Normal-incidence reflection and transmission by uniaxial crystals and crystal plates

1992 J. Phys.: Condens. Matter 4 1387

(http://iopscience.iop.org/0953-8984/4/5/019)

J. Phys.: Condens. Matter 4 (1992) 1387–1398. Printed in the UK

Normal-incidence reflection and transmission by uniaxial crystals and crystal plates

John Lekner
Department of Physics, Victoria University of Wellington, Wellington, New Zealand

Received 9 July 1991, in final form 21 October 1991

Abstract. A 2×2 matrix method is developed and used to calculate the reflection and transmission amplitudes for normal-incidence reflection and transmission by uniaxial crystals and crystal slabs. The reflection is characterized by two reflection amplitudes (r and r') which give the amplitudes of the reflected field components along and perpendicular to the incident polarization direction. The reflection results take the form $r = r_o \cos^2\varphi + r_e \sin^2\varphi$, $r' = (r_o - r_e)\cos\varphi \sin\varphi$, where φ is the angle between the incident polarization and the ordinary electric field vector in the crystal, and r_o and r_e are the reflection amplitudes of isotropic crystals or crystal plates, with appropriate refractive indices. Similar results are obtained for the transmission amplitudes.

1. Introduction

In a recent paper (Lekner 1991) we have given formulae for the reflection and transmission amplitudes of a uniaxial crystal, at an arbitrary angle of incidence. Here we consider the special but important case of normal incidence, and extend the analysis to include the reflection and transmission by a crystal plate. We will show that a 2×2 matrix method can be used at normal incidence to give analytical results for all the required quantities, in contrast to the more general 4×4 matrix formalisms (Teitler and Henvis 1970, Berreman 1971, Yeh 1979, 1988).

We begin by rederiving the results of section 5.4 of Lekner (1991), in order to define the notation and to introduce the 2×2 matrix method.

2. Normal-incidence reflection from a crystal face

The coordinate system to be used is as follows: the z axis is normal to the reflecting crystal surface, which lies in an xy plane, for simplicity taken to be the $z = 0$ plane. The optic axis, $\boldsymbol{c} = (\alpha, \beta, \gamma)$, has direction cosines α, β and γ with respect to the x, y and z axes. Let $\varepsilon_o = n_o^2$ and $\varepsilon_e = n_e^2$ be the ordinary and extraordinary dielectric constants of the crystal, with n_o and n_e the corresponding refractive indices. We set $\Delta\varepsilon = \varepsilon_e - \varepsilon_o$ and define

$$n_\gamma^2 = \varepsilon_\gamma = \varepsilon_o + \gamma^2\,\Delta\varepsilon. \tag{1}$$

For normal incidence the ordinary and extraordinary wavevectors both lie along the inward normal to the reflecting surface, and have magnitudes

$$k_o = n_o\,\omega/c \qquad k_e = (n_o n_e/n_\gamma)\omega/c. \tag{2}$$

The electric fields which can propagate as plane waves in the crystal are orthogonal:

$$\boldsymbol{E}^o = N_o(-\beta, \alpha, 0) \qquad \boldsymbol{E}^e = N_e(\alpha, \beta, -\gamma(1-\gamma^2)\,\Delta\varepsilon/\varepsilon_\gamma). \tag{3}$$

The factors N_o and N_e, which normalize the vectors $\boldsymbol{E}^o$ and $\boldsymbol{E}^e$ to unit magnitude, are given by

$$N_o^{-2} = 1 - \gamma^2 \qquad N_e^{-2} = N_o^{-2}[1 + \gamma^2(1-\gamma^2)(\Delta\varepsilon/\varepsilon_\gamma)^2]. \tag{4}$$

So far we have not defined the orientation of the x and y axes. For oblique incidence these are defined to lie respectively long and perpendicular to the plane of incidence, but at normal incidence the plane of incidence is not uniquely defined, and neither are the s and p polarizations. Let $\boldsymbol{n} = (0, 0, 1)$ be the inward unit normal; since the ordinary and extraordinary wavevectors lie along $\boldsymbol{n}$, the plane of $\boldsymbol{n}$ and $\boldsymbol{c}$ is the principal plane, and its normal $\boldsymbol{n} \times \boldsymbol{c} = (-\beta, \alpha, 0)$ lies in the reflecting plane $z = 0$ (and is parallel to $\boldsymbol{E}^o$). This is one physical direction in the xy plane. Another is the direction of $\boldsymbol{E}_1$ in the (linearly polarized) incident wave, and we take the electric vector $\boldsymbol{E}_1$ to define the x axis. If φ is the angle between the incident polarization and the $\boldsymbol{E}^o$ or $\boldsymbol{n} \times \boldsymbol{c}$ direction,

$$\boldsymbol{E}^o = (\cos\varphi, \sin\varphi, 0) \qquad \boldsymbol{E}^e = (\sin\varphi\cos\delta, -\cos\varphi\cos\delta, \sin\delta) \tag{5}$$

where δ is the angle between the ray and wavevector directions for the extraordinary wave, given by

$$\begin{aligned} \sin\delta &= \boldsymbol{E}^e \cdot \boldsymbol{n} = -\gamma(1-\gamma^2)^{1/2}\,\Delta\varepsilon/[\varepsilon_o^2 + (\varepsilon_e^2 - \varepsilon_o^2)\gamma^2]^{1/2} \\ \tan\delta &= -\gamma(1-\gamma^2)^{1/2}\,\Delta\varepsilon/\varepsilon_\gamma. \end{aligned} \tag{6}$$

We can now write down the incident, reflected and transmitted waves for the case of light normally incident onto an arbitrary crystal face. With the x axis along the direction of the incident plane-polarized electric field, these are

$$\begin{aligned} &\text{incident:} && (e^{ik_1 z}, 0, 0) \\ &\text{reflected:} && (re^{-ik_1 z}, r'e^{-ik_1 z}, 0) \\ &\text{transmitted:} && t_o\boldsymbol{E}^o e^{ik_o z} + t_e\boldsymbol{E}^e e^{ik_e z} \end{aligned} \tag{7}$$

where $k_1 = n_1\omega/c$, n_1 being the refractive index of the medium of incidence. Note that two reflection amplitudes (r and r') and two transmission amplitudes (t_o and t_e) are sufficient to characterize the problem, in contrast to the four reflection and four transmission amplitudes required at general incidence. The continuity of E_x, E_y, $\partial E_x/\partial z$ and $\partial E_y/\partial z$ at the crystal surface $z = 0$ gives the four equations

$$\begin{aligned} 1 + r &= t_o E_x^o + t_e E_x^e \qquad & k_1(1-r) &= k_o t_o E_x^o + k_e t_e E_x^e \\ r' &= t_o E_y^o + t_e E_y^e \qquad & -k_1 r' &= k_o t_o E_y^o + k_e t_e E_y^e. \end{aligned} \tag{8}$$

We define the column vectors

$$\boldsymbol{r} = \begin{pmatrix} r \\ r' \end{pmatrix} \qquad \boldsymbol{t} = \begin{pmatrix} t_o \\ t_e \end{pmatrix} \qquad \boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{9}$$

and the matrices

$$\mathbf{M} = \begin{pmatrix} E_x^o & E_x^e \\ E_y^o & E_y^e \end{pmatrix} \qquad \mathbf{K}_1^{-1} = \begin{pmatrix} k_o/k_1 & 0 \\ 0 & k_e/k_1 \end{pmatrix}. \tag{10}$$

We will call $\mathbf{M}$ the *mode matrix*, since its elements are the transverse components of the modes (or eigenstates) $\boldsymbol{E}^o$ and $\boldsymbol{E}^e$. The four equations (8) can now be written as a pair of coupled vector equations in $\boldsymbol{r}$ and $\boldsymbol{t}$:

$$\boldsymbol{u} + \boldsymbol{r} = \mathbf{M}\boldsymbol{t} \qquad \boldsymbol{u} - \boldsymbol{r} = \mathbf{M}\mathbf{K}_1^{-1}\boldsymbol{t}. \tag{11}$$

These are readily solved, to give

$$\boldsymbol{r} = \mathbf{R}\boldsymbol{u} \qquad \boldsymbol{t} = \mathbf{M}^{-1}(\mathbf{I} + \mathbf{R})\boldsymbol{u} \tag{12}$$

where $\mathbf{I}$ is the identity or unit 2×2 matrix, and

$$\mathbf{A} = \mathbf{M}\mathbf{K}_1\mathbf{M}^{-1} \qquad \mathbf{R} = (\mathbf{A} + \mathbf{I})^{-1}(\mathbf{A} - \mathbf{I}). \tag{13}$$

From (5), (10) and the well-known inverse of a 2×2 matrix,

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}^{-1} = (m_{11}m_{22} - m_{12}m_{21})^{-1} \begin{pmatrix} m_{22} & -m_{12} \\ -m_{21} & m_{11} \end{pmatrix} \tag{14}$$

we have that

$$\mathbf{M} = \begin{pmatrix} \cos\varphi & \sin\varphi\cos\delta \\ \sin\varphi & -\cos\varphi\cos\delta \end{pmatrix} \qquad \mathbf{M}^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi/\cos\delta & -\cos\varphi/\cos\delta \end{pmatrix}. \tag{15}$$

Let $\mathbf{D} = \mathrm{diag}(d_o, d_e)$ be any diagonal 2×2 matrix. Then

$$\mathbf{M}\mathbf{D}\mathbf{M}^{-1} = \begin{pmatrix} d_o\cos^2\varphi + d_e\sin^2\varphi & (d_o - d_e)\cos\varphi\sin\varphi \\ (d_o - d_e)\cos\varphi\sin\varphi & d_o\sin^2\varphi + d_e\cos^2\varphi \end{pmatrix} \tag{16}$$

is a symmetric matrix, and does not contain $\cos\delta$. The matrix $\mathbf{A}$ has this form, and so does $\mathbf{R}$ because $\mathbf{R} = \mathbf{M}\,(\mathbf{K}_1 + \mathbf{I})^{-1}(\mathbf{K}_1 - \mathbf{I})\mathbf{M}^{-1}$:

$$\mathbf{R} = \begin{pmatrix} r_o\cos^2\varphi + r_e\sin^2\varphi & (r_o - r_e)\cos\varphi\sin\varphi \\ (r_o - r_e)\cos\varphi\sin\varphi & r_o\sin^2\varphi + r_e\cos^2\varphi \end{pmatrix} \tag{17}$$

where r_o and r_e are the same as the Fresnel reflection amplitudes for isotropic media with refractive indices n_o and $n_o n_e/n_\gamma$:

$$r_o = (k_1 - k_o)/(k_1 + k_o) \qquad r_e = (k_1 - k_e)/(k_1 + k_e). \tag{18}$$

Thus, from (12),

$$r = r_o\cos^2\varphi + e_e\sin^2\varphi \qquad r' = (r_o - r_e)\cos\varphi\sin\varphi. \tag{19}$$

When the crystal is transparent, r_o and r_e are real, and thus when the incident light is linearly polarized the reflected light is also linearly polarized, with the polarization rotated by an angle $\mathrm{atn}(r'/r)$ relative to the incident polarization. Note that r', which gives the amplitude of the electric field reflected into the y polarization when the incident field $\boldsymbol{E}_1$ is x-polarized, is zero when $\boldsymbol{E}_1$ is either parallel or perpendicular to $\boldsymbol{E}^{\mathrm{o}}$ (or $\boldsymbol{n} \times \boldsymbol{c}$). The magnitude of r' is largest as a function of the azimuthal angle at odd multiples of 45°, and largest as a function of the direction cosine γ at $\gamma = 0$, that is when the optic axis lies in the reflecting plane. The absolute maximum of $|r'|$ is $n_1|n_e - n_o|/(n_1 + n_o)(n_1 + n_e)$. (For calcite in air this takes the value 0.026, about one-tenth of $|r_o|$.) At the other extreme, when $\gamma^2 = 1$ and the reflection is from the basal plane, $r' = 0$ and $r = r_o$, independent of the azimuthal angle.

The formulae (19) are in agreement with the reflection amplitudes given in equations (71) and (73) of Lekner (1991), when we set $\sin\varphi = \alpha/(\alpha^2 + \beta^2)^{1/2}$, $\cos\varphi = -\beta/(\alpha^2 + \beta^2)^{1/2}$. Note the degeneracy in the latter formulae, associated with the lack of uniqueness of s and p polarizations at normal incidence.

The amplitudes of the transmitted waves are obtained from (12), (15) and (17). We find

$$
\begin{aligned}
t_o &= (1 + r_o)\cos\varphi = [2k_1/(k_1 + k_o)]\cos\varphi \\
t_e &= (1 + t_e)\sin\varphi/\cos\delta = [2k_1/(k_1 + k_e)]\sin\varphi/\cos\delta .
\end{aligned}
\tag{20}
$$

If the incident wave is polarized with $\boldsymbol{E}_1$ along the $\boldsymbol{E}^{\mathrm{o}}$ (or $\boldsymbol{n} \times \boldsymbol{c}$) direction, only the ordinary wave will propagate into the crystal, as expected. Only the extraordinary wave will be excited when $\varphi = \pm 90°$. The transmission amplitude t_e contains the divisor $\cos\delta$, which in practice is close to unity. From (6) we find that, for transparent media, $|\tan\delta|$ is greatest when $\gamma^2 = \varepsilon_o/(\varepsilon_o + \varepsilon_e)$, the largest value being $|\Delta\varepsilon|/2n_o n_e$. The corresponding smallest value of $\cos\delta$ is $2n_o n_e/(n_o^2 + n_e^2) = 1 - (n_o - n_e)^2/(n_o^2 + n_e^2)$.

3. Reflection and transmission by a uniaxial plate

We now consider the optical properties of a crystal plate of thickness Δz, bounded by the medium of incidence (of refractive index n_1) and a substrate of refractive index n_2. When light is incident from medium 1 onto the $z = 0$ face of the crystal, and is linearly polarized in the x direction, the electric fields are as follows:

$$
\begin{array}{ll}
\text{incident:} & (\mathrm{e}^{\mathrm{i}k_1 z}, 0, 0) \\
\text{reflected:} & (r\,\mathrm{e}^{-\mathrm{i}k_1 z}, r'\,\mathrm{e}^{-\mathrm{i}k_1 z}, 0) \\
\text{inside crystal:} & (a_o\,\mathrm{e}^{\mathrm{i}k_o z} + b_o\,\mathrm{e}^{-\mathrm{i}k_o z})\boldsymbol{E}^{\mathrm{o}} + (a_e\,\mathrm{e}^{\mathrm{i}k_e z} + b_e\,\mathrm{e}^{-\mathrm{i}k_e z})\boldsymbol{E}^{\mathrm{e}} \\
\text{transmitted:} & (t\,\mathrm{e}^{\mathrm{i}k_2(z-\Delta z)}, t'\,\mathrm{e}^{\mathrm{i}k_2(z-\Delta z)}, 0).
\end{array}
\tag{21}
$$

(The wavevector magnitude in the second medium is $k_2 = n_2\omega/c$.) Note that we have modified the usual definition of the transmission amplitudes t and t' to remove the common phase factor $\exp(-\mathrm{i}k_2\,\Delta z)$. The continuity of E_x, E_y, $\partial E_x/\partial z$, $\partial E_y/\partial z$ (or equivalently of E_x, E_y, B_y, $-B_x$) at $z = 0$ and $z = \Delta z$ gives eight equations in the eight unknowns, $r, r', a_o, b_o, a_e, b_e, t, t'$. At $z = 0$ we obtain

$$\begin{aligned} 1 + r &= s_o E_x^o + s_e E_x^e \qquad & k_1(1 - r) &= k_o d_o E_x^o + k_e d_e E_x^e \\ r' &= s_o E_y^o + s_e E_y^e \qquad & -k_1 r' &= k_o d_o E_y^o + k_e d_e E_y^e \end{aligned} \tag{22}$$

where $s_o = a_o + b_o$ and $d_o = a_o - b_o$, and similarly for the coefficients of the extraordinary wave. At $z = \Delta z$ we find

$$\begin{aligned} t &= s_o' E_x^o + s_e' E_x^e \qquad & k_2 t &= k_o d_o' E_x^o + k_e d_e' E_x^e \\ t' &= s_o' E_y^o + s_e' E_y^e \qquad & k_2 t' &= k_o d_o' E_y^o + k_e d_e' E_y^e \end{aligned} \tag{23}$$

where $s_o' = a_o' + b_o' = a_o\, e^{ik_o\Delta z} + b_o\, e^{-ik_o\Delta z}$, $d_o' = a_o' - b_o'$, with analogous definitions for s_e' and d_e'.

We now reduce this 8×8 problem to four coupled 2×2 equations, which enables a solution to be found in terms of 2×2 matrices. We define the column vectors

$$\boldsymbol{r} = \begin{pmatrix} r \\ r' \end{pmatrix} \qquad \boldsymbol{t} = \begin{pmatrix} t \\ t' \end{pmatrix} \qquad \boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{s} = \begin{pmatrix} s_o \\ s_e \end{pmatrix} \qquad \boldsymbol{d} = \begin{pmatrix} d_o \\ d_e \end{pmatrix}. \tag{24}$$

Then the four equations (22) can be written as

$$\boldsymbol{u} + \boldsymbol{r} = \mathbf{M}\boldsymbol{s} \qquad \boldsymbol{u} - \boldsymbol{r} = \mathbf{M}\mathbf{K}_1^{-1}\boldsymbol{d} \tag{25}$$

where the mode matrix $\mathbf{M}$ and the matrix $\mathbf{K}_1^{-1}$ are defined in (10). For the second set of equations (23) we note that

$$\begin{aligned} s_o' &= a_o\, e^{ik_o\Delta z} + b_o\, e^{-ik_o\Delta z} = s_o \cos(k_o\,\Delta z) + i d_o \sin(k_o\,\Delta z) \\ d_o' &= a_o\, e^{ik_o\Delta z} - b_o\, e^{-ik_o\Delta z} = d_o \cos(k_o\,\Delta z) + i s_o \sin(k_o\,\Delta z) \end{aligned} \tag{26}$$

and similarly for s_e' and d_e'. Thus we can write

$$\boldsymbol{s}' = \mathbf{C}\boldsymbol{s} + i\mathbf{S}\boldsymbol{d} \qquad \boldsymbol{d}' = \mathbf{C}\boldsymbol{d} + i\mathbf{S}\boldsymbol{s} \tag{27}$$

where $\mathbf{C}$ and $\mathbf{S}$ are the diagonal matrices

$$\mathbf{C} = \begin{pmatrix} \cos(k_o\,\Delta z) & 0 \\ 0 & \cos(k_e\,\Delta z) \end{pmatrix} \qquad \mathbf{S} = \begin{pmatrix} \sin(k_o\,\Delta z) & 0 \\ 0 & \sin(k_e\,\Delta z) \end{pmatrix}. \tag{28}$$

The equations (23) can now be stated as

$$\boldsymbol{t} = \mathbf{M}\boldsymbol{s}' = \mathbf{M}(\mathbf{C}\boldsymbol{s} + i\mathbf{S}\boldsymbol{d}) \qquad \boldsymbol{t} = \mathbf{M}\mathbf{K}_2^{-1}\boldsymbol{d}' = \mathbf{M}\mathbf{K}_2^{-1}(\mathbf{C}\boldsymbol{d} + i\mathbf{S}\boldsymbol{s}) \tag{29}$$

where $\mathbf{K}_2^{-1}$ is defined as in (10), with k_2 replacing k_1. The equations (25) and (29) can now be solved as follows: rewrite (25) as $\boldsymbol{s} = \mathbf{M}^{-1}(\boldsymbol{u} + \boldsymbol{r})$ and $\boldsymbol{d} = \mathbf{K}_1\mathbf{M}^{-1}(\boldsymbol{u} - \boldsymbol{r})$ and substitute into (29), equating the two expressions for $\boldsymbol{t}$. The resulting linear equation $\boldsymbol{r}$

in terms of $\boldsymbol{u}$ is then solved, using the fact that $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{C}$ and $\mathbf{S}$ are all diagonal, and thus commute with each other. We thus find

$$\boldsymbol{r} = \mathbf{M}\begin{pmatrix} r_o & 0 \\ 0 & r_e \end{pmatrix}\mathbf{M}^{-1}\boldsymbol{u} = \mathbf{R}\boldsymbol{u} \tag{30}$$

where $\mathbf{R}$ has the form given in (17), with r_o now given by

$$r_o = \frac{k_o(k_1 - k_2)\cos(k_o\Delta z) + \mathrm{i}(k_o^2 - k_1k_2)\sin(k_o\Delta z)}{k_o(k_1 + k_2)\cos(k_o\Delta z) - \mathrm{i}(k_o^2 + k_1k_2)\sin(k_o\Delta z)}. \tag{31}$$

This is the reflection amplitude for a homogeneous isotropic layer of refractive index n_o and thickness Δz (Lekner (1987), equation (2.52)). The amplitude r_e, obtained by replacing k_o by k_e in (31), is likewise the reflection amplitude for an isotropic slab of refractive index n_on_e/n_γ. An alternative form of (31) is

$$r_o = (r_1^o + r_2^o\,\mathrm{e}^{2ik_o\Delta z})/(1 + r_1^or_2^o\,\mathrm{e}^{2ik_o\Delta z}) \tag{32}$$

where

$$r_1^o = (k_1 - k_o)/(k_1 + k_o) \qquad r_2^o = (k_o - k_2)/(k_o + k_2) \tag{33}$$

are the Fresnel amplitudes for reflection at the entry and exit faces of an isotropic slab of refractive index n_o. The same form holds for r_e, with k_e replacing k_o.

Returning to (30), we thus have the direct and orthogonal reflection amplitudes

$$r = r_o\cos^2\varphi + r_e\sin^2\varphi \qquad r' = (r_o - r_e)\cos\varphi\sin\varphi. \tag{34}$$

Note that the direct reflection amplitude r can be zero if r_o and r_e differ in phase by an odd multiple of π and their magnitudes satisfy $|r_o|\cos^2\varphi = |r_e|\sin^2\varphi$.

The transmission amplitudes are found from (29), on substituting for s and d in terms of $\boldsymbol{r}$. The result is

$$\boldsymbol{t} = \mathbf{M}\begin{pmatrix} t_o & 0 \\ 0 & t_e \end{pmatrix}\mathbf{M}^{-1}\boldsymbol{u} = \begin{pmatrix} t_o\cos^2\varphi + t_e\sin^2\varphi & (t_o - t_e)\cos\varphi\sin\varphi \\ (t_o - t_e)\cos\varphi\sin\varphi & t_o\sin^2\varphi + t_e\cos^2\varphi \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{35}$$

where t_o and t_e are the transmission amplitudes of isotropic slabs with indices n_o and n_on_e/n_γ, each with thickness Δz (Lekner (1987), equations (2.53) and (2.59)):

$$t_o = \frac{2k_1k_o}{k_o(k_1 + k_2)\cos(k_o\,\Delta z) - \mathrm{i}\,(k_o^2 + k_1k_2)\sin(k_o\,\Delta z)} = \frac{(1 + r_1^o)(1 + r_2^o)\,\mathrm{e}^{ik_o\Delta z}}{1 + r_1^or_2^o\,\mathrm{e}^{2ik_o\Delta z}}. \tag{36}$$

(The formulae for t_e have k_e replacing k_o.) Thus

$$t = t_o\cos^2\varphi + t_e\sin^2\varphi \qquad t' = (t_o - t_e)\cos\varphi\sin\varphi. \tag{37}$$

We discuss some consequences of the formulae (34) and (37) for reflection and transmission by a homogeneous uniaxial layer in the next two sections. Here we note that t_o and t_e are the exact elements of a transmission matrix which relates the outgoing electric field components along the $\boldsymbol{E}^o$ and $\boldsymbol{E}^e$ directions to those incident on the crystal layer. For layer thickness Δz the approximate matrix used in the Jones calculus (Jones (1941);

see also the reprints of all eight of Jones' papers in Swindell (1975), or the discussion in Yariv and Yeh (1984), ch 5), and the exact matrix, are respectively

$$\begin{pmatrix} \exp(\mathrm{i}k_o\,\Delta z) & 0 \\ 0 & \exp(\mathrm{i}k_e\,\Delta z) \end{pmatrix} \qquad \begin{pmatrix} t_o & 0 \\ 0 & t_e \end{pmatrix}.$$

From (36) we see that the leading term in the difference between the exact elements, which allow for (multiple) reflections at the boundaries of the anisotropic slab, and the approximate elements is $r_1 + r_2$ (superscripts o and e to be added for the o and e waves). When $n_1 = n_2$ this is zero for both polarizations, so the difference becomes second-order in the Fresnel reflection amplitudes and can usually be neglected.

4. Discussion of reflection properties of a uniaxial layer

The reflection amplitudes r and r' are now complex, even for transparent crystals. The azimuthal dependence is as for reflection from a bulk crystal: r takes the values r_o and r_e when the incident polarization is along the $\boldsymbol{E}^o$ and $\boldsymbol{E}^e$ directions, and r' is zero in either case. The magnitude of r' is largest as a function of the azimuthal angle φ at odd multiples of 45°, and then $|r'| = |r_o - r_e|/2$. The variation of r_o, r_e and $r_o - r_e$ with the thickness Δz of the anisotropic layer may be found as follows: both r_o and r_e are of the form

$$(r_1 + r_2 Z)/(1 + r_1 r_2 Z) \qquad Z = \exp(2\mathrm{i}k\,\Delta z) \tag{38}$$

where k takes the values k_o or k_e in r_1, r_2 and Z. As the thickness increases, Z moves on the unit circle in the complex plane (we assume the medium is non-absorbing). The periods in Δz of r_o and r_e are π/k_o and π/k_e. The bilinear (or fractional) transformation, which gives the complex numbers r_o and r_e in terms of Z, transforms circles into circles in the complex plane, and so the loci of r_o and r_e are also circles. Since r_1 is real, these circles are symmetric with respect to reflection in the real axis, and their centres lie on the real axis. We shall first assume the substrate is non-absorbing, so r_2 is real also. The radii and centres may then be found from the intersections with real axis at $Z = \pm 1$. At $Z = 1$, r_o and r_e take the zero-thickness value

$$r^+ = (r_1 + r_2)/(1 + r_1 r_2) = (k_1 - k_2)/(k_1 + k_2). \tag{39}$$

This is the reflection amplitude of a sudden transition between media 1 and 2. At $Z = -1$ the reflection amplitudes are

$$r^- = (r_1 - r_2)/(1 - r_1 r_2) = (k_1 k_2 - k^2)/(k_1 k_2 + k^2). \tag{40}$$

The centre c and radius a are therefore given by

$$\begin{aligned} c &= (r^+ + r^-)/2 = k_2(k_1^2 - k^2)/(k_1 + k_2)(k_1 k_2 + k^2) \\ a &= (r^+ - r^-)/2 = k_1(k^2 - k_2^2)/(k_1 + k_2)(k_1 k_2 + k^2). \end{aligned} \tag{41}$$

The two circles representing r_o and r_e touch at the point r^+ on the real axis. The points r^- lie to the left of r^+ if $k > k_2$, that is if $n_o > n_2$ for r_o and $n_o n_e/n_\gamma > n_2$ for r_e. When both r_o^- and r_e^- are more negative than r^+ the maximum value of $|r_o - r_e|$ as Δz varies is the diameter of the larger of the two circles for r_o and r_e (see figure 1), i.e.

$$|r_o - r_e|_{\max} = 2k_1(k_g^2 - k_2^2)/(k_1 + k_2)(k_1 k_2 + k_g^2) \tag{42}$$

where k_g is the greater of k_o and k_e. The same result holds when r_o^- and r_e^- both lie to the right of r^+. When one of r_o^- and r_e^- is to the left and the other to the right of r^+, the maximum value of $|r_o - r_e|$ as Δz varies is the sum of the diameters of the two circles, i.e.

$$|r_o - r_e|_{\max} = \frac{2k_1}{k_1 + k_2}\left(\frac{k_o^2 - k_2^2}{k_1k_2 + k_o^2} + \frac{k_e^2 - k_2^2}{k_1k_2 + k_e^2}\right). \tag{43}$$

For the maximum of $|r_o - r_e|$ given in (42) to be attained, the crystal plate has to be thicker than

$$(\pi/2)/|k_o - k_e| = (\lambda/4)/|n_o - n_o n_e/n_\gamma| \tag{44}$$

where λ is the vacuum wavelength. The corresponding minimum thickness requirement for (43) is twice the above value.

For thin crystal films ($\Delta z \ll \lambda$) both r_o and r_e tend to r^+, the correction to lowest order in $\omega\,\Delta z/c$ taking the form

$$2\mathrm{i}k_1\,\Delta z(k^2 - k_2^2)/(k_1 + k_2)^2$$

where $k = k_o$ or k_e. Thus to first order in $\omega\,\Delta z/c$ the difference between r_o and r_e is

$$r_o - r_e = \frac{2\mathrm{i}k_1\,\Delta z}{(k_1 + k_2)^2}(k_o^2 - k_e^2) = \frac{-2\mathrm{i}k_1\,\Delta z}{(k_1 + k_2)^2}\left(\frac{\omega}{c}\right)^2(1 - \gamma^2)\frac{\varepsilon_o\,\Delta\varepsilon}{\varepsilon_\gamma}. \tag{45}$$

The reflection amplitude r' is pure imaginary when the substrate is non-absorbing. Its maximum magnitude (at $\varphi = \pm 45°$ and $\gamma = 0$) for thin films is then

$$|r'|_{\max} = \tfrac{1}{2}|r_o - r_e|_{\max} = \frac{2\pi n_1|\Delta\varepsilon|}{(n_1 + n_2)^2}\frac{\Delta z}{\lambda}. \tag{46}$$

The direct reflection amplitude for thin crystal films is

$$r = r_o\cos^2\varphi + r_e\sin^2\varphi = r_+ + [2\mathrm{i}k_1\,\Delta z/(k_1 + k_2)^2](k_o^2\cos^2\varphi + k_e^2\sin^2\varphi - k_2^2) \tag{47}$$

to lower order in the film thickness.

We now consider the case of an absorbing substrate. The Fresnel amplitude $r_2 = (k - k_2)/(k + k_2)$ is now complex for both the o and e waves. The radii and centres of the circles on which r_o move are now given by

$$a = |r_2|(1 - R_1)/(1 - R_1R_2) \qquad c = r_1(1 - R_2)/(1 - R_1R_2) \tag{48}$$

where $R_1 = r_1^2$ and $R_2 = |r_2|^2$. There still exists a point common to the r_o and r_e circles, namely the $Z = 1$ zero-thickness value of r_o and r_e as given in (39), but this is now complex. Figure 2 shows the locii of r_o and r_e for calcite slabs on substrates of Al and Si. We see that in the case of aluminium the reflection from the substrate is so strong as to make the anisotropy unimportant, but that for the silicon substrate the effect of anisotropy is significant.

5. Discussion of transmission properties of uniaxial layers

The formulae (37) give t and t', the transmission amplitudes of light polarized parallel and perpendicular to the incident polarization, in terms of the transmission amplitudes t_o and t_e of isotropic layers with refractive indices n_o and $n_o n_e/n_\gamma$. For incident polarization

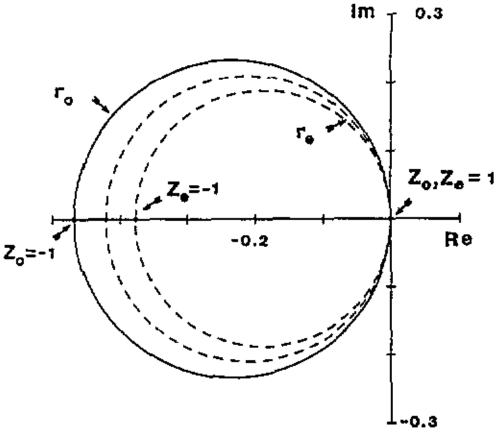

Figure 1. The loci of r_o and r_e, drawn for calcite (n_o = 1.658, n_e = 1.486) in air. As the thickness Δz of the crystal plate increases, r_o and r_e move on circles in the complex plane. The r_o-circle is fixed, while the r_e-circles (dashed) depend on the inclination of the optic axis to the surface. The direction cosine γ of the optic axis relative to the surface normal is zero on the inner r_e-circle, and $\pm 1/\sqrt{2}$ on the outer r_e-circle.

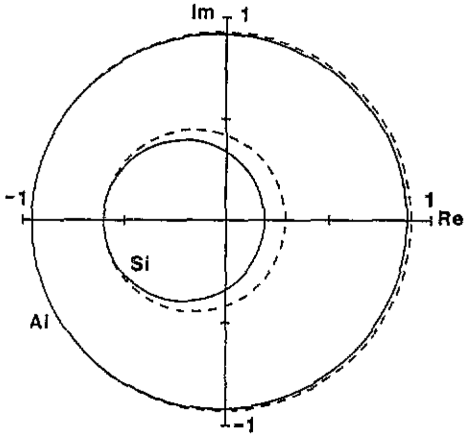

Figure 2. Loci of r_o and r_e in the complex plane, for a calcite slab of variable thickness on a substrate of Al or Si. The refractive indices at 633 nm are: Al, 1.566 + 7.938 i; Si, 4.0 + 0.12 i. The solid circles are r_o, the dashed circles r_e with $\gamma = 0$.

either parallel or perpendicular to $\boldsymbol{n} \times \boldsymbol{c}$ (the direction of E^o), the exit polarization is the same as that on entry. These orientations thus give zero transmission between crossed polarizers. Let us consider the general case of the intensity transmitted through a crystal between polarizer and analyser, with angle χ between their easy axes, as shown in figure 3. The electric field transmitted through the analyser is, for an incident field of unit amplitude,

$$e^{i(k_2 z - \omega t)}(t \cos\chi + t' \sin\chi). \tag{49}$$

Thus the transmitted intensity is given by

$$|t\cos\chi + t'\sin\chi|^2 = |t|^2\cos^2\chi + |t'|^2\sin^2\chi + 2|tt'|\cos\chi\sin\chi\cos(\mathrm{ph}(t'/t)) \tag{50}$$

where we use the notation $\xi = |\xi| \exp(\mathrm{i}\,\mathrm{ph}(\xi))$. When the polarizer and analyser are crossed ($\chi = 90°$), the intensity is

$$|t'|^2 = |t_o - t_e|^2 \cos^2\varphi \sin^2\varphi \tag{51}$$

and is zero when φ is a multiple of 90° (including zero).

The factor $|t_o - t_e|$ depends on γ^2, the square of the cosine of the angle between the plate normal and the optic axis, since t_e is a function of $n_\gamma = (\varepsilon_0 + \gamma^2\,\Delta\varepsilon)^{1/2}$ through $k_e = k_o n_e/n_\gamma$. In addition both t_o and t_e are functions of the thickness Δz of the crystal plate. The main effect on t_o and t_e of the variation of Δz is through their phases, as we shall see shortly. If this were the only effect, then

$$|t_o - t_e|^2 = |t_o|^2 + |t_e|^2 - 2|t_o t_e|\cos(\mathrm{ph}(t_e/t_o)) \tag{52}$$

would have maxima and minima $(|t_o| \pm |t_e|)^2$ when the phase difference $\mathrm{ph}(t_e/t_o)$ passed through odd and even multiples of π, respectively. The phase difference between t_e and

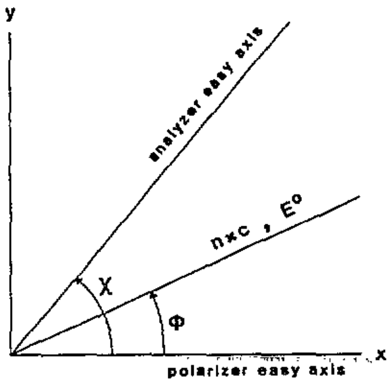

Figure 3. Polarizer, analyser and crystal orientation, in the general case.

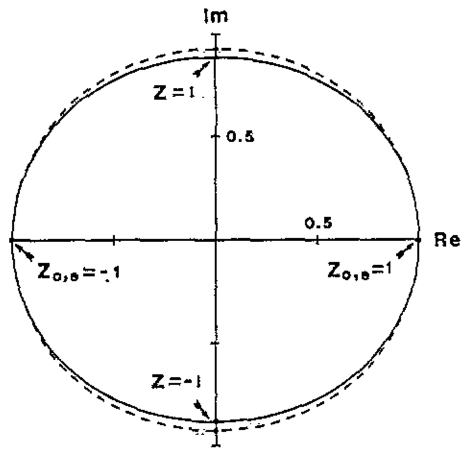

Figure 4. Quartic loci of t_o (solid curve) and t_e (dashed curve). The t_o-quartic is fixed, while the t_e-quartic depends on the direction cosine γ of the optic axis relative to the surface normal. The curves are drawn for calcite in air, as in figure 1, with $\gamma = 0$ (optic axis in the reflecting plane). The values of $Z = \exp(ik\,\Delta z)$ are shown at special points.

t_o is often given as $(k_e - k_o)\,\Delta z$ (see for example Born and Wolf (1965), equation (14.4) (15)), but from (36) we see that it is actually

$$\mathrm{ph}(t_e/t_o) = \mathrm{ph}(t_e) - \mathrm{ph}(t_o) = \mathrm{atn}\left(\frac{k_e^2 + k_1 k_2}{k_e(k_1+k_2)}\tan(k_e\,\Delta z)\right) - \mathrm{atn}\left(\frac{k_o^2 + k_1 k_2}{k_o(k_1+k_2)}\tan(k_o\,\Delta z)\right). \tag{53}$$

The difference between (53) and the approximate phase shift $(k_e - k_o)\,\Delta z$ is due to multiple reflection of light within the crystal. This can be seen from the second form of t_o and t_e shown in (36): if either or both of r_1 or r_2 are reduced to zero by antireflection coating, the phases of t_o and t_e become exactly $k_o\,\Delta z$ and $k_e\,\Delta z$. In the absence of antireflection coatings, the phase increments of the ordinary and extraordinary waves on passing through the crystal plate are different from these values. The phase difference is the optically important quantity, and is given by (53).

We now give an exact treatment of t_o and t_e, and of their difference. The second equality in (36) shows that both t_o and t_e can be written in the form

$$(1+r_1)(1+r_2)Z/(1+r_1 r_2 Z^2) \qquad Z = \exp(ik\,\Delta z) \tag{54}$$

where k takes the values k_o or k_e in r_1, r_2 and Z. We assume the crystal is non-absorbing. Then as Δz increases, Z moves on the unit circle in the complex plane. The transmission amplitudes t_o and t_e also move periodically around loci in the complex plane, with periods $2\pi/k_o$ and $2\pi/k_e$ in Δz. At $Z = \pm 1$ both t_o and t_e take the value $\pm t^+$, where

$$t^+ = 1 + r^+ = 2k_1/(k_1+k_2). \tag{55}$$

At $Z = \pm\mathrm{i}$ the values taken by t_o and t_e are $\pm\mathrm{i}t_o^{\mathrm{i}}$ and $\pm\mathrm{i}t_e^{\mathrm{i}}$, where

$$t_o^{\mathrm{i}} = 2k_1 k_o/(k_1 k_2 + k_o^2) \qquad t_e^{\mathrm{i}} = 2k_1 k_e/(k_1 k_2 + k_e^2). \tag{56}$$

To find the equations of the loci we proceed as in Dorf and Lekner (1987), equations

(30) to (34): the relations between Z and t_o or t_e can be written as quadratics in Z. Elimination of Z by means of $ZZ^* = 1$ gives the following relation between the real and imaginary parts of t_o or t_e, each written as $X + \mathrm{i}Y$:

$$(X^2 + Y^2)^2 = (t^+ X)^2 + (t^\mathrm{i} Y)^2. \tag{57}$$

(We have assumed that the substrate is non-absorbing, so k_2 is real.) In (57) t^+ is given by (55)—the same for t_o and t_e—and the t^i-values are given by (56). Thus t_o and t_e move on quartics in the complex plane, which share the points $(\pm t^+, 0)$ and are symmetric with respect to reflection in both X and Y axes. It is interesting that t_o^{-1} and t_e^{-1} move on ellipses, with semiaxes $(t^+)^{-1}$, and $(t_o^\mathrm{i})^{-1}$, $(t_e^\mathrm{i})^{-1}$. This fact is useful in plotting the loci, shown in figure 4.

We return now to the minima and maxima of $|t_o - t_e|$. First we note that when the optic axis lies along the surface normal, $k_e = k_o$ and $t_e = t_o$ for all thicknesses of the crystal slab. There is then no azimuthal dependence in reflection or transmission, and the crystal behaves as an optically isotropic medium (for normal incidence). For a general orientation of the optic axis, it is still possible for t_e and t_o to be equal, as we saw in figure 4. This occurs if Z_o and Z_e are both $+1$, or both -1. When $k_e \neq k_o$ the minimum non-zero thickness for this to be possible is

$$\Delta z = 2\pi/|k_o - k_e| = \lambda/|n_o - n_o n_e/n_\gamma|. \tag{58}$$

The maximum value of $|t_o - t_e|$ is either $2t^+$ or $t_o^\mathrm{i} + t_e^\mathrm{i}$. The first possibility occurs when $Z_o = +1$ and $Z_e = -1$, or vice versa; the second when $Z_o = +\mathrm{i}$ and $Z_e = -\mathrm{i}$, or vice versa. Since

$$t^+ > t_o^\mathrm{i} \qquad \text{when } (k_1 - k_o)(k_2 - k_o) > 0 \tag{59}$$

with a similar relation for t_e^i, the absolute maximum of $|t_o - t_e|$ is $2t^+ = 4k_1/(k_1 + k_2)$ if k_o and k_e are both greater than either of k_1 and k_2, which is the usual case. If the substrate and the medium of incidence have the same refractive indices, the maximum value attainable by $|t_o - t_e|$ is 2. The minimum thickness for which either of the maxima $2t^+$ or $t_o^\mathrm{i} + t_e^\mathrm{i}$ can be possible is half of that given in (58).

We note in conclusion that the characterization of the transmitted light in terms of the transmission amplitudes t and t' applies to wide beams, which can accurately be represented by the plane wave given in the last line of equation (21). For narrow beams passing through thick crystals there will be complete separation of the o and e rays within the crystal, and two parallel beams will exit the crystal, perpendicularly polarized in the $\boldsymbol{E}^o$ and $\boldsymbol{E}^e$ directions. The transmission amplitudes for these beams are $t_o \cos\varphi$ and $t_e \sin\varphi$. The identity

$$|t_o|^2 \cos^2\varphi + |t_e|^2 \sin^2\varphi = |t_o \cos^2\varphi + t_e \sin^2\varphi|^2 + |t_o - t_e|^2 \cos^2\varphi \sin^2\varphi \tag{60}$$

shows that for given incident power, the transmitted total power in the two beams is the same as it would be in a single very broad beam.

Acknowledgment

The author is grateful to the referee for helpful comments, and in particular for suggesting that reflection from an anisotropic layer on an absorbing substrate be considered (last part of section 4).

References

Berreman D W 1971 *J. Opt. Soc. Am.* **62** 502
Born M and Wolf E 1965 *Principles of Optics* 3rd edn (Oxford: Pergamon)
Dorf M C and Lekner J 1987 *J. Opt. Soc. Am.* A **4** 2096
Jones R C 1941 *J. Opt. Soc. Am.* **31** 488, 493, 500
Lekner J 1987 *Theory of Reflection* (Dordrecht: Nijhoff/Kluwer)
—— 1991 *J. Phys.: Condens. Matter* **3** 6121
Swindell E (ed) 1975 *Polarized Light* (New York: Wiley)
Teitler S and Henvis B W 1970 *J. Opt. Soc. Am.* **60** 830
Yariv A and Yeh P 1984 *Optical Waves in Crystals* (New York: Wiley)
Yeh P 1979 *J. Opt. Soc. Am.* **69** 742
—— 1988 *Optical Waves in Layered Media* (New York: Wiley)