Light in periodically stratified media

John Lekner

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

Received March 11, 1994; revised manuscript received June 1, 1994; accepted June 2, 1994

A pedagogical presentation of the propagation of electromagnetic waves in stratified media is given. The usual 2×2 matrix analysis is simplified and generalized. A Bloch factor eigenvalue equation is obtained, valid for all periodic stratifications. The implications of the band structure of an infinite periodic structure for reflection by a finite structure are demonstrated. Some features of the reflectivity are shown to be universal. In the long-wave limit, a periodic stratification with an arbitrary dielectric function profile $\epsilon(z)$ within a unit cell is shown to be equivalent to a homogeneous anisotropic medium, with ordinary and extraordinary dielectric constants given by $\epsilon_o = \langle\epsilon\rangle$ and $\epsilon_e^{-1} = \langle\epsilon^{-1}\rangle$.

1. INTRODUCTION

Electron wave functions in crystals are modified by interaction of the electrons with the periodic ionic lattice to such an extent that band gaps appear in the spectrum of allowed states. This became clear in the early days of quantum mechanics.[1] The history of wave propagation in periodic structures extends back to Newton, who considered elastic waves on a one-dimensional lattice of masses connected by springs as a model for sound.[2] Rayleigh[3,4] recognized the possibility of what are now known as stop bands or band gaps for waves in periodic structures, particularly in relation to the high reflection (at certain wavelengths and angles of incidence) by periodically stratified media.

The modern optics of stratifications was advanced by Abelès[5]; of special utility are his application of matrices to wave propagation and his theorem that the Nth power of a unimodular (one with unit determinant) 2×2 matrix is given by

$$\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}^N = \begin{bmatrix} m_{11}S_N - S_{N-1} & m_{12}S_N \\ m_{21}S_N & m_{22}S_N - S_{N-1} \end{bmatrix}, \tag{1}$$

where

$$S_N = \frac{\sin(N\phi)}{\sin\phi}, \qquad \cos\phi = \tfrac{1}{2}(m_{11} + m_{22}). \tag{2}$$

One can easily prove this result by induction, on using $m_{11}m_{22} - m_{12}m_{21} = 1$ and the identity (or recurrence relation)

$$2(\cos\phi)S_N - S_{N-1} = S_{N+1}. \tag{3}$$

The matrices used by Abelès link electric- and magnetic-field components at successive layers of the stratification. For nonabsorbing media these matrices are complex, with real diagonal elements and imaginary off-diagonal elements. Matrices that link fields and their derivatives (for example, E and dE/dz for the electromagnetic s or TE wave) are entirely real for nonabsorbing media.[6] This is both simpler and four times faster in numerical work [the matrix product $AB = (A_r + iA_i)(B_r + iB_i)$ requires the evaluation of four products if A and B are complex]. In this paper we shall both simplify and generalize the existing theory of light propagation in periodically stratified media. An expression is given for the matrix of a layer with continuous but otherwise arbitrary dielectric function variation. The eigenvalue equation for the Bloch factor in a periodic system is shown to be determined by the trace of the matrix of a unit cell. When the wavelength is long compared with the period of the stratification, the periodic structure is equivalent to a uniaxial homogeneous medium, with the ordinary dielectric constant equal to the average of the dielectric function and the extraordinary dielectric constant equal to the reciprocal of the average of the reciprocal of the dielectric function.

2. ELECTROMAGNETIC WAVES IN STRATIFIED MEDIA

We consider plane electromagnetic waves incident from a medium of index n_1 onto a nonmagnetic planar stratification, whose optical properties are contained in the dielectric function $\epsilon(z) = n^2(z)$[$n(z)$ is the local value of the refractive index]. For isotropic media, with scalar rather than tensor dielectric function, any plane wave can be written as a superposition of an s (or TE) wave and a p (or TM) wave. The s wave has its electric vector perpendicular to the plane of incidence, and the p wave has its electric vector in the plane of incidence (and its magnetic vector perpendicular to the plane of incidence; hence its designation as a TM, or transverse magnetic, wave). We assume that the medium is stratified in the z direction [so that $\epsilon = \epsilon(z)$], and I further take the plane of incidence to be the (z, x) plane. Then the s wave electric-field vector $\mathbf{E} = (0, E_y, 0)$, and the p wave has magnetic-field vector $\mathbf{B} = (0, B_y, 0)$. It follows directly from the Maxwell curl equations that, for monochromatic waves of angular frequency ω,

$$E_y(z, x, t) = \exp[i(Kx - \omega t)]E(z), \tag{4}$$

$$B_y(z, x, t) = \exp[i(Kx - \omega t)]B(z), \tag{5}$$

where K (the x component of the wave vector) is a separation-of-variables constant whose existence derives

0740-3232/94/112892-08$06.00

from the planar nature of the stratification and whose constancy implies Snell's law. The functions $E(z)$ and $B(z)$ satisfy the ordinary differential equations (see, for example, Sections 1-1 and 1-2 of Ref. 6)

$$\frac{\mathrm{d}^2E}{\mathrm{d}z^2} + q^2E = 0, \qquad \epsilon\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\epsilon}\frac{\mathrm{d}B}{\mathrm{d}z}\right) + q^2B = 0\,, \tag{6}$$

where $q(z)$ is the local value of the normal component of the wave vector, given by

$$q^2(z) = \epsilon(z)\omega^2/c^2 - K^2\,. \tag{7}$$

If θ_1 is the angle of incidence and θ_2 is the angle between the wave vector and the normal in the homogeneous substrate of index n_2, then

$$K = n_1(\omega/c)\sin\,\theta_1 = n_2(\omega/c)\sin\,\theta_2\,, \tag{8}$$

$$q_1 = n_1(\omega/c)\cos\,\theta_1, \qquad q_2 = n_2(\omega/c)\cos\,\theta_2\,. \tag{9}$$

It follows from Eqs. (6) that $\mathrm{d}E/\mathrm{d}z$ and $\epsilon^{-1}\mathrm{d}B/\mathrm{d}z$ are continuous at any discontinuity in $\epsilon(z)$ (otherwise delta-function terms would arise in the second derivatives). *A fortiori*, E and B are continuous at any discontinuity in $\epsilon(z)$.

Consider a stratification extending from $z = a$ to $z = b$, bounded by homogeneous media of indices n_1 and n_2, and suppose at first that $\epsilon(z)$ is continuous for $a < z < b$. For the s wave let $F(z)$ and $G(z)$ be two linearly independent solutions of the second-order differential equation $\mathrm{d}^2E/\mathrm{d}z^2 + q^2E = 0$. Then $E(z)$ may be written as a linear superposition of F and G:

$$E(z) = fF(z) + gG(z)\,, \tag{10}$$

where f and g are constants. We will use a layer matrix $M = \{m_{ij}\}$ that links fields and their derivatives; in the s-wave case it is defined by

$$\begin{pmatrix} E_b \\ E_b' \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{pmatrix} E_a \\ E_a' \end{pmatrix}, \tag{11}$$

where E_a and E_a' represent $E(a+)$ and the derivative of $E(z)$ at $z = a+$, and similarly E_b and E_b' stand for $E(b-)$ and the derivative of E at $z = b-$. From Eq. (10) and its derivative we see that

$$\begin{pmatrix} E_a \\ E_a' \end{pmatrix} = \begin{bmatrix} F_a & G_a \\ F_a' & G_a' \end{bmatrix} \begin{pmatrix} f \\ g \end{pmatrix} \equiv A\begin{pmatrix} f \\ g \end{pmatrix}, \tag{12}$$

$$\begin{pmatrix} E_b \\ E_b' \end{pmatrix} = \begin{bmatrix} F_b & G_b \\ F_b' & G_b' \end{bmatrix} \begin{pmatrix} f \\ g \end{pmatrix} \equiv B\begin{pmatrix} f \\ g \end{pmatrix}. \tag{13}$$

Substitution of Eqs. (12) and (13) into Eq. (11) shows that the layer matrix can be expressed in terms of the fundamental field and derivative values at the boundaries of the layer:

$$M = BA^{-1} = W^{-1}\begin{bmatrix} -(F', G) & (F, G) \\ -(F', G') & (F, G') \end{bmatrix}. \tag{14}$$

Here W is the (constant) Wronskian of the two basic solutions F and G,

$$W = FG' - F'G, \qquad W' = 0\,, \tag{15}$$

and

$$\begin{aligned} (F, G) &\equiv F_aG_b - G_aF_b, & (F, G') &\equiv F_aG_b' - G_aF_b', \\ (F', G) &\equiv F_a'G_b - G_a'F_b, & (F', G') &\equiv F_a'G_b' - G_a'F_b'. \end{aligned} \tag{16}$$

The layer matrix M is unimodular: from the identity (2.31) of Ref. 6,

$$\det M = W^{-2}[(F, G)(F', G') - (F, G')(F', G)] = 1\,. \tag{17}$$

An important example is that of a homogeneous layer, for which $\epsilon(z)$ and $q(z)$ are constant. We can then take $F = \cos(qz)$ and $G = \sin(qz)$, for which $W = q$ and

$$\begin{aligned} (F, G) &= \sin\,\delta, & (F, G') &= q\,\cos\,\delta, \\ (F', G) &= -q\,\cos\,\delta, & (F', G') &= q^2\,\sin\,\delta\,, \end{aligned} \tag{18}$$

where $\delta = q(b - a)$ is the phase increment across the layer. The layer matrix in this case is

$$M = \begin{bmatrix} \cos\,\delta & q^{-1}\sin\,\delta \\ -q\,\sin\,\delta & \cos\,\delta \end{bmatrix} \tag{19}$$

[the same matrix is obtained if we choose $F = \exp(iqz)$ and $G = \exp(-iqz)$].

The matrix elements m_{ij} determine the reflection and transmission amplitudes r_s and t_s of the $n_1|n(z)|n_2$ structure; we have, from the definition of M,

$$\begin{pmatrix} t_s\exp(iq_2b) \\ iq_2t_s\exp(iq_2b) \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \times \begin{pmatrix} \exp(iq_1a) + r_s\exp(-iq_1a) \\ iq_1[\exp(iq_1a) - r_s\exp(-iq_1a)] \end{pmatrix} \tag{20}$$

when $\exp(iq_1z) + r_s\exp(-iq_1z)$ and $t_s\exp(iq_2z)$ are the forms of $E(z)$ in the medium of incidence and in the substrate. It follows from Eq. (20) that (see Section 12-2 of Ref. 6)

$$r_s = \exp(2iq_1a)\,\frac{q_1q_2m_{12} + m_{21} + iq_1m_{22} - iq_2m_{11}}{q_1q_2m_{12} - m_{21} + iq_1m_{22} - iq_2m_{11}}\,, \tag{21}$$

$$t_s = \exp[i(q_1a - q_2b)] \times \frac{2iq_1}{q_1q_2m_{12} - m_{21} + iq_1m_{22} - iq_2m_{11}}\,. \tag{22}$$

The form of the differential equation for $E(z)$ is the same as that of the Schrödinger equation for a particle of mass m in a potential $V(z)$, where

$$\epsilon(z)\omega^2/c^2 \leftrightarrow \frac{2m}{\hbar^2}[E - V(z)]\,. \tag{23}$$

Thus the results derived for the electromagnetic s wave apply also to quantum particle waves in a z-stratified medium, as is well known.

The p-wave layer matrix is defined to link the quantities $B(z)$ and $\epsilon^{-1}\mathrm{d}B/\mathrm{d}z$, which are continuous at discontinuities of ϵ. Let B_a and B_b stand for $B(a+)$ and $B(b-)$ and $\tilde{B}_a$ and $\tilde{B}_b$ represent the values of $\epsilon^{-1}\mathrm{d}B/\mathrm{d}z$ at $z = a+$ and $z = b-$. Then

$$\begin{pmatrix} B_b \\ \tilde{B}_b \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{pmatrix} B_a \\ \tilde{B}_a \end{pmatrix}. \qquad (24)$$

We express $B(z)$ as a linear combination of two independent solutions of Eq. (6), say $C(z)$ and $D(z)$. Then, by the arguments used above, with F, F' and G, G' replaced by C, $\tilde{C}$ and D, $\tilde{D}$,

$$M = U^{-1} \begin{bmatrix} -(\tilde{C}, D) & (C, D) \\ -(\tilde{C}, \tilde{D}) & (C, \tilde{D}) \end{bmatrix}, \qquad (25)$$

where, for example, $(C, \tilde{D}) \equiv C_a\tilde{D}_b - D_a\tilde{C}_b$ and

$$U = C\tilde{D} - \tilde{C}D, \qquad U' = 0 \qquad (26)$$

(the Wronskian $W = CD' - C'D$ is not constant for the p wave; $U = W/\epsilon$ is constant). This matrix is also unimodular, since

$$\det \mathrm{M} = U^{-2}[(C, D)(\tilde{C}, \tilde{D}) - (C, \tilde{D})(\tilde{C}, D)] = 1. \qquad (27)$$

For the homogeneous layer we take $C = \cos(qz)$ and $D = \sin(qz)$ [or $\exp(iqz)$ and $\exp(-iqz)$] to find that $U = Q \equiv q/\epsilon$ and

$$M = \begin{bmatrix} \cos\delta & Q^{-1}\sin\delta \\ -Q\sin\delta & \cos\delta \end{bmatrix}. \qquad (28)$$

The reflection and transmission amplitudes are defined slightly differently for the p wave if one wishes to retain $r_p = r_s$ and $t_p = t_s$ at normal incidence, where there is no physical difference between the s and p waves:

$$\exp(iq_1z) - r_p\exp(-iq_1z) \leftarrow B(z) \rightarrow \frac{n_2}{n_1} t_p \exp(iq_2z). \qquad (29)$$

Thus the equation analogous to Eq. (20) reads, from Eq. (24),

$$\begin{pmatrix} \frac{n_2}{n_1} t_p \exp(iq_2b) \\ iQ_2 \frac{n_2}{n_1} t_p \exp(iq_2b) \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \times \begin{pmatrix} \exp(iq_1a) - r_p\exp(-iq_1a) \\ iQ_1[\exp(iq_1a) + r_p\exp(-iq_1a)] \end{pmatrix}, \qquad (30)$$

where $Q_1 = q_1/\epsilon_1$ and $Q_2 = q_2/\epsilon_2$. This gives

$$-r_p = \exp(2iq_1a) \times \frac{Q_1Q_2m_{12} + m_{21} + iQ_1m_{22} - iQ_2m_{11}}{Q_1Q_2m_{12} - m_{21} + iQ_1m_{22} + iQ_2m_{11}}, \qquad (31)$$

$$\frac{n_2}{n_1} t_p = \exp[i(q_1a - q_2b)] \times \frac{2iQ_1}{Q_1Q_2m_{12} - m_{21} + iQ_1m_{22} + iQ_2m_{11}}. \qquad (32)$$

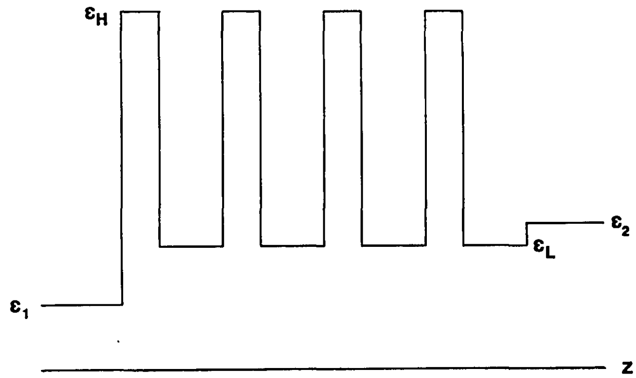

Fig. 1. Dielectric function profile for a (high–low)[4] dielectric mirror, drawn to scale with $n_1 = 1$, $n_h = 2.35$ (ZnS), $n_l = 1.38$ (MgF_2), and $n_2 = 1.5$ (glass). For maximum reflectivity at normal incidence and wavelength λ the layer thicknesses are $d_h = \lambda/4n_h$ and $d_l = \lambda/4n_l$ (a quarter-wave stack).

It is clear from the definition of the layer matrix that a stratification of any number N of layers has the matrix

$$M = M_N M_{N-1} \cdots M_2 M_1. \qquad (33)$$

The results for the reflection and transmission amplitudes given above thus apply to any isotropic stratification. For nonabsorbing media $\epsilon(z)$ is real, and the s- and p-wave basic solutions can be taken to be real (if ψ is a solution of a linear differential equation with real coefficients, then ψ^* is also a solution and so is $\psi + \psi^*$). Thus the matrices are real in the absence of absorption. Energy conservation is then expressed in the algebraic identities

$$R_s + T_s = 1, \qquad R_p + T_p = 1, \qquad (34)$$

where $R_s = |r_s|^2$, $R_p = |r_p|^2$, $T_s = (q_2/q_1)|t_s|^2$, and $T_p = (q_2/q_1)|t_p|^2$. (The reason for the q_2/q_1 factor is discussed in Section 2-1 of Ref. 6; see especially Figure 2-1.)

3. PERIODIC STRUCTURES

We now consider periodic stratifications, such as the high–low multilayer mirror configuration shown in Fig. 1.

We first discuss propagation of waves in an infinite periodic structure. If one period has matrix M, the fields and their derivatives at a corresponding point one period along are given by

$$\begin{pmatrix} \psi_{n+1} \\ \psi_{n+1}' \end{pmatrix} = M \begin{pmatrix} \psi_n \\ \psi_n' \end{pmatrix}, \qquad (35)$$

where ψ represents E or B and ψ' represents E' or $\tilde{B} = \epsilon^{-1}\mathrm{d}B/\mathrm{d}z$. In an infinite structure these positions (one period along from each other) are equivalent, and so the two vectors in Eq. (35) are proportional:

$$\begin{pmatrix} \psi_{n+1} \\ \psi_{n+1}' \end{pmatrix} = \beta \begin{pmatrix} \psi_n \\ \psi_n' \end{pmatrix}. \qquad (36)$$

The Bloch factor β is determined from the condition that Eqs. (35) and (36) together, namely,

$$(M - \beta I)\begin{pmatrix} \psi_n \\ \psi_n' \end{pmatrix} = 0, \qquad (37)$$

have a solution other than zero. [$I = \mathrm{diag}(1, 1)$ is the 2 ×

2 identity matrix.] The condition for nonzero solutions is $\det(M - \beta I) = 0$, which reduces to

$$\beta^2 - 2\beta \cos \phi + 1 = 0\,, \tag{38}$$

when $\operatorname{tr} M = 2 \cos \phi$ and $\det M = 1$ are used. The quadratic (38) has solutions

$$\beta_\pm = \cos \phi \pm (\cos^2 \phi - 1)^{1/2} = \exp(\pm i\phi)\,. \tag{39}$$

Note that $\beta_\pm$ have unit modulus if $\cos^2 \phi < 1$, but that, if $\cos^2 \phi > 1$, the solutions are real and not equal to unity. Thus, if the magnitude of the trace of M exceeds 2, the solutions will grow or decay exponentially; no propagating waves are possible. The condition $\cos^2 \phi > 1$ thus gives the band gaps or the stop bands of the structure. The band edges are given by $\cos^2 \phi = 1$; they occur when ϕ is a multiple of π. When $|\cos \phi| > 1$, ϕ is complex, with the real part a multiple of π and the imaginary part $\xi = \operatorname{Im} \phi$ given by

$$e^\xi = |\cos \phi| + (\cos^2 \phi - 1)^{1/2}\,. \tag{40}$$

In infinite periodic stratifications, the propagation is entirely determined by the trace of the matrix for a single period. We may expect (and we shall shortly show this to be true) that finite periodic structures reflect strongly in the stop bands. The ϕ_s and ϕ_p values for the high–low stack of Fig. 1, repeated to infinity, are shown in Fig. 2. They are calculated from the matrix of the unit cell, a high–low bilayer, which for the s wave is given by

$$\begin{bmatrix} c_l & q_l^{-1}s_l \\ -q_l s_l & c_l \end{bmatrix} \begin{bmatrix} c_h & q_h^{-1}s_h \\ -q_h s_h & c_h \end{bmatrix} = \begin{bmatrix} c_l c_h - q_l^{-1} q_h s_l s_h & q_h^{-1} c_l s_h + q_l^{-1} s_l c_h \\ -q_l s_l c_h - q_h c_l s_h & c_l c_h - q_l q_h^{-1} s_l s_h \end{bmatrix}, \tag{41}$$

where $c_l = \cos \delta_l$, $s_l = \sin \delta_l$, $q_l = (\epsilon_l \omega^2/c^2 - K^2)^{1/2}$, and the phase increment δ_l is $q_l d_l$ with d_l as the thickness of the low-index layer; the parameters for the high-index layer are defined in the same way. Half of the trace of this unit-cell matrix is, for the s wave,

$$\begin{aligned} \cos \phi_s &= c_l c_h - {}^1\!/_2\, s_l s_h (q_l^{-1} q_h + q_h^{-1} q_l) \\ &= \cos(\delta_l + \delta_h) - s_l s_h[(q_l/q_n)^{1/2} - (q_h/q_l)^{1/2}]^2. \end{aligned} \tag{42}$$

For the p wave the arguments of the trigonometric functions remain unchanged, but q_l is replaced by $Q_l = q_l/\epsilon_l$ and q_h is replaced by $Q_h = q_h/\epsilon_h$ in the matrix elements. Thus

$$\begin{aligned} \cos \phi_p &= c_l c_h - {}^1\!/_2 s_l s_h (Q_l^{-1} Q_h + Q_h^{-1} Q_l) \\ &= \cos(\delta_l + \delta_h) - s_l s_h[(Q_l/Q_h)^{1/2} - (Q_h/Q_l)^{1/2}]^2. \end{aligned} \tag{43}$$

Figure 2 shows the real and imaginary parts of ϕ_s and ϕ_p as a function of the angle of incidence θ_1. We see from the figure that, at the design frequency for high reflectivity, the infinite high–low stack does not permit s-wave propagation at any angle of incidence, whereas the p wave can propagate for $\theta_1 \gtrsim 53°$. At $\omega = 1.3\omega_0$ both polarizations can propagate into the stack near normal incidence, but at higher angles the s and p polarizations begin (at different angles of incidence) to reflect totally. The band edges at which this happens are given by the location of $\cos^2 \phi = 1$.

The band structure as a function of frequency is shown in Fig. 3, in which is plotted the real and imaginary parts of ϕ versus ω at normal incidence. The stop band

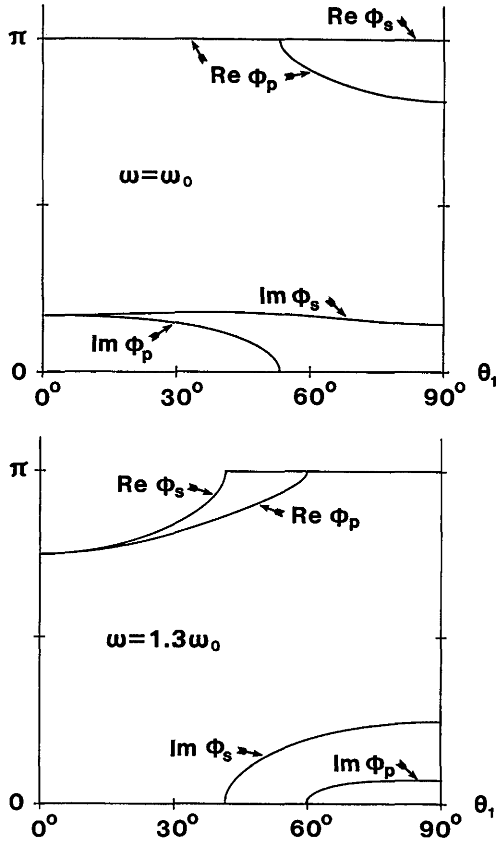

Fig. 2. Real and imaginary parts of ϕ_s and ϕ_p for the ZnS – MgF_2 high–low structure as a function of the angle of incidence. The upper plot is for the design frequency ω_0 for high reflectivity at normal incidence, at which $d_h = \lambda_h/4$ and $d_l = \lambda_l/4$ (the $\lambda/4$ stack) and thus $\delta_h = \pi/2 = \delta_l$. The lower plot is drawn for $\omega = 1.3\omega_0$.

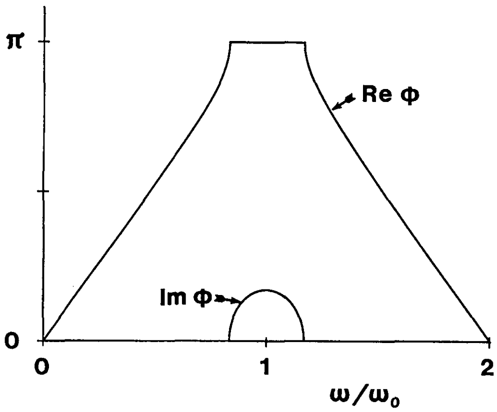

Fig. 3. Band structure parameter $\phi = \arccos({}^1\!/_2 \operatorname{tr} M)$ for a $\lambda/4$ stack as a function of the frequency, drawn for normal incidence onto the high–low stack of Fig. 1. The stop band is centered at the design frequency ω_0, with half-width given by Eq. (44).

($\cos^2 \phi > 1$, ϕ complex) is between $\omega_0 - \Delta\omega$ and $\omega_0 + \Delta\omega$, where

$$\frac{\Delta\omega}{\omega_0} = \frac{2}{\pi} \arcsin\left(\frac{n_h - n_l}{n_h + n_l}\right). \tag{44}$$

At normal incidence the phase increments δ_h and δ_l for the quarter-wave stack are both equal to $(\pi/2)(\omega/\omega_0)$.

Thus $\cos \phi$ is periodic in ω, with period $2\omega_0$. At oblique incidence the s and p waves have different stop bands.

We now look at the optical properties of finite periodic structures. The reflection and transmission amplitudes are given by Eqs. (21) and (22) for the s wave and Eqs. (31) and (32) for the p wave, where in each case the m_{ij} are the matrix elements of the whole structure. Thus, for N periods (for example, N bilayers of the high–low stack), the matrix elements are those of the Nth power of the unit-cell matrix and are given by Eq. (1). It is convenient to define the quantity

$$\sigma_N = \frac{S_{N-1}}{S_N} = \frac{\sin[(N-1)\phi]}{\sin(N\phi)} = \cos \phi - \sin \phi \cot(N\phi), \tag{45}$$

where $\cos \phi$ is half the trace of the unit-cell matrix. Then we have, for the s wave,

$$r_s = \frac{q_1 q_2 m_{12} + m_{21} + iq_1(m_{22} - \sigma_N) - iq_2(m_{11} - \sigma_N)}{q_1 q_2 m_{12} - m_{21} + iq_1(m_{22} - \sigma_N) + iq_2(m_{11} - \sigma_N)}, \tag{46}$$

$$t_s = \frac{2iq_1 S_N^{-1}}{q_1 q_2 m_{12} - m_{21} + iq_1(m_{22} - \sigma_N) + iq_2(m_{11} - \sigma_N)}. \tag{47}$$

(I have omitted the phase factors multiplying r_s and t_s; these are the same for the p-wave reflection and transmission coefficients and do not feature in any experiment that does not compare reflection and transmission phases.[7])

For nonabsorbing media the reflectance and the transmittance are given by

$$R_s = |r_s|^2 = 1 - \frac{4q_1 q_2 S_N^{-2}}{(q_1 q_2 m_{12})^2 + m_{21}^2 + q_1^2(m_{22} - \sigma_N)^2 + q_2^2(m_{11} - \sigma_N)^2 + 2q_1 q_2 S_N^{-2}}, \tag{48}$$

$$T_s = \frac{q_2}{q_1} |t_s|^2 = 1 - R_s. \tag{49}$$

The p wave has a different unit-cell matrix and thus different ϕ and σ_N. The reflection and transmission amplitudes are, again omitting the phase factors,

$$-r_p = \frac{Q_1 Q_2 m_{12} + m_{21} + iQ_1(m_{22} - \sigma_N) - iQ_2(m_{11} - \sigma_N)}{Q_1 Q_2 m_{12} - m_{21} + iQ_1(m_{22} - \sigma_N) + iQ_2(m_{11} - \sigma_N)}, \tag{50}$$

$$\frac{n_2}{n_1} t_p = \frac{2iQ_1 S_N^{-1}}{Q_1 Q_2 m_{12} - m_{21} + iQ_1(m_{22} - \sigma_N) + Q_2(m_{11} - \sigma_N)}. \tag{51}$$

For nonabsorbing media the reflectance is $R_p = 1 - T_p$, where

$$T_p = \frac{q_2}{q_1} |t_p|^2 = \frac{4Q_1 Q_2 S_N^{-2}}{(Q_1 Q_2 m_{12})^2 + m_{21}^2 + Q_1^2(m_{22} - \sigma_N)^2 + Q_2^2(m_{11} - \sigma_N)^2 + 2Q_1 Q_2 S_N^{-2}}. \tag{52}$$

The forms for the reflectance and the transmittance have been obtained by use of the facts that $\det M = 1$ and $\mathrm{tr}\, M = 2 \cos \phi$ (M is the unit-cell matrix) and the identity

$$\sigma_N^2 - 2\sigma_N \cos \phi + 1 = [(\sin \phi)/\sin(N\phi)]^2 = S_N^{-2}. \tag{53}$$

When $N\phi$ is a multiple of π and $(N-1)\phi$ is not, σ_N is infinite and

$$r_s \rightarrow \frac{q_1 - q_2}{q_1 + q_2}, \qquad -r_p \rightarrow \frac{Q_1 - Q_2}{Q_1 + Q_2}. \tag{54}$$

These are the reflection amplitudes of the bare substrate. When $(N-1)\phi$ is a multiple of π and $N\phi$ is not, σ_N is zero and r_s and r_p are the same as the reflection amplitudes of a single period of the structure (supported by the substrate). Thus, for large N, there will be many passes of the reflectivity through the bare substrate and single-period values as the wavelength or the angle of incidence varies.

At the band edges, where $\cos \phi = \pm 1$ and ϕ is a multiple of π, $S_N^2 = N^2$. Thus the transmittance goes to zero as N^{-2} at the band edges, and the reflectance is $1 - O(N^{-2})$. Within the band gaps $\cos \phi = \frac{1}{2} \mathrm{tr}\, M$ has magnitude greater than unity, and ϕ is a multiple of π plus an imaginary part given by Eq. (40):

$$\mathrm{Im}\, \phi = \log[|\cos \phi| + (\cos^2 \phi - 1)^{1/2}]. \tag{55}$$

Then S_N^2 increases exponentially with N,

$$S_N^2 = \left[\frac{\sinh(N\, \mathrm{Im}\, f)}{\sinh(\mathrm{Im}\, f)}\right]^2, \tag{56}$$

and thus the s and p transmittances tend to zero exponentially with the number of periods.

The results (46)–(56) hold for waves in any finite periodic stratification. In particular, two facts are universal: $1 - R = O(N^{-2})$ at the band edges, and R approaches unity exponentially with N inside the stop bands. The construction of the matrices does not assume homogeneity within parts of a unit cell (as is assumed in Refs. 8 and 9, for example). We do not have to assume plane-wave eigenstates. When the unit cell consists of two homogeneous layers, the matrices used here are simpler; compare Eq. (41) with the matrices in Section II of Ref. 8.

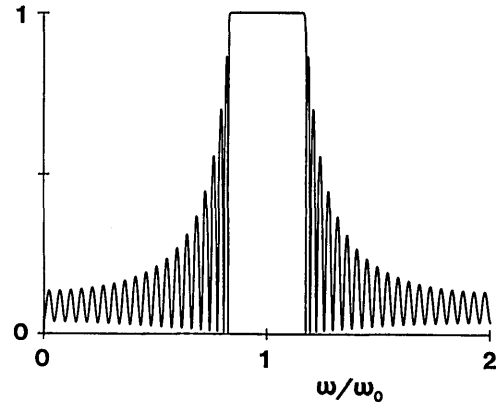

Fig. 4. Frequency dependence of the normal-incidence reflectivity of a dielectric multilayer, drawn for 20 ZnS–MgF_2 bilayers on glass. The multilayer is tuned for high reflectivity at $\omega = \omega_0$ (it is a quarter-wave stack at the design frequency).

Figure 4 shows the normal-incidence reflectivity for a 20-bilayer high–low stack as a function of frequency, and Fig. 5 shows the s and p reflectivities for the same 20-bilayer stack at $\omega = \omega_0$ and $\omega = 1.3\omega_0$, both as a function of the angle of incidence. The stack parameters are the same as those in Figs. 1–3. The corresponding curves for a 4-bilayer stack can be found in Figs. 12-3 and 12-4 of Ref. 6.

4. FORM BIREFRINGENCE

A pencil of light entering an anisotropic material, such as a crystal of calcite, is in general split into two beams: calcite is doubly refracting, or birefringent. The optical properties of anisotropic materials are characterized by a tensor dielectric function.[10,11] For a given angle of incidence of a plane electromagnetic wave onto a given crystal face, two plane-wave modes are possible within the crystal. For a uniaxial crystal, such as calcite, these are called the ordinary and extraordinary modes. The configuration of interest in relation to planar stratified media is one in which the optic axis of a uniaxial material coincides with the surface normal. Then the ordinary and extraordinary modes have wave vectors $(K, 0, q_o)$ and $(K, 0, q_e)$, where[12,13]

$$q_o^2 = \epsilon_o \omega^2/c^2 - K^2, \qquad q_e^2 = \epsilon_o \omega^2/c^2 - (\epsilon_o/\epsilon_e)K^2, \tag{57}$$

where $\epsilon_o = n_o^2$ and $\epsilon_e = n_e^2$ are the ordinary and extraordinary dielectric constants for the crystal. The electric-field vectors of the ordinary and extraordinary modes are along the directions

$$\mathbf{E}_o \sim (0, 1, 0), \qquad \mathbf{E}_e \sim [q_e, 0, -(\epsilon_o/\epsilon_e)K]. \tag{58}$$

From Maxwell's equation for the curl of $\mathbf{E}$ we find that $\mathbf{B}_e \sim (0, 1, 0)$. Thus the normal mode o and e field directions in the crystal correspond to the s and p wave characterizations used in isotropic media. (This holds only when the optic axis is normal to the reflecting surface of the crystal.)

When a narrow beam is incident onto the crystal, it is refracted into two beams whose directions are those of $\mathbf{E} \times \mathbf{B}$ (i.e., along the Poynting vector). The ordinary-mode ray direction always coincides with that of the ordinary wave vector $(K, 0, q_o)$. The extraordinary-mode ray direction does not coincide with $(K, 0, q_e)$ in general. When the optic axis is normal to the reflecting surface, the ray direction of the extraordinary wave is along $[(\epsilon_o/\epsilon_e)K, 0, q_e]$.

We now consider waves in a periodic stratification made up of isotropic component layers. Form birefringence is the name given to the way in which such a structure behaves like an anisotropic homogeneous medium in the limit when the wavelength is large compared with the period (see Ref. 9, Section 14.5.2 of Ref. 10, and Section 6.8 of Ref. 11). The equivalent homogeneous medium is uniaxial, with optic axis normal to the stratifications. To see this, we write the Bloch factor β [which, according to Eq. (39), has eigenvalues $\exp(\pm i\phi)$] as $\exp(\pm iqd)$, where d is the thickness of one unit cell of the stratification and q is interpreted as the normal component of the effective wave vector, which is thus $(K, 0, q_s)$ and $(K, 0, q_p)$ for the s(TE) and p(TM) polarizations. Since q_s and q_p are different (being determined by the trace of M_s and M_p), we have a correspondence with the normal components q_o and q_e of the ordinary and extraordinary waves.

For the high–low stack, $\cos \phi_s$ and $\cos \phi_p$ are given by Eqs. (42) and (43). In the long-wave limit we have $q_l d_l \ll 1$ and $q_h d_h \ll 1$; expansion of Eqs. (42) and (43) in powers of $q_l d_l$ and $q_h d_h$ gives, with

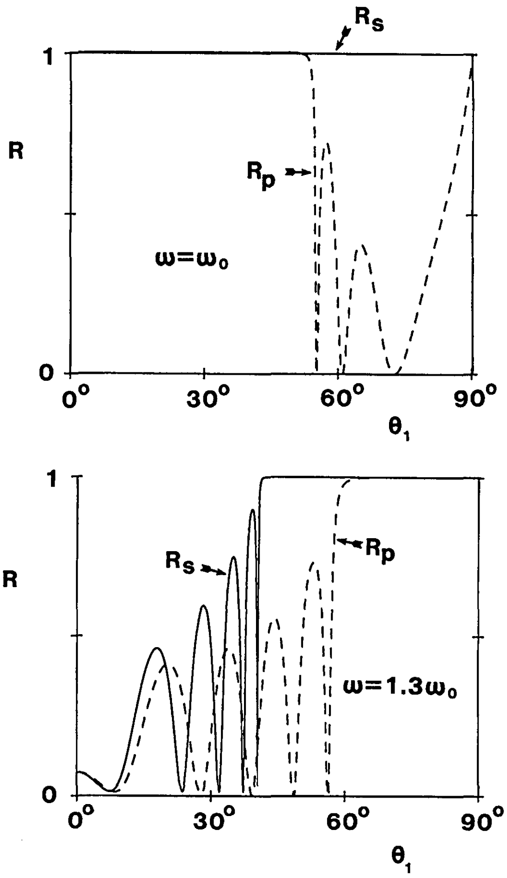

Fig. 5. Reflectivities of a 20-bilayer high–low stack as a function of the angle of incidence at (top) $\omega = \omega_0$ and (bottom) $\omega = 1.3\omega_0$. The parameters are the same as those in Fig. 2. The stop-band edges are at 53.1° for the p waves when $\omega = \omega_0$ and at 41.4° for the s wave and 59.7° for the p wave when $\omega = 1.3\omega_0$.

$$\phi_s = q_s(d_h + d_l) = q_s d\,, \qquad q_s{}^2 \equiv \epsilon_s\omega^2/c^2 - K^2, \tag{59}$$

$$\phi_p = q_p(d_h + d_l) = q_p d\,, \qquad q_p{}^2 \equiv \epsilon_s\omega^2/c^2 - (\epsilon_s/\epsilon_p)K^2 \tag{60}$$

for the TE and TM waves [compare with the expressions for $q_o{}^2$ and $q_e{}^2$ in Eqs. (57)], that

$$\epsilon_s = f_h\epsilon_h + f_l\epsilon_l, \qquad \epsilon_p = \frac{\epsilon_h\epsilon_l}{f_h\epsilon_l + f_l\epsilon_h}, \tag{61}$$

where

$$f_h = \frac{d_h}{d_h + d_l}, \qquad f_l = \frac{d_l}{d_h + d_l} \tag{62}$$

are the fractions of total volume occupied by the high- and low-index in the medium. The expressions (61) have been obtained by electrostatic considerations[10] or by a Bloch-wave dynamical argument[9,11] as in Section 3. Note that the effective anisotropy $\epsilon_e - \epsilon_o$ cannot be positive:

$$\epsilon_p - \epsilon_s = -\frac{(\epsilon_h - \epsilon_l)^2 f_h f_l}{f_h\epsilon_l + f_l\epsilon_h}. \tag{63}$$

Thus the ordinary s or TE wave experiences a larger effective refractive index than does the extraordinary p or TM wave. An experimental demonstration of form birefringence may be seen in Refs. 14 and 15, for example.

The previously known results of the last paragraph apply only to the case in which the unit cell is composed of two homogeneous layers. The long-wave limit can be generalized to an arbitrary dielectric function profile within the unit cell. To second order in the cell thickness divided by the wavelength, the single-period matrix for the s wave is [Eq. (12.96) of Ref. 6]

$$M_s = \begin{bmatrix} 1 - \int_a^b \mathrm{d}z\, q^2(z)(b - z) & b - a \\ -\int_a^b \mathrm{d}z\, q^2(z) & 1 - \int_a^b \mathrm{d}z\, q^2(z)(z - a) \end{bmatrix} \tag{64}$$

(the unit cell extends from $z = a$ to $z = b = a + d$). Thus

$$\cos\phi_s = {}^1\!/\!_2 \operatorname{tr} M_s = 1 - {}^1\!/\!_2(b - a)\int_a^b \mathrm{d}z\, q^2(z) + \ldots. \tag{65}$$

We expand $\cos\phi_s$ as $1 - {}^1\!/\!_2 q_s{}^2(b - a)^2 + \ldots$ and put $q_s{}^2 = \epsilon_s\omega^2/c^2 - K^2$ and $q^2(z) = \epsilon(z)\omega^2/c^2 - K^2$. Then Eq. (65) gives

$$\epsilon_s = \langle\epsilon\rangle \equiv \frac{1}{b - a}\int_a^b \mathrm{d}z\, \epsilon(z). \tag{66}$$

For the p wave the unit-cell matrix is given by Eq. (12.100) of Ref. 6 to second order in the cell thickness:

$$M_p = \begin{bmatrix} 1 - \int_a^b \mathrm{d}z\, [q^2(z)/\epsilon(z)]\int_z^b \mathrm{d}\zeta\, \epsilon(\zeta) & \int_a^b \mathrm{d}z\, \epsilon(z) \\ -\int_a^b \mathrm{d}z\, q^2(z)/\epsilon(z) & 1 - \int_a^b \mathrm{d}z\, \epsilon(z)\int_z^b \mathrm{d}\zeta\, q^2(\zeta)/\epsilon(\zeta) \end{bmatrix}. \tag{67}$$

We expand $\cos\phi_p = {}^1\!/\!_2 \operatorname{tr} M_p$ as $1 - {}^1\!/\!_2 q_p{}^2(b - a)^2 + \ldots$ and set $q_p{}^2 = \epsilon_s\omega^2c^2 - (\epsilon_s/\epsilon_p)K^2$ to find the same $\epsilon_s = \langle\epsilon\rangle$ as above and

$$\frac{\epsilon_s}{\epsilon_p}(b - a)^2 = \int_a^b [\mathrm{d}z/\epsilon(z)]\int_z^b \mathrm{d}\zeta\, \epsilon(\zeta) + \int_a^b \mathrm{d}z\, \epsilon(z)\int_z^b \mathrm{d}\zeta/\epsilon(\zeta), \tag{68}$$

which reduces to

$$\frac{1}{\epsilon_p} = \left\langle\frac{1}{\epsilon}\right\rangle \equiv \frac{1}{b - a}\int_a^b \frac{\mathrm{d}z}{\epsilon(z)}. \tag{69}$$

The expressions (66) and (69) for the ordinary and extraordinary dielectric constants of the equivalent homogeneous but anisotropic medium reduce to Eqs. (61) in the special case of a unit cell made up of two homogeneous layers.

We have thus shown that, in the long-wave limit, any periodically stratified isotropic medium can be replaced by a homogeneous uniaxial medium, with optic axis normal to the stratification, and that $\epsilon_o = \langle\epsilon\rangle$ and $\epsilon_e{}^{-1} = \langle\epsilon^{-1}\rangle$. Since the harmonic mean of a set of positive quantities is never more than its arithmetic mean, it follows that ϵ_e will not exceed ϵ_o, provided that $\epsilon(z)$ is positive everywhere.

The reader may have noticed a curious feature of the proofs given above: we have used the periodicity of the stratification to define a Bloch wave vector through $\phi = qd$, where the cosine of ϕ is half of the trace of the matrix for a unit cell, but the thickness d of the unit cell drops out of the expressions for the equivalent ordinary and extraordinary indices of the equivalent homogeneous medium in the long-wave limit. Could it be that the $\epsilon_o = \langle\epsilon\rangle$ and $\epsilon_e = \langle\epsilon^{-1}\rangle^{-1}$ results apply also to disordered finely layered media? The following argument suggests that they do: consider a stratification that appears disordered on a fine scale (e.g., the nanometer scale) but is actually periodic on a larger scale (e.g., the period is in the tens of nanometers range). The above proof then applies, provided that the wavelength of the radiation is larger still (e.g., hundreds of nanometers). It seems plausible that nonperiodic finely layered media can be represented in the long-wave limit by an effective uniaxial medium with ϵ_o and ϵ_e given by Eqs. (66) and (69); the only difference is that disordered media will scatter more: they will show reflection from variations in the dielectric function $\epsilon(z)$, even in the long-wave limit.

5. SUMMARY

We have used matrices that link fields and their derivatives to simplify the usual treatment of light propagation in periodically stratified media. An arbitrary variation of the dielectric function within a unit cell of the stratification is permitted, instead of the piecewise constant form previously assumed. The band structure and the optical properties are determined by the trace of the unit-cell matrices for the s and p polarizations. The reflectivity shows universal properties with the number of layers N; for example, the reflectivity at the band edges (which occur at values of frequency, wavelength, or angle of incidence for which the trace of the unit-cell matrices has magnitude 2) differs from unity by a term of the order of N^{-2}. The existing results for the form birefringence of periodically stratified media in the long-wavelength limit, valid for a piecewise constant dielectric function, are generalized, and we find that the equivalent ordinary and extraordinary dielectric constants are given by $\epsilon_o = \langle\epsilon\rangle$ and $\epsilon_e^{-1} = \langle\epsilon^{-1}\rangle$. It is suggested that the same relations remain valid in disordered finely layered media.

ACKNOWLEDGMENT

The author is grateful to David Beaglehole for a stimulating conversation about waves in random media.

REFERENCES

1. See, for example, N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Clarendon, Oxford, 1936; reprinted by Dover, New York, 1958).
2. See L. Brillouin, *Wave Propagation in Periodic Structures* (McGraw-Hill, New York, 1946; reprinted by Dover, New York, 1953), Chap. I, for a brief historical review.
3. J. W. S. Rayleigh, "On the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure," Phil. Mag. **24,** 145–159 (1887).
4. J. W. S. Rayleigh, "On the reflection of light from a regularly stratified medium," Proc. R. Soc. London Ser. A **93,** 565–577 (1917).
5. F. Abelès, "Recherches sur la propagation des ondes électromagnétiques sinusoidales dan les milieux stratifiés. Application aux couches minces," Ann. Phys. (Paris) **5,** 596–640, 706–782 (1950).
6. J. Lekner, *Theory of Reflection of Electromagnetic and Particle Waves* (Nijhoff/Kluwer, Dordrecht, The Netherlands, 1987), Chap. 12.
7. J. Lekner, "The phase relation between reflected and transmitted waves, and some consequences," Am. J. Phys. **58,** 317–320 (1990).
8. P. Yeh, A. Yariv, and C. S. Hong, "Electromagnetic propagation in periodic stratified media. I. General theory," J. Opt. Soc. Am. **67,** 423–438 (1977).
9. A. Yariv and P. Yeh, "Electromagnetic propagation in periodic stratified media. II. Birefringence, phase matching, and x-ray lasers," J. Opt. Soc. Am. **67,** 438–448 (1977).
10. M. Born and E. Wolf, *Principles of Optics*, 3rd ed. (Pergamon, Oxford, 1965), Chap. 14.
11. A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley, New York, 1984).
12. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977), pp. 354–355.
13. J. Lekner, "Reflection and refraction by uniaxial crystals," J. Phys. Condens. Matter **3,** 6121–6133 (1991), Sec. 5.1.
14. J. P. van der Ziel, M. Ilegems, and R. M. Mikulyak, "Optical birefringence of thin GaAs–AlAs multilayer films," Appl. Phys. Lett. **28,** 735–737 (1976).
15. M. Kitagawa and M. Tateda, "Form birefringence of SiO_2/Ta_2O_5 periodic multilayers," Appl. Opt. **24,** 3359–3363 (1985).