ on a side. Normally the beam waist exceeds the wavelength, but this was not done here so that the phase structure would be easy to interpret. The interface line is drawn for reference; the index of refraction on both sides of it is 1.5, thus there is no interface. The beam is focused on the origin, incident on the "interface" at an angle of 45°. At the focus the phase front is flat and zero, and becomes curved as one leaves the origin in the direction of propagation.

A sequence of six top views presenting a beam incident at a glancing angle of 8° with an index of 1.5 on the incident side and a decreasing target index is presented in Fig. 4. The magnitudes and phases of the six fields are grouped separately for easier comparison; Fig. 4(a) comprises the image in the upper left of the magnitude images and the image in the upper left of the phase images. The 8° angle looks much larger than 8° because the z distance represented in the vertical height of these pictures is $320\lambda_0$, while the y distance represented (horizontal length) is only $82\lambda_0$. The index of refraction of the target medium, the associated critical incident glancing angle, and the magnitude of the reflection coefficient (calculated by numerically integrating the reflected and incident fields) are listed for each solution in the caption.

The index of the target medium falls steadily as one progresses from Fig. 4(a) to (f). As the target medium's index falls, the critical angle for TIR increases until the incident beam (held at 8°) is coming in under the critical angle. Over this range, the reflectance changes rapidly with incident angle. The transmitted beam scans across, while the reflected beam rises out of the ether at a fixed angle that

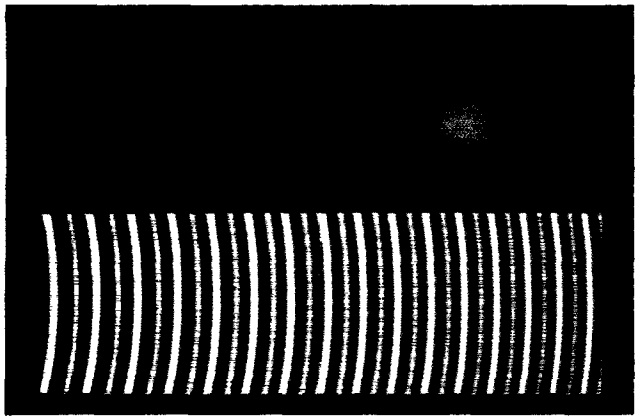


FIG. 3. The initial condition for 8° glancing angle incidence. The slice is taken at $z = -160\lambda_0$, through a Gaussian beam rotated through 8°. If this view was from a point on the beam's axis, the phase would appear as a set of concentric rings, with more rings appearing as the view point moved away from the beam focus. At the beam focus (on the axis), the phase portrait would be entirely red (flat phase front). The beam cross section appears elliptical (as opposed to circular) in this slice and others like it because the resolution in the y dimension is 1.65 times higher than that in the x dimension, a condition necessitated by the index discontinuity in the y dimension. The simple mapping between grid variable and pixel then results in more pixels per λ_0 in the y dimension than in the x dimension.

looks the same as the incident angle. The interference fringes between the incident and reflected field are clearly visible in the incident medium. If one examines the phase closely, one can see the correspondence between the fringe peaks and the regions where the fields are in phase. The red–cyan and blue–yellow transitions occur at the same place (constructive interference) for both fields, while the yellow–red and blue–cyan transitions are in opposition (destructive interference). Throughout the sequence the square of the y component of the transmitted field's propagation vector (as seen in the phase) gets smaller as the target index decreases. Finally, it becomes negative, and the transmitted field becomes a decaying exponential. Conversely, as this y component decreases, the z component of the propagation vector increases, as more nonuniform plane waves join the evanescent surface wave propagating along the z axis. Thus the phase on the transmitted side starts showing a z component, and the phase fronts tilt off the vertical.

In these top views the paraxial approximation of Eq. (7) is made graphically clear. In Fig. 2(a), one sees the E of Eq. (7) complete with properly curved phase fronts, oriented along the beam axis. In all other top views one will see the envelope variable Φ, which, relative to E, is missing $\cos(\varphi)k_0 n_1$ from the z component of its propagation vector. As previously described, this value is the z component of the propagation vector of the beam's main plane-wave component. One result of this is that along the interfaces in these top views, where one would expect to see fast phase oscillation of the beam, one sees the phase change only slightly and only as one proceeds to the wings of the beam (where the dominant plane waves have propagation vectors with z components that differ from the one which was extracted). The phase is continuous across the interface, as it should be in any case. In general, the field's variation with z is much smaller than the y variation (keeping in mind the z dimension has been greatly compressed in these pictures). The phase fronts are no longer curved nor oriented down the beam axis. As argued before, as long as one seeks paraxial solutions, then the only difference between Φ and E is the phase term $\exp(-ik_{Pz}z)$. If one wanted pictures of E instead of Φ, the computer could reinsert this term into the phase pictures (the magnitude pictures would not change). But, like the time dependence $\exp(i\omega t)$, it is left suppressed since its reintroduction into the solution would only serve to make things more complicated and confuse the more interesting y dependence.

In the z dimension, the phase flips 180° to demarcate the region where the incident field dominates from the region where the reflected field dominates. In Fig. 4(f) (TIR), this line is straight enough to define the y axis. In Fig. 4(a) (near total transmission) the line is curved show-

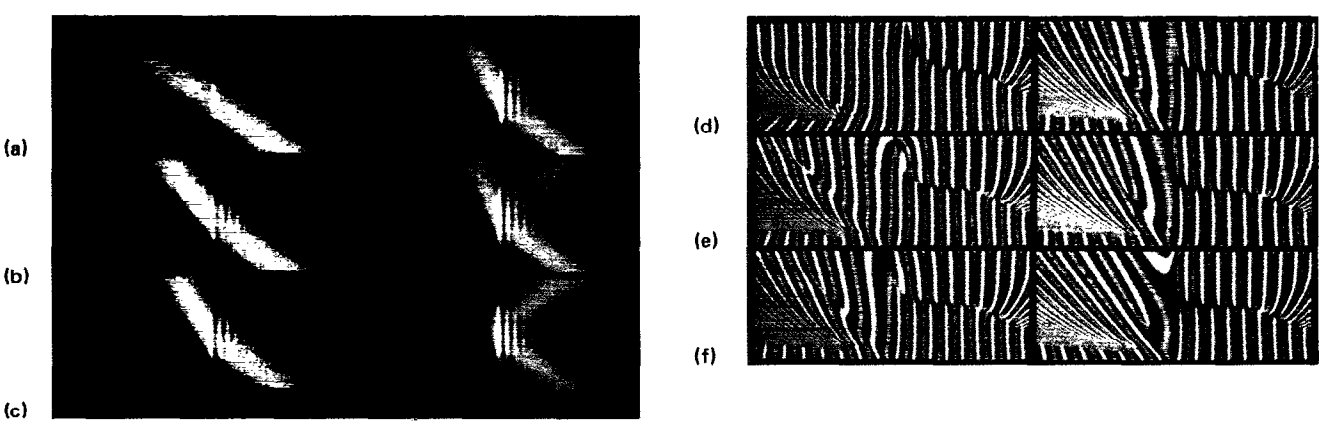


FIG. 4. A sequence of six top views (magnitude and phase grouped separately) presenting a beam incident at a glancing angle of 8° from a medium with an index of 1.5 onto a target with a steadily decreasing index. The 8° angle looks much larger than 8° because the z distance covered in the vertical height of these pictures is $320\lambda_0$, while the y distance covered (horizontal length) is only $82\lambda_0$. The index of refraction of the target, the associated critical incident glancing angle, and the calculated reflection coefficient are as follows: (a) 1.504, undefined, 0.003; (b) 1.491, 6.245°, 0.067; (c) 1.488, 7.198°, 0.255; (d) 1.487, 7.602°, 0.435; (e) 1.485, 8.000°, 0.651; (f) 1.482, 8.885°, 0.958.

ing where the weak reflected field begins. In this case, since the reflected field does not appear in the magnitude portrait, the phase is the only indication that it exists.

In Fig. 4(e), the incident angle is at the critical angle; thus this represents a case for TIR. Clearly, there is still a great deal of transmission, since half of the beam's plane-wave components are still coming in above the critical angle. In Fig. 4(f), the critical angle well exceeds the incident angle of 8°, and one can see that very little light transmits. The phase singularities (whirlpools) in both of these pictures shows the point where the *propagating* transmitted field begins to dominate over the *evanescent* transmitted field. If the index on the other side were to be lowered further, more of these singularities appear, and they move closer to the interface because the evanescent field falls off faster.

In Figs. 2–8, the magnitude of the field outside the beam path is hundreds of decibels below the peak intensity. In these regions, the phase of the field may be as much an artifact of the machine representation as it is a feature of the true solution, so the phase information in these *low*-intensity regions should be interpreted with care.

The solution in Fig. 4(d) is presented from two more perspectives in Fig. 5. In the end view of Fig. 5(a) the refracted field (left) and reflected field (right) are evident. Close comparison of the reflected beam to the slice through the incident beam (Fig. 3) at the same distance ($160\lambda_0$) from the reflection point at the origin shows more differences besides a loss of intensity due to partial transmission. The beam has clearly changed shape as a result of the reflection, a consequence of the various shifting mechanisms associated with the nonlinear reflectance. Figure 5(b) shows the entire three-dimensional beam. Six transparent intensity contour levels were chosen, with the remainder of the beam suppressed. The interference fringes appear as concentric toroids of intensity, and the distribution of in-

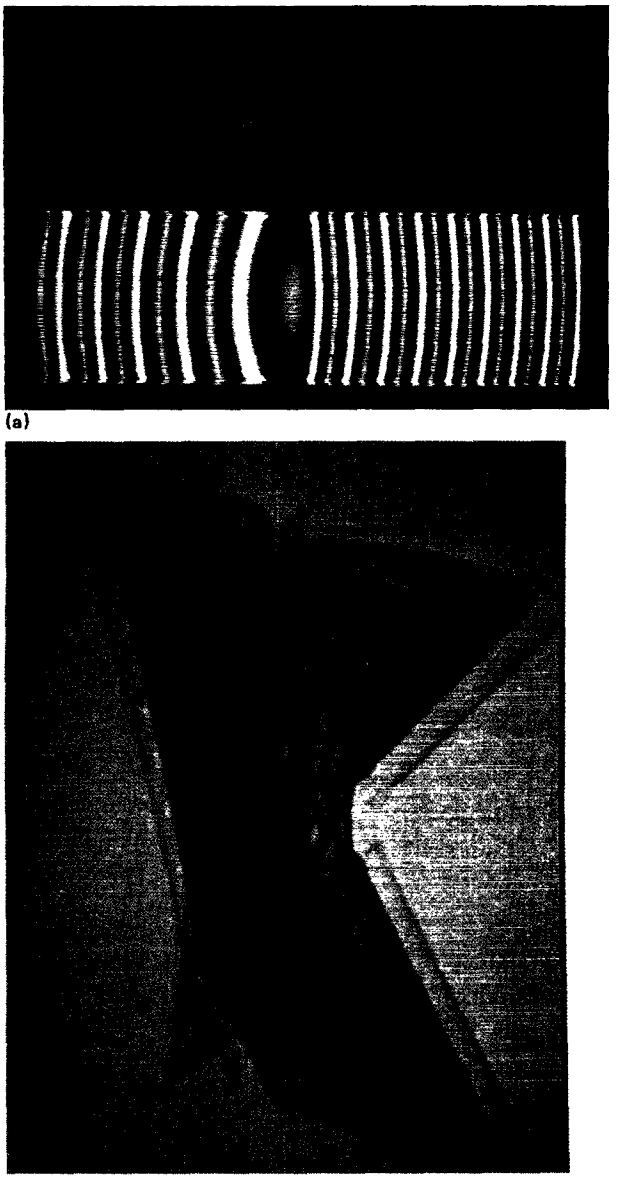


FIG. 5. (a) An end view ($z = 160\lambda_0$) and (b) a 3-D view of the field corresponding to the partial transmission case of Fig. 4(d). The glancing angle of 8° is just over the critical angle of 7.6° associated with this interface, and just over half of the beam is transmitted. Compare reflected beam profile to the incident beam profile in Fig. 3. A geometrically reflected beam would only appear dimmer, but in this case the beam shape is distorted due to the nonlinear relationship between the reflectivity and the incident angle of the plane waves which the beam comprises. In the 3-D view one sees in detail how the structure of the transmitted (upper left) and reflected (upper right) beams differ from that of the incident beam (lower right).

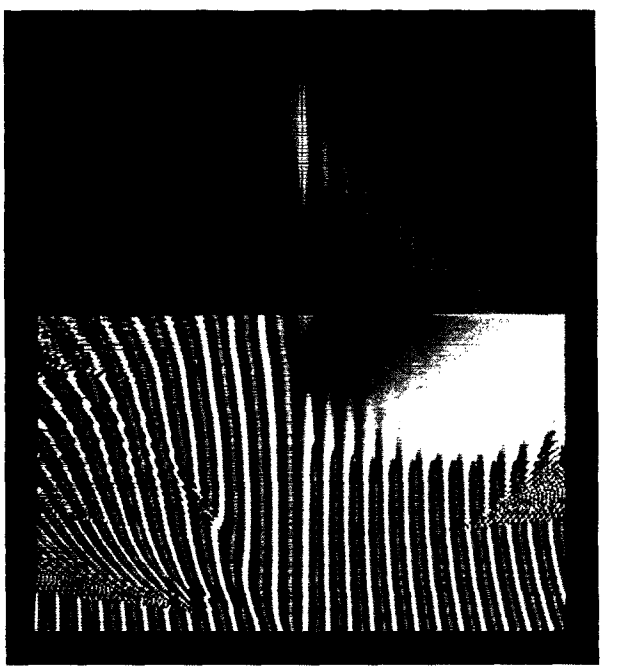


FIG. 6. A scalar subtraction of a "geometrically" reflected reference beam from an actual reflected beam. The geometrically reflected reference beam is made by originating a mirror image of the usual 8° incident beam on the left side of the interface (at $-8°$), and permitting it to transmit unchanged through the interface. Where the reflected beam is stronger than the reference beam, the magnitude is green. The magnitude is purple when the reference beam dominates. Black indicates the two are equal. The phase of the reference field is simply subtracted from the phase of the reflecting beam.

tensity in the transmitted and reflected field can be seen. There is a profound difference between the overall field distribution of the reflected beam and transmitted beam from the incident beam which the 2-D theory outlined in this paper is ill-equipped to characterize. The shape of the intensity contours of the transmitted beam suggest a beam whose normalized waist is smaller than the incident beam from which it originates, resulting in a faster diffraction of the beam as it propagates. A "geometrical" transmission would result only in an attenuation.

To emphasize the difference between a reflected beam and a geometrically reflected beam, and to observe the aforementioned shifts, one can subtract one beam from the other and observe the difference. To obtain a geometrically reflected beam, the problem was solved for a beam incident at $-8°$ (from the negative y region) with no interface present. With no index discontinuity at $y = 0$, the beam transmits through unchanged to take up the role of a geometrically reflected reference beam. The reference beam magnitude should be matched to the peak magnitude of the reflected beam to permit a fair comparison. This reference field is then removed from the reflecting field via a *scalar* subtraction, resulting in Fig. 6. The magnitude shows green where the reflected field magnitude dominates, and purple where the reference beam magnitude dominates. Black indicates the magnitudes are equal. The phase of the reflecting field is simply subtracted from the phase of the reference beam. The y components of the propagation vector of the incident beam (lower right) and the reference beam (lower left) have opposite signs, thus subtracting one from the other results in a phase difference varying at twice the frequency. The y component of the reflected beam's propagation vector, however, roughly matches that of the reference beam (in the upper right), thus the difference is a constant. The reflecting beam was incident at 8° (from the positive y region) just under the critical angle of 8.009° for the interface in question. For a plane wave this situation would yield TIR, exactly along the path of the reference beam. But for a finite beam a good portion of the beam will transmit, and the portion reflected should exhibit a GH shift.

In Fig. 6 it is clear that the reflected field has shifted down the interface. If the beam were geometrically reflected, the upper right of the image would be black, indicating

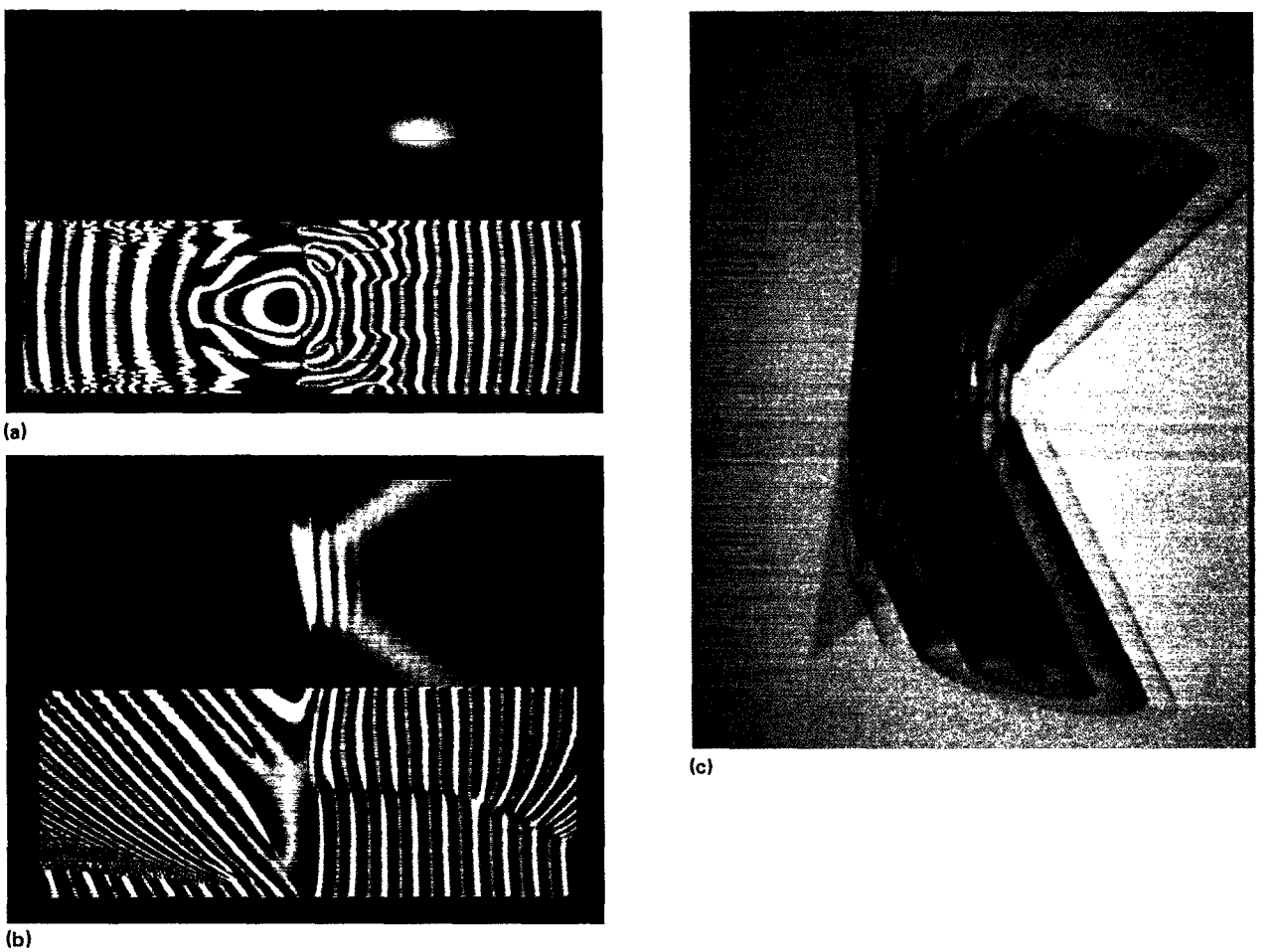
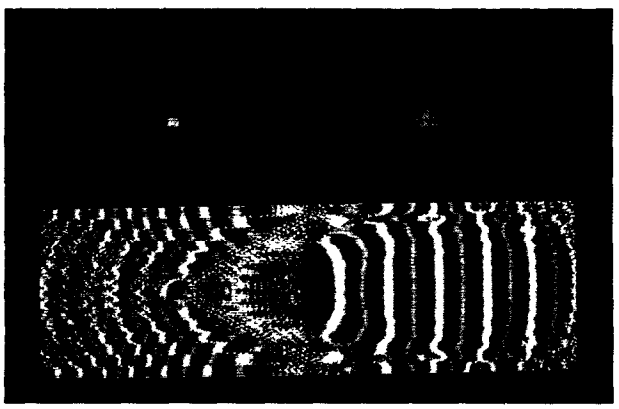


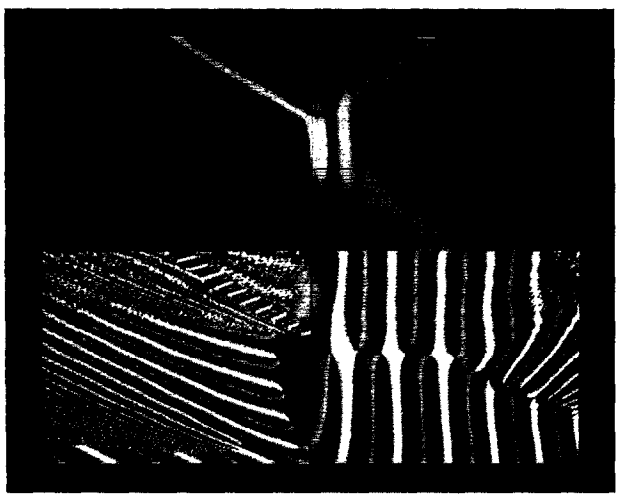
FIG. 7. Example of a nonlinear interface; (a) end view; (b) top view; (c) 3-D view. The index in the target depends on the intensity of the field ($n_{nl} = 0.004$), thus the reflectivity an incident plane wave sees depends not only on its incident angle but on its magnitude as well. Both the reflected and transmitted fields are profoundly deformed. The incident angle is 8°, which is well under the critical angle of 9.37° for this interface. In the linear case, the beam would be almost completely reflected [more so than in Fig. 4(f)], but since the evanescent field increases the index in the target slightly, a significant amount of light transmits. The reflectivity is calculated at 0.865.

perfect cancellation, instead of showing a black stripe bounded by regions of significant difference. The reflected field dominates in a region shifted forward along the interface. The peak in the calculated profile of the reflected beam suggests it reflects at an angle of $6.511° \pm 0.088°$ (the uncertainty is due to the discrete mesh). The same technique used on the reference beam indicates it transmits through at $7.989° \pm 0.087°$, which simply verifies the solver is working properly since the beam was incident at $-8°$. Drawing an 8° line back from the peak in the reflected field to the interface suggests that the reflection point shifts from the origin to $z = 27\lambda_0$, which is a very large GH shift for a beam with $\omega_0 = 10\lambda_0$. This figure is inflated by the fact that the dominant plane wave of the reflected beam is not propagating away at 8°, but rather at some smaller angle due to a forward angular shift. Thus the reflection point (defined as the intersection point of the propagation vectors of the two dominant plane waves in the incident and reflected beams) falls somewhere in between $z = 0$ and $z = 27\lambda_0$, and the GH shift is not as large after all. Equation (10) predicts a shift of $323\lambda_0$, but this beam is so close to the critical angle that Eq. (10) is not applicable. In the phase portrait, evidence for a focal shift can be seen. If the phase of the two fields were in perfect step, the color in the upper right would be flat red. Instead, one sees a pattern that suggests the reflected field has been shifted along its axis slightly. Near the interface and near the center of the beams the phase difference is indeed zero, but the difference increases as one moves to the wings of the beam or down the beam axis.

IX. THE NONLINEAR INTERFACE

The analysis so far has assumed that the intensity of the light does not matter, and that if two fields provide a solution to the Helmholtz equation for some boundary conditions, then the sum of these fields and any scalar multiple of the fields are also a solution. This is the linearity assumption, and without it, none of the analysis above holds. In all materials the permittivity (or permeability) changes with the electromagnetic field the material experiences, but for most materials this effect can be neglected (the material can be considered linear) unless the fields are very strong. There are, however, some materials that are nonlinear even


(a)


(b)


(c)

FIG. 8. Another nonlinear interface; (a) end view; (b) top view; (c) 3-D view. The incident angle is reduced from 8° to 5° (note the decrease in the y component of the propagation vector and the correspondingly fatter interference fringes) and the nonlinearity is increased to 0.011. The transmitted beam forms a filament of light in the nonlinear target. The tendency of the beam to diffract is offset by the nonlinear self-focusing. In this case, as opposed to the case in Fig. 7, there is enough power in the transmitted beam to have a collapse to an intense singularity. This collapse is prevented by using a hard-limited nonlinearity in the model.

at relatively low intensities, which may soon find practical use in photonics applications. When the field dependence is simple enough to encapsulate into a linear dependence of a scalar index of refraction n on the local field intensity, one has what is known as a Kerr optical nonlinearity. The field induced index change in nonlinear materials can have many interesting effects. A beam can create a lensing action in the medium and either focus itself into an intense filament or defocus to a broader beam. New nonlinear propagation modes can exist. One such mode, called a dark soliton, has many useful properties that could be applied in communication.

A nonlinear interface is the interface between a linear and nonlinear region. Such an interface poses a difficult analytical problem, and has been researched extensively (see Refs. 11–14 and references therein). Though the analytical treatment is more difficult, the numerical solution of the nonlinear interface is no more difficult than that for the linear case. One may simply permit the index in the target media to vary as a function of the field intensity. An example of the nonlinear interface is presented in Figs. 7 and 8. The beam is incident from a region of index $n = n_0$ onto a target medium with an index $n = n_0 - \Delta + n_{nl}|E|^2$. In the cases presented here, $n_0 = 1.5$ and $\Delta = 0.02$, yielding a low-intensity critical angle of 9.37°. In Fig. 7, the nonlinear parameter $n_{nl} = 0.004$. The beam comprises plane waves, each of which sees a reflectivity that depends not only on the wave's incident angle, but on the wave's *intensity* as well. The beam is incident at 8°, far enough under the critical angle [deeper in the TIR regime than the beam in Fig. 4(f)] that there should be almost no transmission. But the evanescent field in the target increased the target index to permit more light to transmit. Both the reflected and transmitted beams are profoundly deformed, in defiance of simple categorizations like the shifts described above. In Fig. 8 the nonlinear parameter is increased to $n_{nl} = 0.011$, and the incident angle is decreased to 5°, deeper into the TIR regime. Decreasing the incident angle is equivalent to reducing the y component of the propagation vector, as is evident in the slower phase fluctuation and fatter interference fringes. Aided by a higher nonlinearity, a portion of the beam manages to punch through, despite the lower angle. This beam would self-focus and collapse to a filament if the nonlinearity were not hard limited by the model, stopping the positive feedback cycle. A saturation of the nonlinearity would be expected in a physical situation.

In general, for a normally TIR situation, as the nonlinearity of the target is increased first the reflected beam exhibits enhanced deformation, then transmission begins. When enough light manages to transmit, self-focusing will occur leading to a collapse of the transmitted beam into filaments. This phenomenon is examined further in Ref. 14.

ACKNOWLEDGMENTS

Computations were performed at the National Center for Supercomputing Applications (NCSA) and at the Cornell Theory Center. Two-dimensional visualization of the numerical solutions was facilitated with NCSA's Image program and associated Hierarchical Data Format. Three-dimensional visualization was done at the Image Analysis Facility at the University of Iowa.

REFERENCES

1. F. Goos and H. Hänchen, Ann. Phys. **1**, 333 (1947).
2. H. K. V. Lotsch, J. Opt. Soc. Am. **58**, 551 (1967).
3. B. R. Horowitz and T. Tamir, J. Opt. Soc. Am. **61**, 586 (1971).
4. M. McGuirk and C. K. Carniglia, J. Opt. Soc. Am. **67**, 103 (1977).
5. C. K. Carniglia and K. R. Brownstein, J. Opt. Soc. Am. **67**, 121 (1977).
6. J. W. Ra, H. L. Bertoni, and L. B. Felsen, SIAM J. Appl. Math. **24**, 396 (1973).
7. Y. M. Antar and W. M. Boerner, Can. J. Phys. **52**, 962 (1974).
8. I. A. White, A. W. Synder, and C. Pask, J. Opt. Soc. Am. **67**, 703 (1977).
9. C. Chiu Chan and T. Tamir, Opt. Lett. **10**, 378 (1985).
10. T. Tamir, J. Opt. Soc. Am. A **3**, 558 (1986).
11. W. J. Tomlinson, J. P. Gordon, P. W. Smith, and A. E. Kaplan, Appl. Opt. **21**, 2041 (1982).
12. P. W. Smith, W. J. Tomlinson, P. J. Maloney, and J. P. Hermann, IEEE J. Quantum Electron. **QE-17**, 340 (1981).
13. P. W. Smith and W. J. Tomlinson, IEEE J. Quantum Electron. **QE-20**, 30 (1984).
14. D. R. Andersen and J. J. Regan, J. Opt. Soc. Am. A **6**, 1484 (1989).