Optical properties of a uniaxial layer

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

Received 21 January 1994, in final form 28 April 1994

Abstract. The transmission and reflection properties of a homogeneous anisotropic uniaxial layer are discussed. The layer may be transparent or absorbing. Analytic expressions are given for the elements of a $4 \times 4$ mode matrix $M$ characterizing such a layer, for any angle of incidence and for arbitrary orientation of the optic axis. The reflection and transmission amplitudes are given in terms of elements of a layer matrix $L = MPMP^{-1}$, where the diagonal phase matrix $P$ has as elements the phase factors for the ordinary and extraordinary waves as they traverse the layer in the forward and backward directions. Analytic expressions for the reflection and transmission amplitudes are given explicitly for the case when the optic axis of the layer lies in the plane of incidence, when the layer is thin, and when the layer anisotropy is weak. Application is made to anisotropic antireflection coatings, and to the modelling of slightly rough surfaces by anisotropic layers.

1. Introduction

In two recent papers ([1, 2]) the author has given analytic expressions for the optical coefficients of uniaxial crystals, and of crystal plates illuminated at normal incidence. These results will be extended here to the optical properties of a uniaxial crystal plate, bounded by isotropic media of dielectric constants $\varepsilon_1 = n_1^2$ and $\varepsilon_2 = n_2^2$. We will use a $4 \times 4$ matrix method, which is not different in fundamentals from that in current use (see for example [3] for references and discussion of three formulations [4–6]). The new feature is that all results given here are analytic.

Light is incident from medium 1 onto a uniaxial layer of thickness $\Delta z$. For mathematical convenience we will take $z = 0$ and $z = \Delta z$ as the boundary planes of the layer. The plane of incidence is taken as the $zx$ plane. The direction cosines of the optic axis with respect to the $x$, $y$ and $z$ axes are $\alpha$, $\beta$ and $\gamma$; thus $e = (x, \beta, \gamma)$ is the unit vector giving the direction of the optic axis. When a plane monochromatic wave of angular frequency $\omega$ is incident at angle $\theta_1$ to the normal, all components of the electric and magnetic vectors in the medium of incidence, the anisotropic layer and the substrate will have $x$ and $t$ dependence contained in the factor $\exp i(Kx - \omega t)$, where

$$K = n_1 \left( \frac{\omega}{c} \right) \sin \theta_1 = n_2 \left( \frac{\omega}{c} \right) \sin \theta_2$$

is the $x$-component of all the wavevectors, and $\theta_2$ is the angle to the normal in the substrate. The $y$-component of the wavevectors is zero (by choice of coordinates, and translational invariance). The $z$-component of the wavevector of the incident wave is
Figure 1. Reflection geometry for a uniaxial layer (refractive indices $n_o$ and $n_e$) resting on an isotropic substrate of index $n_i$. The medium of incidence has index $n_i$, and the angle of incidence is $\theta_i$. The plane of incidence is the $z-x$ plane, $z$ is the inward normal and $\epsilon$ (the broken line) is the optic axis of the uniaxial layer.

$$q_1 = n_1 \left( \frac{\omega}{c} \right) \cos \theta_i$$

and it is $-q_1$ for the reflected wave, and $q_2$ for the transmitted wave, where

$$q_2^2 = \varepsilon_2 \frac{\omega^2}{c^2} - K^2 = k_2^2 - K^2.$$  

Thus the $z$-dependence of the incident, reflected and transmitted waves is given by $\exp(iq_1z)$, $\exp(-iq_1z)$ and $\exp(iq_2z)$.

The situation within the crystal layer is more complicated. Let $\varepsilon_o = n_o^2$ and $\varepsilon_e = n_e^2$ be the ordinary and extraordinary dielectric constants of the uniaxial layer. There are four plane waves that can propagate within this crystalline layer for a given incident plane wave $\exp[i(Kx + q_z - \omega t)]$. All have the $\exp[i(Kx - \omega t)]$ dependence, as stated above. The ordinary wave propagating down into the crystal layer has $z$-dependence $\exp(iq_o z)$, where

$$q_o^2 = \varepsilon_o \frac{\omega^2}{c^2} - K^2 = k_o^2 - K^2.$$  

There is a backward (or upward) propagating ordinary wave, with $z$-dependence $\exp(-iq_o z)$. The corresponding extraordinary plane waves are $\exp(iq_e^\pm z)$, where

$$q_e^\pm = \pm \tilde{q} - \chi \gamma K \Delta \varepsilon / \varepsilon_i$$

with

$$\Delta \varepsilon = \varepsilon_e - \varepsilon_o \quad \varepsilon_i = \varepsilon_o + \gamma^2 \Delta \varepsilon = n_i^2$$

$$\tilde{q}^2 = \varepsilon_0 \left\{ \varepsilon_e \gamma \frac{\omega^2}{c^2} - K^2 (\varepsilon_e - \beta^2 \Delta \varepsilon) \right\} / \varepsilon_i.$$
The four plane waves which can propagate in the crystalline layer have electric field vectors which depend on $K$, on the direction cosines $\alpha$, $\beta$, $\gamma$ of the optic axis, and on the values of the appropriate $z$-component of the wavevector, namely $\pm q_o$ and $\pm q_e$. From (4) to (6) we see that $q_o$ depends on $K$, while $q_e^\pm$ depends on $\alpha$, $\beta$ and $\gamma$, as well as on $K$. The electric field vector of the ordinary wave is given by

$$E_o = N_o(-\beta q_o, \alpha q_o - \gamma K, \beta K)$$

for the forward propagating mode. $N_o$ is a normalization constant. The backward mode has the sign of $q_o$ reversed. For both modes $E_o$ is perpendicular to the optic axis and to the appropriate wavevector $(K, 0, \pm q_o)$. The electric field vectors of the extraordinary waves are given by

$$E_e = N_e(q_o^2 - \gamma q_e K, \beta k_e^2, \gamma(k_e^2 - q_e^2) - \alpha q_e K)$$

where $q_e$ takes the values $q_e^+$ for the forward and $q_e^-$ for the backward propagating waves, and $k_e^2 = \varepsilon_e \omega^2/c^2$. These results all follow from section 3 of [1], where bounds on $q_e^\pm$ and expressions for the scalar product of $E_o$ and $E_e$ and for the direction of the extraordinary ray may also be found.

2. Mode, phase and layer matrices

The optical properties of an anisotropic layer may be characterized by four reflection amplitudes $r_{ss}$, $r_{sp}$, $r_{pp}$, $r_{ps}$ and four transmission amplitudes $t_{ss}$, $t_{sp}$, $t_{pp}$, $t_{ps}$. For example, $r_{sp}$ is the amplitude of the wave reflected into the $p$ polarization when the incident wave is $s$ polarized. The method by which these amplitudes are determined is simple in principle: the continuity of $E_x$, $E_y$, $\partial E_x/\partial z - iKE_z$ and $\partial E_y/\partial z$ (that is, the continuity of the tangential components of the vectors $E$ and $B$) is applied at the boundaries of the layer. We shall begin by evaluating the four amplitudes $r_{ss}$, $r_{sp}$, $t_{ss}$, $t_{sp}$ (incident $s$ polarization). The electric fields, with the factor $\exp[i(Kx - \omega t)]$ suppressed, are

incoming: $(0, \exp(iq_1 z), 0)$

reflected: $\exp(-iq_1 z)(r_{sp} \cos \theta_1, r_{ss}, r_{sp} \sin \theta_1)$

within layer: $a_o \exp(iq_o z)E_o^+ + b_o \exp(-iq_o z)E_o^- + a_e \exp(iq_e z)E_e^+ + b_e \exp(iq_e z)E_e^-$

transmitted: $\exp[iq_2(z - \Delta z)](t_{ss} \cos \theta_2, t_{ss}, -t_{sp} \sin \theta_2)$.

The continuity of $E_x$, $E_y$, $\partial E_x/\partial z - iKE_z$ and $\partial E_y/\partial z$ at $z = 0$ gives

$$r_{sp} \cos \theta_1 = a_o X_o^+ + b_o X_o^- + a_e X_e^+ + b_e X_e^-$$

$$1 + r_{ss} = a_o Y_o^+ + b_o Y_o^- + a_e Y_e^+ + b_e Y_e^-$$

$$-k_1 r_{sp} = a_d(q_o X_o^+ - KZ_o^+) - b_d(q_o X_o^- + KZ_o^-) + a_e(q_e X_e^+ - KZ_e^+) + b_e(q_e X_e^- - KZ_e^-)$$

$$q_1(1 - r_{ss}) = a_o q_o Y_o^+ - b_o q_o Y_o^- + a_e q_e Y_e^+ + b_e q_e Y_e^-$$

where $k_1 = n_1 \omega/c$ and $X_o^+$ is the $x$-component of $E_o^+$, etc. At $z = \Delta z$ the boundary
conditions lead to
\[
\begin{align*}
t_{sp} \cos \theta_2 &= q_o X_0^+ + b_o X_0^- + a_e X_e^+ + b_e X_e^- \\
t_{ss} &= a_o Y_0^+ + b_o Y_0^- + a_e Y_e^+ + b_e Y_e^- \\
k_2 t_{sp} &= a_o(q_o X_0^+ - KZ_o^+) - b_o(q_o X_0^- + KZ_o^-) + a_e(q_e X_e^+ - KZ_e^+) \\
&\quad + b_e(q_e X_e^- - KZ_e^-) \\
k_2 t_{ss} &= a_o q_o Y_0^+ - b_o q_o Y_0^- + a_e q_e Y_e^+ + b_e q_e Y_e^-
\end{align*}
\]
(11)
where \( k_2 = n_2 \omega / c \), and
\[
\begin{align*}
a_o' &= \exp(iq_o \Delta z) a_o \\
b_o' &= \exp(-iq_o \Delta z) b_o \\
a_e' &= \exp(iq_e \Delta z) a_e \\
b_e' &= \exp(iq_e \Delta z) b_e.
\end{align*}
\]
(12)

The structure of (10) and (11) leads us to define the mode matrix \( \mathbf{M} \) (so named because its elements are determined by the components of the electric fields of the propagating plane wave modes),
\[
\mathbf{M} = \begin{pmatrix}
X_0^+ & X_0^- & X_e^+ & X_e^- \\
Y_o^+ & Y_o^- & Y_e^+ & Y_e^- \\
(q_o X_0^+ - KZ_o^+) & -(q_o X_0^- + KZ_o^-) & q_e X_e^+ - KZ_e^+ & q_e X_e^- - KZ_e^- \\
q_o Y_o^+ & -q_o Y_o^- & q_e Y_e^+ & q_e Y_e^-
\end{pmatrix}
\]
(13)
and the column vectors
\[
\begin{pmatrix}
\cos \theta_1 t_{sp} \\
1 + t_{ss} \\
-k_1 t_{sp} \\
q_1(1 - t_{ss})
\end{pmatrix}
\]
(14)
\[
\begin{pmatrix}
t_{ss} \\
k_2 t_{sp}
\end{pmatrix}
\]
and
\[
\begin{pmatrix}
a_o \\
b_o \\
a_e \\
b_e
\end{pmatrix}
\]
(15)

Then (10) and (11) can be written in matrix notation as
\[
\begin{align*}
\mathbf{r}_s &= \mathbf{M} \mathbf{a} \\
\mathbf{t}_s &= \mathbf{M} \mathbf{a}' = \mathbf{M} \mathbf{P} \mathbf{a}
\end{align*}
\]
where the diagonal phase matrix \( \mathbf{P} \) is given by
\[
\begin{pmatrix}
\exp(iq_e \Delta z) & 0 & 0 & 0 \\
0 & \exp(-iq_o \Delta z) & 0 & 0 \\
0 & 0 & \exp(iq_e \Delta z) & 0 \\
0 & 0 & 0 & \exp(iq_e \Delta z)
\end{pmatrix}
\]
(16)
The diagonal elements of the phase matrix give the phase change of the four plane wave modes on propagating through the layer thickness \( \Delta z \). The unknown coefficients \( a_o, b_o, a_e, b_e \) may be eliminated from (15):
\[
\begin{align*}
\mathbf{t}_s &= \mathbf{M} \mathbf{P} \mathbf{a} = \mathbf{M} \mathbf{P} \mathbf{M}^{-1} \mathbf{r}_s
\end{align*}
\]
(17)
Optical properties of a uniaxial layer

which is now a set of four simultaneous linear equations in the four wanted unknowns \( r_{ss}, r_{sp}, t_{ss}, t_{sp} \).

Before discussing the solution of this set, we look at the case of incoming \( p \) polarization. The electric fields are now

\[
\begin{align*}
\text{incoming:} & \quad \exp(iq_1 z)(\cos \theta_1, 0, -\sin \theta_1) \\
\text{reflected:} & \quad \exp(-iq_1 z)(r_{pp} \cos \theta_1, r_{ps}, r_{pp} \sin \theta_1) \\
\text{transmitted:} & \quad \exp[iq_2(z - \Delta z)](t_{pp} \cos \theta_2, t_{ps}, -t_{pp} \sin \theta_2).
\end{align*}
\]

(The definition of \( r_{pp} \) is such that \( r_{pp} \) and \( r_{ss} \) are equal at normal incidence onto an isotropic layer.) The form of the electric field within the layer is the same as given in (9), being made up of the four plane wave modes. Thus the boundary conditions lead to a similar set of equations to that obtained above for \( s \) polarization incident, with the same mode matrix \( \mathbf{M} \) and the same phase matrix \( \mathbf{P} \):

\[
\mathbf{r}_p = \mathbf{Ma}, \quad \mathbf{t}_p = \mathbf{Ma}^\prime = \mathbf{MPa}.
\]

Here the column vector \( \mathbf{a} \) is as defined in (14), and

\[
\begin{align*}
\mathbf{r}_p &= \begin{pmatrix} \cos \theta_1 (1 + r_{pp}) \\ r_{ps} \\ k_1 (1 - r_{pp}) \\ -q_1 r_{ps} \end{pmatrix} & \begin{pmatrix} \cos \theta_2 t_{pp} \\ t_{ps} \\ k_2 t_{pp} \\ q_2 t_{ps} \end{pmatrix} = \mathbf{t}_p.
\end{align*}
\]

It follows that the optical amplitudes for both polarizations are given by an equation of the form \( \mathbf{t} = \mathbf{Lr} \), with the same matrix

\[
\mathbf{L} = \mathbf{MPM}^{-1}.
\]

We call the \( 4 \times 4 \) matrix \( \mathbf{L} \) the layer matrix, since it depends on the properties of the layer through the electric field modes and the phase shifts they experience in propagating through the layer. The elements of \( \mathbf{L} \) are independent of the polarization of the incident wave; they depend on the angle of incidence through the \( x \)-component of all the wavevectors, \( K \) (since the \( z \)-components of the wavevectors and the electric field mode vectors depend on \( K \)).

We now look at the matrices \( \mathbf{M}, \mathbf{P} \) and \( \mathbf{L} \) in more detail. First we note that the normalization factors \( N_o^\pm \) and \( N_e^\pm \) for the electric field vectors may be absorbed into the coefficients \( a_o, b_o \) and \( a_e, b_e \). Thus we may set the normalization factors in (7) and (8) equal to unity without loss of generality. The resulting mode matrix is

\[
\mathbf{M} = \begin{pmatrix} -\beta q_o & \beta q_o & xq_o^2 - \gamma q_e^+ K & xq_o^2 - \gamma q_e^- K \\ xq_o - \gamma K & -xq_o - \gamma K & \beta k_o^2 & \beta k_o^- \\ -\beta k_o^2 & -\beta k_o^- & (xq_e^+ - \gamma K)k_o^2 & (xq_e^- - \gamma K)k_o^2 \\ (xq_o - \gamma K)q_o & (xq_o + \gamma K)q_o & \beta q_e^+ k_o^2 & \beta q_e^- k_o^- \end{pmatrix}.
\]

The determinant of this matrix is independent of the signs of the direction cosines of
the optic axis:

\[
det \mathbf{M} = 4k_0^2 q_0 q \left[ \left( 1 - \gamma^2 \right) k_0^2 + \left( 1 - \beta^2 \right) K^2 \right] - 4x^2 K^2 k_0^2 \]

(23)

(We note that all physical results must be invariant to the simultaneous change of sign of all the direction cosines, since \( c \) and \(-c\) are equivalent directions; \( det \mathbf{M} \) is invariant to the independent change of sign of any direction cosine.) The determinant of \( \mathbf{M} \) is non-negative, and can be zero only if the optic axis lies in the plane of incidence (\( \beta = 0 \)), and at the same time \( K^2 = x^2 k_0^2 \). This degenerate case will be considered in conjunction with the special case \( \beta = 0 \) in section 4.

3. The reflection and transmission amplitudes

The relation \( t_r = L r_s \) represents four simultaneous equations in the four unknowns \( r_{ss}, r_{sp}, t_{ss}, t_{sp} \). We solve these to find the reflection amplitudes in terms of the matrix elements \( L_{ij} \):

\[
\begin{align*}
 r_{ss} &= \frac{B_1 S_2 - S_1 B_2}{A_1 B_2 - B_1 A_2} \\
 r_{sp} &= \frac{S_1 A_2 - A_1 S_2}{A_1 B_2 - B_1 A_2}
\end{align*}
\]

(24)

where

\[
\begin{align*}
 S_1 &= k_2 (L_{12} + q_1 L_{14}) - \cos \theta_2 (L_{32} + q_1 L_{34}) \\
 A_1 &= k_2 (L_{12} - q_1 L_{14}) - \cos \theta_2 (L_{32} - q_1 L_{34}) \\
 B_1 &= k_2 (\cos \theta_1 L_{11} - k_1 L_{13}) - \cos \theta_2 (\cos \theta_1 L_{31} - k_1 L_{33}) \\
 S_2 &= q_2 (L_{22} + q_1 L_{24}) - L_{42} - q_1 L_{44} \\
 A_2 &= q_2 (L_{22} - q_1 L_{24}) - L_{42} + q_1 L_{44} \\
 B_2 &= q_2 (\cos \theta_1 L_{21} - k_1 L_{23}) - \cos \theta_1 L_{41} + k_1 L_{43}.
\end{align*}
\]

(25)

Similarly, the relation \( t_r = L r_s \) can be solved for \( r_{pp}, r_{ps}, t_{ps} \) and \( t_{pp} \). The reflection amplitudes are given by

\[
\begin{align*}
 r_{pp} &= \frac{P_1 A_2 - A_1 P_2}{A_1 B_2 - B_1 A_2} \\
 r_{ps} &= \frac{B_1 P_2 - P_1 B_2}{A_1 B_2 - B_1 A_2}
\end{align*}
\]

(26)

where \( A_i \) and \( B_i \) are as defined in (25), and

\[
\begin{align*}
 P_1 &= k_2 (\cos \theta_1 L_{11} + k_1 L_{13}) - \cos \theta_2 (\cos \theta_1 L_{31} + k_1 L_{33}) \\
 P_2 &= q_2 (\cos \theta_1 L_{21} + k_1 L_{23}) - \cos \theta_1 L_{41} - k_1 L_{43}.
\end{align*}
\]

(27)

We note that the reflection coefficients have a common denominator, as they do for reflection by a bulk crystal [1]. Note also the close correspondence between \( B_1, P_1 \) and between \( B_2, P_2 \).

As the thickness of the layer tends to zero, the phase matrix \( \mathbf{P} \) tends to the identity matrix, and so does the layer matrix \( \mathbf{L} = \mathbf{MPM}^{-1} \). Then we regain the reflection amplitudes appropriate to an interface between isotropic media of indices \( n_1 \) and \( n_2 \).
Optical properties of a uniaxial layer

\[
\begin{align*}
  r_{ss} &\rightarrow \frac{q_1 - q_2}{q_1 + q_2} & r_{sp} &\rightarrow 0 & r_{ps} &\rightarrow 0 & r_{pp} &\rightarrow \frac{Q_2 - Q_1}{Q_2 + Q_1} \\
\end{align*}
\]  

(28)

where \( Q_1 = q_1/\epsilon_1 \) and \( Q_2 = q_2/\epsilon_2 \).

When the layer is isotropic, with \( \pm q_o \) and \( q^\pm \) tending to \( \pm q \), we regain the familiar formulae

\[
\begin{align*}
  r_{ss} &\rightarrow \frac{s_1 + s_2 e^{2i\Delta z}}{1 + s_1 s_2 e^{2i\Delta z}} & r_{sp} &\rightarrow 0 & r_{ps} &\rightarrow 0 & r_{pp} &\rightarrow \frac{p_1 + p_2 e^{2i\Delta z}}{1 + p_1 p_2 e^{2i\Delta z}} \\
\end{align*}
\]

(29)

where \( s_1 = (q_1 - q)/(q_1 + q) \), \( s_2 = (q - q_2)/(q + q_2) \), \( p_1 = (Q - Q_1)/(Q + Q_1) \) and \( p_2 = (Q_2 - Q)/(Q_2 + Q) \) are the s and p Fresnel reflection amplitudes at the boundary of the layer with media 1 and 2 (\( Q = q/\epsilon \) where \( \epsilon \) is the dielectric constant of the isotropic layer).

At normal incidence we regain the simple results ([2], (31)–(34))

\[
\begin{align*}
  r_o = \frac{\alpha^2 r_o + \beta^2 r_e}{\alpha^2 + \beta^2} & r_{pp} = \frac{\alpha^2 r_e + \beta^2 r_o}{\alpha^2 + \beta^2} & r_{sp} = r_{ps} = \frac{\alpha \beta}{\alpha^2 + \beta^2} (r_e - r_o) \\
\end{align*}
\]

(30)

where \( r_o \) and \( r_e \) are the reflection amplitudes of isotropic layers of indices \( n_o \) and \( n_o n_e/n_s \) (\( n_s \) was defined in equation (6)):

\[
\begin{align*}
  r_o = \frac{s_1^o + s_2^o e^{2i\kappa \Delta z}}{1 + s_1^o s_2^o e^{2i\kappa \Delta z}}, \ s_1^o = \frac{k_1 - k_o}{k_1 + k_o}, \ s_2^o = \frac{k_o - k_2}{k_o + k_2} \\
\end{align*}
\]

(31)

with \( r_e \) being obtained by replacing \( k_o = n_o \omega/\epsilon \) by \( k_e = (n_o n_e/n_s) \omega/c \). Note that the plane of incidence is not defined at normal incidence, except by a limiting process, which is the way the formulae (30) and (36) are obtained. When one is considering normal incidence only, it is better to work in terms of reflection amplitudes \( r \) and \( r' \), which give the reflection parallel and perpendicular to the incident polarisation, as was done in [2]. At normal incidence \( r \) and \( r' \) give all the reflection information, and similarly \( t \) and \( t' \) all the transmission information, in terms of the amplitudes \( r_o \) and \( r_s \), and \( t_o \) and \( t_s \).

The transmission amplitudes \( t_{ss} \) and \( t_{sp} \) can be obtained from \( t_s = L r_s \) in terms of \( r_{ss} \) and \( r_{sp} \):

\[
\begin{align*}
  t_{ss} &= L_{23} r_{sp} \cos \theta_1 + L_{22}(1 + r_{ss}) - L_{23} k_1 r_{sp} + L_{23} q_1(1 - r_{ss}) \\
  k_2 t_{sp} &= L_{31} r_{sp} \cos \theta_1 + L_{32}(1 + r_{ss}) - L_{33} k_1 r_{sp} + L_{34} q_1(1 - r_{ss}).
\end{align*}
\]

(32)

Similarly, \( t_{pp} \) and \( t_{ps} \) can be found from \( r_p = L r_p \):

\[
\begin{align*}
  k_2 t_{pp} &= L_{31}(1 + r_{pp}) \cos \theta_1 + L_{32} r_{ps} + L_{33} k_1(1 - r_{pp}) - L_{34} q_1 r_{ps} \\
  t_{ps} &= L_{21}(1 + r_{pp}) \cos \theta_1 + L_{22} r_{ps} + L_{23} k_1(1 - r_{pp}) - L_{24} q_1 r_{ps}.
\end{align*}
\]

(33)

As the layer thickness tends to zero, \( L_{ij} \rightarrow \delta_{ij} \) and we regain the companion formulae of (28):

\[
\begin{align*}
  t_{ss} &\rightarrow \frac{2q_1}{q_1 + q_2} & t_{sp} &\rightarrow 0 & t_{ps} &\rightarrow 0 & t_{pp} &\rightarrow \frac{n_1}{n_2} \frac{2Q_1}{Q_1 + Q_2}.
\end{align*}
\]

(34)
For an isotropic layer we set \( q_o \to q \), \( q_e \to q_e \), and obtain (compare (29))

\[
\begin{align*}
\tau_{ss} & \to \frac{(1 + s_1^2)(1 + s_2^2)e^{i\Delta_z}}{1 + s_1^2s_2^2e^{2i\Delta_z}} \quad \tau_{sp} \to 0 \quad \tau_{ps} \to 0 \\
\tau_{pp} & \to \frac{n_1}{n_2} \frac{(1 - p_1^2)(1 - p_2^2)e^{2i\Delta_z}}{1 + p_1p_2e^{2i\Delta_z}}
\end{align*}
\]

(35)

At normal incidence we find (compare equations (35) to (37) of [2])

\[
\begin{align*}
t_{ss} & = \frac{\alpha^2 t_o + \beta^2 t_e}{\alpha^2 + \beta^2} \\
t_{pp} & = \frac{\alpha^2 t_e + \beta^2 t_o}{\alpha^2 + \beta^2} \\
t_{ps} & = t_{sp} = \frac{\alpha \beta}{\alpha^2 + \beta^2} (t_e - t_o)
\end{align*}
\]

(36)

where \( t_o \) and \( t_e \) are normal incidence transmission amplitudes for isotropic layers of indices \( n_o \) and \( n_o n_e/n_i \), respectively:

\[
\begin{align*}
t_o & = \frac{(1 + s_1^2)(1 + s_0^2)e^{i\Delta_z}}{1 + s_1^2s_0^2e^{2i\Delta_z}} \\
t_e & = \frac{(1 + s_1^2)(1 + s_2^2)e^{i\Delta_z}}{1 + s_1^2s_2^2e^{2i\Delta_z}}
\end{align*}
\]

(37)

(see (31) for the definitions of \( k_o, k_e \), and of the reflection amplitudes \( s_1^2 \), etc).

In the following sections we will give detailed results for some special cases. Here we mention some especially simple results: \( r_{sp} = -r_{ps} \) when \( \alpha = 0 \) (optic axis perpendicular to the \( x \) axis), while \( r_{sp} = r_{ps} \) when \( \gamma = 0 \) (which obtains when the optic axis lies in the plane of the reflecting surface).

When the optic axis is perpendicular to the plane of incidence (\( \beta^2 = 1 \)), both \( \alpha \) and \( \gamma \) are zero and \( q_e^\pm = \pm (\varepsilon_o \omega^2/c^2 - K^2)^{1/2} \). The s polarization has \( E \) parallel to the optic axis and converts completely to the extraordinary mode. The p polarization has \( E \) perpendicular to the optic axis and converts completely to the ordinary mode. Thus \( r_{sp} = 0 = r_{ps} \); \( t_{sp} = 0 = t_{ps} \); \( r_{ss}, r_{pp} \) and \( t_{ss}, t_{pp} \) have the isotropic layer forms (29) and (35), with \( q = (\varepsilon_o \omega^2/c^2 - K^2)^{1/2} \) in the definitions of \( s_1 \) and \( s_2 \) and in the exponents of \( r_{ss} \) and \( t_{ss} \), while \( q = q_o \) and \( Q = q_o/v_0 \) in the definitions of \( p \) and \( p \) in and the exponents of \( r_{pp} \) and \( t_{pp} \).

4. Optic axis in the plane of incidence (\( \beta = 0 \))

When the optic axis lies in the plane of incidence (the \( zx \) plane) the ordinary electric field vector is perpendicular to the plane of incidence. Thus the incident s-polarized wave has electric field along the \( E_o \) direction, and we may expect the s polarization to convert fully to the ordinary mode, which implies that \( r_{sp} \) and \( t_{sp} \) are zero, and that \( r_{ss} \) is the same as that of the s reflection amplitude for an isotropic layer of index \( n_o \):

\[
r_{ss} = \frac{s_1^2 + s_2^2 Z_o}{1 + s_1 s_2 Z_o} \quad Z_o = \exp(2i q_o \Delta z)
\]

(38)

where \( s_1 \) and \( s_2 \) are the Fresnel s wave reflection amplitudes at the front and back faces of the crystal for the ordinary wave:

\[
s_1 = \frac{q_1 - q_0}{q_1 + q_0} \quad s_2 = \frac{q_0 - q_2}{q_0 + q_2}
\]

(39)
To find the other amplitudes we use the matrix $M$ of equation (22) simplified by setting $\beta = 0$ (this makes 8 of the 16 elements zero), and

$$v_\gamma q^\pm = \pm n_\gamma n_e q_\gamma - \gamma K \Delta \varepsilon$$  \hspace{1cm} (40)

where

$$q^2_\gamma = v_\gamma \omega^2/c^2 - K^2.$$  \hspace{1cm} (41)

This substitution produces common factors among the remaining non-zero elements, since now

$$v_\gamma (\alpha q^+ - \gamma K) = n_\gamma (\alpha q_\gamma n_\gamma - \gamma K n_e)$$

$$v_\gamma (\alpha q^- - \gamma K) = -n_\gamma (\alpha q_\gamma n_\gamma + \gamma K n_e)$$  \hspace{1cm} (42)

$$v_\gamma (\alpha q^2 - \gamma K q^+ q^-) = q_\gamma n_\gamma (\alpha q_\gamma n_\gamma - \gamma K n_e)$$

$$v_\gamma (\alpha q^2 - \gamma K q^+ q^-) = q_\gamma n_\gamma (\alpha q_\gamma n_\gamma + \gamma K n_e).$$

These relations enable us to simplify $r_{pp}$ to the form taken for an isotropic layer, namely

$$r_{pp} = \frac{p_1 + p_2 Z}{1 + p_1 p_2 Z} \quad Z = \exp(2i\gamma \Delta z)$$  \hspace{1cm} (43)

where $\tilde{q} = n_\gamma n_e q_\gamma/v_\gamma$ is the value of $\tilde{q}$ when $\beta = 0$, and

$$p_1 = \frac{Q - Q_1}{Q + Q_1} \quad p_2 = \frac{Q_2 - Q}{Q_2 + Q}$$  \hspace{1cm} (44)

have the same form as Fresnel p-wave reflection amplitudes. Here

$$Q_1 = q_1/v_1, \quad Q_2 = q_2/v_2 \quad Q = q_\gamma/n_\gamma n_e.$$  \hspace{1cm} (45)

For non-absorbing crystals $Z$ lies on the unit circle, and $r_{pp}$ can be zero if $Z = 1$ or $-1$. When $Z = 1$, $r_{pp}$ is zero if $p_1 + p_2 = 0$, which happens when $Q_2 = Q_1$. This equality holds at the substrate Brewster angle, $\theta_i = \tan^{-1}(n_2/n_1)$. When $Z = -1$, $r_{pp}$ is zero if $p_1 = p_2$, which happens when

$$Q^2 = Q_1 Q_2 \quad \text{or} \quad q^2_\gamma = \frac{q_1 q_2}{v_\gamma v_2}.$$  \hspace{1cm} (46)

This condition leads to a quadratic in $K^2$ (for given $\gamma$), or to a linear equation in $\gamma^2$ (for given angle of incidence). At normal incidence (46) is satisfied if

$$\gamma^2 = \frac{v_\gamma - n_1 n_2}{n_1 n_2}. \quad \frac{v_\gamma - n_1 n_2}{v_\gamma}$$  \hspace{1cm} (47)

The transmission amplitude $t_{ss}$ has the same form as for an isotropic layer:

$$t_{ss} = \frac{(1 + s_1)(1 + s_2)e^{i\gamma \Delta z}}{1 + s_1 s_2 e^{2i\gamma \Delta z}},$$  \hspace{1cm} (48)
(s_1 and s_2 are defined in (39)). The p to p transmission has a similar form, but with an additional phase factor:

\[ t_{pp} = e^{-i2\pi K z \Delta z/c} \frac{n_1 (1 - p_1)(1 - p_2)e^{i\Delta z}}{n_2 (1 + p_1 p_2 e^{2i\Delta z})} \]

where \( p_1 \) and \( p_2 \) are given in (44).

The remaining amplitudes, \( r_{pp} \) and \( t_{ps} \), are zero. Thus there is no cross-linking between s and p polarizations when the optic axis lies in the plane of incidence, either in reflection or in transmission.

When \( \beta = 0 \) the extraordinary electric field vector lies in the plane of incidence: for the forward wave we have

\[ E_c = N_c(e, q, 0, -[n_0 n_e K + x(q, \Delta z)]). \]

This is perpendicular to both the ordinary electric field, proportional to \((0, 1, 0)\), and to the extraordinary ray vector, which also lies in the plane of incidence:

\[ \text{ray vector} \sim (n_0 n_e K + x(q, \Delta z), 0, e, q). \]

We noted in section 2 (see equation (23)) that the determinant of the mode matrix \( M \) is zero if \( \beta = 0 \) and at the same time \( K^2 = \alpha^2 k_0^2 \). Then \( q_0^2 = \gamma^2 k_0^2 \), and the wavevector \((K, 0, q_0)\) is parallel to the optic axis \( c = (\alpha, 0, \gamma) \) if \( \alpha \) and \( \gamma \) have the same sign. (The wavevector of the internally reflected wave, \((K, 0, -q_0)\) is parallel to \( c \) if \( \alpha \) and \( \gamma \) have the opposite sign). From (42) and the fact that \( q_0 = \gamma(n_0/n_e)k_0 \) for this special case, we see that four more of the elements of \( M \) become zero, eight already being zero because they are proportional to \( \beta \). This leaves only four non-zero elements of \( M \), which now has two columns of zeros. Thus \( M \) cannot be inverted, and the matrix method fails.

This is related to the fact that when \( \beta = 0 \) and simultaneously \( \alpha q_0 = \pm \gamma K \) the ordinary electric field vector is indeterminate for the forward or backward propagating waves, respectively (see equation (7)). In a biaxial crystal, indeterminacy of one of the extraordinary electric fields leads to conical refraction [7, 8]. In this case \( K = \alpha^2 k_0^2 \) being considered here, however, it is the ordinary electric field vector which is indeterminate. Since the ray vector always coincides with the wavevector for the ordinary mode, there is no conical refraction in this case. The formulae given above for the reflection and transmission amplitudes remain valid when \( \beta = 0 \) and \( K^2 = \alpha^2 k_0^2 \).

5. Optical properties of a thin layer

A thin isotropic layer, that is one whose thickness is small compared to the wavelength, so that \( \omega \Delta z/c = 2\pi \Delta z/\lambda \ll 1 \), has reflection and transmission amplitudes which are given by (28) and (34), plus power series in \( \omega \Delta z/c \). Of most interest is the ratio \( r_p/r_s \), which can be obtained by ellipsometry, and differs from the zero thickness value

\[
(r_p/r_s)_0 = \left( \frac{Q_2 - Q_1}{Q_2 + Q_1} \right) \left( \frac{q_1 + q_2}{q_1 - q_2} \right) = \frac{(q_1 + q_2)^2}{\varepsilon_1 \varepsilon_2 (Q_1 + Q_2)^2} \left[ 1 - \left( \frac{cK}{\omega} \right)^2 \left( \frac{1}{\varepsilon_1} + \frac{1}{\varepsilon_2} \right) \right]
\]

\[
= \frac{q_1 q_2 - K^2}{q_1 q_2 + K^2}
\]

(52)
by a term linear in $\Delta z$, namely
\[
(r_p/r_s)_1 = -2iQ_1(cK/\omega)^2 \frac{(q_1 + q_2)^2}{\varepsilon_1 \varepsilon_2 (Q_1 + Q_2)^2} \frac{\varepsilon_1 - \varepsilon}{\varepsilon_1 - \varepsilon_2} \Delta z.
\] (53)

For non-uniform isotropic layers, where $\varepsilon = \varepsilon(z)$, the same form holds, with
\[\varepsilon^{-1}(\varepsilon_1 - \varepsilon)(\varepsilon - \varepsilon_2)\Delta z \text{ replaced by}
\]
\[I_1 = \int dz \frac{(\varepsilon_1 - \varepsilon)(\varepsilon - \varepsilon_2)}{\varepsilon}.
\] (54)

This well known result has been extended to uniaxial non-uniform layers in which the optic axis is along the surface normal ([9], section 7-2). In this case, which corresponds to $\gamma^2 = 1$ in our notation, $r_{sp}$ and $r_{ps}$ remain zero, and
\[
(r_{pp}/r_{ss})_1 = -2iQ_1(cK/\omega)^2 \frac{(q_1 + q_2)^2}{\varepsilon_1 \varepsilon_2 (Q_1 + Q_2)^2} \frac{I_1}{\varepsilon_1 - \varepsilon_2}
\] (55)

where now
\[
I_1 = \int dz \left\{ \varepsilon_1 + \varepsilon_2 - \frac{\varepsilon_1 \varepsilon_2}{\varepsilon} - \varepsilon_o \right\}.
\] (56)

The reader may have noticed the factor $(\varepsilon_1 - \varepsilon_2)^{-1}$ in (53) and (55). For unsupported films this appears to lead to a divergence in $r_p/r_s$ or $r_{pp}/r_{ss}$. In fact there is no divergence when $\varepsilon_1 = \varepsilon_2$, but there is an enhancement of the ellipsometric ratio when $\varepsilon_2$ is close to $\varepsilon_1$. The isotropic case is discussed in [10].

For general anisotropic layers, the reflection is characterized by four reflection amplitudes $r_{ss}$, $r_{sp}$, $r_{pp}$, $r_{ps}$, and reflection ellipsometry measures the quantities
\[
(r_{pp} + r_{sp} \tan P)/(r_{ps} + r_{ss} \tan P) \quad \text{or} \quad (r_{pp} + r_{ps} \tan A)/(r_{sp} + r_{ss} \tan A)
\] (57)

depending on whether the compensator or modulator is placed between the sample and the analyser, or between the polarizer and the sample [11]. ($P$ and $A$ are angles defining the orientation of the polarizer and analyser.)

We will find the first-order corrections to (28) and (34). The relative phases of the reflection amplitudes are experimentally measurable, but the overall common phase depends on the choice of origin. Here we choose the front face of the anisotropic layer to be at $z = 0$ for convenience. This fixes the common phase, and we can write down expressions for the separate reflection coefficients, not just for their ratios.

When $\Delta z \to 0$, the matrix $P$ tends to the identity matrix $I = \text{diag}(1, 1, 1, 1)$. To first order in $\omega \Delta z/c$,
\[
P = I + i\Delta z \text{diag}(q_o, -q_o, q_e^+, q_e^-) \equiv I + D
\] (58)

Thus the layer matrix $L = MPM^{-1}$ becomes, to this order,
\[
L = I + MDM^{-1}.
\] (59)

The identity part of $L$ leads to the bare-substrate reflection amplitudes given in (28). The first-order corrections come from the $MDM^{-1}$ part of $L$. In the absence of absorption in the crystal or the substrate (we always assume lack of absorption in the medium of incidence), the first-order terms are pure imaginary, as can be seen from (58) and (59).
Particularly simple results are obtained when \( \beta = 0 \) (optic axis in the plane of incidence), and we will give these first. We find

\[
    r_{ss}(\beta = 0) = \frac{q_1 - q_2}{q_1 + q_2} + 2i\Delta z \frac{q_1(q_0^2 - q_2^2)}{(q_1 + q_2)^2} + O(\Delta z)^2
\]

(60)

\[
    r_{pp}(\beta = 0) = \frac{Q_2 - Q_1}{Q_2 + Q_1} + 2i\Delta z \frac{Q_1(q_0^2 - Q_2^2)\epsilon_o \epsilon_e}{(Q_1 + Q_2)^2\epsilon_e} + O(\Delta z)^2
\]

(61)

where \( Q_1, Q_2 \) and \( Q \) are defined in (45). The results (60) and (61) may also be obtained from (38) and (43). The cross-reflection amplitudes \( r_{sp} \) and \( r_{ps} \) are identically zero when the optic axis is in the plane of incidence. The transmission coefficients are

\[
    t_{ss}(\beta = 0) = \frac{2q_1}{q_1 + q_2} + 2i\Delta z \frac{q_1(q_1q_2 + q_0^2)}{(q_1 + q_2)^2} + O(\Delta z)^2
\]

(62)

\[
    t_{pp}(\beta = 0) = e^{-i\gamma/k\Delta z c/e} \frac{n_1}{n_2} \left\{ \frac{2Q_1}{Q_1 + Q_2} + 2i\Delta z \frac{\epsilon_e \epsilon_o Q_1(Q_1 + Q_2)}{\epsilon_e(Q_1 + Q_2)^2} \right\} + O(\Delta z)^2.
\]

(63)

We now give the reflection amplitudes to first order in \( \Delta z \), in the general case. These are

\[
    r_{ss} = \frac{q_1 - q_2}{q_1 + q_2} + 2i\Delta z \frac{(q_0^2 - q_2^2 + \beta^2 k_0^2 \Delta \epsilon/\epsilon_e)}{(q_1 + q_2)^2} + O(\Delta z)^2
\]

(64)

\[
    r_{sp} = \frac{2i\beta q_1 \Delta z(xQ_1k_0^2 + \gamma K \omega^2/c^2) \Delta \epsilon}{k_1(q_1 + q_2)(Q_1 + Q_2)\epsilon_e} + O(\Delta z)^2
\]

(65)

\[
    r_{ps} = \frac{-2i\beta q_1 \Delta z(xQ_2k_0^2 - \gamma K \omega^2/c^2) \Delta \epsilon}{k_1(q_1 + q_2)(Q_1 + Q_2)\epsilon_e} + O(\Delta z)^2
\]

(66)

\[
    r_{pp} = \frac{Q_2 - Q_1}{Q_2 + Q_1} + 2i\Delta z \frac{(Q_0^2 - Q_2^2)\epsilon_e - \beta^2 Q_2^2 \Delta \epsilon}{(Q_1 + Q_2)^2\epsilon_e} + O(\Delta z)^2.
\]

(67)

We note that when \( \beta = 0 \) the \( r_{ss} \) and \( r_{pp} \) amplitudes reduce to (60) and (61), while \( r_{sp} \) and \( r_{ps} \) become zero. We stated at the end of section 3 that \( r_{sp} = -r_{ps} \) when \( x = 0 \), and \( r_{sp} = r_{ps} \) when \( \gamma = 0 \). This is verified to first order in the layer thickness by (65) and (66).

The transmission amplitudes to first order in the layer thickness are given by

\[
    t_{ss} = \frac{2q_1}{q_1 + q_2} \left\{ 1 + i\Delta z \frac{[q_1q_2\epsilon_e + q_0^2\epsilon_0 + (\beta^2 k_0^2 + \gamma^2 q_0^2)\Delta \epsilon]}{(q_1 + q_2)^2} \right\} + O(\Delta z)^2
\]

(68)

\[
    t_{sp} = \frac{2i\beta q_1 \Delta z(xQ_1k_0^2 - \gamma K \omega^2/c^2) \Delta \epsilon}{k_2(q_1 + q_2)(Q_1 + Q_2)\epsilon_e} + O(\Delta z)^2
\]

(69)

\[
    t_{ps} = \frac{-2i\beta q_1 \Delta z(xQ_2k_0^2 + \gamma K \omega^2/c^2) \Delta \epsilon}{k_1(q_1 + q_2)(Q_1 + Q_2)\epsilon_e} + O(\Delta z)^2
\]

(70)

\[
    t_{pp} = \frac{n_1}{n_2} \frac{2Q_1}{Q_1 + Q_2} \left\{ 1 + i\Delta z \frac{\epsilon_0(xQ_2 - xQ_2 K(Q_1 + Q_2)\Delta \epsilon + q_0^2)}{(Q_1 + Q_2)\epsilon_e} \right\} + O(\Delta z)^2.
\]

(71)
In all cases the reflection and transmission amplitudes (to first order in $\Delta z$) have the form $u + iv$, where $v$ is proportional to the film thickness, and $u$ and $v$ are real if there is no absorption. Thus the reflectances and transmittances are all of the form $u^2 + v^2$, with $u^2$ the zero-thickness value, and $v^2$ the correction term which is second order in $\Delta z$.

The formulae of this section are intended for use in the ellipsometry of thin anisotropic layers; another application (in section 7) is to weakly rough surfaces.

6. Weak anisotropy

The isotropic layer has reflection and transmission amplitudes given by (29) and (35). When the uniaxial layer is weakly anisotropic, we expect that these relations are corrected by terms proportional to $\Delta \varepsilon = \varepsilon_e - \varepsilon_o$. This section will give these first-order terms. From (6) we find that

$$
q_e^+ = q_o + \frac{[k_o^2 - (\varepsilon_k + \varepsilon q_o)^2]\Delta \varepsilon}{2\varepsilon_q q_o} + O(\Delta \varepsilon)^2
$$

$$
q_e^- = -q_o - \frac{[k_o^2 - (\varepsilon_k - \varepsilon q_o)^2]\Delta \varepsilon}{2\varepsilon_q q_o} + O(\Delta \varepsilon)^2.
$$

Let us denote by $\Delta q$ the differences $q_e^+ - q_o$ and $q_e^- + q_o$. There are two kinds of correction terms: those of order $\Delta q / q_o$, and those of order $\Delta q \Delta z$. It is clear that the latter (which gives the phase shift between the ordinary and extraordinary waves on traversing the layer) can be large for thick films, even if the anisotropy is weak. We will separate the two types of terms in what follows.

The isotropic values of $r_{ss}$ and $r_{sp}$ are given by (29). The first-order corrections are

$$
r_{ss}^{(1)} = \frac{\beta^2 q_1 [(1 + s_2^2 Z)(Z - 1) + 4i s_2 Z q_o \Delta z](\omega/c)^2 \Delta \varepsilon}{q(q_1 + q)^2 [1 + s_1 s_2 Z]^2}
$$

$$
r_{sp}^{(1)} = \frac{q [f_f f_- (1 + p_2^2 Z)(Z - 1) + 4i p_2 (\varepsilon q^2 + \varepsilon K^2) Z q_o \Delta z] \Delta \varepsilon}{q e_k^+ k_o^+ (Q_1 + Q)^2 [1 + p_1 p_2 Z]^2}
$$

where $Z = \exp(2i q \Delta z)$, $q_1$, $s_1$, $s_2$, $p_1$ and $p_2$ are the Fresnel amplitudes defined below (29), and we have dropped the suffix o on $q_o$, as we did in (29). (We shall also drop the o suffix on $\kappa_o$ and $\varepsilon_o$ in the remainder of this section.) The functions $f_+$ and $f_-$ are defined by

$$
f_{\pm} = \varepsilon q \pm \gamma K,
$$

and we make use of the identity

$$(\varepsilon q \pm \gamma K)^2 + \beta^2 k^2 = k^2 - (\varepsilon K \pm \gamma q)^2.
$$

The first-order terms in $\Delta \varepsilon$ for the cross-coupling reflection amplitudes are more complicated:

$$
r_{sp}^{(1)} = \frac{\beta k q_1 [Z - 1] e_2 q + q_2)(Q + Q_2)[f_+ + s_2 p_2 f_+ Z] + 4i(e - e_2)(\varepsilon q^2 q_2 + \gamma K^2) Z q_o \Delta z] \Delta \varepsilon}{q(q_1 + q)(q + q_2)(Q_1 + Q)(Q_2) e_1 e_2 e^2 [1 + s_1 s_2 Z][1 + p_1 p_2 Z]}
$$

(77)
\[ r_{ps}^{(1)} = \frac{\beta k_1 q_1 (Z - 1) \varepsilon_2 (q + q_2) (Q + Q_2) [f_s + s_1 p_2 f_s Z] + 4i(\varepsilon - \varepsilon_2) K^2 \Delta z Z q_2 \Delta \varepsilon}{q (q_1 + q + q_2) (Q + Q_2) [1 + s_1 s_2 Z] [1 + p_1 p_2 Z]} \]

(78)

We see that both the s to p and the p to s reflection amplitudes are zero to first order in the anisotropy when the optic axis lies in the plane of incidence (i.e. when \( \beta = 0 \)).

The isotropic values of the transmission coefficients are given by (35). The first-order corrections are

\[ t_{ss}^{(1)} = -2 \beta^2 q_1 e^{i \Delta z} (Z - 1) \varepsilon_2 (q + q_2 - \varepsilon_2) - i(q_1 + q) (Q + Q_2) [1 - s_1 s_2 Z] q_2 \Delta \varepsilon (\varepsilon_1 / \varepsilon_2) (Q + Q_2) \]

\[ q (q_1 + q + q_2) [1 + s_1 s_2 Z] [1 + p_1 p_2 Z] \]

(79)

\[ t_{pp}^{(1)} = \frac{2 q_1 e^{i \Delta z} (Z - 1) (Q_1 Q_2 - Q^2) + i(q_1 + q) (Q + Q_2) [f_s - p_1 p_2 f_s Z] q_2 \Delta \varepsilon}{n_1 n_2 \varepsilon_2 q (q_1 + q + q_2) (Q + Q_2) [1 + p_1 p_2 Z]} \]

(80)

\[ 2 \beta q_1 k_2 e^{i \Delta z} (Z - 1) [x q^2 (Q_1 - Q_2) + \gamma K (Q_1 q_2 - Q q)] \]

\[ + i(q_1 + q) (Q + Q_2) [f_s - p_1 s_2 f_s Z] q_2 \Delta \varepsilon \]

\[ q (q_1 + q + q_2) (Q + Q_2) [1 + s_1 s_2 Z] [1 + p_1 p_2 Z] \]

(81)

\[ t_{sp}^{(1)} = \frac{2 \beta q_1 k_1 e^{i \Delta z} (Z - 1) [x q^2 (Q_2 - Q_1) + \gamma K (Q_1 q_2 - Q q)]}{e \varepsilon_2 q (q_1 + q) (Q + Q_2) [1 + s_1 s_2 Z] [1 + p_1 p_2 Z]} \]

(82)

where

\[ Q'_1 = q_1 / \varepsilon \quad Q'_2 = q_2 / \varepsilon. \]

(83)

We note that the first-order parts of \( t_{ss}^{(1)} \), \( t_{sp}^{(1)} \) and \( t_{ps}^{(1)} \) are all zero when \( \beta = 0 \). Thus only \( t_{pp}^{(1)} \) and \( t_{pp}^{(1)} \) are non-zero when the optic axis lies in the plane of incidence.

In the orthogonal configuration, when \( \beta^2 = 1 \) and the optic axis is perpendicular to the plane of incidence, only \( t_{ps}^{(1)} \) and \( t_{ps}^{(1)} \) are non-zero (the incident s polarization converts to the extraordinary wave, since \( E_\varepsilon \) is along the optic axis).

7. Discussion and applications

We have seen that analytic results may be given for the reflection and transmission properties of a uniaxial layer. We have not assumed that the layer or the substrate is non-absorbing: if either is absorbing the formulae given here remain valid, with complex values of \( \varepsilon_\alpha \), \( \varepsilon_\varepsilon \) and \( \varepsilon_\varepsilon \), as appropriate, and thus with complex values of \( q_\alpha \), \( q_\varepsilon \) and \( q_\varepsilon \).
However, the formulae apply to absorbing uniaxial crystals only when the principal axes of the real and imaginary parts of $\varepsilon$ coincide.

The matrix elements given here are for arbitrary crystal orientation, angle of incidence, and degree of anisotropy. Explicit formulae for the reflection and transmission amplitudes have, however, been given only for special cases, because the general formulae are not compact enough to be useful. One important special case is missing: that of nearly normal incidence. The optical properties of a normally illuminated uniaxial layer are beautifully simple (see [2] and formulae (30) and (36) of this paper), but I have not so far been able to reduce the nearly-normal-incidence formulae to usable size.

The utility of analytic formulae lies in providing immediate answers to physical questions. We give two examples. Consider first the optical properties of an antireflection coating which is anisotropic. Equations (30) and (31) give the reflection amplitudes at normal incidence in terms of the basic amplitudes $r_o$ and $r_e$. For an isotropic layer we can in general make one zero and the other small; only when the optic axis is coincident with the normal to the layer do we have $r_o = r_e$, for then $k_o = k_e$. Suppose for example that $r_o$ is zero, and $r_e$ is not. Then zero reflection will obtain only when the incident polarization coincides with the $E_o$ direction in the layer.

As the second example, consider the reflection properties of rough surfaces. For liquid surfaces it has been demonstrated that the effect of a small amount of roughness (on a scale small compared to the wavelength), is similar to that of an adsorbed layer [12, 13]. Recent interest lies in the shift of the Brewster angle as a function of the roughness parameters [14–16]. The calculated shift is toward smaller angles of incidence, and increases in magnitude linearly with the mean square of the displacement of the surface from flatness. For an adsorbed layer of thickness $\Delta z$ the shift in the Brewster angle is also proportional to $(\Delta z)^2$. One must make clear which definition of the Brewster angle one uses. The ellipsometric definition of $\theta_B$ is that angle at which $\text{Re}(r_p/r_e) = 0$, and the general formula for a film of arbitrary dielectric function profile is given in [9], equation (3.53). In [14–15] the definition is the location of $\text{Re}(r_p) = 0$; this is not experimentally defined until a reference phase has been specified, the natural one being to take $r_p$ to be real for a flat surface. With this definition one finds, to second order in the thickness $\Delta z$ of an isotropic homogeneous layer,

$$\Delta \theta_B = \frac{\varepsilon^2 - \varepsilon(\varepsilon_1 + \varepsilon_2) + \varepsilon_1 \varepsilon_2}[\varepsilon^2 + \varepsilon(\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2] \varepsilon^{1/2} \varepsilon^{3/2} \left\{ \frac{2 \varepsilon^2 (\varepsilon_1^2 - \varepsilon_2^2)}{[\varepsilon(\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2]} \right\} (q_B \Delta z)^2 \tag{84}$$

where $\varepsilon$ is the dielectric constant of the layer between media 1 and 2, and $q_B$ is the value of $q = (\omega \varepsilon^2 / c^2 - K^2)^{1/2}$ at the zero-thickness Brewster angle given by $\tan^2 \theta_B = \varepsilon_2 / \varepsilon_1$. Assuming $\varepsilon_2 > \varepsilon_1$ and that $\varepsilon$ lies between $\varepsilon_1$ and $\varepsilon_2$, $\Delta \theta_B$ will be negative.

We may expect that a homogeneous layer mimicking the behaviour of a slightly rough surface would be anisotropic; for random roughness the optic axis must coincide with the surface normal, by symmetry. (This is not the case for some special models; see for example figure 2 of [15].) When the optic axis is normal to the surface, we can obtain $r_{pp}$ from the $\beta = 0$ formula (43) by setting $\gamma^2 = 1$. We find

$$\Delta \theta_B = \frac{\varepsilon_0 \varepsilon_2 - \varepsilon_0(\varepsilon_1 + \varepsilon_2) + \varepsilon_1 \varepsilon_2}{\varepsilon_0 \varepsilon_2 + \varepsilon_0(\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2} \varepsilon^{1/2} \varepsilon^{3/2} \left\{ \frac{2 \varepsilon_0 \varepsilon_2 (\varepsilon_1^2 - \varepsilon_2^2)}{[\varepsilon(\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2]} \right\} (q_B \Delta z)^2 \tag{85}$$
where now $q_B$ is the value of $(\varepsilon_a \omega^2/c^2 - K^2)^{1/2}$ at the Brewster angle. The sign of $\Delta \theta_B$ now depends on how different $\varepsilon_o$ and $\varepsilon_e$ are. (When they are the same, (84) reduces to (85) and $\Delta \theta_B$ will be negative under the conditions stated in the discussion following (84).) Thus a positive $\Delta \theta_B$ would indicate strong anisotropy in the effective layer representing the roughness. We note that (84) and (85) predict an enhancement of the shift in Brewster's angle when $\varepsilon_1 \approx \varepsilon_3$. This index-matching enhancement occurs also for the ellipsometric ratio $r_p/r_s$ [17].

For a slightly rough surface, as for a homogeneous thin layer on a transparent substrate, the p to p reflectivity is no longer zero at its minimum. The minimum reflectivity is proportional to $(\Delta z)^2$, where $\Delta z$ represents the thickness of the adsorbed layer, or the root-mean-square deviation from flatness. The magnitude of $R_{pp}$ at the minimum for a uniaxial layer with its optic axis normal to its surface may be obtained from (61) or (67) by setting $\gamma^2 = 1$ and taking the absolute square of the first-order term (the zero- and second-order terms add to zero at the (shifted) Brewster angle, as defined above). We find, to lowest order in $\Delta z$,

$$R_{pp}(\text{min}) = \frac{(\varepsilon_o \varepsilon_1 + \varepsilon_2 - \varepsilon_2 \varepsilon_1)^2}{4 \varepsilon_2^2 (\varepsilon_1 + \varepsilon_2)} \left( \frac{\omega \Delta z}{c} \right)^2.$$

Note that, on using the definition (56) for a homogeneous layer, this may be written as

$$R_{pp}(\text{min}) = \frac{(I_1 \omega/c)^2}{4(\varepsilon_1 + \varepsilon_2)}.$$

Thus the minimum p to p reflectivity is proportional to the square of

$$I_1 = \left( \varepsilon_1 + \varepsilon_2 - \frac{\varepsilon_1 \varepsilon_2}{\varepsilon_e} \right) \Delta z.$$

while, from (55), the ellipsometric ratio is proportional to $I_1$. Ellipsometry and minimum intensity measurement thus determine $I_1$ and its square. The Brewster angle shift gives additional information, since, from (85), $\Delta \theta_B$ contains factors additional to $I_1$:

$$\Delta \theta_B = - \frac{I_1 \left[ \varepsilon_o \varepsilon_1 + \varepsilon_e (\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2 \right] \varepsilon_e^1/2 \varepsilon_2^3/2 q_B^3 \Delta z}{2 \varepsilon_o (\varepsilon_2^2 - \varepsilon_1^2) [\varepsilon_e (\varepsilon_1 + \varepsilon_2) - \varepsilon_1 \varepsilon_2]}.$$

Acknowledgments

The paper arose from a suggestion by Dr John McGilp, and the applications in the last section were suggested by Dr Jean-Jacques Greffet. The author is grateful for these suggestions, and also for those supplied by a referee.

References

Optical properties of a uniaxial layer