Aust. J. Phys., 1980, 33, 319–35

Shifts of Coherent Light Beams on Reflection at Plane Interfaces between Isotropic Media

R. G. Turner

Department of Physics, Monash University, Clayton, Vic. 3168.

Abstract

Previous experimental and theoretical work on both longitudinal and transverse shifts of light beams at totally reflecting interfaces is briefly reviewed and the discrepancy between the predictions of the two principal theoretical approaches is discussed. A theoretical treatment, valid for an interface between any two media, is presented. The intensity profile of the reflected beam is the same as that of the incident beam (albeit shifted in the reflecting interface) only for certain polarization states of the incident beam and provided that the reflection parameters of the interface meet certain conditions. If these conditions are not met the reflected beam profile suffers distortion and, possibly, deviation from its expected direction. Because the polarization state of a beam is, in general, altered by reflection, measurements of the shifts over a large range of angles of incidence at a single reflection are needed in order to verify the predictions.

Introduction

Although the first observations of a shift, along the reflecting surface, of a light beam undergoing total internal reflection were made in 1947 (Goos and Hanchen 1947), it was predicted much earlier (Picht 1929). Later observations and measurements of the shift have been made by Goos and Hanchen (1949), Mazet *et al.* (1971), Levy and Imbert (1972) and Green *et al.* (1973). The shift is longitudinal (in the plane of incidence). Subsequently, an additional transverse shift (perpendicular to the plane of incidence) was predicted (Fedorov 1955) and observed (Imbert 1969, 1970, 1972; Levy and Imbert 1975). In each case the shifts are generally small (of the order of a wavelength or less) and their magnitudes depend on the angle of incidence, the refractive indices of the media and the polarization state of the beam. The topic has been developed theoretically by a number of authors and a comprehensive review of early work has been given by Lotsch (1970/71).

Because the beam shifts (particularly the longitudinal shift) are largest at or close to the critical angle for total internal reflection (and have only been measured at such angles) attention has been concentrated on the magnitudes of the shifts over a small range of angles of incidence in this region. As has been pointed out by Pavageau (1969), because of abrupt changes in certain reflection parameters at the critical angle, the profile of the reflected beam is distorted when the incident beam is at or very close to this angle, and the magnitude of a shift needs careful definition. (Studies of the magnitudes of the shifts close to the critical angle have been made by Horowitz and Tamir (1971, 1973), Ricard (1973, 1974) and Horowitz (1974).)

The complication just mentioned, however, can be avoided by supposing the angular spread of the incident beam to be more and more restricted as the critical angle is approached and, in this paper, we shall only consider the reflection of incident beams for which this supposition is valid. In these circumstances, expressions for the shifts can be developed which are independent of the profile and angular spread of the incident beam. Among authors dealing with the topic in this way, theoretical investigations have fallen generally into two classes, here called energy-flux conservation treatments and stationary-phase treatments. Both treatments yield similar results for the longitudinal shift at angles of incidence close to the critical angle and because, as mentioned earlier, interest has been concentrated on such angles, little attention has been paid to the fact that they yield different results at other angles of incidence. For the transverse shift, the two treatments yield different results at all angles of incidence. Unfortunately, in addition to the fact that measurements of both shifts have only been made at angles of incidence close to the critical angle, the interpretation of the principal experimental measurements of transverse shifts so far made (Imbert 1969, 1970, 1972) has been called into question (Julia and Neveu 1973; Boulware 1973; Ashby and Miller 1973, 1976). It is clear that more experimental work on both shifts over a large range of angles of incidence is needed.

It is the purpose of this paper to review the bases of the two classes of theoretical treatment and to discuss their validity. It will be pointed out that, whereas the assumptions made in stationary-phase treatments are unobjectionable, energy-flux conservation treatments, as given by previous authors, have involved inappropriate simplifications.

All previous experimental work and most theoretical work has been concerned with the special case of total internal reflection in all-dielectric systems, and there is need for a more general treatment. Such a treatment, valid for all plane reflecting surfaces, will be given here. It will be based on arguments which are extensions of those used in stationary-phase analyses and will yield expressions for both longitudinal and transverse shifts. It will be shown that the reflected beam profile is free of distortion (to a first order) only for particular polarization states of the incident beam.

Previous Work

In this section the application by previous authors of energy-flux conservation and stationary-phase arguments to the evaluation of both longitudinal and transverse shifts will be reviewed and comparisons made between their predictions (they do not, in general, agree). It will be pointed out that energy-flux conservation treatments have involved gross simplifications and have, in any event, only treated the case of total internal reflection at interfaces between dielectrics. While stationary-phase treatments do appear to be valid for the special case of total internal reflection, they are not applicable to the general case of reflection. The possibility of extending the basic ideas involved in stationary-phase treatments to the treatment of reflection at general interfaces will be pointed out.

Longitudinal Shifts

Although the first predictions of the shift of a light beam on total reflection were made about 50 years ago by Picht (1929) using energy-flux conservation arguments, specific expressions for its magnitude were not obtained until about 20 years later by

Artmann (1948), who used a stationary-phase approach, and by von Fragstein (1949), who used energy-flux conservation principles.

Energy-flux conservation arguments depend on the fact that the time-averaged Poynting vector in the evanescent wave above a totally reflecting interface is not zero. The standard energy-flux conservation approach may be taken as that given by Renard (1964) and may be outlined as follows.

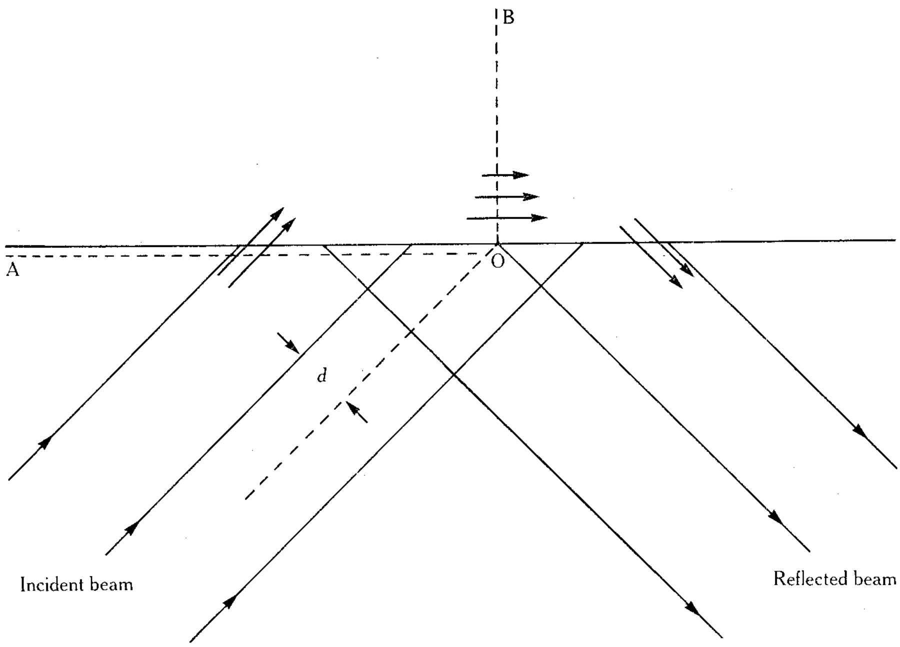

Fig. 1. Energy-conservation requirements can only be met if the reflected beam is displaced so that the average energy flow through OA equals that through OB.

Fig. 1 shows an incident beam and the totally reflected beam; the point O is at the centre of the reflected beam. This latter beam is shown as displaced from its expected position by the distance d for the following reason. There is an energy-flux crossing the plane OB from left to right. If energy is to be conserved, there must be an equal net average energy flux upward through the plane OA. Since the incident and reflected beams have the same intensity, this can only occur if the reflected beam is displaced as shown, leaving an upward energy flux at the left-hand side of the incident beam which is not compensated by a downward flux due to the reflected beam. It is argued that the difference between the upward energy flux due to the incident beam and the downward energy flux due to the reflected beam through the plane OA is equal to the flux carried in a strip of width d near to the centre of the incident beam and so this must equal the total flux over the plane OB due to the evanescent wave. Expressions for these are written, assuming the electric and magnetic fields near to the beam centre, both above and below the interface, are approximated sufficiently closely by those in a plane wave.

The results obtained show that the components of the incident beam linearly polarized perpendicular and parallel to the plane of incidence undergo different shifts $d_\perp$ and $d_\parallel$, which are given by

$$d_\perp = \frac{\lambda}{\pi}\frac{\sin\theta\cos^2\theta}{(1-n^2)(\sin^2\theta-n^2)^{\frac{1}{2}}}, \qquad d_\parallel = \frac{\lambda}{\pi}\frac{\sin\theta\cos^2\theta}{(1-n^2)(\sin^2\theta-n^2)^{\frac{1}{2}}}\frac{n^2}{\sin^2\theta-n^2\cos^2\theta},$$

where λ is the wavelength in the medium containing the incident and reflected waves, θ is the angle of incidence and n is the refractive index of the second medium relative to the first (thus $n < 1$).

A stationary-phase treatment of the problem was first given by Artmann (1948) who was primarily interested in the magnitude of the shift at angles of incidence close to the critical angle. A treatment applicable to other angles of incidence was subsequently given by Schilling (1965). In this approach the incident beam is looked upon as a superposition of plane waves whose propagation vectors cover a range of directions and which, accordingly, undergo different phase changes at reflection. Consequently, at any point P on the reflecting interface, the phase relationships in the reflected ensemble of plane waves are not the same as those in the incident ensemble. Another point on the reflecting interface can be found, however, at which the phase relationships in the reflected ensemble are, at least to a first order, the same as those in the incident ensemble at P. The distance between these two points is the beam shift.

The results of the stationary-phase arguments again show that the components of the incident beam linearly polarized perpendicular and parallel to the plane of incidence undergo different shifts:

$$d_\perp = \frac{\lambda}{\pi}\frac{\sin\theta}{(\sin^2\theta-n^2)^{\frac{1}{2}}}, \qquad d_\parallel = \frac{\lambda}{\pi}\frac{\sin\theta}{(\sin^2\theta-n^2)^{\frac{1}{2}}}\frac{n^2}{\sin^2\theta-n^2\cos^2\theta}.$$

In each case these are different from the expressions resulting from energy-flux conservation treatments by the factor $(1-n^2)/\cos^2\theta$.

Transverse Shifts

It was pointed out by Fedorov (1955) that, since the time-averaged Poynting vector in the evanescent wave in the low-index medium during total internal reflection has, in general, a component perpendicular to the plane of incidence, a transverse shift of the beam is also to be expected. Such shifts were subsequently observed (Imbert 1969, 1970, 1972; Levy and Imbert 1975). An expression for the magnitude of the shift has been developed by Imbert (1968) using energy-flux conservation arguments. It involves the same assumptions as those used in the energy-flux conservation treatment of longitudinal shifts. The magnitude of the transverse shift is found to depend on the polarization state of the incident beam. For the simple case of a circularly polarized incident beam, the transverse shift y^c is given by

$$y^c = \pm\frac{\lambda}{\pi}\frac{\sin^3\theta\cos\theta}{(1-n^2)(\sin^2\theta-n^2\cos^2\theta)},$$

where λ, θ and n are as defined above. The plus and minus signs apply to right-handed and left-handed circularly polarized incident waves.

Stationary-phase analyses of the situation in most of the experimental work, in which the beam undergoes many reflections inside a prism, have been given by Julia and Neveu (1973), Boulware (1973) and Ashby and Miller (1973, 1976). These authors conclude that the measured shift in an experiment of this type is not due to the addition of transverse shifts at each reflection. Indeed, it has been claimed (Canals-Frau 1975) that stationary-phase arguments do not predict any transverse shift at a single reflection. This, however, is at variance with the findings of several other authors (Schilling 1965; Ricard 1974, 1976; Hugonin and Petit 1977). This latter group of authors show that the magnitude of the shift depends on the polarization state of the incident beam and on the angle of incidence and that it occurs at angles of incidence both above and below the critical angle for total internal reflection. They obtain essentially identical results and give expressions for the shifts which are generally applicable to all polarization states of the incident beam and all angles of incidence. For the purpose of illustrating the difference between their results and those provided by energy-flux conservation arguments it will suffice to quote the result for the case of a circularly-polarized incident beam undergoing total internal reflection. It is

$$y^{\mathrm{c}} = \pm\frac{\lambda}{\pi}\frac{\sin^3\theta\cos\theta}{\sin^2\theta - n^2\cos^2\theta},$$

which differs from the energy-flux conservation expression by the factor $(1-n^2)$.

Comparison of Bases of Energy-flux Conservation and Stationary-phase Treatments

There is disagreement between the predictions of energy-flux conservation and stationary-phase treatments, and consequently the bases of each will now be examined more closely. The possibility of a treatment that is more general than any previously given will then be discussed.

In the standard energy-flux conservation treatments the Poynting vector, in the region of overlap of the incident and reflected beams, is thought of as simply the sum of the Poynting vectors in the two beams taken separately. This is not in general true; there is no principle of superposition for Poynting vectors. The only valid procedure is to evaluate the integral of the time-averaged Poynting vector over the whole plane OA in Fig. 1. Obviously, the beams cannot then be thought of as simple plane waves; indeed the concept of a shift for such a wave has no meaning. This same criticism of the standard energy-flux conservation treatment has recently been made by Agudin and Platzeck (1978).

By contrast, stationary-phase arguments involve no assumptions other than the validity of the procedure of resolving the beam into an ensemble of plane waves and the linearity of the electromagnetic wave equation and of the boundary condition on the electric and magnetic fields at the interface.

While a stationary-phase condition is adequate to treat the case of total internal reflection at an all-dielectric interface, the general case of reflection at any interface requires a further condition; in addition to the phase relationships, the amplitude ratios in the incident and reflected plane-wave ensembles must also be constant. A treatment incorporating both of these conditions may be termed a stationary-amplitude treatment (complex amplitude). Such a treatment for the reflection, at any single plane interface, of a beam in any fully polarized spatially coherent state will be given in the next section, and will yield expressions for both the longitudinal and transverse shifts. The longitudinal shift is due essentially to the dispersion of the phase change at reflection. The transverse shift is essentially a geometrical effect and results from a mixing of the linearly polarized components of the plane waves constituting the beam, due to their different planes of incidence. The results will show that the profile of the reflected beam is, in general, undistorted only if the range of angles of incidence covered by the incident plane-wave ensemble is small enough for both real and imaginary parts of the reflection coefficient to be regarded as varying linearly with angle of incidence; and, further, that the conditions on the polarization state of the incident beam for undistorted shifts, even of the principal linearly polarized components of the beam are, in general, much more stringent than has been pointed out previously.

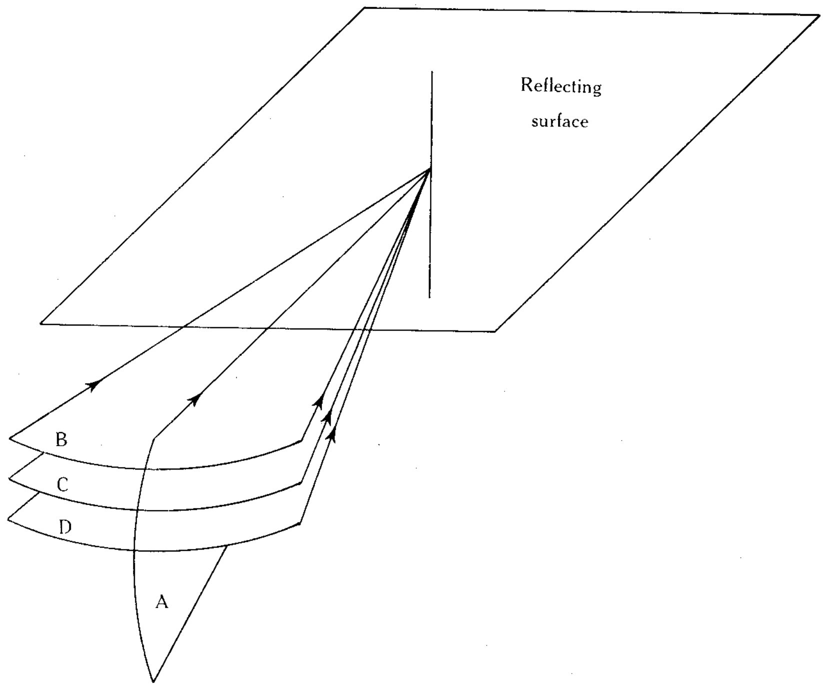

Fig. 2. Planes containing the propagation vectors of the principal group (A) and three minor groups (B, C, D) of the plane-wave components of an incident beam.

Stationary-amplitude Analysis

In this section, the conditions will be sought under which the profile of a light beam reflected from a plane interface between any two media is the same as that of the incident beam. It will be shown that, in general, this does not occur but that within a first-order approximation (ignoring effects dependent on higher powers of the beam's angular divergence) the two components of the reflected beam polarized perpendicular and parallel to the plane of incidence may individually retain the profile of the incident beam (although each is shifted by a different distance along the interface), provided that the reflection parameters of the interface and the polarization of the incident beam fulfil certain conditions. The approach will be to consider the incident beam as an ensemble of plane waves and to seek a point in the interface at which the amplitude ratios and phase differences between pairs of plane waves in the reflected ensemble are the same as those between all corresponding pairs in the incident ensemble.

The incident light beam is considered as a superposition of plane waves whose propagation vectors occupy only a small angular range, and one of these, whose propagation-vector direction is approximately central within the ensemble, is nominated as the principal wave. If the beam shape is symmetrical it is of course possible to define a principal wave more precisely but, in general, more exact definition is unnecessary. The plane of incidence of the principal wave will be called the principal plane of incidence and the group of plane waves which have this plane as their plane of incidence will be called the principal group. A group of plane waves whose propagation vectors are parallel to any given plane perpendicular to the principal plane of incidence will be called a minor group.

Fig. 2 illustrates the meaning of the terms 'principal group' and 'minor group'. Plane A is the principal plane of incidence and contains the propagation vectors of the principal group. Planes B, C and D are perpendicular to plane A and each contains the propagation vectors of a minor group.

If the intensity profile of the reflected beam is to be identical with that of the incident beam except for an inversion with respect to the reflecting surface (and a possible uniform reduction in intensity), the relative phases and amplitudes and the angular relationships between the field vectors of the plane waves constituting the beam must be unchanged by reflection. This condition cannot be met in the most general sense because the s and p (TE and TM) components of a plane wave undergo different phase changes at reflection (resulting in the polarization state of the reflected wave differing, in general, from that of the incident wave). As will be shown, however, it can be met (at least within a first-order approximation) separately for the components of the plane waves whose magnetic vectors are parallel to the principal plane of incidence and the components whose electric vectors are parallel to this plane. The electric vectors of these two components will be called $E_\perp$ and $E_\parallel$ respectively. For the plane waves of the principal group, $E_\perp$ and $E_\parallel$ are s and p components respectively at reflection. For waves not in the principal group, however, this will not be so.

The cartesian coordinate system to be used is shown in Fig. 3. The x and y axes lie in the reflecting surface. The incident plane-wave ensemble is represented by the cone of propagation-vector directions at the approximate centre of which is the propagation vector of the principal wave. The angle of incidence of the principal wave is θ. The axes are oriented so as to make the z–x plane the principal plane of incidence.

Expressions for $E_\perp$ and $E_\parallel$ for a typical plane wave of the ensemble have the form (omitting the factor $e^{-i\omega t}$ which plays no part in the analysis)

$$E_\perp = E_0 \exp\big(i\{k(\alpha x+\beta y+\gamma z)+\phi\}\big), \qquad E_\parallel = R\,E_0 \exp\big(i\{k(\alpha x+\beta y+\gamma z)+\phi+\psi\}\big), \tag{1}$$

where k is the wave number $(2\pi/\lambda)$, α, β, γ are the direction cosines of the propagation vector ($\sin\theta$, 0, $\cos\theta$ for the principal wave), ϕ is the phase of the wave relative to that of the principal wave at the origin, ψ is the phase difference between the two orthogonally polarized components, and E_0 and R are real. Any fully polarized plane wave can be expressed in this form by suitable choices of R and ψ.

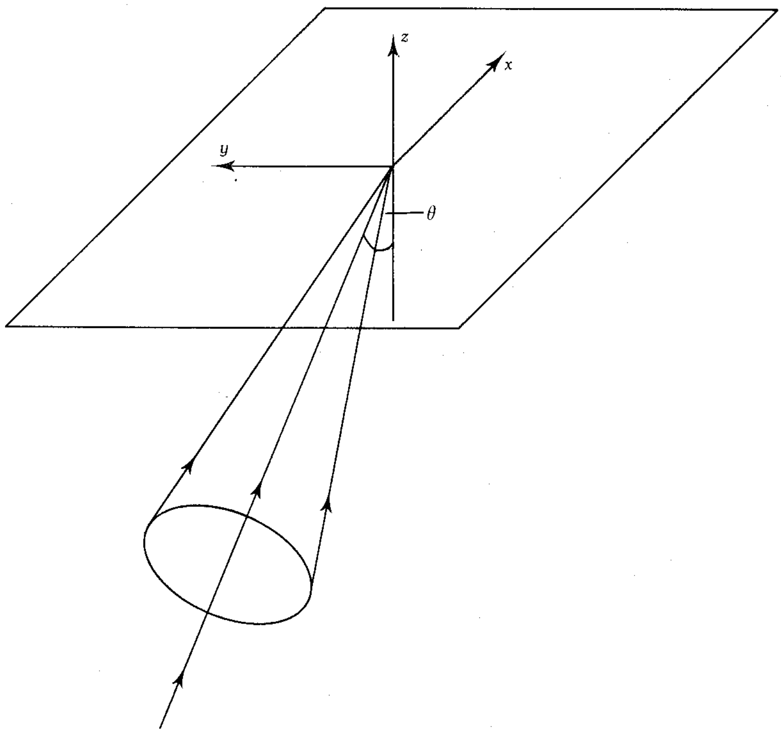

Fig. 3. Cone of propagation-vector directions of the plane-wave components of a beam incident at a principal angle of incidence θ on a reflecting surface.

As stated above, the two components (1) are not, in general, s and p components at the reflecting surface. Suppose $E_\perp$ makes an angle ε with the plane of the surface. The electric vectors of the s and p components are then

$$\begin{aligned} E_s &= E_\perp \cos\varepsilon + E_{\|} \sin\varepsilon \\ &= E_0 \exp\big(\mathrm{i}\{k(\alpha x+\beta y+\gamma z)+\phi\}\big)\,\big(\cos\varepsilon + R\exp(\mathrm{i}\psi)\sin\varepsilon\big), \\ E_p &= E_{\|}\cos\varepsilon - E_\perp \sin\varepsilon \\ &= E_0 \exp\big(\mathrm{i}\{k(\alpha x+\beta y+\gamma z)+\phi\}\big)\,\big(R\exp(\mathrm{i}\psi)\cos\varepsilon - \sin\varepsilon\big). \end{aligned}$$

Introducing reflection coefficients $\rho_s \exp(\mathrm{i}\delta_s)$ and $\rho_p \exp(\mathrm{i}\delta_p)$, where ρ is real and δ is the phase delay at the reflection, we find that the electric vectors of the s and p components of the plane wave after reflection are

$$\begin{aligned} E'_s &= E_0 \exp\big(\mathrm{i}\{k(\alpha x+\beta y-\gamma z)+\phi\}\big)\,\big(\cos\varepsilon + R\exp(\mathrm{i}\psi)\sin\varepsilon\big)\rho_s \exp(\mathrm{i}\delta_s), \\ E'_p &= E_0 \exp\big(\mathrm{i}\{k(\alpha x+\beta y-\gamma z)+\phi\}\big)\,\big(R\exp(\mathrm{i}\psi)\cos\varepsilon - \sin\varepsilon\big)\rho_p \exp(\mathrm{i}\delta_p). \end{aligned}$$

In order to compare the structures of the incident and reflected beams, expressions are needed for the magnitudes $E'_\perp$ and $E'_\parallel$ of the electric vectors of the components of the reflected wave whose magnetic and electric vectors respectively are parallel to the principal plane of incidence. These expressions are

$$E'_\perp = E'_s \cos\varepsilon + E'_p \sin\varepsilon, \qquad E'_\parallel = E'_p \cos\varepsilon - E'_s \sin\varepsilon$$

and so, defining the quantities $r_\perp$ and $r_\parallel$ by

$$r_\perp = E'_\perp / E_{\perp 0}, \qquad r_\parallel = E'_\parallel / E_{\parallel 0},$$

where $E_{\perp 0}$ and $E_{\parallel 0}$ are the magnitudes of the electric-vector components of the incident wave at the origin, we find

$$r_\perp = \exp\big(ik(\alpha x+\beta y-\gamma z)\big)\big\{\big(\cos\varepsilon + R\exp(i\psi)\sin\varepsilon\big)\rho_s \exp\big(i\delta_s\big)\cos\varepsilon + \big(R\exp(i\psi)\cos\varepsilon - \sin\varepsilon\big)\rho_p \exp\big(i\delta_p\big)\sin\varepsilon\big\}, \tag{2a}$$

$$r_\parallel = \exp\big(ik(\alpha x+\beta y-\gamma z)\big)\big\{\big(\cos\varepsilon - R^{-1}\exp(-i\psi)\sin\varepsilon\big)\rho_p \exp\big(i\delta_p\big)\cos\varepsilon - R^{-1}\exp(-i\psi)(\cos\varepsilon + \sin\varepsilon)\rho_s \exp\big(i\delta_s\big)\sin\varepsilon\big\}. \tag{2b}$$

The quantities $r_\perp$ and $r_\parallel$ can be looked upon as effective reflection coefficients for the plane waves.

If the relative phases and amplitudes of the perpendicular and parallel components of the plane waves in the reflected ensemble are to be the same as those in the incident ensemble, $r_\perp$ and $r_\parallel$ must be independent of the propagation-vector direction. This is ensured by requiring the constancy of $r_\perp$ and $r_\parallel$:

(i) for the waves of the principal group;

(ii) for the waves of a minor group.

It will be shown that these two conditions yield expressions for the longitudinal and transverse shifts respectively of the beam.

Longitudinal Shift

For the planes waves constituting the principal group, we have $\beta = 0$ (and $\varepsilon = 0$) and so $r_\perp$ and $r_\parallel$ can be expressed as functions of α only by putting $\gamma = (1-\alpha^2)^{\frac{1}{2}}$ and $\varepsilon = 0$ in equations (2a) and (2b). Writing α_0, $r_{\perp 0}$ and $r_{\parallel 0}$ for the values of α, $r_\perp$ and $r_\parallel$ for the principal wave, we can expand $r_\perp$ and $r_\parallel$ as a Taylor series in $\alpha-\alpha_0$:

$$r_\perp - r_{\perp 0} = (\alpha-\alpha_0)\,\mathrm{d}r_\perp/\mathrm{d}\alpha + \tfrac{1}{2}(\alpha-\alpha_0)^2\,\mathrm{d}^2 r_\perp/\mathrm{d}\alpha^2 + \ldots, \tag{3a}$$

$$r_\parallel - r_{\parallel 0} = (\alpha-\alpha_0)\,\mathrm{d}r_\parallel/\mathrm{d}\alpha + \tfrac{1}{2}(\alpha-\alpha_0)^2\,\mathrm{d}^2 r_\parallel/\mathrm{d}\alpha^2 + \ldots. \tag{3b}$$

For the expressions (3) to vanish exactly, all derivatives of $r_\perp$ and $r_\parallel$ with respect to α must vanish. However, provided $\alpha-\alpha_0$ is always sufficiently small (i.e. the angular plane-wave spectrum of the beam is sufficiently narrow) all terms except the first may be ignored. Then, for the wave components with magnetic vectors parallel to the principal plane of incidence, $\mathrm{d}r_\perp/\mathrm{d}\alpha$ vanishes when

$$i\left\{k\left(x + \frac{\alpha z}{(1-\alpha^2)^{\frac{1}{2}}}\right) + \frac{\mathrm{d}\delta_s}{\mathrm{d}\alpha}\right\} + \frac{1}{\rho_s}\frac{\mathrm{d}\rho_s}{\mathrm{d}\alpha} = 0, \tag{4a}$$

and, for the components with electric vectors parallel to the principal plane of incidence, $\mathrm{d}r_{\parallel}/\mathrm{d}\alpha$ vanishes when

$$\mathrm{i}\left\{k\left(x+\frac{\alpha z}{(1-\alpha^2)^{\frac{1}{2}}}\right)+\frac{\mathrm{d}\delta_p}{\mathrm{d}\alpha}\right\}+\frac{1}{\rho_p}\frac{\mathrm{d}\rho_p}{\mathrm{d}\alpha}=0. \tag{4b}$$

The shifts along the reflecting surface of the reflected beam are found by putting $z = 0$ (a comment on this will be made in the Discussion below):

$$x_\perp = -\frac{1}{k}\frac{\mathrm{d}\delta_s}{\mathrm{d}\alpha}+\frac{\mathrm{i}}{k\rho_s}\frac{\mathrm{d}\rho_s}{\mathrm{d}\alpha}, \qquad x_\parallel = -\frac{1}{k}\frac{\mathrm{d}\delta_p}{\mathrm{d}\alpha}+\frac{\mathrm{i}}{k\rho_p}\frac{\mathrm{d}\rho_p}{\mathrm{d}\alpha}. \tag{5}$$

Transverse Shift

For plane waves of a minor group, $r_\perp$ and $r_\parallel$ can be expressed as functions of β only by putting $\alpha = \alpha_0(1-\beta^2)^{\frac{1}{2}}$ and $\gamma = \gamma_0(1-\beta^2)^{\frac{1}{2}}$ into equations (2a) and (2b), where $\alpha_0, 0, \gamma_0$ are now the direction cosines of the propagation vector of the wave of this minor group which is also a member of the principal group. The symbols $r_{\perp 0}$ and $r_{\parallel 0}$ are now used to denote the values of $r_\perp$ and $r_\parallel$ for this latter wave. The quantities $r_\perp$ and $r_\parallel$ can be expanded as a Taylor series in β as

$$r_\perp - r_{\perp 0} = \beta\,\mathrm{d}r_\perp/\mathrm{d}\beta + \tfrac{1}{2}\beta^2\,\mathrm{d}^2r_\perp/\mathrm{d}\beta^2 + \ldots, \tag{6a}$$

$$r_\parallel - r_{\parallel 0} = \beta\,\mathrm{d}r_\parallel/\mathrm{d}\beta + \tfrac{1}{2}\beta^2\,\mathrm{d}^2r_\parallel/\mathrm{d}\beta^2 + \ldots. \tag{6b}$$

Assuming again that the angular plane-wave spectrum is sufficiently narrow that only the first terms of the series need be retained, we find that the expressions (6) vanish when, for the wave components with magnetic vectors parallel to the principal plane of incidence,

$$\mathrm{i}ky+R\left(\exp(\mathrm{i}\psi)+\frac{\rho_p}{\rho_s}\exp\{\mathrm{i}(\psi+\delta_p-\delta_s)\}\right)\frac{\mathrm{d}\varepsilon}{\mathrm{d}\beta}=0, \tag{7a}$$

and, for the components with electric vectors parallel to the principal plane of incidence,

$$\mathrm{i}ky-\frac{1}{R}\left(\exp(-\mathrm{i}\psi)+\frac{\rho_s}{\rho_p}\exp\{-\mathrm{i}(\psi+\delta_p-\delta_s)\}\right)\frac{\mathrm{d}\varepsilon}{\mathrm{d}\beta}=0. \tag{7b}$$

It can be shown that $\cos\varepsilon = \alpha_0(\alpha_0^2+\beta^2\gamma_0^2)^{-\frac{1}{2}}$ and so

$$\mathrm{d}\varepsilon/\mathrm{d}\beta = \gamma_0/\alpha_0.$$

As long as the angular width of the beam's plane-wave spectrum is small, α_0 and γ_0 can be taken as the values of α and γ for the principal wave and so $\mathrm{d}\varepsilon/\mathrm{d}\beta \approx \cot\theta$. The conditions (7a) and (7b) then become

$$y_\perp = \frac{\mathrm{i}R}{k}\left(\exp(\mathrm{i}\psi)+\frac{\rho_p}{\rho_s}\exp(\mathrm{i}\psi')\right)\cot\theta, \tag{8a}$$

$$y_\parallel = -\frac{\mathrm{i}}{kR}\left(\exp(-\mathrm{i}\psi)+\frac{\rho_s}{\rho_p}\exp(-\mathrm{i}\psi')\right)\cot\theta, \tag{8b}$$

where ψ' has been written for $\psi+\delta_p-\delta_s$, the phase difference between the orthogonal components of the waves after reflection.

Application to Dielectric Media

Equations (5) and (8) are generally applicable to reflection at a plane interface between any two media. Since, however, previous authors have dealt exclusively with the case of total internal reflection in all-dielectric systems, it is appropriate for purposes of comparison to examine the form taken by equations (5) and (8) in such a situation. Before doing so, however, the simplification resulting from the assumption that only the medium containing the incident wave is a dielectric will be examined. The simplification results from the fact that k is real. These results are applicable to cases of practical interest such as the reflection of light beams at metallic surfaces.

Incident Waves in Dielectric Medium

The wave number k is real and, since ρ and δ are real by definition, the equations (5) yield the pairs of conditions

$$x_\perp = -k^{-1}\,\mathrm{d}\delta_s/\mathrm{d}\alpha, \qquad \mathrm{d}\rho_s/\mathrm{d}\alpha = 0\,; \tag{9a}$$

$$x_\parallel = -k^{-1}\,\mathrm{d}\delta_p/\mathrm{d}\alpha, \qquad \mathrm{d}\rho_p/\mathrm{d}\alpha = 0\,. \tag{9b}$$

The second of each pair is the condition for the amplitude ratios between the plane waves in the reflected ensemble to be the same as those in the incident ensemble. They are, of course, only satisfied if ρ is stationary with respect to angle of incidence at the angle of incidence θ. As has been pointed out by White *et al.* (1977), if this is not so, the reflected beam suffers distortion of its profile and deviation from its expected direction. These effects can be reduced to any desired extent (unless ρ undergoes a discontinuity within the angular plane-wave spectrum) by restricting the angular width of the beam's plane-wave spectrum (consequently increasing the beam's linear width). With the assumption that the second of each pair of conditions is satisfied or may be ignored, the first of each pair (the condition for the phase relationships between the plane waves to be the same in the incident and reflected ensembles) implies a shift, in the plane of incidence, of the reflected beam from its expected position.

Equations (8) yield, when k is real, the pairs of conditions

$$y_\perp = -\frac{R}{k}\left(\sin\psi + \frac{\rho_p}{\rho_s}\sin\psi'\right)\cot\theta, \qquad R\left(\cos\psi + \frac{\rho_p}{\rho_s}\cos\psi'\right) = 0\,; \tag{10a}$$

$$y_\parallel = -\frac{1}{Rk}\left(\sin\psi + \frac{\rho_s}{\rho_p}\sin\psi'\right)\cot\theta, \qquad \frac{1}{R}\left(\cos\psi + \frac{\rho_s}{\rho_p}\cos\psi'\right) = 0\,. \tag{10b}$$

The second of each pair (the amplitude-ratio condition) determines the value of ψ if the profile of the reflected beam is to be undistorted. The first of each pair (the phase condition) indicates a shift in the reflected beam in a direction perpendicular to the principal plane. Equations (9) and (10) show that, if the centre of the incident beam falls at the origin, the centre of the reflected-beam component polarized perpendicular to the plane of incidence will leave the surface at $x_\perp, y_\perp$ and the component polarized parallel to the plane of incidence at $x_\parallel, y_\parallel$.

Since all experimental work has been performed on plane interfaces between two dielectric materials, expressions for the shifts will now be found for this case.

Plane Interface between Two Dielectric Media

It is convenient to consider separately two ranges of α:

Range 1. $0 \leqslant \alpha \leqslant \sin\theta_c$, where θ_c is the critical angle. In this range we have

$$\mathrm{d}\delta_s/\mathrm{d}\alpha = \mathrm{d}\delta_p/\mathrm{d}\alpha = 0$$

(excluding the discontinuity in δ_p at the Brewster angle), and

$$\rho_s = \frac{(1-\alpha^2)^{\frac{1}{2}}-(n^2-\alpha^2)^{\frac{1}{2}}}{(1-\alpha^2)^{\frac{1}{2}}+(n^2-\alpha^2)^{\frac{1}{2}}}, \qquad \rho_p = \frac{n^2(1-\alpha^2)^{\frac{1}{2}}-(n^2-\alpha^2)^{\frac{1}{2}}}{n^2(1-\alpha^2)^{\frac{1}{2}}+(n^2-\alpha^2)^{\frac{1}{2}}},$$

where n has been written for the ratio of the refractive indices of the second and first media. To facilitate comparison with the results of previous authors, the values of x from equations (9) will be converted to shifts, in a direction perpendicular to the beam's propagation direction, by multiplying by $\cos\theta$. Equations (9) and (10) then give

$$d_\perp = d_\parallel = 0, \tag{11a}$$

$$y_\perp = \pm\frac{\lambda R}{\pi}\left(\frac{\sin\theta\cos\theta}{\sin^2\theta+\cos\theta\,(n^2-\sin^2\theta)^{\frac{1}{2}}}\right), \tag{11b}$$

$$y_\parallel = \pm\frac{\lambda}{R\pi}\left(\frac{\sin\theta\cos\theta}{\sin^2\theta-\cos\theta\,(n^2-\sin^2\theta)^{\frac{1}{2}}}\right), \tag{11c}$$

the plus or minus signs in equations (11b) and (11c) corresponding to $\psi = \mp\frac{1}{2}\pi$. The second of the pairs of conditions in equations (9) is not satisfied and the reflected beam is distorted and deviated.

Range 2. The second range of α (applicable only for $n < 1$) is $\sin\theta_c < \alpha \leqslant 1$ (the range of total internal reflection). In this range

$$\rho_s = \rho_p = 1,$$

$$\tan\tfrac{1}{2}\delta_s = -\{(\alpha^2-n^2)/(1-\alpha^2)\}^{\frac{1}{2}}, \qquad \tan\tfrac{1}{2}\delta_p = -n^{-2}\{(\alpha^2-n^2)/(1-\alpha^2)\}^{\frac{1}{2}}.$$

Equations (9) and (10) then give

$$d_\perp = \frac{\lambda}{\pi}\frac{\sin\theta}{(\sin^2\theta-n^2)^{\frac{1}{2}}}, \qquad d_\parallel = \frac{\lambda}{\pi}\frac{\sin\theta}{(\sin^2\theta-n^2)^{\frac{1}{2}}}\frac{n^2}{\sin^2\theta-n^2\cos^2\theta}; \tag{12a}$$

$$y_\perp = \pm\frac{\lambda R}{\pi}\frac{\sin\theta\cos\theta}{(\sin^2\theta-n^2\cos^2\theta)^{\frac{1}{2}}}, \qquad y_\parallel = \pm\frac{\lambda}{R\pi}\frac{\sin\theta\cos\theta}{(\sin^2\theta-n^2\cos^2\theta)^{\frac{1}{2}}}, \tag{12b}$$

the plus or minus signs in equations (12b) corresponding to $\psi = \mp\frac{1}{2}\pi-\frac{1}{2}(\delta_p-\delta_s)$.

Discussion

The foregoing analysis, resulting in equations (5) and (8), is applicable to any plane reflecting surface. It is a first-order analysis in the sense that all terms except the first in the Taylor series of equations (3) and (6) have been assumed negligible. In fact, the analysis could have been extended, in the case of the longitudinal shift, by requiring both $\mathrm{d}r/\mathrm{d}\alpha$ and $\mathrm{d}^2r/\mathrm{d}\alpha^2$ in equations (3) to vanish. It is easy to show that, when ρ is constant, $\mathrm{d}^2r/\mathrm{d}\alpha^2$ vanishes when, in addition to equations (4), we have

$$z = -k^{-1}(1-\alpha^2)^{3/2}\,\mathrm{d}^2\delta/\mathrm{d}\alpha^2\,.$$

This implies a shift not wholly parallel to the reflecting surface (i.e. a shift of the reflected beam along its propagation direction). It does not, of course, affect the validity of equations (5) (which give the value of x when $z = 0$) or of any equations derived from them. An effect of this sort has been pointed out previously by Julia and Neveu (1973) and by McGuirk and Carniglia (1977).

In general, terms higher than the second in the series of equations (3) and the first in those of equations (6) will not vanish simultaneously. Consequently, if the angular plane-wave spectrum of the beam straddles angles of incidence over which rapid changes of the reflection parameters or their derivatives occur, it would be expected that the profile of the reflected beam would be distorted and that, as shown by White *et al.* (1977), it might be deviated from its expected direction (the results of Read *et al.* (1978), who used microwave beams, appear to be, at least in part, a demonstration of this).

Another fact which the present analysis makes clear is that, even if the distorting effects referred to in the previous paragraph are negligible, the profile of the reflected beam as a whole is, in general, distorted since the components of the beam linearly polarized perpendicular and parallel to the plane of incidence undergo different shifts. Moreover, even if one such component in the reflected beam is isolated, only certain polarization states of the incident beam can result in an undistorted reflected beam (in the particular case of total internal reflection, equations (12) show that there is a unique angle of incidence, namely $\arccos\{(1-n^2)/(1+n^2)\}^{\frac{1}{2}}$, at which, for two particular elliptically polarized states of the incident beam, all components suffer the same shift). The results referred to in this paragraph accord with the findings of Hugonin and Petit (1977) in their study of the quality of images formed by reflection at totally reflecting surfaces and of Costa de Beauregard and Imbert (1972, 1973).

The results of the present analysis will now be compared with those of previous authors. These have mainly treated specifically the special case of total reflection at an all-dielectric interface (indeed the energy-flux conservation treatment, in the simple form presented by authors using this approach, cannot deal with the case of reflection by an absorbing medium). A comparison will therefore be made between the equations (12) and the corresponding expressions previously obtained.

Longitudinal Shifts

The values of $d_\perp$ and $d_\parallel$ in equations (12a) are the same as those obtained by previous authors using stationary-phase methods. They also agree with those of Agudin (1968) who uses Fermat's principle to evaluate the shifts. They are, however, greater by the factor $(1-n^2)/\cos^2\theta$ than expressions obtained using the standard energy-flux conservation treatment. The two treatments thus agree in the immediate neighbourhood of the critical angle but not at greater angles of incidence. In particular, $d_\perp$ and $d_\parallel$ in equations (12a) approach the values

$$d_\perp = \lambda/\pi(1-n^2)^{\frac{1}{2}}, \qquad d_\parallel = n^2\lambda/\pi(1-n^2)^{\frac{1}{2}}$$

as the grazing incidence is approached, whereas the energy-flux conservation treatment predicts that $d_\perp$ and $d_\parallel$ both approach zero. It has been argued (e.g. Renard 1964) that nonzero values of the shift at grazing incidence are unreasonable on the grounds that the incident and reflected beams should be indistinguishable at this angle. In fact this is untrue; the reflected beam would be distinguished from the incident beam in that its phase would be reversed. More importantly, however, since the concept of angle of incidence only has meaning for a plane wave, the notion of truly grazing incidence for a beam of finite width is unrealizable because of the range of propagation directions of the plane waves into which it must be considered as resolved. The beam can never suffer a zero deviation and, consequently, the objection is groundless.

Since the shifts are large at angles of incidence only slightly greater than the critical angle, experimental measurements have been confined almost entirely to such angles at which, unfortunately, the disagreement between the results from energy-flux conservation and stationary-phase treatments is extremely small. Rhodes and Carniglia (1977) claim to have evidence for a nonzero shift near to grazing incidence from studies of the positions of interference fringes in a Lloyd's-mirror experiment in which the mirror was a totally reflecting interface. Care is needed, however, in interpreting these results as evidence of a longitudinal shift; the position of a fringe depends on the phase change at reflection rather than on its derivative, and an alteration in fringe position as n alters merely indicates an alteration in this phase change.

Transverse Shifts

Expressions for the transverse shift of a beam suffering total reflection at an all-dielectric interface have been derived by several authors (e.g. Imbert 1968; Ricard 1970) using what is essentially an energy-flux conservation argument. The expression developed by Imbert (1968) for the transverse shift suffered by a circularly polarized beam incident at the critical angle is

$$y^c = \pm\lambda/(\pi\sin\theta\cos\theta),$$

the plus and minus signs applying to right and left circularly polarized beams. This expression differs from that derived from equations (12) in that it is larger by the factor $\sec^2\theta$. The view has been expressed (Costa de Beauregard and Imbert 1972, 1973) that circularly polarized evanescent waves can be regarded as eigenmodes and that the totally reflecting interface, in effect, resolves the incident wave into two elliptically polarized components which excite these modes. The present analysis does not support this view; the polarization states into which the incident beam is resolved by the interface are essentially linearly polarized perpendicular and parallel to the plane of incidence. An incident beam, for example, linearly polarized either perpendicular or parallel to the plane of incidence would, from the eigenmode view-point, be expected to excite both modes and the reflected beam would be expected to emerge as two components shifted transversely in opposite directions. Equations (12), however, show that the shifts for such beams would be purely longitudinal. It is true that (except for the linearly polarized beams just mentioned) the whole reflected beam will suffer a uniform transverse shift only if the evanescent wave is circularly polarized. However, equations (12) show that perpendicular and parallel components of the beam separately undergo undistorted shifts (both longitudinal and transverse) as long as the evanescent wave is elliptically polarized with a principal axis of the ellipse parallel to the interface (incident waves linearly polarized perpendicular and parallel to the plane of incidence are limiting cases of this condition).

Stationary-phase treatments of the transverse shift at a single reflecting surface have been given by Schilling (1965), Ricard (1974, 1976) and Hugonin and Petit (1977). Their results differ from those of equations (10) in that, since a stationary-phase and not a stationary-amplitude treatment is used, only the first of each pair of conditions in equations (10) is obtained and no restriction is placed on ψ. In addition, separate shifts for the perpendicular and parallel components of the beam are not distinguished and the shift for the whole beam is given as an intensity-weighted mean of $y_\perp$ and $y_\parallel$, valid for all values of ψ.

The only experimental measurements of the transverse shift to which the present analysis might, at first sight, apply appear to be those of Imbert (1969, 1970, 1972). These have involved the use of a prism inside which the beam undergoes many reflections while following a helix-like path. In view of the nonplanar path of the beam it is difficult to analyse the results in terms of the effect of a single reflection. In addition, because the polarization state of the beam is only preserved (approximately) between successive reflections at angles of incidence very close to the critical angle, it is impossible to measure the shift over a range of angles of incidence in this way. Stationary-phase analyses of this type of experiment have been given by Julia and Neveu (1973), Boulware (1973) and Ashby and Miller (1973, 1976). These authors consider the propagation of polarization eigenstates through the successive reflections and reach the conclusion that the helix-like path gives rise to a resultant shift which is composed, at least partly, of components of the longitudinal shift at each reflection. They disagree with the general expressions for the shift given by Imbert and claim that the agreement between his theoretical and experimental results is fortuitous.

The transverse and longitudinal shifts at a quasi-single reflection have been observed by Levy and Imbert (1972, 1975). For the transverse shift, the 'surface' was a four-layer dielectric stack and, for the longitudinal shift, a two-layer stack. The layers were deposited on a transparent substrate, and their thicknesses were chosen so that the phase change at reflection varied very rapidly (and the energy-flux in the evanescent wave became very large) over a small range of angles of incidence ($\sim 0\cdot 01°$). In view of the large variations in the phase change and its derivatives over the range of angles of incidence occupied by the plane waves constituting the beam, the approximations used in the present analysis are invalid in this situation.

Conclusions

A method has been presented for evaluation of the longitudinal and transverse shifts of a light beam on reflection at a plane interface between any two media. The predictions of the longitudinal shifts for the case of total internal reflection at an interface between dielectrics agree, as expected, with the predictions of stationary-phase treatments; they do not agree with the predictions of the simple energy-flux conservation treatments so far presented. The predicted transverse shifts for the case of total internal reflection at an interface between dielectrics agree in part with those of the previous published stationary-phase treatments but not with the general interpretation of those treatments; they do not agree with the predictions of authors using energy-flux conservation treatments.

Energy-flux conservation treatments have involved the simplification of treating both the incident and evanescent waves as plane waves. By contrast, the stationary-amplitude treatment depends solely on the validity of the procedure of resolving the beam into a set of plane waves and on the linearity of the boundary conditions to which the electric and magnetic fields are subject at the interface, and it is difficult to see how any error can arise in its application. It is equally difficult, however, to believe that either approach is wrong in principle and it is likely that more rigorous energy-flux conservation treatments would yield results agreeing with those obtained by the stationary-amplitude method.

For situations involving very rapid changes in the reflection parameters (the dielectric stacks of Levy and Imbert (1972, 1975) or a single dielectric interface at the critical angle or Brewster angle), analytic treatments fail and resort must be had to numerical computation of the intensity distribution in the reflected beam. In such cases the beam suffers distortion of its profile and possibly deviation from its expected direction and the notion of a 'shift' becomes indistinct.

Experimental measurements of both types of shift have not, so far, been carried out over a range of angles of incidence sufficient to resolve the disagreement between the two treatments. It is clear that measurements of the shifts over a large range of angles of incidence are needed. Such measurements (at least for transverse shifts) must be carried out using single reflecting surfaces because of the change in the polarization state of the beam at each reflection.

Acknowledgments

The author is indebted to Dr G. J. Troup and Dr J. L. A. Francey of Monash University and Mr R. M. Sillitto of Edinburgh University for stimulating discussions. He also wishes to thank Mr R. C. Tobin and Dr A. C. McLaren for reading the manuscript and suggesting improvements.

References

Agudin, J. L. (1968). *Phys. Rev.* **171**, 1385.
Agudin, J. L., and Platzeck, A. M. (1978). *J. Opt.* (*Paris*) **9**, 101.
Artmann, K. (1948). *Ann. Phys.* (*Leipzig*) **2**, 87.
Ashby, N., and Miller, S. C. (1973). *Phys. Rev.* D **7**, 2383.
Ashby, N., and Miller, S. C. (1976). *Phys. Rev.* D **13**, 3219.
Boulware, D. G. (1973). *Phys. Rev.* D **7**, 2375.
Canals-Frau, D. (1975). *Nouv. Rev. Opt.* **6**, 203.
Costa de Beauregard, O., and Imbert, C. (1972). *Phys. Rev. Lett.* **28**, 1211.
Costa de Beauregard, O., and Imbert, C. (1973). *Phys. Rev.* D **7**, 3555.
Fedorov, F. I. (1955). *Dokl. Akad. Nauk SSSR* **105**, 465.
von Fragstein, C. (1949). *Ann. Phys.* (*Leipzig*) **4**, 271.
Goos, F., and Hanchen, H. (1947). *Ann. Phys.* (*Leipzig*) **1**, 333.

Goos, F., and Hanchen, H. (1949). *Ann. Phys. (Leipzig)* **5**, 251.
Green, M., Kirkby, P., and Timsit, R. S. (1973). *Phys. Lett.* A **45**, 259.
Horowitz, B. R. (1974). *Appl. Phys.* **3**, 411.
Horowitz, B. R., and Tamir, T. (1971). *J. Opt. Soc. Am.* **61**, 586.
Horowitz, B. R., and Tamir, T. (1973). *Appl. Phys.* **1**, 31.
Hugonin, J. P., and Petit, R. (1977). *J. Opt. (Paris)* **8**, 73.
Imbert, C. (1968). *C. R. Acad. Sci.* B **267**, 1401.
Imbert, C. (1969). *C. R. Acad. Sci.* B **269**, 1227.
Imbert, C. (1970). *C. R. Acad. Sci.* B **270**, 529.
Imbert, C. (1972). *C. R. Acad. Sci.* B **274**, 1213.
Julia, B., and Neveu, A. (1973). *J. Phys. (Paris)* **34**, 335.
Levy, Y., and Imbert, C. (1972). *C. R. Acad. Sci.* B **275**, 723.
Levy, Y., and Imbert, C. (1975). *Opt. Commun.* **13**, 43.
Lotsch, H. K. V. (1970/71). *Optik (Stuttgart)* **32**, 116, 189, 299, 553.
McGuirk, M., and Carniglia, C. K. (1977). *J. Opt. Soc. Am.* **67**, 103.
Mazet, A., Imbert, C., and Huard, S. (1971). *C. R. Acad. Sci.* B **273**, 592.
Pavageau, J. (1969). *C. R. Acad. Sci.* B **268**, 737.
Picht, J. (1929). *Ann. Phys. (Leipzig)* **3**, 433.
Read, L. A. A., Wong, M., and Reesor, G. E. (1978). *J. Opt. Soc. Am.* **68**, 319.
Renard, R. H. (1964). *J. Opt. Soc. Am.* **54**, 1190.
Rhodes, D. J., and Carniglia, C. K. (1977). *J. Opt. Soc. Am.* **67**, 679.
Ricard, J. (1970). *Nouv. Rev. Opt. Appl.* **1**, 273.
Ricard, J. (1973). *Nouv. Rev. Opt.* **4**, 63.
Ricard, J. (1974). *Nouv. Rev. Opt.* **5**, 7.
Ricard, J. (1976). *Nouv. Rev. Opt.* **7**, 1.
Schilling, H. (1965). *Ann. Phys. (Leipzig)* **16**, 122.
White, I. A., Snyder, A. W., and Pask, C. (1977). *J. Opt. Soc. Am.* **67**, 703.

Manuscript received 16 August 1979, accepted 25 October 1979