Pure Appl. Opt. 3 (1994) 307–321. Printed in the UK

Normal incidence transmission ellipsometry of anisotropic layers

John Lekner

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

Received 21 January 1993, in final form 28 May 1993

Abstract. We consider various ellipsometric arrangements, consisting of polarizer, sample and analyser, with a compensator or polarization modulator inserted before or after the sample. These configurations enable the determination of the real and complex parts of the ratio of two basic transmission amplitudes. The polarizer–sample–analyser configuration can determine the absolute square and the real part of the ratio of the two basic transmission amplitudes. The transmission amplitudes are obtained by methods previously developed by the author for uniaxial crystals. The method is extended to include biaxial crystals. Transmission properties of free and supported anisotropic layers are discussed.

1. Introduction

In a recent paper (Lekner 1992) we have shown that the normal incidence reflection and transmission properties of a uniaxial crystal layer are contained in two reflection and two transmission amplitudes, r_o, r_e and t_o, t_e. These were found by a 2×2 matrix method. This is in contrast to the situation at oblique incidence, where four reflection and four transmission coefficients are needed, and where the usual approach is via a 4×4 matrix formalism. (See, for example, Wöhler *et al* 1988 or Eidner *et al* 1989 for recent work and further references.) Here we give the theory of what is measured in normal incidence transmission ellipsometry, for various experimental arrangements. Normal incidence transmission ellipsometry can be considered to be complementary to the technique of reflectance anisotropy, in which the difference between normal incidence reflectances is measured for light polarized along the two principal axes of the surface (Acher and Drévillon 1992).

Section 2 summarizes the results for transmission by uniaxial layers, and the appendix extends these results to biaxial layers. Sections 3 to 7 analyse five ellipsometer configurations. In section 8 we give some properties of the measured transmission amplitude ratio for a single anisotropic layer (e.g. mica in air), and for an anisotropic layer on an isotropic substrate.

2. Normal-incidence transmission by a uniaxial layer

A plane monochromatic wave normally incident onto a uniaxial crystal splits into two components which travel in the crystal as the plane waves

$$\boldsymbol{E}_o \exp \mathrm{i}(k_o z - \omega t) \qquad \boldsymbol{E}_e \exp \mathrm{i}(k_e z - \omega t). \tag{1}$$

The subscripts o and e stand for *ordinary* and *extraordinary*. $\boldsymbol{E}_o$ and $\boldsymbol{E}_e$ are electric field vectors, and are orthogonal for normal incidence (but not in general: see Lekner 1991, equation (29)). The wavevector magnitudes k_o and k_e are given by

$$k_o = n_o\omega/c \qquad k_e = k_o n_e/(\varepsilon_o \sin^2\psi + \varepsilon_e \cos^2\psi)^{1/2} \tag{2}$$

where n_o and n_e are the ordinary and extraordinary refractive indices, $\varepsilon_o = n_o^2$ and $\varepsilon_e = n_e^2$ are the corresponding dielectric constants, and ψ is the angle between the optic axis $\boldsymbol{c}$ and the normal to the crystal. (The inward normal $\boldsymbol{n} = (0, 0, 1)$ coincides with the z axis.) The ordinary electric field vector $\boldsymbol{E}_o$ is perpendicular to $\boldsymbol{n}$, being along the $\boldsymbol{n} \times \boldsymbol{c}$ direction. The extraordinary electric field vector $\boldsymbol{E}_e$ is perpendicular $\boldsymbol{E}_o$, but is not perpendicular to $\boldsymbol{n}$ unless ψ is zero or 90°. Nevertheless one can designate by o and e two perpendicular directions in the plane normal to $\boldsymbol{n}$, the first being $\boldsymbol{E}_o$, and the second the projection of $\boldsymbol{E}_e$ onto this plane. The latter direction coincides with that of $\boldsymbol{E}_o \times \boldsymbol{n}$.

Now consider a crystal plate of thickness d, between the medium of incidence of refractive index n_1, and the substrate of index n_2. It is shown in Lekner 1992 that the transmission properties are characterized by two transmission amplitudes, t_o and t_e, corresponding to incident polarization along the o and e directions, respectively. These are given by

$$t^o = \frac{t_1^o t_2^o\, e^{ik_o d}}{1 + r_1^o r_2^o\, e^{2ik_o d}} \tag{3}$$

$$t_1^o = 1 + r_1^o = \frac{2k_1}{k_1 + k_o} \qquad t_2^o = 1 + r_2^o = \frac{2k_o}{k_o + k_2} \tag{4}$$

and a similar set of formulae for t_e, with k_e replacing k_o. The wavevector magnitudes k_1 and k_2 are $n_1\omega/c$ and $n_2\omega/c$, respectively. The coefficients r_1^o and t_1^o are the Fresnel reflection and transmission amplitudes for the interface between media of indices n_1 and n_o, and likewise r_2^o and t_2^o are those for the interface between n_o and n_2. Thus (30) is identical to the transmission amplitude for an isotropic slab of index n_o, while the formula for t_e is the same as for an isotropic slab of index

$$\frac{n_o n_e}{[(n_o \sin\psi)^2 + (n_e \cos\psi)^2]^{1/2}}. \tag{5}$$

When the incident wave is linearly polarized with its electric field vector at angle ϕ to the $\boldsymbol{E}_o$ direction, the transmission amplitudes

$$t = t_o \cos^2\phi + t_e \sin^2\phi \qquad t' = (t_o - t_e)\cos\phi \sin\phi \tag{6}$$

give the components of the transmitted field along and perpendicular to the incident field direction.

The above formulae are for uniaxial crystals. In the appendix we show that results of the same form as equations (3), (4) and (6) hold also for biaxial crystals, with k_o and k_e replaced by k_+ and k_-, the positive square roots of a quadratic equation for k^2, equation (A5). The directions o and e are simultaneously to be replaced by the directions e^+ and e^- of the projections of the electric fields $\boldsymbol{E}^+$ and $\boldsymbol{E}^-$ onto the reflecting plane. In the body of this paper we will continue to use the uniaxial o and e notation for convenience, with the understanding that the results apply (with the above substitutions) to the biaxial case also.

To calculate the output of a transmission ellipsometer we need an extension of the results of Lekner (1992) to the case where the substrate is of finite thickness d_2, since the usual experimental situation is that the light originates and ends in the same ambient medium of index n_1 (usually air or vacuum). Consider first the case of light incident with polarization along the o direction. The reflection–transmission problem is then the same as for isotropic media with indices n_1, n_o, n_2, n_1. The reflection and transmission amplitudes r_o and t_o for this system can be found by matrix methods. From Lekner (1987), section 12.2, we find

$$\left.\begin{aligned} r_o &= \frac{k_1^2 m_{12} + m_{21} + ik_1(m_{22} - m_{11})}{k_1^2 m_{12} - m_{21} + ik_1(m_{22} + m_{11})} \\ t_o &= \frac{2ik_1}{k_1^2 m_{12} - m_{21} + ik_1(m_{22} + m_{11})} \end{aligned}\right\} \tag{7}$$

where the matrix elements m_{ij} are given by

$$\left.\begin{aligned} m_{11} &= c_2 c_o - \frac{k_o}{k_2} s_2 s_o \\ m_{12} &= \frac{c_2 s_o}{k_o} + \frac{s_2 c_o}{k_2} \\ m_{21} &= -k_2 s_2 c_o - k_o s_o c_2 \\ m_{22} &= c_2 c_o - \frac{k_2}{k_o} s_2 s_o \end{aligned}\right\} \tag{8}$$

$$c_o = \cos k_o d \qquad s_o = \sin k_o d \qquad c_2 = \cos k_2 d_2 \qquad s_2 = \sin k_2 d_2.$$

It follows from the structure of (7) and (8) that t_o is the same whether the layer of index n_2 comes before or after the anisotropic layer, in accord with the general theorem of section 2-1 of Lekner 1987, equation (2.14).

For incident polarization along the e direction, corresponding results are obtained by substituting e for o in the subscripts of the formulae (7) and (8). It is clear from formulae (3) and (7) that the phase difference between the transmitted o and e waves is not simply $(k_o - k_e)d$, as is sometimes assumed (see for example Born and Wolf (1965), section 14.4.3). As explained in Holmes (1964) and Lekner (1992), the latter expression does not allow for multiple reflections inside the crystal plate. Holmes considered the case where the principal dielectric axes of the unsupported plate are aligned parallel and perpendicular to the plate normal. The phase difference given in equation (53) of Lekner (1992) is for arbitrary orientation of the optic axis of an unsupported uniaxial plate, and the appendix of the present paper shows how this may be generalized to biaxial plates, again of arbitrary orientation. The experimental aspects of ellipsometry with non-ideal compensators have been considered by Archer and Shank (1967) and Yolken *et al* (1967). See also Azzam and Basharta (1987), section 5.2.1.2.

3. Polarizer–sample–analyser

Figure 1 shows this arrangement, with the anisotropic layer on a substrate of index n_2. For this case, and also when the isotropic layer precedes the anisotropic layer, t_o is given by formulae (7) and (8), and t_e by the same formulae with k_o replaced by k_e. If

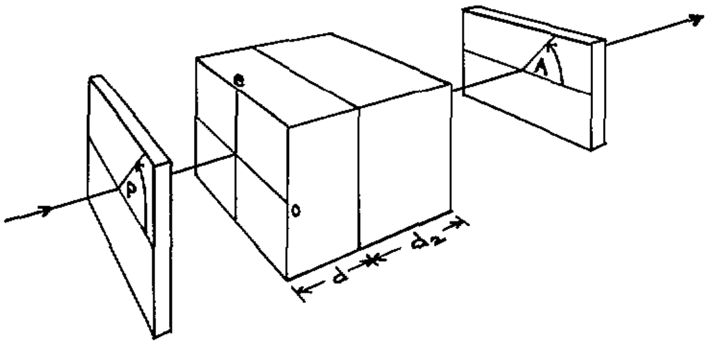

Figure 1. *The polarizer–sample–analyser arrangement. The angles P and A give the directions of the easy axes of the polarizer and analyser, relative to the o direction of the crystal.*

the substrate is absent, t_o is given by equations (3) and (4) with k_2 replaced by k_1, and t_e likewise, with k_o replaced by k_e.

We consider the electric field amplitude, resolved along the o and e directions of the crystal:

	along o	along e
after polarizer:	$\cos P$	$\sin P$
after sample:	$t_o \cos P$	$t_e \sin P$

After the analyser the amplitude is thus

$$t_o \cos P \cos A + t_e \sin P \sin A. \tag{9}$$

For an isotropic material we would have $t_o = t_e$, and hence the amplitude transmitted by the analyser proportional to $\cos(P - A)$, which is zero when the polarizer and analyser easy axes are at right angles. This is the crossed-polar configuration used in polarization microscopes (see for example Gribble and Hall 1985). When $A = P \pm 90°$, formula (9) becomes

$$\pm(t_e - t_o) \cos P \sin P. \tag{10}$$

Thus the transmitted intensity is zero when the polarizer and analyser easy axes are at right angles, with one along the o direction and the other along the e direction. Extinction thus determines the o–e pair of axes, but does not distinguish between the o and e directions.

The intensity, obtained as the absolute square of formula (9), is proportional to

$$|t_o|^2 \cos^2 P \cos^2 A + 2\,\mathrm{Re}\,(t_o t_e^*) \cos P \cos A \sin P \sin A + |t_e|^2 \sin^2 P \sin^2 A. \tag{11}$$

There are three unknowns: $|t_o|$, $|t_e|$, and the phase difference between t_o and t_e. Three intensity measurements at different polarizer or analyser settings, plus an intensity measurement with the sample absent, are in principle sufficient to find the absolute magnitudes of the transmission amplitudes, and their relative phase.

The expression (9) for the final field amplitude was obtained by using the fundamental t_o and t_e transmission amplitudes, and resolving along the o and e directions. We can check that the same expression results on using formula (6): the field transmitted by the sample has components along and perpendicular to the polarizer easy direction equal to

$$t_o \cos^2 P + t_e \sin^2 P \qquad (t_o - t_e) \cos P \sin P. \tag{12}$$

The field transmitted by the analyser is thus

$$(t_o \cos^2 P + t_e \sin^2 P) \cos(P - A) + (t_o - t_e) \cos P \sin P \sin (P - A)$$
$$= t_o \cos P \cos A + t_e \sin P \sin A \tag{13}$$

in agreement with formula (9).

4. Polarizer–compensator–sample–analyser

We consider now the effect of inserting a compensator (also known as a waveplate, or a retarder) between the polarizer and the sample. The compensator is a crystal plate, or combination of plates, which produces a known phase difference between two orthogonal components of the transmitted electric field. We will call these orthogonal directions o′ and e′, and C the angle between the o′ and o directions of compensator and sample (see figure 2).

We first resolve along the o′ and e′ directions of the compensator. After the polarizer, the respective field components are $\cos (P - C)$ and $\sin (P - C)$. After the compensator the o′ and e′ components of the electric field are $t'_o \cos (P - C)$ and $t'_e \sin (P - C)$, where t'_o and t'_e are the complex compensator transmission amplitudes for light polarized along the o′ and e′ directions.

We now resolve along the o and e directions of the crystal. The electric field components are

$$\left.\begin{aligned} E_o &= t'_o \cos (P - C) \cos C - t'_e \sin (P - C) \sin C \\ E_e &= t'_o \cos (P - C) \sin C + t'_e \sin (P - C) \cos C \end{aligned}\right\}. \tag{14}$$

After transmission through the sample the electric field components along o and e are $t_o E_o$ and $t_e E_e$, and after the analyser the final field is

$$t_o E_o \cos A + t_e E_e \sin A. \tag{15}$$

The intensity is proportional to the absolute square of this quantity. We consider the

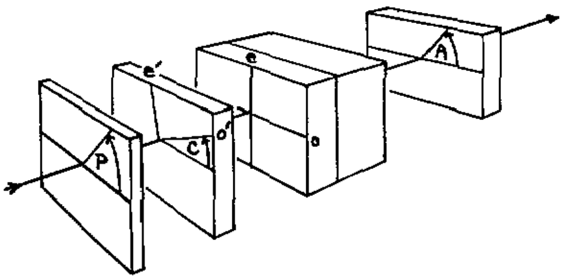

Figure 2. The polarizer–compensator–sample–analyser transmission ellipsometer arrangement.

particular case of a null setting of the ellipsometer, in which the intensity is made zero (in practice minimized). The intensity will be zero when the real and imaginary parts of the equality

$$t_o/t_e = -\frac{E_e}{E_o}\tan A \tag{16}$$

are satisfied. From equations (14) we see that

$$\frac{E_e}{E_o} = \frac{t'_o \tan C + t'_e \tan(P-C)}{t'_o - t'_e \tan C \tan(P-C)} \equiv \tan(C+D) \tag{17}$$

where $D = D_r + iD_i$ is a complex angle defined by

$$\tan D = \frac{t'_e}{t'_o}\tan(P-C) \tag{18}$$

Thus a null setting determines the complex ratio t_o/t_e in terms of the compensator transmission amplitude ratio and the angles P, C and A:

$$t_o/t_e = -\tan(C+D)\tan A\,. \tag{19}$$

5. Polariser–sample–compensator–analyser

This ellipsometer arrangement is as in the previous section, except that the compensator follows the sample. We resolve first along the sample o and e directions. As in section 3, the o and e components of the field after passing through the sample are $t_o \cos P$ and $t_e \sin P$. We now resolve along the o′ and e′ directions of the compensator. The electric field components are

$$\left.\begin{aligned} E'_o &= t_o \cos P \cos C + t_e \sin P \sin C \\ E'_e &= -t_o \cos P \sin C + t_e \sin P \cos C \end{aligned}\right\}. \tag{20}$$

After passing through the compensator the o′ and e′ components are $t'_o E'_o$ and $t'_e E'_e$. After the analyser the field (along the analyser easy direction) is

$$t'_o E'_o \cos(A-C) + t'_e E'_e \sin(A-C)\,. \tag{21}$$

We again define a complex angle related to the ratio t'_e/t'_o:

$$\tan D' = \frac{t'_e}{t'_o}\tan(A-C)\,. \tag{22}$$

A null setting is obtained when

$$\frac{E'_o}{E'_e} = -\tan D' \tag{23}$$

which gives the complex ratio t_o/t_e in terms of the compensator transmission amplitude ratio and the angles P, C and A:

$$\frac{t_o}{t_e} = -\tan P \tan(C+D')\,. \tag{24}$$

6. Polarizer–modulator–sample–analyser

We now turn to polarization modulation ellipsometry, which has been extensively used in reflection studies of interfaces (Jasperson and Schnatterly 1969, Beaglehole 1980). The use of a birefringence modulator to determine the polarization properties of light has been described in detail by Badoz *et al* (1977). The polarization state of the light is varied sinusoidally, with synchronous (lock-in) detection of the intensity. The modulator may be, for example, a block of fused quartz glued to a similar block of crystalline quartz, to which is applied an oscillatory electric field. The piezoelectric stress produces an oscillatory birefringence in the fused quartz, which then acts as a modulator with an oscillatory phaseshift

$$\frac{t'_e}{t'_o} \approx e^{i\delta} \qquad \delta(t) \approx M \sin \Omega t. \tag{25}$$

Here M is the maximum phaseshift, and $\Omega/2\pi$ is the modulation frequency. The first equality in (25) is approximate, since the ratio t'_e/t'_o does not have unit modulus exactly, as noted in the discussion at the end of section 2. For most birefringent modulators $|t'_e/t'_o|$ is very close to unity, and we set it equal to unity for simplicity. The second equality in (25) is also approximate, as has been shown by Acher *et al* (1988), who considered the effect of residual strain in the modulator, and of higher harmonics in the modulation. Again, we assume (25) to be true here.

Since the modulator is a compensator with a sinusoidally varying phase, we can use part of the analysis of section 5, making the substitution (25). The field passing the analyser has amplitude given by (14) and (15), which we write as

$$t'_o t_e \cos(P-C)\cos C \cos A\{\tau + \tan C \tan A + e^{i\delta}\tan(P-C)[\tan A - \tau \tan C]\} \tag{26}$$

where

$$t_o/t_e = \tau = \tau_r + i\tau_i \tag{27}$$

is the quantity to be determined. The absolute square of the expression inside the braces in equation (26) is

$$\begin{aligned}(\tau_r^2 + \tau_i^2)&[1 + \tan^2(P-C)\tan^2 C] + 2\tau_r \tan C \tan A[1 - \tan^2(P-C)]\\ &+ \tan^2 A[\tan^2(P-C) + \tan^2 C]\\ &+ 2\tau_i \sin\delta \tan(P-C)\tan A(1 + \tan^2 C) - 2\cos\delta \tan(P-C)\\ &\times \{(\tau_r^2 + \tau_i^2)\tan C + \tau_r \tan A(\tan^2 C - 1) - \tan C \tan^2 A\}.\end{aligned} \tag{28}$$

Now sinusoidal functions of sinusoidal arguments are periodic, and can thus be expressed as Fourier series, the coefficients of which are Bessel functions (Watson 1966, section 2.22):

$$\left.\begin{aligned}\cos(M\sin\Omega t) &= J_0(M) + 2\sum_{n=1}^{\infty} J_{2n}(M)\cos(2n\Omega t)\\ \sin(M\sin\Omega t) &= 2\sum_{n=0}^{\infty} J_{2n+1}(M)\sin((2n+1)\Omega t)\end{aligned}\right\}. \tag{29}$$

Thus if we write expression (28) as $a + b\sin\delta + c\cos\delta$, the DC, Ω and 2Ω parts of the intensity are proportional to

$$\left.\begin{array}{ll} \text{DC:} & a + cJ_0(M) \\ \Omega: & 2bJ_1(M)\sin\Omega t \\ 2\Omega: & 2cJ_2(M)\cos 2\Omega t \end{array}\right\}. \tag{30}$$

The Ω signal is proportional to the imaginary part of t_o/t_e, while the DC and 2Ω signals depend on both the absolute square and the real part of this ratio.

In considering actual intensities, rather than ratios of intensities, it is better to include the factor $[\cos(P-C)\cos C\cos A]^2$ (see expression (26)). The intensity then becomes $\frac{1}{2}|t'_o t_e|^2\ (a' + b'\sin\delta + c'\cos\delta)$, where

$$\left.\begin{array}{l} a' = (\tau_r^2 + \tau_i^2)\cos^2 A[1 + \cos 2(P-C)\cos 2C] + \tau_r\sin 2C\sin 2A\cos 2(P-C) \\ \qquad + \sin^2 A[1 - \cos 2(P-C)\cos 2C] \\ b' = \tau_i\sin 2(P-C)\sin 2A \\ c' = \sin 2(P-C)\{\sin 2C\sin^2 A + \tau_r\cos 2C\sin 2A - (\tau_r^2 + \tau_i^2)\sin 2C\} \end{array}\right\}. \tag{31}$$

When $P - C = 45°$ this simplifies to

$$\left.\begin{array}{l} a' = (\tau_r^2 + \tau_i^2)\cos^2 A + \sin^2 A \\ b' = \tau_i\sin 2A \\ c' = \sin 2C\sin^2 A + \tau_r\cos 2C\sin 2A - (\tau_r^2 + \tau_i^2)\sin 2C \end{array}\right\}. \tag{32}$$

We note that rotating the sample gives the reference o or e direction by the position at which the Ω signal is zero, since b' is zero when the o or e direction coincides with the analyser easy axis (i.e. when $A = 0$ or $90°$).

7. Polarizer–sample–modulator–analyser

The final configuration we consider is a polarization modulation ellipsometer with the modulator between the sample and the analyser. From expressions (20) and (21) of section 5, the field amplitude passing the analyser can be written as

$$t'_o t_e\cos P\cos C\cos(A-C)\{\tau + \tan P\tan C + e^{i\delta}\tan(A-C)[\tan P - \tau\tan C]\}. \tag{33}$$

The absolute square of the quantity inside the braces is

$$\begin{aligned} &(\tau_r^2 + \tau_i^2)[1 + \tan^2(A-C)\tan^2 C] + 2\tau_r\tan P\tan C[1 - \tan^2(A-C)] \\ &\quad + \tan^2 P[\tan^2 C + \tan^2(A-C)] \\ &\quad + 2\tau_i\tan P\tan(A-C)(1 + \tan^2 C)\sin\delta - 2\tan(A-C) \\ &\quad \times\{(\tau_r^2 + \tau_i^2)\tan C + \tau_r\tan P(\tan^2 C - 1) - \tan^2 P\tan C\}\cos\delta. \end{aligned} \tag{34}$$

If again we write this as $a + b\sin\delta + c\cos\delta$, the DC, Ω and 2Ω signals will be given by expressions (30). Again the $\sin\Omega$ part is proportional to τ_i, while both $\tau_r^2 + \tau_i^2$ and τ_r enter into the DC and $\cos 2\Omega t$ parts. We also note that expression (34) can be obtained from expression (28) by interchanging A and P.

The actual intensities are $\frac{1}{2}|t_o' t_e|^2(a' + b' \sin\delta + c' \cos\delta)$, where

$$\left.\begin{aligned} a' &= (\tau_r^2 + \tau_i^2)\cos^2 P[1 + \cos 2(A - C)\cos 2C] + \tau_r \sin 2P \sin 2C \cos 2(A - C) \\ &\qquad + \sin^2 P[1 - \cos 2C \cos 2(A - C)] \\ b' &= \tau_i \sin 2P \sin 2(A - C) \\ c' &= \sin 2(A - C)\{\sin^2 P \sin 2C + \tau_r \sin 2P \cos 2C - (\tau_r^2 + \tau_i^2)\sin 2C\} \end{aligned}\right\}. \tag{35}$$

When $A - C = 45°$ the coefficients simplify to

$$\left.\begin{aligned} a' &= (\tau_r^2 + \tau_i^2)\cos^2 P + \sin^2 P \\ b' &= \tau_i \sin 2P \\ c' &= \sin^2 P \sin 2C + \tau_r \sin 2P \cos 2C - (\tau_r^2 + \tau_i^2)\sin 2C \end{aligned}\right\}. \tag{36}$$

In this case rotating the sample and noting the position at which the Ω signal is zero gives the o or e direction as then coincident with the easy axis of the polarizer.

8. Properties of the transmission ratio t_o/t_e

We begin with the simplest case, that of an unsupported crystal plate of thickness d. From equations (3) and (4) we have (setting $k_2 = k_1$)

$$t_o = \frac{4k_1 k_o Z_o}{(k_1 + k_o)^2 - (k_1 - k_o)^2 Z_o^2} \qquad Z_o = e^{ik_o d}. \tag{37}$$

The formula for t_e is obtained by replacing k_o by k_e. Thus

$$\tau = t_o/t_e = \frac{k_o Z_o}{k_e Z_e} \times \frac{(k_1 + k_e)^2 - (k_1 - k_e)^2 Z_e^2}{(k_1 + k_o)^2 - (k_1 - k_o)^2 Z_o^2}. \tag{38}$$

The ratio $Z_o/Z_e = \exp i(k_o - k_e)d$ gives the phase change which is usually assumed to hold between the o and e waves. For thin crystal plates we set $Z^2 = 1 + 2ikd + \cdots$ (with subscripts o and e), and find that to first order in the plate thickness

$$\tau = 1 + id(k_o - k_e)(k_o + k_e)/2k_1 + \dots. \tag{39}$$

Thus multiple reflections inside the crystal alter the first-order term by the factor $(k_o + k_e)/2k_1$. (When the optic axis is perpendicular to the surface normal, this factor equals $(n_o + n_e)/2n_1$, which is 1.57 for calcite in air.) For small thicknesses, measurement of the imaginary part of τ gives the quantity

$$(k_o^2 - k_e^2)d/2k_1 = \frac{\pi d}{\lambda_1} \times \frac{n_o^2(n_o^2 - n_e^2)\sin^2\psi}{n_o^2 \sin^2\psi + n_e^2 \cos^2\psi} \tag{40}$$

where λ_1 is the wavelength in the medium of incidence, ψ is the angle between the optic axis and the normal to the crystal plate, and equation (5) has been used.

If we imagine the thickness d of the crystal to increase continuously, $Z_o = e^{ik_o d}$ and $Z_e = e^{ik_e d}$ move around the unit circle in the complex plane, with periods in d equal to $2\pi/k_o$ and $2\pi/k_e$. The transmission amplitude ratio $\tau = t_o/t_e$ also follows a path in the complex plane as d increases. We first consider the behaviour of τ for thin layers. We can rewrite equation (38) as

$$\tau = \frac{t_o}{t_e} = \frac{k_o}{k_e} \times \frac{2c_e k_1 k_e - \mathrm{i}s_e(k_1^2 + k_e^2)}{2c_o k_1 k_o - \mathrm{i}s_o(k_1^2 + k_o^2)} \tag{41}$$

where $c_o = \cos(k_o d)$, $s_o = \sin(k_o d)$, and similarly for c_e and s_e. Thus the real and imaginary parts of τ are

$$\left.\begin{aligned} \tau_r &= \left(\frac{k_o}{k_e}\right) \frac{4c_o c_e k_1^2 k_o k_e + s_o s_e (k_1^2 + k_e^2)(k_1^2 + k_o^2)}{4k_1^2 k_o^2 + s_o^2 (k_1^2 - k_o^2)^2} \\ \tau_i &= \left(\frac{2k_1 k_o}{k_e}\right) \frac{s_o c_e k_e (k_1^2 + k_o^2) - s_e c_o k_o (k_1^2 + k_e^2)}{4k_1^2 k_o^2 + s_o^2 (k_1^2 - k_o^2)^2} \end{aligned}\right\}. \tag{42}$$

Figure 3 shows how τ_r and τ_i vary with crystal thickness, up to about $1\frac{1}{4}$ wavelengths. We see that the imaginary part is linear in the thickness only for very thin samples. Both τ_r and τ_i are non-monotonic for larger thicknesses, so that a given value of either may not be associated with a unique thickness. We shall see shortly that even when both the real and imaginary parts are measured the thickness is not fixed uniquely.

We now consider the path of τ in the complex plane as the thickness of the crystal plate increases. From equation (38) we see that τ is made up of functions with periods π/k_o, π/k_e and $2\pi/|k_o - k_e|$ in the layer thickness d. The path of τ as d increases is like that of a satellite circling a planet which orbits the sun. The motion is not periodic unless k_o/k_e is rational, but τ crosses the positive real axis each time that d is close to a multiple of $2\pi/|k_o - k_e|$. We note that the path crosses itself many times during one orbit (see figure 4). At each crossing the same value of τ results from two nearby values of d.

The absolute square of τ can be written as

$$|\tau|^2 = \frac{k_o^2}{k_e^2} \times \frac{(k_1^2 + k_e^2)^2 - (k_1^2 - k_e^2)^2 \cos^2(k_e d)}{(k_1^2 + k_o^2)^2 - (k_1^2 - k_o^2)^2 \cos^2(k_o d)}. \tag{43}$$

Thus τ must lie within the annulus formed by two circles centred on the origin:

$$\frac{2k_1 k_o}{k_1^2 + k_o^2} \leqslant |\tau| \leqslant \frac{k_1^2 + k_e^2}{2k_1 k_e}. \tag{44}$$

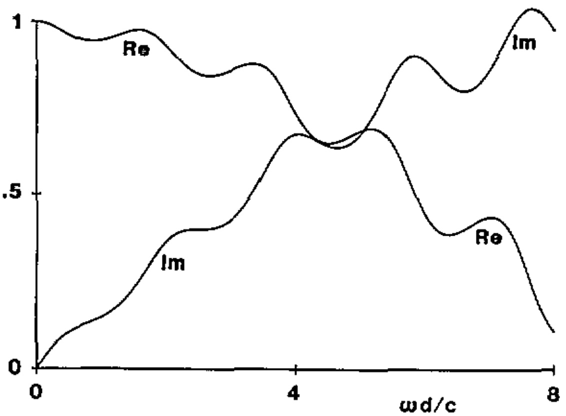

Figure 3. The real and imaginary parts of $\tau = t_o/t_e$ for a thin calcite layer in air, as a function of the layer thickness. The optic axis is taken to lie parallel to the layer.

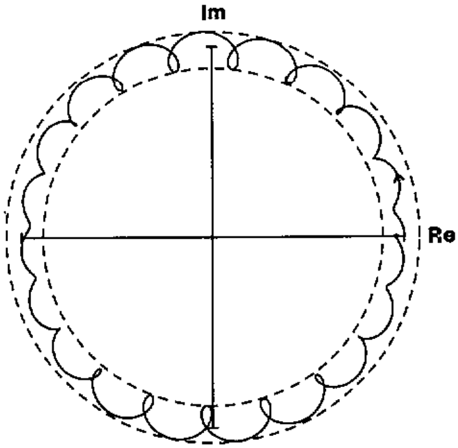

Figure 4. The path traced by $\tau = t_o/t_e$ in the complex plane as the thickness of the anisotropic layer increases from zero to one 'orbital period' $2\pi/|k_o - k_e|$. The curve is drawn for calcite, with its optic axis perpendicular to the surface normal. The arrow shows the direction of increasing layer thickness. The real and imaginary axes are each drawn from -1 to $+1$. The 'orbit' starts at $\tau = 1$ and ends at the point indicated by a dot. The broken curves are the inner and outer bounding circles defined in equation (44).

This spread of values of the magnitude of τ is attained when its phase is near $\pm\pi/2$, but a narrower spread occurs when the phase of τ is close to 0 or π.

We now consider an anisotropic layer on a substrate of index n_2 and thickness d_2. (The discussion following expressions (8) noted that the transmission properties are the same when the light enters the supporting layer first.) The ratio $\tau = t_o/t_e$ is obtained from expressions (7), (8) and the corresponding equations with k_e replacing k_o. For a thin anisotropic layer the transmission amplitude becomes, to first order in d

$$\tau = 1 + \frac{(k_o^2 - k_e^2)\{ik_1 s_2 - k_2 c_2\} d}{2ik_1 k_2 c_2 + (k_1^2 + k_2^2) s_2} + O(d^2) \tag{45}$$

where $c_2 = \cos k_2 d_2$ and $s_2 = \sin k_2 d_2$, as in (8). This expression reduces to (39) when $d_2 \to 0$, and also when $k_2 \to k_1$ (in either case the substrate has no optical effect). Note that, in contrast to (39), the first-order term is complex rather than imaginary. The real and imaginary parts are seen by writing the coefficient of d in (45) as

$$(k_o^2 - k_e^2)\frac{\{(k_1^2 - k_2^2)k_2 s_2 c_2 + i[2k_2^2 + s_2^2(k_1^2 - k_2^2)]\}}{(2k_1 k_2)^2 + (k_1^2 - k_2^2)^2 s_2^2}. \tag{46}$$

We see that the real part is zero when $\sin 2k_2 d_2$ is zero, but that the imaginary part is non-zero for all values of $k_2 d_2$.

The 'substrate' may in fact be a thin isotropic layer (for example a liquid film) supported by the crystal. Thus it is interesting to examine the behaviour of τ as d_2 tends to zero. We find that τ is the sum of (41) plus a series in powers of d_2, the coefficient of d_2 being

$$\frac{k_o}{k_e}\frac{(k_2^2 - k_1^2)\{k_1^2(k_o c_o s_e - k_e c_e s_o) + k_o k_e(k_o s_o c_e - k_e s_e c_o) + i(k_e^2 - k_o^2)k_1 s_o s_e\}}{[(k_1^2 + k_o^2)s_o + 2ik_1 k_o c_o]^2}. \tag{47}$$

As one would expect, this becomes zero when the anisotropy tends to zero, and also when $n_2 \to n_1$.

We comment finally on *absorbing* media. The fundamental formulae derived here are valid for absorbing crystals and/or substrates. Absorption is allowed for by means of complex values of the dielectric constants ε_o, ε_e and ε_2, or of the refractive indices n_o, n_e and n_2. The wavevectors k_o, k_e and k_2 will be complex if the media are absorbing. Throughout this paper we have made simplifications based on the tacit assumption that k_o, k_e and k_2 are real. For example, (41) is general but (42), (43) and (44) assume that the ks are real. Likewise (45) is general but in deducing (46) we have assumed that k_2 is real.

Appendix. Normal incidence onto biaxial crystals

When a wave is incident onto any planar stratified system which is invariant with respect to translation along the system surface, the wavevector components parallel to the surface will be conserved. For normal incidence this implies that the wavevector $\boldsymbol{k}$ inside the anisotropic medium will be along the inward surface normal. (The propagation of energy, with direction given by $\boldsymbol{E} \times \boldsymbol{B}$, is not in general along $\boldsymbol{k}$, as is well known.)

To find the possible values of $|\boldsymbol{k}|$ inside a homogeneous anisotropic medium, we eliminate the magnetic field $\boldsymbol{B}$ from the Maxwell curl equations. For plane waves, with space and time dependence $\exp \mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)$, these read

$$\boldsymbol{k} \times \boldsymbol{E} = \frac{\omega}{c}\boldsymbol{B} \qquad \boldsymbol{k} \times \boldsymbol{B} = -\frac{\omega}{c}\varepsilon\boldsymbol{E} \tag{A1}$$

where ε is the dielectric tensor. On eliminating $\boldsymbol{B}$ by taking the vector product of the first equation in (A1) with $\boldsymbol{k}$ and substituting from the second, one obtains (see, for example, Born and Wolf 1965, section 14.2.1)

$$(\varepsilon\omega^2/c^2 - k^2)\boldsymbol{E} + \boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{E}) = 0. \tag{A2}$$

We look at the possible solutions of (A2) in the principal dielectric axes frame, in which

$$\varepsilon = \begin{pmatrix} \varepsilon_a & 0 & 0 \\ 0 & \varepsilon_b & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix}. \tag{A3}$$

The three equations (A2) then read

$$\left.\begin{aligned} (\varepsilon_a\omega^2/c^2 - k^2)E_a + k_a(k_aE_a + k_bE_b + k_cE_c) = 0 \\ (\varepsilon_b\omega^2/c^2 - k^2)E_b + k_b(k_aE_a + k_bE_b + k_cE_c) = 0 \\ (\varepsilon_c\omega^2/c^2 - k^2)E_c + k_c(k_aE_a + k_bE_b + k_cE_c) = 0 \end{aligned}\right\}. \tag{A4}$$

This is a set of three homogeneous equations in the unknowns E_a, E_b and E_c, and a solution with non-zero $\boldsymbol{E}$ exists if the determinant of the coefficients of E_a, E_b and E_c is zero. This leads to 'Fresnel's equation of wave normals' (Born and Wolf 1965, section 14.2.2), which we write in the form

$$(\alpha^2\varepsilon_a + \beta^2\varepsilon_b + \gamma^2\varepsilon_c)\left(\frac{ck}{\omega}\right)^4 - \{(\alpha^2+\beta^2)\varepsilon_a\varepsilon_b + (\beta^2+\gamma^2)\varepsilon_b\varepsilon_c + (\gamma^2+\alpha^2)\varepsilon_c\varepsilon_a\}\left(\frac{ck}{\omega}\right)^2 + \varepsilon_a\varepsilon_b\varepsilon_c = 0 \tag{A5}$$

where α, β and γ are the direction cosines of $\boldsymbol{k}$ relative to the principal dielectric axes $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, so that

$$\boldsymbol{k} = k(\alpha, \beta, \gamma) \qquad \alpha^2 + \beta^2 + \gamma^2 = 1\,. \tag{A6}$$

Equation (A5) is a quadratic in k^2, giving the magnitude of the wavevector as a function of its direction. We label the solutions $k_\pm^2$, according to the sign chosen in taking the square root of the discriminant of equation (A.5). We will note some properties of the solutions. Suppose (as is conventional) the axes are chosen so that

$$\varepsilon_a \leqslant \varepsilon_b \leqslant \varepsilon_c\,. \tag{A7}$$

Then bounds on $k_\pm^2$ are

$$\varepsilon_a \leqslant \left(\frac{ck_-}{\omega}\right)^2 \leqslant \varepsilon_b \qquad \varepsilon_b \leqslant \left(\frac{ck_+}{\omega}\right)^2 \leqslant \varepsilon_c\,. \tag{A8}$$

Equality in (A8) is attained when $\boldsymbol{k}$ is parallel or perpendicular to one of the dielectric axes. When $\alpha^2 = 1$, for example, $(ck/\omega)^2$ takes the values ε_b and ε_c, while when $\alpha = 0$ it takes the values ε_a and $\varepsilon_b\varepsilon_c/(\beta^2\varepsilon_b + \gamma^2\varepsilon_c)$.

The discriminant of the quadratic for $(ck/\omega)^2$ can be written as

$$\alpha^4\varepsilon_a^2(\varepsilon_b-\varepsilon_c)^2 + \beta^4\varepsilon_b^2(\varepsilon_c-\varepsilon_a)^2 + \gamma^4\varepsilon_c^2(\varepsilon_a-\varepsilon_b)^2 + 2\alpha^2\beta^2\varepsilon_a\varepsilon_b(\varepsilon_c-\varepsilon_a)(\varepsilon_c-\varepsilon_b) + 2\beta^2\gamma^2\varepsilon_b\varepsilon_c(\varepsilon_a-\varepsilon_b)(\varepsilon_a-\varepsilon_c) + 2\gamma^2\alpha^2\varepsilon_c\varepsilon_a(\varepsilon_b-\varepsilon_c)(\varepsilon_b-\varepsilon_a)\,. \tag{A9}$$

From inequality (A7), all the terms are non-negative except for the last. When $\boldsymbol{k}$ is perpendicular to the $\boldsymbol{b}$ axis ($\beta = 0$), three of the non-negative terms are zero, and the negative term remains. Thus the minimum of the discriminant occurs when $\beta = 0$, its value then being

$$\{\alpha^2\varepsilon_a(\varepsilon_c-\varepsilon_b) - \gamma^2\varepsilon_c(\varepsilon_b-\varepsilon_a)\}^2. \tag{A10}$$

The discriminant is thus zero when the direction cosines α, β, γ of $\boldsymbol{k}$ are given by

$$\alpha^2 = \frac{\varepsilon_c(\varepsilon_b-\varepsilon_a)}{\varepsilon_b(\varepsilon_c-\varepsilon_a)} \qquad \beta = 0 \qquad \gamma^2 = \frac{\varepsilon_a(\varepsilon_c-\varepsilon_b)}{\varepsilon_b(\varepsilon_c-\varepsilon_a)}. \tag{A11}$$

Thus there are four directions of $\boldsymbol{k}$, given by equations (A11), for which the discriminant is zero and the roots $k_\pm^2$ of equation (A5) are equal. These directions, by definition, determine the two optic axes, along which the phase speeds of the two modes are equal ($\boldsymbol{k}$ and $-\boldsymbol{k}$ are taken to be equivalent in the definition of optic axes). The common eigenvalue of equation (A5) when equations (A11) hold is $k^2 = \varepsilon_b\omega^2/c^2$. Note that the optic axes lie in the $\boldsymbol{a}$, $\boldsymbol{c}$ plane, and have directions $(|\alpha|, 0, \pm|\gamma|)$, where α^2 and γ^2 are given by equations (A11). The cosine of the angle between the optic axes is therefore

$$\frac{\varepsilon_a\varepsilon_b + \varepsilon_b\varepsilon_c - 2\varepsilon_c\varepsilon_a}{\varepsilon_b(\varepsilon_c - \varepsilon_a)}. \tag{A12}$$

Thus the principal dielectric axis a is the acute bisectrix of the optic axes if ε_b is greater than the harmonic mean of ε_a and ε_c.

Once equation (A5) is solved for the magnitude of k, the direction of the corresponding electric field vector can be found from (A2):

$$E_a = \frac{k_a(\boldsymbol{k}\cdot\boldsymbol{E})}{k^2 - \varepsilon_a\omega^2/c^2} \qquad \text{etc.} \tag{A13}$$

Thus, for given direction (α, β, γ) of $\boldsymbol{k}$, the electric field components are in the ratio

$$E_a{:}E_b{:}E_c = \frac{\alpha}{(ck/\omega)^2 - \varepsilon_a}:\frac{\beta}{(ck/\omega)^2 - \varepsilon_b}:\frac{\gamma}{(ck/\omega)^2 - \varepsilon_c} \tag{A14}$$

where k^2 is one of the solutions k_+^2, k_-^2 of equation (A5). (In the cases where one or more of the denominators in equation (A14) is zero, the directions of $\boldsymbol{E}$ must be found directly from expression (A2), to avoid division by zero in equation (A13).)

$\boldsymbol{E}^+$ and $\boldsymbol{E}^-$ are not orthogonal in general, but we will prove that their projections onto the reflecting surface are orthogonal. To verify this we show that, for normal incidence,

$$(\boldsymbol{n} \times \boldsymbol{E}^+)\cdot(\boldsymbol{n} \times \boldsymbol{E}^-) = 0 \tag{A15}$$

where $\boldsymbol{n}$ is the inward surface normal. This is the same as the unit vector in the $\boldsymbol{k}$ direction, i.e. $\boldsymbol{n} = (\alpha, \beta, \gamma)$, so

$$\boldsymbol{n} \times \boldsymbol{E} = (\beta E_c - \gamma E_b, \gamma E_a - \alpha E_c, \alpha E_b - \beta E_a). \tag{A16}$$

From equation (A14), this is proportional to the vector

$$\beta\gamma(\varepsilon_b - \varepsilon_c)[\varepsilon_a - (ck/\omega)^2], \gamma\alpha(\varepsilon_c - \varepsilon_a)[\varepsilon_b - (ck/\omega)^2], \alpha\beta(\varepsilon_a - \varepsilon_b)[\varepsilon_c - (ck/\omega)^2]. \tag{A17}$$

Thus the scalar product in equation (A15) is the cyclic sum of terms like

$$\beta^2\gamma^2(\varepsilon_b - \varepsilon_c)^2(\varepsilon_a - \varepsilon_+)(\varepsilon_a - \varepsilon_-) \tag{A18}$$

where $\varepsilon_\pm = (ck_\pm/\omega)^2$ are the solutions of equation (A5). The quadratic equation (A5), of the form

$$A(ck/\omega)^3 + B(ck/\omega)^2 + C = 0 \tag{A19}$$

has solutions ε_+ and ε_- whose product and sum are given by

$$\varepsilon_+\varepsilon_- = C/A \qquad \varepsilon_+ + \varepsilon_- = -B/A. \tag{A20}$$

Thus the product of the final two terms in (A18) is equal to

$$(A\varepsilon_a^2 + B\varepsilon_a + C)/A. \tag{A21}$$

On using $\alpha^2 + \beta^2 + \gamma^2 = 1$ and the values of A, B and C as given in (A5), it follows from the algebraic identity

$$\beta^2\gamma^2(\varepsilon_b - \varepsilon_c)^2(A\varepsilon_a^2 + B\varepsilon_a + C) + \text{cyclic terms} = 0 \tag{A22}$$

that the projections of $\boldsymbol{E}^+$ and $\boldsymbol{E}^-$ onto the plane perpendicular to $\boldsymbol{n}$ are orthogonal. Thus, at normal incidence, biaxial crystals have both modes with wavevector normal to the surface, of magnitudes k_+ and k_-, and electric field components in the plane of the surface are orthogonal. The continuity of the tangential components of $\boldsymbol{E}$ and of their derivatives at the boundaries of the crystal, which determines the reflection and transmission amplitudes, can thus be applied to biaxial crystals in the same way as to uniaxial crystals.

We note finally that the vector identity

$$(\boldsymbol{a} \times \boldsymbol{b})\cdot(\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d}) - (\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c}) \tag{A23}$$

applied to (A15) shows that at normal incidence

$$\boldsymbol{E}^+\cdot\boldsymbol{E}^- = (\boldsymbol{n}\cdot\boldsymbol{E}^+)(\boldsymbol{n}\cdot\boldsymbol{E}^-). \tag{A24}$$

That is, the scalar product of the two electric fields is equal to the product of the components of $\boldsymbol{E}^+$ and $\boldsymbol{E}^-$ along the direction of propagation.

Acknowledgment

The author is grateful to Professor David Beaglehole for suggesting this problem, and for stimulating discussions.

References

Acher O, Bigan E and Drévillon, B 1989 *Rev. Sci. Instrum.* **60** 65
Acher O and Drévillon, B 1992 *Rev. Sci. Instrum.* **63** 5332
Archer R J and Shank C V 1967 *J. Opt. Soc. Am.* **57** 191
Azzam R M A and Bashara N M 1987 *Ellipsometry and Polarized Light* (Amsterdam: North-Holland)
Badoz J, Billardon M, Cauit J and Russel M F 1977 *J. Optique* **8** 373
Beaglehole D 1980 *Physica* **B 100** 163
Born M and Wolf E 1965 *Principles of Optics* (Oxford: Pergamon)
Eidner K, Mayer G, Schmidt M and Schniedel H 1989 *Mol. Cryst. Liq. Cryst.* **172** 191
Gribble C D and Hall A J 1985 *A Practical Introduction to Optical Mineralogy* (London: Allen and Unwin)
Holmes D A 1964 *J. Opt. Soc. Am.* **54** 1115
Jasperson S N and Schnatterly S E 1969 *Rev. Sci. Instrum.* **40** 761
Lekner J 1987 *Theory of Reflection* (Dordrecht: Nijhoff/Kluwer)
—— 1991 *J. Phys. Condens. Matter* **3** 6121
—— 1992 *J. Phys. Condens. Matter* **4** 1387
Watson G N 1966 *Theory of Bessel Functions* (Cambridge: Cambridge University Press)
Wöhler H, Hass G, Fritsch M and Mlynski D A 1988 *J. Opt. Soc. Am.* **A 5** 1554
Yolken H T, Waxler R M and Kruger J 1967 *J. Opt. Soc. Am.* **57** 283–284